

Whole genome shotgun sequences of *Magnaporthe oryzae* wheat blast isolates from Zambia

Batiseba Tembo¹, Suwilanji Sichilima¹, Rabson M. Mulenga¹, Adeline Harant², Vincent Were², Thorsten Langner², Pawan K. Singh³, Alison Bentley³, Nicholas J. Talbot², Sophien Kamoun², Joe Win²

¹Zambia Agricultural Research Institute, Mt. Makulu Central Research Station, Chilanga, Lusaka, Zambia

²The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, Norwich, United Kingdom

³International Maize and Wheat Improvement Center, Global Wheat Program, CIMMYT, El Bata'n, Texcoco, Mexico

We report whole genome shotgun sequences of 13 *Magnaporthe oryzae* (syn. *Pyricularia oryzae*) isolates collected in Zambia from 2018 to 2020 during outbreaks of wheat blast. The genome sequences add to the existing genomics resources for the study of the wheat blast fungus. The resource should prove critical in identifying the precise identity and potential origin of the Zambian population and in identifying variants of concern among this population. We provide the dataset as an open science resource as part of the OpenWheatBlast initiative. We welcome analyses and feedback from the community.

Introduction

The wheat blast fungus *Magnaporthe oryzae* (syn. *Pyricularia oryzae*), originated in Brazil, spread to Bangladesh in 2016 with damaging effects on wheat production in that country (Islam, et al., 2016). It has persisted in Bangladesh and has been detected every year since 2016. More recently in 2018, the fungus has been detected for the first time in Africa, more precisely in Zambia, as reported in September 2019 (Tembo et al., 2020). To monitor the spread of the fungus, we recently described a set of oligonucleotide primers flanking 84 single nucleotide polymorphisms (SNPs) in conserved genes that can differentiate between blast fungus lineages (Tembo et al., 2021). We used a multiplex amplicon sequencing approach to determine the SNP genotype of 186 isolates of *M. oryzae*, including 13 isolates from Zambia (Tembo et al., 2021). Phylogenetic and clustering analyses of these genotype data revealed that the wheat blast fungus detected in Zambia belongs to the “B71” clonal lineage that has caused epidemics in Bangladesh and traces its origin to South America (Latorre et al., 2021; Win et al., 2021).

To validate the genotyping results and gain a deeper understanding of the genome evolution and potential adaptative variation in the B71 clonal lineage across different geographic locations, we sequenced the genomes of 13 *M. oryzae* isolates collected in Zambia over three years (2018-2020). The genome sequences add to existing genomics resources for the study of the wheat blast fungus, including a recently released set of genome sequences by Were et al. (2021).

Results and Conclusions

High molecular weight genomic DNA was isolated from 13 single-spore fungal isolates (Table 1) using the protocol described by Schwessinger (2016). We used Illumina sequencing service provided by Novogene (Cambridge, UK) to obtain 150 bp paired-end short sequence reads with >65x coverage of the ~43 Mbp *M. oryzae* genome (Table 1). The reads were of high quality (error rates below 0.03%, and 93.6% of reads had better than or equal Q30 (Illumina per base sequence quality score) and average coverage of ~72.5x (65.1-83.7x) (Table 1). We conclude that the sequence reads are of high quality and can be used for further analysis.

Data availability

The sequence data are provided as an open science resource via European Nucleotide Archive (ENA). This project is part of the [OpenWheatBlast](#) initiative and the [Zenodo OpenWheatBlast Community](#).

Statement on continuing work

The datasets are being shared prior to formal publication and should be considered preliminary. We encourage and welcome analyses and feedback from the community. Please use the [Zenodo OpenWheatBlast Community](#) to share any analysis reports or get in touch with JW and SK. We will post our own analyses of this dataset as soon as completed.

References

- Islam, M.T., Croll, D., Gladioux, P., Soanes, D.M., Persoons, A., Bhattacharjee, P., Hossain, M.S., Gupta, D.R., Rahman, M.M., Mahboob, M.G., Cook, N., Salam, M.U., Surovy, M.Z., Sancho, V.B., Maciel, J.L., NhaniJunior, A., Castroagudin, V.L., Reges, J.T., Ceresini, P.C., Ravel, S., Kellner, R., Fournier, E., Tharreau, D., Lebrun, M.H., McDonald, B.A., Stitt, T., Swan, D., Talbot, N.J., Saunders, D.G., Win, J., and Kamoun, S. (2016). Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. *BMC Biol* 14, 84.
- Tembo, B., Mahmud, N.U., Paul, S.K., Asume, S., Harant, A., Langner, T., Reyes-Avila, C.S., Chanclud, E., Were, V., Sichilima, S., Mulenga, R.M., Gupta, D.R., Meheub, M.S., Muzahid, A.N.M., Rabby, M.F., Singh, P.K., Bentley, A., Tosa, Y., Croll, D., Lamour, K., Islam, T., Talbot, N.J., Kamoun, S., and Win, J. (2021). Multiplex amplicon sequencing dataset for genotyping pandemic populations of the wheat blast fungus [Data set]. Zenodo <http://doi.org/10.5281/zenodo.4605959>.
- Tembo, B., Mulenga, R.M., Sichilima, S., M'Siska K, K., Mwale, M., Chikoti, P.C., Singh, P.K., He, X., Pedley, K.F., Peterson, G.L., Singh, R.P., and Braun, H.J. (2020). Detection and characterization of fungus (*Magnaporthe oryzae* pathotype Triticum) causing wheat blast disease on rain-fed grown wheat (*Triticum aestivum* L.) in Zambia. *PLOS One* 15, e0238724.
- Latorre, S.M., and Burbano, H.A. (2021). The emergence of wheat blast in Zambia and Bangladesh was caused by the same genetic lineage of *Magnaporthe oryzae*. Zenodo <http://doi.org/10.5281/zenodo.4619405>.
- Were, V., Mwongera, D.T., Soanes, D.M., Shrestha, R.-K., Ryder, L., Foster, A.J., et al. (2021). Genome sequences of sixty *Magnaporthe oryzae* isolates from multiple host plant species [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4627043>.
- Win, J., Malmgren, A., Langner, T., and Kamoun, S. (2021). A pandemic clonal lineage of the wheat blast fungus. Zenodo <http://doi.org/10.5281/zenodo.4618522>.
- Schwessinger, B. (2016). High quality DNA from fungi for long read sequencing e.g. PacBio. protocol.io <https://dx.doi.org/10.17504/protocols.io.evmbe46>.

Table 1. *Magnaporthe oryzae* isolates from Zambia used for genome sequencing and summary from Illumina sequencing runs

Sample Name	Wheat variety collected from ²	Collection date	Location	No. raw reads ³	Fold cov. ⁴	Library name	Error (%) ⁵	Q20 (%) ⁵	Q30 (%) ⁵	ENA Reads accession ⁶
ZMW18_06 ¹	RPWYT-35	Mar, 2018	Mpika, Zambia	21156786	74.4	FDSW210069331	0.03	97.69	93.49	ERR5525603
ZMW18_08 ¹	COUCAL	Mar, 2018	Mpika, Zambia	22292720	76.7	FDSW210069333	0.03	97.64	93.39	ERR5525605
ZMW18_10 ¹	HON-9	Mar, 2018	Mpika, Zambia	21856972	76.7	FDSW210069335	0.03	97.83	93.82	ERR5525607
ZMW18_11 ¹	KWALE	Mar, 2018	Mpika, Zambia	21136340	74.4	FDSW210069336	0.03	97.77	93.69	ERR5525608
ZMW19_09	RNWYT-18	Mar, 2019	Mt. Makulu, Zambia	21874332	76.7	FDSW210069334	0.03	97.75	93.66	ERR5525606
ZMW19_13	COUCAL	Mar, 2019	Mt. Makulu, Zambia	20061624	69.8	FDSW210069337	0.03	97.82	93.77	ERR5525609
ZMW19_17	29TH HRSWN-8	Mar, 2019	Mt. Makulu, Zambia	23981696	83.7	FDSW210069340	0.03	97.72	93.59	ERR5525613
ZMW20_03	COUCAL	Mar, 2020	Mpika, Zambia	18689410	65.1	FDSW210069329	0.03	97.79	93.75	ERR5525601
ZMW20_04	RAWYT-1-26	Mar, 2020	Mpongwe, Zambia	18887346	65.1	FDSW210069330	0.03	97.76	93.66	ERR5525602
ZMW20_07	19TH ESBWYT-45	Mar, 2020	Mpongwe, Zambia	21091332	74.4	FDSW210069332	0.03	97.65	93.42	ERR5525604
ZMW20_14	20TH FHBSN-50	Mar, 2020	Mt. Makulu, Zambia	18766188	65.1	FDSW210069338	0.03	97.83	93.82	ERR5525610
ZMW20_15	19TH EBSWYT-18	Mar, 2020	Mpongwe, Zambia	20997088	72.1	FDSW210069339	0.03	97.73	93.58	ERR5525611
ZMW20_16	19TH EBSWYT-82	Mar, 2020	Mpika, Zambia	19548178	67.4	FDSW210069328	0.03	97.78	93.74	ERR5525612

¹ These samples were previously reported by Tembo et al. (2020) PLOS ONE, 15, e0238724.

² *M. oryzae* samples were collected by Dr. Batiseba Tembo (Zambia Agricultural Research Institute).

³ Sequencing was performed by Novogene using Illumina Novaseq 6000 in lane 4 of a flowcell (ID: HWGV3DSXY). Paired-end reads with 150 bp length were acquired to achieve >60x coverage of the *M. oryzae* genome

⁴ Fold coverage was calculated by dividing the product from (no. raw reads X 150 bp) by approximate genome size of *M. oryzae* (4.3 Mbp)

⁵ Quality scores were provided by Novogene

⁶ Sequence reads are publicly available from European Nucleotide Archive (ENA) under Study accession PRJEB43836.