

Draft Genome Sequence of *Endocarpon pusillum* Strain KoLRILF000583

Sook-Young Park,^a  Jaeyoung Choi,^b Gir-Won Lee,^c Chan-Ho Park,^a Jung A Kim,^a Soon-Ok Oh,^a Yong-Hwan Lee,^{b,c,d} Jae-Seoun Hur^a

Korean Lichen Research Institute, Suncheon National University, Suncheon, South Korea^a; Center for Fungal Pathogenesis, Fungal Bioinformatics Laboratory, Department of Agricultural Biotechnology, Seoul National University, Seoul, South Korea^b; Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul, South Korea^c; Center for Fungal Genetic Resources and Plant Genomics and Breeding Institute, Seoul National University, Seoul, South Korea^d

S.-Y.P., J.C., and G.-W.L. contributed equally to this work.

Lichen *Endocarpon pusillum* is a highly desiccation-tolerant and dominant species in biological soil crusts in arid and semi-arid regions. We report the draft genome sequence of a lichen-forming fungus, *E. pusillum* strain KoLRILF000583. The draft genome assembly has a size of 37,173,200 bp with a GC content of 49.71%, consisting of 40 scaffolds.

Received 23 April 2014 Accepted 1 May 2014 Published 22 May 2014

Citation Park S-Y, Choi J, Lee G-W, Park C-H, Kim JA, Oh S-O, Lee Y-H, Hur J-S. 2014. Draft genome sequence of *Endocarpon pusillum* strain KoLRILF000583. *Genome Announc.* 2(3):e00452-14. doi:10.1128/genomeA.00452-14.

Copyright © 2014 Park 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 Yong-Hwan Lee, yonglee@snu.ac.kr, or Jae-Seoun Hur, jshur1@suncheon.ac.kr.

Biological soil crusts (BSCs) are ubiquitous microbial communities covering most terrestrial surfaces, especially in arid and semi-arid regions. In desert regions, the soil crust serves to restore desert ecosystems by stabilizing the soil against water and wind erosion (1, 2) and increasing carbon and nitrogen fixation (3). Indeed, the soil crusts effectuate the reduction of sand and dust storms and increase soil fertility (4–8). Lichens are common constituents of crust communities, alongside various other nonlichen fungi, bacteria, bryophytes, and algae. Lichens are composite organisms formed by symbiotic associations between a fungus (mycobiont) and an alga or a cyanobacterium (photobiont). As lichens are self-sufficient, autotrophic organisms, lichens are good candidates for the initial colonization and restoration of sandy soils in desert regions.

The lichen *Endocarpon pusillum* is a dominant species among the BSC community in desert regions (9). A number of studies have been conducted where *E. pusillum* was allowed to colonize sandy soils to prevent sand and dust storms and to increase soil fertility (9, 10), because *E. pusillum* exhibits stronger desiccation tolerance than other lichens. However, the information currently available is insufficient to understand the desiccation tolerance of *E. pusillum* at the molecular level. Although a strain of *E. pusillum* has been sequenced and analyzed previously (11), sequencing the genomes of several isolates may help us to understand the evolution of desiccation tolerance within this species.

The strain *E. pusillum* KoLRILF000583 was isolated from soil in Cuiliugou county (N37°25'49.6" N104°34'59.5"), Ningxia Province, China, in 2010. DNA from axenic culture of the fungus was extracted using a DNeasy minikit (Qiagen, Valencia, CA). Draft sequencing was performed by the Illumina HiSeq 2000 system using a whole-genome shotgun strategy (Macrogen, Inc., Seoul, South Korea). The total length of the assembled genome of *E. pusillum* strain KoLRILF000583 was 37,173,200 bp, with a GC content of 49.71%, representing a 556-fold coverage. The genome was assembled into 469 contigs using Allpaths-LG as-

sembler (version 47503) (12). The generated contigs were assembled into 40 scaffolds ($\geq 1,000$ bp) by using SPPACE (version 2.0) (13). Subsequent gene prediction analysis using MAKER (14) yielded a total of 12,062 protein-coding genes. Using the previously developed three gene family pipelines (15–17), 790 transcription factor (TF) genes, 91 cytochrome P450 genes, and 2,031 genes encoding secretory proteins were predicted. In addition, 15 putative polyketide synthase genes, containing ketoacyl synthase, acyltransferase, and acyl carrier domains, were predicted by domain search (18). The genome sequence of *E. pusillum* strain KoLRILF000583 would be a valuable resource for identifying desiccation-tolerance-related genes. Furthermore, the genome sequence will also serve as a platform to facilitate comparative genomics with other lichen-forming fungi as well as other species in the phylum *Ascomycota*.

Nucleotide sequence accession numbers. The draft genome sequence of *E. pusillum* KoLRILF000583 has been deposited at GenBank under accession number [JFDM00000000](https://www.ncbi.nlm.nih.gov/nuccore/JFDM00000000). The version described in this article is the first version, JFDM01000000. The scaffold sequences were also deposited in GenBank under accession numbers [KK106941](https://www.ncbi.nlm.nih.gov/nuccore/KK106941) to [KK106980](https://www.ncbi.nlm.nih.gov/nuccore/KK106980) (40 scaffolds).

ACKNOWLEDGMENTS

This work was supported by grants from the Korea National Research Resource Center Program through the National Research Foundation of Korea (2012M3A9B8021726), Forest Science and Technology Projects (project S111212L030100) provided by the Korea Forest Service, the National Research Foundation of Korea grant funded by the Ministry of Education (2011-0019465) and the Korean government (2008-0061897, 2013-003196, 2013R1A1A2013062, and 2013R1A1A2060561), and the Cooperative Research Program for Agriculture Science and Technology Development (project PJ00821201), Rural Development Administration, Republic of Korea.

REFERENCES

- Eldridge DJ, Kinnel PIA. 1997. Assessment of erosion rates from microphyte-dominated calcareous soils under rain-impacted flow. *Aust. J. Soil Res.* 35:475–489. <http://dx.doi.org/10.1071/S96072>.
- Chaudhary VB, Bowker MA, O'Dell TE, Grace JB, Redman AE, Rillig MC, Johnson NC. 2009. Untangling the biological contributions to soil stability in semiarid shrublands. *Ecol. Appl.* 19:110–122. <http://dx.doi.org/10.1890/07-2076.1>.
- Belnap J, Eldridge DJ. 2001. Disturbance and recovery of biological soil crusts. Springer Verlag, Berlin, Germany.
- Eldridge DJ, Greene RSB. 1994. Microbiotic soil crusts: a view of their roles in soil and ecological processes in the rangelands of Australia. *Aust. J. Soil Res.* 32:389. <http://dx.doi.org/10.1071/SR9940389>.
- Lays JF, Eldridge DJ. 1998. Influence of microbiotic crust disturbance to wind erosion on sand and loam rangeland soils. *Earth Surf. Process. Landforms* 23:963–974. [http://dx.doi.org/10.1002/\(SICI\)1096-9837\(199811\)23:11<963::AID-ESP914>3.3.CO;2-O](http://dx.doi.org/10.1002/(SICI)1096-9837(199811)23:11<963::AID-ESP914>3.3.CO;2-O).
- Belnap J. 2003. The world at your feet: desert biological soil crusts. *Front. Ecol. Environ.* 1:181. [http://dx.doi.org/10.1890/1540-9295\(2003\)001\[0181:TWAYFD\]2.0.CO;2](http://dx.doi.org/10.1890/1540-9295(2003)001[0181:TWAYFD]2.0.CO;2).
- Belnap J, Gillette DA. 1997. Disturbance of biological soil crust: impacts on potential wind erodibility of sandy desert soils in southeastern Utah. *Land Degrad. Dev.* 8:355–362. [http://dx.doi.org/10.1002/\(SICI\)1099-145X\(199712\)8:4<355::AID-LDR266>3.0.CO;2-H](http://dx.doi.org/10.1002/(SICI)1099-145X(199712)8:4<355::AID-LDR266>3.0.CO;2-H).
- Belnap J, Lange OL. 2001. Biological soil crusts: characteristics and distribution. Springer Verlag, New York, NY.
- Ding L, Zhou Q, Wei J. 2013. Estimation of *Endocarpon pusillum* Hedwig carbon budget in the Tengger Desert based on its photosynthetic rate. *Sci. China Life Sci.* 56:848–855. <http://dx.doi.org/10.1007/s11427-013-4526-9>.
- Li X-R, He M-Z, Zerbe S, Li X-J, Liu L-C. 2010. Microgeomorphology determines community structure of biological soil crusts at small scales. *Earth Surf. Process. Landforms* 35:932–940. <http://dx.doi.org/10.1002/esp.1963>.
- Wang YY, Liu B, Zhang XY, Zhou QM, Zhang T, Li H, Yu YF, Zhang XL, Hao XY, Wang M, Wang L, Wei JC. 2014. Genome characteristics reveal the impact of lichenization on lichen-forming fungus *Endocarpon pusillum* Hedwig (*Verrucariales, Ascomycota*). *BMC Genomics* 15:34. <http://dx.doi.org/10.1186/1471-2164-15-S2-P34>.
- 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. U. S. A.* 108:1513–1518. <http://dx.doi.org/10.1073/pnas.1017351108>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding preassembled contigs using SSPACE. *Bioinformatics* 27:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
- Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18:188–196. <http://dx.doi.org/10.1101/gr.6743907>.
- Park J, Park J, Jang S, Kim S, Kong S, Choi J, Ahn K, Kim J, Lee S, Park B, Jung K, Kang S, Lee YH, Kim S, Kang S, Lee YH. 2008. FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. *Bioinformatics* 24:1024–1025. <http://dx.doi.org/10.1093/bioinformatics/btn058>.
- Choi J, Park J, Kim D, Jung K, Kang S, Lee YH. 2010. Fungal secretome database: integrated platform for annotation of fungal secretomes. *BMC Genomics* 11:105. <http://dx.doi.org/10.1186/1471-2164-11-105>.
- Moktali V, Park J, Fedorova-Abrams ND, Park B, Choi J, Lee YH, Kang S. 2012. Systematic and searchable classification of cytochrome P450 proteins encoded by fungal and oomycete genomes. *BMC Genomics* 13:525. <http://dx.doi.org/10.1186/1471-2164-13-525>.
- Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJ, Scheremetjew M, Tate J, Thimmajananathan M, Thomas PD, Wu CH, Yeats C, Yong SY. 2012. Interpro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res.* 40:D306–D312. <http://dx.doi.org/10.1093/nar/gkr948>.