DOI 10.18699/SBB-2022-346

Whole genome sequencing and development of KASP markers for genotyping the West Siberian wheat stem rust population

Kelbin V.N.*, Muterko A.F., Skolotneva E.S., Sergeeva E.M., Salina E.A.
Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
* *kelbin@bionet.nsc.ru*

Key words: wheat stem rust, *Puccinia graminis*, genomic sequencing, KASP-markers

Motivation and Aim: *Puccinia graminis* f. sp. *tritici* is a biotrophic pathogenic fungus that causes the stem rust disease of bread wheat and leads to significant yield loss. During the last decade the multiple episodes of stem rust infection have been noted in the West Siberia region. The ability of the fungal pathogen to infect wheat plants is determined by the combination of virulence genes (race), at the same time either resistance or susceptibility of the host plant is provided by corresponding resistance genes (*Sr*). During the *P. graminis* life cycle its genome undergoes the recombination events, leading to the different combinations of virulence genes and together with multiple allelism make the pathogen population highly heterogeneous. The estimating of the pathogen race composition and the migration routes of virulent races in West Siberia is of great importance for the breeding of resistant bread wheat varieties. The traditional phytopathology methods are laborious and time-consuming. Therefore, there is a need for diagnostic set of genotyping PCR markers for pathogen population. The aim of the study is the development of allele-specific KASP (Kompetitive Allele Specific PCR) genotyping markers for *P. graminis* f. sp. *tritici* in the West Siberia region. For this reason, the whole genome sequencing and comparative transcriptomic analysis of the West Siberia stem rust lines was carried out for the first time.
Methods and Algorithms: Four lines of wheat stem rust from the Novosibirsk and Omsk regions with different combinations of virulence genes were selected for the study: Wh_18_6 and Wh_18_7 (differing by ability to infect the plants with *Sr6*, *Sr11*, *SrTmp* genes), and Duet_sp_1, Duet_sp_2 (*Sr6*, *Sr30*). The genome sequencing for each line was carried out as 2x150 bp reads on NextSeq550 (Illumina) at ICG SB RAS Genomics Core Facility Novosibirsk, Russia. The representative transcriptome, used for development of KASP markers, included predicted for *P. graminis* coding sequences [1], which are completely and unambiguously mapped with BLAT [2] to the reference genome of this species [3]. The paired reads were trimmed by quality, error corrected with Musket [4] and merged. Obtained fragments were mapped to the representative transcriptome using custom mapper, based on Minimap [5] and SSW library [6], and alignments were filtered. SNP searching was carried out during analysis of the alignments of all samples with the use of SAMtools mpileup [7]. The designed KASP-markers (LGC, Biosearch Technologies, United Kingdom) were tested on stem rust samples from Western Siberia and adjacent regions of the Russian Federation. KASP amplifications and allelic discriminations were performed using QuantStudio 5 (Applied Biosystems) Real-Time PCR machine and the data was analyzed using the QuantStudioTM Design and Analysis Software V2.6.0 (Applied Biosystems). Genotyping reactions were performed following the guideline of “Guide to running KASP genotyping on the ABI 7500 instrument”.

Results: The Illumina reads obtained in this work (for Wh_18_6, Wh_18_7, Duet_sp_1, Duet_sp_2 lines) were assembled in the the 135 contigs with a total length of 62.3 Mbp (*P. graminis* reference genomic sequence is of 85 Mbp). From the 379923 polymorphic sites, total of 16610 SNPs were selected for the development of 1061 KASP markers, whereof the 20 markers approximately evenly distributed across the genome were selected. Since the aim of this work was to produce KASP markers useful to monitor the phytosanitary situation and determine the races of the pathogen fungi, the KASP tests were performed to genotype 120 accessions from Western Siberia and other Russian regions. All studied KASP assays yielded consistent results, confirming the ability of markers to clearly distinguish between the Wh_18_6, Wh_18_7 and Duet_sp_1, Duet_sp_2 lines, indicating to correctly selected SNPs. Many of the KASP assays were diagnostic and polymorphic between samples from different regions of Russia (Fig. 1).

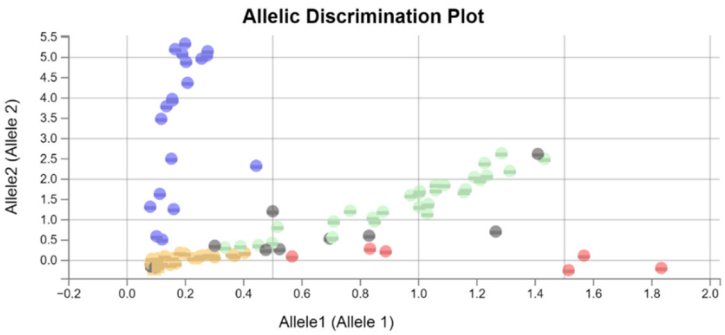

Fig. 1. Allelic Discrimination plot for selected KASP assays showing clustering of genotypes stem rust from regions of the Russian Federation on the Y- and X-axes. Genotypes colored red have a HEX-type allele; genotypes colored blue have a FAM-type allele; green-colored have both types of alleles (heterozygotes); black and orange dots represent non-template control

Conclusion: The genome sequencing of the stem rust lines from the West Siberia, having contrasting combinations of virulence genes, was carried out for the first time. By comparison of genomic data obtained from our samples with the *P. graminis* reference transcriptome the 20 KASP-markers were selected. The results obtained showed good resolution of developed KASP markers to characterize the pathogen population in the Western Siberia and other regions of Russia. The entire panel of markers is successful for diagnostic tasks and can be used in the phytosanitary monitoring and to determine the races of the pathogen fungi.

Acknowledgements: The study is supported by the Budget Project FWNR-2022-0007.

References

1. Upadhyaya N.M. et al. Comparative genomics of Australian isolates of the wheat stem rust pathogen *Puccinia graminis* f. sp. *tritici* reveals extensive polymorphism in candidate effector genes. *Front Plant Sci*. 2015;5:759.
2. Kent W.J. BLAT--the BLAST-like alignment tool. *Genome Res.* 2002;12(4):656-664.
3. Duplessis S. et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *PNAS.* 2011;108(22):9166-9171.
4. Liu Y., Schröder J., Schmidt B. Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. *Bioinformatics.* 2013;29(3):308-315.
5. Li H. Minimap and miniasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics.* 2016;32(14):2103-2110.
6. Zhao M., Lee W.-P., Garrison E.P., Marth G.T. SSW Library: An SIMD Smith-Waterman C/C++ Library for Use in Genomic Applications. *PLoS ONE.* 2013;8(12):e82138.
7. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011;27(21):2987-2993.