1 **Genomic sequencing of the aquatic *Fusarium* spp. QHM and BWC1 and their**
2 **potential application in environmental protection**

4 Running title: Genomic sequences of *Fusarium* spp. QHM and BWC1

6 Hongfei Zhu,^{a,#} Long Zhu,^a and Ning Ding^a

7 ^aCollege of Environmental Science and Engineering, Liaoning Technical University

8 No. 47 Zhonghua Road, Xihe District, Fuxin, Liaoning 123000, China.

9 #Address correspondence to Hongfei Zhu, hongfei.zhu@yahoo.com.

12 **ABSTRACT**

Fusarium species are distributed widely in ecosystems of a wide pH range and play a pivotal role in the aquatic community through the degradation of xenobiotic compounds and secretion of secondary metabolites. The elucidation of their genome would therefore be highly impactful with regard to the control of environmental pollution. Therefore, in this study, two indigenous strains of aquatic *Fusarium*, QHM and BWC1, were isolated from a coal mine pit and a subterranean river respectively, cultured under acidic conditions, and sequenced. Phylogenetic analysis of these two isolates was conducted based on the sequences of internal transcript (ITS1 and ITS4) and encoding β -microtubulin (TUB2), translation elongation factors (TEFs) and the

second large sub-unit of RNA polymerase (RPB2). *Fusarium*, QHM could potentially represent a new species within the *Fusarium fujikuroi* species complex. *Fusarium* BWC1 were found to form a clade with *Fusarium subglutinans* NRRL 22016, and predicted to be *Fusarium subglutinans*. Shot-gun sequencing on the Illumina HiSeq×10 Platform was used to elucidate the draft genomes of the two species. Gene annotation and functional analyses revealed that they had bio-degradation pathways for aromatic compounds; further, their main pathogenic mechanism was found to be the efflux pump. To date, the genomes of only a limited number of aquatic species from the *Fusarium fujikuroi* species complex, especially from the aquatic species, have been sequenced. Therefore, the present findings are novel and have important potential for the future in terms of environmental control.

IMPORTANCE *Fusarium* genus has over 300 species and were distributed in a variety of ecosystem. Increasing attention has been drawn to *Fusarium* due to the importance in aquatic community, pathogenicity and environmental protection. The genomes of the strains in this work isolated in acidic condition, were sequenced. The analysis has indicated that the isolates were able to biodegrade xenobiotics, which makes it potentially function as environmental bio-agent for aromatic pollution control and remediation. Meanwhile, the virulence and pathogenicity were also predicted for reference of infection control. The genome information may lay foundation for the fungal identification, disease prevention resulting from these

isolates and other “-omics” research. The isolates were phylogenetically classified into *Fusarium fujikuroi* species complex by means of concatenated gene analysis, serving as new addition to the big complex.

Keywords: isolation; phylogenetics; *Fusarium*; genome; sequencing; degradation

Introduction

Species of the genus *Fusarium* are an important group of fungi that are distributed widely in below-ground and above-ground habitats (1, 2). Certain *Fusarium* species can even be found in aquatic habitats, including coal mine pits and subterranean rivers (2). These species are also characterized by their ability to grow in environments with a wide pH range (1, 2). The genus *Fusarium* is important from the perspective of environmental protection and phytopathogenicity: On the one hand, *Fusarium* spp. have potential as bio-control agents, on account of their ability to degrade a variety of xenobiotics, such as aromatic compounds from water or soil. On the other hand, they infect a broad spectrum of crops, thereby resulting in huge economic losses or mycotoxin contamination (3). Given the complex characteristics of these species, it is imperative to elucidate the *Fusarium* genome sequence for the elimination of

63 pollutants while ensuring the control of its pathogenicity. Further, as the *Fusarium*
 64 genus forms a “species complex” with the *Aspergillus* genus on account of closely
 65 related species, species-level identification of *Fusarium* is also necessary (4).
 66 Whole-genome sequencing would be ideal, as this method provides more information
 67 than other sequencing approaches in terms of the elucidation of acidic mechanisms,
 68 biodegradation and microbial identification. Fungus-specific genes, such as
 69 *ITS1/ITS4*, *TUB2*, *TEF* and *RPB2*, and their sequences have been universally used for
 70 species identification and analysis of evolutionary relationships (4). The simultaneous
 71 use of the fungus-specific genes may phylogenetically give more resolution to over
 72 90% gene identity within *Fusarium* species complex.

73 The *Fusarium* spp. QHM and BWC1 have been isolated and purified from the
 74 9KG medium culture in the process of isolation of *Acidiphilium cryptum* from coal
 75 mine water and underground river-water. Based on the isolation environment,
 76 *Fusarium* spp. QHM and BWC1 were thought to be indigenous and predicted to the
 77 best solution for the *in situ* remediation of acidic coal mine water (5). These fungal
 78 isolates were experimentally proved to be tolerant in a wide pH range of 3.0 to 8.5;
 79 therefore, they also have the potential to act as bio-agents for the treatment of acid
 80 mine drainage (5, 6). Accordingly, the genomes of both *Fusarium* QHM and
 81 *Fusarium* BWC1 isolates were sequenced, annotated and analyzed using
 82 bioinformatics methods. The information obtained is highly valuable, as genomic
 83 sequences from aquatic fungal isolates, especially from coal mine water, are scant.

The yielded genomic sequences and functions will prove useful for research endeavors in the field of “-omics”, microbial communities and water pollution management. Importantly, this genomic information could be used to investigate the biodegradation of aromatic compounds, such phenol and benzo[a]pyrene, for water pollution control and bio-remediation of soil contaminated with polycyclic aromatic hydrocarbons (7).

In this work, we present two draft genomes of *Fusarium* spp. QHM and BWC1. Based on the genes sequenced and their annotational and functional analysis, the evolutionary relationships between these species and their aromatic metabolism were investigated. The findings are expected to have future applications in the management and remediation of coal mine water or underground river water.

Results

1. Isolation of *Fusarium* spp. from the water samples

Two isolates, *Fusarium* spp. QHM and BWC1, were obtained from the water samples. These strains exhibit growth that is visible to the naked eye when cultured in 9 KG medium at a pH of 3.5 to 8.5. The optimum pH for growth is 5.0 to 6.0, so these species are considered to be acidophiles. The overall pH tendency during the growth was to slide down to lower level in the first two or three days. From inoculation to the exponential stage, pH 5.0-6.0 may go down to pH 2.5-3.5. *Fusarium* sp. QHM and *Fusarium* sp. BWC1 are currently stored in the in-house laboratory of the university.

2. Characterization

The sequences obtained after PCR amplification of the ITS1, TUB2, TEF and RPB2 genes were submitted to the GenBank of the National Centre of Biological Information (NCBI) under accession No. MK791252, MK850849, MK850848, MK850847, MK898823, MK907693, MK907692 and MK907691. The BLAST results are provided in Table 1.

3. Sequencing data

The Whole Genome Shotgun project sequences have been deposited in GenBank under accession numbers SWCP000000000 (*Fusarium* sp. QHM) and SWCQ000000000 (*Fusarium* sp. BWC1). In total, 479 and 2352 scaffolds were obtained for *Fusarium* spp. QHM and BWC1 respectively. The overall results of the assemblage are summarized in Table 2. The GC content of the genomic sequences of both species was 44% to 60%, which is in line with the sequencing data. The predicted coding genes are shown in Table 3.

4. Gene annotation results

The gene annotation results for *Fusarium* spp. QHM and BWC1 were similar, which suggested that they may evolve from a later common ancestor because of a speciation. The minor divergence derived from duplication. As shown in Figure 1, based on Gene Ontology (GO) analysis, both genes were categorized into three main groups that were further divided into various subcategories: biological processes, cellular components and molecular functions. The majority of genes were assigned to

metabolic processes, and the second highest number of genes were assigned to catalytic activity. The findings indicate that genes involved in primary and secondary metabolic processes, including secretion, degradation and catalysis, were identified in both strains. Information from the five databanks was used as reference to determine the functions of all the genes from both samples. The annotation results are summarized in Figure 1. The fundamental annotations for both strains in the five public databanks are provided in two excel files, Table S1 and Table S2 in Supplemental Material.

Histograms of the findings obtained from Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, as shown in Figure 2 and 3, show the classification of the genes from both species into six categories: metabolism, human disease, organismal systems, genetic information processing, environmental information processing and cellular processes. The number of genes involved in metabolic processes was 3243 and 3423 in *Fusarium* spp. QHM and BWC1 respectively. The main pathways identified were xenobiotic biodegradation, carbohydrate metabolism and amino acid metabolism. Gene enrichment analysis of the transport- and catabolism-related genes indicated that a high number of genes were involved in these functions at the cellular level. Further, the number of translation-associated genes identified in both species indicate that these genes played a dynamic role in genetic information processing in both strains.

There are six main enzyme families that are involved in the synthesis, metabolism

and recognition of complex carbohydrates: glycoside hydrolases, glycosyltransferases, polysaccharide lyases, carbohydrate esterases and non-catalytic carbohydrate-binding modules. Carbohydrate hydrolases, which form the majority of the enzyme components, comprise 41.87% and 41.88% of the genes associated with complex carbohydrates in *Fusarium* spp. QHM and BWC1, respectively, as shown in Figure 4 and 5. Overall, the difference between *Fusarium* spp. QHM and BWC1 genomes with regard to these six gene families is inconspicuous.

5. Phylogenetic analysis

Simultaneous use of *TUB2*, *TEF* and *RPB2* genes effectively overcame the conservativeness of *ITS* in the phylogenetic analysis. Based on the evolutionary tree shown in Figure 6, it was predicted that *Fusarium* spp. QHM and BWC1 are essentially the different *Fusarium* species that were classified into distinct clades. *Fusarium* sp BWC1 and *Fusarium subglutinans* NRRL 2016, a type strain, coexisted in same one clade and were in the closest phylogenetic relationship other than any species. Thereof *Fusarium* sp BWC1 was predicted to be a strain of *Fusarium subglutinans* at bootstrap value 1.00. *Fusarium pininemorale* CBS 137240, *Fusarium marasasianum* CBS 137238, *Fusarium parvisorum* CBS 137236 and *Fusarium fracticaudum* CBS 137234 exhibited close phylogenetic relationship with it in the *Fusarium fujikuroi* species complex (FFSC). *Fusarium* sp. QHM contained in a separated clade shown in bold font distinctly differ from other taxa members of *Fusarium*. Thereof *Fusarium* sp. QHM was putative new species of FFSC. *Fusarium*

oxysporum CBS 77497 is considered as a non-FFSC or outgroup species. Taken together, the multiple loci in the phylogenetic tree topologies shown in Figure 6 are in concordance with previously reported findings for FFSC (4, 24).

6. Genes associated with virulence and pathogenicity

The virulence and pathogenicity genes were identified by referencing the genes deposited in the fungal virulence factors (DFVF), comprehensive antibiotic research (CAR) and pathogen host interactions (PHI) databases. A total of 107 genes of *Fusarium* sp. QHM and 106 genes of *Fusarium* sp. BWC1 were found to play a role in the efflux pump and confer resistance against antibiotics such as glycopeptides, rifampin and tetracycline (17). The efflux pump is the main mechanism for antibiotic resistance adopted by these two fungi. Other underlying mechanisms included molecular bypass, enzymatic antibiotic inactivation, alterations in the antibiotic target, cell permeability, horizontal transfer of resistance genes and modulation of antibiotic efflux. Comprehensive data related to the antibiotic resistance mechanisms of both strains are provided in Figure 7. With regard to pathogenicity, 1487 and 1494 unaffected pathogenic genes were identified in *Fusarium* sp. QHM and *Fusarium* sp. BWC1, respectively, while 1197 and 1219 reduced virulence genes were identified. The virulent genes of clinical interest may be selected as candidate targets for control measures and novel way detection.

7. Xenobiotic degradation pathway analysis

Genes involved in the degradation of a variety of xenobiotics were identified in

the KEGG pathways. Based on the findings, *Fusarium* spp. QHM and BWC1 were considered to be responsible for the bio-degradation or catalysis of aromatic compounds such as chloroalkane, styrene, atrazine, chlorocyclohexane and dioxins. A detailed observation of these pathways might reveal the presence of relevant enzymes such as 2-haloacid dehalogenase, carboxymethylenebutenolidase, catechol-1,2-dioxygenase, salicylate hydroxylase, DDT-dehydrochlorinase, and biphenyl-2,3-dioxygenase. The key enzymes involved in the degradation pathway and the KEGG orthology identifiers are summarized in Table 4.

8. Syntenic analysis

The dense ligatures and concentrated genes in the DNA fragment alignments between the genomes of *Fusarium* spp. QHM (reference) and BWC1 (query) indicate a high degree of synteny, which is believed to be derived from the sequence conservation. This finding indicates the close phylogenetic relationship between the two strains, even though they were contained in distinct clade within FFSC in Figure 6. The difference in homology and orthology was proposed to be responsible for the individual adaption mechanism. Gene loss and horizontal gene transfer may somehow occurred under the natural selection pressure in the evolution.

Discussion

The present study aims to present the draft genome sequences of the previously isolated *Fusarium* spp. QHM and BWC1 from acidic aquatic environments. The

genomes were analyzed to identify the main xenobiotic degradation and pathogenic mechanisms of these species. Moreover, the phylogenetic relationship between the two species and other species in the genus was elucidated. The two species were found to be closely related phylogenetically. *Fusarium* spp. QHM and QHM 03 forms a separate clade (based on the degree of synteny) within FFSC.

Here, functional analyses of the genes and sequences revealed that the majority of the metabolism-associated genes in both species were involved in xenobiotic biodegradation, carbohydrate metabolism (the main enzyme component being carbohydrate hydrolases) and amino acid metabolism. Further analysis of the xenobiotic degradation pathways showed that *Fusarium* spp. QHM and BWC1 played a role in the bio-degradation or catalysis of several important aromatic compounds. The xenobiotic degradation ability of *Fusarium* spp. QHM and BWC1 would be potentially useful for the treatment of water that is polluted by anthropogenic activities. The use of these fungal species for tackling the pollution of water and soil would also be more effective and safer than the use of other chemical approaches. Further, the plasticity observed in *Fusarium* spp. QHM and BWC1 to adapt to low pH conditions suggests that it can be utilized in acid water control and heavy metal removal from the environment. Thus, *Fusarium* spp. QHM and BWC1 would be highly useful as bio-control agents within the indigenous microbial community in coal mine water.

There are some limitations to this study on the degradation analysis of xenobiotic,

due to the lack of experimental evidence. It has been well documented that the cytochrome P450 monooxygenase plays pivotal role in the degradation process by the oxygen atom addition to the bonds between carbon and hydrogen, or among the carbon atoms. Future studies should focus on the investigation of P450 transcriptomics of these two isolates in different pH and aromatic concentration.

With regard to their pathogenic mechanisms, the present findings revealed that the efflux pump was the major mechanism of antibiotic resistance commonly adopted by both bacterium and fungus. The efflux pump components should be taken as drug targets in the fungal disease control or prevention. It has been reported that *Fusarium subglutinans* may function as a maize pathogen and the phyto-virulence is in need of further investigation of this strain. The potential infectivity and pathogenicity of the *Fusarium* community, especially for the environmental resource are to be fully aware of.

Multigene phylogenetic analysis generally provides stronger resolution than analysis of a single conservative gene. The identity of ITS shared by *Fusarium* spp. QHM and BWC1 is 100%, which makes it necessary to adopt to multigene phylogenetic analysis. Under the analysis, *Fusarium. pseudonygamai* CBS 41897, *Fusarium pseudoanthophilum* NRRL 25206, *Fusarium lactis* NRRL 25200 type strain, *Fusarium brevicatenuatum* CBS 40497, and *Fusarium napiforme* CBS 74897 were found to be closely related to *Fusarium* sp. QHM in FFSC. The two *Fusarium* species were well separated from the others and formed a separate clade (QHM) with

high bootstrap support for three or four-gene datasets. Therefore, multigene phylogenetic analysis provides genealogical exclusivity for molecular fungal taxonomy. Moreover, genomic information may provide insight into the growth, spread, and infection control of this fungus. In particular, information about the genome sequence of *Fusarium* spp. QHM and BWC1 would have important implications for aquatic quarantine and water quality assessment.

In conclusion, the findings of this study have important implications for future research on the potential use of these species as biocontrol agents for the treatment of polluted water bodies and soil. The genomic information provides opportunity and tool for the control of their pathogenicity or the etiological detection.

Materials and methods

1. Isolation and cultivation

Water samples were obtained from Fuxin Qinghemmen (for QHM) coal mine, which is located 700 m underground, and Benxi water cave (for BWC), Liaoning, China. Both the QHM and BWC sampling sites are 20 cm below the river water surface. Each water sample (5 mL) was added to 100 mL of 9 KG liquid medium at pH 3.0–3.5, and this solution was used for *A. cryptum* cultivation. The 9 KG medium has the following composition: $(\text{NH}_4)_2\text{SO}_4$, 3.0 g; K_2HPO_4 , 0.5 g; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.5 g; $\text{Ca}(\text{NO}_3)_2$, 0.01 g; tryptone, 0.1 g; $\text{C}_6\text{H}_{12}\text{O}_6$, 1 g (sterilizing filtered); Agar, 20 g (if solid) in 1000 mL (8, 9). The pH of the solid medium was 4.5. All the ingredients

(with the exception of glucose) were sterilized for 20 min, as per the standard protocol.

The inoculated flask was shaken to turbidity at 100 rpm (rotation per minute) at 30°C for 24–36 h. Subsequently, the liquid culture was spread on a solid plate so as to enable the isolation of a single colony. The single colony, which appeared white, smooth and round (approximately 1 mm in diameter), was selected for enrichment on the solid medium. The solid culture derived from the single colony was inoculated in a liquid culture to increase growth for identification and storage.

2. Characterization

A series of 9 KG liquid media, across a pH gradient of 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 8.0 and 9.0, was prepared to investigate the acidic tolerance of the strains under the same cultural conditions as described above. Culture growth under the gradient pH was manually observed.

Polymerase chain reaction (PCR) targeting the ITS1, TUB2, TEF and RPB2 genes was performed primarily for species identification through BLAST (Basic Local Alignment Search Tool) with the reference genes obtained from the relevant public databases. The PCR primers used in the present work and their sequences are provided in Table 1.

3. Multigene phylogenetic analysis

Since the ITS sequence is known to be relatively well conserved within the fungal clade of the species complex, the phylogenetic relationship of the two isolates with

other species in the *Fusarium* genus was analyzed based on multiple gene loci, including TUB2, TEF and RPB2 (10-12). The simultaneous use with multiple gene loci may reduce the phylogenetic bias and provide a comprehensive understanding of the phylogenetic relationships (3). The PhyML software was used to construct the maximum-likelihood tree for the TUB2, TEF and RPB2 gene sequences with the neighbor-joining method (13).

4. Genome sequencing

Genomic DNA extraction for *Fusarium* spp. QHM and BWC1 was carried out with the universal DNA purification kit from Tiangen Biotech (Beijing) Co. Ltd. Shot-gun sequencing on the Illumina Hiseq×10 Platform was conducted by Shanghai Majorbio.

5. Sequence assembly and gene function prediction

SOAP denovo v.2.04 (<http://soap.genomics.org.cn/>) was used for de novo assembly of the short sequences obtained from the clean data. Trimmomatic, SeqPrep, Sickel and FastqTotalHighQualityBase.jar etc. had been used to acquire clean data. Maker 2, Barnap 0.4.2, tRNAscan-SE v1.3.1 and RepeatMasker were used to predict the gene function of coding sequences, rRNA, tRNA genes and repeat sequences. All these softwares were provided by the i-sanger cloud platform (<https://www.i-sanger.com/>).

6. Gene annotation

The first step is fundamental annotation based on sequences and functions

deposited in the five most well-known databanks, i.e., Non-Redundant Protein Database (NRPD), Swiss-Prot, Pfam, Clusters of Orthologous Groups of proteins (COG) and GO. The gene function of the coding sequences was predicted by means of BLAST (14, 15).

In the next step, KEGG analysis was used to identify the pathways that the genes were involved in; the encoded proteins were primarily analyzed using the Diamond software. The analysis includes pathways involved in metabolism, diseases, organismal systems, genetic information processing, environmental information processing and cellular processes.

With regard to carbohydrate metabolism, the carbohydrate active enzymes (CAZy) database with the Diamond software was used (16). For pathogenicity and antibiotic resistance, DFVF, CAR and PHI databases were used to determine the functions of virulent or resistant genes (17, 18). The Diamond software was extensively used to identify the virulence genes, pathogen–host interactions, secretion proteins, transport proteins and trans-membrane proteins through alignment of the obtained genomes (16, 19).

7. Bioinformatics-based analysis of the genes involved in the degradation of xenobiotics

From the environmental perspective, the metabolic pathways of aromatic compounds, such as chloroalkane and styrene, are of importance (20). A variety of enzymes present in these pathways play a role in the degradation of xenobiotics by the

Fusarium spp. QHM and BWC1. The key enzymes involved in xenobiotic degradation were manually identified based on the reference pathways in the KEGG database (21-23).

8. Syntenic analysis of the genomes of the *Fusarium* spp. QHM and BWC1

Syntenic analysis, which is one of the methods of comparative genomics, is the main method used to investigate homology between species. The sequences of the *Fusarium* sp. QHM, which served as a reference, were aligned against those of the *Fusarium* sp. BWC1 to determine the homology and orthologues between the species. The orthologous and homologous regions between the contigs of the two species were aligned to make a contiguous block within the range. The Sibelia and Circos softwares installed in i-sanger platform were used to carry out the comparison.

Acknowledgments

This research did not receive grants from any funding agency in the public, commercial, or not-for-profit sectors.

References

- 353 1. Rosset J, Barlocher F. 1985. Aquatic Hyphomycetes - Influence of Ph, Ca-2+
354 and Hco3- on Growth-Invitro. Transact Brit Mycol Soci 84:137–145.
- 355 2. Moussa TAA, Al-Zahrani HS, Kadasa NMS, Ahmed SA, de Hoog GS,
356 Al-Hatmi AMS. 2017. Two new species of the *Fusarium fujikuroi* species
357 complex isolated from the natural environment. Antonie Van Leeuwenhoek
358 110:819–832.
- 359 3. Niehaus EM, Munsterkotter M, Proctor RH, Brown DW, Sharon A, Idan Y,
360 Oren-Young L, Sieber CM, Novak O, Pencik A, Tarkowska D, Hromadova K,
361 Freeman S, Maymon M, Elazar M, Youssef SA, El-Shabrawy EM, Shalaby
362 ABA, Houterman P, Brock NL, Burkhardt I, Tsavkelova EA, Dickschat JS,
363 Galuszka P, Guldener U, Tudzynski B. 2016. Comparative "Omics" of the
364 *Fusarium fujikuroi* Species Complex Highlights Differences in Genetic
365 Potential and Metabolite Synthesis. Gen Biol Evol 8:3574–3599.
- 366 4. Moussa TAA, Al-Zahrani HS, Kadasa NMS, Ahmed SA, de Hoog GS,
367 Al-Hatmi AMS. 2017. Two new species of the *Fusarium fujikuroi* species
368 complex isolated from the natural environment. Antonie Van Leeuwenhoek
369 Internat J Gen Mol Microbiol 110:819–832.
- 370 5. Sanchez-Rangel D, Hernandez-Dominguez EE, Perez-Torres CA, Ortiz-Castro
371 R, Villafan E, Rodriguez-Haas B, Alonso-Sanchez A, Lopez-Buenfil A,
372 Carrillo-Ortiz N, Hernandez-Ramos L, Ibarra-Laclette E. 2018. Environmental

373 pH modulates transcriptomic responses in the fungus *Fusarium* sp. associated
374 with KSHB Euwallacea sp. near fornicatus. BMC Genomics 19:721.

375 6. Zinger L, Coissac E, Choler P, Geremia RA. 2009. Assessment of microbial
376 communities by graph partitioning in a study of soil fungi in two Alpine
377 meadows. Appl Environ Microbiol 75:5863–5870.

378 7. Tigini V, Prigione V, Di Toro S, Fava F, Varese GC. 2009. Isolation and
379 characterisation of polychlorinated biphenyl (PCB) degrading fungi from a
380 historically contaminated soil. Microb Cell Fact 8:5.

381 8. Niu J, Deng J, Xiao Y, He Z, Zhang X, Van Nostrand JD, Liang Y, Deng Y, Liu
382 X, Yin H. 2016. The shift of microbial communities and their roles in sulfur
383 and iron cycling in a copper ore bioleaching system. Sci Rep 6:34744.

384 9. Latorre M, Cortes MP, Travisany D, Di Genova A, Budinich M, Reyes-Jara A,
385 Hodar C, Gonzalez M, Parada P, Bobadilla-Fazzini RA, Cambiazo V, Maass
386 A. 2016. The bioleaching potential of a bacterial consortium. Bioresour
387 Technol 218:659–666.

388 10. Xu JP. 2016. Fungal DNA barcoding. Genome 59:913–932.

389 11. Guarro J, GeneJ, Stchigel AM. 1999. Developments in fungal taxonomy. Clin
390 Microbiol Rev 12:454–500.

391 12. Toju H, Tanabe AS, Yamamoto S, Sato H. 2012. High-coverage ITS primers
392 for the DNA-based identification of ascomycetes and basidiomycetes in
393 environmental samples. PLoS One 7:e40863.

394 13. Richards TA, Leonard G, Wideman JG. 2017. What Defines the "Kingdom"
395 Fungi? Microbiology Spectrum 5.

396 14. Issotta F, Galleguillos PA, Moya-Beltran A, Davis-Belmar CS, Rautenbach G,
397 Covarrubias PC, Acosta M, Ossandon FJ, Contador Y, Holmes DS,
398 Marin-Eliantonio S, Quatrini R, Demergasso C. 2016. Draft genome sequence
399 of chloride-tolerant *Leptospirillum ferriphilum* Sp-Cl from industrial
400 bioleaching operations in northern Chile. Stand Genomic Sci 11:19.

401 15. Valdes J, Ossandon F, Quatrini R, Dopson M, Ds. H. 2011. Draft genome
402 sequence of the extremely acidophilic biomining bacterium *Acidithiobacillus*
403 *thiooxidans* ATCC 19377 provides insights into the evolution of the
404 *Acidithiobacillus* genus. J Bacteriol 193:7003–7004.

405 16. Stajich JE. 2017. Fungal Genomes and Insights into the Evolution of the
406 Kingdom. Microbiology Spectrum 5.

407 17. Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro
408 A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S,
409 Shim WB, Woloshuk C, Xie XH, Xu JR, Antoniw J, Baker SE, Bluhm BH,
410 Breakspear A, Brown DW, Butchko RAE, Chapman S, Coulson R, Coutinho
411 PM, Danchin EGJ, Diener A, Gale LR, Gardiner DM, Goff S,
412 Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira
413 CD, Koehrsen M, Kumar L, Lee YH, Li LD, Manners JM, Miranda-Saavedra
414 D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan

415 MC, Sain D, et al. 2010. Comparative genomics reveals mobile pathogenicity
416 chromosomes in *Fusarium*. *Nature* 464:367–373.

417 18. Bakour S, Sankar SA, Rathored J, Biagini P, Raoult D, Fournier PE. 2016.
418 Identification of virulence factors and antibiotic resistance markers using
419 bacterial genomics. *Future Microbiol* 11:455–466.

420 19. Duba A, Goriewa-Duba K, Wachowska U. 2018. A Review of the Interactions
421 between Wheat and Wheat Pathogens: *Zymoseptoria tritici*, *Fusarium* spp. and
422 *Parastagonospora nodorum*. *Int J Mol Sci* 19.

423 20. Maphosa MN, Steenkamp ET, Wingfield BD. 2016. Genome-Based Selection
424 and Characterization of *Fusarium circinatum*-Specific Sequences. *G3-Gen*
425 *Gen* 6:631–639.

426 21. Fetzner S. 2012. Ring-cleaving dioxygenases with a cupin fold. *Appl Environ*
427 *Microbiol* 78:2505-2514.

428 22. Wang Y, Li J, Liu A. 2017. Oxygen activation by mononuclear nonheme iron
429 dioxygenases involved in the degradation of aromatics. *J Biol Inorg Chem*
430 22:395–405.

431 23. Zhang X, Niu J, Liang Y, Liu X, Yin H. 2016. Metagenome-scale analysis
432 yields insights into the structure and function of microbial communities in a
433 copper bioleaching heap. *BMC Genet* 17:21.

434 24. Al-Hatmi AMS, Hagen F, Menken SBJ, Meis JF, de Hoog GS. 2016. Global
435 molecular epidemiology and genetic diversity of *Fusarium*, a significant

436 emerging group of human opportunists from 1958 to 2015. Emerging
 437 Microbes & Infections 5.

438 25. Karlsson I, Edel-Hermann V, Gautheron N, Durling MB, Kolseth AK,
 439 Steinberg C, Persson P, Friberg H. 2016. Genus-Specific Primers for Study of
 440 *Fusarium* Communities in Field Samples. Appl Environ Microbiol
 441 82:491–501.

442 26. Martin-Rodriguez AJ, Reyes F, Martin J, Perez-Yepe J, Leon-Barrios M,
 443 Couttolenc A, Espinoza C, Trigos A, Martin VS, Norte M, Fernandez JJ. 2014.
 444 Inhibition of bacterial quorum sensing by extracts from aquatic fungi: first
 445 report from marine endophytes. Mar Drugs 12:5503–5526.

446 **Table 1. Identification genes and characteristics of the yielded sequences of the analyzed genes in *Fusarium* spp. QHM and BWC1**

Seq. no.	Identification gene	Primer sequence (5'-3')	Query length (bp)	Stain name	Accession number of the yielded sequences	Similar gene/sequence(s), strain(s) (Genbank accession No.)	%Identity	Putative name	References
1	ITS (Internal Transcribed Spacer 1)	ITS1(TCCGTAGGTGAACCTGCGG), ITS4(TCCTCCGCTTATTGATATGC)	462	BWC1	MK791252	Small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer2, (MH862670.1)	462/463(99.78%)	<i>Fusarium napiforme</i> strain CBS 748.97	(12, 26)
2	TUB2 (β -Tubulin)	T1(AACATGCGTGAGATTGTAAGT), T2(TAGTGACCCTTGGCCCAGTTG)	1075	BWC1	MK850849	Beta-tubulin (TUB2) gene, partial cds	1074/1078(99.63%)	<i>Fusarium subglutinans</i> strain M16084	(4, 10)
3.	TEF (Translation Elongation	EF1(ATGGGTAAGGARGACAAGAC), EF2(GGARGTACCAGTSATCATG)	642	BWC1	MK850848	Translation elongation factor 1-alpha (tef1)	642/642(100%)	<i>Fusarium subglutinans</i> voucher GFYX224	(2, 24)

	Factor)					gene, partial cds (JX456583.1)			
4	RPB2 (Subunit B of RNA Polymerase II)	RPB2_5F2(GGGGWGAYCAGAAGAAG GC), RPB2_11AR(GCRTGGATCTTRTCRTCS ACC)	957	BWC1	MK850847	RNA polymerase II (RPB2) gene, partial cds (MH582187.1)	957/957(100%)	<i>Fusarium subglutinans</i> strain MRC 627	(2, 24)
5	ITS	ITS1(TCCGTAGGTGAACCTGCGG), ITS4(TCCTCCGCTTATTGATATGC)	480	QHM	MK898823	Isolate M108 small subunit ribosomal RNA gene, partial sequence (MK250069.1)	480/480 (100%)	<i>Fusarium fujikurio/proliferatum</i> strain TF1	(12, 26)
6	TUB2	T1(AACATGCGTGAGATTGTAAGT), T2(TAGTGACCCTTGGCCCAGTTG)	521	QHM	MK907693	strain MUCL 56019, isolate B10-XB14 (LT575139.1)	521/521 (100%)	<i>Fusarium proliferatum</i> strain BCD/verticillioides MUCL	(4, 10)
7	TEF	EF1(ATGGGTAAGGARGACAAGAC), EF2(GGARGTACCAGTSATCATG)	626	QHM	MK907692	isolate Fus_005 translation elongation factor 1-alpha (TEF1-alpha) gene (MH496632.1)	625/626 (99%)	<i>Fusarium verticillioides</i> isolate Fus	(2, 24)

447
448

8	RPB2	RPB2_5F2(GGGGWGAYCAGAAGAAG GC), RPB2_11AR(GCRTGGATCTTRTCRTCS ACC)	939	QHM	MK907691	strain MRC 930 RNA polymerase II (RPB2) gene (MH582221.1)	929/929 (100%)	<i>Fusarium verticillioides</i> strain MRC 930	(2, 24)
---	------	--	-----	-----	----------	--	----------------	---	---------

449 **Table 2. Overall results of genome assemblage**

450

	Genomic statistics															
			No. of large scaffolds	No. of bases in large scaffolds (bp)	Length of largest scaffold (bp)	Scaffold N50 (bp)	Scaffol d N90 (bp)	G+C (%)	N rate (%)	No. of contigs	No. of bases in contigs (bp)	No. of large contigs	No. of bases in large contigs (bp)	Length of the largest contig (bp)	Ctg N50 (bp)	Ctg N90 (bp)
	No. of scaffolds	No. of bases in scaffolds (bp)														
Sample																
<i>Fusarium</i>																
m sp.																
QHM	479	42164867	248	42030945	2038123	478828	129729	48.78	0.029	818	42152448	457	41974141	814451	231468	56903
<i>Fusarium</i>																
m sp.																
BWC1	2352	42063116	1765	41741716	228966	50807	11922	48.87	0.057	3634	42038762	2597	41539615	172141	32773	7241

451

452

453

454

455

456

457

458

459

460

461 **Table 3. Predictive functions of the coding genes of *Fusarium* spp. QHM and BWC1**

Sample	No. of genes	Total length of the genes (bp)	Average length of the genes (bp)	Gene density (kb)	GC content in the gene region (%)	Gene/Genome (%)	Length of the intergenetic region (bp)	GC content of the intergenetic region (%)	Intergenetic length/Genome (%)
Fusarium sp.QH									
M	14814	24109110	1627.45	0.35	50.92	57.18	18055757	45.77	42.82
Fusarium sp.BWC1									
	15295	23939480	1565.18	0.36	50.96	56.91	18123636	45.88	43.09

485 **Table 4. Xenobiotic biodegradation pathways and key enzymes of *Fusarium* spp. QHM and BWC1**

486

Xenobiotic	Pathway ID	Key enzymes involved in degradation and metabolism
Chloroalkane	ko00625	aldehyde dehydrogenase (NAD ⁺), 2-haloacid dehalogenase, alcohol dehydrogenase
Aminobenzoate	ko 00627	benzoate 4-monooxygenase, phenol 2-monooxygenase, 4-nitrophenyl phosphatase
Naphthalene	ko00626	salicylate hydroxylase, alcohol dehydrogenase, S-(hydroxymethyl)glutathione dehydrogenase
Atrazine	ko00791	Cyanamide hydratase, urease, urea carboxylase
Benzoate	ko00362	dihydropyrimidinase, xanthine dehydrogenase, carboxylesterase 2
Styrene	ko00643	homogentisate 1,2-dioxygenase, phenylacetate 2-hydroxylase, amidase
Fluorobenzoate	ko00364	catechol 1,2-dioxygenase, carboxymethylenebutenolidase
Polycyclic aromatic hydrocarbon	ko00624	salicylate hydroxylase, dibenzothiophene dihydrodiol dehydrogenase, PAH dioxygenase
Caprolactam	ko00930	3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase, gluconolactonase
Toluene	ko00623	phenol 2-monooxygenase, carboxymethylenebutenolidase, catechol 1,2-dioxygenase
Dioxin	ko00621	salicylate hydroxylase, DDT-dehydrochlorinase, biphenyl 2,3-dioxygenase
Chlorocyclohexane and chlorobenzene	ko00361	2-haloacid dehalogenase, carboxymethylenebutenolidase, catechol 1,2-dioxygenase

487

488 “ko” is the abbreviation of KEGG Orthology.

489

490

491

505

506

507 **Figure 1. *Fusarium* spp. QHM and BWC1 bar charts obtained from GO analysis**

508 The gene orthologs of *Fusarium* spp. QHM and BWC1 are similar generally. All genes were categorized into three main groups that were further divided into various
509 subcategories: biological processes, cellular components and molecular functions. The majority of genes belongs to metabolic processes, and the second highest
510 number of genes were assigned to catalytic activity group as usual.

511

512

513

514

515

516

517

518 **Figure 2. *Fusarium* sp. QHM histogram obtained from KEGG data**

530 **Figure 3. *Fusarium* sp. BWC1 histogram obtained from KEGG data**

Figure 4. *Fusarium* sp. QHM CAZy analysis chart

Figure 5. *Fusarium* sp. BWC1 CAZy analysis chart

567 **Figure 6. Maximum likelihood phylogeny tree**

568 The tree was constructed based on analysis of the TUB2, TEF and RPB2 genes of *Fusarium* spp. QHM and BWC1 of the *Fusarium fujikuroi* species complex
569 (FFSC).

570 The names in bold font represent the two analyzed strains, that is, *Fusarium* sp. QHM and *Fusarium* sp. BWC1. Bootstrap values are provided at the branch nodes.

571 *Fusarium oxysporum* CBS 774 was designated as an outgroup species in the analysis.

572

573

574

575

576

577

578

Figure 7. Comprehensive data on antibiotic resistance in *Fusarium* spp. QHM and BWC1

Overall mechanism of the antibiotic resistance were similar for both strains. 107 genes of *Fusarium* sp. QHM (left) and 106 genes of *Fusarium* sp. BWC1 (right) conferred the antibiotic resistance in a way of efflux pump.

591

592 **Figure 8. Syntenic analysis of the *Fusarium* spp. QHM and BWC1 genomes**

593 “query” stands for QHM genome; “ref” stands for BWC1.