ILVO
Flanders Research Institute for Agriculture, Fisheries and Food

Development and Evaluation of an MLSA-ONT Barcoding Workflow for '*Ca.* Phytoplasma' Diagnosis

ILVO Plant Sciences Unit
Funded by FPS Health
METAMINSURV Project (RF 22/6353)

Ellen Everaert, Kris De Jonghe, Dieter Slos, Yoika Foucart, Maaike Heyneman, Annelies Haegeman

ILVO

Diagnosing '*Ca.* Phytoplasma': current challenges

Phloem-restricted, unculturable bacteria, significant crop damage worldwide

- **Low & uneven titers**
  Difficult to detect in plant tissue; seasonal and host-dependent variation
- **Genetic complexity**
  37 ribosomal groups, >150 subgroups
  ~ 50 recognized species - a fraction of all described phytoplasmas
- **Nomenclature inconsistencies**
  Many species, despite having official names, are still referred to by English common names reflecting host symptoms

- **Current diagnostic framework**

  Bertaccini et al. (2022) guidelines:

  - 16S rRNA as primary marker
    - Demarcation: 98.65% sequence identity
    - Fragment length: ≥1500 bp
  - ANI / MLSA with fixed thresholds
  - Sanger sequencing as standard method

  Translating these broad guidelines into standardized routine diagnostics remains difficult

Study objective: To operationalize the species-assignment guidelines for '*Ca.* Phytoplasma' detection and identification through a structured, proof-of-concept diagnostic scheme, that integrates Multi-locus Sequence Analysis (MLSA), Oxford Nanopore Sequencing Technology (ONT), barcoding and read-based quantification.

ILVO

Terminology

Barcoding

Technique to identify species based on ‘short’, standardized genetic sequences that are matched against a reference database.

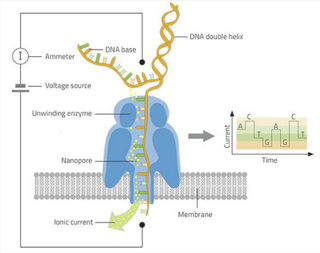
DNA Extraction → Amplification of a gene fragment using universal primers (18S, COI, etc.) → ONT / Sanger sequencing → 1 DNA sequence → *Sphaerechinus granularis*

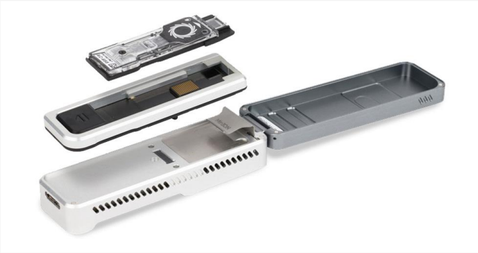
Multi-Locus Sequence Analysis (MLSA)

The simultaneous analysis of multiple specific DNA regions (loci), rather than relying on a single gene.

Oxford Nanopore Sequencing Technology (ONT)

A modern, high-throughput sequencing technology that reads long DNA fragments in real time through microscopic protein pores, allowing multiple barcodes and samples in a single run and enabling detection of mixed infections with pathogens such as ‘*Ca.* Liberibacter’.





Study design — Three phases

PHASE 1

MLSA-ONT barcoding workflow development

- Sample preparation
- Barcode & primer selection
- Reference database construction
- MLSA-ONT workflow development

PHASE 2

Guidelines translation into a structured workflow for '*Ca.* Phytoplasma' detection and identification

- Develop detection & identification scheme
- Integrate MLSA-ONT barcoding with read-based quantification
- Barcode selection guidance
- Reevaluating demarcation criteria

PHASE 3

Evaluation of the MLSA-ONT barcoding workflow

- Sensitivity, specificity, repeatability
- ONT vs Sanger comparison
- PCR multiplexing assessment

ILVO

Phase 1

MLSA-ONT Barcoding Workflow Development

Sample preparation • Barcode & primer selection • Reference database construction • MLSA-ONT workflow development

Initial Steps in MLSA-ONT Workflow Development

Sample panel

- 22 naturally infected samples — 15 '*Ca.* Phytoplasma' species
- Spiked dilution series: '*Ca.* P. asteris' in carrot + FD in grapevine
- Bacterial MOCK community: 9 culturable bacteria for specificity testing (exclusivity)

Barcode selection

Selection of 4 barcodes

16S rRNA	*rplV*
secA	*tufB*

Database reconstruction

- Conservative approach
- Based on almost exclusively officially recognized species (Bertaccini et al. 2022)

Primer development and evaluation

- Literature review
- *In silico* evaluation & new primer design
- Laboratory validation

ILVO

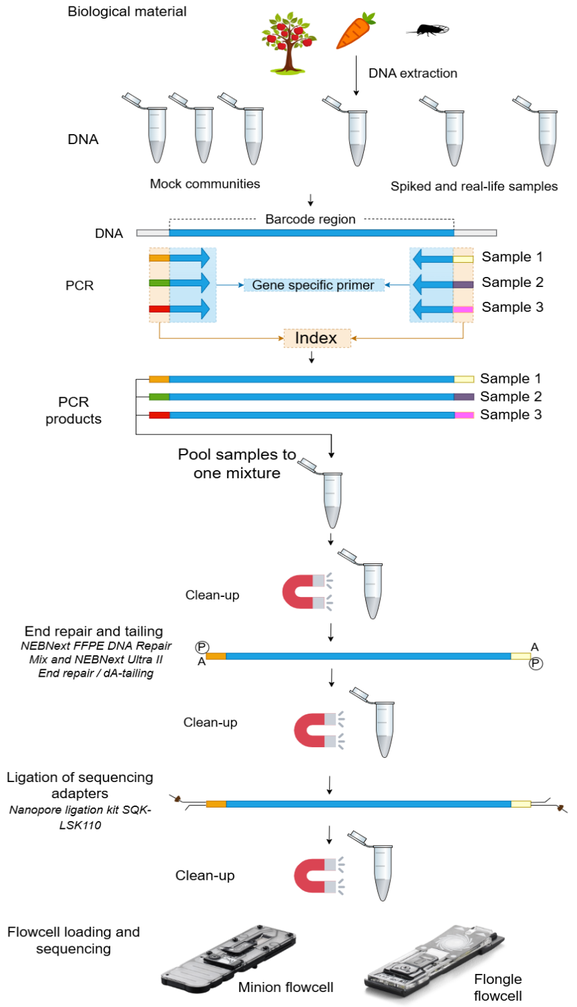
‘*Ca.* Phytoplasma’ primers

Barcode	Primer	Sequence (5'-3')	Amplicon length (bp)	Reference
16S	P1	AAGAGTTTGATCCTGGCTCAGGATT	~1800	Deng & Hiruki (1991); Schneider et al. (1995)
	P7	CGTCCTTCATCGGCTCTT		
	R16F2n	GAAACGACTGCTAAGACTGG	~1200	Gundersen & Lee (1996)
	R16R2	TGACGGGCGGTGTGTACAAACCCCG		Lee et al. (1993)
secA	SecARev3_Adj_F	GGNATGACAGGWACTGCTAAAAC	~360	Hodgetts et al. (2008)
	SecA_ILVO3_R	GCCATATTAGTWGCWATAGT		ILVO (this study)
rplV	rplV_1_F	AAGCACGTTTAGTTGYTGATTT	~430	ILVO (this study)
	rplV_3_1_R	CATTAGGATTACTCTTTTGACCC		
	rplV_3_2_R	CATTAGGATTAGATTTTTGTCCC		
tufB	Tuf340a	GCTCCTGAAGAAARAGAACGTGG	~550	Makarova et al. (2012); EPPO (2021)
	Tuf340b	ACTAAAGAAGAAAAAGAACGTGG		
	Tuf890ra	ACTTGDCCTCTTTCKACTCTACCAGT		
	Tuf890rb	ATTTGTCCTCTTTCWACACGTCCTGT		
	Tuf890rc	ACCATTCCTCTTTCAACACGTCCAGT		

ILVO

MLSA-ONT Barcoding Workflow

Wet-lab protocol



PCR with indexed primers

Pooling of PCR products

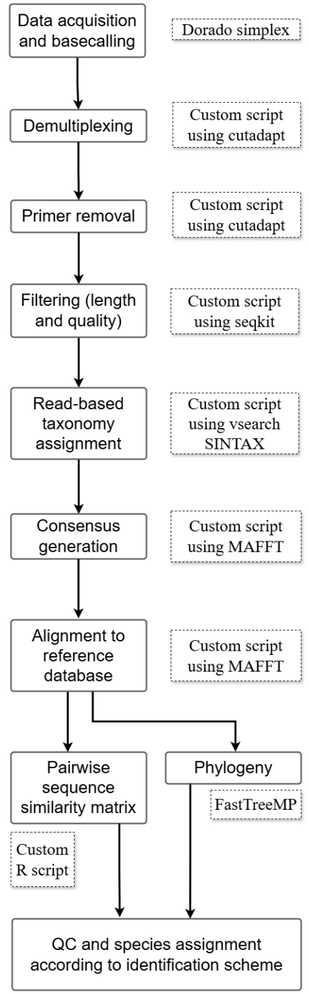
Clean-up with magnetic beads

End repair + dA tailing (NEB kit, instruct. Nanopore)

Ligation of adapters (Nanopore ligation V114 kit)

Flowcell loading + sequencing (R10.4.1 flowcell)

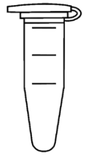
Bio-informatics workflow



MLSA-ONT barcoding workflow

Simplified Wet Lab Protocol

Sample and control selection + DNA extraction

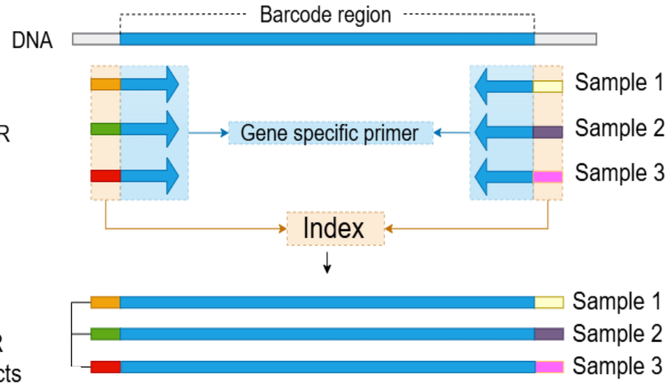


10x

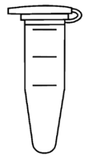
Barcode PCR assays with indexed primers

Barcodes:

- 16S
- *rplV*
- *secA*
- *tufB*



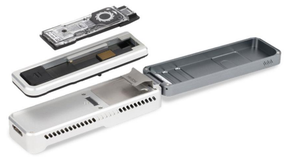
Pooling of PCR product



1x

Library preparation

Ligation Sequencing Kit V14 SQK-LSK114

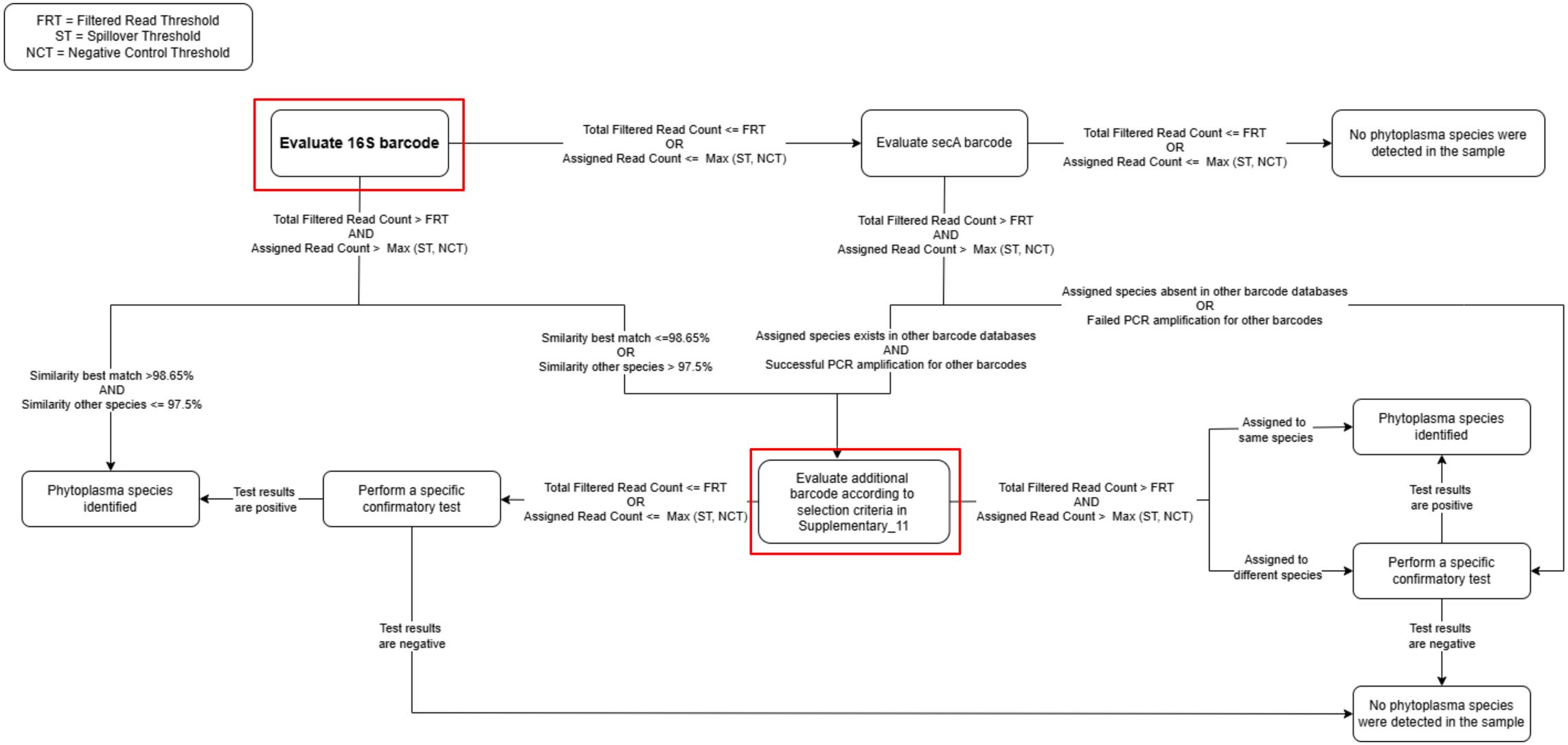


Phase 2

Guidelines translation into a structured workflow for ‘*Ca.* Phytoplasma’ detection and identification

Develop detection & ID scheme • Integrate ONT + read-based quantification • Barcode selection guidance • Re-evaluate demarcation criteria

Operationalizing the guidelines — a structured diagnostic scheme



Operationalizing the guidelines — a structured diagnostic scheme

16Sr Group	*Number of 'Ca. Phytoplasma' Species*	*'Ca. Phytoplasma' Species*	16S				Additional Barcode Suggestion	rplV				*secA*				*tufB*			
			PCR amplification	Database presence	Highest similarity percentage with other species	Closest related speces		PCR amplification	Database presence	Highest similarity percentage with other species	Closest related speces	PCR amplification	Database presence	Highest similarity percentage with other species	Closest related speces	PCR amplification	Database presence	Highest similarity percentage with other species	Closest related speces
16SrI: Aster yellows group	3	'Ca. Phytoplasma asteris'	Strong	Yes	99,5	'Ca. P. tritici'	*secA*	Strong	Yes	99,4	'Ca. P. tritici'	Strong	Yes	94,7	'Ca. P. tritici'	Strong	Yes	97,5	'Ca. P. tritici'
		'Ca. Phytoplasma lycopersici'		Yes	96,6	'Ca. P. asteris'			No				No				No		
		'Ca. Phytoplasma tritici'	Strong	Yes	99,5	'Ca. P. asteris'	*secA*	Strong	Yes	99,4	'Ca. P. asteris'	Strong	Yes	94,7	'Ca. P. asteris'	Strong	Yes	97,5	'Ca. P. asteris'
16SrII: Peanut witches' broom group	1	'Ca. Phytoplasma aurantifolia'	Strong	Yes	97,3	'Ca. P. brasiliense'	*rplV / secA / tufB*	Strong	Yes	73,9	'Ca. P. palmae' ('Ca. P. brasiliense' not in database)	Strong	Yes	82	'Ca. P. fraxini' ('Ca. P. brasiliense' not in database)	Weak	Yes	71,8	'Ca. P. trifolii' ('Ca. P. brasiliense' not in database)
	* Abolished	'Ca. Phytoplasma australasia'		No					No				No				No		
16SrIII: X-disease group	1	'Ca. Phytoplasma pruni'	Strong	Yes	94,8	'Ca. P. sacchari' / 'Ca. P. cynodontis'	*rplV / secA / tufB*	Strong	Yes	72,8	'Ca. P. noviguineense' (67,8% 'Ca. P. sacchari'; 67,3% 'Ca. P. cynodontis')	Strong	Yes	78,6	'Ca. P. rubi' / 'Ca. P. fraxini' (77,4% 'Ca. P. sacchari'; 75,8% 'Ca. P. cynodontis')	Strong	Yes	78,5	'Ca. P. convolvuli' (70,8% 'Ca. P. sacchari'; 70,9% 'Ca. P. cynodontis')
16SrIV: Coconut lethal yellows group	2	'Ca. Phytoplasma palmae'	Absent	Yes	97	'Ca. P. dypsidis'	*rplV*	Strong	Yes	82,4	'Ca. P. palmicola' (72,4% 'Ca. P. dypsidis')	Strong	Yes	84	'Ca. P. rubi' / 'Ca. P. vitis' ('Ca. P. dypsidis' not in database)	Strong	Yes	80,1	'Ca. P. fraxini' ('Ca. P. dypsidis' not in database)
		'Ca. Phytoplasma cocostanzaniae'		Yes	96,9	'Ca. P. palmae'			No				No				No		
16SrV: Elm yellows group	4	'Ca. Phytoplasma ulmi'	Strong	Yes	99,8	'Ca. P. vitis' (99,4% 'Ca. P. rubi')	*tufB*	Strong	Yes	97,9	'Ca. P. vitis'	Strong	Yes	98,3	'Ca. P. rubi' (97,8% 'Ca. P. vitis')	Strong	Yes	97,7	'Ca. P. rubi' (97,5% 'Ca. P. vitis')
		'Ca. Phytoplasma vitis'	Strong	Yes	99,8	'Ca. P. ulmi'	*tufB*	Strong	No	99,1	Ca. P. rubi' (97,9% 'Ca. P. ulmi')	Strong	Yes	98,9	'Ca. P. rubi' (97,8% 'Ca. P. ulmi')	Strong	Yes	98,7	'Ca. P. rubi (97,5% 'Ca. P. ulmi')
		'Ca. Phytoplasma ziziphi'		Yes	99,2	'Ca. P. ulmi'	*tufB*		Yes	97,3	'Ca. P. ulmi'		Yes	96,3	'Ca. P. vitis' (96,1% 'Ca. P. ulmi')		Yes	93	'Ca. P. vitis' / 'Ca. P. rubi' (91,9% 'Ca. P. ulmi')
		'Ca. Phytoplasma rubi'		Yes	99,4	'Ca. P. vitis' / 'Ca. P. ulmi'	*tufB*		Yes	99,1	Ca. P. vitis' (97,6% 'Ca. P. ulmi')		Yes	98,9	'Ca. P. vitis' (98,3% 'Ca. P. ulmi')		Yes	98,7	'Ca. P. vitis' (97,7% 'Ca. P. ulmi')
		'Ca. Phytoplasma balanitae'		Yes	98,5	'Ca. P. ziziphi'	*rplV*		Yes	94,1	'Ca. P. ziziphi'		No				No		
16SrVI: Clover proliferation group	2	'Ca. Phytoplasma trifolii'	Strong	Yes	98	'Ca. P. malaysianum'		Absent	No			Strong (aspecific amplification)	No			Absent		88	'Ca. P. fraxini' ('Ca. P. malaysianum' not in database)
		'Ca. Phytoplasma sudamericanum'		Yes	97,5	'Ca. P. fraxini'			No				No				No		
16SrVII: Ash yellows group	97	'Ca. Phytoplasma fraxini'		Yes	97,9	'Ca. P. trifolii'	*tufB*		Yes	84,9	'Ca. P. ziziphi' ('Ca. P. trifolii' not in database)		Yes	89,6	'Ca. P. ulmi' ('Ca. P. trifolii' not in database)		Yes	88	'Ca. P. trifolii'
16SrVIII: Loofah witches' broom group	1	'Ca. Phytoplasma luffae'		Yes	97,7	'Ca. P. stylosanthis'	*rplV / secA / tufB*		Yes	83,4	'Ca. P. malaysianum' (79,8% 'Ca. P. stylosanthis')		Yes	83,4	'Ca. P. rubi' (82,3% 'Ca. P. stylosanthis')		Yes	86,4	'Ca. P. stylosanthis'
16SrIX: Pigeon pea witches' broom group	1	'Ca. Phytoplasma phoenicium'	Strong	Yes	96,3	'Ca. P. omanense'	*rplV / secA*	Strong	Yes	77,8	'Ca. P. noviguineense' ('Ca. P. omanense' not in database)	Strong	Yes	83,4	'Ca. P. vitis' / 'Ca. P. ulmi' ('Ca. P. omanense' not in database)	Very weak	Yes	79,5	'Ca. P. trifolii' ('Ca. P. omanense' not in database)

ILVO

Barcode Evaluation: Controls & read-based quantification

Quantitative analysis of barcode-specific read count data from samples and controls

Three control types per ONT run

NC	**Negative Control** Purified water — monitors background contamination
PC	**Positive Control** Low-concentration non-sample target — determines LOD & sequencing efficiency
AC	**Alien Control** High-concentration non-sample target — monitors cross-contamination

Three quantitative thresholds

FRT	**Filtered Read Threshold** Filtering background noise before taxonomic assignment
ST	**Spillover Threshold** Cross-contamination: alien or sample to sample contamination
NCT	**Negative Control Threshold** Contamination from samples to negative control

Reliable assignment:

- Filtered reads must exceed FRT
- Assigned reads must exceed Max(ST, NCT)

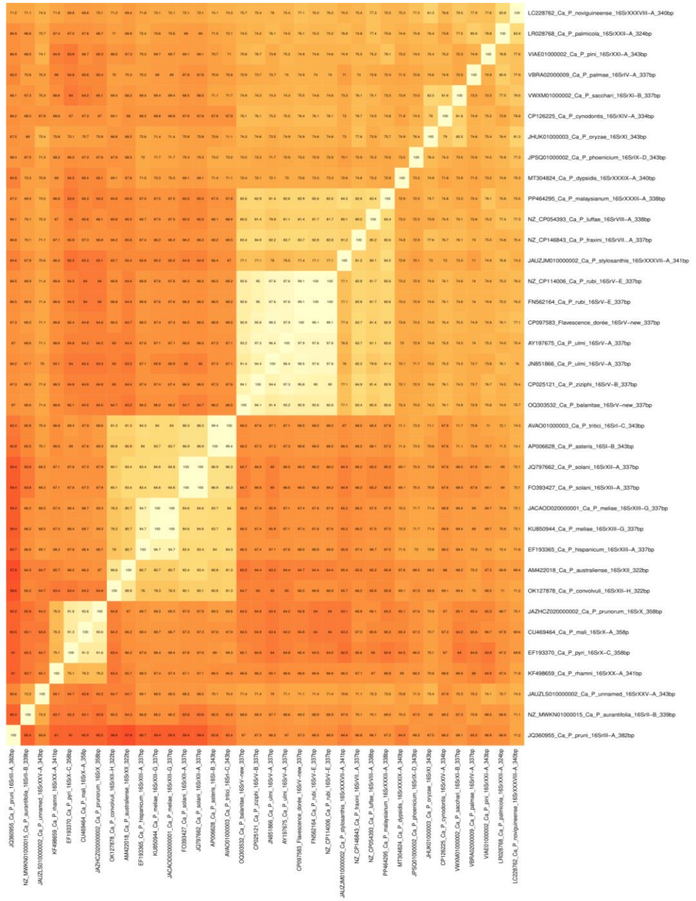
ILVO

Barcode Evaluation: Taxonomic validation

Consensus generation:

When: The number of reads assigned to a taxon exceeds 5% of the total filtered reads

How: Top 100 quality reads → alignment → majority nucleotide per position → identity matrix for validation.



16S rRNA barcode

Fixed threshold approach

→ >98.65% similarity to best match

→ <=97.5% similarity to all other species

→ Conservative dual-threshold reduces misassignment risk

secA • *rplV* • *tufB*

Best-match principle

→ Assigned to species with highest similarity

→ No fixed thresholds — avoids arbitrary cutoffs

→ Species accepted when ≥2 barcodes agree

Phase 3

Evaluation of the MLSA-ONT Barcoding Workflow

Sensitivity • Specificity • Repeatability • Comparison with Sanger • PCR Multiplexing

Sensitivity, specificity & repeatability

Sensitivity

- Spiked dilutions: as expected, read counts decreased with increasing dilution across all barcodes.
- ONT sensitivity comparable to or better than Sanger: producing results even without visible PCR signal
- *rplV* showed lowest sensitivity

Specificity

- Complete sample panel analyzed
- All target species detected and correctly assigned
- Exception: 16SrI group — assignment challenges noted

Repeatability

- 4 samples analyzed twice across different runs
- Taxonomic assignments consistent across replications

The ONT workflow demonstrated robust performance across all three diagnostic validation criteria

ILVO

ONT vs Sanger sequencing & PCR multiplexing

Sequence generation success

% of sample-barcode combinations yielding a usable sequence
(22 samples × 4 barcodes = 88 combinations)

86%
ONT overall

78%
Sanger overall

Sequence accuracy

% similarity between ONT consensus and Sanger sequence
(65 combinations where both methods yielded usable sequences)

47/65 100% match **11/65** 99–100% **7/65** –99%

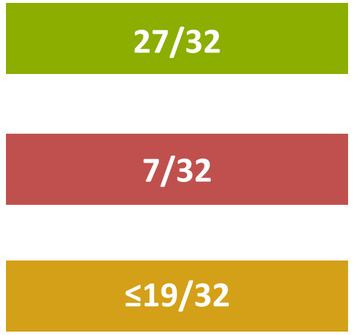
PCR multiplexing evaluation

8 samples across 4 runs:

Run 1: Simplex
Each barcode amplified individually at optimal conditions

Run 2: Full multiplex
All 7 barcodes in a single PCR reaction

Run 3–4: Split multiplex
Barcodes split across 2 PCRs at different annealing temps



Key insight: PCR multiplexing reduced sensitivity

ILVO

Key conclusions

1 Successful MLSA-ONT barcoding workflow development and evaluation

Preliminary steps successfully completed:

- Target taxa identified; 4 barcodes selected
- Curated reference databases constructed
- Primers assessed and optimized
- Successful development of MLSA-ONT barcoding workflow

Robust ONT evaluation:

- Sensitivity, specificity and repeatability comparable to or exceeding Sanger sequencing
- PCR multiplexing reduced sensitivity

2 Structured diagnostic workflow

Operationalizes the established
'*Ca.* Phytoplasma' species-assignment guidelines:

- Best-match principle replaces fixed thresholds for non-16S barcodes
- Read-based quantification with defined controls (NC, PC, AC) and thresholds (FRT, ST, NCT) to increase diagnostic reliability
- Structured barcode selection guidance per 16Sr group
- The '*Ca.* Phytoplasma' scheme is also compatible with Sanger seq.

ILVO

Limitations & future outlook

Current limitations

PCR multiplexing sensitivity
Higher degrees of PCR multiplexing reduce sensitivity, not yet viable for routine throughput advantage

Incomplete databases
Reference databases constrain identification, especially for non-16S barcodes

Higher operational cost
ONT currently more time-consuming and costly per sample than Sanger

Proof-of-concept status of diagnostic scheme
Increased complexity. Not intended as the next gold standard, but a structured implementation for complex cases

Future directions

Automate bioinformatics
Streamline the pipeline for routine diagnostic use

Expand reference databases
Fill gaps in non-16S barcode coverage across species

Broader validation
Test across a wider range of phytoplasma species and host plants

Broader application of the MLSA-ONT barcoding strategy
This approach can be extended to diagnose other pathosystems

Thank you!

QUESTIONS?

Contact:

Flanders Research Institute for Agriculture, Fisheries and Food
Ellen Everaert
ILVO – Plant Sciences UnitMerelbeke – Belgium
T + 32 (0)9 272 24 19

ellen.everaert@ilvo.vlaanderen.be
www.ilvo.vlaanderen.be

ILVO