

Genome assemblies for *Pyricularia* species and related genera isolated from diverse host plants

Avery Meyer,¹ Bram Dutch,² Ashab Ahmed,¹ Anna Baloh,³ Matt Bartholomai,¹ John Boggess,¹ Amelia Burnett,¹ Colin Carver,¹ Evan Courtwright,¹ Chase Eastham,^{4,5} Darrin Egan,¹ Knox Garland,¹ Kelly Claire Gray,¹ Audrey Harper,¹ Megan Johar,¹ Tanner Jones,¹ Leighanne Lyvers,¹ Jacob Marquez,¹ Summer McCune,¹ Matt Mitchell,¹ Carson Moylan,¹ Luke Olsen,¹ Tucker Overton,¹ Oluwatofunmi Oyetan,¹ Khumbo Phiri,¹ Kevin Ramirez,¹ Gavin Robinson,¹ Connor Talbot,¹ Bruno Athie Teruel,¹ Parker Thompson,⁴ Lark Wuetcher,³ Wesley Yang,¹ Andrew Tapia,² Jerzy Jaromczyk,² Mark Farman⁵

AUTHOR AFFILIATIONS See affiliation list on p. 4.

ABSTRACT We report the availability and preliminary analyses of genomic resources for 35 fungal strains representing *Pyricularia* and related genera. These data will provide new insights into genome expansions, as well as mechanisms driving colonization of new plant hosts.

KEYWORDS whole-genome shotgun, blast fungus

The ascomycete fungus *Pyricularia oryzae* causes blast diseases in rice and wheat but also infects diverse Gramineae (1, 2). Recently, we discovered novel host-specialized populations of *P. oryzae* in Brazil (3) and determined that admixture has significantly influenced host range expansion (4). The objective of this study was to develop genomic resources for understudied *Pyricularia* population/species and related genera to fill critical gaps in the species' evolutionary history.

Fungal strains, whose species affiliations were previously ascertained using sequence-based markers (3, 5–7), were provided by other researchers or isolated from leaf lesions. Cultures were established on oatmeal agar, single spored on 3% water agar, and grown in liquid complete medium with shaking (8). DNA was extracted using a phenol/chloroform method and purified with the Genomic Clean and Concentrator kit (Zymo Inc.) (8). Library construction used a 60-minute tagmentation step but otherwise followed the Illumina DNA Prep protocol (Illumina Inc.). Sequences (150-bp paired-end) were generated using the NovaSeq 6000 platform (Novogene Corp. Inc.).

Undergraduates in an “Applied Bioinformatics” class taught by Farman and Jaromczyk performed the data processing which is detailed in a GitHub repository (9). Sequence quality was assessed with FastQC v0.11.9 (10), and trimming/adaptor removal was performed using Trimmomatic v0.38 (11). VelvetOptimiser v2.2.6 was used to identify optimized assemblies by testing k-mer values between 69 and 129. Genome completeness was evaluated using BUSCO v5.7.0 (odb10_ascomycota database) (12). Annotation utilized MAKER v2.31.11 (13) to perform *de novo* gene predictions with SNAP v2.68.4 (14) and AUGUSTUS v3.4.0 (15) using a custom hidden Markov model. Gene modeling incorporated alignments to reference proteins from *P. oryzae* (9) and the nr database. Putative secreted proteins were identified using SignalP v5.0 (16). Single-nucleotide polymorphisms (SNPs) were called using iSNPcaller (17) based on BLAST alignments between each genome assembly and the B71 genome reference ([GCA_004785725.2](https://doi.org/10.1101/004785725.2)). Fasta alignment files were generated using custom scripts. Phylogenetic analyses were performed using RAxML 8.2.12 (18) with the GTRCAT substitution model, and bootstrap replications were determined using the autoMRE function. Trees were plotted using ggtree and ggtreeExtra (19, 20).

Editor Jason E. Stajich, University of California Riverside, Riverside, California, USA

Address correspondence to Mark Farman, mark.farman@uky.edu.

Avery Meyer and Bram Dutch contributed equally to this article. Author order was determined based on breadth of contribution (research plus writing).

The authors declare no conflict of interest.

Received 8 April 2025

Accepted 19 September 2025

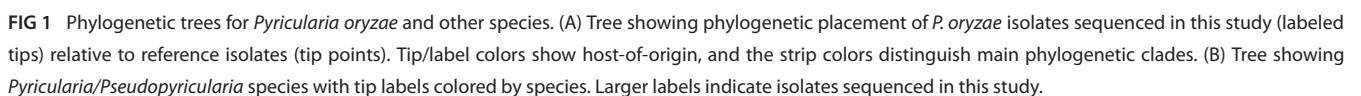
Published 24 November 2025

Darrin Egan, Jacob Marquez, and Kevin Ramirez have seen and approved an earlier version of the manuscript but could not be reached for approval of the final version.

Copyright © 2025 Meyer et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

TABLE 1 Genome assembly statistics

Sample name	Species	Host	Genome accession	SRA accession	Genome size (bp)	# Contigs	N50 (bp)	Coverage	%GC	BUSCO score	# Predicted proteins	# Secreted proteins	Reference
BP1	<i>Pyricularia oryzae</i>	<i>Stenotaphrum secundatum</i>	JASWIG0000000000	SRR23400923	44,269,911	8,402	38,801	28	50.7	96.6%	14,617	1,775	(3)
CD0086	<i>Pyricularia penniseticola</i>	<i>Pennisetum typhoides</i>	JAVHJQ0000000000	SRR23499064	37,718,974	9,300	9,584	41	50.8	97.8%	11,816	1,425	(6)
Cr9010	<i>Pseudopyricularia cyperi</i>	<i>Cyperus rotundifolia</i>	JASAOQ0000000000	SRR23498685	35,206,045	7,256	18,270	28	52.1	97.3%	10,465	1,223	(7)
CrA8401	<i>Pseudopyricularia javanica</i>	<i>Cyperus rotundifolia</i>	JARHUK0000000000	SRR23199414	32,715,629	5,506	17,480	53	53.6	98.0%	10,631	1,244	(7)
RN001	<i>Pyricularia zingibericola</i>	<i>Zingiber officinalis</i>	JASWJ0000000000	SRR23476220	37,907,990	6,067	24,913	57	49.9	96.2%	10,856	1,306	(6)
T21	<i>Pyricularia oryzae</i>	<i>Triticum aestivum</i>	JAVHTW0000000000	SRR23388169	43,329,010	2,709	88,629	24	49.9	98.0%	12,571	1,551	(3)
T29	<i>Pyricularia oryzae</i>	<i>Triticum aestivum</i>	JBMKKG0000000000	SRR27892070	42,317,924	1,784	64,165	37	50.0	98.1%	12,983	1,686	(3)
T6	<i>Pyricularia oryzae</i>	<i>Triticum aestivum</i>	JASWIK0000000000	SRR23501853	42,058,027	2,590	55,184	102	50.1	96.9%	13,375	1,781	(3)
U247	<i>Pyricularia oryzae</i>	<i>Festuca arundinacea</i>	JBCEWT0000000000	SRR27892087	44,649,226	6,941	66,992	87	50.0	98.0%	13,610	1,788	(3)
U248	<i>Pyricularia oryzae</i>	<i>Lolium multiflorum</i>	JBIZY0000000000	SRR28533131	43,867,247	4,388	67,659	84	50.0	97.6%	13,723	1,717	(3)
U249	<i>Pyricularia oryzae</i>	<i>Panicum sabolorum</i>	JBCARD0000000000	SRR28533132	40,914,617	3,426	104,435	88	50.8	98.4%	13,049	1,751	(3)
U270	<i>Pyricularia oryzae</i>	<i>Eleusine indica</i>	JAVHUH0000000000	SRR23567522	43,877,699	7,125	13,674	35	49.9	96.9%	14,580	1,793	(3)
U276	<i>Pyricularia oryzae</i>	<i>Leersia hexandra</i>	JAUKTU0000000000	SRR23388404	41,408,883	7,474	18,175	41	50.7	98.2%	13,832	1,759	(3)
UFVPY108	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JAVHTU0000000000	SRR23491016	39,330,571	6,352	17,944	21	51.7	98.1%	13,521	1,770	(3)
UFVPY110	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JAVHTV0000000000	SRR23498602	40,155,654	3,820	38,087	30	51.6	98.1%	13,487	1,764	(3)
UFVPY113	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JBEURV0000000000	SRR27893664	43,332,199	1,906	70,336	58	49.9	98.0%	13,243	1,753	(3)
UFVPY118	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JBPCDH0000000000	SRR23385021	40,034,794	4,807	25,618	71	51.7	98.4%	13,551	1,759	(3)
UFVPY119	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JAVHSR0000000000	SRR23386668	40,329,958	3,686	44,150	88	51.5	98.5%	13,488	1,783	(3)
UFVPY121	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JASWIF0000000000	SRR23387524	40,535,769	3,922	38,129	58	51.5	97.9%	13,580	1,784	(3)
UFVPY125	<i>Pyricularia grisea</i>	<i>Urochloa brizantha</i>	JASWII0000000000	SRR23456574	41,712,136	5,154	55,303	71	49.4	96.9%	11,942	1,551	(3)
UFVPY158	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JAVHTT0000000000	SRR23455553	36,015,173	11,335	5,656	36	52.0	97.9%	14,397	1,503	(3)
UFVPY166	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JBXLVL0000000000	SRR27893665	39,680,865	3,154	84,342	82	50.7	97.7%	12,750	1,722	(3)
UFVPY171	<i>Pyricularia urashimae</i>	<i>Urochloa brizantha</i>	JASWIH0000000000	SRR23388381	39,770,362	1,978	118,202	41	51.0	97.3%	12,320	1,766	(3)
UFVPY174	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JAVHTX0000000000	SRR23404000	37,458,192	1,835	57,082	52	51.8	98.0%	11,897	1,477	(3)
UFVPY184	<i>Pyricularia oryzae</i>	<i>Melinis repens</i>	JBCNVJ0000000000	SRR27893520	44,926,346	5,626	62,152	90	49.2	98.0%	13,231	1,749	(3)
UFVPY198	<i>Pyricularia oryzae</i>	<i>Urochloa brizantha</i>	JBMNVX0000000000	SRR27918925	38,944,860	1,755	98,072	82	51.0	98.2%	12,566	1,744	(3)
UFVPY202	<i>Pyricularia urashimae</i>	<i>Urochloa brizantha</i>	JBFYDN0000000000	SRR27891905	42,200,512	2,052	76,417	69	50.4	98.0%	12,729	1,770	(3)
UFVPY204	<i>Pyricularia urashimae</i>	<i>Urochloa brizantha</i>	JBCAWL0000000000	SRR27918216	41,269,248	1,859	77,297	56	50.4	98.0%	12,429	1,763	(3)
UFVPY210	<i>Pyricularia urashimae</i>	<i>Panicum maximum</i>	JBCARE0000000000	SRR28051826	41,100,774	1,892	80,230	64	50.5	98.0%	12,338	1,747	(3)
UFVPY218	<i>Pyricularia oryzae</i>	<i>Triticum aestivum</i>	JBCEVJ0000000000	SRR27891917	43,094,738	5,694	34,577	48	50.5	97.7%	13,682	1,782	(3)
UFVPY231	<i>Pyricularia oryzae</i>	<i>Melinis repens</i>	JBCASK0000000000	SRR27893663	44,428,863	2,078	101,661	67	49.2	98.3%	12,925	1,762	(3)
UFVPY232	<i>Pyricularia urashimae</i>	<i>Urochloa brizantha</i>	JBCEV5000000000	SRR27891887	41,802,667	1,976	83,582	63	50.5	98.0%	12,223	1,693	(3)
UFVPY63	<i>Pyricularia oryzae</i>	<i>Melinis repens</i>	JBBULU0000000000	SRR27892088	43,402,530	2,270	75,381	79	49.1	97.7%	12,646	1,676	(3)
UFVPY657	<i>Pyricularia urashimae</i>	<i>Urochloa brizantha</i>	JBDQYV0000000000	SRR29065219	43,027,175	3,161	66,750	80	50.3	97.9%	12,455	1,696	(3)
UFVPY677	<i>Pyricularia urashimae</i>	<i>Urochloa brizantha</i>	JBLIOR0000000000	SRR27917628	41,859,759	2,173	87,678	83	50.5	98.3%	12,143	1,692	(3)

10.1128/mra.00091-25 **3**

with approximately 12.5% being classified as candidate secreted proteins. A phylogenetic tree based on 327,024 SNPs (Fig. 1) indicated that 20 of 23 *P. oryzae* isolates clustered according to their respective hosts of origin—consistent with *P. oryzae*'s status as a host-specialized pathogen (3). The three exceptions were U249 and U276—reportedly from *Panicum* and *Leersia*, respectively—which clustered with *Luziola* pathogens inside the *Urochloa*3 clade, and BP1 from *Stenotaphrum*, which grouped with the *Setaria* pathogens. A tree built using 2,223,388 SNPs in the non-*P. oryzae* isolates (Fig. 1) also showed grouping patterns consistent with host-of-origin. UFVPY125 from *Digitaria* grouped with *Pyricularia grisea*, and the isolates from *Urochloa* and *Panicum maximum* belonged to *P. urashimae*. Due to the absence of reference genomes, the others occupied singleton branches whose lengths revealed high divergence from *P. oryzae* ($\geq 10\%$), supporting these strains' prior classifications as distinct *Pyricularia* or *Pseudopyricularia* species (*Pyricularia penniseticola*, *Pyricularia zingibericola*, *Ps. javanica*, and *Ps. cyperi*) (6, 21).

These 35 genomes provide valuable resources for studying *Pyricularia* evolution and will offer insights into genetic divergence, host adaptation, and genome dynamics.

ACKNOWLEDGMENTS

Students worked in a cloud-based training environment running on virtual machines in the University of Kentucky OpenStack cluster (<https://www.ccs.uky.edu/services-and-support/hardware/openstack-clusters/>). Genome assembly and BUSCO analysis were conducted on the Lipscomb and Morgan High Performance Computing clusters. We thank Bhushan Chitre and Vikram Gazula for assistance in setting up accounts and installing necessary software.

AUTHOR AFFILIATIONS

¹Computer Sciences Program, University of Kentucky, Lexington, Kentucky, USA

²Department of Computer Sciences, University of Kentucky, Lexington, Kentucky, USA

³Department of Horticulture, University of Kentucky, Lexington, Kentucky, USA

⁴Agricultural and Medical Biotechnology Program, University of Kentucky, Lexington, Kentucky, USA

⁵Department of Plant Pathology, University of Kentucky, Lexington, Kentucky, USA

AUTHOR ORCIDs

Mark Farman  <http://orcid.org/0000-0003-0410-2082>

AUTHOR CONTRIBUTIONS

Avery Meyer, Formal analysis, Investigation, Writing – original draft | Bram Dutch, Data curation, Formal analysis, Investigation, Software | Ashab Ahmed, Formal analysis, Investigation | Anna Baloh, Formal analysis, Investigation | Matt Bartholomai, Formal analysis, Investigation | John Boggess, Formal analysis, Investigation | Amelia Burnett, Formal analysis, Investigation | Colin Carver, Formal analysis, Investigation | Evan Courtwright, Data curation, Formal analysis, Investigation | Chase Eastham, Formal analysis, Investigation | Darrin Egan, Formal analysis, Investigation | Knox Garland, Formal analysis, Investigation | Kelly Claire Gray, Formal analysis, Investigation | Audrey Harper, Formal analysis, Investigation | Megan Johar, Formal analysis, Investigation | Tanner Jones, Formal analysis, Investigation | Leighanne Lyvers, Formal analysis, Investigation | Jacob Marquez, Formal analysis, Investigation | Summer McCune, Formal analysis, Investigation | Matt Mitchell, Formal analysis, Investigation | Carson Moylan, Formal analysis, Investigation | Luke Olsen, Formal analysis, Investigation | Tucker Overton, Formal analysis, Investigation | Oluwatofunmi Oyetan, Formal analysis, Investigation | Khumbo Phiri, Formal analysis, Investigation | Kevin Ramirez, Formal analysis, Investigation | Gavin Robinson, Formal analysis, Investigation | Connor Talbot, Formal analysis, Investigation | Bruno Athie Teruel, Formal analysis, Investigation | Parker Thompson,

Formal analysis, Investigation | Lark Wuetcher, Formal analysis, Investigation | Wesley Yang, Formal analysis, Investigation | Andrew Tapia, Conceptualization, Software, Supervision | Jerzy Jaromczyk, Project administration, Supervision | Mark Farman, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Raw read data and assembled genomes are deposited in NCBI under BioProject [PRJNA926786](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA926786) with the individual accessions listed in Table 1. Genome annotations (GFF files) and variant call data (VCF files) are accessible in Zenodo (22). Detailed methods and custom scripts are provided on GitHub (9).

REFERENCES

- Baudin M, Le Naour-Vernet M, Gladieux P, Tharreau D, Lebrun M-H, Lambou K, Leys M, Fournier E, Césari S, Kroj T. 2024. *Pyricularia oryzae*: lab star and field scourge. *Mol Plant Pathol* 25:e13449. <https://doi.org/10.1111/mpp.13449>
- Ou SH. 1985. Rice diseases. Commonwealth Agricultural Bureaux, Slough, UK.
- Ascari JP, Cazón LI, Rahnama M, Lamour K, Fernandes JMC, Farman ML, Ponte EMD. 2024. *Pyricularia* are mostly host-specialized with limited reciprocal cross-infection between wheat and endemic grasses in Minas Gerais, Brazil. *Phytopathology* 114:226–240. <https://doi.org/10.1094/PHYTO-01-23-0024-R>
- Rahnama M, Condon B, Ascari JP, Dupuis JR, Del Ponte EM, Pedley KF, Martinez S, Valent B, Farman ML. 2023. Recent co-evolution of two pandemic plant diseases in a multi-hybrid swarm. *Nat Ecol Evol* 7:2055–2066. <https://doi.org/10.1038/s41559-023-02237-z>
- Farman ML, Ascari JP, Rahnama M, Ponte EMD, Pedley KF, Martinez S, Fernandes JMC, Valent B. 2024. A re-evaluation of phylogenomic data reveals that current understanding in wheat blast population biology and epidemiology is obfuscated by oversights in population sampling. *Phytopathology* 114:220–225. <https://doi.org/10.1094/PHYTO-01-23-0025-R>
- Klaubauf S, Tharreau D, Fournier E, Groenewald JZ, Crous PW, de Vries RP, Lebrun M-H. 2014. Resolving the polyphyletic nature of *Pyricularia* (*Pyriculariaceae*). *Stud Mycol* 79:85–120. <https://doi.org/10.1016/j.simyco.2014.09.004>
- Borromeo ES, Nelson RJ, Bonman JM, Leung H. 1993. Genetic differentiation among isolates of *Pyricularia* infecting rice and weed hosts. *Phytopathology* 83:393. <https://doi.org/10.1094/Phyto-83-393>
- Farman M. 2025. Isolation of *Pyricularia* and preparation of DNA for Illumina Sequencing v1. protocols.io. <https://doi.org/10.17504/protocols.io.14egnr7pl5d/v1>
- Farman M. 2025. *Pyricularia* Genomes (v1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.16812621>
- Andrews S. 2010. A quality control tool for high throughput sequence data, v0.12.0. <https://github.com/s-andrews/FastQC>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Cantarel BL, Korf I, Robb SMC, 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. <https://doi.org/10.1101/gr.6743907>
- Korf I. 2004. Gene finding in novel genomes. *BMC Bioinformatics* 5:59. <https://doi.org/10.1186/1471-2105-5-59>
- Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res* 33:W465–7. <https://doi.org/10.1093/nar/gki458>
- Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37:420–423. <https://doi.org/10.1038/s41587-019-0036-z>
- Farman M. 2023. iSNPcaller. <https://github.com/drdna/iSNPcaller/tree/v.1.0.0>
- Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Xu S, Li L, Luo X, Chen M, Tang W, Zhan L, Dai Z, Lam TT, Guan Y, Yu G. 2022. Ggtree: a serialized data object for visualization of a phylogenetic tree and annotation data. *Imeta* 1:e56. <https://doi.org/10.1002/imt2.56>
- Xu S, Dai Z, Guo P, Fu X, Liu S, Zhou L, Tang W, Feng T, Chen M, Zhan L, Wu T, Hu E, Jiang Y, Bo X, Yu G. 2021. ggtreeExtra: compact visualization of richly annotated phylogenetic data. *Mol Biol Evol* 38:4039–4042. <https://doi.org/10.1093/molbev/msab166>
- Marin-Felix Y, Groenewald JZ, Cai L, Chen Q, Marincowitz S, Barnes I, Bensch K, Braun U, Camporesi E, Damm U, et al. 2017. Genera of phytopathogenic fungi: GOPHY 1. *Stud Mycol* 86:99–216. <https://doi.org/10.1016/j.simyco.2017.04.002>
- Meyer A, Dutch B, Ahmed A, Baloh A, Batholomai M, Boggess J, Burnett A, Carver C, Courtwright E, Eastham C, et al. 2025. Variant call and genome annotation files for the project “*Pyricularia* genome assemblies”. Zenodo. <https://doi.org/10.5281/zenodo.16812823>