Euphresco

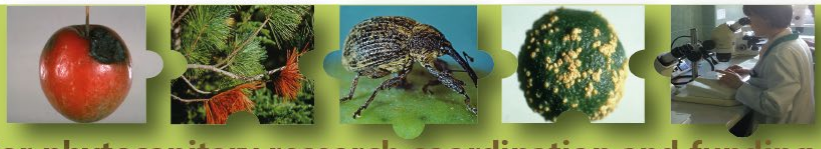
Final Report

Project title (Acronym)
Faster, cheaper identification of emerging virus problems (VIRFAST)

Project duration:

Start date:	2018-09-01
End date:	2021-02-28

Contents



1. Research consortium partners

Coordinator – Partner 1			
Organisation	Université de Liège – Gembloux Agro-Bio Tech (Ulg)		
Name of Contact (incl. Title)	Sébastien Massart (Prof. Dr. ir.)	**Gender**	M
Job Title	Professor		
Postal Address	Passage des déportés 2, 5030 Gembloux, Belgium		
E-mail	sebastien.massart@uliege.be		
Phone	+32 81622431		

Partner 2			
Organisation	Flanders research institute for agriculture, fisheries and food (ILVO)		
Name of Contact (incl. Title)	Kris De Jonghe (Dr. ir.)	**Gender**	M
Job Title	Doctor		
Postal Address	Burgemeester Van Gansberghelaan 96, 9820 Merelbeke, Belgium		
E-mail	kris.dejonghe@ilvo.vlaanderen.be		
Phone	+32 92722448		

Partner 3			
Organisation	Université Catholique de Louvain (UCL), Institut de Recherche Expérimentale et Clinique (IREC), Centre de Technologies Moléculaires Appliquées (CTMA)		
Name of Contact (incl. Title)	Jean-Luc Gala (Prof. Dr.)	**Gender**	M
Job Title	Professor		
Postal Address	Clos Chapelle-aux-Champs, 30 bte B1.30.24, B-1200-Bruxelles, Belgium		
E-mail	jean-luc.gala@uclouvain.be		
Phone	+32 27643165		

Partner 4			
Organisation	Fera Science Limited (FERA)		
Name of Contact (incl. Title)	Ian Adams; Adrian Fox	**Gender**	M; M
Job Title	Scientists		
Postal Address	Biotech Campus, York YO41 1LZ, United Kingdom		
E-mail	Ian.Adams@fera.co.uk; Adrian.Fox@fera.co.uk		
Phone	-		

Partner 5			
Organisation	Canadian Food Inspection Agency (CFIA)		
Name of Contact (incl. Title)	Anna-Mary Schmidt (MSc)	**Gender**	F
Job Title	Head, Grapevine Diagnostics at Canadian Food Inspection Agency		
Postal Address	1400 Merivale Rd Tower 1 -Ottawa, ON K1A 0Y9, Canada		
E-mail	Anna-Mary.Schmidt@inspection.gc.ca		
Phone	-		

Partner 6			
Organisation	US Department of Agriculture, Animal and Plant Health Inspection Service (APHIS)		
Name of Contact (incl. Title)	Gang Wei	**Gender**	M
Job Title	scientist		
Postal Address	4700 River Rd, Riverdale Park, MD 20737, United States		
E-mail	Gang.wei@aphis.usda.gov		
Phone	+1 2027997132		

Partner 7			
Organisation	National Institute of Biology (NIB)		
Name of Contact (incl. Title)	Natasa Mehle; Denis Kutnjak; Maja Ravnikar	**Gender**	F; M; F
Job Title	Senior Research Associate, post-doctoral researcher, professor		
Postal Address	Večna pot 111, 1000 Ljubljana, Slovenia		
E-mail	natasa.mehle@nib.si; maja.ravnikar@nib.si; denis.kutnjak@nib.si		
Phone	+386 59232808		

Partner 8			
Organisation	French Agency for Food, Environmental and Occupational Health & Safety (ANSES)		
Name of Contact (incl. Title)	Delphine Masse	**Gender**	F
Job Title	Scientist		
Postal Address	14 rue Pierre et Marie Curie - 94701 Maisons-Alfort Cedex, France		
E-mail	delphine.masse@anses.fr		
Phone	+33 149771350		

Partner 9			
Organisation	Science and Advice for Scottish Agriculture (SASA)		
Name of Contact (incl. Title)	Emma Back; Christophe Lacomme	**Gender**	F; M
Job Title	Scientists		
Postal Address	1 Roddinglaw road, Edinburgh, EH12 9FJ, UK		
E-mail	emma.back@sasa.gov.scot ; christophe.lacomme@sasa.gov.scot		
Phone	+44 131 2448916		

2. Short project report

2.1. Short executive summary

Plant pest diagnosis is performed by official laboratories upon request of National Plant Protection Organizations (NPPOs), growers or traders, in samples that inspectors have collected *in situ* (a consignment, a place of production, an outbreak area, a buffer zone, *etc*.). Resources allocated to official laboratories have decreased over time, while trade in plants and plant products, and consequently the material to be tested, have increased steadily. As indicated in the Euphresco Strategic Research Agenda (priority R-6), on-site detection and identification tests, that are both high throughput and scalable at contained costs should be developed and validated to accelerate diagnosis (especially in the case of perishable goods) and to relieve pressure on laboratories. In particular, the application of on-site detection and surveillance methods for plant viruses and viroids is needed for the quick assessment of the health status of plant material. New protocols and technologies based on sequencing are under development (such as Oxford nanopore direct RNA sequencing - cDNA sequencing), but need to be tested, optimised and validated in order to be applicable in routine on-site testing. Furthermore, current barriers for the use of novel on-site technologies need to be addressed.

The VIRFAST project explored the possibility of on-site virus diagnosis with the MinION technology (Oxford Nanopore sequencing). Early results indicated that the traditional RNA extraction protocol could be bypassed resulting in a quicker and low-tech solution for practical diagnostics on-site. However, skipping the extraction needs additional development for diagnostic purposes, as the sequencing threshold and, consequently, the sensitivity of the test is greatly reduced. However, it should be noted that validation was performed using dehydrated samples and moreover some participating laboratories used the Oxford nanopore sequencer on crude extract material for the first time. Better analytical sensitivity is expected when working with fresh tissue and if the users are familiar with the use of crude extract material.

Two sample preparation kits were also compared starting from purified RNA extracts. In summary, better results for virus/viroid detection when using cDNA-PCR library preparation approach, compared to direct RNA sequencing approach, due to the higher relative amount of sequencing reads of viral origin. The increased analytical sensitivity when using the cDNA-PCR library preparation opens the possibility of using this method to monitor an outbreak in the field, in a middle-tech laboratory or in remote locations with limited access to plant diagnostics facilities.

A survey of 26 NPPO laboratories showed that while only 15% were using high-throughput sequencing (HTS) in their diagnostic workflows in 2019, almost all wished to adopt it and that improved staff training and reduced costs would be essential for this uptake.

The results of a survey on on-site testing allowed also to observe interesting trends. Twelve laboratories were already using on-site testing with an estimated volume ranging from 25 to 2,500 tests per year. The main context for using on-site testing kits was field inspection (63%), followed by screening test in the laboratory (19%). Time saving was the most important advantage for use. The laboratories also highlighted the ease of use. The three main limitations for a wider use of HTS were the lack of validation data, the limited sensitivity and the cost.

2.2. Project aims

The project had several aims:

- To select and develop innovative detection/identification method(s) that can be used for fast, reliable and cost-effective on-site detection of (un)known and emerging harmful viruses on plants and plant products (sample preparation, sequencing, data-analysis)
- To compare the validated detection/identification method(s) with methods that are currently being used
- To increase awareness on the use of HTS among future users and risk managers
- To validate methods and tools for data analysis: software, their parameters and database
- To identify barriers and recommendations to adopt these methodologies in the current legal framework R2000/29 (and R2016/2031-R625/2017 as from 12/2019) as an official test method.

2.3. Description of the main activities

2.3.1. Testing new protocols using the Oxford Nanopore Technology (ONT)

The aim of the task was to assess the possibility to perform sequencing on crude extract without the RNA extraction step. The advantages of this method are the possibility to deliver a rapid on-site diagnosis and the reduced cost which could be helpful for the adoption of the method in developing countries. All the samples used in the experiments were previously sequenced by Illumina; a standard RNA extraction (RNeasy by Qiagen) was also added as a comparison in all runs.

The evaluation on crude extract preparation vs. Qiagen RNeasy extraction kit was carried out in 8 different laboratories. The following plants and viruses were included in the validation:

Plant sample	Infecting virus	Viral genus
Prunus spp.	Little cherry virus-1, Cherry virus A	Ampelovirus, Capillovirus
Potato	Potato virus Y	Potyvirus
Banana	Banana bract mosaic virus	Potyvirus
Potato	Tomato chlorosis virus, Tomato leaf curl New Delhi virus, Potato virus S and Potato virus Y	Crinivirus, Begomovirus, Carlavirus, Potyvirus
Sweet potato	Sweet potato leaf curl virus (SPLCV)	Begomovirus
Hydrangea sp.	Hydrangea ringspot virus (HRSV)	Potexvirus
Oxalis sp.	Arabis mosaic virus (ArMV)	Nepovirus
Prunus sp.	Plum pox virus (PPV)	Potyvirus
Tomato	Pepino mosaic virus (PepMV)	Potexvirus
Yam	Yam virus Y (YVY)	Unclassified Betaflexiviridae
Sweet potato	Sweet potato feathery mottle virus, Sweet potato virus G	Potyvirus

Different library preparation protocols were used before sequencing the samples multiplexed on a Minion flowcell (ONT).

A test performance study (TPS) was organized with six laboratories (in addition to the Belgian organizers). Virus-infected plant leaves were lyophilized, homogenized and tested using the same direct cDNA sequencing protocol to validate the presence and the detectability of the

viruses. The use of dehydrated samples enhanced the homogenization and simplified the shipping. All the samples were sent pre-weighted and ground in fine powder. The participating laboratories were asked to run this crude extract in parallel to RNA extracted using the RNeasy extraction kit (Qiagen). The laboratories could run the libraries obtained from crude extract and from RNA extract a single flowcell or choose to run them on separate flowcells. Three dehydrated samples of the same origin were sent in sextuplicate (six copies), half to be sequenced from crude extract, and the other half to be sequenced after RNA extraction. The plant-virus selected was Hydrangea spp. infected with HRSV (Potexvirus), tomato sample infected with Pepino mosaic virus (Potexvirus), and sweet potato infected with Sweet potato virus G (Potyvirus). Samples were sent in double-blind. The evaluation of the library preparation protocol was carried out on RNA samples previously characterized: https://doi.org/10.3389/fmicb.2017.01998. For RNA isolation the RNeasy kit (Qiagen) was used with DNAse treatment following the manufactured instructions. Additionally, polyA tailing was implemented using *Escherichia coli* Poly(A) Polymerase. The samples were sequenced on MinION using two nanopore sequencing library preparation methods: Direct RNA and cDNA-PCR. More details on the laboratory protocols and the different tests carried out for crude extract evaluation are presented in Appendix 1. Bioinformatics analyses were carried out using the algorithms implemented in the participant laboratories to detect viruses in sequencing datasets.

2.3.2. Surveys of diagnostic laboratories

Two surveys were carried out in collaboration with the Valitest project: the first one was focused on on-site testing while the second one was focused on the use of HTS by plant pest diagnostic laboratories.

2.4. Main results

The results of the evaluation of the use of crude extracts are detailed in Appendix 2. This report includes the results of the TPS carried out with six laboratories.

The overall results of the TPS are difficult to compare as the yield obtained by the laboratories was highly divergent and the repeatability was variable too. The number of nucleotides mapping the control, highlights the better detection with the RNA extraction over the crude extract, but at the same time it shows the variability between hosts and viruses with sweet potato being more difficult to detect. Table 1 summarizes the results obtained using a threshold of 1 000 nucleotides mapped onto the control to call a positive. This TPS showed a great variability of response between laboratories but overall, a good virus detection wen using the RNA extraction method (91% detection). The results showed that the six laboratories that had not used the protocol before were able to achieve a detection of 50% which raised to 62.5% (the results from L07 were excluded because of an electric power incident during the run that compromised the results). It should be noted that the threshold used was very low. The results clearly show that the method cannot be used for diagnostics, but detection of plant virus from crude extract direct cDNA sequencing with ONT is possible.

Table N1: Viruses detected by the different laboratories in the frame of the TPS. The numbers (from 1 to 3) represent the number of sample(s) where 1 000 nucleotides or more mapped the correct virus from the three replicates. The percentage correspond to the diagnostic sensitivity (DSE) obtained from RNA extracts or crude extracts.

	Hydrangea crude	*Hydrangea RNA*	*Sweet potato crude*	*Sweet potato RNA*	*Tomato crude*	*Tomato RNA*	DSE *RNA*	*DSE CRUDE*
L01	2	2	0	3	1	2	78%	33%
L03	3	3	0	3	3	3	100%	67%
L04	1	3	1	3*	3	3*	100%	56%
L05	3	3	3	3	3*	3*	100%	100%
L06	0	2	1	1	3	3	67%	44%
L07	0**	3	0**	3*	0**	3*	100%	0%
Average DSE	50%	89%	28%	89%	72%	94%	91%	50 %

* indicates that at least 1 000 nucleotides of another sample mapped onto this virus (contamination)
** indicates that the run was impacted by a power incident

The results of the comparison between the sensitivity obtained with direct RNA vs. cDNA - PCR preparation kits are detailed in Appendix 2.

The results obtained during the surveys on on-site diagnostics and HTS technologies are detailed in Appendix 3. In total, 26 laboratories across 20 countries participated in the surveys. Only four of the laboratories reported the use of HTS for routine diagnostics with two other laboratories that use it for research purposes. All the laboratories that use HTS routinely reported using it for “Resolving unknown etiology of symptoms” either as an alternative of, or after more classical molecular tests. All four laboratories identified the interpretation of results as the most challenging aspect of the procedure. Data analysis was also identified as challenging by three laboratories. Lack of validation and cost were identified as the bottle necks to greater adoption of the technology. Amongst the 22 laboratories that do not use HTS for diagnostics, 73% identified the cost as the main reason for not adopting HTS. Other reasons included the complexity (50%) and the lack of expertise (50%). One laboratory stated that HTS was not relevant for their organization. These laboratories were asked what could be changed to increase their likelihood of using HTS. Training and lower price were rated essential by the most laboratories. Faster turnaround as well as frameworks and guidelines for validation were considered essential of important by the majority of the labs.

For the rapid testing, twelve laboratories answered they use on-site testing with an estimated volume ranging from 25 to 2500 tests per year. The main context for using on-site testing kits was field inspection (63%), followed by screening test in the laboratory (19%). This means that these kits allowing immediate response have a significant use in the laboratory. On-site tests in the laboratory can save time and are easy to use. However, the ability to take immediate decision was considered an important criterion for choosing on-site tests for less than half of

the responders. The tree main limitations for a wider use of the test were the lack of validation data, the limited analytical sensitivity and their cost. The absence of tests for some pests also limits its use. Participants identified the pests for which they consider tests should be developed/validated: for viruses, plum pox virus (3 responses), for fungi and oomycetes *Phytophtora ramorum* (4 responses), and for bacteria *Erwinia amylovora*, *Ralstonia solanacearum* and *Xylella fastidiosa* (4 responses each).

2.5. Conclusions and recommendations to policy makers

The results of the TPS, using challenging matrix (lyophilized plant tissues) showed a high analytical sensitivity in normal conditions (i.e. with RNA extraction). The comparison test between different libraries showed a 100% virus/viroid detection across the Baltimore classification when using the cDNA-PCR of total RNA kit.

Results indicated that the preparation of crude extract instead of purified RNA allowed the detection of viruses but lowered the analytical sensitivity of the test. However, the tests were done in challenging conditions, as validation was performed using dehydrated samples and moreover some participating laboratories used the Oxford nanopore sequencer on crude extract material for the first time. Better analytical sensitivity is expected when working with fresh tissue and if the users are familiar with the use of crude extract material. The library preparation kit used in this project was a time-consuming process, but new developments for rapid library preparation with the use of enzymes resistant to PCR-inhibitors have the potential to improve the performance of the technology. These enhancements should open to the possibility of using the Oxford nanopore sequencer to monitor an outbreak in the field, in a molecular biology laboratory or in remote locations with limited access to plant diagnostics facilities.

A survey of 26 NPPO laboratories showed that while only 15% were currently using HTS in their diagnostic workflows, almost all wished to adopt it and that improved staff training and reduced costs would facilitate this uptake. The results of a survey on on-site sampling and testing allowed also to observe interesting trends in the use of the kits, highlighting the most popular uses, defining the current bottleneck and identifying candidate pests for kit development.

There is a growing interest and progressive use of HTS by plant health laboratories. Nevertheless, each laboratory has its own laboratory and bioinformatics protocols without any harmonization for the validation and the routine use of HTS. There is therefore a strong need for guidelines to be written and implemented. For the use of on-site testing, at present, the drawbacks are the cost (LAMP technology) and/or the analytical sensitivity (LFD technology). They could be used as an early warning system. There is a strong need of validation data for the available tests to determine in which cases on-site result must be confirmed in the laboratory.

2.6. Benefits from trans-national cooperation

The project generated new collaborations nationally (ILVO, ULiege and the CTMA - UCL) and internationally (USDA, NIB, SASA, FERA, CFIA and ANSES) between project partners but also for other volunteering laboratories.

The collaboration between ULg, ILVO and CTMA that emerged through the Euphresco frame will be continued through national projects for better understanding of plant and animal viruses in Belgium.

3. Publications

3.1. Article(s) for publication in the EPPO Bulletin

None.

3.2. Article for publication in the EPPO Reporting Service

None.

3.3. Article(s) for publication in other journals

The consortium is currently evaluating the opportunity to publish the results of the TPS and of the methodological evaluation in scientific peer-review journals.

4. Open Euphresco data

None.

Appendix 1 Testing new protocols using the Oxford Nanopore technology (ONT)

T2.1 Testing the influence of Nucleic Acid preparation on the sequencing quality

The aim of the task is to assess the possibility to perform a ONT sequencing run on crude extract without the RNA extraction step in order to deliver a rapid on the on-site diagnostic (minimum point of care). Furthermore, an economical and simple tool (without RNA extraction step) could be helpful to monitor a disease outbreak in developing countries. All the samples used in the experiments were already sequenced by Illumina; a standard RNA extraction (RNeasy by Qiagen) was also added as a comparison in all runs.

The **first assay** explored the possibility of detecting the presence of a polyadenylated virus from crude extract.

Three plants selected:

1. *Prunus* infected with cherry virus A (CVA, Capillovirus) and little cherry virus 1 (LChV-1, Ampelovirus) [non-polyA]),
2. Potato infected with potato virus Y (PVY, Potyvirus)
3. Banana infected with banana bract virus mosaic virus (BBrMV, Potyvirus).

Three extraction methods assessed:

1. RNAclean XP (Agencourt): crude extract + RNA clean-up on magnetic beads-1/10 dilution,
2. Crude extract (dilution of tissue extract with Kaji buffer– 1/400 dilution)
3. Standard RNA extraction with RNeasy Mini Kit or RNeasy plus mini kit (Qiagen).

For the RNeasy (Qiagen) extraction, the standard protocol was followed using the RLT buffer. For the RNAClean and the crude extract, samples were ground in Kaji buffer (137 mM NaCl; 8 mM Na2HPO4; 1.5 mM KH2PO4; 2.7 mM KCl; 80 mM Na2SO3; 3 mM NaN3; 0.05% Tween, pH7.2) at the ratio 1/10 (100 mg in 1 mL) using a FastPrep-24™ tissue homogenizer (Lysing Matrix D). Samples were then span at maximum speed for 5 min (4°C). Extracts were filtered (45 µm syringe adaptor) and RNA was concentrated using RNAclean XP or diluted a further 40x in Kaji buffer (1/400). Library preparation was performed for the tree extraction methods following the Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104) protocol. The initial RNA input was measured for the Qiagen RNA to 1,5 µg. For the crude extract, the RNA could not be quantified and 7,5 µL was used, the maximum volume allowed by the protocol. In brief, the RNA was reverse transcribed to cDNA and strand switched. After an RNase treatment and a DNA clean up (AMPure XP, Beckman), the second strand was synthesized and the DNA was cleaned up again. Sample were left overnight at 4°C. The double-stranded cDNA was then end-repaired and A-tailed followed by an AMPure XP beads (Beckman Coulter) purification. The samples were then barcoded using Native Barcoding Expansion 1-12 Kit (EXP-NBD104). Samples were then pooled together and ONT sequencing adaptors were ligated followed by an AMPure beads clean-up before loading into a MinION flow cell (FLO-MIN106D R9) for an 48h run. The MinION was connected to a MinIT.

The **second assay** assessed the effect of a dilution series of crude extract on the total and viral sequencing throughput.

Two plants selected:

1. Potato sample infected with non-polyA viruses (Tomato chlorosis virus, ToCV, Crinivirus; Tomato leaf curl New Delhi, ToLCNDV, Begomovirus), and polyA viruses potato virus S (PVS, Carlavirus) and PVY (Potyvirus)
2. Banana infected with BBrMV.

Four conditions assessed:

1. Crude extraction in Kaji buffer – 1/10 dilution
2. Crude extraction in Kaji buffer – 1/50 dilution
3. Crude extraction in Kaji buffer – 1/100 dilution
4. Qiagen RNeasy extraction

Extraction protocols were unchanged from the first assay. Samples were ground in Kaji buffer (100 mg in 1 mL) using a FastPrep-24™ tissue homogenizer (Lysing Matrix D). Samples were then centrifuged (10,000g on a benchtop centrifuge) for 5 min (4°C). Crude extract 1/10 were used directly, diluted 5x (1/50) and 10x (1/100). The filtration step was omitted. RNA from the banana sample was extracted the same day following the same protocol used for the first run (RNeasy, Qiagen). The Kaji buffer was the same as the one used in run 1 and had been kept in the fridge for the time (2 months). Sample preparation was identical to the run 1 except for the homogenization of the banana tissue. After the centrifugation the tissue observed was not ground well and it was ground again on the FastPrep-24™ tissue homogenizer (1 min and then re-centrifugation). Despite this the extract observed was very pale compared with the one of the potato sample. Library preparation was identical to the first run: ONT Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104). Pooled libraries were loaded into a MinION flow cell (FLO-MIN106D R9) for an 48h run. The MinION was connected to a MinIT with a new version of the software (MinIT v19.05.02 using MinKNOW v3.3.2 and Guppy v3.0.3).

The **third assay** (run 3 and 4) compared the sequencing of the RNA extracted and the crude extract on two different flow-cells to eliminate a potential competition effect. All the samples were prepared in duplicate (eight barcodes loaded on each flowcell).

Four plants selected:

1. *Prunus* infected with CVA
2. Potato sample infected with non-polyA viruses (ToCV, and ToLCNDV), and polyA viruses PVS and PVY
3. Banana infected with BBrMV
4. Sweet potato infected with sweet potato leaf curl virus (SPLCV, Begomovirus).

Two conditions assessed:

1. Crude extraction in Kaji buffer – 1/10 dilution
2. Qiagen RNeasy extraction

Extraction protocols were unchanged from the first assay. Crude extracts were ground in Kaji buffer and used at the concentration of 1/10 and the purified RNA was prepared with the

RNeasy (Qiagen) extraction standard protocol. The same library preparation was used (ONT Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104)). One library was performed for the crude extract, and one for the purified RNA extract each loaded on distinct flowcell. Pooled libraries were loaded into a MinION flow cell (FLO-MIN106D R9) for a 48h run.

The **fourth assay** (run 5) was a test of a different library preparation with a PCR-cDNA sequencing kit, to improve the cDNA amount loaded onto the flowcell from the crude extract.

Six plants selected:

1. Hydrangea infected with hydrangea ringspot virus (HRSV, Potexvirus)
2. Oxalis infected with Arabis mosaic virus (ArMV, Nepovirus)
3. *Prunus* infected with plum pox virus (PPV, Potexvirus)
4. Sweet potato co-infected with sweet potato feathery mottle virus (SPFMV, Potyvirus) and sweet potato virus G (SPVG, Potyvirus)
5. Tomato infected with pepino mosaic virus (PepMV, Potexvirus)
6. Yam infected with yam virus Y (YVY, Unclassified Betaflexiviridae)

Two conditions assessed:

1. Crude extraction in Kaji buffer – 1/10 dilution
2. Qiagen RNeasy extraction

The extraction protocols remained unchanged, but a new library preparation was tested the PCR-cDNA sequencing kit (SQK-PCS109) following the instructions from the manufacturer. Twelve cycles were selected for the PCR amplification. Pooled libraries were loaded into a MinION flow cell (FLO-MIN106D R9) for an 48h run.

The **fifth assay** (run 7) compared the sequencing of fresh and dehydrated tissues using the RNA extracted and the crude extract on the same flow-cells. For this run, the original library ONT Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104) was used.

Three plants selected:

1. Hydrangea infected HRSV
2. Sweet potato co-infected with SPFMV and SPVG
3. Tomato infected with PepMV

Four conditions assessed:

1. Qiagen RNeasy extraction on frozen tissue
2. Qiagen RNeasy extraction on dehydrated tissue
3. Crude extraction on frozen tissue
4. Crude extraction on dehydrated tissue

In order to simplify the sampling, all the dehydrated tissue had been lyophilized in FastPrep® Lysis Beads & Matrix Tubes with the fresh weight noted on the tube. Sample preparation and sequencing followed the methods used in the third assay.

The **sixth assay** (run 8) was a ring-test across seven international laboratories. Samples were anonymised and randomised. A total of 18 samples were supplied to each participant. Three plants were sampled and dehydrated in FastPrep® Lysis Beads & Matrix Tubes and ground (dry) with the fresh weight of tissue recorded on the tubes. The samples were divided in two identical lots of nine tubes, one for each treatment (RNA extraction and crude extract). For each treatment, the samples were in triplicates.

Three plants selected:

1. Hydrangea infected with HRSV
2. Sweet potato infected with SPVG
3. Tomato infected with PepMV

Two conditions assessed:

1. Qiagen RNeasy extraction on dehydrated tissue
2. Crude extraction on dehydrated tissue

The seven laboratories participating (randomly named L01 to L07 further on) were:

A. USDA, S&T Beltsville Laboratory, USA
B. NIB, Ljubljana, Slovenia
C. SASA, Edinburgh, Scotland
D. FERA, York, United Kingdom
E. ANSES, Reunion Island, France
F. ILVO, Merelbeke
G. CTMA, Bruxelle, Belgium

The same instructions were given to the participants:

"*You will receive nine samples in duplicate (a total of 18 tubes), all made of eq. 100 mg fresh weight of dehydrated powder. Nine samples will be for your standard RNA extraction protocol (e.g. RNeasy, Qiagen), and the other nine for a crude extract protocol. The samples will be sent in fast prep tubes containing the ceramic beads and the grinding will be done prior to the shipment (you will have plant dry powder with beads ready to add the buffer). The fresh weight of tissue added will be indicated on each tube*."

The 18 samples can be prepared for one single flowcell, or divided in one flowcell extraction type:

1. Run crude extracts (CE) and the purified RNA (PR) samples on two distinct flowcells. Add as much CE cDNA as possible on a single flowcell (normalized to the lowest concentration) and add the optimal amount of cDNA from the PR on a second flowcell.

2. Run the 18 samples on a single flowcell. Quantify the molarity of all the samples (bioanalyzer for the size and fluorometer (e.g. Qubit) for the concentration) and set the molarity to the lowest sample of CE (unless far too low), then halve this amount for the cDNA obtained from PR. In that case, the amount of cDNA loaded is far from optimal loading concentration recommended by ONT. For example, in our last run we normalized our cDNA from CE to 10 fmols, and to 5

fmols for the PR to load a total of 75 fmols (for 12 samples, 6 PR + 6 CE including 2 samples with less than 10 fmols of cDNA). The optimal amount to load is 200 fmols.

The library(ies) can be prepared in two days as the protocol includes multiple time-consuming Ampure DNA clean up steps.

For the CE protocol, use the following buffer: 137 mM NaCl; 8 mM Na2HPO4; 1.5 mM KH2PO4; 2.7 mM; KCl; 80 mM Na2SO3; 3 mM NaN3; 0.05% Tween, adjusted to 7,2 with HCl. It is critical to have the shortest time between the addition of the buffer and the beginning of the library (prepare the first mix in advance and keep on ice).

1. Spin the tubes before opening.
2. Add 10 volumes of buffer to the fresh weight of the sample (+/- 1 mL) and mix well (vortex 1 minute).
3. Centrifuge 3 minutes at maximum speed (15K g) at 4 °C.
4. From the supernatant, add 7,5 µL (maximum volume possible) to the RT reaction
5. Then follow the protocol as described for all the samples (alongside with the PR samples).

Follow the manufacturer instruction then on (Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104 and EXP-NBD114) Version: DCB_9091_v109_revL_14Aug2019).

T 2.2: Evaluating different protocols for sample preparation

Materials & Methods

The aim of the task was to sequence RNA plant samples infected by diverse viruses using two existing protocols: cDNA-PCR sequencing and direct RNA sequencing from Oxford Nanopore Technologies.

For this purpose, we selected sample of tomato plant, infected with 5 different plant viruses (among those, one viral species with two different strains) and a viroid. The sample was previously also very well characterized (Pecman et al., Frontiers in Microbiology, 2017; https://doi.org/10.3389/fmicb.2017.01998).

For RNA isolation the RNeasy kit (Qiagen) was used with DNAse treatment by following the manufactured instructions. Additionally, polyA tailing was implemented using E. coli Poly(A) Polymerase (NEB# M0276).

The sample was sequenced on MinION using following nanopore sequencing library preparation approaches: Direct RNA (SQK-RNA002) and cDNA-PCR (SQK-PCS108). Each prepared library was sequenced on a separate flow cell for 48 h using standard MinKNOW software (v18.12.6). Generated reads were basecalled using Guppy v3.1.5 and command: rna_r9.4.1_70bps_hac.cfg/ dna_r9.4.1_bps450_hac.cfg --device auto --u_substitution false.

The statistic and quality of MinION-generated fastq reads were checked using NanoQC v0.8.1, NanoStat v1.1.2 and NanoPlot v1.20.1 programs. Plots generated by the program NanoPlot v1.20.1. were inspected for each sample individually, following by the decision how to trim reads (reads length, trimming head of the reads and tail of the reads) using program NanoFilt

2.5.0. Trimmed/filtered fastq reads were again checked using NanoStat v1.1.2 and mapped to viral RefSeq (NCBI database, updated 19.05.2019) using minimap2 (v2.16-r922) and commands: minimap2 -ax splice -uf -k14 for directRNA reads and minimap2 -ax map-ont for cDNA-PCR reads. The reads were also analysed using Diamond (v0.9.22) with --frameshift 15 --range-culling --sensitive command option. Additionally, all reads were de novo assembled using the combination of minimap2, miniasm and racon and the output (contigs) were analysed with using blastN against NCBI *nt* database. All results were individually inspected before final conclusions.

In the next step all reads were mapped (minimap2 (v2.16-r922) to the corresponding consensuses (obtained with Illumina reads; Pecman et al., 2017) for further analysis: e.g., the calculation of mapped reads % and average mapped reads identity.

Appendix 2 Testing new protocols using the Oxford Nanopore Technology (ONT) & Testing ONT technology on plant samples

First assay on crude extract preparation

The **first assay** explored the possibility of detecting the presence of a polyadenylated virus from crude extract and compare it with standard RNA extraction (Qiagen) and crude extract combined with RNAclean XP (Agencourt) cleanup. Optimum DNA loading concentrations could not be reached for most samples. The Qiagen RNA extraction from *Prunus* was very low to start with while the potato and banana were fine.

Table 1: final DNA concentration for each library (target 50 fmols)

	RNAclean XP	Crude extract	RNeasy (Qiagen)
Prunus	41.6 fmols	undetectable	6.1 fmols
Potato	26.6 fmols	undetectable	56.2 fmols
Banana	8.4 fmols	undetectable	17.3 fmols

The libraries were loaded completely on the MinION (no normalization). The sequenced were filtered at Q7 which retained only 3% of the total reads. Overall, a lot more data were obtained from the RNAclean XP and Qiagen (>10,000 reads/sample) than in the crude extract samples (<1000 reads/sample).

Table 2: Sequencing throughput and average sequence length obtain for each library (total and viral) during the first MinION run

	RNAclean XP		**Crude extract**		**RNeasy (Qiagen)**	
	Total number of reads (read length)	Number of **viral** reads (read length)	Number of reads (read length)	Number of viral reads (read length)	Number of reads (read length)	Number of **viral** reads (read length)
Prunus	21,918 (1013)	56 (1179)	383 (1056)	1 (1828)	6,954 (1346)	2 (2380)
Potato	28,172 (781)	92 (1072)	36 (1303)	2 (802)	68,532 (1470)	620 (3465)
Banana	11,088 (1039)	33 (869)	84 (1152)	0 (na)	17,765 (1149)	90 (1763)

Only the polyadenylated viruses were detected across extraction methods, and PVY was the only virus detected in the crude extract of potato.

Figure 1: representation of the read number obtained after sequencing RNA from RNAclean XP (Agencourt) (left, labelled beads), crude extract (middle, labelled unpurified) and Qiagen (right) on *Prunus* (blue) potato (grey) or banana (red). Read number is expressed in log scale.

Conclusion: All viruses were detected from the RNA extraction protocol as well as with the bead's extraction method. Very low sequencing reads were obtained from the crude extract, but two viruses were detected. The initial dilution of the extract (400x) was selected as it is used in some current RT-PCR protocols. However, the poor sequencing results would suggest trialling a dilution series.

Second assay on crude extract preparation

The second assay aimed to assess the effect of the dilution of crude extract on de detectability of polyadenylated viruses.

Two plants were tested (banana and potato) and four treatments (crude extract diluted 10, 50 and 100 x and RNA extracted with RNeasy Qiagen). The banana sample used was collected from the same plant as run 1, the potato sample is a different one with multiple virus infection.

Similar amount of final library was measured from RNeasy (Qiagen) purified samples (around 20-100 fmols) as for the run1 while the concentration of final library obtained for the crude extract samples (1/10 1/50 and 1/100) were too low to be quantified and therefore the representation of the RNeasy (Qiagen) sample was higher. Around 20 fmols of DNA was loaded to the flowcell.

The sequence throughput was considerably higher than run 1 (6174.21 Mb vs 186.77 Mb) with a better representation of all samples.

Due to the difference experiment settings, data from run 1 and 2 could not be directly compared.
As observed before the RNeasy (Qiagen) RNA extraction gave the best throughput (1665 and 5375 Mb), but it was also the only one with optimum final library quantity loaded into the flow cell for the run.
The banana sample diluted 10 times had the lowest throughput which may correspond to a material lost during the second AMPure XP wash (beads lost during washing steps).
With the exception of this banana sample, the throughput was to be correlated with the dilution steps and the potato sample showed better throughput than the banana sample (banana 2.8 - 26.35>11.18 Mb; potato 139.25 > 31.14 > 23.14 Mb). Overall, those yields were sufficient for the identification of the plant pathogens using MinION Nanopore sequencing.
For the potato sample (RNeasy (Qiagen) purification), the read length distribution histogram shows a pick at around 8 kb probably corresponding to the genome size of the PVS present in very high concentration in this sample (Figure 2). No obvious pick of viral origin was observed in the other samples.
Using Kraken2 (v.2.0.7), a program of taxonomic assignment of metagenomics data, more than 25% of the potato reads are of viral origin where the banana samples are below 1%. The lower viral titer in banana was also observed in the RNeasy (Qiagen) sample and is similar to what was observed in run 1.

Table 3: Sequencing throughput and average sequence length obtain for each library (total and viral) of the second MinION run

		Banana	**Potato**
Crude extract 1/10 dilution	Total number of reads [Throughput](read length)	4166 [2.8 Mb] (671 bp)	106477 [139.25] (1308 bp)
	Number of viral reads [Throughput] (read length)	20 [0.04 Mb] (1803 bp)	6454 [695 Mb] (1078 bp)
Crude extract 1/50 dilution	Total number of reads [Throughput](read length)	29480 [26.35 Mb] (894 bp)	31685 [31.14 Mb] (983 bp]
	Number of viral reads [Throughput] (read length)	30 [0.06 Mb] (2124 bp)	10048 [11.11 Mb] (1105 bp)
Crude extract 1/100 dilution	Total number of reads [Throughput](read length)	13796 [11.18 Mb] (810 bp)	21893 [23.14 Mb] (1057 bp)
	Number of viral reads [Throughput] (read length)	24 [0.07 Mb] (2765 bp)	6573 [8.79 Mb] (6573 bp)
RNeasy (Qiagen)	Total number of reads [Throughput](read length)	1440135 [1665.17 Mb] (1156 bp)	3002437 [4275.18 Mb] (1424 bp)
	Number of viral reads [Throughput] (read length)	4368 [9.62 Mb] (2202 bp)	139730 [483.28 Mb] (3459 bp)

Several viruses identified from the potato sample are related to each other (genus Begomovirus e.g. tomato leaf curl viruses, bitter gourd yellow vein virus, luffa yellow mosaic virus, Macroptilium common mosaic virus, and papaya leaf crumple virus). Some of them are detected in high coverage but for those with lower read number some sequencing error could result in wrong taxonomic assignation. Those begomoviruses have a DNA genome and no poly A tail and were not expected to be sequenced. The virus PVY was detected in high read number but it is interesting that the PVS was much more prevalent than the PVY as the opposite was observed in the Illumina sequencing results (Table 4, Figure 4). However, a second analysis of the Illumina data confirmed the presence of the begomoviruses detected by MinION. One important point is that the data generated by Illumina was performed on the mother plant where the MinION was done on one of the progeny. This could explain the discrepancy between the results observed.

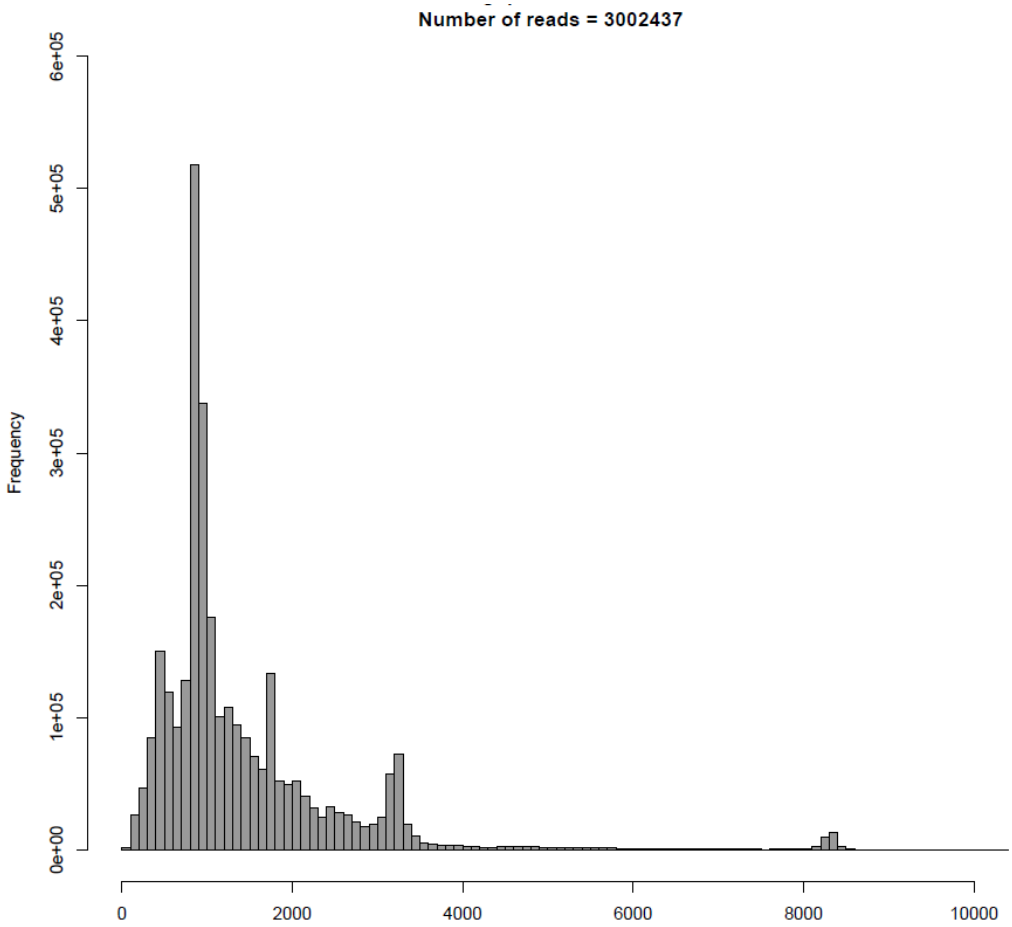


Figure 2: Read length distribution Potato sample, Qiagen extraction (RNeasy). The pick observed at 8000 bp is caused by the very high level of the virus potato virus S.

Table 4: difference of viral observed between the Illumina and minION sequencing

Virus	Illumina	MinION
ToCV	0.02%	Not detected
PVS	1.5%	Very high
PVY	23.3%	moderate
ToLCNDV	1.3%	Multiple species within genus

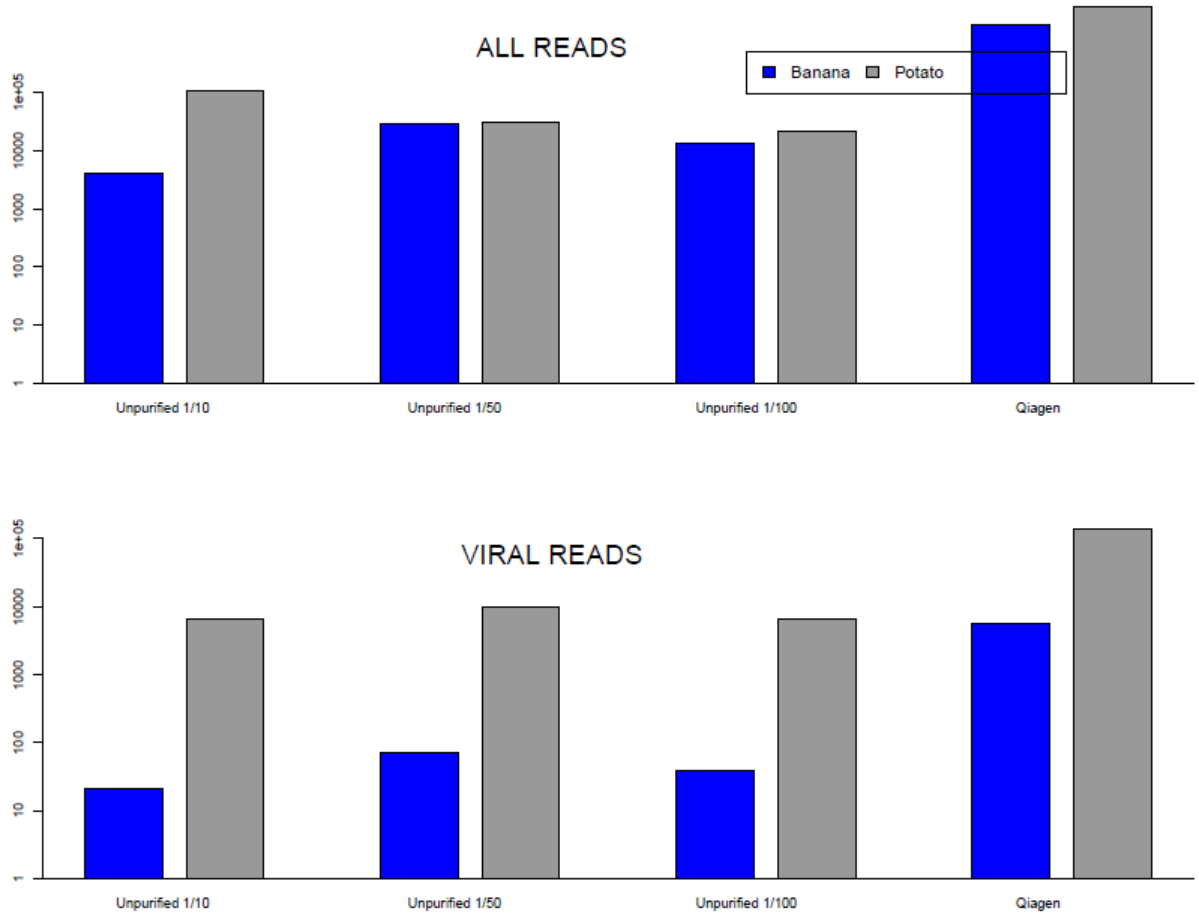


Figure 3: representation of the read number obtained after sequencing RNA from crude extract diluted 10x (left, labelled Unpurified 1/10), 50x (labelled Unpurified 1/50), 100x (labelled Unpurified 1/100), and RNeasy (Qiagen) (right) on banana (blue) and potato (grey). Read number is expressed in log scale.

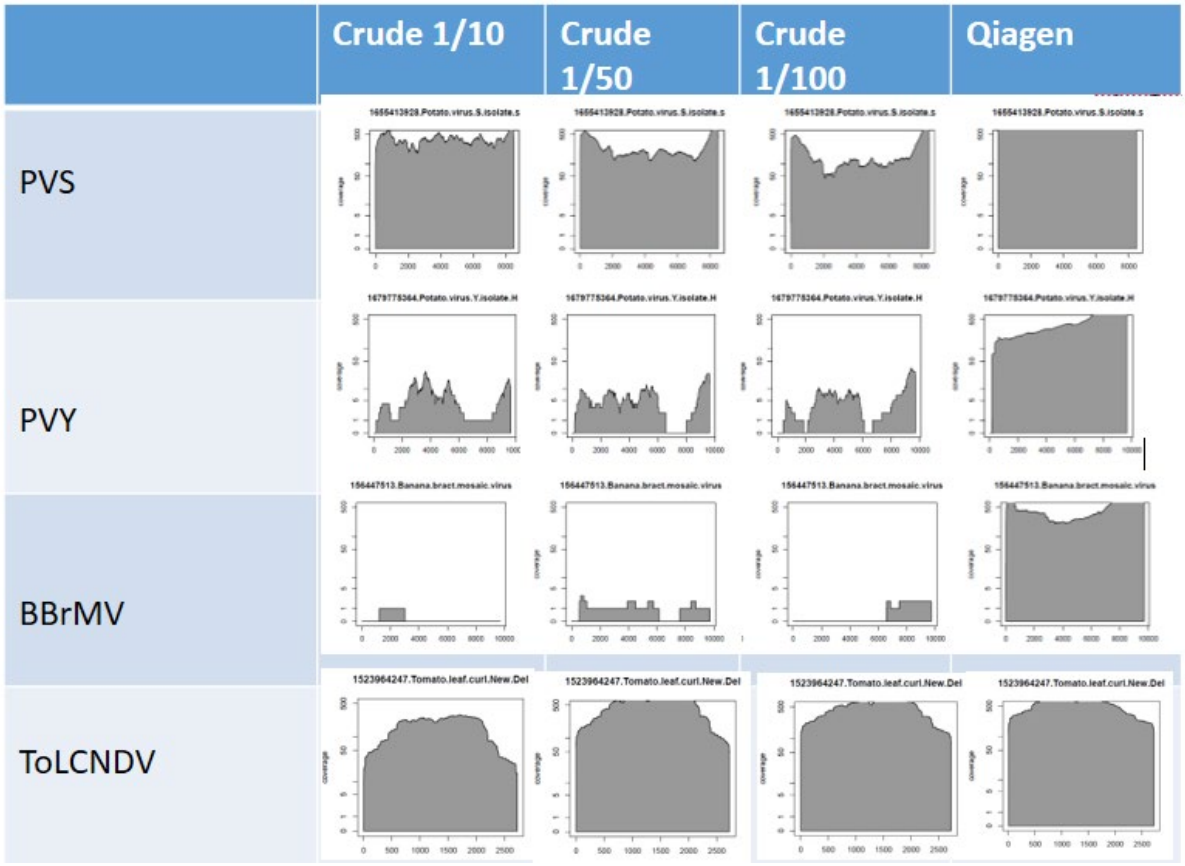


Figure 4: Coverage of the main viruses (PVS, PVY, BBrMV and ToLCNDV) in banana (BBrMV) or in potato (PVS, PVY and ToLCNDV) from a crude extract 1/10, 1/50 or 1/50 and from a standard RNA extraction (RNeasy (Qiagen)).

Conclusion: The initial dilution of 1/400, used in the first assay, was based on current practice of PCR on crude extract, where polymerase inhibition is observed with a lower dilution. Results observed in this assay suggest that these inhibitors have no effect on the reverse-transcriptase and are washed off during the DNA clean-up (AMPure XP Beckman) post RT. From the three dilutions used, the 1/10 gave the best and allowed the detection of multiple viruses. PVS was the most represented virus present in read number and as seen on the first assay, it was also responsible to cross contamination to other sample (some reads were recovered from the banana sample). Lastly, the throughput obtained from RNA extraction preparation were much higher than the one obtained from the crude extract. There is a possibility that the library prepared from crude extract is carrying some contaminant that would compete for the pore or that its DNA quantification is overestimated by fluorometric quantification.

Third assay on crude extract preparation

The third assay (run 3 and 4) compared the sequencing of the RNA extracted and the crude extract on two different flow-cells to eliminate a potential competition effect.

The first flowcell (run 3) was loaded with the library prepared from crude extract preparation (1/10 dilution) while the second flowcell was loaded with the library made from the Qiagen RNeasy preparation.
From the crude extract, cDNA concentrations after the ligation of the barcodes were very low (0.13 ng/µl for *Prunus*, 1,3 ng/µl for potato, 0,6 ng/µl for banana and 0,4 ng/µl for sweet potato. The maximum of the final library obtained for each sample (25 µl/sample) was loaded (no quantification done).
From the RNA extracted, the concentration of cDNA loaded was much higher except for the *Prunus* sample. In total, almost 100 fmols was loaded, normalized between the samples at nearly 15 fmols per sample except for the *Prunus* (4-5 fmols loaded).
The global yield obtained from the crude extract preparation was about half of the preparation from the RNA extraction preparation (3,39 M reads /5,78 Gb and 5,97 M reads /10,39 Gb).
As expected, the throughput from the *Prunus* sample was much lower in both flowcells but comparable in the other samples (Figure 5).

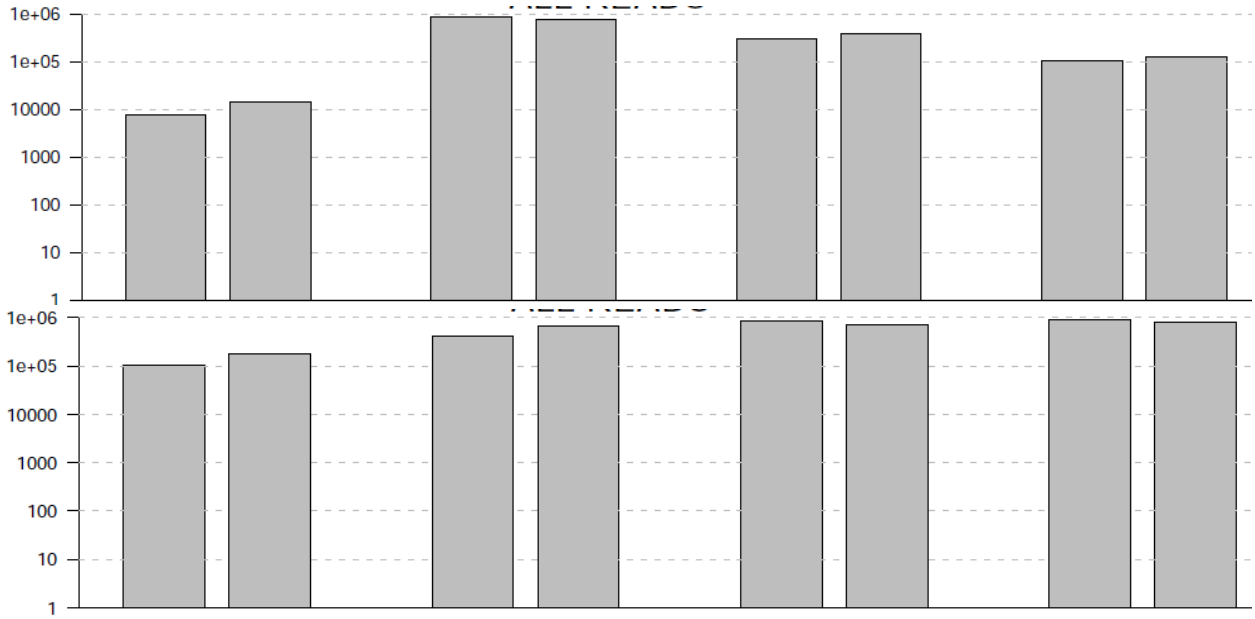


Figure 5: representation of the yield of all samples obtained from the crude sample, run 3 (**top**) and the RNA extraction run 4 (**bottom**). The order of the samples is the same with two replicates for all the samples, from left to right, *Prunus*, potato, banana and sweet potato

Virus detection in the library from RNA extraction was: CVA detected in *Prunus*, PVS, PVY, and multiple members of the genus Begomovirus for the potato; BBrMV and BSV in banana and Sweet potato leaf curl Lanzarote virus (SPLCLV, Begomovirus), sweet potato leaf curl virus (SPLCV, Begomovirus) in sweet potato. The difference between the two libraries was the absence of detection of PVY in the crude extract potato and of CVA in the crude extract of *Prunus*.

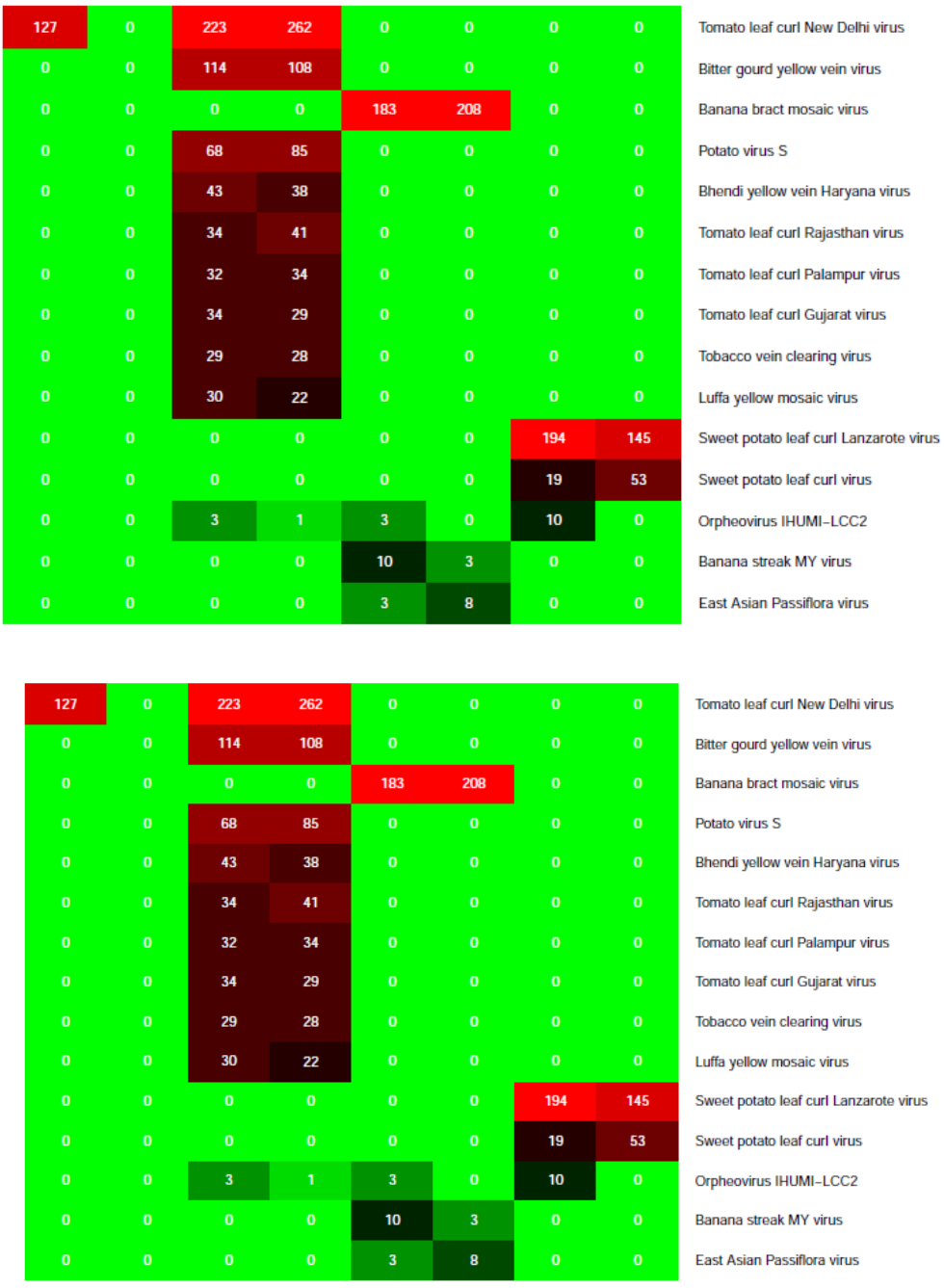


Figure 6: Heat map generated after Kraken2 (v.2.0.7) taxonomic assignation of the sequenced obtained from the crude extract preparation (run 3-top) and the RNA extract preparation (RNeasy Qiagen, run 4- bottom). Samples were loaded in the same order with the *Prunus* in the first two columns, the potato in column 3 and 4, the banana in column 5 and 6, and the sweet potato in the last two columns. Numbers correspond the number of reads assigned to this virus normalized in count per million.

Conclusion: This run harvested the highest throughput from crude extract. The viruses detected were similar with the run obtained from the RNA extracts, except for PVY. Despite those good results, the quantity of cDNA obtained from the crude extract samples after the library preparation was extremely low and this could explain the low yield. Library preparation with a PCR step could solve this problem.

Fourth assay on crude extract preparation

The fourth assay (run 5) was a test of a new library, quicker than the cDNA direct used previously and that included a PCR amplification that was selected to improve the yield obtained from the crude extract preparation.

The library preparation followed the manufacturer protocol. Before loading the library onto the MinION, the Bioanalyzer profile revealed that the length distribution of the DNA was extremely short when compared with the previous run with a pick at around 900 bp (Figure 7). By comparison, using the ONT Direct cDNA Native Barcoding, the average length size measured with the Bioanalyzer was ~2500 bp for the run 3 (Qiagen RNA). As a result, the yield was low in read number (4,5 M reads) but very low in base sequenced (1,6 Gb). The previous runs from the third assay yielded 2,7 M reads and 4,3 Gb from the crude extract. Almost all the sequences obtained were from the RNA extracted samples with only about 10,000 reads harvested from the crude extract preparation. The average size of the reads obtained was around 300 bp.

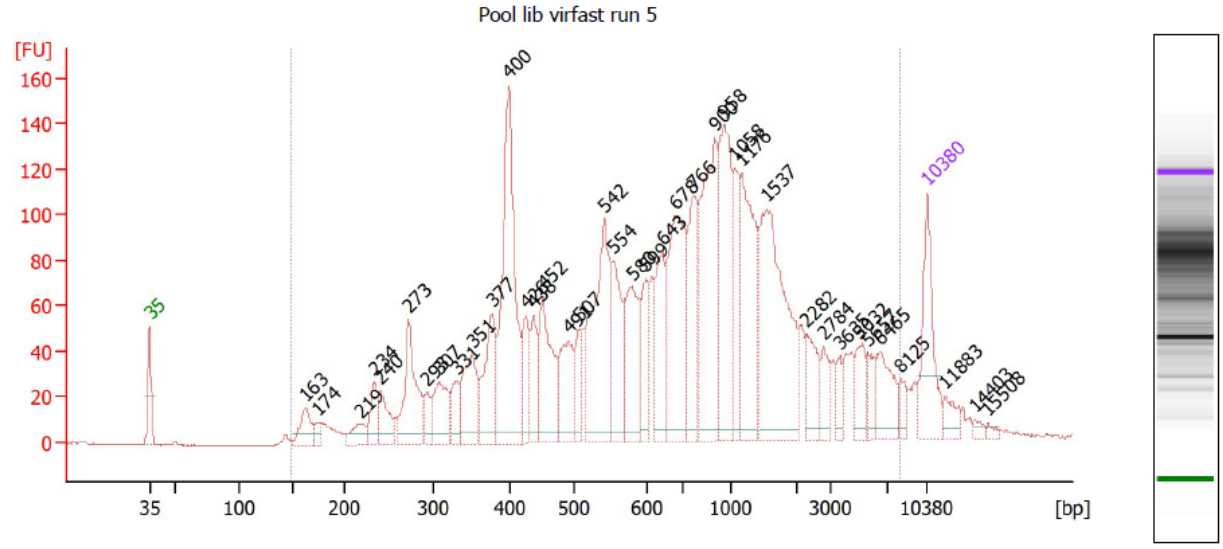
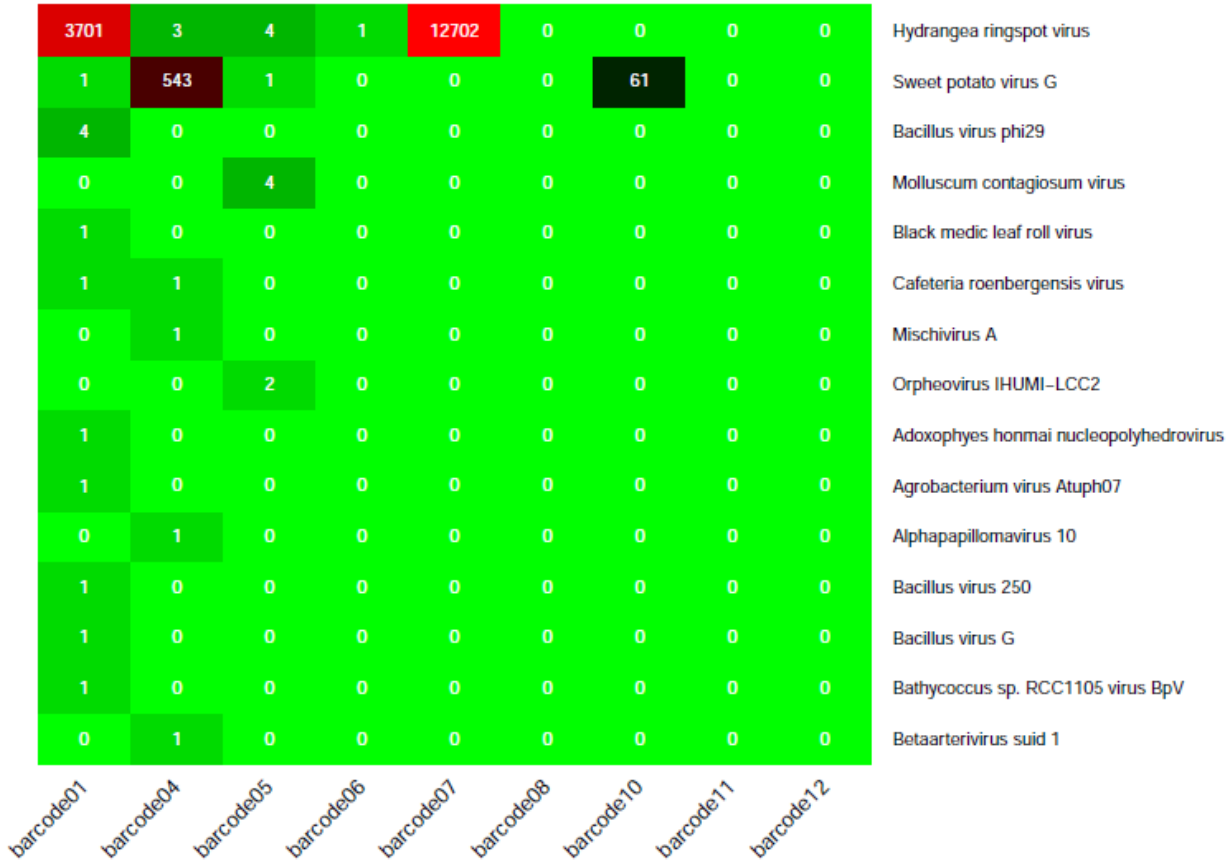


Figure 7: cDNA size profile of the library preparation of the fourth assay (Bioanalyzer)

The *Prunus* samples failed in both preparations and the Oxalis failed in the RNA extraction preparation. Those samples were not loaded. No viruses were detected in the tomato sample, but further lab analyses revealed that the plant was not infected (inoculation with PepMV failed). The Yam viruses were not detected in any preparation. The RNA preparation successfully detected the virus in the Hydrangea (HRSV) and the viruses of sweet potato

(SPVG). Despite the very low yield, the sample from the crude extract preparation were also successful at detecting (with only a few reads) the same virus.



Due to the low yield and sequence length, several options were considered to improve the output:

- the low yield from the crude extract could be explained by an overestimation of the quantification of cDNA, and/or an effect of plant PCR inhibitor. Those inhibitors are removed from the Direct cDNA Native barcoding protocol by an extra DNA clean step using AmPure purification after the initial RT. This purification step could be added, post RT.

- The read length could be affected by the PCR cycle condition (number of cycle and extension time).

However, due to the lack of time, it was decided to revert to the cDNA Native barcoding protocol.

The fifth assay (run 7) RNA extracted and the crude extract on dehydrated samples.

The fifth assay (run 7) compared the sequencing of the RNA extracted and the crude extract on the same flow-cells. For this run, the original library ONT Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104) was used.

The manufacturer instructions were followed and the Bioanalyzer profiles were compared. The crude extract and RNA extract preparation were loaded on the same flowcell and more cDNA from the crude extract were loaded to compensate for a potential overestimation from those samples as observed before. The RNA extract preparations were normalized to 5 fmols while the samples obtained from the crude extract were normalized to 10 fmols, when possible (tomato sample frozen and dehydrated). For the frozen samples, 8,3 fmols (hydrangea) and 9,7 fmols (sweet potato) were loaded but only 3 fmols (hydrangea) and 4 fmols (sweet potato) could be loaded from the dehydrated samples.

In total the run yielded almost 3 M reads and 4,5 Gb. The normalization between the crude and the RNA extract preparations was excellent with about +/- 1 Gb obtained for each of the four conditions (RNA/crude, frozen/dehydrated) and all the samples harvested between 200 and 500 Mb except for the hydrangea (barcode07) and the tomato (barcode09) crude extract frozen with 55 and 980 Mb respectively (Figure 9).

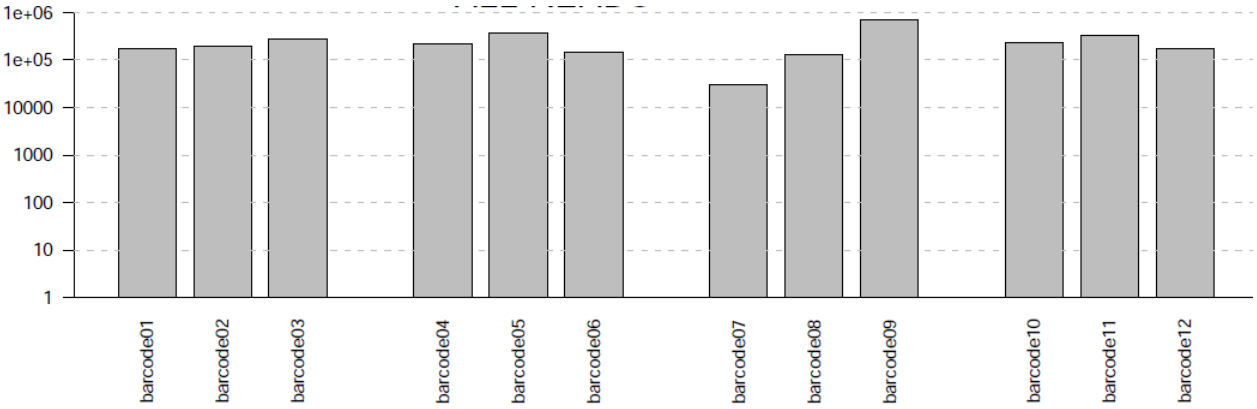
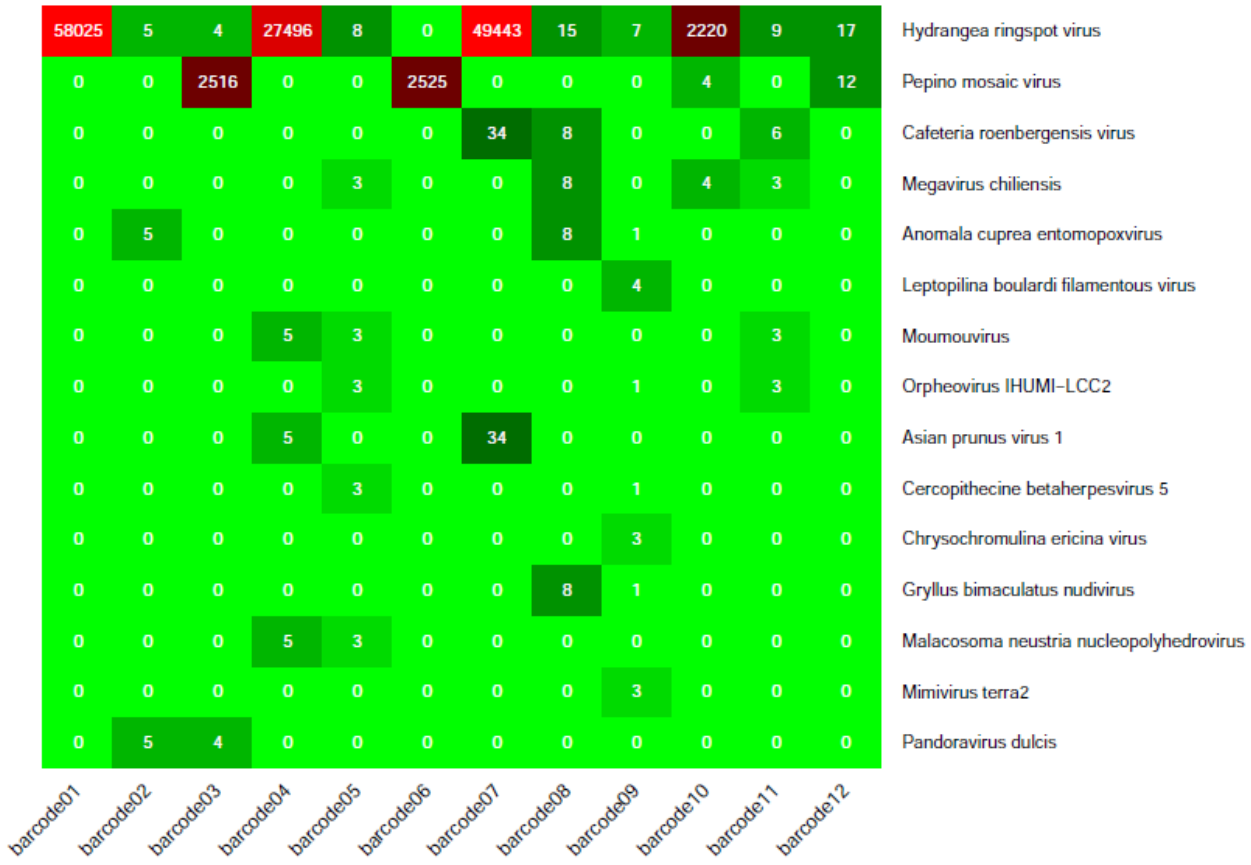


Figure 9: representation of the yield of all samples obtained from the fifth assay's samples. Samples follow the order of hydrangea, sweet potato and tomato using RNA extraction (barcode 1 to 6) or crude extract (barcode 7 to 12). Samples were frozen (barcode 1 o 3 and 7 to 9) or dehydrated (sample 4 to 6 and 10 to 12).

The detection of HRSV in Hydrangea was successful under all the four conditions, yet the virus was also detected in the other samples (cross contamination or barcode jumping in the library process). No viruses were detected in the sweet potato sample. Finally, the tomato virus PepMV was easily found from the RNA extract samples, but barely detected from the crude extract (>1,000 read per million detected from the RNA extracts and only 12 reads per million in the crude extract dehydrated).



Conclusion: This run was successful at normalizing the throughput across all the samples and conditions, by increasing the amount of DNA obtained from the crude extract samples at the expense of the RNA extracted samples. This assay also showed that the detection of plant viruses was possible from crude extract of dehydrated samples. This parameter should help the homogenisation of plant tissue (in case virus distribution is not systemic) and the preservation and shipping of tissue to collaborating partners overseas. However, working with dehydrated samples is an extra hurdle for virus detection.

Sixth assay Test performance study (TPS)

The sixth assay (run 8) was the TPS involving six additional international laboratories. All the samples were anonymised and randomised.

The results obtained were highly variable from the different laboratories. All of them used two flowcells with the exceptions of the L02 and L05 who loaded all the samples onto a single

flowcell. However, the laboratory L02 has no results (> 1000 reads in total) and they were not included in the following sections.
The overall throughput of the different laboratories is presented in the Figure 11. In all cases, the yield is two to seven-fold better from the RNA extract preparation than from the crude extract, with one exception being L06 where the yield is better from the crude extract.

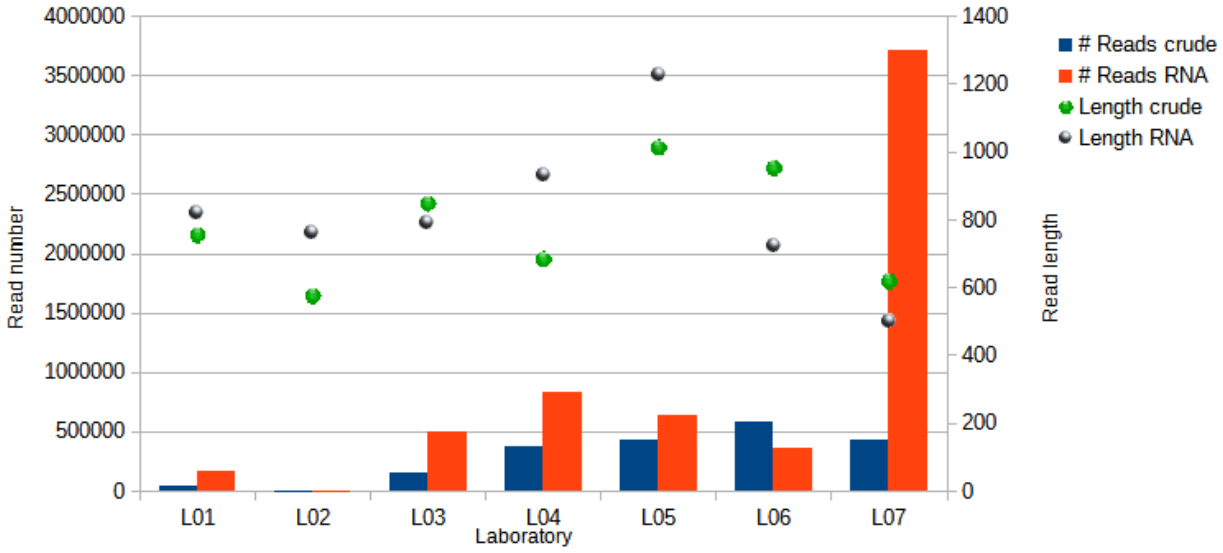


Figure 11: Overview of the sequencing yield across the seven laboratories participating to the TPS. Yield obtained from the crude extract is presented in blue bars, the yield obtained from the RNA extract is presented in orange bar. Read length is presented in green (crude extract preparation) and grey (RNA extract preparation) dots. L02 and L05 loaded the cDNA prepared from the crude extract and from the RNA extraction onto the same library and flowcell.

For the viral detection, mapping was used to calculate the number of nucleotides mapping the virus targeted. A threshold of 1,000 nucleotides mapped onto the genome of the selected viruses was chosen to count a positive detection for this virus. The laboratory L02 was not included in the analyses. From the sample of hydrangea, 89% of the RNA extracted samples and 50% of the crude extract samples were positive for the virus HRSV. From the sweet potato, 89% RNA extracted samples and 28% of the crude extraction were positive for the detection of the virus SPVG. Lastly for the tomato samples, 94% of the RNA extracted sample and 72% of the crude extract samples were positive for the PepMV. It should be mentioned that L07 had a power incident while sequencing the crude extract library and that problem could explain the lack of detection of any viruses from that run.

Table 5: Overview of the virus detection obtained by the different laboratories. The number represent the number of sample(s) where 1,000 nucleotides or more mapped the correct virus out of three replicates.

	Hydrangea crude	*Hydrangea RNA*	*Sweet potato crude*	*Sweet potato RNA*	*Tomato crude*	*Tomato RNA*	*RNA*	*CRUDE*
L01	2	2	0	3	1	2	78%	33%
L03	3	3	0	3	3	3	100%	67%
L04	1	3	1	3*	3	3*	100%	56%
L05	3	3	3	3	3*	3*	100%	100%
L06	0	2	1	1	3	3	67%	44%
L07	0**	3	0**	3*	0**	3*	100%	0%
	50%	89%	28%	89%	72%	94%		

* indicates that at least 1,000 nucleotides of another sample mapped onto this virus (contamination)

** indicates that the run was impacted by a power incident

The overall data are difficult to compare as the yield per lab was highly divergent (Figure 8) and the repetition within laboratories were variable too. The number of nucleotides mapping the control (Figure 12), highlights the better detection with the RNA extraction over the crude extract, but also it shows the variability between hosts and viruses with sweet potato being more difficult to detect.

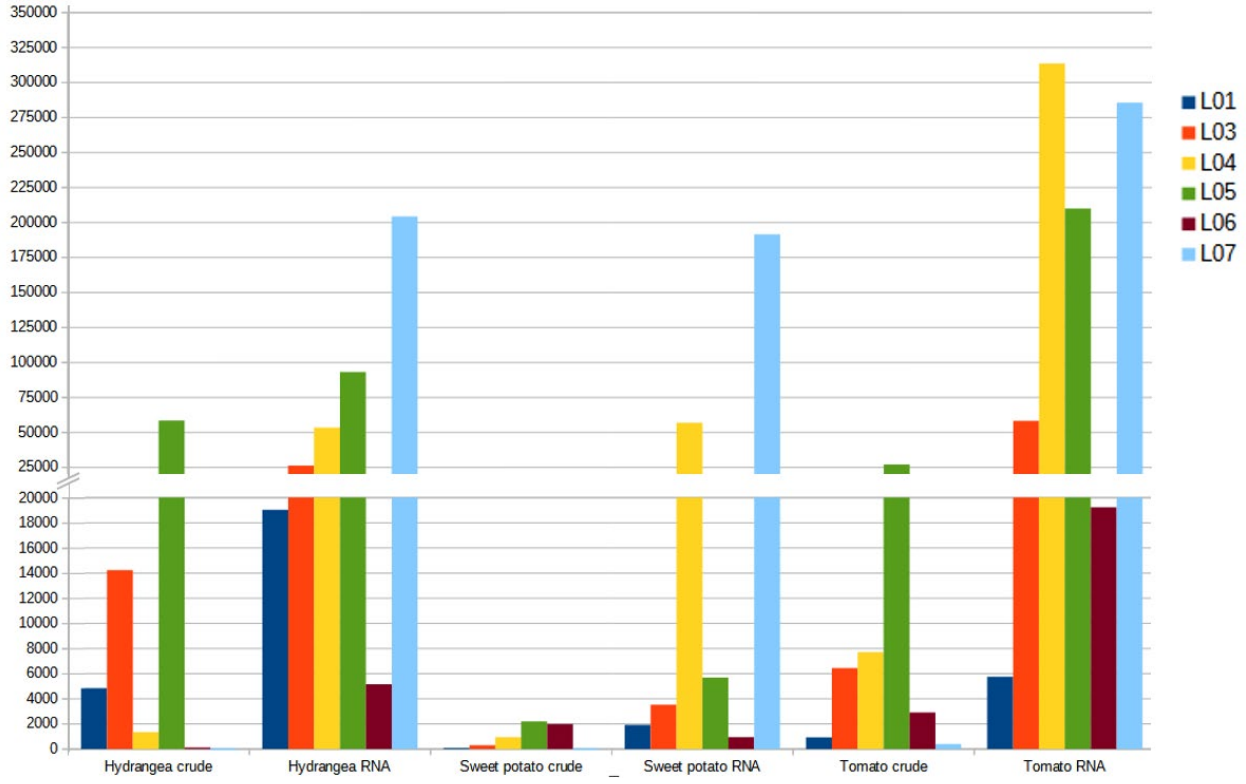


Figure 12: Average number of nucleotides of the three replicates mapping to the viral reference sequence for each of the sample (Hydrangea ringspot virus in Hydrangea, Sweet potato virus G in sweet potato and pepino mosaic virus in tomato). Colour bars represent the results of each of the participating laboratories (L02 was not included).

This assay has shown a great variability of response between laboratories but overall, a good virus detection wen using the RNA extraction method (91% detection). Sequencing from crude extract is an exciting prospect for the future in order to reduce the time of diagnostic and the point of care. The results show that despite working with six laboratories that had not used the protocol before, 50% of the detection was made (up to 62.5% if we remove the L07 crude extract run), albeit that the threshold used was very low.
For now, these results clearly show that the method is not ready to be used for diagnostics, but the detection of plant virus from crude extract direct cDNA sequencing with ONT is possible.

Results on sample preparation comparison

General summary reported by NanoStat v1.1.2 (Table 1) showed better sequencing output when using Direct RNA for library preparation. Using this approach higher number of reads with higher mean/median read length and mean read quality was achieved.

Table 1: General summary reported by NanoStat for nanopore sequencing using directRNA of totRNA and cDNA-PCR of totRNA approaches.

General summary:	nanopore sequencing directRNA of totRNA	nanopore sequencing cDNA-PCR of totRNA
Mean read length:	989.4	955.7
Mean read quality:	10.1	7.8
Median read length:	1,280.00	414
Median read quality:	10.6	8.1
Number of reads:	1,085,202	198,916
Read length N50:	1,303	1,889
Total bases:	1,073,653,862	190,112,771

The comparative analysis of Direct RNA versus cDNA-PCR libraries for sequencing on MinION showed some differences regarding virus detection. Using direct RNA sequencing approach we were not able to detect one virus (tomato mosaic virus, ToMV) out of 5 in the sample, however, this virus was present in the sample in very low titer. Furthermore, the %age of mapped reads to the specific genome consensuses of viral/viroid reads was higher for all investigated viruses/viroid when using cDNA-PCR sequencing approach (Table 2). When investigating the average MinION reads identity against the Illumina generated references, there was no significant differences among the compared approaches. In summary, the results indicated better potential of virus/viroid detection when using cDNA-PCR library preparation approach (compared to direct RNA sequencing approach), due to the higher relative number of viral reads derived using this approach (Table 2).

Table 2: The % of mapped reads and average mapped MinION reads identity against Illumina-generated reference sequences using directRNA of totRNA and cDNA-PCR of totRNA nanopore sequencing approaches.

Virus/viroid (Baltimore classification)	nanopore sequencing directRNA of totRNA		nanopore sequencing cDNA-PCR of totRNA	
	% of mapped reads	Average MinION reads identity	% of mapped reads	Average MinION reads identity
Tomato yellow leaf curl virus - TYLCV (ssDNA)	0.002	89.63	0.0486	89.38
Tomato chlorosis virus - ToCV (ssRNA+)	0.0007	90.53	0.0773	89.48
Pepino mosaic virus (Chile 2 strain) - PepMV-CH (ssRNA+)	1.1528	86.05	27.1932	87.07
Pepino mosaic virus (EU strain) - PepMVEU (ssRNA+)	0.5638	85.45	15.7794	86.73
Tomato mosaic virus - ToMV (ssRNA+)	0	/	0.0847	87.38
Southern tomato virus - STV (dsRNA)	0.0001	90.82	0.0006	90.43
Columnea latent viroid - CLVd (viroid)	0.0005	88.65	0.004	90.85

Validation and Implementation of HTS technology for routine testing in the Sidney Laboratory Diagnostic Unit (CFIA)

A test method and associated standard operating procedures for the routine use of NGS testing in the CFIA Sidney Laboratory's Diagnostic Unit have been written and reviewed internally. The method has not yet been approved since work that arose from the review was delayed due to the lab shut-down during the COVID-19 pandemic. These final points are being addressed now and the method is due to be approved in early 2022. A workflow has been drafted and will be completed once the method is finalized. The Diagnostic Unit hopes to use NGS as a routine test method for a portion of its testing starting in April 2022.

Appendix 3 Summary of the information obtained in the frame of the two surveys carried out on (i) high throughput sequencing technologies for plant health diagnostics and (ii) on-site test use

Evaluating the use of HTS technologies by NPPOs and diagnostic laboratories

During 2019 the opportunity was taken to partner with the EU Valitest project to carry out a joint survey of National plant health laboratories and the results of this survey relating to NGS technology use are presented below.

26 labs across 20 countries responded to the survey. Only four of the labs reported using HTS for routine diagnostics with two others using it for research purposes. Of the four using HTS routinely, all reported using it for "Resolving unknown etiology of symptoms" either instead of, or after classical testing. All four reported using the Illumina platform with 2 also using Oxford Nanopore and one Ion Torrent. All four identified interpreting the results as the most challenging aspect of the procedure with three also identifying data analysis. Two of the four labs stated that they were planning to request ISO accreditation. Lack of validation and cost were identified as the bottle necks to greater adoption.

Turning to the 22 labs which didn't state they were currently using HTS for diagnostics 73% named "cost" as a reason for not adopting HTS. Other reasons included "complexity" (50%) and "a lack of expertise" (50%). Only one lab stated it was "not relevant for their organization". These labs were then asked what could be changed to increase their likelihood of using HTS. Figure 13 summarizes their responses. More than half of the labs rated every question as either essential or important with Training (59%) and lower price (50%) being rated essential by the most labs. Faster turnaround (14%) and frameworks and guidelines for validation (23% each) were considered essential by the least number of labs although 45% of the labs then rated these three factors as important.

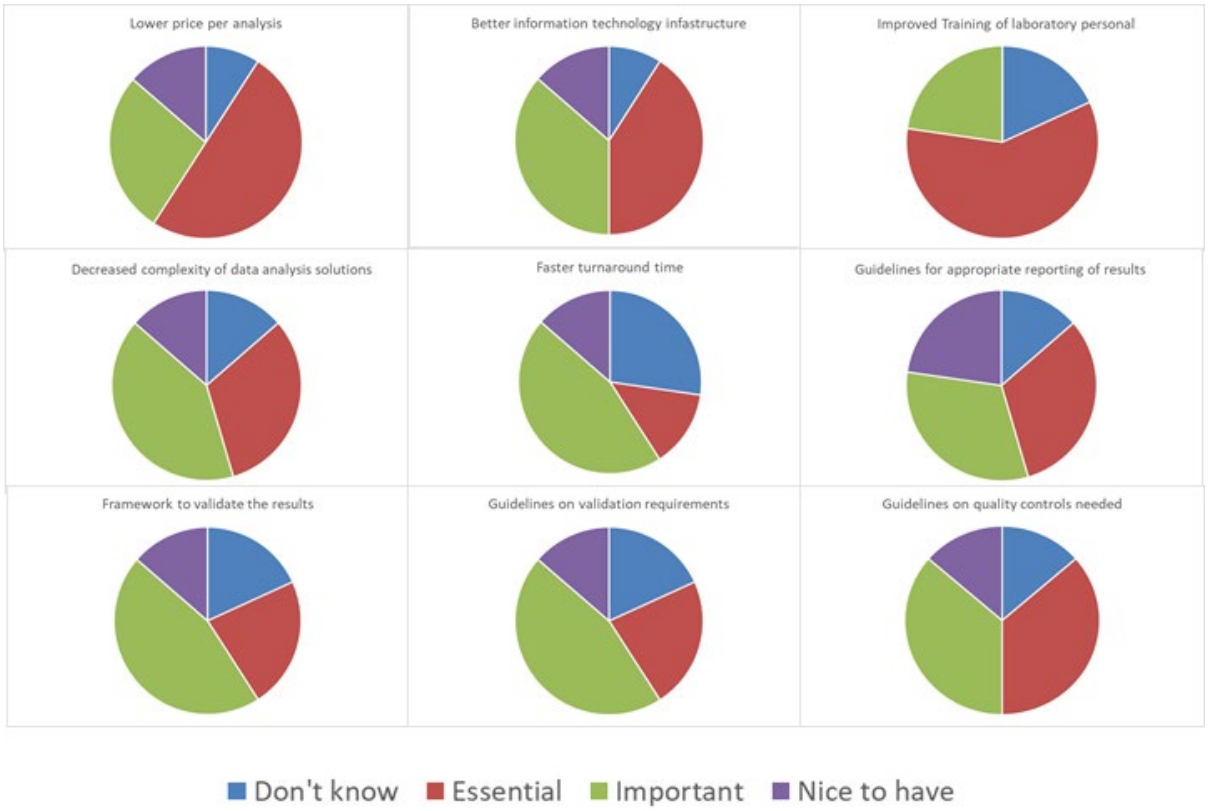


Figure 13. Questions on actions why might improve uptake of HTS.

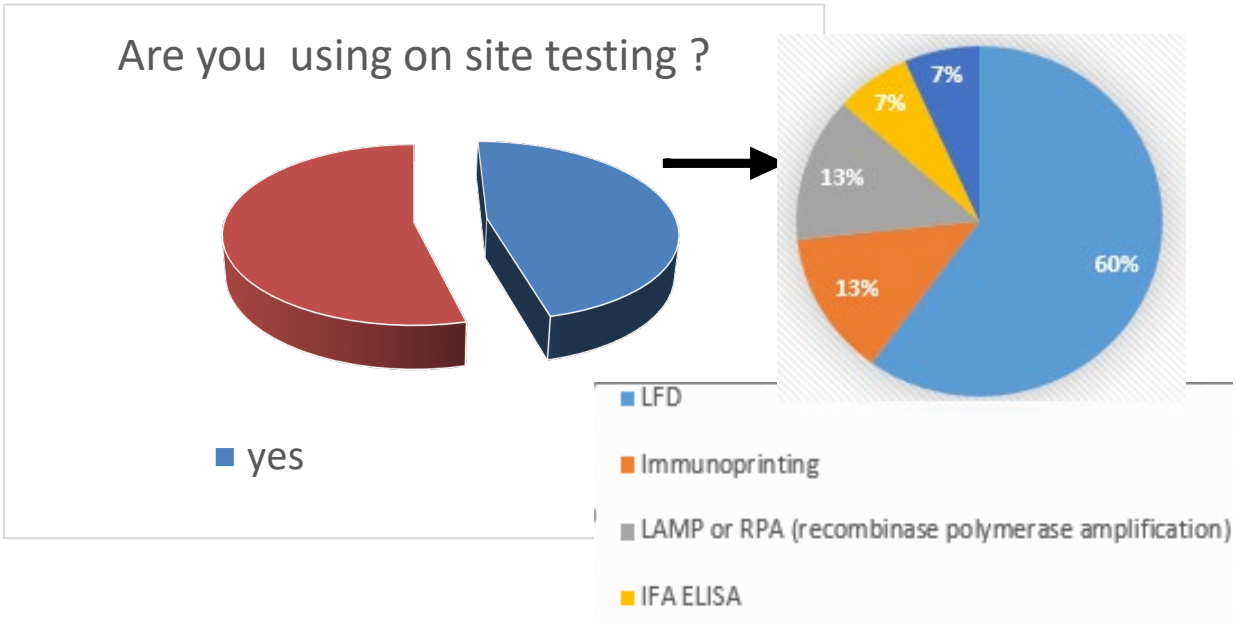
Summary of findings from survey**:** Only 15% of the laboratories questioned where actually using HTS for diagnostic work and this was for looking at diseases of unknown etiology and was used both after and instead of classical tests. Of the remaining labs, almost all were interested in using HTS and the reasons they gave for not adopting it were cost (73%), complexity (50%) and lack of expertise (50%). They were then asked what would help them to adopt and “better staff training” and “lower costs” were identified as key factors.

2. Opportunities for on-site surveillance on plant viruses

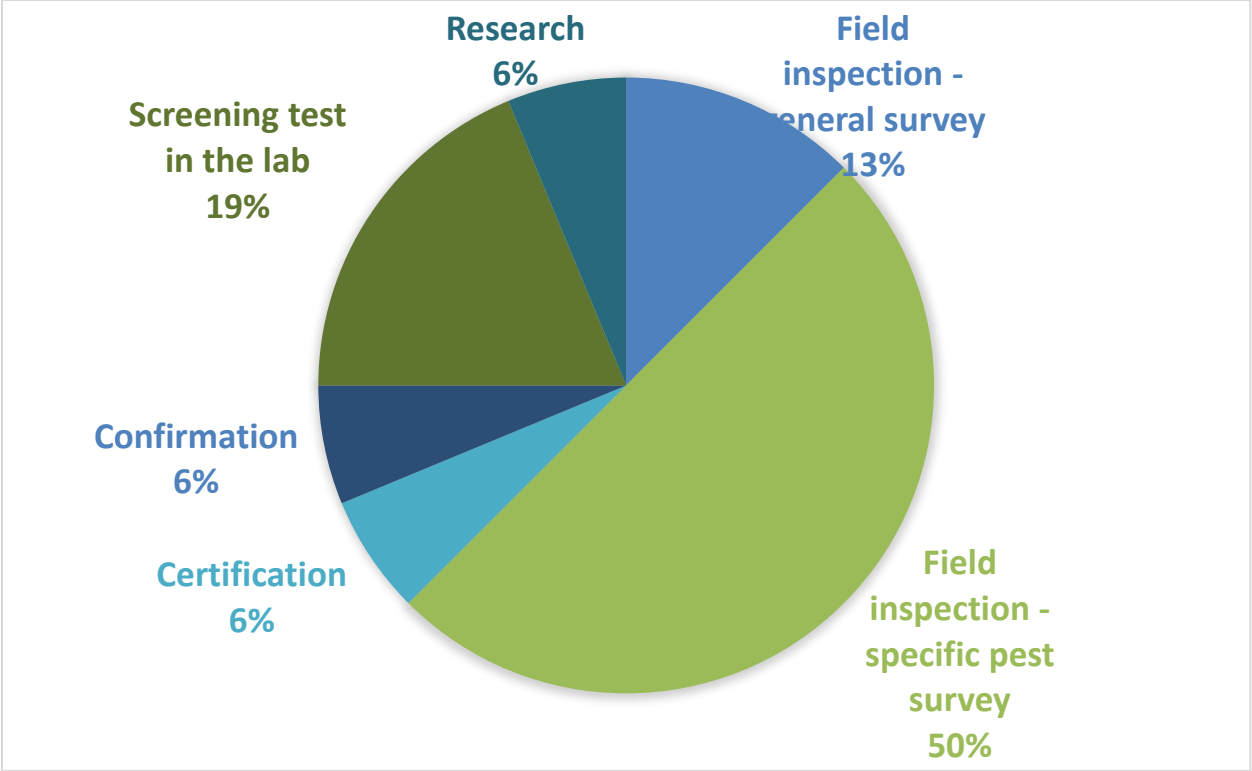
During the first project year, there was interaction with the project partners of a parallel H2020 project, VALITEST. Within VALITEST, a survey for the use of on-site diagnostic methods was also aimed for. A joint effort was set up, not to duplicate the efforts within two different projects.

Results of the survey:

- 26 labs replied
- Representing 20 countries
- Only 12 already use on-site testing



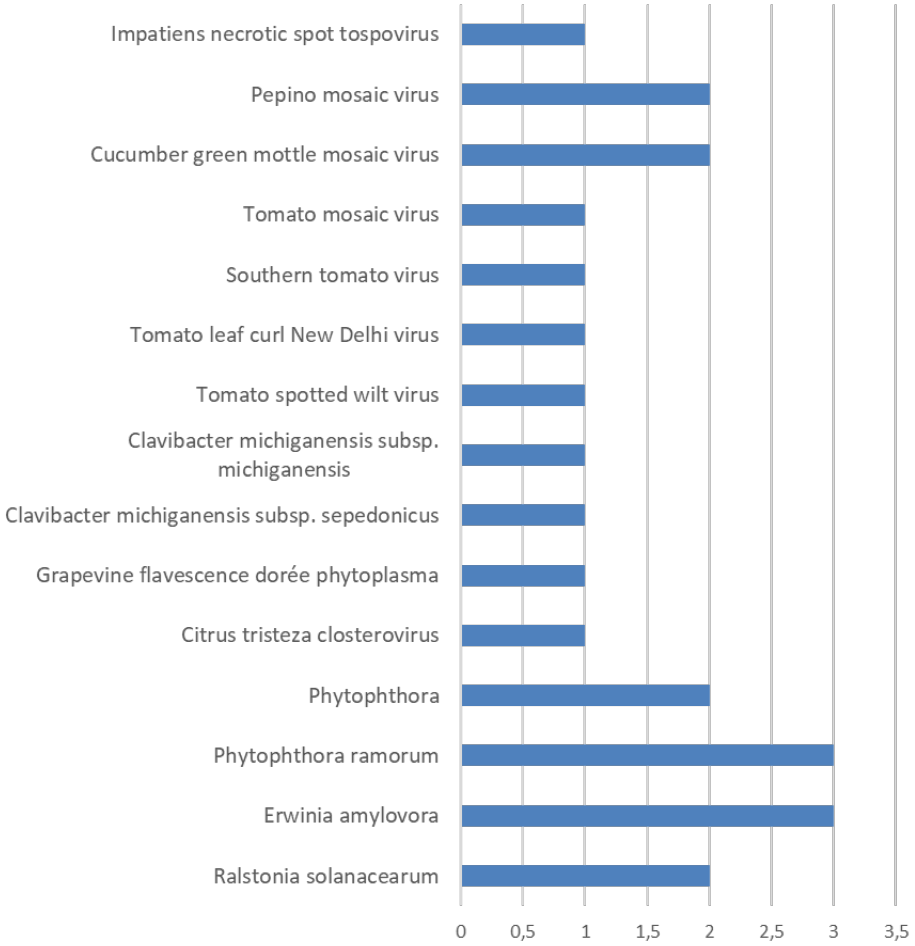
In which context are the labs using on-site testing?



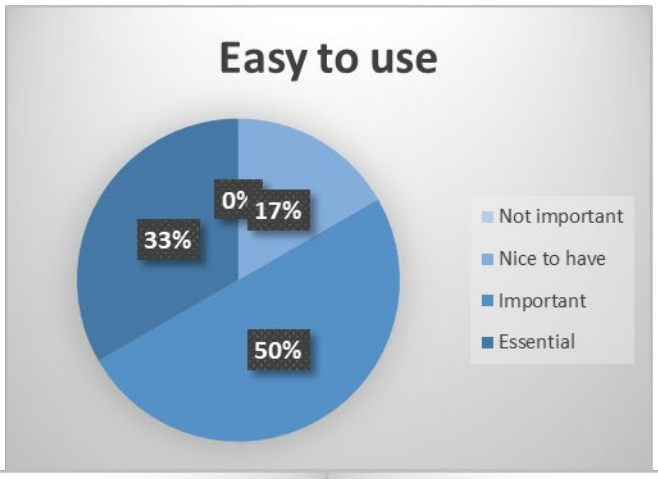
What are the total number of samples tested /year?

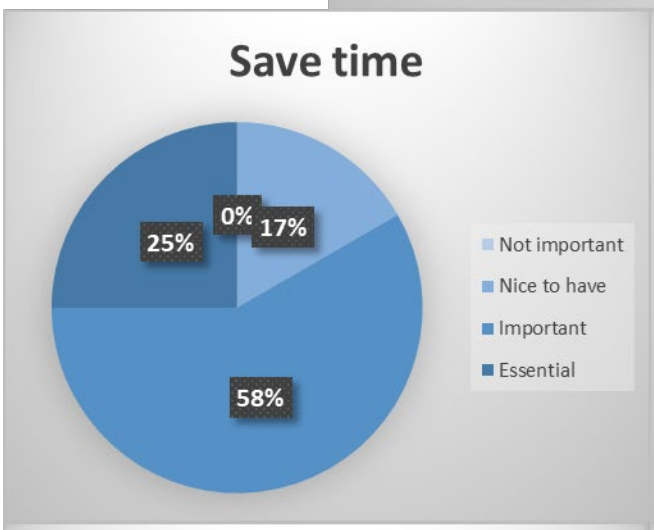
Country	Total number of tests	Specifications
Austria	100	Screening in the lab
Belgium	1500	LFD in the lab
Bulgaria	260	Specific pest survey
Finland	80	Specific pest survey (50); Screening tests in the lab 30
Greece	2500	General surveillance 1500; specific pest survey 1000)
Italy	300	certification
RE	100	Specific pest survey
Slovenia	102	LAMP 100; Specific pest survey 2
Spain	25	Specific pest survey
Switserland	30	general surveillance 20; specific pest survey 10

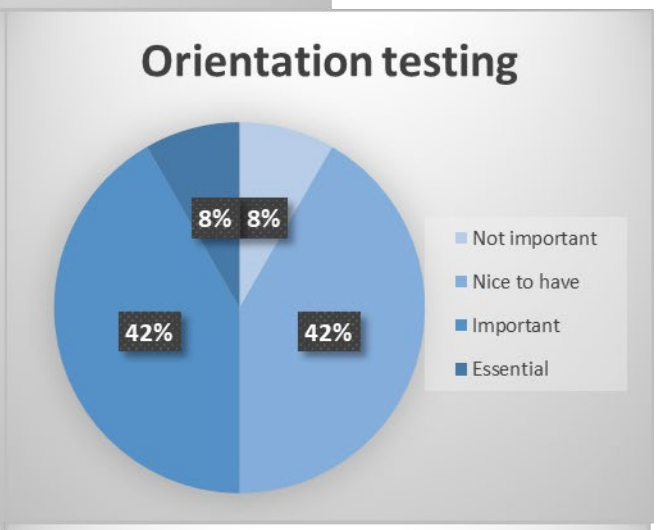
What are the most targeted pests for on-site testing?

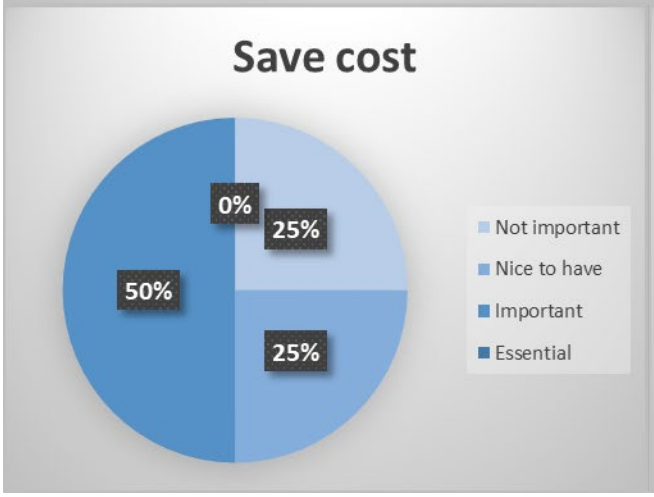


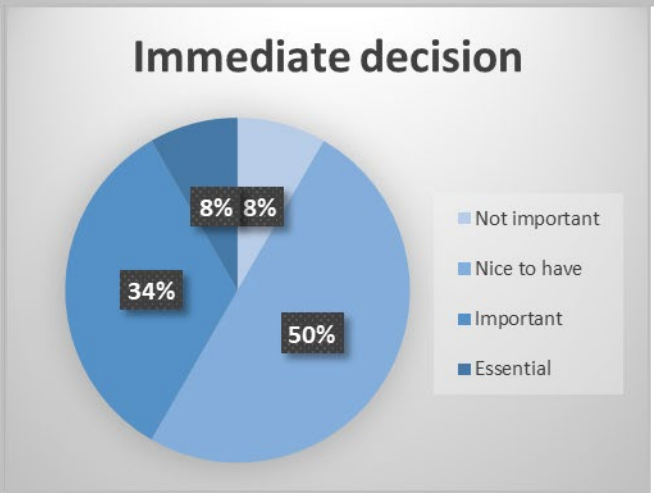
How were the benefits for on-site testing estimated ?



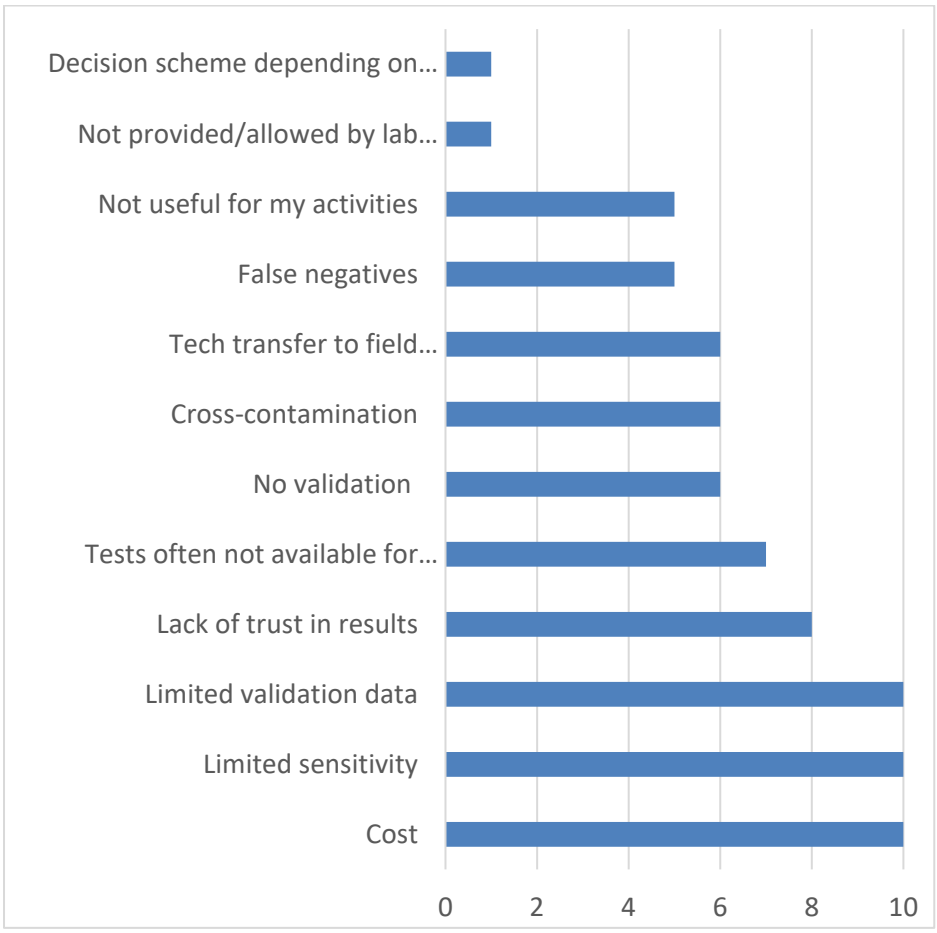




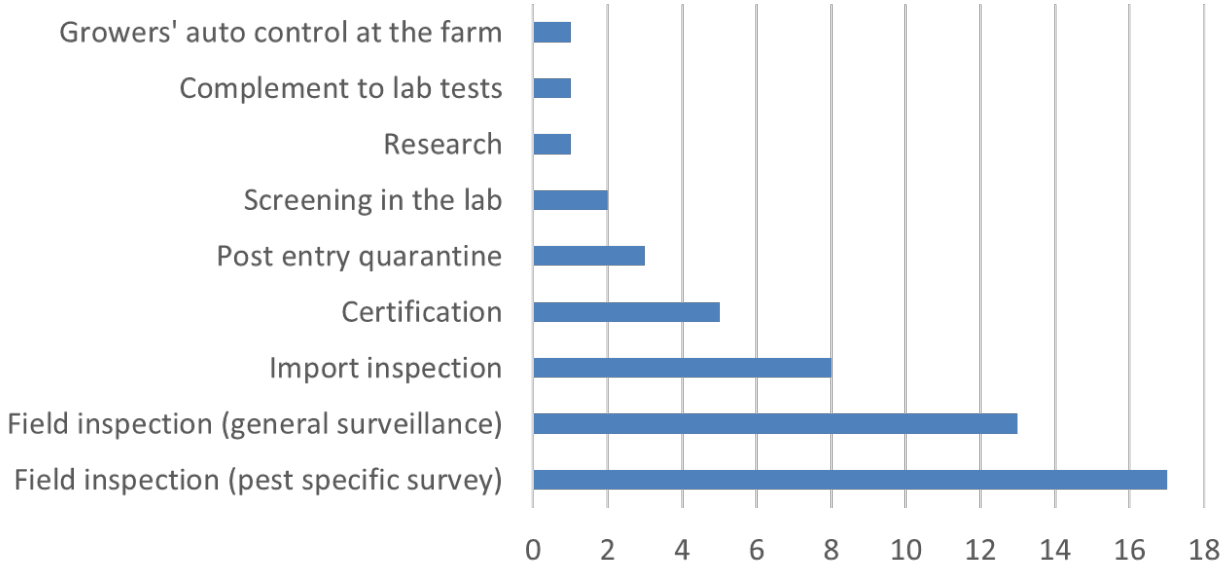




Which are the current limitations to use on-site testing?



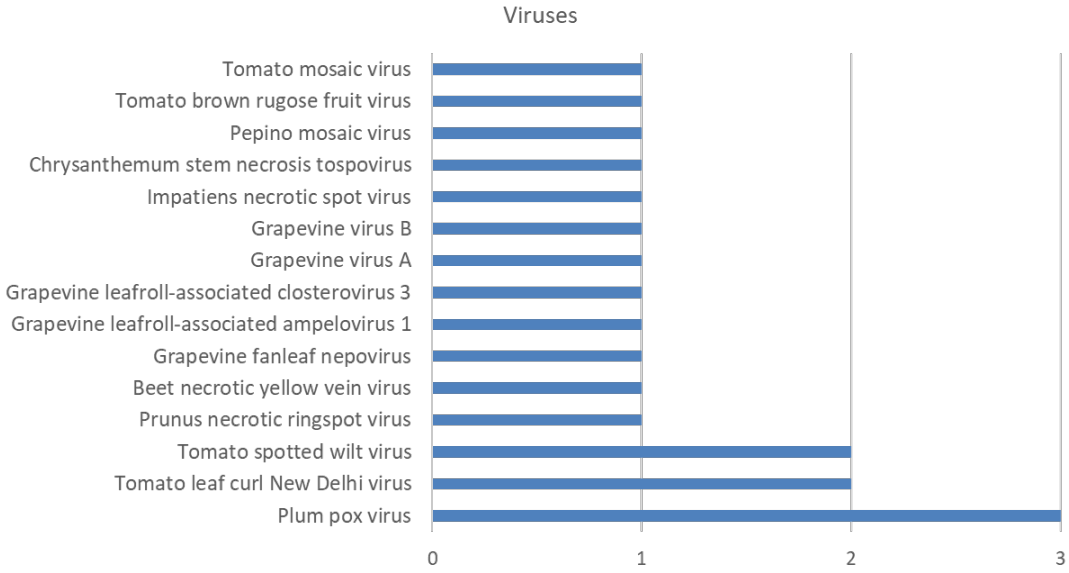
In which context do you see opportunities for future use of on-site testing?

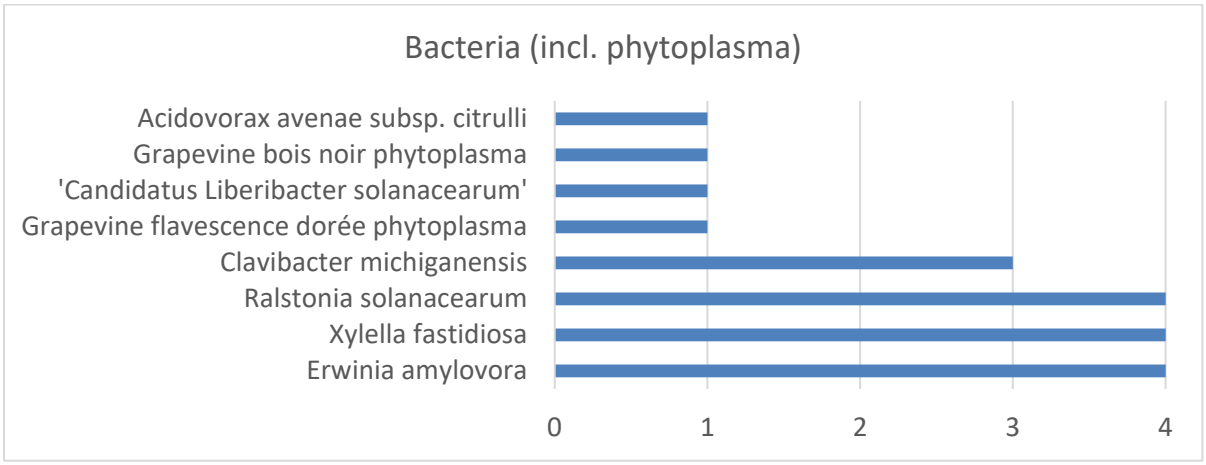


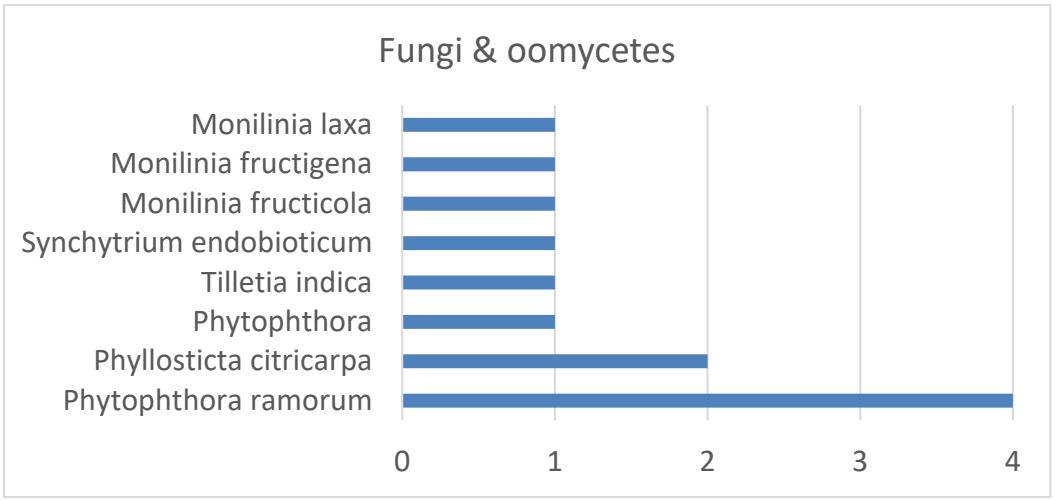
For which targets are on-site testing kits needed the most?

For nematodes, only *Globodera rostochiensis* was mentioned once.

For the other organism groups:







Summary of the output from the survey: In total, 27 laboratories from 20 EU countries replied to the survey. Twelve laboratories were already using on-site testing with an estimated volume ranging from 25 to 2,500 tests per year. The main context for using on-site testing kits was field inspection (63%), followed by screening test in the laboratory (19%). This means that these kits allowing immediate response have a significant use in the laboratory itself. This is in line with the stated benefits with the time saving evaluated as the most important factor. After, laboratories underlined the ease of use in order to make orientation testing. On the other side, the ability to take immediate decision was considered as essential of important for less than half of the responders. The tree main limitations for a wider use of the test were the lack of validation data, the limited sensitivity and the cost. The absence of tests for some pests limited also a wider use and participants gave their priorities for further development. The priorities were quite scattered but the main pests per category were plum pox virus for viruses (3 responses), *Phytophtora ramorum* for fungi and oomycetes (4 responses), *Erwinia amylovora*, *Xylella fastidiosa* and *Ralstonia solanacearum* for bacteria (4 responses each). In addition, the laboratories saw opportunities of development mainly for field inspection (n=17) and import inspection (n=8).