Data Article

Data on genetic polymorphism of flax (Linum usitatissimum L.) pathogenic fungi of Fusarium, Colletotrichum, Aureobasidium, Septoria, and Melampsora genera

Roman O. Novakovskiy a,*, Ekaterina M. Dvorianinova a,b,#, Tatiana A. Rozhmina a,c, Ludmila P. Kudryavtseva c, Aleksey A. Gryzunov d, Elena N. Pushkova a, Liubov V. Povkhova a,b, Anastasiya V. Snezhkina a, George S. Krasnov a, Anna V. Kudryavtseva a, Nataliya V. Melnikova a,*, Alexey A. Dmitriev a

a Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 119991, Russia
b Moscow Institute of Physics and Technology, Dolgoprudny 141701, Russia
c Federal Research Center for Bast Fiber Crops, Torzhok 172002, Russia
d All-Russian Scientific Research Institute of Refrigeration Industry – branch of V.M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences, Moscow 127422, Russia

A R T I C L E   I N F O

Article history:
Received 3 March 2020
Revised 28 April 2020
Accepted 8 May 2020
Available online 15 May 2020

Keywords:
Fungi
Phytopathogens
Fusarium
Genetic polymorphism
Deep sequencing
Amplicon sequencing
Flax
Linum usitatissimum L

A B S T R A C T

Being a valuable agricultural plant, flax (Linum usitatissimum L.) is used for oil and fiber production. However, the cultivation of this agriculture faces an urgent problem of flax susceptibility to fungal diseases. The most destructive ones are caused by the representatives of Fusarium, Colletotrichum, Aureobasidium, Septoria, and Melampsora genera, reducing flax yields significantly. To combat such pathogens effectively, it is of high importance to assess their genetic diversity that can be used to develop molecular markers to distinguish fungal genera and species. Morphological analysis traditionally carried out for fungal identification requires a given amount of time and tends to be difficult. In the present work, we

https://doi.org/10.1016/j.dib.2020.105710
2352-3409/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)
determined the DNA sequences that are frequently used for phylogenetic studies in fungi – internal transcribed spacer (ITS) and beta-tubulin (tub2), translation elongation factor 1-alpha (tef1), RNA polymerase II largest subunit (RPB1), RNA polymerase II second largest subunit (RPB2), and minichromosome maintenance protein (MCM7) genes – for 203 flax fungal pathogens of Fusarium oxysporum, F. avenaceum, F. solani, F. sporotrichiella, F. moniliforme, F. culmorum, F. semitectum, F. gibbosum, Colletotrichum lini, Aureobasidium pullulans, Septoria linicola, and Melampsora lini species. The sequencing was performed using the Illumina MiSeq platform with a 300+300 bp kit, and on average, about 2350 reads per sample were obtained that allows accurate identification of the genetic polymorphism. Raw data are stored at the Sequence Read Archive under the accession number PRJNA596387. The obtained data can be used for fungal phylogenetic studies and the development of a PCR-based test system for flax pathogen identification.

© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

### Specifications table

| Subject                  | Genetics          |
|--------------------------|-------------------|
| Specific subject area    | Molecular Genetics, Agricultural Mycology |
| Type of data             | Amplicon sequence data |
| How data were acquired   | Illumina MiSeq, 300+300 bp paired-end reads |
| Data format              | Raw sequence reads, fastq format |
| Parameters for data collection | DNA was extracted from 203 flax (Linum usitatissimum L.) fungal pathogen samples of Fusarium oxysporum, F. avenaceum, F. solani, F. sporotrichiella, F. moniliforme, F. culmorum, F. semitectum, F. gibbosum, Colletotrichum lini, Aureobasidium pullulans, Septoria linicola, and Melampsora lini species obtained from the phytopathogen collection of the Institute for Flax (Torzhok, Russia) |
| Description of data collection | Amplicon libraries of fragments of internal transcribed spacer (ITS) and beta-tubulin (tub2), translation elongation factor 1-alpha (tef1), RNA polymerase II largest subunit (RPB1), RNA polymerase II second largest subunit (RPB2), and minichromosome maintenance protein (MCM7) genes of flax fungal pathogens of Fusarium, Colletotrichum, Aureobasidium, Septoria, and Melampsora genera were prepared using two-step PCR and sequenced on Illumina MiSeq with a 600-cycle kit |
| Data source location     | Institute for Flax, Torzhok, Russia |
| Data accessibility       | One can access amplicon sequence data at the NCBI Sequence Read Archive (SRA) under the accession number PRJNA596387 (https://www.ncbi.nlm.nih.gov/sra/PRJNA596387) |

### Value of the data

- The dataset could be actively used for the assessment of genetic diversity and phylogenetic investigations of fungi belonging to Fusarium, Colletotrichum, Aureobasidium, Septoria, and Melampsora genera.
• The data can support those working in the field of molecular genetics of fungi and those who work with plant fungal pathogens.
• The kinship and evolution of fungi of *Fusarium, Colletotrichum, Aureobasidium, Septoria*, and *Melampsora* genera could be estimated using the provided dataset.
• The generated data open up an opportunity to create a test system for the identification of flax pathogens basing on the information on genetic polymorphism.

1. Data Description

Being a source of oil and fiber [1], flax is affected by numerous fungal pathogens. The most serious diseases of this plant are caused by the representatives of such species as *Fusarium oxysporum* f. sp. *lindi*, *F. avenaceum*, *F. culmorum*, *Melampsora* *linci*, *Colletotrichum linci*, *Septoria linciola*, and *Aureobasidium pullulans* [2,3]. Information on the genetic diversity of the listed species is lacking but is of high importance for the determination of their kinship and evolution, as well as for the development of effective ways to combat these flax pathogens.

For 203 samples of flax pathogens of *Fusarium, Colletotrichum, Aureobasidium, Septoria*, and *Melampsora* genera, we sequenced the fragments of the regions that are frequently used for phylogenetic studies in fungi: internal transcribed spacer (ITS) [4,5], beta-tubulin (*tub2*), translation elongation factor 1-alpha (*tefl*), RNA polymerase II largest subunit (*RPB1*), RNA polymerase II second largest subunit (*RPB2*), and minichromosome maintenance protein (*MCM7*) [6-8]. Amplicon libraries were prepared using two-stage PCR (primers for the first stage of PCR are listed in Supplementary Table S1) and deep sequenced on the Illuma MiSeq platform. On average, 2350 (range – 400-6000) paired-end reads (300+300 bp) per sample were obtained in the fastq format and deposited to the Sequence Read Archive (SRA) under the accession number PRJNA596387. These data play a key role in assessing the genetic diversity of flax fungal pathogens and the development of time-saving and accurate test systems for their identification. Basing on the given dataset, future experiments are likely to give further insight into the question of the kinship and evolution of fungi of the studied genera.

2. Experimental Design, Materials, and Methods

2.1. Material

203 samples of 12 fungal species were provided by the Institute for Flax (Torzhok, Russia) from their phytopathogen collection in 2018. Among them were *Fusarium oxysporum* (72 samples), *F. avenaceum* (23), *F. solani* (5), *F. sporotrichiella* (3), *F. moniliforme* (8), *F. culmorum* (5), *F. semitectum* (2), *F. gibbosum* (4), *Colletotrichum linci* (46), *Aureobasidium pullulans* (17), *Septoria linciola* (8), and *Melampsora linci* (10) species.

2.2. DNA extraction and quality control

DNA was extracted from fungal mycelium according to the standard CTAB protocol. Agarose gel electrophoresis (2% agarose) and the Qubit 2.0 fluorometer (Thermo Fisher Scientific, USA) were used to control DNA quality and evaluate DNA quantity.

2.3. DNA library preparation and sequencing

We amplified DNA fragments of 203 samples of fungi comprising ITS and *tub2, tefl, RPB1, RPB2*, and *MCM7* genes. Amplicon libraries were prepared according to the protocol [9] with
two-stage PCR as we described earlier [10]. Amplification of target sequences using primers that comprised target-specific sequences [8,11] and overhang adapters was performed in the first step (Supplementary Table S1). For each sample, amplicons were then equimolarly pooled and the second PCR was performed with primers consisted of dual-index barcodes and sequencing adapters. Next, all PCR-products were equimolarly pooled and the quality of the library was evaluated by the 2100 Bioanalyzer (Agilent Technologies, USA), while the quantity – by the Qubit 2.0 fluorometer (Thermo Fisher Scientific). The library was sequenced using the MiSeq platform (Illumina, USA) and the Illumina MiSeq Reagent Kit v3 (2 × 300 bp reads).

Acknowledgments

Part of this work (Sequencing of Fusarium oxysporum strains) was financially supported by the Russian Science Foundation [grant number 16-16-00114]. Part of this work (Sequencing of the strains of Fusarium, Colletotrichum, Aureobasidium, Septoria, and Melampsora genera) was financially supported by the Russian President Grant [grant number MK-5828.2018.4]. The maintenance of flax pathogen collection is carried out under the financial support of the Ministry of Science and Higher Education of the Russian Federation [state assignment number 075-00853-19-00]. We thank the Center for Precision Genome Editing and Genetic Technologies for Biomedicine, EIMB RAS for providing the techniques for targeted deep sequencing. The sequencing was performed using the equipment of EIMB RAS “Genome” center (http://www.eimb.ru/rul/ckp/ccu_genome_ce.php).

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships which could have, or could be perceived to have, influenced the work reported in this article.

Supplementary materials

Supplementary article associated with this article can be found in the online version, at doi: 10.1016/j.dib.2020.105710.

References

[1] C.A. Cullis, Genetics and Genomics of Linum, Springer, Cham, 2019.
[2] R.F. Nyvall, Diseases of flax, in: R.F. Nyvall (Ed.), Field Crop Diseases Handbook, Springer, Boston, 1989, pp. 251–264.
[3] E. Gruzdievienė, I. Brazauskienė, J. Repėkšlienė, A. Lugauskas, The occurrence of pathogenic fungi during flax growing season in Central Lithuania, J Plant Prot Res 48 (2008) 255–266, doi:10.2478/v10045-008-0029-2.
[4] C.L. Schoch, K.A. Seifert, S. Huhndorf, V. Robert, J.L. Spouge, C.A. Levesque, W. Chen, C. Fungal Barcoding Consortium, Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi, Proc Natl Acad Sci U S A 109 (2012) 6241–6246, doi:10.1073/pnas.1117018109.
[5] T.M. Pryce, S. Palladino, I.D. Kay, G.W. Coombs, Rapid identification of fungi by sequencing the ITS1 and ITS2 regions using an automated capillary electrophoresis system, Med Mycol 41 (2003) 369–381, doi:10.1080/13693780310001600435.
[6] P.W. Croux, D.L. Hawksworth, M.J. Wingfield, Identifying and naming plant-pathogenic fungi: past, present, and future, Annu Rev Phytopathol 53 (2015) 247–267, doi:10.1146/annurev-phyto-080614-120245.
[7] K. O’Donnell, D.A. Sutton, M.G. Rinaldi, B.A. Sarver, S.A. Balajee, H.J. Schroers, R.C. Summerbell, V.A. Robert, P.W. Croux, N. Zhang, T. Aoki, K. Jung, J. Park, Y.H. Lee, S. Kang, B. Park, D.M. Geiser, Internet-accessible DNA sequence database for identifying fusaria from human and animal infections, J Clin Microbiol 48 (2010) 3708–3718, doi:10.1128/JCM.00989-10.
[8] H.A. Raja, A.N. Miller, C.J. Pearce, N.H. Oberlies, Fungal Identification Using Molecular Tools: A Primer for the Natural Products Research Community, J Nat Prod 80 (2017) 756–770, doi:10.1021/acs.jnatprod.6b01085.
[9] Illumina Support Center, 16S Metagenomic Sequencing Library Preparation. support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html, 2019 (Accessed 20 November 2019).
[10] N.V. Melnikova, A.V. Kudryavtseva, E.V. Borkhert, E.N. Pushkova, M.S. Fedorova, A.V. Snezhkina, G.S. Krasnov, A.A. Dmitriev, Sex-specific polymorphism of MET1 and ARR17 genes in Populus x sibirica, Biochimie 162 (2019) 26–32, doi:10.1016/j.biochi.2019.03.018.

[11] P.B. Matheny, Z. Wang, M. Binder, J.M. Curtis, Y.W. Lim, R.H. Nilsson, K.W. Hughes, V. Hofstetter, J.F. Ammirati, C.L. Schoch, E. Langer, G. Langer, D.J. McLaughlin, A.W. Wilson, T. Froslav, Z.W. Ge, R.W. Kerrigan, J.C. Slot, Z.L. Yang, T.J. Baroni, M. Fischer, K. Hosaka, K. Matsuura, M.T. Seidl, J. Vauras, D.S. Hibbett, Contributions of rpb2 and tef1 to the phylogeny of mushrooms and allies (Basidiomycota, Fungi), Mol Phylogenet E 43 (2007) 430–451, doi:10.1016/j.ympev.2006.08.024.