WAGENINGEN UR
For quality of life

Full genome characterisation of 10 tospoviruses by next generation sequencing

Annette M. Dullemans (1), Petra J. van Bekkum (1), Annelien (J.W.) Roenhorst (2), Richard Kormelink (1) and René A.A. van der Vlugt (1)

Background and objectives

The genus *Tospovirus* currently contains 11 officially recognized species (ICTV, 2015) however, more tentative species have been described. For most tospoviruses no specific antisera are available and only partial sequence data have been described. This hampers reliable detection of these agronomical important viruses. To facilitate the development of diagnostics and ensure the future availability of reference isolates the full RNA genomes of 10 tospovirus isolates were determined, their sequences and biological data included in Q-bank and the individual isolates included in the physical plant virus collection at Wageningen UR.

Methods

Isolates of the different tospoviruses were obtained from several plant virus collections and inoculated on indicator plants. Total plant RNA extracts were used for Illumina RiboZero 125 base paired-end library preparations with individual MIDs and subsequently run in batch on a Illumina HiSeq 2500. After MID splitting, individual datasets were fed into custom designed workflows within CLC Genomics Workbench (Qiagen, Denmark). These workflows comprised 'De novo' and reference-assemblies with and without substraction of plant-related reads. Resulting contigs were analysed by BlastN and BlastX against the NCBI database to identify tospovirus related sequences. RT-PCR, 5'- and 3'-RACE and Sanger sequencing was used to determine complete tospovirus genomes.

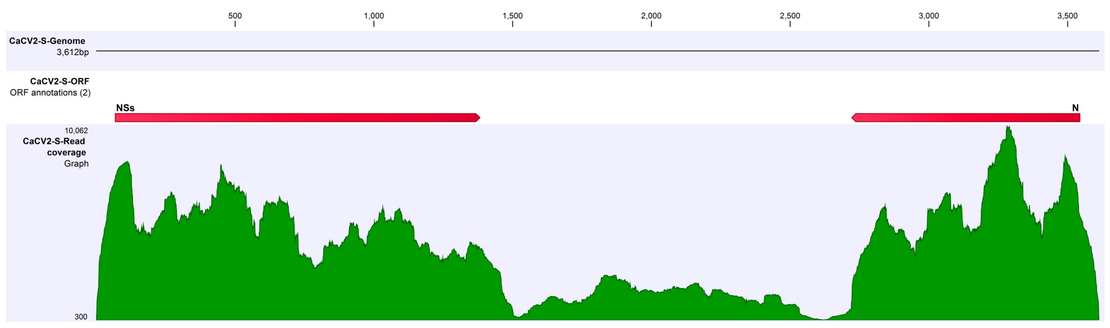

Figure 1. An example of remapping of reads on the assembled CaCV-PD S-segment.

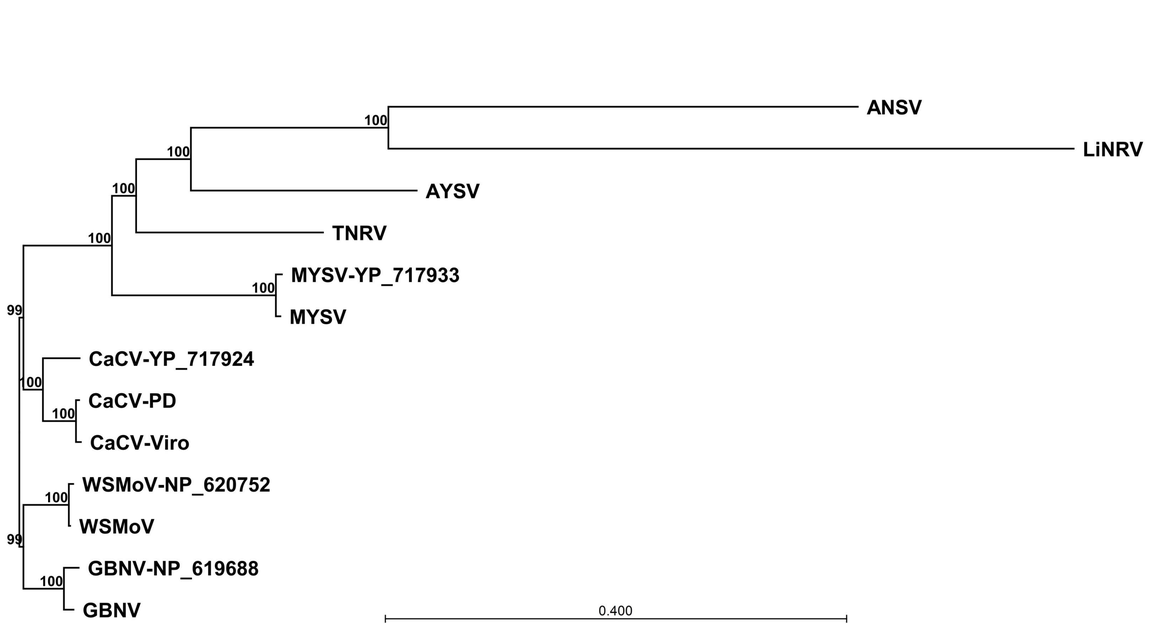

Figure 2. Phylogenetic tree of the L proteins of the sequenced tospoviruses and the available reference isolate sequences from NCBI.

Results

Table 1. Obtained lengths (nts) for the sequenced virus segments. The two AYSV isolates sequenced were identical. Sequences which weren't available at NCBI are marked in red.

Virus isolate	Acronym	L	M	S
Alstroemeria necrotic streak virus	ANSV	8967	4871	3137
Alstroemeria yellow spot virus	AYSV	8865	4797	2734
Capsicum chlorosis virus-Viro	CaCV-Viro	8916	4826	3611
Capsicum chlorosis virus-PD	CaCV-PD	8913	4822	3612
Ground bud necrosis virus	GBNV	8914	4813	3054
Lisianthus necrotic ringspot virus	LiNRV	8958	4698	2747
Melon yellow spot virus	MYSV	8918	4904	3230
Tomato necrotic ringspot virus	TNRV	8858	4725	3017
Watermelon silver mottle virus	WSMoV	8917	4881	3533

Conclusions

- NGS proofed to be a quick and relatively easy method to determine nearly the complete sequences of a significant number of tospoviruses simultaneously.
- Sequence reads were not equally spread over the RNA segments; IGR's were less covered or missing (Fig. 1).
- All genome sequences were finished by 5'- and 3'-RACE, RT-PCR and Sanger sequencing (Table 1).
- Obtained consensus sequences were compared with available sequences from NCBI (Fig. 2 and 3).
- The full length ANSV and AYSV genomes were sequenced for the first time.
- For LiNRV and TNRV the missing genome segments were completed.

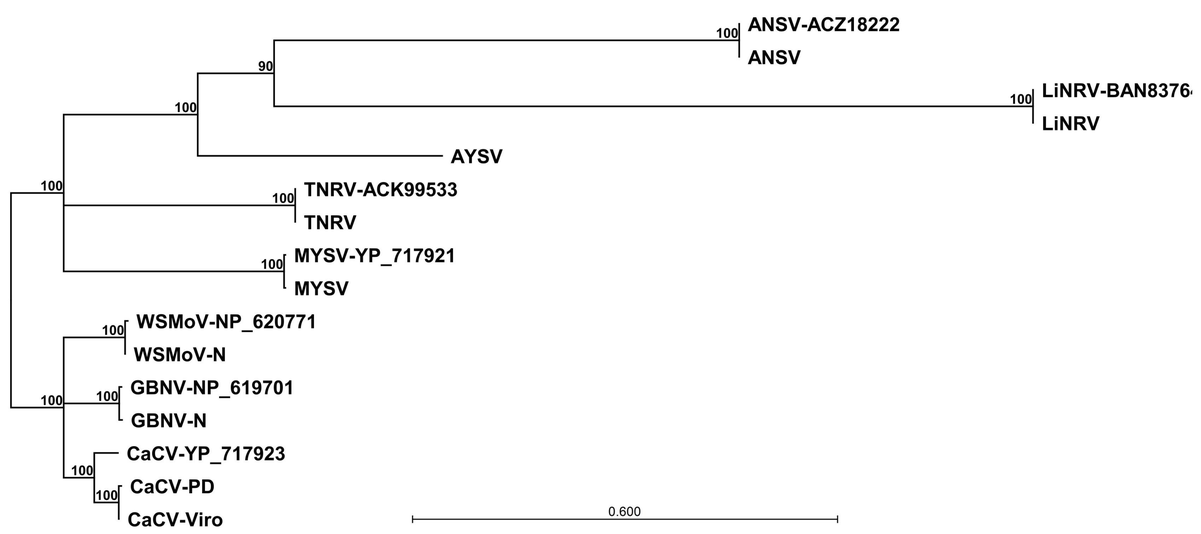

Figure 3. Phylogenetic tree of the N proteins of the sequenced tospoviruses and the available reference isolate sequences from NCBI.

Acknowledgements

This research was funded by the Dutch programme PPS Fundament Fytosanitair

1. Wageningen University and Research, P.O. Box 16, 6700 AB Wageningen, The Netherlands
Contact: rene.vandervlugt@wur.nl, T + 31 (0)317 48 06 75
www.WageningenUR.nl

2. Dutch Plant Protection Organisation (NPPO), Wageningen, The Netherlands