



*Netherlands Institute for Vectors,
Invasive plants and Plant health*

Netherlands Food and Consumer
Product Safety Authority
*Ministry of Agriculture,
Nature and Food Quality*



Validation of a high throughput sequencing test within an ISO17025 accredited plant health laboratory

How high throughput sequencing can be used as a generic tool to identify viruses
in a regulatory framework

Marcel Westenberg





Why using HTS in a regulatory framework

Official Controls Regulation (EU) 2017/625

Laboratories designated by the competent authorities to carry out analyses, tests and diagnoses on samples taken in the context of official controls and other official activities should possess the expertise, equipment, infrastructure and staff to carry out such tasks to the highest standards. To ensure sound and reliable results, those laboratories should be accredited for the use of these methods according to standard EN ISO/IEC 17025 on 'General requirements for the competence of testing and calibration laboratories'. The accreditation should be delivered by a national accreditation body operating in accordance with Regulation (EC) No 765/2008 of the European Parliament and of the Council ⁽¹⁾.

Commission implementing regulation (EU) 2021/2285

'ANNEX II

List of Union quarantine pests and their respective codes assigned by EPPO

TABLE OF CONTENTS

Part A : Pests not known to occur in the Union territory

Part B: Pests known to occur in the Union territory



57	virus species
1	begomovirus genus (>400 species)
2	viroids
40	phytoplasmas strains

One single HTS test, instead of implementing and validating specific tests (ELISAs, PCRs) for each regulated pest



HTS at NIVIP

- > Since 2014, HTS (Illumina sequencing) is used in applied research
- > Since 2015, virus detection and identification in diagnostic samples
- > Pipeline optimization due to PT/TPS organized in frame of international research projects
- > Since 2019 from research applications to standardized diagnostics

COST Action: FA1407 - Application of next generation sequencing for the study and diagnosis of plant viral diseases in agriculture

Phytopathology • 2019 • 109:488-497 • <https://doi.org/10.1094/PHYTO-02-18-0067-R>

Virology

e-Xtra*

Virus Detection by High-Throughput Sequencing of Small RNAs: Large-Scale Performance Testing of Sequence Analysis Strategies

Sebastien Massart,[†] Michela Chiumenti, Kris De Jonghe, Rachel Glover, Annelies Haegeman, Igor Koloniuk, Petr Kominek, Jan Kreuze, Denis Kutnjak, Leonidas Lotos, François Maclot, Varvara Maliogka, Hans J. Maree, Thibaut Olivier, Antonio Olmos, Mikhail M. Pooggin, Jean-Sébastien Reynard, Ana B. Ruiz-García, Dana Safarova, Pierre H. H. Schneeberger, Noa Sela, Sílvia Turco, Eeva J. Vainio, Eva Varallyay, Eric Verdin, Marcel Westenberg, Yves Brostaux, and Thierry Candresse

EUPHRESKO 2015-F-172: The application of next-generation sequencing technology for the detection and diagnosis of non-cultural organisms: Viruses and viroids (NGS-detect)

 pathogens



Article

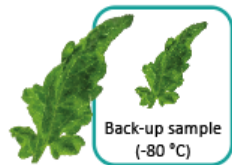
Interlaboratory Comparison Study on Ribodepleted Total RNA High-Throughput Sequencing for Plant Virus Diagnostics and Bioinformatic Competence

Yahya Z. A. Gaafar¹, Marcel Westenberg², Marleen Botermans², Krizbai László³, Kris De Jonghe⁴, Yoika Foucart⁴, Luca Ferretti⁵, Denis Kutnjak⁶, Anja Pecman^{6,7}, Nataša Mehle⁶, Jan Kreuze⁸, Giovanna Muller⁸, Nikolaos Vakirlis⁹, Despoina Beris⁹, Christina Varveri⁹ and Heiko Ziebell^{1,*}



Defining the tests (2)

- > A test is a specific application of a method (EPPO PM7/76)
- > Includes everything from nucleic acid extraction to data interpretation



Sampling and packaging

1



RNA extraction

2

Outsourced at ISO17025 accredited sequence provider



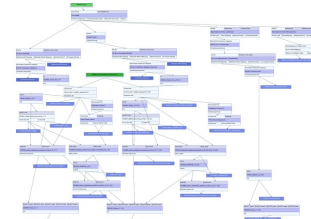
Library preparation

3



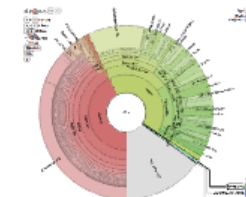
Illumina Sequencing

4



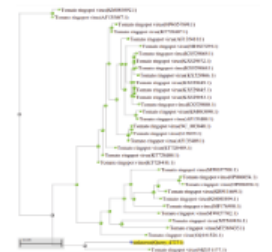
Bioinformatic pipeline

5



Selection and analysis of putative viral contigs

6



Sequence analysis and reporting

7

versie	datum	beschrijving
4	12-04-2021	Werkstructuur is opgesteld in twee onderdelen (A en B) met een paar verspreiden in vier delen. Onderdeel generatie Sequencing is nieuw toegevoegd
3	19-12-2017	Aanpassen by-out (onderdeel 7) en 8 binnen Moleculaire Biologie.
2	11-03-2016	Hopelijkheids toevoeging om aan te gevang worden ingedeeld.

Monstertromen Moleculaire Biologie

- Onderwerp**
Het beschrijven van de monstertromen naar en binnen (MIB) van het Nationaal Referentiecentrum (NRC).
- Principe**
Het labgeleid MIB is voor moleculaire biologie niet geschikt voor plantentesten en plagenprofielen, monstertromen van de organismische vijandige (Cephus-moeders) veldvossen, Entomologie, Dierenwet en Virologie. Om het uitvoeren van deze verschillende verschillende vakgebieden in goede banen te leiden voor het aanpakken van analyses, het is behalve de nu van vakgebieden en de rapportage van resultaten, verschillende stappen binnen moleculaire biologie met betrekken medewerkers.

RNA extractie met behulp van de RNeasy

- Onderwerp**
Het oorspronken van RNA uit plantemateriaal met behulp (Qiagen) om eenzijdig virus RNA te isoleren.
- Principe**
De RNeasy Mini Kit is ontworpen voor de analyse (genooms, proteom, enz.) vanuit plantemateriaal, welke geschikt voor (post-trans) (RT-PCR, Southern) analyses. Om het uitvoeren van deze verschillende verschillende vakgebieden in goede banen te leiden voor het aanpakken van analyses, het is behalve de nu van vakgebieden en de rapportage van resultaten, verschillende stappen binnen moleculaire biologie met betrekken medewerkers.

Monstertromen	Monstertromen	Monstertromen
1	2	3
4	5	6
7	8	9
10	11	12
13	14	15
16	17	18
19	20	21
22	23	24
25	26	27
28	29	30
31	32	33
34	35	36
37	38	39
40	41	42
43	44	45
46	47	48
49	50	51
52	53	54
55	56	57
58	59	60
61	62	63
64	65	66
67	68	69
70	71	72
73	74	75
76	77	78
79	80	81
82	83	84
85	86	87
88	89	90
91	92	93
94	95	96
97	98	99
100	101	102

versie	datum	beschrijving
4	15-10-2022	- RNC Dikwyl/VIR toegevoegd - RNC Hlwaag toegevoegd in RNC Hlwaag - Inhoudelijke naar CLC veranderd omringel
3	19-07-2021	RNC veranderd in RNC
2	07-03-2021	- Nucleotiden worden in plaats van - Benadrukt worden toegevoegd - Opname worden voor digitale data en - Volume worden 20 op 25 af - Opname van monstertromen aang - Periode 3.0 Inladen naar data in CLC en 1 - Files in CLC zijn veranderd en terug te

Uitbestellen Next Generation Sequencing

- Onderwerp**
Het uitvoeren van monstertromen met Next Generation Sequencing (NGS). Deze instructie beschrijft hoe de gegevens worden opgeslagen en geïmporteerd in CLC met behulp van de volgende stappen.
- Principe**
Er wordt gebruik gemaakt van Illumina HiSeq 2500 (HiSeq) met behulp van de volgende stappen.
- Begrippen**
Niet van toepassing.
- Benodigdheden**
4.1 Apparatuur
• Azure Labcloud werkpost met CLC Genomic Server
- Helpmiddelen**
• CLC Genomic Workbench software
- Concepten**
• Niet van toepassing
- Oplossingen en media**
Niet van toepassing

Werkten met Azure cloud

- Onderwerp**
In deze instructie wordt beschreven hoe te werken met de Azure Labcloud werkpost.
- Principe**
De vakgroep Moleculaire Biologie voert verschillende moleculaire testen uit. Sommige van deze testen maken gebruik van het laboratorium in uitgerust, nog geïmporteerd worden middelen specifieke software. Deze software is niet beschikbaar onder de CLC licentie, maar kan worden geïmporteerd in de Azure cloud omgeving. De applicatie is geïmporteerd met het gebruik van de Labcloud om informatie te verschaffen rond de implementatie.
- Begrippen**
Niet van toepassing.
- Benodigdheden**
4.1 Apparatuur
• Azure Labcloud werkpost met CLC Genomic Server
- Helpmiddelen**
• CLC Genomic Workbench software
- Concepten**
• Niet van toepassing
- Oplossingen en media**
Niet van toepassing

versie	datum	beschrijving
4	15-10-2022	- RNC Dikwyl/VIR toegevoegd - RNC Hlwaag toegevoegd in RNC Hlwaag - Inhoudelijke naar CLC veranderd omringel
3	19-07-2021	RNC veranderd in RNC
2	07-03-2021	- Nucleotiden worden in plaats van - Benadrukt worden toegevoegd - Opname worden voor digitale data en - Volume worden 20 op 25 af - Opname van monstertromen aang - Periode 3.0 Inladen naar data in CLC en 1 - Files in CLC zijn veranderd en terug te

Detection and identification of viruses, viroids and phytoplasmas by High Throughput Sequencing

- Type**
In this instruction the various steps for the detection and identification of viruses, viroids and phytoplasmas by high throughput sequencing (HTS) are described.
- Principe**
This instruction is divided in two parts. The standard HTS workflow is described and the detection and identification of viruses, viroids and phytoplasmas is described. In some cases the various steps can be used for or instead of the other assembly workflow (part 2). In particular cases are described in paragraph 4.

versie	datum	beschrijving
4	15-10-2022	- RNC Dikwyl/VIR toegevoegd - RNC Hlwaag toegevoegd in RNC Hlwaag - Inhoudelijke naar CLC veranderd omringel
3	19-07-2021	RNC veranderd in RNC
2	07-03-2021	- Nucleotiden worden in plaats van - Benadrukt worden toegevoegd - Opname worden voor digitale data en - Volume worden 20 op 25 af - Opname van monstertromen aang - Periode 3.0 Inladen naar data in CLC en 1 - Files in CLC zijn veranderd en terug te

Detailed sequence data analysis for virological samples

- Type**
This instruction is divided in two parts. The standard HTS workflow is described and the detection and identification of viruses, viroids and phytoplasmas is described. In some cases the various steps can be used for or instead of the other assembly workflow (part 2). In particular cases are described in paragraph 4.
- Principe**
This instruction is divided in two parts. The standard HTS workflow is described and the detection and identification of viruses, viroids and phytoplasmas is described. In some cases the various steps can be used for or instead of the other assembly workflow (part 2). In particular cases are described in paragraph 4.

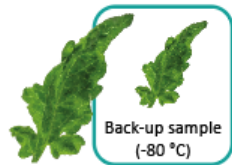
versie	datum	beschrijving
4	15-10-2022	- RNC Dikwyl/VIR toegevoegd - RNC Hlwaag toegevoegd in RNC Hlwaag - Inhoudelijke naar CLC veranderd omringel
3	19-07-2021	RNC veranderd in RNC
2	07-03-2021	- Nucleotiden worden in plaats van - Benadrukt worden toegevoegd - Opname worden voor digitale data en - Volume worden 20 op 25 af - Opname van monstertromen aang - Periode 3.0 Inladen naar data in CLC en 1 - Files in CLC zijn veranderd en terug te

Detailed sequence data analysis for virological samples

- Type**
This instruction is divided in two parts. The standard HTS workflow is described and the detection and identification of viruses, viroids and phytoplasmas is described. In some cases the various steps can be used for or instead of the other assembly workflow (part 2). In particular cases are described in paragraph 4.
- Principe**
This instruction is divided in two parts. The standard HTS workflow is described and the detection and identification of viruses, viroids and phytoplasmas is described. In some cases the various steps can be used for or instead of the other assembly workflow (part 2). In particular cases are described in paragraph 4.



One control to serve many purposes



Sampling and packaging

1



RNA extraction

2

Outsourced at ISO17025 accredited sequence provider



Library preparation

3



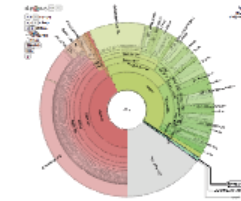
Illumina Sequencing

4



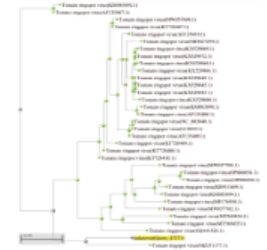
Bioinformatic pipeline

5



Selection and analysis of putative viral contigs

6



Sequence analysis and reporting

7



Healthy *N. benthamiana*



20 ng/ μ l RNA Spiked (99:1) with ERCC RNA Spike-In Mix (ThermoFischer)

Positive Process Control batch

- Should pass entry QC and library QC

- Should produce 2Gb of data at >80% PHRED30

- Should complete the bioinformatic pipeline with all output (check on bioinformatic pipeline)
- Identification of plant species based on *rbcl* sequence (only *Nicotiana* should be identified, serves as positive control and contamination control)
- Recovery of 4 selected ERCC constructs (MRD>10) from positive process control sample (positive control)
- Limited reads (<10) of most abundant ERCC should be identified in diagnostic samples in the same batch (contamination control)
- Assessment of rRNA depletion success (<10% rRNA reads or >1.8 Gb non-rRNA reads)



Validating tests

What to validate or verify?

- › Well described analytical standard operating procedures (A-SOP)

Why should we validate or verify?

- › To assess if a test (A-SOP) is fit for purpose (scope)

How to validate or verify?

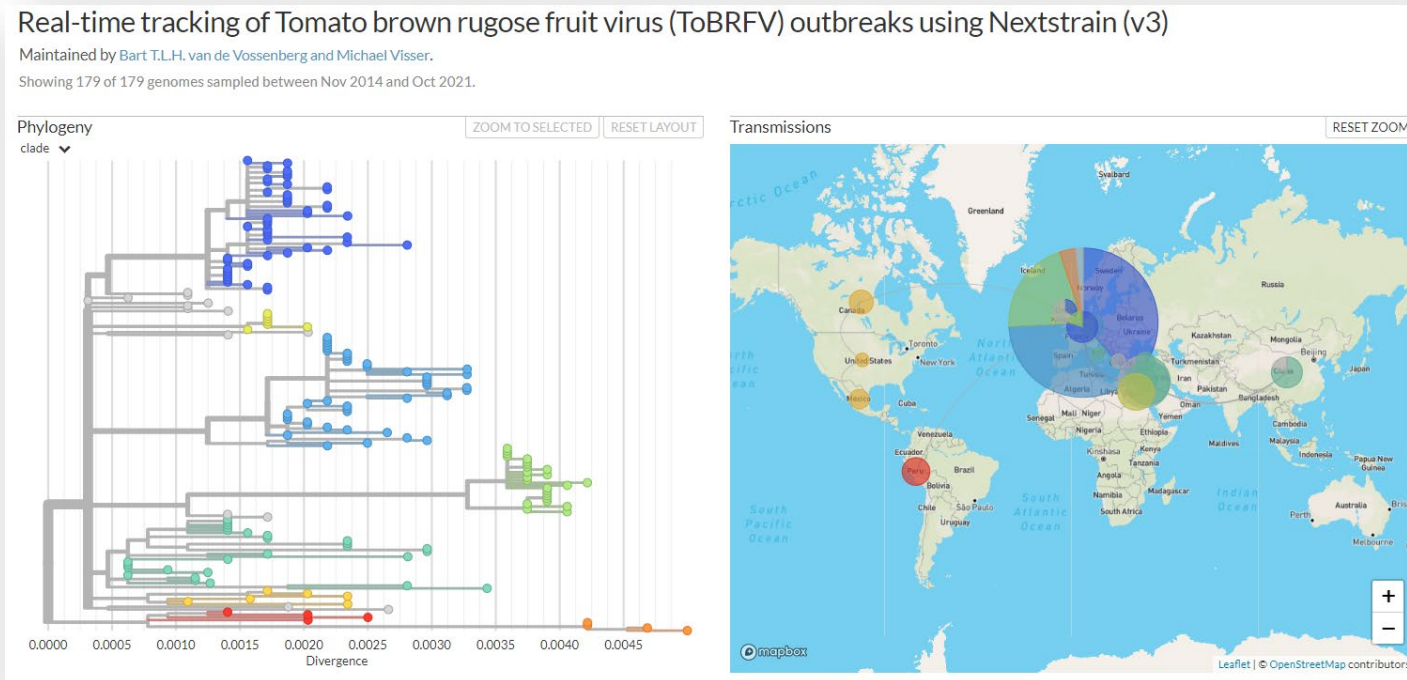
- › Determination or verification of performance criteria (EPPO PM 7/98)
- › Analytical sensitivity, Analytical specificity, Selectivity, Repeatability, Reproducibility
- › Well characterized (biological) material





Validation scope

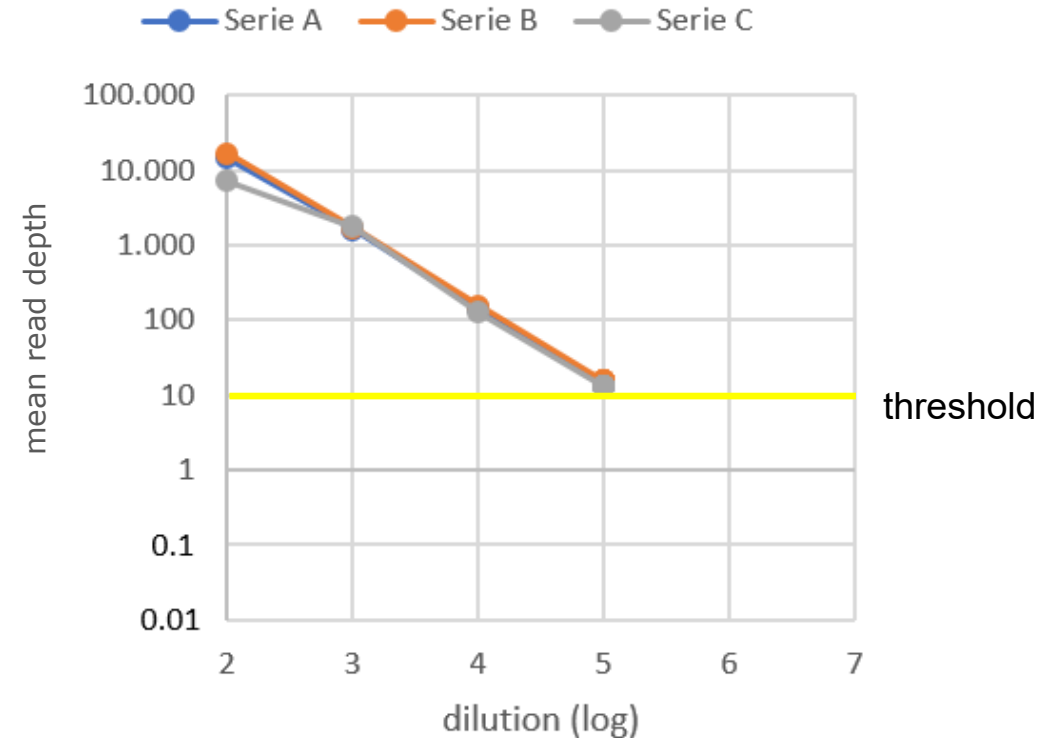
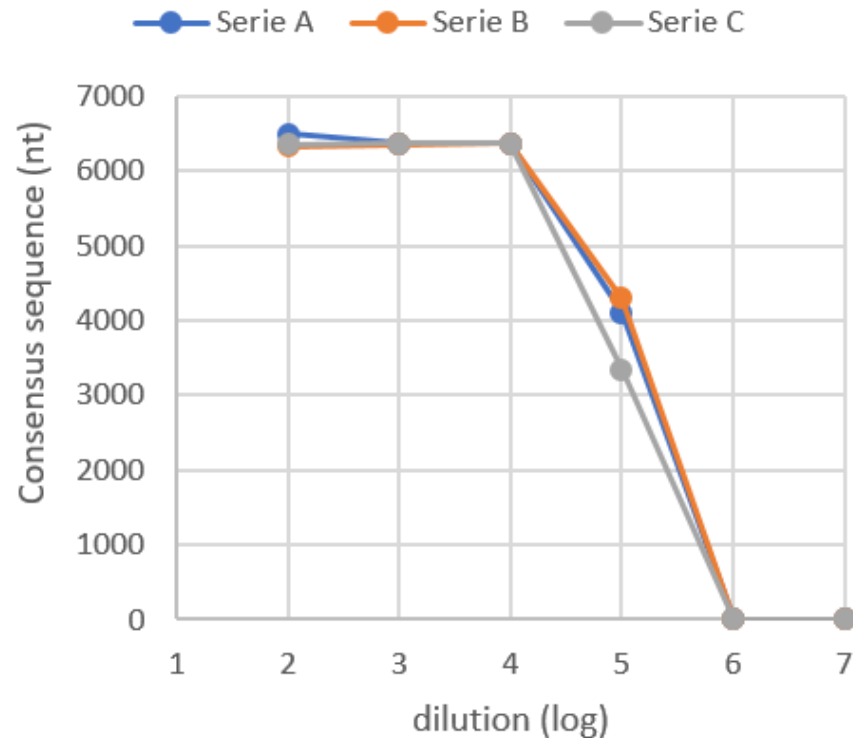
- › Detection and identification of ToBRFV in tomato leaf material by HTS (Illumina sequencing)





Analytical sensitivity

- > Full length sequences were obtained till a dilution of 10^4
- > Species identification still possible with partial sequences at dilution 10^5
- > Analytical sensitivity: relative dilution rate of 10^5
- > Correlation between average read coverage and dilution.





Analytical specificity

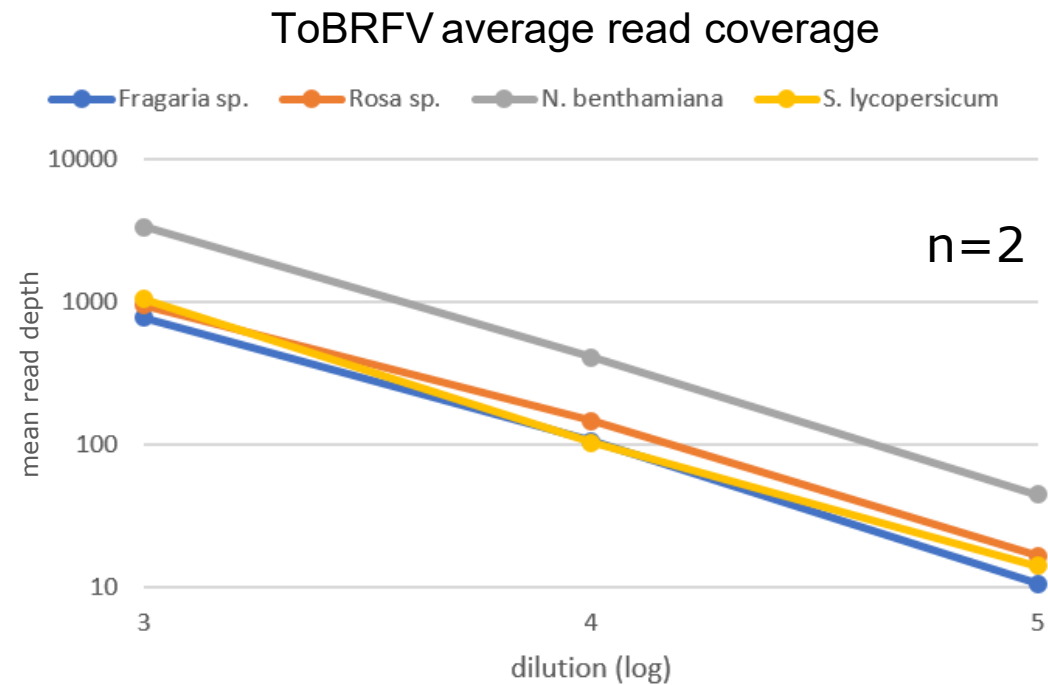
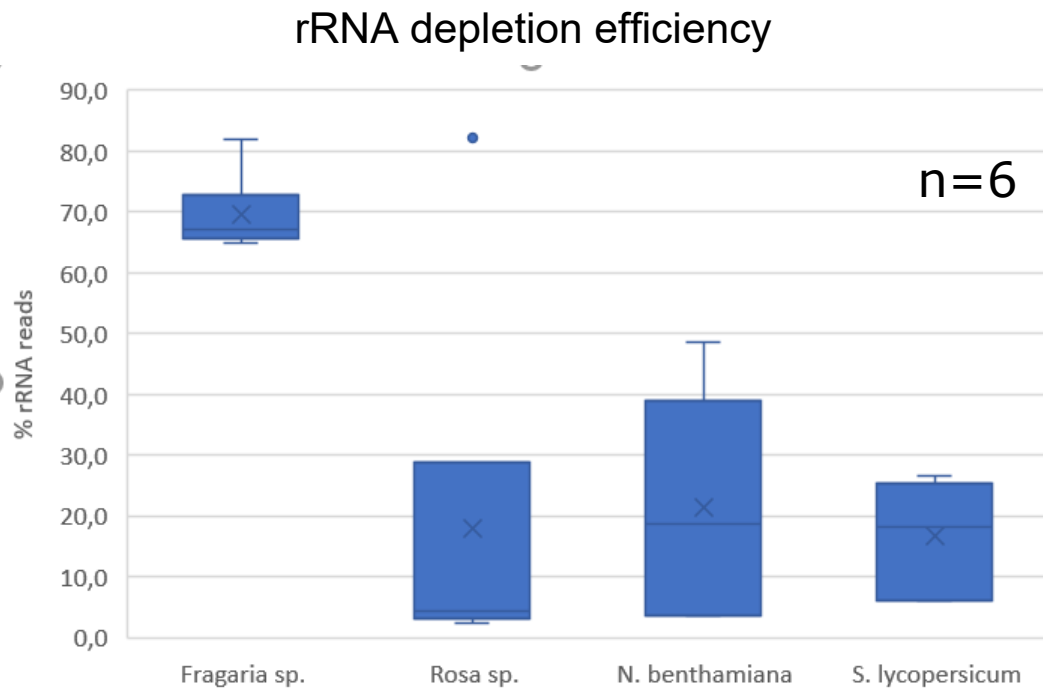
- › Between 2016-2022 detected and identified over 180 viruses/viroids
- › Belonging to 47 genera, 25 families
- › All known viral nucleic acid types:
ssRNA(+), ssRNA(-), dsRNA, CssRNA, dsDNA(-RT), ssDNA
- › In a variety of hosts, and plant parts (e.g. leaf, fruit, seed)





Selectivity

- > Affect of difficult matrices (*Fragaria* sp., *Rosa* sp.)
- > Dilution of ToBRFV-infected plantsap in plantsap of different matrices



matrix has no effect on the analytical sensitivity



Repeatability and reproducibility

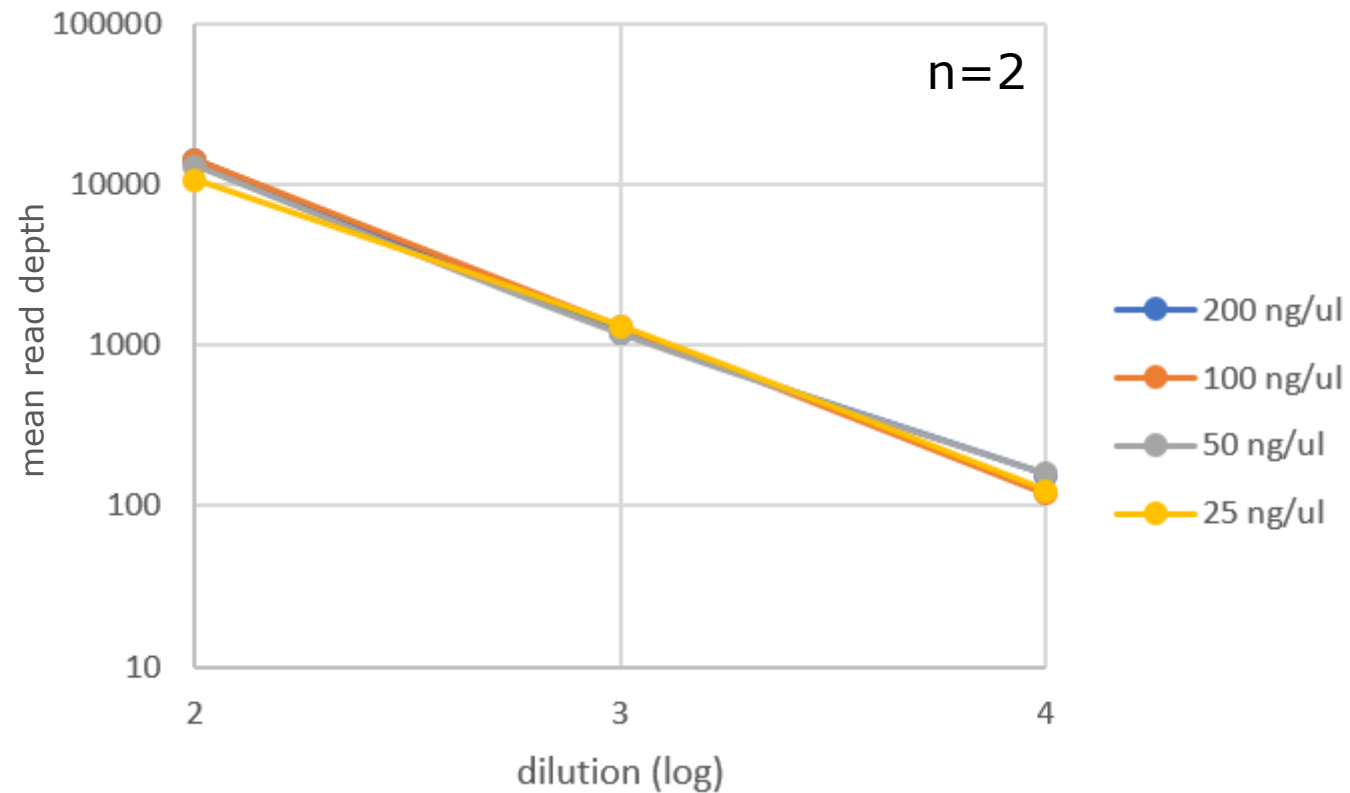
- › Reproducibility of the bio-informatics pipeline
 - Data from one serial dilution (10^2 - 10^5):
100% reproducible (max. 3 nt difference in ToBRFV sequence length)
- › Repeatability and reproducibility with biological samples
 - At low and medium dilutions (10^2 - 10^4):
A single ToBRFV contig (6379-6353 nt) was obtained, all 100% identical
 - At a high dilution (10^5):
11-13 ToBRFV contigs (Σ =3349-5107 nt) were obtained, all leading to a positive ToBRFV identification





Robustness

- > Effect of RNA concentration on virus detection



RNA concentration had no effect on average read coverage



Validation of the HTS test for other viruses/viroids

- > Only the analytical sensitivity should be determined
- > Make use of
 - correlation between mean read depth (MRD) and dilution
 - minimal required data (2 Gb)
 - minimal read depth of 10
- > Example: received 2.5 Gb data (yield) and obtained viral contig has an MRD of 1500
analytical sensitivity = (2 Gb/2.5 Gb) x (1500/10) = 0.8 x 150 = relative dilution rate of 120
- > So far validated the test for 75 viruses/viroids





Conclusions

- › A HTS test can be validated in a similar way as done for other molecular methods
- › In 2022 we became ISO17025 accredited for detection and identification of ToBRFV by HTS.
- › Due to transferable performance characteristics of the HTS test we were able to validate the HTS test for 74 other viruses/viroids, which were added to our flexible accreditation scope in 2023
- › The ISO17025 accredited HTS test allows us, as an official laboratory, to detect and identify potentially all EU-regulated viruses and viroids
- › The single HTS test is providing an efficient and appropriate tool to screen for multiple viruses and viroids in surveys





Acknowledgement

Virology

- › Pier de Koning
- › Marleen Botermans
- › Annelien Roenhorst



Molecular Biology

- › Bart van de Vossenberg
- › Micheal Visser
- › Lucas van der Gouw





Thank you

 **ISTA ANNUAL MEETING 2024**  **01-04 JULY CAMBRIDGE, UNITED KINGDOM**

