

Resource Announcement

**High quality genome of the tree pathogen *Phytophthora plurivora* - a novel resource for epidemiological research**

T. Tsykun<sup>1,4\*</sup>, B. Mishra<sup>2</sup>, S. Ploch<sup>2</sup>, RI Alcalá Briseño<sup>3</sup>, S. Prospero<sup>4</sup>, N. J. Grünwald<sup>5</sup>, M. Thines<sup>1,2,6</sup>

<sup>1</sup> Institute of Ecology, Evolution and Diversity, Faculty of Biosciences, Goethe University Frankfurt am Main, Max-von-Laue-Str. 9, D-60438 Frankfurt am Main, Germany

<sup>2</sup> Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Senckenberganlage 25, D-60325 Frankfurt am Main, Germany

<sup>3</sup> Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, USA

<sup>4</sup> Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

<sup>5</sup> Horticultural Crops Research Laboratory, USDA Agricultural Research Service, Corvallis, OR, USA

<sup>6</sup> LOEWE Centre for Translational Biodiversity Genomics (TBG), Georg-VoigtStr. 14-16, 60325 Frankfurt am Main, Germany

\*Corresponding author: T. Tsykun; [tetyana.tsykun@wsl.ch](mailto:tetyana.tsykun@wsl.ch)

**Keywords:** oomycetes, chromosome level, disease causing agent, *Fagus sylvatica*

**Abstract**

*Phytophthora plurivora* can affect a range of ecologically and silviculturally important tree species, including European beech (*Fagus sylvatica*), a common late successional tree species native to Europe. Here, we report on the high-quality genome of *P. plurivora* strain TJ71 (CBS 124093). We sequenced it using Oxford Nanopore MinION and PacBio Sequel II long-read sequencing with 80x coverage, chromatin conformation capture (Hi-C) sequencing with 400x coverage and DNBSEQ 150bp paired-end short reads sequencing with 200x coverage. This complex sequencing approach allowed assembly of the

genome at the chromosome level. Specifically, the *Phytophthora plurivora* genome resulted in 18 scaffolds of 47 Mbp total size with 95% completeness of the eukaryotic gene set as implemented in BUSCO. This is a considerable improvement relative to the previous NCBI reference genomes *P. plurivora* (NMPK00000000.1) with ~41 Mbp organized in 1,898 scaffolds with 93.8% eukaryotic BUSCO completeness. This high quality genome provides a valuable resource for further evolutionary, epidemiological and population genomic studies.

### Genome Announcement

A number of broadleaf tree genera are currently threatened worldwide by pathogens in the oomycete genus *Phytophthora* (Goheen et al. 2009; Jung et al. 2016; Lamour 2013). European beech (*Fagus sylvatica*) is of special concern among common forest forming tree species in Europe (Houston Durrant et al. 2016). Until the beginning of this century, European beech was considered to be resistant to pathogens and pests. However, we now are observing a widespread decline of juvenile as well as mature trees across Europe and North America caused mainly by two *Phytophthora* species, i.e. *P. plurivora* and *P. x cambivora* (Cleary et al. 2017; Jung and Burgess 2009; Jung et al. 2016; Ruffner et al. 2019). These two pathogens are, however, also associated with oak (*Quercus* spp.) decline and damage to others deciduous trees (Jung et al. 2006; Jung et al. 2016; Mrazkova et al. 2013).

Whole genome sequencing enables novel population genomic and epidemiological insights into plant pathogen evolution (Grünwald 2012; Thines and Kamoun 2010). In particular, genome-wide association studies and linkage mapping can identify genomic regions associated with virulence or genes responsible for pathogenic success of newly emerged genotypes (Dalman et al. 2013; Talas and McDonald 2015). Genome scans of regions under diversifying selection in different populations have the potential to identify candidate genes previously unknown to be involved in virulence, host specialization or local adaptation (Cooke et al. 2012; Grünwald et al. 2016). A crucial prerequisite for these studies is a high quality reference genome of disease causing agents and their structural and functional annotation

(Grünwald et al. 2017; Grünwald et al. 2016). Novel, high-quality genomes of *P. plurivora* would provide a keystone for further understanding of the underlying patterns leading to the emergence of aggressive genotypes; this would also be of critical importance for biosecurity risk assessment and management of future epidemics of soil-borne *Phytophthora* species.

The *P. plurivora* strain TJ71 was isolated from a diseased *Fagus sylvatica* tree in Germany (deposited as CBS 124093) and was first reported in Corcobado et al. (2022). For DNA extraction, the strain was grown at room temperature for one week on Standard Agar Medium (Kruse et al. 2017) supplemented with the antibiotics rifampicin and pimaricin (25 µg ml<sup>-1</sup> each). A double-sterilized cellophane foil was placed on the agar surface of the inoculated plates. High molecular weight DNA was extracted from the mycelium, which was scratched from the cellophane and immediately frozen in liquid nitrogen. The modified from Francis and Clair (1993) DNA extraction procedure was used. Prior to Oxford Nanopore library preparation, the samples were purified with magnetic beads (Mag-Bind TotalPure NGS, Omega BIO-TEK, USA) using a 0.6x magnetic bead:DNA ratio to remove low molecular DNA fragments and residual inhibitors. The DNA was used for Nanopore library construction using the SQK-LSK110 kit following the manufacturer's instructions (ONT, Oxford, UK) and then sequenced on a MinION Mk1b with a FLO-MIN106 flowcell (both ONT, Oxford, UK). An aliquot of the frozen DNA was sent to the commercial sequencing provider BGI (Hong Kong, China) for DNBSEQ 150bp paired-end reads and for PacBio Sequel II sequencing. In addition, the deep-frozen mycelium was sent to the same the commercial sequencing company for library preparation and Hi-C sequencing.

The base calling and processing was done using Guppy version 6.0.1 (Oxford Nanopore Technologies, Oxford, UK) to obtain fastq files with a Phred quality score of more than 7 and length more than 500 bp.

The short 150 bp paired-end reads of DNBSEQ were filtered to remove adapters, reads with an average quality less than 25 and shorter than 70 bp using Trimmomatic v0.39 (Bolger et al. 2014). Long-reads were self-corrected, trimmed and assembled using PacBio software Canu v2.1.1 (Koren et al. 2017) separately for Nanopore, PacBio reads and a dataset combining reads from both technologies. Each of

the three assemblies were improved using three cycles of subsequent 150 bp paired-end reads alignment using Bwa mem v0.7.17 (Li 2013) and filter for quality level of at least 40. Sorting and indexing was done using Samtools v1.12 and a final assembly correction was achieved with Pilon v1.24 (Walker et al. 2014). The completeness and continuity of the raw assemblies was assessed using the BUSCO eukaryotic genes set (Seppey et al. 2019), N50 contigs length, the number and sizes of contigs. The assembly with optimal statistical characteristics was chosen for further assembly of its contigs into chromosome-level scaffolds with the Hi-C reads.

HiCUP v0.8.3 (Wingett et al. 2015) was used for filtering and aligning Hi-C reads. Digestion sites of the Hi-C enzymes (ATC, DpnII) were located in the raw assembly, and reads were cleaned by termination at the restriction enzyme recognition sites. Finally, reads were mapped to the raw assembly independently using Bowtie2 with default parameters (Langmead and Salzberg 2012), retaining only reads with both partners of each pair uniquely mapped. Other pairs, i.e. those generated from contiguous sequences, dangling ends, circularization, re-ligation, PCR duplication, and fragments of unexpected size ( $\leq 80$  bp and  $\geq 700$  bp) were filtered out. The Hi-C alignments and raw assemblies were used for further scaffolding applying the ALLHiC pipeline following Zhang et al. (2019). Apparent problematic regions in the assembly were improved using the juicebox pipelines (Durand et al. 2016). The final assembly was further corrected using three cycles of Pilon polishing with the 150 bp paired-end short reads. The completeness of the assembled genomes were evaluated using benchmarking universal single-copy orthologs (BUSCO) version 5.2.2 (Seppey et al. 2019) with the eukaryota\_odb10 (70 genomes, 255 genes) database.

The BRAKER v 3.0.2 pipeline was used to predict genes (Brůna et al. 2021). The assembly was soft-masked for repeats using RepeatMasker v 4.1.2-p1 (Smit et al. 2015). A custom library for masking was generated by concatenating repeats found *ab initio* with RepeatModeler v 2.0.3 (Smit et al. 2015) and repeats from GIRI RepBase (Bao et al. 2015) in the *Phytophthora* lineage. In addition, proteomes from *Phytophthora* spp. clades 1 to 8 were downloaded from NCBI database and protein hints predicted with

101 ProtHin (Brůna et al. 2020) were supplied to BRAKER (Brůna et al. 2021). Then, genes were predicted  
102 using the dataset training with GeneMark-EP+ (Brůna et al. 2020; Lomsadze et al. 2005), AUGUSTUS  
103 protein predictions (Stanke et al. 2006), DIAMOND and Spaln2 alignments (Buchfink et al. 2015; Gotoh  
104 et al. 2014; Iwata and Gotoh 2012). The resulting gene predictions were subset with BRAKER to three  
105 structural annotations: (1) with full proteome support in the assembly, (2) partial findings of the proteins  
106 and (3) *de novo* gene annotations. Final gene predictions were converted to gff3, and basic statistics were  
107 generated with AGAT v1.0.0 (Dainat 2020), translated to protein files using BRAKER (Stanke et al.  
108 2008). Functional annotation of the predicted genes was conducted with Interproscan v5.56-89.0 (Jones  
109 et al., 2014).

110 The assembly of the *P. plurivora* genome resulted in 18 pseudo-molecules with 95% BUSCO eukaryotic  
111 gene completeness (Table 1) and 22 unscaffolded contigs. Each unscaffolded contig was shorter than  
112 0.12 Mbp and had no additional BUSCO genes. The analysis of 21 bp k-mer frequencies and coverage  
113 distributions in 150 bp paired-end reads performed with Jellyfish (Marçais and Kingsford 2011) and  
114 GenomeScope 2.0 (Ranallo-Benavidez et al. 2020) suggested that *P. plurivora* has a diploid genome  
115 with an expected low heterozygosity of 0.06%, similar to other homothallic *Phytophthora* species  
116 (Thines et al. 2020; Tsykun et al. 2022) with ~20% repetitive sequences. The last is congruent with the  
117 results we obtained with RepeatMasker analysis for the assembled 18 scaffolds (Table 1). The number  
118 of genes (12733 - 14888, Table 1) called for the assembly and annotated (Table 2) is congruent to the  
119 number (11,741 genes) reported in Vetukuri et al. (2018). However, our genome is a substantial  
120 improvement with respect to contiguity and completeness relative to the previously reported genome for  
121 *P. plurivora* (NMPK00000000.1) with ~41 Mbp genome size organized into 1,898 with 93.8%  
122 completeness with respect to core eukaryotic genes as implemented in BUSCO (Vetukuri et al. 2018).  
123 Furthermore, we achieved a near-chromosome level assembly with oomycete and plant specific  
124 telomeric motifs (TTTAGGG)<sub>n</sub> or/and (TTTAGG)<sub>n</sub> (Fulnečková et al. 2013) present in both ends of 13  
125 scaffolds and in one end of the other 5 scaffolds.

126 The high quality genome of the important tree pathogen *Phytophthora plurivora*, which was *de novo*  
 127 assembled and annotated here, will serve as a valuable resource for future discovery of virulence factors.  
 128 This genome provides a novel resource for understanding the recent emergence, evolution, ecology and  
 129 adaptation of *P. plurivora*. The raw sequencing reads and the genome assembly data are deposited in  
 130 NCBI BioProject the accession number PRJNA962935.

### 131 **Acknowledgements**

132 We gratefully acknowledge the German Science Foundation (DFG) for financial support of the TT  
 133 research project (TS 421/1-1), as well as the MT project funding by the LOEWE excellence initiative of  
 134 the government of Hesse, in the framework of the LOEWE Centre for Translational Biodiversity  
 135 Genomics (TBG). NJG work was supported by NIFA grant 2018-67013-27823 and USDA ARS 2072-  
 136 22000-041-000-D. We thank the Swiss Federal Institute for Forest, Snow and Landscape Research  
 137 (WSL), the Center for Quantitative Life Sciences at Oregon State University, and the LOEWE TBG  
 138 laboratory center for access to high-performance computing facilities. We thank Thomas Jung for  
 139 providing the reference strains.

140 **Table 1.** Genome assembly statistics of *Phytophthora plurivora* strain TJ71 (CBS 124093 isolate)

Statistical characteristic	Value
Genome size in scaffolds (Mbp)	46.88
Scaffolds (counts)	18
Scaffold N50 (Mbp)	3.00
Scaffold count L50 (counts)	6
GC Content, %	51.89
Longest scaffold, Mbp	6.64

Shortest scaffold, Mbp	1.06
Number (% to total size) of unscaffolded contigs	22 (2.2 %)
Complete BUSCOs, % (counts)	94.50 (241 / 241)*
Complete and single-copy BUSCOs (S), % (counts)	92.90 (237 / 237)
Complete and duplicated BUSCOs (D), % (counts)	1.60 (4 / 4)
Fragmented BUSCOs (F), % (counts)	2.00 (5 / 5)
Missing BUSCOs (M), % (counts)	3.50 (9 / 9)
Total BUSCO groups searched (counts)	255
Predicted ploidy	2n
Heterozygosity, %	0.07
Repeats, % among them identified:	21.15
Retroelements	5.70
DNA transposons	4.81

Satellites	0.003
Simple repeats	0.57
Low complexity	0.04
Number of genes (mRNA) among them has matches in the proteome dataset used:	16619 (17321)
Partial	14880 (15484)
Full	10994 (11188)
No	1739 (1837)

141

142 \* Numbers of BUSCO genes counted in scaffolds excluding and including unscaffolded contigs are  
 143 shown in brackets respectively.

144 **Table 2.** Functional annotation of the *P. plurivora* genes with the complete proteome dataset support

Total number of protein families / number of protein sequences	17.84 K / 130 K
Among them related to pathogenicity:	
Aspartyl proteases	9 / 26
Serine carboxypeptidases	7 / 116

Cysteine proteinases	22 / 314
Glycosyl hydrolases	94 / 975
Pectin esterases	7 / 124
Pectate lyases	5 / 110
Lipases	22 / 63
Phospholipases	37 / 180
Protease inhibitors	61 / 583
Cytochrome P450s	10 / 374
ABC transporters	27 / 796
Necrosis inducing proteins	7 / 96
Elicitin-like proteins	5 / 230
Phytotoxin (PcF) protein	1 / 2**
RXLR cytoplasmic effectors	2 / 23 (2 / 27)*
CRN cytoplasmic effectors	1 / 21 (1 / 26)*

145

146 \*Functional annotation of genes that has no orthologs (no support) in the protein database used in  
147 BRAKER analysis

148

## 149 Literature Cited

- 150 Bao, W., Kojima, K. K., and Kohany, O. 2015. Repbase Update, a database of repetitive  
151 elements in eukaryotic genomes. *Mobile DNA* 6:11.
- 152 Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina  
153 sequence data. *Bioinformatics* 30:2114-2120.
- 154 Brůna, T., Lomsadze, A., and Borodovsky, M. 2020. GeneMark-EP+: eukaryotic gene  
155 prediction with self-training in the space of genes and proteins. *NAR genomics and*  
156 *bioinformatics* 2.
- 157 Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., and Borodovsky, M. 2021. BRAKER2:  
158 automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS  
159 supported by a protein database. *NAR genomics and bioinformatics* 3.
- 160 Buchfink, B., Xie, C., and Huson, D. H. 2015. Fast and sensitive protein alignment using  
161 DIAMOND. *Nature Methods* 12:59-60.
- 162 Cleary, M. R., Blomquist, M., Vetukuri, R. R., Böhlenius, H., and Witzell, J. 2017.  
163 Susceptibility of common tree species in Sweden to *Phytophthora cactorum*,  
164 *P. cambivora* and *P. plurivora*. *Forest Pathol.* 47:n/a-n/a.
- 165 Cooke, D. E., Cano, L. M., Raffaele, S., Bain, R. A., Cooke, L. R., Etherington, G. J., Deahl,  
166 K. L., Farrer, R. A., Gilroy, E. M., and Goss, E. M. 2012. Genome analyses of an  
167 aggressive and invasive lineage of the Irish potato famine pathogen. *PLoS pathogens*  
168 8:e1002940.
- 169 Corcobado, T., Milenković, I., Saiz-Fernández, I., Kudláček, T., Plichta, R., Májek, T., Bačová,  
170 A., Ďatková, H., Dálya, L. B., Trifković, M., Mureddu, D., Račko, V., Kardošová, M.,  
171 Ďurkovič, J., Rattunde, R., and Jung, T. 2022. Metabolomic and Physiological  
172 Changes in *Fagus sylvatica* Seedlings Infected with *Phytophthora plurivora* and the A1  
173 and A2 Mating Types of *P. &times;cambivora*. *Journal of Fungi* 8:298.

174 Dainat, J. 2020. AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF  
175 format. (Version v1.0.0). Zenodo. <https://www.doi.org/10.5281/zenodo.3552717>.

176 Dalman, K., Himmelstrand, K., Olson, A., Lind, M., Brandstrom-Durling, M., and Stenlid, J.  
177 2013. A genome-wide association study identifies genomic regions for virulence in the  
178 non-model organism *Heterobasidion annosum* s.s. Plos One 8:e53525.

179 Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., and  
180 Aiden, E. L. 2016. Juicebox Provides a Visualization System for Hi-C Contact Maps  
181 with Unlimited Zoom. Cell systems 3:99-101.

182 Fulnečková, J., Ševčíková, T., Fajkus, J., Lukešová, A., Lukeš, M., Vlček, Č., Lang, B. F.,  
183 Kim, E., Eliáš, M., and Sýkorová, E. 2013. A Broad Phylogenetic Survey Unveils the  
184 Diversity and Evolution of Telomeres in Eukaryotes. Genome Biology and Evolution  
185 5:468-483.

186 Goheen, E. M., Frankel, S. J., and coords., t. 2009. Phytophthora in Forests and Natural  
187 Ecosystems. Page 334 in: Proceedings of the fourth meeting of the International Union  
188 of Forest Research Organizations (IUFRO) Working Party S07.02.09, Gen. Tech. Rep.  
189 PSW-GTR-221. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific  
190 Southwest Research Station.

191 Gotoh, O., Morita, M., and Nelson, D. R. 2014. Assessment and refinement of eukaryotic  
192 gene structure prediction with gene-structure-aware multiple protein sequence  
193 alignment. BMC Bioinformatics 15:189.

194 Grünwald, N. J. 2012. Genome sequences of *Phytophthora* enable translational plant disease  
195 management and accelerate research. Canadian Journal of Plant Pathology 34:13-19.

196 Grünwald, N. J., McDonald, B. A., and Milgroom, M. G. 2016. Population genomics of fungal  
197 and oomycete pathogens. Annu. Rev. Phytopathol. 54:323-346.

- 198 Grünwald, N. J., Everhart, S. E., Knaus, B. J., and Kamvar, Z. N. 2017. Best Practices for  
199 Population Genetic Analyses. *Phytopathology* 107:1000-1010.
- 200 Houston Durrant, T., de Rigo, D., and Candullo, G. 2016. *Fagus sylvatica* and other beeches  
201 in Europe: distribution, habitat, usage and threats in San Miguel Ayanz. J., de Rigo, D.,  
202 Candullo, G., Houston Durrant, T., Mauri, A.(eds.) *European Atlas of Forest Tree*  
203 *Species*. Publication Office of the European Union, Luxembourg, pp. e012b90.
- 204 Iwata, H., and Gotoh, O. 2012. Benchmarking spliced alignment programs including Spaln2,  
205 an extended version of Spaln that incorporates additional species-specific features.  
206 *Nucleic Acids Research* 40:e161-e161.
- 207 Jung, T., and Burgess, T. 2009. Re-evaluation of *Phytophthora citricola* isolates from multiple  
208 woody hosts in Europe and North America reveals a new species, *Phytophthora*  
209 *plurivora* sp. nov. *Persoonia-Molecular Phylogeny and Evolution of Fungi* 22:95-110.
- 210 Jung, T., Hudler, G. W., Jensen-Tracy, S. L., Griffiths, H. M., Fleischmann, F., and Osswald,  
211 W. 2006. Involvement of *Phytophthora* species in the decline of European beech in  
212 Europe and the USA. *Mycologist* 19:159-166.
- 213 Jung, T., Orlikowski, L., Henricot, B., Abad-Campos, P., Aday, A. G., Aguín Casal, O.,  
214 Bakonyi, J., Cacciola, S. O., Cech, T., Chavarriaga, D., Corcobado, T., Cravador, A.,  
215 Decourcelle, T., Denton, G., Diamandis, S., Doğmuş-Lehtijärvi, H. T., Franceschini, A.,  
216 Ginetti, B., Green, S., Glavendekić, M., Hantula, J., Hartmann, G., Herrero, M., Ivic, D.,  
217 Horta Jung, M., Lilja, A., Keca, N., Kramarets, V., Lyubenova, A., Machado, H.,  
218 Magnano di San Lio, G., Mansilla Vázquez, P. J., Marçais, B., Matsiakh, I., Milenkovic,  
219 I., Moricca, S., Nagy, Z. Á., Nechwatal, J., Olsson, C., Oszako, T., Pane, A.,  
220 Paplomatas, E. J., Pintos Varela, C., Prospero, S., Rial Martínez, C., Rigling, D.,  
221 Robin, C., Rytkönen, A., Sánchez, M. E., Sanz Ros, A. V., Scanu, B., Schlenzig, A.,  
222 Schumacher, J., Slavov, S., Solla, A., Sousa, E., Stenlid, J., Talgø, V., Tomic, Z.,

223 Tsopelas, P., Vannini, A., Vettraino, A. M., Wenneker, M., Woodward, S., and Pérez-  
224 Sierra, A. 2016. Widespread *Phytophthora* infestations in European nurseries put  
225 forest, semi-natural and horticultural ecosystems at high risk of *Phytophthora* diseases.  
226 Forest Pathol. 46:134-163.

227 Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017.  
228 Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and  
229 repeat separation. Genome research 27:722-736.

230 Kruse, J., Doehlemann, G., Kemen, E., and Thines, M. 2017. Asexual and sexual morphs of  
231 Moesziomyces revisited. IMA Fungus 8:117-129.

232 Lamour, K. 2013. Phytophthora: a global perspective. Cabi.

233 Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. Nature  
234 Methods 9:357-359.

235 Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-  
236 MEM. arXiv preprint arXiv:1303.3997.

237 Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. O., and Borodovsky, M. 2005. Gene  
238 identification in novel eukaryotic genomes by self-training algorithm. Nucleic Acids  
239 Research 33:6494-6506.

240 Marçais, G., and Kingsford, C. 2011. A fast, lock-free approach for efficient parallel counting  
241 of occurrences of k-mers. Bioinformatics 27:764-770.

242 Mrazkova, M., Černý, K., Tomšovský, M., Strnadova, V., Gregorová, B., Holub, V., Pánek, M.,  
243 Havrdova, L., and Hejna, M. 2013. Occurrence of *Phytophthora multivora* and  
244 *Phytophthora plurivora* in the Czech Republic. Plant Prot Sci 49:155-164.

245 Ranallo-Benavidez, T. R., Jaron, K. S., and Schatz, M. C. 2020. GenomeScope 2.0 and  
246 Smudgeplot for reference-free profiling of polyploid genomes. Nature Communications  
247 11:1432.

- 248 Ruffner, B., Rigling, D., and Schoebel, C. N. 2019. Multispecies *Phytophthora* disease  
249 patterns in declining beech stands. *Forest Pathol.*
- 250 Seppey, M., Manni, M., and Zdobnov, E. M. 2019. BUSCO: Assessing Genome Assembly  
251 and Annotation Completeness. Pages 227-245 in: *Gene Prediction: Methods and*  
252 *Protocols*. M. Kollmar, ed. Springer New York, New York, NY.
- 253 Smit, A., Hubley, R., and Green, P. 2015. RepeatMasker Open-4.0. 2013–2015.
- 254 Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. 2006. Gene prediction in  
255 eukaryotes with a generalized hidden Markov model that uses hints from external  
256 sources. *BMC Bioinformatics* 7:62.
- 257 Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. 2008. Using native and syntenically  
258 mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 24:637-  
259 644.
- 260 Talas, F., and McDonald, B. A. 2015. Genome-wide analysis of *Fusarium graminearum* field  
261 populations reveals hotspots of recombination. *BMC Genomics* 16:996.
- 262 Thines, M., and Kamoun, S. 2010. Oomycete–plant coevolution: recent advances and future  
263 prospects. *Curr. Opin. Plant Biol.* 13:427-433.
- 264 Thines, M., Buaya, A., Ploch, S., Naim, Y. B., and Cohen, Y. 2020. Downy mildew of lavender  
265 caused by *Peronospora belbahrii* in Israel. *Mycol. Prog.* 19:1537-1543.
- 266 Tsykun, T., Prospero, S., Schoebel, C. N., Rea, A., and Burgess, T. I. 2022. Global invasion  
267 history of the emerging plant pathogen *Phytophthora multivora*. *BMC Genomics*  
268 23:153.
- 269 Vetukuri, R. R., Tripathy, S., Malar C, M., Panda, A., Kushwaha, S. K., Chawade, A.,  
270 Andreasson, E., Grenville-Briggs, L. J., and Whisson, S. C. 2018. Draft Genome  
271 Sequence for the Tree Pathogen *Phytophthora plurivora*. *Genome biology and*  
272 *evolution* 10:2432-2442.

273 Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A.,  
274 Zeng, Q., Wortman, J., Young, S. K., and Earl, A. M. 2014. Pilon: An Integrated Tool  
275 for Comprehensive Microbial Variant Detection and Genome Assembly Improvement.  
276 PLOS ONE 9:e112963.

277 Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and  
278 Andrews, S. 2015. HiCUP: pipeline for mapping and processing Hi-C data.  
279 F1000Research 4:1310.

280 Zhang, X., Zhang, S., Zhao, Q., Ming, R., and Tang, H. 2019. Assembly of allele-aware,  
281 chromosomal-scale autoploid genomes based on Hi-C data. Nature Plants 5:833-  
282 845.

283