DOI 10.18699/SBB-2022-354

Pangenome construction and analyses for a population of flax parasitic fungus *Fusarium oxysporum* f. sp. *lini*

Logachev A.[1]*, Kanapin A.[2], Samsonova A.[2], Bankin M.[1], Stanin V.[1], Rozhmina T.[3], Samsonova M.[1]

[1] *Mathematical Biology & Bioinformatics Laboratory, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

[2] *Center for Computational Biology, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

[3] *Laboratory of Breeding Technologies, Federal Research Center for Bast Fiber Crops, Torzhok, Russia*

** loga4ov@gmail.com*

Key words: population genomics, pangenome, Fusarium, fusariosis, flax, virulence

Pangenome usually defined as a full gene repertoire of a species, infraspecific taxa, or population [1]. Upon this conception pangenome divides on core genes, shared between all individuals, and accessory genes, which are absent in one or more individuals [2]. Application of this term to microorganisms is very useful since some of them have significant variation in gene content between individuals within a species. Sordariomycete, *Fusarium oxysporum* Shlecht., is widely distributed soil-born fungus that may appear as saprophyte, or have parasitic lifestyle, infecting different plants, including economically important crops. Distinction between saprotrophic and parasitic forms of this species lies in different gene content possessing by this forms. Phytopathogenic special forms (*forma specialis – f.sp.*) of *F. oxysporum* have very strict specificity to their host-plants, that is also determined by differences in genomes, and more than 200 of special forms are described by today. Earlier we reported about sequencing and analysis of 5 *Foli* genomes with chromosome-level assemblies and described that karyotype of *Foli* divides at least on two parts: conserved chromosomes composed on sequences with shared synteny to other special forms, and variable chromosomes that lack of such synteny [3, 4].

Here we report the pangenome construction and characteristics for *F. oxysporum* f.sp. *lini* (*Foli*), that infects flax and cause fusariose wilt of plant, based on full genome sequences of 13 strains. We conducted a population genomic analysis of different pangenome gene categories. We determined a repertoire of *SIX* (*Secreted In Xylem*) genes, and virulence of these strains. We also determined MAT-idiomorphs and vegetative compatibility (VC) groups. To discover the relations between observed *Foli* strains and to other special forms we inferred phylogenies on selected single genes, SNPs, and on protein-coding genes.

Methods: We sequenced 3 genomes of low-virulent *Foli* isolates with Illumina technology, and one of them was also sequenced with PacBio technology. Together with 5 genomes we published earlier [3] and 5 other available *Foli* genomes [5], these 3 new genomes were included in population genomic analysis. All investigated 13 isolates were provided by Federal Research Center for Bast Fiber Crops. Pangenome construction was carried out based on genes identified in complete genomes of 13 *Foli* isolates. We performed resampling of all possible combinations of 1–13 genomes and modelled the core genome and pangenome size curves [2]. We also conducted population genomic

analyses of different gene categories according to [6, 7]. Chromosomal synteny between the genomes was analysed by blastn pairwise alignment and visualised with R.
SIX genes (*SIX1 – SIX14*), coding small effector proteins, were determined by blastn searching for homologous sequences on 13 *Foli* genomes. The results of the search were then confirmed by PCR with respective primers for each of 14 *SIX* genes on genomic DNA of 11 strains. To investigate the expression level of *SIX* genes during colonisation of plant we used RNAseq data on mycelium culture, 3rd and 5th day of flax infection, obtained earlier from infection experiments with flax seedlings. Phylogenetic trees were constructed for each *SIX* gene found in our 13 strains with IQ-TREE.
MAT-idiomorphs were determined by blastn search through all 13 genomes with known sequences of MAT-locuses for *F. oxysporum* (GeneBank: AB011379 for MAT1-1 idiomorph, and AB011378 for MAT1-2 idiomorph) [8]. Phylogeny was inferred on *MAT1-2-1* gene sequences with JC69 model, and maximum likelihood tree was constructed with IQ-TREE. Expression level of *MAT1-2-1* gene in culture and during infection was determined by RNAseq data. Phylogenomic analyses of 13 *Foli* strains completed with 2 other *Fusarium* genomes used as outgroups were performed with protein sequences found only in one copy in each genome based on a cluster analysis. The identified single-copy orthologs then were aligned and ML phylogeny were inferred using RaxML.
Results: Pangenome size estimated on 13 genomes comprises of 14283 genes. We found that the core genome stabilised at approximately 967 genes. Number of accessory genes did not stabilise and increased almost linearly, and was equal to 10337 on 13 genomes. Number of singletons in each genomes varied from 138 to 453. CAZymes were mostly belong to core genes, genes coding secretory proteins relate to core and accessory genes while effectors belong to accessory genes and even singletons. Repertoire of *SIX* genes in *Foli* pangenome comprises *SIX1, SIX7, SIX10, SIX12* and *SIX13*. Two analysed isolates had no *SIX* genes at all, four isolates did not possess *SIX7* and *SIX12*, one isolate lack of only *SIX7* gene. Phylogenetic analysis on found *SIX* genes shown clear clusterization of strains by their special forms.
We found that 12 of 13 *Foli* isolates shared MAT1-2 locus with only one intact *MAT1-2-1* gene with 2 introns exactly as appear in other heterothallic sordariomycetes. One isolate had MAT1-1 idiomorph and possessed 3 genes inside the locus: *MAT1-1-1, MAT1-1-2* and *MAT1-1-3*. Phylogeny that was inferred for 12 MAT1-2 strains on *MAT1-2-1* gene depicts real origin of investigated strains and demonstrates 2 large distinct clades that include 11 *Foli*, and 1 small clade with only one *Foli* strain. Moreover, we demonstrate that *MAT1-2-1* gene do expressed in mycelium of MI39 isolate despite *Foli* does not have reported sexual reproduction.
Conclusions: We found that in *Foli* pangenome accessory genes occupied approximately 72 % of all genes, while core genes comprises only 6 %, and remaining 22 % were singletons. This fact reveals that *Foli* has highly plastic genome with prevalence of genes not shared with all individuals in the population. Isolates with high virulence have set of 5 different *SIX* genes while low-virulence strains lack at least *SIX7* gene. Phylogenomic analyses shown that investigated *Foli* population divides on 2 big distinct clades and 2 small clades each consisted each only of one *Foli* isolate. MAT-locus analyses reveal that all individuals possessed MAT1-2 idiomorph and the population is quite homogenous on mating type. Our first attempt to construct pangenome for *F. oxysporum f.sp. lini* provides new insights into genomic studies on populations and give rise to further investigations on phytopathogenic fungi.

Acknowledgements: The study is supported by the Russian Science Foundation, grant No. 19-16-00030.

References

1. Hurgobin B., Edwards D. SNP discovery using a pangenome: has the single reference approach become obsolete? *Biology.* 2017;6(1):21.
2. Plissonneau C., Hartmann F.E. Croll D. Pangenome analyses of the wheat pathogen Zymoseptoria tritici reveal the structural basis of a highly plastic eukaryotic genome. *BMC Biology.* 2018;16(1):1-16.
3. Kanapin A., Samsonova A., Rozhmina T., Bankin M., Logache A., Samsonova M. The genome sequence of five highly pathogenic isolates of *Fusarium oxysporum* f. sp. lini. *Mol Plant-Microbe Interactions.* 2020;33(9):1112-1115.
4. Samsonova A., Kanapin A., Bankin M., Logachev A., Gretsova M., Rozhmina T., Samsonova M. A genomic blueprint of flax fungal parasite *Fusarium oxysporum* f. sp. lini. *Int J Mol Sci.* 2021;22(5):2665.
5. Dvorianinova E.M., Pushkova E.N., Novakovskiy R.O., Povkhova L.V., Bolsheva N.L., Kudryavtseva L.P., Rozhmina T.A., Melnikova N.V., Dmitriev A.A. Nanopore and Illumina genome sequencing of *Fusarium oxysporum* f. sp. lini strains of different virulence. *Front Genet.* 2021;12:662928.
6. Badet T., Oggenfuss U., Abraham L., McDonald B.A., Croll D. A 19-isolate reference-quality global pangenome for the fungal wheat pathogen *Zymoseptoria tritici*. *BMC Biology.* 2020;18(1):1-18.
7. Alouane T., Rimbert H., Bormann J. et al. Comparative genomics of eight fusarium graminearum strains with contrasting aggressiveness reveals an expanded open pangenome and extended effector content signatures. *Int J Mol Sci*, 2021;22(12):6257.
8. Yun S.H., Arie T., Kaneko I., Yoder O.C., Turgeon B.G. Molecular organization of mating type loci in heterothallic, homothallic, and asexual Gibberella/Fusarium species. *Fungal Genet Biol.* 2000;31(1):7-20.