



## Data in Brief

# The complete mitochondrial genome of the acid-tolerant fungus *Penicillium* ShG4C



Andrey V. Mardanov<sup>a,\*</sup>, Lubov B. Glukhova<sup>b</sup>, Eugeny V. Gruzdev<sup>a</sup>, Alexey V. Beletsky<sup>a</sup>,  
Olga V. Karnachuk<sup>b</sup>, Nikolai V. Ravin<sup>a</sup>

<sup>a</sup> Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, 119071 Moscow, Russia

<sup>b</sup> Tomsk State University, 634050 Tomsk, Russia

## ARTICLE INFO

## Article history:

Received 26 October 2016

Accepted 9 November 2016

Available online 10 November 2016

## ABSTRACT

Complete mitochondrial genome of the acid-tolerant fungus *Penicillium* ShG4C, isolated from oxidized sediments of an abandoned polymetallic mine site, has been sequenced using high-throughput sequencing approach. The mitochondrial genome represents a circular DNA molecule with size of 26,725 bp. It encodes a usual set of mitochondrial genes, including 15 protein coding genes, large and small ribosomal RNAs and 27 tRNA genes. All genes are located on H-strand DNA and transcribed in one direction. Taxonomic analysis based on concatenated sequences of mitochondrial proteins confirmed taxonomic position of this fungus within the genus *Penicillium*. The sequence of the complete mitochondrial genome of *Penicillium* ShG4C was deposited in DDBJ/EMBL/GenBank under accession number KX931017.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Specifications	
Organism/cell line/tissue	<i>Penicillium</i> ShG4C
Sex	N/A
Sequencer or array type	Illumina HiSeq2000
Data format	Analyzed
Experimental factors	N/A
Experimental features	Whole genome sequencing of <i>Penicillium</i> ShG4C resulted in identification of single circular contig representing the complete mitochondrial genome.
Consent	N/A
Sample source location	Open-cast polymetallic ore mine Sherlovaya Gora, Eastern Siberia, Russia (50° 31'27"; 116° 19'8").

## 1. Direct link to deposited data

<https://www.ncbi.nlm.nih.gov/nucore/KX931017>.

## 2. Introduction

Fungi of the genus *Penicillium* of *Trichocomaceae* family of *Pezizomycotina* order (filamentous fungi) of *Ascomycetes* are widely

distributed in nature. Generally they are saprophyte and use dead organic material for feed. Eukaryotic organisms could inhabit extreme environments, for example, fungi of the genus *Penicillium* sp. were found in acidic Rio Tinto in Spain and Iron Mountains in California [9,10], hot soil of Yellowstone Part in USA [12]. A few *Ascomycetes* fungi were found in Zanzjan province (Iran) as potential species for cadmium removal from soils [11]. Evolutionary adaptation of fungi to metal-contaminated soils is a well-documented phenomenon, particularly because it is one of the most striking examples of microevolution driven by edaphic factors [8]. They have considerable potential in the solution of bioremediation tasks [2]. These fungi can absorb different metals from soil and water [3], they are easily isolated, grow quickly and adjust to environmental conditions, so they have a major potential for bioremediation [2]. In this paper we present the results of sequencing and analysis of the mitochondrial genome of acid tolerant strain of fungus *Penicillium* ShG4C. The obtained data will be useful for further research in the field of taxonomy and evolution of filamentous fungi.

## 3. Experimental design, materials and methods

### 3.1. Features of the mitochondrial genome of *Penicillium* ShG4C.

In 2013 the fungal strain ShG4C belonging to the genus *Penicillium* was isolated from wastes of the ore mining deposit "Sherlovaya Gora" located in Transbaikal region, Eastern Siberia, Russia. Chemical analysis of water at the sampling site showed low pH value (1.9) and high

\* Corresponding author.

E-mail address: [mardanov@biengi.ac.ru](mailto:mardanov@biengi.ac.ru) (A.V. Mardanov).

concentrations of iron (320 mM), arsenic (39 mM), zinc (41 mM), aluminum (100 mM), and copper (31 mM). Strain ShG4C is able to grow in presence of high concentration of arsenic and metals. Due to these properties, strain ShG4C is a potential object for development of new bioremediation approaches.

Genomic DNA was extracted from mycelium by modified protocol described in works of [1]. Whole genomic DNA was sequenced using Illumina HiSeq2500 platform (10 millions of 100-bp long reads). The sequencing reads were de novo assembled into contigs using the Spades v. 3.7.1 [4]. A single circular contig with an average 1158 X coverage representing the mitochondrial genome was identified based on sequence similarity to the mitochondrial genome sequence of *Penicillium polonicum* (KU530219). Identification of protein-coding genes, ribosomal and tRNA genes was carried out using Mitos server [5] and tRNAscan-SE [6]. The obtained automatic annotation was checked and corrected manually using BLAST search against the NCBI sequence database (<http://www.ncbi.nlm.nih.gov/genbank/>).

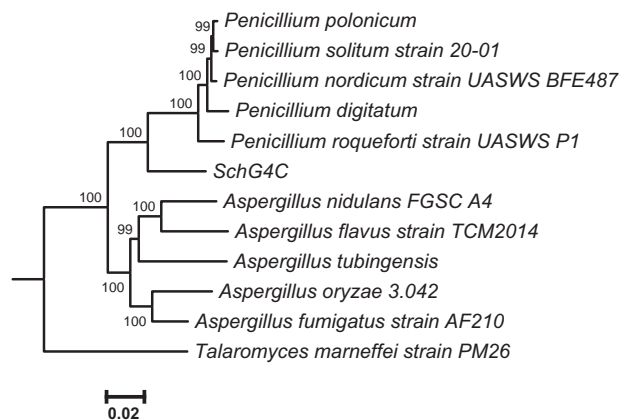
The complete mitochondrial genome of *Penicillium* ShG4C was a circular 26,725 bp long DNA. Its size is comparable to mtDNA of other closely related fungi of genus *Penicillium*, e.g. mtDNA of *Penicillium polonicum* – 28,192 bp (NC\_030172), and *Penicillium roqueforti* – 29,908 bp (KR952335). The standard set of genes, including 15 protein-coding genes, 27 tRNA genes and 2 genes of ribosomal RNA is

**Table 1**  
Mitochondrial genome of *Penicillium* ShG4C.

Gene	Start	Stop	Length, bp
rrnL	113	4705	4593 <sup>a</sup>
RPS <sup>b</sup>	2861	4072	1212
trnT-UGU	4753	4823	71
trnE-UUC	4861	4934	74
trnV-UAC	4936	5008	73
trnM-CAU	5010	5080	71
trnM-CAU	5081	5153	73
trnL-UAA	5158	5239	82
trnA-UGC	5245	5316	72
trnF-GAA	5564	5636	73
trnL-UAG	5649	5731	83
trnQ-UUG	5744	5816	73
trnM-CAU	5820	5891	72
trnC-GCA	5915	5985	71
trnH-GUG	6118	6188	71
COX1	6376	8088	1713
ATP9	8452	8676	225
trnN-GUU	8731	8801	71
NAD3	8990	9397	408
COX2	9537	10,298	762
trnR-ACG	10,404	10,474	71
NAD4L	10,744	11,013	270
NAD5	11,013	12,992	1980
NAD2	13,047	14,735	1689
COB	16,400	17,557	1158
trnY-GUA	17,648	17,713	66
NAD1	17,897	18,952	1056
NAD4	19,213	20,679	1467
trnR-UUC	20,749	20,819	71
trnN-GUU	20,851	20,921	71
ATP8	21,045	21,191	147
ATP6	21,356	22,129	774
rrnS	22,660	24,049	1390
trnY-GUA	24,177	24,261	85
NAD6	24,349	25,002	654
COX3	25,052	25,861	810
trnK-UUU	25,900	25,971	72
trnG-ACC	26,014	26,084	71
trnG-UCC	26,105	26,175	71
trnD-GUC	26,188	26,260	73
trnS-GCT	26,314	26,394	81
trnW-UCA	26,395	26,466	72
trnI-GAU	26,483	26,554	72
trnS-UGA	26,559	26,644	86

<sup>a</sup> Contains intron (2517–4200).

<sup>b</sup> Gene encoding ribosomal protein, located within rrnL intron.



**Fig. 1.** Phylogenetic analysis of representatives of the genera *Aspergillus* and *Penicillium*.

encoded by the mitochondrial genome of *Penicillium* ShG4C (Table 1). All identified genes are encoded on H-strand of mtDNA. All protein-encoding genes have the same start codon ATG, except for COX1 gene with TTG start codon. NAD6 gene has stop codon TAG and the other genes have stop codon TAA. The GC content of *Penicillium* ShG4C mtDNA is 25%, like in other mitochondrial genomes of *Penicillium* (e.g. *Penicillium polonicum* – 25.56%, NC\_030172).

To verify the taxonomic position of the new strain, we made phylogenetic analysis based on concatenated sequences of 14 mitochondrial proteins, including the respiratory chain (COX1–COX3, COB), three subunits of ATPase (ATP6, ATP8, and ATP9) and seven subunits of NADH-quinone reductase (NAD1–NAD4, NAD4L, NAD5 and NAD6). Alignments of amino acid sequences and construction of phylogenetic tree by a Maximum-likelihood method were performed using MEGA6 package [7]. Phylogenetic analysis confirmed affiliation of the strain ShG4C with the genus *Penicillium*, where it forms distinct species-level lineage at the base of this branch (Fig. 1).

The maximum-likelihood phylogenetic tree was constructed using concatenated amino acid sequences of 14 mitochondrial proteins. Numbers at nodes indicates bootstrap support values from 500 replicates. The tree is drawn to scale, with branch lengths measured by the number of substitutions per site. The following mtDNAs were used to build the tree: *Aspergillus oryzae* (JX129489), *Aspergillus tubingensis* (DQ217399), *Aspergillus nidulans* (JQ435097), *Aspergillus flavus* (KP725058), *Aspergillus fumigatus* (JQ346809), *Penicillium solitum* (JN696111), *Penicillium digitatum* (HQ622809), *Penicillium polonicum* (KU530219), *Penicillium nordicum* (KR952336), *Penicillium roqueforti* (KR952335), and *Talaromyces marneffeii* (AY347307).

In conclusion, we have sequenced the complete mitochondrial genome of the acid-tolerant fungal strain *Penicillium* ShG4C. The obtained genomic data will be useful for further taxonomic and evolutionary studies of filamentous fungi.

## Conflict of interest

The authors declare no conflicts of interest in this study.

## Acknowledgments

This work was funded by grant of the President of Russian Federation for Young Scientists (project MD-7668.2016.4).

## References

- [1] E.M. Möller, G. Bahnweg, H. Sandermann, H.H. Geiger, A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic Acids Res.* 20 (1992) 6115–6116.
- [2] A. Pal, S. Ghosh, A.K. Paul, Biosorption of cobalt by fungi from serpentine soil of Andaman. *Bioresour. Technol.* 97 (10) (2006) 1253–1258.

- [3] D.S. Hibbett, J.W. Taylor, Fungal systematics: is a new age of enlightenment at hand? *Nat. Rev. Microbiol.* 11 (2) (2013) 129–133.
- [4] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19 (5) (2012) 455–477.
- [5] M. Berni, A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsche, J. Pütz, M. Middendorf, P.F. Stadler, MITOS: improved de novo metazoan mitochondrial genome annotation *Mol. Phylogenet. Evol.* 69 (2013) 313–319.
- [6] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25 (1997) 955–964.
- [7] K. Tamura, G. Stecher, D. Peterson, A. Filipinski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30 (2013) 2725–2729.
- [8] J.V. Colpaert, P. Vandenkoornhuys, K. Adriaensen, J. Vangronsveld, Genetic variation and heavy metal tolerance in the ectomycorrhizal basidiomycete *Suillus luteus*. *New Phytol.* 147 (2000) 367–379.
- [9] Z. Amaral, F. Gómez, E. Zettler, B.G. Keenan, R. Amils, M.L. Sogin, Microbiology: eukaryotic diversity in Spain's River of Fire. *Nature* 417 (6885) (2002) 137.
- [10] B.J. Baker, M.A. Lutz, S.C. Dawson, P.L. Bond, J.F. Banfield, Metabolically active eukaryotic communities in extremely acidic mine drainage. *Appl. Environ. Microbiol.* 70 (10) (2004) 6264–6271.
- [11] M. Mohammadian Fazli, N. Soleimani, M. Mehrasbi, S. Darabian, J. Mohammadi, A. Ramazani, Highly cadmium tolerant fungi: their tolerance and removal potential. 2015. *J. Environ. Health Sci. Eng.* 13 19, <http://dx.doi.org/10.1186/s40201-015-0176-0>.
- [12] R.S. Redman, A. Litvintseva, K.B. Sheehan, J.M. Henson, R.J. Rodriguez, Fungi from geothermal soils in Yellowstone National Park. *Appl. Environ. Microbiol.* 65 (12) (1999) 5193–5197.