An endolichenic fungus, *Xylaria grammica* strain EL000614, showed strong nematicidal effects against plant pathogenic nematode, *Meloidogyne incognita* by producing grammicin. We report genome assembly of *X. grammica* EL000614 comprised of 25 scaffolds with a total length of 54.73 Mb, N50 of 4.60 Mb, and 99.8% of BUSCO completeness. GC contents of this genome were 44.02%. Gene families associated with biosynthesis of secondary metabolites or regulatory proteins were identified out of 13,730 gene models predicted.
assembled genome of long-reads was polished based on Illumina paired-end library (70,717,638 reads) using Pilon v 1.22 [6]. The final genome consists of 25 contigs and one mitochondrial genome with a total length of 54,730,220 bp and 196,363 bp, respectively. The assembly has N50 being 4.60 Mbp, and the GC content of the genome was 44.02%. Genome completeness was assessed using BUSCO v. 5.0.0 against the “fungi_odb10” database, showing 99.8% of 758 fungal gene sets were complete (Table 1) [7].

Gene prediction was performed by BRAKER2 [8] using constructed model from RNA-seq and protein models of five close species in the NCBI nonredundant (nr) database. A total of 13,730 gene models were predicted, of which 12,900 genes (93.96%) had hits in UniProt, nr, or InterPro database. Ninety-eight Biosynthetic gene clusters (BGCs), including 45 polyketide synthase (PKS) and 42 non-ribosomal protein synthetase (NRPS), 21 terpene synthase, were identified through antiSMASH v5.0 [9]. Among these were two orthologs encoding PatK, a PKS backbone gene in patulin biosynthesis [10]. Further, 476 genes encoding transcription factor and 222 genes encoding cytochrome P450 were cataloged through Fungal Transcription Factor Database (FTFD) v1.0 [11], and Fungal Cytochrome P450 Database v1.0 [12], respectively. Genome was composed of 16.36% of repetitive sequences such as retrotransposons, DNA transposons, simple sequence repeats, and unclassified repeats through the combining of RebBase [13,14] search and de novo identification using RepeatModeler 1.0.8 and RepeatMasker 4.0.6 [15]. A total of 531 non-coding RNAs, including 43 rRNA from BLASTn search against NCBI, 225 tRNA through tRNAscan-SE [16], 121 microRNA and 142 small nucleolar RNA

Table 1. Genome assembly and annotation statistics of Xylaria grammica EL000614.

| Statistics                      | X. grammica EL000614 |
|---------------------------------|----------------------|
| NCBI accession                  | NGZP0000000          |
| Sequence Reads Archive          | SRR14073447–SRR14073457 |
| Coverage (X)                    | 513.77               |
| Number of scaffolds             | 25                   |
| Genome size (bp)                | 54,730,220           |
| N50 (bp)                        | 3,878,448            |
| GC content (%)                  | 44.02                |
| BUSCO completeness (%)           | 99.8                 |
| Number of predicted gene models | 13,730               |
| Number of functional annotation | 12,900               |
| Repettitive sequence (%)        | 16.36                |
| Number of ncRNA                 | 531                  |

Figure 1. Morphological characteristics of Xylaria grammica strain EL000614. (A) Two weeks-old colony on the upper side of PDA. (B) Dark red pigmentation on the reverse side of PDA. Four-day-old colony on the upper side (left) and the reverse side (right) of PDA under constant dark condition (C) and under constant light condition (D).
through Infernal search against Rfam 11.0 [17], were also identified. This draft genome of *X. grammica* EL000614 will help further characterize the genes involved in grammicin biosynthesis and their regulation.

The sequence data and RNA-seq data of *X. grammica* EL000614 were deposited at Sequence Read Archive (SRA) under the accession number SRR14073447 to SRR14073457. The assembly data also deposited GenBank under the accession no. NGZP00000000. The version described in this article is NGZP02000000 (https://www.ncbi.nlm.nih.gov/nuccore/NGZP00000000.2). The associated Bioproject number is PRJNA368785

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by the National Institute of Biological Resources, funded by the Ministry of Environment of the Republic of Korea (projects NIBR201921101 and NIBR202021101).

ORCID
Sook-Young Park [http://orcid.org/0000-0003-1267-1111](http://orcid.org/0000-0003-1267-1111)
Yong-Hwan Lee [http://orcid.org/0000-0003-2462-1250](http://orcid.org/0000-0003-2462-1250)
Soonok Kim [http://orcid.org/0000-0003-1654-3643](http://orcid.org/0000-0003-1654-3643)

References

[1] Kellogg J, Raja HA. Endolichenic fungi: a new source of rich bioactive secondary metabolites on the horizon. Phytochem Rev. 2017;16(2):271–293.

[2] Kim TY, Jang JY, Yu NH, et al. Nematicidal activity of grammicin produced by *Xylaria grammica* KCTC 13121BP against *Meloidogyne incognita*. Pest Manag Sci. 2018;74(2):384–391.

[3] Edwards RL, Maitland DJ, Pittayakajhonwut P, et al. Metabolites of the higher fungi. Part 33. Grammicin, a novel bicyclic C$_7$H$_6$O$_4$ furanopyranol from the fungus *Xylaria grammica* (Mont.) Fr. Perkin Trans. 2001;1(11):1296–1299.

[4] Wang A, Wang Z, Li Z, et al. BAUM: improving genome assembly by adaptive unique mapping and local overlap-layout-consensus approach. Bioinformatics. 2018;34(12):2019–2028.

[5] Koren S, Walenz BP, Berlin K, et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27(5):722–736.

[6] Walker BJ, Abeel T, Shea T, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE. 2014;9(11):e112963.

[7] Waterhouse RM, Seppey M, Simao FA, et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol Biol Evol. 2018;35(3):543–548.

[8] Brůna T, Hoff KJ, Lomsadze A, et al. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP + and AUGUSTUS supported by a protein database. NAR Genom Bioinform. 2021;3:108.

[9] Blin K, Shaw S, Steinke K, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 2019;47(W1): W81–W87.

[10] Artigot MP, Loiseau N, Laffitte J, et al. Molecular cloning and functional characterization of two CYP619 cytochrome P450s involved in biosynthesis of patulin in *Aspergillus clavatus*. Microbiology. 2009;155(Pt 5):1738–1747.

[11] Park J, Park J, Jang S, et al. FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. Bioinformatics. 2008;24(7):1024–1025.

[12] Park J, Lee S, Choi J, et al. Fungal cytochrome P450 database. BMC Genomics. 2008;9:402.

[13] Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob Dna. 2015;6:11.

[14] Kapitonov VV, Jurka J. A universal classification of eukaryotic transposable elements implemented in Repbase. Nat Rev Genet. 2008;9(5):411–412.

[15] Smit AFA, Hubley R, Green P. RepeatMasker. 2015. Available from: http://repeatmasker.org.

[16] Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol. 2019;1962:1–14.

[17] Burge SW, Daub J, Eberhardt R, et al. Rfam 11.0: 10 years of RNA families. Nucleic Acids Res. 2013;41(Database issue):D226–D232.