

# Whole-genome analyses of 286 *Magnaporthe oryzae* genomes suggest that an independent introduction of a global pandemic lineage is at the origin of the Zambia wheat blast outbreak

Daniel Croll<sup>1</sup>

<sup>1</sup>Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, Switzerland.  
<https://pathogen-genomics.org> - [@danielcroll](#)

**The emergence of agricultural pests is a severe threat to global food security. The recent wheat blast outbreak in Bangladesh had devastating consequences for the country. More recently, a wheat blast outbreak in Zambia was assigned to the same tight group of genotypes originally introduced from South America into Bangladesh. Here, I analyzed publicly available whole-genome sequencing datasets shared by the OpenWheatBlast initiative. The aim is to contrast the Zambian outbreak strains with genome-wide differentiation of *Magnaporthe oryzae* infecting different hosts. The genome-wide SNP analyses show that the Zambian outbreak strains are indeed highly similar to the strains responsible for the outbreak in Bangladesh as recently reported. The genomic data suggest that the Zambian outbreak may have originated from South America rather than a secondary introduction from the South Asian gene pool. The genome analyses confirm that wheat blast has emerged in recent years as a pandemic lineage across three continents. I provide the analyses as an open science contribution to the OpenWheatBlast initiative.**

## Introduction

Wheat blast caused by the fungus *Magnaporthe oryzae* (syn. *Pyricularia oryzae*) was first described in Brazil where it affects most wheat-producing areas. A recent introduction to Bangladesh was caused most likely by a singular lineage (Islam et al. 2016). More recently, a report from Zambia indicated that wheat blast has also become established in Africa (Tembo et al. 2020). A genotyping assay of 84 SNPs enabled rapid screening of Zambian isolates to contrast genotypes with wheat blast from South America and Bangladesh (Tembo et al. 2021a). Phylogenetic analyses of the SNP dataset revealed that the Zambian strain genotypes are indistinguishable from genotypes recovered in Bangladesh and share high similarity to Brazilian genotypes (Latorre et al., 2021; Win et al., 2021). This raises the worrying prospect of a globally successful lineage of wheat blast potentially spreading across Africa and other continents.

To ascertain the phylogenetic placement of the Zambian outbreak strains, Tembo et al (2021b) sequenced 13 *M. oryzae* genomes from Zambian isolates. In addition to these very recent publicly shared genomic resources, Were et al. (2021) made a broad range of *M. oryzae* genomes from Bangladesh and elsewhere available. Here, I combine these latest whole-genome datasets with publicly available *M. oryzae* genomes to identify the most likely invasion route.

## Methods

All Illumina whole-genome sequencing datasets were retrieved from the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>; Table 1). Of particular relevance for my analyses were two sets of recently shared datasets from Tembo et al. (2021b) covering 13 *M. oryzae* genomes from Zambian isolates and the dataset shared by Were et al. (2021) covering a range of wheat and rice blast isolates including genomes of Bangladesh strains.

Procedures for read trimming, alignment to the *M. oryzae* 70-15 reference genome, SNP discovery and SNP filtering, as well as phylogenetic network construction are detailed on <https://github.com/crolllab/wheat-blast>.

## Results and Conclusions

1.) The read alignment and variant calling steps produced a total of 1,454,322 high-quality SNPs. The SNP dataset included all known lineages of *M. oryzae* and the outgroup *M. grisea* for a total of 286 genomes (Table 1).

2.) The reconstruction of phylogenetic distances within *M. oryzae* was based on SplitsTree 5.0 (Huson et al. 2006), which implements a NeighborNet algorithm accounting for reticulation (*i.e.* a network structure). Given the significant degree of recombination in some lineages, a network representation provides a more accurate assessment of genetic distances.

3.) The phylogenetic network resolves clearly all major lineages of *M. oryzae* including genotypes typically associated with rice and wheat blast (Figure 1).

4.) The recently shared whole-genome sequencing datasets from Bangladesh are part of the OpenWheatBlast initiative (Were et al. 2021). The genomes from Bangladesh strains collected in 2016 and 2017 cluster tightly including a Brazilian genotype (Figure 2). The high level of resolution pinpoints the most likely donor genotypes being from South America. This confirms the previous genotyping approach based on field pathogenomics (*i.e.* transcriptome sequencing; Islam et al. 2016). However, thanks to whole-genome analyses and a deeper genomic sampling of Brazilian diversity revealed a strain collected in 2012 in Rio Grande do Sul (isolate 12.1.181) as sharing the highest known similarity to the Bangladesh genotypes (Figure 2).

5.) Zambian wheat blast genotypes cluster as a singular lineage with a direct connection to the gene pool of Brazilian wheat blast (Figure 2). This suggests that the introduction occurred possibly directly from South America into Zambia instead of originating from the South Asian wheat blast gene pool. Both options were explicitly mentioned by the recent SNP assay analyses (Latorre et al., 2021; Win et al., 2021).

6.) The genome sequencing data analyses strongly support the identification of a global pandemic lineage now present on three continents (Latorre et al., 2021; Win et al., 2021). Urgent efforts to contain further spread of this lineage are needed.

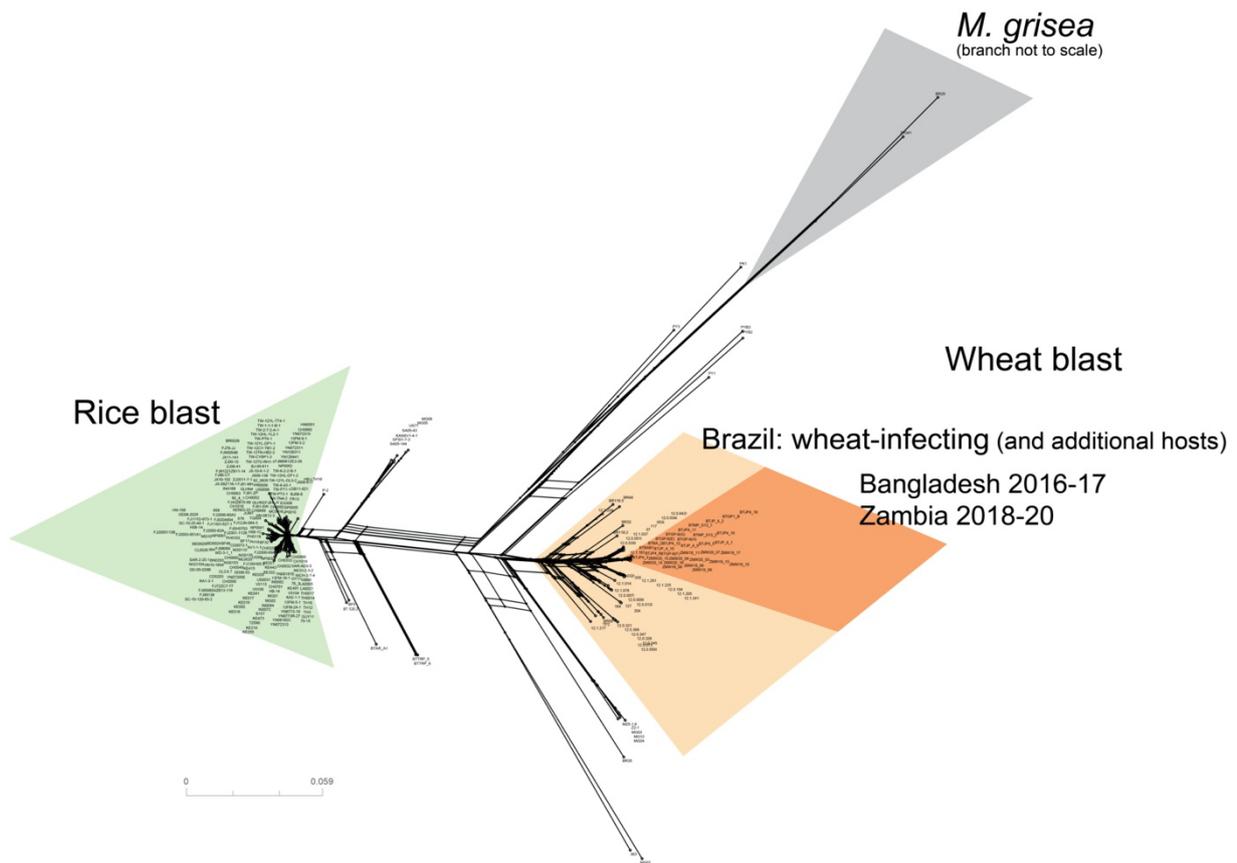


Figure 1: Phylogenetic network (SplitsTree) of genome-wide polymorphism of the known genetic diversity of *Magnaporthe oryzae* including wheat and rice blast lineages. The total SNP dataset was subset to a single randomly selected SNP every 1000 bp to reduce computational load and improve the representativity of the entire genome for phylogenetic analyses. The subsampling retained 34,458 SNPs for phylogenetic analyses.

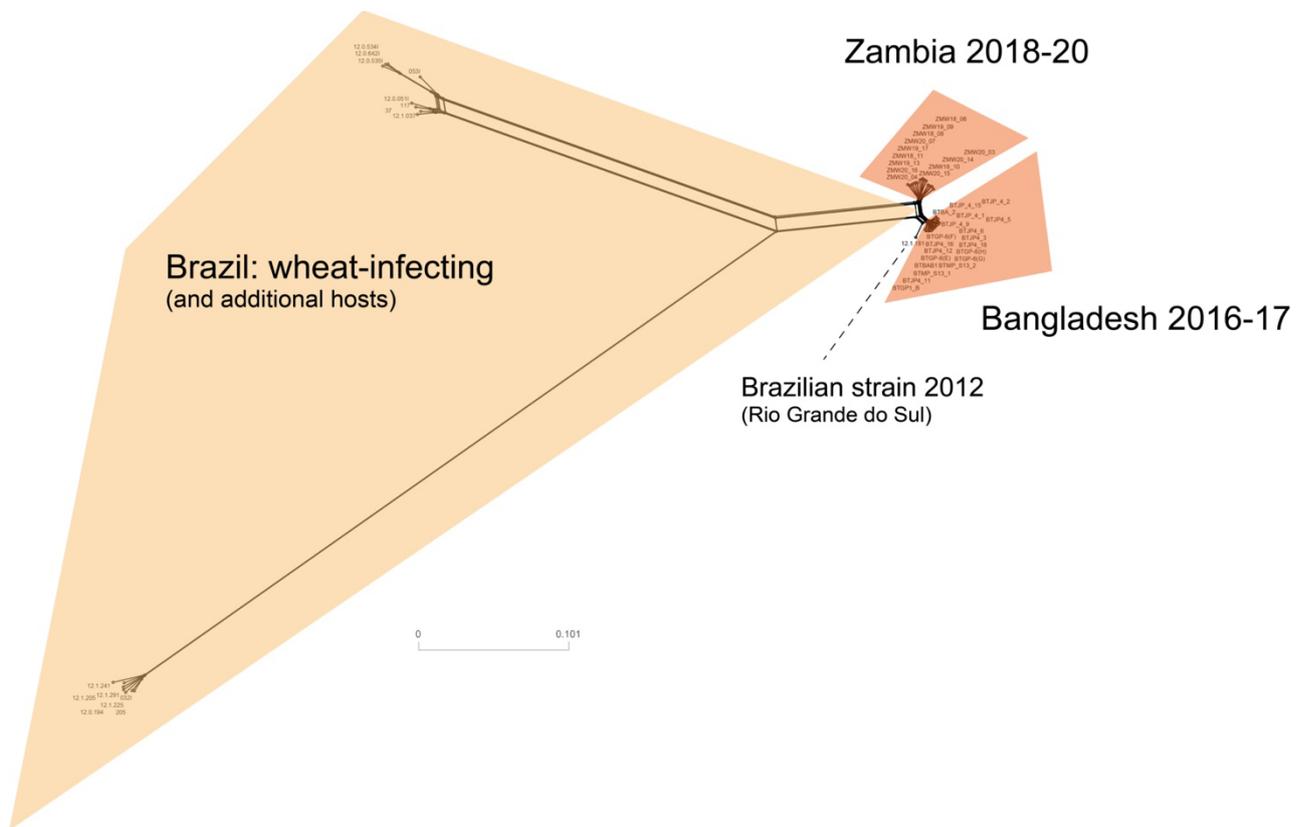


Figure 2: Phylogenetic network (SplitsTree) of genome-wide polymorphism focusing only on the genotypes from Zambia, Bangladesh, and the closest known genotypes from Brazil. The total SNP dataset included 51,121 SNPs (no subsampling).

## Acknowledgements

None of the described work above would have been possible without the outstanding openness of the worldwide community of researchers contributing to OpenWheatBlast - an initiative created by among others by Tofazzal Islam and Sophien Kamoun in 2016. I am particularly grateful for the recent whole-genome sequencing datasets made available by Vincent Were and Batiseba Tembo with colleagues.

## References

- Huson, D. H. and Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies, *Mol. Biol. Evol.*, 23(2):254-267.
- Islam, M.T., Croll, D., Gladieux, 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.
- 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 <https://doi.org/10.5281/zenodo.4619405>
- Tembo, B., Sichilima, S., Mulenga, R.M., Harant, A., Were, V., Langner, T., Singh, P.K., Bentley, A., Talbot, N.J., Kamoun, S., and Win, J. (2021b). Whole genome shotgun sequences of *Magnaporthe oryzae* wheat blast isolates from Zambia [Data set]. Zenodo <https://doi.org/10.5281/zenodo.4637175>
- Tembo, B., Mahmud, N.U., Paul, S.K., Asuke, 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. (2021a). Multiplex amplicon sequencing dataset for genotyping pandemic populations of the wheat blast fungus [Data set]. Zenodo <https://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.
- 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. <https://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 <https://doi.org/10.5281/zenodo.4618522>