Use of deep sequencing to study the genetic diversity of flax pathogens

R.O. Novakovskiy[1], G.S. Krasnov[1], E.N. Pushkova[1], L.P. Kudryavtseva[2], T.A. Rozhmina[1, 2], P. Kezimana[1, 3], L.V. Povkhova[1, 4], N.V. Melnikova[1], A.A. Dmitriev[1]*

[1] *Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia*
[2] *Federal Research Center for Bast Fiber Crops, Torzhok, Russia*
[3] *Peoples' Friendship University of Russia (RUDN University), Moscow, Russia*
[4] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*

Abstract: Pathogens decrease flax yield and reduce product quality. The development of molecular markers is necessary for fungus identification and application of proper defense actions. In our work, we present the method based on deep sequencing for studying the genetic diversity of fungal pathogens of flax. ITS regions of rRNA genes and regions of genes encoding beta-tubulin (tub2), translation elongation factor 1-alpha (tef1), and RNA polymerase II subunits (RPB1 and RPB2) were studied using deep sequencing in 100 pathogen strains of *Fusarium*, *Colletotrichum*, *Melampsora*, *Aureobasidium*, and *Septoria*. Our method can be used for the characterization of genetic diversity of pathogens and determination of DNA sequences that are the most suitable for identification of fungus using molecular markers.

Key words: deep sequencing; flax pathogens; genetic diversity.

1. Introduction

Flax is a valuable crop which has been attracting more and more attention in recent years. Flax fiber is used to produce high-quality fabrics and composite materials (Costa et al., 2018). Flax seed has a complex of useful properties and is used in the pharmaceutical and food industries, as well as in the production of feed, high-quality paints, and polymeric materials (Singh et al., 2011). Diseases, primarily caused by fungal pathogens, lead to significant crop losses and reduce product quality. The main pathogens of flax are *Fusarium oxysporum*, *Melampsora lini*, *Colletotrichum lini*, *Septoria linicola*, and *Aureobasidium pullulans* (Muir, Westcott, 2003).

To combat flax diseases, effective measures are necessary, among which the identification of pathogen for choosing plant protection actions is an important one. Molecular markers are an effective tool for identifying pathogens (Crous et al., 2015), but the choice of nucleotide sequences that will be used for genetic diagnosis of pathogens is crucial. In the present work, the method for assessment of polymorphism of particular genes in fungal pathogens of flax using deep sequencing was proposed.

2. Materials and methods

Strains of *F. oxysporum*, *M. lini*, *C. lini*, *S. linicola*, *A. pullulans*, *Fusarium moniliforme*, *Fusarium solani*, *Fusarium culmorum*, *Fusarium gibbosum*, *Fusarium sporotrichiella*, *Fusarium avenaceum*, and *Fusarium semitectum* were obtained from the Institute for Flax (Torzhok, Russia). To isolate the DNA, the fungal samples were ground using a MagNA Lyser homogenizer (Roche, Switzerland) in a CTAB buffer and phenol-chloroform purification was performed. One hundred DNA samples were isolated, the quality and concentration of which were determined by agarose gel electrophoresis and a Qubit 2.0 fluorometer (Life Technologies, USA). The internal

transcribed spacer (ITS) regions of the rRNA genes and regions of genes encoding beta-tubulin (tub2), translation elongation factor 1-alpha (tef1), and RNA polymerase II subunits (RPB1 and RPB2) were chosen to study the genetic diversity of flax pathogens with. The analysis of the literature data was performed and the sequences of primers that allow amplification of selected sequences from various species and genera of pathogens causing fungal plant diseases and are actively used in phylogenetic studies were found (Raja et al., 2017). To prepare DNA libraries of selected genes for deep sequencing, universal adapters were added to the target-specific primers. This is necessary for sequencing and performing the second PCR, which is needed for indexing the samples and adding the sequences required for sequencing. The Illumina protocol was taken as a basis (support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html).

For the second PCR, universal Nextera XT v2 primers were used. The PCR conditions were optimized with selected primer pairs. The best results were obtained with amplification in 20 μl of the reaction mixture containing 2 units of high-precision Tersus polymerase (Evrogen, Russia), 1× Tersus polymerase buffer (Evrogen), 200 μM dNTP (Thermo Fisher Scientific, USA), 0.5 μM of each primer (Evrogen), and 20 ng of DNA under the following conditions: 3 min at 95 °C; 35 cycles for the first PCR and 12 cycles for the second PCR: 15 s at 95 °C, 30 s at 58 °C, 30 s at 72 °C; 3 min at 72 °C. After the first PCR and after the second PCR, purification of PCR products using CleanMag DNA (Evrogen) magnetic beads was performed. According to the developed method, ITS, *tub2*, *tef1*, *RPB1*, and *RPB2* sequences were amplified for 100 samples of pathogens causing fungal flax diseases, and DNA libraries were prepared for deep sequencing. The quality and concentration of the obtained DNA libraries were evaluated on an Agilent 2100 bioanalyzer (Agilent, USA) and a Qubit 2.0 fluorometer. The described sample preparation

allowed us to obtain high-quality DNA libraries, which were equimolarly mixed before sequencing.

Using the MiSeq sequencer (Illumina, USA), we performed sequencing of ITS and *tub2*, *tef1*, *RPB1*, and *RPB2* genes for 100 samples of the fungal flax pathogens: *F. oxysporum*, *F. avenaceum*, *F. culmorum*, *F. moniliforme*, *F. gibbosum*, *F. semitectum*, *F. sporotrichiella*, *F. solani*, *M. lini*, *C. lini*, *S. linicola*, and *A. pullulans*. The read length was 300 + 300 nucleotides. The pipeline was developed for the analysis of the sequencing data. This work was performed using the equipment of "Genome" Center of Engelhardt Institute of Molecular Biology (http://www.eimb.ru/rus/ckp/ccu_genome_ce.php).

3. Results and discussion

Although the use of ITS to determine the species of fungi is the most promising approach (Pryce et al., 2003; Schoch et al., 2012), sometimes the use of only these sequences has an insufficient resolution. In this case, it is advisable to use them in combination with protein-coding genes, such as *tub2* and *tef1*, and to perform an additional analysis of *RPB1* and *RPB2* also having a high interspecific diversity but poorly amplified for a number of fungi (Schoch et al., 2012; Crous et al., 2015; Raja et al., 2017).

Approaches for identifying fungal pathogens are being actively developed and used for different plant species (McCartney et al., 2003; Atkins, Clark, 2004; Divakara et al., 2014). However, for pathogens of flax, such work has not yet been performed. This is largely due to the lack of data on polymorphism of nucleotide sequences that are the most suitable for the identification of pathogens. The data on polymorphism of such genes on representative sample sets of strains of fungi of various species would allow molecular diagnostics of flax pathogens.

The proposed approach based on deep sequencing enables the analysis of up to 384 pathogen samples in one run of a high-throughput sequencer. The samples will be separated by bioinformatics analysis due to the presence of double indexes. The approach is especially effective when sequencing several amplicons for a large number of samples takes place (in our case, 5 amplicons for 100 samples). This allows one to reduce both financial and labor costs, as well as to obtain a high coverage (a thousand-fold or more) to accurately identify DNA polymorphism.

4. Conclusions

The approach based on deep sequencing makes it possible to effectively assess the genetic diversity of pathogens causing fungal flax diseases and to determine DNA sequences that are the most promising for use as genetic markers for identification of plant pathogens.

References

Atkins S.D., Clark I.M. Fungal molecular diagnostics: a mini review. *J Appl Genet*. 2004;45(1):3–15.

Costa S.M., Ferreira D.P., Ferreira A. et al. Multifunctional flax fibres based on the combined effect of silver and zinc oxide (Ag/ZnO) *Nanostructures. Nanomaterials* (*Basel*). 2018;8(12). DOI 10.3390/nano8121069.

Crous P.W., Hawksworth D.L., Wingfield M.J. Identifying and naming plant-pathogenic fungi: past, present, and future. *Annu Rev Phytopathol*. 2015;53:247–267. DOI 10.1146/annurev-phyto-080614-120245.

Divakara S.T., Santosh P., Aiyaz M. et al. Molecular identification and characterization of Fusarium spp. associated with sorghum seeds. *J Sci Food Agric*. 2014;94(6):1132–1139. DOI 10.1002/jsfa.6380.

McCartney H.A., Foster S.J., Fraaije B.A. et al. Molecular diagnostics for fungal plant pathogens. *Pest Manag Sci*. 2003;59(2):129–142. DOI 10.1002/ps.575.

Muir A.D., Westcott N.D. Flax: The genus Linum: CRC Press, 2003.

Pryce T.M., Palladino S., Kay I.D. et al. Rapid identification of fungi by sequencing the ITS1 and ITS2 regions using an automated capillary electrophoresis system. *Med Mycol*. 2003;41(5):369–381. DOI 10.1080/13693780310001600435.

Raja H.A., Miller A.N., Pearce C.J. et al. Fungal identification using molecular tools: A primer for the natural products research community. *J Nat Prod*. 2017;80(3):756–770. DOI 10.1021/acs.jnatprod.6b01085.

Schoch C.L., Seifert K.A., Huhndorf S. et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci USA*. 2012;109(16):6241–6246. DOI 10.1073/pnas.1117018109.

Singh K.K., Mridula D., Rehal J. et al. Flaxseed: a potential source of food, feed and fiber. *Crit Rev Food Sci Nutr*. 2011;51(3):210–222. DOI 10.1080/10408390903537241.

Conflict of interest. The authors declare no conflict of interest.