Proceedings of the LXIV SIGA Annual Congress
Online, 14/16 September, 2021
ISBN: **978-88-944843-2-8**

Poster Communication Abstract – 1.13

CAS9-MEDIATED ENRICHMENT COUPLED TO NANOPORE SEQUENCING PROVIDES A VALUABLE TOOL FOR DE-NOVO ASSEMBLY OF CULTIVAR-SPECIFIC GENOMIC REGIONS

ALFANO M.*, MAESTRI S.*, PAPA R.***, DI VITTORI V.**, MARCOLUNGO L.*, MARTURANO G.*, COSENTINO E.*, LAVEZZARI D.*, BELLUCCI E.**, NANNI L.**, BITOCCHI E.**, DELLEDONNE M.*, ROSSATO M.*

*) Department of Biotechnology, University of Verona, Strada le Grazie 15, 37134 Verona, Italy
**) Department of Scienze Agrarie, Alimentari e Ambientali, via Brecce Bianche, 60131 Ancona, Italy
***) Genartis srl, via IV Novembre 24, 37126 Verona, Italy

cultivar-specific genome, de-novo assembly, CRISPR-Cas9-based enrichment, Nanopore sequencing

The availability of well-assembled genomes is critical for accurate variant calling and the identification of candidate genes/variants responsible of a certain phenotype. However, one single reference genome may be not sufficient, especially in plants, where different cultivars can vary consistently not only at single-nucleotide level (SNV), but also at the structural level (SV), as recently demonstrated with the tomato pan-genome. In turn, de-novo assembly of cultivar-specific reference genomes can be very costly.

Long-DNA fragment capture in combination with long-read sequencing may provide a cost-efficient approach to accurately reconstruct genomic regions of interest for in depth analysis. As proof-of-concept, we have applied the Cas9-mediated enrichment coupled to nanopore sequencing to reconstruct a 250 Kb region on chromosome 5 of *P.vulgaris* genome. The region presented a large amount of SNV and SV in the cultivar Midas, as compared to the reference genome. Five tiled sub-regions of 50Kb each were cut with high efficiency (>70%) from Midas genomic DNA, using target-specific guide RNAs designed on conserved coding regions. Sequencing on a MinION device yielded good amount of data (˜150X coverage and ˜130-fold enrichment, on average)

that were assembled de-novo, generating a single contig spanning the whole 250Kb target region. The target region captured with Cas9 was properly reconstructed and shared a 99.5% identity with the one assembled using a traditional approach based on whole-genome-sequencing (nanopore data, 50X average coverage). Finally, Illumina data derived from Midas-inbred lines showed a consistently improved mapping quality on the *de-novo* assembled locus, as compared to the *P.vulgaris* reference genome.

In conclusion, the Cas9-mediated target enrichment “tiling” approach represents a valuable alternative to whole genome sequencing to assemble ultra-long target regions, with consistent cost-saving. In the future, this approach can allow the fine characterization of cultivar-specific regions of interest, especially in plants with very large genomes where whole genome *de-novo* assembly is little affordable.