Deep Sequencing of a Local Grapevine Variety in Turkey for Diagnosis of Viruses and Viroids via Next Generation Sequencing Techniques

Serkan Önder[1], Z. Neslihan Öztürk Gökçe[2], Çiğdem Ulubaş Serçe[3]*

[1] Manisa Viticulture Research Institute, Plant Health Department, 45125 Yunusemre, Manisa.

[2] Department of Agricultural Genetic Engineering, Ayhan Şahenk Faculty of Agricultural Sciences and Technologies, Niğde Ömer Halisdemir University, 51240, Niğde, Turkey.

[3] Department of Plant Production and Technologies, Ayhan Şahenk Faculty of Agricultural Sciences and Technologies, Niğde Ömer Halisdemir University, 51240, Niğde, Turkey.

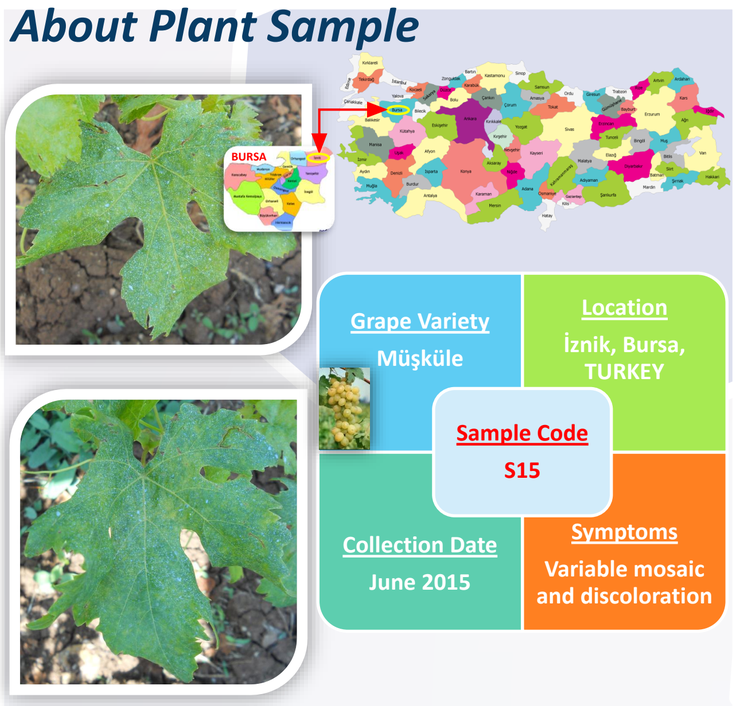
*Corresponding author e-mail: culubas@gmail.com

Abstract

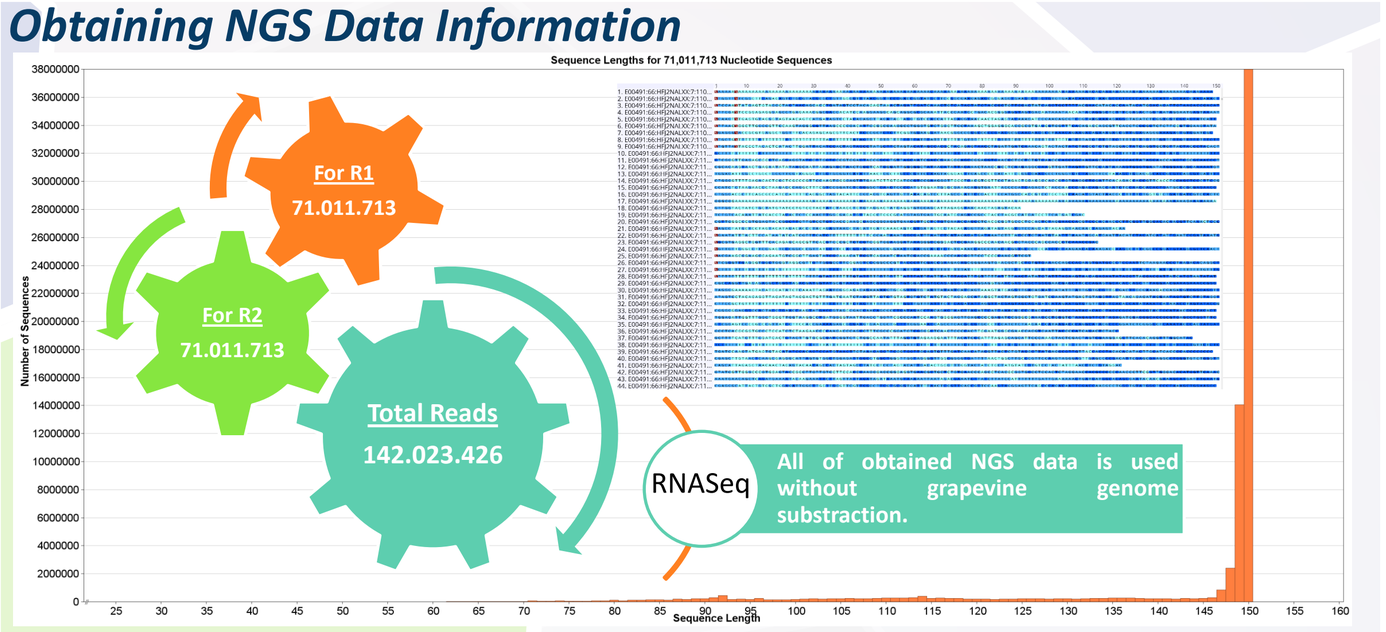
Next generation sequencing (NGS) technologies also known as high-throughput sequencing, provides a platform for the detection of all viral agents in a sample with simultaneous sequencing of millions of nucleic acids. Samples from a local table grape variety, Müşküle, exhibiting variable mosaic symptoms and discoloration were collected from Bursa province in Turkey for NGS analysis. Total RNA were extracted and rRNA depletion was performed by treatment with Ribo-Zero rRNA Plant Removal Kit. NEBNext® Ultra™ RNA Library Prep Kit was used for library preparation. Deep sequencing was performed using Illumina Hiseq2000 RNAseq technology with 2x150 read length and 40 million depths for each read. At the end of the deep sequencing analysis, around 142 million sequences were derived. Bioinformatic analysis was performed using Geneious R11 software. *De novo* assembly was performed with Tadpole assembler and 29.051 contigs were obtained. All derived contigs were analyzed by blastn against NCBI viral RefSeq database to detect known viruses and viroids. After the blastn analysis, mapping was performed with all reads against reference sequences. As a result of genome mapping analysis, *Grapevine deformation virus* (GDefV) RNA 2 segment, *Grapevine leafroll-associated virus-2* (GLRaV-2) and *Grapevine fanleaf virus* (GFLV) RNA 2 segment genomes were recovered 32.4%, 18.9% and 6.3% respectively; whereas GDefV RNA 1 segment, GFLV RNA 1 segment, *Grapevine virus A* (GVA), *Grapevine roditis leaf discoloration-associated virus* (GRLDaV), *Grapevine leafroll-associated virus-1* (GLRaV-1) and *Arabis mosaic virus* (ArMV) RNA 2 segment were not assembled.

The preliminary mapping results showed that obtaining full genome of viral agents detected by NGS is highly dependent on the virus and part of the viral segment obtained by sequencing. In addition, further studies should be performed for reliable detection of unknown viral agents.

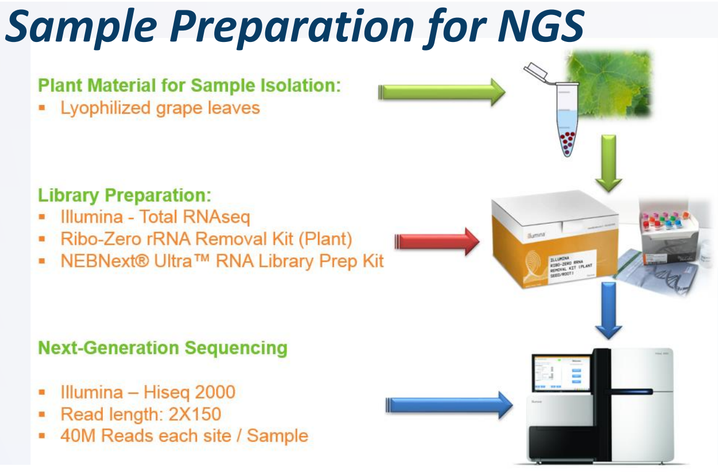
About Plant Sample



Obtaining NGS Data Information



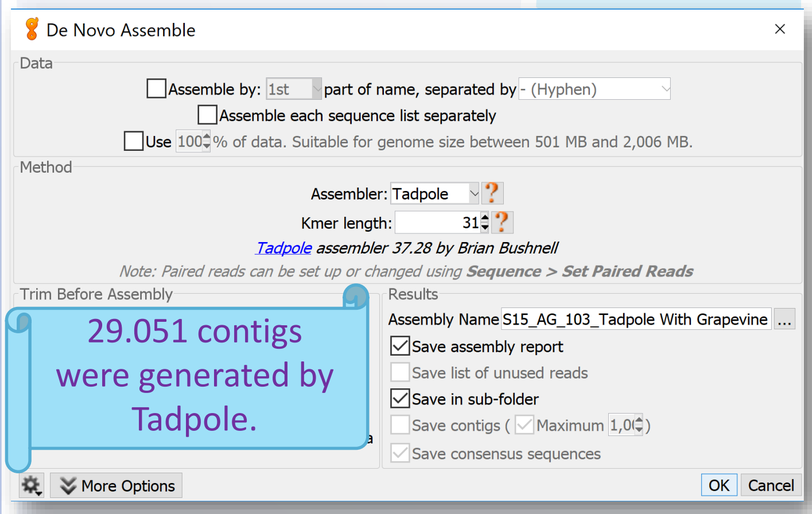
Sample Preparation for NGS



NCBI blastn Results Against Tadpole Contigs

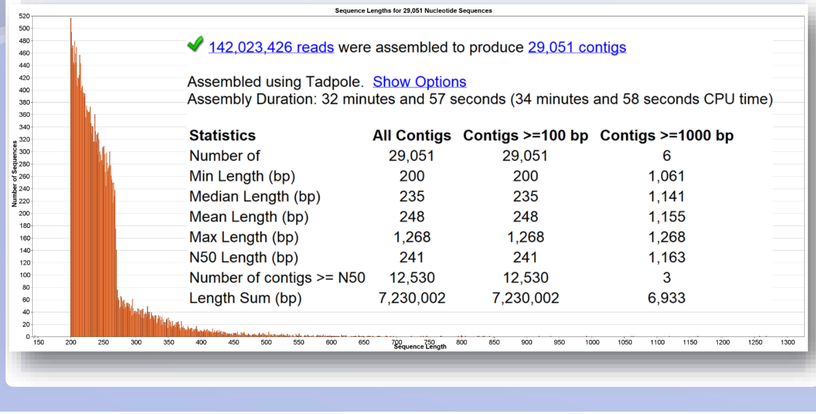
Viruses	RefSeq Accession No	Range of the Matched Sequence Length (bp)	% Pairwise Identity Range of Sequences
GDefV RNA 2	NC_017938	29 to 823	82.1 to 94.6
GLRaV-2	NC_007448	227 to 316	87 to 99.6
GFLV RNA 2	NC_003623	208 to 623	86.5 to 87.0
GDefV RNA 1	NC_017939	222-673	80 to 90.4
GFLV RNA 1	NC_003615	206-489	70.6 to 90.6
GVA	NC_003604	24-343	81.2 to 91.7
GRLDaV	NC_027131	23-264	85.2 to 95.7
GLRaV-1	NC_016509	51-226	72.6 to 86.3
ArMV RNA 2	NC_006056	352	85.8

De novo Assembly

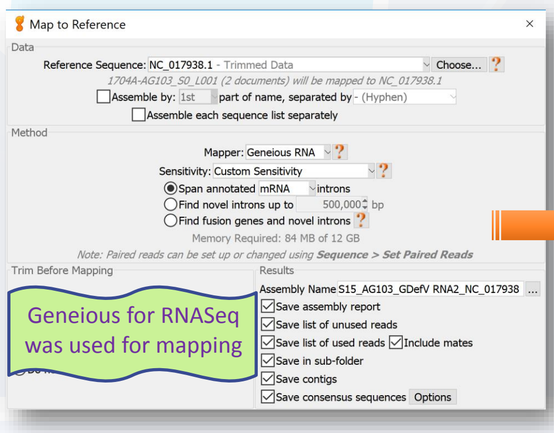


29.051 contigs were generated by Tadpole.

- Tadpole • Extremely fast and optimized
- Tadpole • Needs less memory
- Tadpole • Low misassembly rate

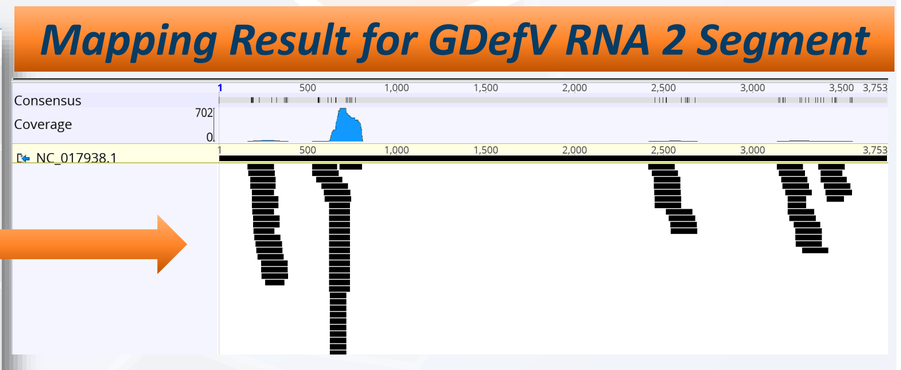


Map to Reference



Geneious for RNASeq was used for mapping

Mapping Result for GDefV RNA 2 Segment



Viruses	RefSeq Length (bp)	Coverage of RefSeq (bp)	% Coverage of RefSeq
GDefV RNA 2	3753	1216	32.4
GLRaV-2	16494	3113	18.9
GFLV RNA 2	3774	236	6.3
GDefV RNA 1	7386	Not assembled	
GFLV RNA 1	7342	Not assembled	
GVA	7351	Not assembled	
GRLDaV	6988	Not assembled	
GLRaV-1	18659	Not assembled	
ArMV RNA 2	3820	Not assembled	

Acknowledgement

This study supported by the project TUBITAK-115 O 014 and COST-DIVAS of COST FA-1407 Action.