Draft Genome Sequence of *Aspergillus persii* NIBRFGC000004109, Which Has Antibacterial Activity against Plant-Pathogenic Bacteria

Jung A. Kim,[a] Jongbum Jeon,[b] Sook-Young Park,[c] (ID) Ki-Tae Kim,[b] Gobong Choi,[b] Hoa Thi Nguyen,[d] Sun Jeong Jeon,[d] Hyang Burm Lee,[c] Chang-Hwan Bae,[a] Hee-sun Yang,[a] Joo-Hong Yeo,[a] Jin-Cheol Kim,[c] Yong-Hwan Lee,[b] Soonok Kim[a]

Genetic Resources Assessment Division, National Institute of Biological Resources, Incheon, South Korea[a]; Department of Agricultural Biotechnology, Interdisciplinary Program in Agricultural Genomics, Center for Fungal Genetic Resources, and Center for Fungal Pathogenesis, Seoul National University, Seoul, South Korea[b]; Department of Plant Medicine, Sunchon National University, Suncheon, South Korea[c]; Department of Agricultural Chemistry, Chonnam National University, Gwangju, South Korea[d]

ABSTRACT The fungus *Aspergillus persii* strain NIBRFGC000004109 is capable of producing penicillic acid and showed antibacterial activity against various plant-pathogenic bacteria, including *Xanthomonas arboricola* pv. pruni. Here, we report the first draft whole-genome sequence of *A. persii*. The assembly comprises 38,414,373 bp, with 12 scaffolds.

Bioactive compounds from fungi provide an important resource as a lead molecule for the development of novel chemicals. *Aspergillus* spp. are promising fungi for the production of various useful chemicals. We recently found that *Aspergillus persii* strain NIBRFGC000004109 showed strong antibacterial activity, with penicillic acid as an active compound, especially against plant-pathogenic bacteria, including *Agrobacterium tumefaciens*, *Ralstonia solanacearum*, *Xanthomonas arboricola* pv. pruni, and *Xanthomonas oryzae* pv. oryzae (1). The genome sequence of *A. persii* will help us understand the genetic machinery of penicillic acid biosynthesis. Here, we report the first draft genome sequence of *Aspergillus persii* NIBRFGC000004109, which was isolated from barley grain (*Hordeum vulgare*) harvested in 2012 (1).

Genomic DNA was extracted from germlings grown in yeast extract broth by shaking at 150 rpm overnight at 23°C, using the DNeasy minikit (Qiagen, Valencia, CA, USA). A total of 24.62 Gb of genome sequences were produced using 1 paired-end and 2 mate pair libraries on HiSeq 2000 (Illumina) and 8 cells on PacBio RSII platforms at the Theragen Etex Bio Institute (Suwon, South Korea), representing 616.57-fold coverage. Reads containing the nuclear genome were assembled into 12 scaffolds (501 contigs), with an N_{50} value of 4.60 Mb, using a mixed pipeline of FALCON, an assembler of PacBio reads, and SOAP*denovo* version 2, an assembler of Illumina reads (2). The respective assembled reads were merged with HaploMerger2 (3). The total length of the assembled genome was 38,414,373 bp, with a G+C content of 48.64%.

A total of 12,328 protein-coding genes were identified after gene prediction and annotation based on the AUGUSTUS software (4). Using the previously developed gene family pipelines (5–10), we found 673 transcription factor genes, 159 cytochrome P450 genes, 1,079 genes encoding secretory proteins, 61 genes encoding plant cell wall-degrading enzymes, 8 genes encoding laccases, and 33 genes encoding peroxidases. Additionally, 68 polyketide synthase and 77 nonribosomal peptide synthetase genes were identified using the antiSMASH software (11, 12). This is the first report of the genome sequence of *A. persii* and its analysis. This draft genome sequence will provide

valuable information to identify genes for the biosynthesis of penicillic acid and will also serve as a platform to facilitate comparative genomics with other *Aspergillus* fungi for evolutionary biology.

Accession number(s). The draft genome sequence of *Aspergillus persii* NIBRFGC000004109 has been deposited in GenBank under accession no. NGZO00000000. The version described in this article is the first version, NGZO01000000.

REFERENCES

1. Nguyen HT, Yu NH, Jeon SJ, Lee HW, Bae CH, Yeo JH, Lee HB, Kim IS, Park HW, Kim JC. 2016. Antibacterial activities of penicillic acid isolated from *Aspergillus persii* against various plant pathogenic bacteria. Lett Appl Microbiol 62:488–493. https://doi.org/10.1111/lam.12578.
2. 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 SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1:18. https://doi.org/10.1186/2047-217X-1-18.
3. Huang S, Chen Z, Huang G, Yu T, Yang P, Li J, Fu Y, Yuan S, Chen S, Xu A. 2012. HaploMerger: reconstructing allelic relationships for polymorphic diploid genome assemblies. Genome Res 22:1581–1588. https://doi.org/10.1101/gr.133652.111.
4. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a Web server for gene finding in eukaryotes. Nucleic Acids Res 32:W309–W312. https://doi.org/10.1093/nar/gkh379.
5. Park J, Park J, Jang S, Kim S, Kong S, Choi J, Ahn K, Kim J, Lee S, Kim S, Park B, Jung K, Kim S, Kang S, Lee YH. 2008. FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. Bioinformatics 24:1024–1025. https://doi.org/10.1093/bioinformatics/btn058.
6. 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. https://doi.org/10.1186/1471-2164-11-105.
7. 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. https://doi.org/10.1186/1471-2164-13-525.
8. Choi J, Détry N, Kim KT, Asiegbu FO, Valkonen JP, Lee YH. 2014. fPoxDB: fungal peroxidase database for comparative genomics. BMC Microbiol 14:117. https://doi.org/10.1186/1471-2180-14-117.
9. Choi J, Kim KT, Jeon J, Lee YH. 2013. Fungal plant cell wall-degrading enzyme database: a platform for comparative and evolutionary genomics in fungi and oomycetes. BMC Genomics 14(Suppl 5):S7. https://doi.org/10.1186/1471-2164-14-S5-S7.
10. Park J, Lee S, Choi J, Ahn K, Park B, Park J, Kang S, Lee YH. 2008. Fungal cytochrome P450 database. BMC Genomics 9:402. https://doi.org/10.1186/1471-2164-9-402.
11. 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, Thimmajanarthanan 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. https://doi.org/10.1093/nar/gkr948.
12. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0–a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.