APPLICATION OF NEXT-GENERATION SEQUENCING IN THE STUDY OF VARIOUS FORMS OF *BLUMERIA GRAMINIS*

Tomasz Ociepa, Sylwia Okoń

Institute of Plant Genetics, Breeding and Biotechnology, University of Life Sciences, Lublin, Poland

The biotrophic fungus *Blumeria graminis* causing powdery mildew of cereals and grasses has evolved into nine special forms constituting subspecies infecting one species of the host plant. So far, the genomes of 3 special forms have been known or studied: *hordei* , *tritici* and *triticale*. The rest of the special forms have not been sequenced, including f.sp. *avenae*, which attacks oats.

Spores of *Blumeria graminis* on leaves of various cereals

Genetic and evolutionary studies of *Blumeria graminis avenae* have shown that it different from the other subspecies. Sequencing of the known genomes of *Blumeria graminis* showed that this genome is larger than that of other fungal pathogens (Table 1.) and contains 60 to 90% of evenly distributed repeat elements.

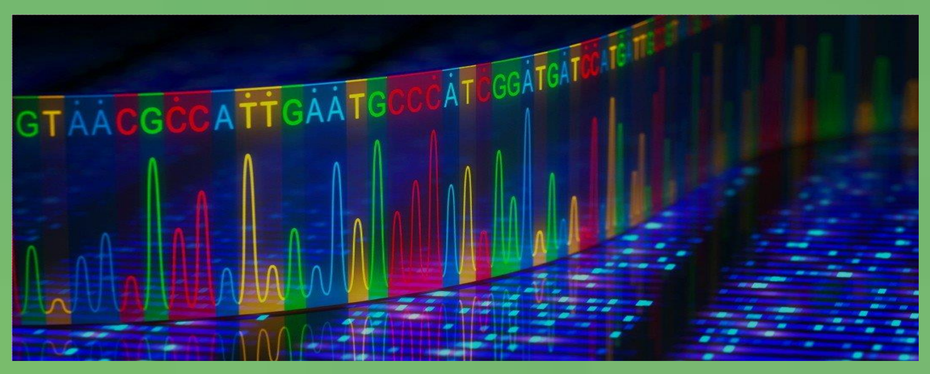

Table 1. Fungal cereal pathogens including genome size from MycoCosms and NCBI.

Genus species	Strain	Genome sizes (Mbp)	Phylum
Blumeria graminis hordei	DH14	124,49	Ascomycota
Blumeria graminis tritici	96224	158,94	Ascomycota
Blumeria graminis triticale	THUN12	141,40	Ascomycota
Blumeria graminis avenae	?	?	Ascomycota
Puccinia graminis	-	88,64	Basidiomycota
Puccinia triticina	1-1 BBBD	135,34	Basidiomycota
Fusarium graminearum	PH-1	36,45	Ascomycota
Fusarium oxysporum	Fox64	52,24	Ascomycota
Mycosphaerella graminicola	-	39,70	Ascomycota

The presence of a large number of repetitive sequences in the genome makes it difficult to read, assemble and annotate. Consequently, first assemblies of the barley and wheat powdery mildew genomes were highly fragmented. Recent advances in long-read sequencing technology (i.e. Pacific Biosciences and Oxford Nanopore Technologies), along with new scaffolding methods, have enabled resolution of chromosome-scale assemblies of an increasing number of plant pathogens genomes.

Examples of next-generation sequencers (left Illumina MiSeq; right PacBio RSII)

Only genomes of high contiguity allow the addressing of topics such as gene space organization and copy number variation (CNV). These analyses are essential to be able to cover the entire diversity of the candidate effector complement of a pathogen.

Examples of bioinformatics programs for NGS data analysis