

# Whole-Genome SNP Analysis for Typing the Pandemic Pathogen *Fusarium Graminearum* Sensu Stricto

**Tomasz Kulik** (✉ [tomaszkulik76@gmail.com](mailto:tomaszkulik76@gmail.com))

University of Warmia and Mazury in Olsztyn <https://orcid.org/0000-0002-9951-0873>

**Tomasz Molcan**

Institute of Biochemistry and Biophysics Polish Academy of Sciences: Instytut Biochemii i Biofizyki Polskiej Akademii Nauk

**Grzegorz Fiedorowicz**

University of Warmia and Mazury in Olsztyn: Uniwersytet Warmińsko-Mazurski w Olsztynie

**Anne van Diepeningen**

Wageningen University and Research Wageningen Plant Research

**Alexander Stakheev**

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry Russian Academy of Sciences: FBGUN Institut bioorganiceskoj himii im akademikov M M Semakina i U A Ovcinnikova Rossijskoj akademii nauk

**Kinga Treder**

University of Warmia and Mazury in Olsztyn: Uniwersytet Warmińsko-Mazurski w Olsztynie

**Jacek Olszewski**

University of Warmia and Mazury in Olsztyn: Uniwersytet Warmińsko-Mazurski w Olsztynie

**Katarzyna Bilska**

University of Warmia and Mazury in Olsztyn: Uniwersytet Warmińsko-Mazurski w Olsztynie

**Marco Beyer**

Luxembourg Institute of Science and Technology Environmental Research and Innovation Research Department

**Matias Pasquali**

University of Milan: Università degli Studi di Milano

**Sebastian Stenglein**

CONICET Tandil

---

## Research

**Keywords:** *Fusarium graminearum* sensu stricto, *F. graminearum* complex, whole-genome sequencing

**Posted Date:** January 17th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1249675/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Recent improvements in microbiology and molecular epidemiology were largely stimulated by whole-genome sequencing (WGS), which provides an unprecedented resolution in discriminating highly related genetic backgrounds. WGS is becoming the method of choice in epidemiology of fungal diseases, but its application is still in a pioneer stage, mainly due to the limited number of available genomes. Fungal pathogens often belong to complexes composed of numerous cryptic species. Detecting cryptic diversity is fundamental to understand the dynamics and the evolutionary relationships underlying disease outbreaks. In this study, we explore the value of whole-genome SNP analyses in identification of the pandemic pathogen *Fusarium graminearum* sensu stricto (*F.g.*). This species is responsible for cereal diseases and negatively impacts grain production worldwide. The fungus belongs to the monophyletic fungal complex referred to as *F. graminearum* species complex including at least sixteen cryptic species, among them a few may be involved in cereal diseases in certain agricultural areas. We analyzed WGS data from a collection of 99 *F.g.* strains and 33 strains representing all known cryptic species belonging to the FGSC complex. As a first step, we performed a phylogenomic analysis to reveal species-specific clustering. A RAxML maximum likelihood tree grouped all analyzed strains of *F.g.* into a single clade, supporting the clustering-based identification approach. Although, phylogenetic reconstructions are essential in detecting cryptic species, a phylogenomic tree does not fulfill the criteria for rapid and cost-effective approach for identification of fungi, due to the time-consuming nature of the analysis. As an alternative, analysis of WGS information by mapping sequence data from individual strains against reference genomes may provide useful markers for the rapid identification of fungi. We provide a robust framework for typing *F.g.* through the web-based PhaME workflow available at EDGE bioinformatics. The method was validated through multiple comparisons of assembly genomes to *F.g.* reference strain PH-1. We showed that the difference between intra- and interspecies variability was at least two times higher than intraspecific variation facilitating successful typing of *F.g.* This is the first study which employs WGS data for typing plant pathogenic fusaria.

## Introduction

Fungi are among the most diverse group of eukaryotic organisms on earth and include a wide range of taxonomic groups with various morphological and phenotypic characters, ecologies and life cycles (Hyde et al. 2019; Naranjo-Ortiz and Gabaldón 2020). With their complex and often cryptic nature, fungi are widespread from polar to tropical habitats acting as decomposers of organic matter and symbionts of algae and plants (Frac et al. 2018; Bonfante 2019). Mushrooms are used to make foods rich in mycoproteins, dietary fibers, vitamins and antioxidants (Valverde et al. 2015; Ulian et al. 2020). Fungi are also well known as producers of a vast array of secondary metabolites reflecting their adaptation to a diversity of environments (Boruta 2018; Tralamazza et al. 2019). Besides their beneficial roles for food production, medicine and other industries, fungi can be harmful to the health of humans and animals as well as for the environment (Wu and Hao 2019). Around 625 fungi were found to affect vertebrates, and approximately 700 of them could be associated with humans, either as commensals, microbiome

community members or as pathogens (Fisher et al. 2020; de Hoog et al. 2021). From about 100,000 known species, around 20% are involved in 70–80% of plant diseases (Persley 1993; Ray et al. 2017). Besides, many fungi are responsible for significant qualitative loss of food and feed by mycotoxin contamination (Marin et al. 2013; Ukwuru et al. 2017), which has negative effects on humans and animals (Zain 2011).

Undoubtedly, taxonomic information is crucial for understanding fungal diversity and evolution. Taxonomy provides a common language for scientists and the regulatory community (Raja et al. 2017a; Raja et al. 2017b), and is especially critical in clinical, environmental, food and biological applications (Kulik et al. 2020). Classical taxonomic methods enable for the determination of the species based on morphological characteristics. However, laborious culture-based methods are often neglected by scientists mostly due to morphological/phenotypic variability and pleomorphism (Capote et al. 2012; Sharma and Sharma 2016). The field of mycology has come to accept DNA-based technology, which has long been used for identification purposes. The most prevalent approaches are based on polymerase chain reaction (PCR), which targets specific DNA regions allowing species determination. Numerous molecular approaches have been developed to detect races, formae speciales, mating types and mycotoxin genotypes among different fungal taxa. However, the necessity for multiple laboratory tests for fungi remains laborious, expensive, and time-consuming (Kulik et al. 2020).

More universal DNA barcoding promises to be a valuable approach for high-throughput identification of fungi. Defining species through DNA barcoding involves sequence comparison of short and unique DNA barcodes (ca. 400–800 bp) which are amplified and sequenced with the use of universal primers. Today, the internal transcribed spacer (ITS) region in the ribosomal RNA gene cluster has been validated as the best DNA barcode marker for fungi. However, its reliability as taxonomic marker for fungi has been criticized, because of: (i) the presence of divergent intragenomic ITS sequences in several fungal lineages (Kiss 2012) and (ii) lack of sufficient sequence polymorphism facilitating recognition of some closely related species (Seifert 2009; Kiss 2012). For many phytopathogenic genera, e.g., *Alternaria*, *Botryosphaeria*, *Cercospora*, *Diaporthe*, and *Fusarium* ITS sequence alone cannot be used to identify most species (Seifert 2009; Sharma et al. 2015; Kashyap et al. 2017).

A number of fungi previously considered to be a single species are species complexes composed of multiple cryptic species. Identification of these diverged species is usually achieved through multilocus sequence typing (MLST). This approach involves the amplification of housekeeping genes by PCR, followed by DNA sequencing of amplicons and defining an allelic profile corresponding to a Sequence Type (ST) (Taylor and Fisher 2003). The use of MLST allows us to detect the emergence of novel genotypes or sequence variants that can be additionally stored in a publicly accessible database. The obvious limitation of MLST can be linked to its failure to fully identify genomic diversity. Analysis of a limited number of loci in the genome is often not discriminative enough to be useful for outbreak detection (Teatero et al. 2015; Uelze et al. 2020). In addition, MLST analysis of various fungal complexes appears to be more complicated than bacterial pathogens because of the less conserved MLST loci which require multiple genus- or even species complex-specific primers for successful amplification. The

choice of suitable primer sets is thereby critical and often dependent on the ITS sequence (Wickes and Romanelli 2020).

Recent advances in genomic sequence analysis and SNP discovery have empowered researchers to explore microbial diversity in more detail. The increasing amount of sequenced fungal genomes opens up the availability to better understand the tremendous diversity of fungi and their evolution. Fungal genomes are relatively small in comparison to plant and animal genomes, but vary from several to nearly 1000 megabases (Mb) (Stajich 2017). Genomes of fungi are dynamic in nature and rapid progress in high throughput sequencing methods pinpoint to different mechanisms of genome evolution (Mohanta and Bae 2015; Priest et al. 2020). Among them gene duplication, polyploidy, chromosomal rearrangements, interspecific hybridization, introgression and horizontal gene transfer are considered to be the main mechanisms shaping genetic diversity of fungi (Albertin and Marullo 2012; Priest et al. 2020). In addition, fungal genomics demonstrate significant divergence between fungal genomes that can be linked to host/niche specialization and lifestyles, and allows the research community to gain insight into the fundamental aspects of fungal biology (Gladieux et al. 2014). Nowadays, the rapid generation of genomic data contributes significantly to our understanding of fungal diseases, fungicide resistance, toxicology, and to the discovery of biosynthetic gene clusters that underlie the capacity to produce secondary metabolites (Tralamazza et al. 2019; Pasquali et al. 2020; Rampersad 2020). The whole-genome strategies allow for determination of fungal species, support fungal taxonomy and dispersal patterns (Araujo 2014; Araujo and Sampaio-Maia 2018). Among these approaches, genome-wide SNPs provide increased resolution to evaluate fungal diversity (Araujo 2014; Araujo and Sampaio-Maia 2018). SNPs display relatively low mutation rates and are evolutionarily stable (Leekitcharoenphon et al. 2012). The uniform rate of accumulation of selectively neutral SNPs makes them ideal to establish divergence between species as well as strains (Schork et al. 2000; Filliol et al. 2006; Shakya et al. 2020). Usually, SNPs are identified by mapping sequence data from individual strains against a closely related reference genome. Such analysis is based on a set of generated core SNPs, which is covered by all studied genomes including reference genome. SNP distance matrices determined from combinations of pairwise SNP distances enable further phylogenetic analysis (Uelze et al. 2020). It is worth noting, however, that the application of genome-based approaches in mycology is still in a pioneer stage. Validation of diagnostic tools requires the incorporation of genomic data covering a large fraction of cryptic diversity within the groups of morphologically indistinguishable species. Unfortunately, many fungal complexes still lack sufficient number of genomes to fully explore their efficiency for diagnostic purposes.

Our previous large scale sequencing project was launched in order to obtain genome sequences from a large collection of strains of the pandemic pathogen *Fusarium graminearum* sensu stricto (*F.g.*) (Wyrębek et al. 2021). This cryptic species is mainly responsible for two cereal diseases: Fusarium Head Blight (FHB) of wheat and barley and Fusarium Ear Rot (FER) of maize. Both diseases have led to major economic losses for the cereal-based feed and food supply chains worldwide (van der Lee et al. 2015). *F.g.* belongs to the *F. graminearum* species complex (FGSC), which includes at least sixteen cryptic species ((Sarver et al. 2011)), among them a few may be involved in cereal diseases in certain

agricultural areas (van der Lee et al. 2015). Determination of cryptic diversity within field populations enables us to better understand their dynamics and the evolutionary relationships underlying disease outbreaks (van der Lee et al. 2015; Wyrębek et al. 2021)). We have previously explored the value of mitochondrial sequences for diagnostic purposes of *F.g.* We showed that mitochondrial-based SNP analysis are useful for typing most, but not all, strains of *F.g.* We also underlined the limitations of clustering-based identification approaches for FGSC species using mitochondrial sequences (Wyrębek et al. 2021).

In this paper, sequence data from previous WGS project were first used to assemble genomes from a total of 99 *F.g.* strains and 33 strains representing all known cryptic species from the FGSC complex. We constructed a phylogenomic tree to demonstrate species-specific clustering for species recognition. We also show that the web-based calculation of the number of SNPs via comparison to reference PH-1 strain of *F.g.* provides an easy method of typing *F.g.* This is the first report on typing plant pathogenic fusaria through WGS data.

## Materials And Methods

### Fungal Strains and Genome Assembly

In total, 136 strains were analyzed in this study, with 99 *F.g.* strains and 33 strains representing all known cryptic species from the FGSC complex: *Fusarium acaciae-mearnsii* (three strains), *Fusarium aethiopicum* (one strain), *Fusarium asiaticum* (three strains), *Fusarium austroamericanum* (two strains), *Fusarium boothii* (four strains), *Fusarium brasiliicum* (one strain), *Fusarium cortadariae* (three strains), *Fusarium gerlachii* (two strains), *Fusarium louisianense* (two strains), *Fusarium meridionale* (two strains), *Fusarium mesoamericanum* (one strain), *Fusarium nepalense* (two strains), *Fusarium ussurianum* (four strains), *Fusarium vorosii* (two strains), and strain CBS 123663 (NRRL34461), which lacks a Latin binomial. In addition, four single strains from the closely related morphospecies *F. culmorum*, *F. pseudograminearum*, *F. sambucinum* and *F. venenatum* were incorporated into SNP analysis. Detailed information on all strains used in this study are included in Supplementary File S1.

### Whole-genome Sequencing And Genome Assembly

Whole fungal genomic DNA was sequenced by Macrogen (Seoul, South Korea). Libraries were prepared using the KAPA HyperPlus Kit (Roche Sequencing Solutions, Pleasanton, CA, USA). An Illumina HiSeq X Ten was used to sequence the genomes using a paired-end read length of 2x150 bp with an insert size of 350 bp. The sequencing quality was assessed via FastQC (ver. 0.11.9) Kliknij lub naciśnij tutaj, aby wprowadzić tekst.(Andrews 2010). Low-quality reads were trimmed using Trimmomatic (v.0.36) (Bolger et al. 2014) and the genome was assembled via SPAdes (v.3.13.2) (Nurk et al. 2013). The completeness of the assembly was evaluated using BUSCO (Simão et al. 2015). Genome statistics including genome length, GC content and N50 values were calculated using QUAST v. 5.0.2 (Mikheenko et al. 2018). The

project was submitted to NCBI BioProject under accession: PRJNA677929. GenBank accession numbers are listed in Supplementary File S1.

## Phylogenomic Analysis

Locally installed PhaME (Shakya et al. 2020) was used to reconstruct phylogeny of the strains from FGSC. The phylogenomic tree was inferred using the RAxML maximum likelihood method. The *F. pseudograminearum* strain CS3096 was used as an outgroup.

## SnP Analysis

To extract core SNPs, we used the whole-genome SNP-based phylogeny tool PhaME, which is integrated into the EDGE bioinformatics platform (Li et al., 2017) (<https://edgebioinformatics.org/>). PhaME extracts SNPs from the core genome identified through comparison of assembly genome to the reference PH-1 strain of *F.g.* (GenBank assembly accession: GCA\_900044135.1). The total number of SNPs is counted as the number of positions that are variable between two genomes.

## Results And Discussion

### A phylogenomic approach facilitates clustering-based identification of *F.g.*

The pattern of clustering observed on the phylogenomic tree (Fig. 1) sheds new light on the phylogeographic structure and genetic relationship among the species, which plays a significant role in their identification. Species-specific clustering was evident in case of all species represented by more than one strain (Fig. 1). In general, two major sister clades were resolved. The first clade occupied the basal position of the tree and grouped all species endemic to Africa (*F. aethiopicum*, *F. acaciae-mearnsii*), Asia (*F. asiaticum*, *F. nepalense*, *F. ussurianum* and *F. vorosii*) and Australia (*F. acaciae-mearnsii*). The second larger clade was more diverse and included species endemic to North (*F. gerlachii*, *F. louisianense*) and South America (*F. austroamericanum*, *F. mesoamericanum*, *F. brasiliicum*) and New Zealand (*F. cortadariae*). *F. boothii* and *F. meridionale* reported from diverse regions of Asia, Africa, and Latin America (van der Lee et al. 2015) were also clustered in this large clade.

To minimize the impact of geographic variation that may interfere with the results, we incorporated a large set of geographically diverse strains. Such a strategy is especially critical for pandemic species, which are often subdivided into genetically distinct populations (Shakya et al. 2021). *F. graminearum* formed a peripheral clade on the tree separating the species into thirteen major subclades named I to XIII. We observed geographic overlap in these subclades. Five out of nine strains originating from either North or South America were members of subclade I (n=9). Its basal position in the *F.g.* clade may suggest that subclade I includes genotypes with the highest genetic relationship to ancestors of *F.g.* Interestingly, subclade I included the oldest known strains of *F.g.*: un1 (CBS 185.32) and un2 (CBS 104.09), which were

isolated/deposited in the collection of the Westerdijk Fungal Biodiversity Institute in 1932 and 1904, respectively. Unfortunately, associated metadata of these two strains do not include information on their geographic origin. The first three subclades (II, III and IV) diverging from subclade I were represented by single strains, among which two: ar4 (114-2) and ar1 (CBS\_139514) come from South America and one (sy1, CS3005) from Australia. Most European strains were grouped into clusters VI-XIII. Interestingly, subclades VI (n=7), VII (n=1), VIII (n=10) and IX (n=2) grouped mostly Polish strains. A small subclade X clustered only two strains: ne5 (79E1) originating from the Netherlands and po11 (16-390-z) originating from Poland. This small clade was located between Polish subclade IX and west European subclade XI grouping strains (n=34) mostly from West Europe. The observed high degree of geographic clustering of the strains was not observed in subclade XIII (n=27), which diverged from west European subclade XI. It grouped strains from diverse geographic locations such Argentina, Brazil, Italy, Germany, the Netherlands, Poland, Russia and Serbia, pointing out evidence of their most recent spread to new geographic locations. Geographic expansion in accessible directions may be also indicated by clustering two European strains ge3 (CS10007) (from Germany) and ru13 (70725) (from Russia) into subclade I. Other examples come from subclade XI, which included single strain from South Africa (sa1, CBS 119799), and subclade VI, which included strain ir1 (CBS 110263) from Iran.

### **A phylogenomic approach enables the detection and clarification of the incorrect taxonomic status of historical strains held in fungal collections**

Among the *F.g.* strains, one strain (sa2, CBS 119800) unexpectedly clustered outside *F.g.* clade and was grouped in *F. boothii* clade. This strain is held in the fungal collection Westerdijk Fungal Biodiversity Institute (Utrecht, The Netherlands) as *F. graminearum* and according to available metadata it was isolated from maize in South Africa where *F. boothii* has been frequently reported (van der Lee et al. 2015). However, our previous mitochondrial-based comprehensive studies did not indicate incorrect taxonomic status of this strain (Kulik et al. 2015; Brankovics et al. 2018; Wyrębek et al. 2021). The complete mitogenome of CBS 119800 (NCBI accession no. KP966554) displays 100% sequence identity to *F.g.* strain CBS 104.09 (NCBI accession no. KR011238). Partial (659 nt) *tef* gene sequence of CBS 119800 (NCBI accession no. KT855180) shows 100% identity to *F.g.* (NCBI accession no. KP267345), indicating that *tef* alone could not definitively confirm its taxonomic affiliation. Blast search against NCBI non-redundant (nr) nucleotide database with an e-value cutoff of  $\leq e-0.0$ , 100% identity and 100% coverage also yielded nearly 30 hits to *F. boothii*. To further clarify its taxonomic status, we retrieved the complete sequence of the topoisomerase 1 (*top1*) and phosphoglycerate kinase (*pgk*) genes, which have become widely used taxonomic markers for *Fusaria* (Stielow et al. 2015). Blast searching with the *top1* gene as a query (e-value cutoff of  $\leq e-0.0$  and 100% coverage) yielded two hits to *F. boothii* (NCBI accession no. KY952952 and KY952951) with 100% sequence identity. However, blast searching with *pgk* gene as a query did not produce hits with 100% sequence identity, presumably due to the lack of *pgk* sequences from *F. boothii* in GenBank database. To determine sequence similarity of *pgk* between CBS 119800 and other strains of *F. boothii*, we retrieved its sequence from genome assembly of three strains of *F. boothii*: CBS 316.73, CBS 110251 and CBS 119170. Subsequent sequence comparison (data not

shown) revealed that all four strains shared 100% identity in the *pgk* gene, thus, supporting its identity as *F. boothii*.

Our results of phylogenomic analysis may be also helpful in resolving the uncertain taxonomic status of CBS 110260. This strain was isolated from maize in Nepal and has been assigned to either *F. asiaticum* or *F. meridionale* or a hybrid strain (O'Donnell et al. 2000; Ward et al. 2002; O'Donnell et al. 2004; Starkey et al. 2007; Yang et al. 2008). Most recent studies by (Walkowiak et al. 2016) based on analyses of SNPs and indels suggested that this strain shows 99% sequence identity to *F. meridionale*. Indeed, our phylogenomic approach grouped CBS 110260 together with the strain CBS 110249 (fme1, Fig. 1), which supports its taxonomic assignment as *F. meridionale*. Moreover, positioning of *F. meridionale* on the second large clade suggests that this cryptic species is more closely related to especially *F. cortadariae*, *F. austroamericanum*, *F. brasiliicum* and even to *F.g.* than to *F. asiaticum*. Additional whole-genome SNP analyses (Table 1) confirmed the above findings. Notably, the number of SNPs (126,189) between CBS 110260 and *F. meridionale* strain (CBS 110249) is in range of intraspecific variability found for *F.g.* (discussed in later sections), which is indicative of its taxonomic assignment as *F. meridionale*.

Table 1  
Results of whole-genome SNP analyses by mapping of assembly genomes to reference CBS 110260 strain

Assembly genomes		<i>F. meridionale</i>	<i>F. cortadariae</i>	<i>F. austroamericanum</i>	<i>F. brasiliicum</i>	<i>F. graminearum</i>	<i>F. asiaticum</i>
		CBS 110249	CBS 119183	CBS 110244	CBS 119179	PH-1	CBS 110258
CBS 110260	SNPs	126,189	336,338	335,339	342,238	567,931	577,489
	Linear coverage	98.5%	97.4%	97.4%	97.1%	95.6%	93.5%

### Assessment of genome-wide SNP counts allow successful identification of *F.g.*

One of the major drawbacks of clustering-based strategies for microbial identification is that they are time consuming, require bioinformatics skills, and specialized software and equipment. The remedy for these limitations is intuitive, user-friendly web-based platforms enabling fast and easy processing of next generation sequencing data through numerous cutting-edge tools. We used the PhaME workflow available at EDGE bioinformatics enabling fast counting of the total number of SNPs for determination of *F.g.*. The method was validated through multiple comparisons of assembly genomes to reference PH-1 strain. We estimated intra and interspecific differences in the number of SNPs facilitating species recognition (Supplementary File S2). Intraspecific variation calculated via comparisons of *F.g.* strains ranged from around 86,000 to nearly 158,000 SNPs, with one exception. The exceptional result was found for the strain CBS 119173, which yielded nearly 209,000 variable SNPs. However, the increased variability of CBS 119173 is not unexpected. CBS 119173 belongs to Gulf Coast population of *F.g.* with higher divergence, as evaluated via previous phylogenetic analyses of multilocus DNA sequence data

(Starkey et al. 2007). The increased genetic divergence of CBS 119173 was also depicted on phylogenomic tree (Fig. 1) by clustering of this strain separately from the remaining *F.g.* strains.

For cryptic species from the FGSC complex, interspecific variation ranged from 304,164 to 706,454 SNPs being from at least, nearly two times higher than intraspecific variation. Unsurprisingly, higher intraspecific variation was found for closely related morphospecies *F. culmorum*, *F. pseudograminearum*, *F. sambucinum* and *F. venenatum*, and ranged from 1,032,686 to 1,955,620 SNPs being more than 6.5 times higher than intraspecific variation.

## Conclusion

Calculation of whole-genome SNP variation allows to determine strains of *Fusarium graminearum* sensu stricto (*F.g.*) based on intra- and interspecific differences in the number of SNPs. Analysis of genome comparison does not require specialized software and can be rapidly performed in a user-friendly web-based workflows.

## Abbreviations

*F.g.* - *Fusarium graminearum* sensu stricto

FGSC – *Fusarium graminearum* complex

WGS – whole-genome sequencing

SNP – single nucleotide polymorphism

ITS – internal transcribed spacer

## Declarations

### Ethics approval and consent to participate

No animal or human volunteers were used in this study.

### Adherence to national and international regulations

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and material

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material. Assembled genomes of *Fusarium* spp. can be accessed at the NCBI bioproject under the accession numbers PRJNA677929 and PRJNA730356.

### Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Funding

This research was funded by the National Science Center, Poland, grant number 2018/31/B/NZ9/00546.

### Authors' contributions

TK conceived the idea, acquired funding, designed the study and prepared the draft manuscript; TM reconstructed phylogeny of the strains; GF performed assessment of genome-wide SNP counts; AD, AS, JO, KB, MB, MP and SS provided fungal strains for analysis, KT, KB and GF provided the technical support during the study; All authors read the manuscript, critically revised and agreed to the content of the manuscript. All authors read and approved the final manuscript.

### Acknowledgements

Not applicable.

## References

1. Albertin W, Marullo P (2012) Polyploidy in fungi: Evolution after whole-genome duplication. *Proceedings of the Royal Society B: Biological Sciences* 279:2497–2509
2. Andrews S (2010) Babraham Bioinformatics - FastQC A quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 4 Jan 2022
3. Araujo R (2014) Towards the genotyping of fungi: methods, benefits and challenges. *Current Fungal Infection Reports* 8:203–210
4. Araujo R, Sampaio-Maia B (2018) Fungal genomes and genotyping. *Advances in Applied Microbiology*. Academic Press Inc., pp 37–81
5. 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>
6. Bonfante P (2019) Algae and fungi move from the past to the future. *eLife*. [8https://doi.org/10.7554/eLife.49448](https://doi.org/10.7554/eLife.49448)
7. Boruta T (2018) Uncovering the repertoire of fungal secondary metabolites: From Fleming's laboratory to the international space station. *Bioengineered* 9:12–16

8. Brankovics B, Kulik T, Sawicki J, Bilska K, Zhang H, de Hoog GS, van der Lee TAJ, Waalwijk C, van Diepeningen AD (2018) First steps towards mitochondrial pan-genomics: Detailed analysis of *Fusarium graminearum* mitogenomes. PeerJ 2018:e5963. <https://doi.org/10.7717/PEERJ.5963/SUPP-4>
9. Capote N, Mara A, Aguado A, Snchez-Torres P (2012) Molecular tools for detection of plant pathogenic fungi and fungicide resistance. Plant Pathol. <https://doi.org/10.5772/38011>
10. de Hoog GS, Guarro J, Ahmed S, Al-Hatmi AMS, Figueras MJ, Vitale RG (2021) The Atlas of Clinical Fungi, 4th edn
11. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbón MH, Del Valle MB, Fyfe J, García-García L, Rastogi N, Sola C, Zozio T, Guerrero MI, León CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendón A, Sifuentes-Osornio J, De Ponce A, Cave MD, Fleischmann R, Whittam TS, Alland D (2006) Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: Insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J Bacteriol 188:759–772. <https://doi.org/10.1128/JB.188.2.759-772.2006>
12. Fisher MC, Gurr SJ, Cuomo CA, Blehert DS, Jin H, Stukenbrock EH, Stajich JE, Kahmann R, Boone C, Denning DW, Gow NAR, Klein BS, Kronstad JW, Sheppard DC, Taylor JW, Wright GD, Heitman J, Casadevall A, Cowen LE (2020) Threats posed by the fungal kingdom to humans, wildlife, and agriculture. <https://doi.org/10.1128/mBio.00449-20>. mBio 11
13. Frac M, Hannula SE, Belka M, Jędryczka M (2018) Fungal biodiversity and their role in soil health. Front Microbiol 9:707
14. Gladieux P, Ropars J, Badouin H, Branca A, Aguilera G, De Vienne DM, De La Vega RC, Branco S, Giraud T (2014) Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. Molecular Ecology 23:753–773. <https://doi.org/10.1111/mec.12631>
15. Hyde KD, Xu J, Rapior S, Jeewon R, Lumyong S, Grace Niego AT, Abeywickrama PD, Aluthmuhandiram S, Brahamanage JV, Brooks RS, Chaiyasen S, Thilini Chethana A, Chomnunti KW, Chepkirui P, Chuankid C, de Silva B, Doilom NI, Faulds M, Gentekaki C, Gopalan E, Kakumyan V, Harishchandra P, Hemachandran D, Hongsanan H, Karunarathna S, Karunarathna A, Khan SC, Kumla S, Jayawardena J, Liu RS, Liu J-K, Luangharn N, Patrick Macabeo T, Marasinghe AG, Meeks DS, Mortimer D, Mueller PE, Nadir P, Nataraja S, Nontachaiyapoom KN, Penkhrue S, Phukhamsakda W, Shaanker Ramanan C, Rathnayaka U, Sadaba AR, Sandargo RB, Samarakoon B, Tennakoon BC, Siva DS, Suwunwong R, Thongbai T, Thongklang B, Wei N, Nuwanthika Wijesinghe D, Winiski S, Yan J, Yasanthika J, Stadler E M (2019) The amazing potential of fungi: 50 ways we can exploit fungi industrially. Fungal Divers 97:1–136. <https://doi.org/10.1007/s13225-019-00430-9>
16. Kashyap PL, Rai P, Kumar S, Chakdar H, Srivastava AK (2017) DNA barcoding for diagnosis and monitoring of fungal plant pathogens. fungal biology. Springer International Publishing Switzerland, pp 87–122

17. Kiss L (2012) Limits of nuclear ribosomal DNA internal transcribed spacer (ITS) sequences as species barcodes for Fungi. *Proc Natl Acad Sci USA* 109:E1811
18. Kulik T, Biliska K, Żelechowski M (2020) Promising perspectives for detection, identification, and quantification of plant pathogenic fungi and oomycetes through targeting mitochondrial DNA. *International Journal of Molecular Sciences* 2020 21:2645. <https://doi.org/10.3390/IJMS21072645>
19. Kulik T, Ostrowska A, Buśko M, Pasquali M, Beyer M, Stenglein S, Załuski D, Sawicki J, Treder K, Perkowski J (2015) Development of an FgMito assay: A highly sensitive mitochondrial based qPCR assay for quantification of *Fusarium graminearum* sensu stricto. *Int J Food Microbiol* 210:16–23. <https://doi.org/10.1016/J.IJFOODMICRO.2015.06.012>
20. Leekitcharoenphon P, Kaas RS, Thomsen MCF, Friis C, Rasmussen S, Aarestrup FM (2012) snpTree -a web-server to identify and construct SNP trees from whole genome sequence data. *BMC Genomics* 13. <https://doi.org/10.1186/1471-2164-13-s7-s6>. Suppl 7:S6
21. Marin S, Ramos AJ, Cano-Sancho G, Sanchis V (2013) Mycotoxins: Occurrence, toxicology, and exposure assessment. *Food Chem Toxicol* 60:218–237
22. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A (2018) Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics (Oxford England)* 34:i142–i150. <https://doi.org/10.1093/BIOINFORMATICS/BTY266>
23. Mohanta TK, Bae H (2015) The diversity of fungal genome. *Biological Procedures Online* 17
24. Naranjo-Ortiz MA, Gabaldón T (2020) Fungal evolution: cellular, genomic and metabolic complexity. *Biol Rev* 95:1198–1232. <https://doi.org/10.1111/brv.12605>
25. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA (2013) Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol* 20:714–737. <https://doi.org/10.1089/CMB.2013.0084>
26. O'Donnell K, Kistler HC, Tacke BK, Casper HH (2000) Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. *Proceedings of the National Academy of Sciences* 97:7905–7910. <https://doi.org/10.1073/PNAS.130193297>
27. O'Donnell K, Ward TJ, Geiser DM, Kistler HC, Aoki T (2004) Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genet Biol* 41:600–623. <https://doi.org/10.1016/J.FGB.2004.03.003>
28. Pasquali M, Pallez-Barthel M, Beyer M (2020) Searching molecular determinants of sensitivity differences towards four demethylase inhibitors in *Fusarium graminearum* field strains. *Pestic Biochem Physiol* 164:209–220. <https://doi.org/10.1016/J.PESTBP.2020.02.006>
29. Persley D (1993) *Diseases of fruit crops* / edited by Denis Persley. Queensland Department of Primary Industries

30. Priest SJ, Yadav V, Heitman J (2020) Advances in understanding the evolution of fungal genome architecture. *F1000Research* 9:776
31. Raja HA, Baker TR, Little JG, Oberlies NH (2017a) DNA barcoding for identification of consumer-relevant mushrooms: A partial solution for product certification? *Food Chem* 214:383–392. <https://doi.org/10.1016/j.foodchem.2016.07.052>
32. Raja HA, Miller AN, Pearce CJ, Oberlies NH (2017b) Fungal identification using molecular tools: a primer for the natural products research community. *J Nat Prod* 80:756–770
33. Rampersad SN (2020) Pathogenomics and management of fusarium diseases in plants. *Pathogens* 9
34. Ray M, Ray A, Dash S, Mishra A, Achary KG, Nayak S, Singh S (2017) Fungal disease detection in plants: Traditional assays, novel diagnostic techniques and biosensors. *Biosensors and Bioelectronics* 87:708–723
35. Sarver BAJ, Ward TJ, Gale LR, Broz K, Corby Kistler H, Aoki T, Nicholson P, Carter J, O'Donnell K (2011) Novel Fusarium head blight pathogens from Nepal and Louisiana revealed by multilocus genealogical concordance. *Fungal Genet Biol* 48:1096–1107. <https://doi.org/10.1016/j.fgb.2011.09.002>
36. Schork NJ, Fallin D, Lanchbury JS (2000) Single nucleotide polymorphisms and the future of genetic epidemiology. *Clin Genet* 58:250–264. <https://doi.org/10.1034/j.1399-0004.2000.580402.x>
37. Seifert KA (2009) Progress towards DNA barcoding of fungi. *Molecular Ecology Resources* 9:83–89. <https://doi.org/10.1111/j.1755-0998.2009.02635.x>
38. Shakya M, Ahmed SA, Davenport KW, Flynn MC, Lo CC, Chain PSG (2020) Standardized phylogenetic and molecular evolutionary analysis applied to species across the microbial tree of life. *Sci Rep* 10:1–15. <https://doi.org/10.1038/s41598-020-58356-1>
39. Shakya SK, Grünwald NJ, Fieland VJ, Knaus BJ, Weiland JE, Maia C, Drenth A, Guest DI, Liew ECY, Crane C, Chang TT, Fu CH, Minh Chi N, Quang Thu P, Scanu B, von Stowasser ES, Durán Á, Horta Jung M, Jung T (2021) Phylogeography of the wide-host range panglobal plant pathogen *Phytophthora cinnamomi*. *Mol Ecol* 30:5164–5178. <https://doi.org/10.1111/MEC.16109>
40. Sharma P, Sharma S (2016) Paradigm Shift in Plant Disease Diagnostics: A Journey from Conventional Diagnostics to Nano-diagnostics. 237–264. [https://doi.org/10.1007/978-3-319-27312-9\\_11](https://doi.org/10.1007/978-3-319-27312-9_11)
41. Sharma R, Polkade AV, Shouche YS (2015) “Species concept” in microbial taxonomy and systematics. *Curr Sci* 108:1804–1814. <https://doi.org/10.18520/cs/v108/i10/1804-1814>
42. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E, 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>
43. Stajich JE (2017) Fungal genomes and insights into the evolution of the kingdom. The fungal kingdom. *American Society of Microbiology*, pp 619–633

44. Starkey DE, Ward TJ, Aoki T, Gale LR, Kistler HC, Geiser DM, Suga H, Tóth B, Varga J, O'Donnell K (2007) Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity. *Fungal genetics and biology: FG & B* 44:1191–1204.  
<https://doi.org/10.1016/J.FGB.2007.03.001>
45. Stielow JB, Lévesque CA, Seifert KA, Meyer W, Irinyi L, Smits D, Renfurm R, Verkley GJM, Groenewald M, Chaduli D, Lomascolo A, Welti S, Lesage-Meessen L, Favel A, Al-Hatmi AMS, Damm U, Yilmaz N, Houbraken J, Lombard L, Quaedvlieg W, Binder M, Vaas LAI, Vu D, Yurkov A, Begerow D, Roehl O, Guerreiro M, Fonseca A, Samerpitak K, van Diepeningen AD, Dolatabadi S, Moreno LF, Casaregola S, Mallet S, Jacques N, Roscini L, Egidi E, Bizet C, Garcia-Hermoso D, Martín MP, Deng S, Groenewald JZ, Boekhout T, de Beer ZW, Barnes I, Duong TA, Wingfield MJ, de Hoog GS, Crous PW, Lewis CT, Hambleton S, Moussa TAA, Al-Zahrani HS, Almaghrabi OA, Louis-Seize G, Assabgui R, McCormick W, Omer G, Dukik K, Cardinali G, Eberhardt U, de Vries M, Robert V (2015) One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia* 35:242–263. <https://doi.org/10.3767/003158515X689135>
46. Taylor JW, Fisher MC (2003) Fungal multilocus sequence typing—it's not just for bacteria. *Curr Opin Microbiol* 6:351–356. [https://doi.org/10.1016/S1369-5274\(03\)00088-2](https://doi.org/10.1016/S1369-5274(03)00088-2)
47. Teatero S, McGeer A, Li A, Gomes J, Seah C, Demczuk W, Martin I, Wasserscheid J, Dewar K, Melano RG, Fittipaldi N (2015) Population structure and antimicrobial resistance of invasive serotype IV group B *Streptococcus*, Toronto, Ontario, Canada. *Emerg Infect Dis* 21:585–591.  
<https://doi.org/10.3201/eid2014.140759>
48. Tralamazza SM, Rocha LO, Oggenfuss U, Corrêa B, Croll D (2019) complex evolutionary origins of specialized metabolite gene cluster diversity among the plant pathogenic fungi of the *Fusarium graminearum* species complex. *Genome Biol Evol* 11:3106–3122.  
<https://doi.org/10.1093/gbe/evz225>
49. Uelze L, Grützke J, Borowiak M, Hammerl JA, Juraschek K, Deneke C, Tausch SH, Malorny B (2020) Typing methods based on whole genome sequencing data. *One Health Outlook* 2.  
<https://doi.org/10.1186/S42522-020-0010-1>
50. Ukwuru MU, Ohaegbu CG, Muritala A (2017) An overview of mycotoxin contamination of foods and feeds. *Journal of Biochemical and Microbial Toxicology* 1:101.  
<https://doi.org/10.1002/9780470744307.gat157>
51. Ulian T, Diazgranados M, Pironon S, Padulosi S, Liu U, Davies L, Howes MR, Borrell JS, Ondo I, Pérez-Escobar OA, Sharrock S, Ryan P, Hunter D, Lee MA, Barstow C, Łuczaj Ł, Pieroni A, Cámara-Leret R, Noorani A, Mba C, Nono Womdim R, Muminjanov H, Antonelli A, Pritchard HW, Mattana E (2020) Unlocking plant resources to support food security and promote sustainable agriculture. *PLANTS, PEOPLE, PLANET* 2:421–445. <https://doi.org/10.1002/ppp3.10145>
52. Valverde ME, Hernández-Pérez T, Paredes-López O (2015) Edible mushrooms: Improving human health and promoting quality life. *International Journal of Microbiology* 2015

53. van der Lee T, Zhang H, van Diepeningen A, Waalwijk C (2015) Biogeography of *Fusarium graminearum* species complex and chemotypes: a review. Taylor & Francis 32:453–460. <https://doi.org/10.1080/19440049.2014.984244>
54. Walkowiak S, Rowland O, Rodrigue N, Subramaniam R (2016) Whole genome sequencing and comparative genomics of closely related Fusarium Head Blight fungi: *Fusarium graminearum*, *F. meridionale* and *F. asiaticum*. BMC Genomics 17:1–15. <https://doi.org/10.1186/S12864-016-3371-1/FIGURES/6>
55. Ward TJ, Bielawski JP, Corby Kistler H, Sullivan E, O'Donnell K (2002) Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*. Proc Natl Acad Sci USA 99:9278–9283. <https://doi.org/10.1073/PNAS.142307199>
56. Wickes BL, Romanelli AM (2020) Diagnostic mycology: xtreme challenges. Journal of Clinical Microbiology 58. <https://doi.org/10.1128/JCM.01345-19>
57. Wu B, Hao W (2019) Mitochondrial-encoded endonucleases drive recombination of protein-coding genes in yeast. Environ Microbiol 21:4233–4240. <https://doi.org/10.1111/1462-2920.14783>
58. Wyrębek J, Molcan T, Myszczyński K, van Diepeningen AD, Stakheev AA, Żelechowski M, Biliska K, Kulik T (2021) Uncovering diagnostic value of mitogenome for identification of cryptic species *Fusarium graminearum* sensu stricto. Front Microbiol 12:714651. <https://doi.org/10.3389/FMICB.2021.714651/FULL>
59. Yang L, van der Lee T, Yang X, Yu D, Waalwijk C (2008) Fusarium populations on Chinese barley show a dramatic gradient in mycotoxin profiles. Phytopathology 98:719–727. <https://doi.org/10.1094/PHYTO-98-6-0719>
60. Zain ME (2011) Impact of mycotoxins on humans and animals. Journal of Saudi Chemical Society 15:129–144. <https://doi.org/10.1016/j.jscs.2010.06.006>

## Figures

