


BEXYL
BEYOND XYLELLA

# Evaluation of molecular approaches for absolute quantification, in-situ detection and molecular typing of *Xylella fastidiosa* in plants and vectors

Deliverable 2.2

29 August 2025

Author(s): Blanca B. Landa, Ana Sánchez, María P. Velasco-Amo, Concepción Olivares, Miguel Román-Écija, Juan A. Navas-Cortés (P1-CSIC), Ester Marco-Noales (P2-IVIA), Amoia S. Serena, Nicoletta Contaldo, Pasquale Saldarelli, Angelantonio Minafra, Angela Ligorio, Vincenzo Cavalieri, Donato Boscia, Maria Saponari, Giuliana Loconsole (P5-CNR), Milica Pastar, Stéphane Compant (P13-AIT), Carlos Chacón (P16-UCR)

Funded by the European Union

# DELIVERABLE INFORMATION

| | |
|---|---|
| **Start of the project:** | 01 Sep 2022 |
| **Duration:** | 48 months |
| **Coordinator:** | Blanca B. Landa |
| **Deliverable title:** | Evaluation of molecular approaches for absolute quantification, in-situ detection and molecular typing of *Xylella fastidiosa* in plants and vectors |
| **Deliverable n°:** | 2.2 |
| **Nature of the deliverable:** | Report |
| **Dissemination level:** | Public |
| **WP contributing:** | WP2 |
| **Lead beneficiary:** | IVIA |
| **Due date of deliverable:** | 31 Aug 2025 |
| **Actual submission date:** | 29 August 2025 |

**Deliverable status**

| Ver. | Status | Date | Author(s) |
|---|---|---|---|
| 1.0 | Final | 29 August 2025 | Blanca B. Landa, María P. Velasco-Amo, Concepción Olivares, Miguel Román-Écija, Juan A. Navas-Cortés (P1-CSIC), Ester Marco-Noales (P2-IVIA), Amoia S. Serena, Nicoletta Contaldo, Pasquale Saldarelli, Angelantonio Minafra, Angela Ligorio, Vincenzo Cavalieri, Donato Boscia, Maria Saponari, Giuliana Loconsole (P5-CNR), Milica Pastar, Stéphane Compant (P13-AIT), Carlos Chacón (P16-UCR) |

**Disclaimer**

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.



Funded by the European Union

# Content

1. Introduction ..... 6

2. Detection of *Xylella fastidiosa* in Host Plants and Insect Vectors by Droplet Digital PCR ..... 9

*2.1 Background* ..... *9*

*2.2 Materials and Methods* ..... *9*

2.2.1 Bacterial Strain ..... 9

2.2.2 Plant Sources and Insects ..... 10

2.2.3 Sample Preparation and DNA Extraction ..... 10

2.2.4 Quantitative PCR Assay ..... 10

2.2.5 Droplet Digital PCR Assay ..... 11

2.2.6 Evaluation of ddPCR Performance in Naturally Infected Plant and Insect Samples 11

2.2.7 Data Analysis ..... 12

*2.3 Results* ..... *12*

2.3.1 Optimization of the ddPCR Assay ..... 12

2.3.2 Linearity and Analytical Sensitivity of qPCR Assay ..... 13

2.3.3 Linearity and Analytical Sensitivity of the ddPCR Assay ..... 14

2.3.4 Comparison between ddPCR and qPCR Assays in Artificially Contaminated Samples ..... 16

2.3.5 Evaluation of the ddPCR Assay in Field Samples ..... 18

3. Naked-Eye Molecular Testing for the Detection of *Xylella fastidiosa* in Almond Orchards by Colorimetric LAMP ..... 19

*3.1 Background* ..... *19*

*3.2 Materials and Methods* ..... *19*

*3.3 Results* ..... *22*

4. Modeling the Accuracy of *Xylella fastidiosa* Molecular Diagnostic Tests in Naturally Infected Almond Tree Samples ..... 24

*4.1 Background* ..... *24*

*4.2 Materials and Methods* ..... *25*

4.2.1 Plant material ..... 25

4.2.2 Visual assessment of ALS symptoms ..... 25

4.2.3 DNA extraction ..... 25

4.2.4 DNA standard curves ..... 26

4.2.5 Quantitative real-time polymerase chain reaction assays (q-PCR) ..... 26



Funded by the European Union

4.2.6 Recombinase-polymerase amplification assay (RPA) ..... 27
4.2.7 Droplet digital polymerase chain reaction (ddPCR) ..... 27
4.2.8 Statistical Analysis ..... 28
*4.3 Results ..... 29*
4.3.1 Reproducibility, efficiency and sensitivity of qPCR, RPA and ddPCR tests ..... 29
4.3.2 Detection of *Xf* in naturally infected almond samples ..... 33
4.3.3 Performance and agreement between diagnostic tests ..... 34
4.3.4 Diagnostic accuracy of diagnostic tests ..... 36
4.3.5 Latent Class Model (LCM) Analysis ..... 37

5. High-Resolution Diagnosis of *Xylella fastidiosa* Using Targeted NGS Enrichment ..... 39
*5.1 Background ..... 39*
*5.2 Materials and Methods ..... 40*
5.2.1 *Xylella fastidiosa* strains and DNA isolation ..... 41
5.2.2 Mock inoculated samples of plants and insect vectors ..... 41
5.2.3 Development of *Xf* DNA standard curves ..... 41
5.2.4 Naturally-infected plants and insect vectors samples ..... 42
5.2.5 Selection of *Xylella fastidiosa* target sequences and baits design ..... 44
5.2.6 Illumina library preparation of studied samples ..... 45
5.2.7 *Xylella fastidiosa*-TSCE and sequencing ..... 45
5.2.8 *In silico* bioinformatic pipeline ..... 46
*5.3 Results ..... 47*
5.3.1 Limit of detection of the *Xf*-Targeted Sequence Capture Enrichment ..... 47
5.3.2 Effect of library pooling and concentration of Xylella fastidiosa on Xf-TSCE results 47
5.3.3 *Xylella fastidiosa*-TSCE in naturally infected samples ..... 50
5.3.4 Phylogenetic reconstruction using *Xf*-TSCE ..... 51

6. A FLOW-FISH Method for Detection and Quantification of *Xylella fastidiosa* ..... 57
*6.1 Background ..... 57*
*6.2 Materials and Methods ..... 58*
6.2.1 Bacterial culture and plant material ..... 58
6.2.2 Optimization of FISH and FLOW-FISH protocol using *X. fastidiosa*-specific and eubacteria probes ..... 59
*6.3 Results ..... 60*



Funded by the European Union

6.3.1 Evaluation of probe specificity and cross-reactivity for FLOW-FISH optimization 60

6.3.2 Optimization of XF-KO210 and Eub338 probe concentrations and interaction analysis 62

6.3.3 Performance evaluation of *Xylella fastidiosa* FLOW-FISH in coffee matrices spiked with *X. fastidiosa* cells.....63

6.3.4 Detection limit of *Xylella fastidiosa* FLOW-FISH in coffee extracts.....63

7. Conclussions.....65

8. Publications from this deliverable.....65

9. References.....66

ANNEX 1. Supplementary figures.....73

ANNEX 2. Supplementary tables.....81



Funded by the European Union

# 1. INTRODUCTION

*Xylella fastidiosa* (*Xf*) is a xylem-inhabiting plant pathogenic bacterium classified as a priority pest in the European Union (EU) due to its potential for devastating impact on economically important crops as well as culturally valued ornamental plants and landscape vegetation. This pathogen causes a variety of severe diseases, leading to substantial yield losses and posing significant threats to the heritage, landscape and biodiversity of affected regions (EFSA et al. 2023; Sánchez et al. 2019). The destructive potential of the bacterium is most evident when introduced into non-native areas with a conducive combination of a susceptible host plant, favourable environmental conditions and established insect vector populations. Prominent examples include the Olive Quick Decline Syndrome (OQDS) epidemic in Apulia, Italy, and the Almond Leaf Scorch (ALS) epidemic in Alicante, mainland Spain.

Early detection is essential for preventing the spread of diseases caused by this pathogen and remains a top research priority, particularly in Europe. The development and implementation of robust preventive management strategies, including surveillance, monitoring, and accurate diagnosis of *Xf* and its vectors are critical. These measures are key not only in mitigating its immediate impacts, but also in assisting timely eradication or optimizing containment practices for its control (Almeida 2016; De la Fuente et al. 2024; Zarco-Tejada et al. 2018). One of the largest eradication campaigns for *Xf* in Europe is currently underway in Alicante, mainland Spain, targeting the ALS outbreak caused by *Xf* subsp. *multiplex*. In this area, the Demarcated Area (DA) spans over 156,864 ha and the Infected Zone (IZ) 3,589 ha, with more than 242,840 almond trees already eradicated since the first detection in the summer of 2017 (Cendoya et al. 2024; De La Fuente et al. 2024; GVA, 2024).

Conventional diagnostic methods for *Xf,* based on isolation, biochemical, and serological tests are time-consuming, labor-intensive, and often lack sensitivity (Baldi and La Porta 2017). Currently, a variety of molecular detection tests, including thermal and isothermal amplification techniques, are available for *Xf*, each with varying limits of detection (LoDs). The EPPO Diagnostic protocol (EPPO 2023) for *Xf* is the most complete source of standardized guidelines for sampling, testing, and reporting *Xf*. This standard also provides a detailed description of various screening tests for *Xf* diagnosis, completed with validation data, and serves as a cornerstone in global efforts to control and manage *Xf* outbreaks. It supports effective surveillance, early detection, and prompt response efforts across Europe. In fact, the EU has established stringent regulations for the detection and diagnosis of *Xf* (e.g., Commission Implementing Regulation (EU) 2020/1201, amended and corrected by Commission Implementing Regulation (EU) 2024/2507) that specify which tests must be used for official surveys and the identification of *Xf* and its subspecies. These regulations have evolved in parallel with multiple updates to the EPPO protocol, benefiting from contributions by numerous experts and diagnostic laboratories to enhance diagnostic accuracy and reliability.

Among the diagnostic procedures available for *Xf*, molecular tests are the most recommended due to their higher analytical sensitivity and specificity (EPPO 2023; Roy et al. 2010). Several conventional specific PCR tests have been developed for *Xf* diagnosis, targeting various genes such as the β-subunit polypeptide of the DNA gyrase (*gyr*B) (Rodrigues et al. 2003), the RNA polymerase sigma-factor 70 (*rpo*D) (Minsavage et al. 1994), or the seven housekeeping genes



Funded by the European Union

*cys*G, *glt*T, *hol*C, *leu*A, *mal*F, *nuo*L, *pet*C used for Multilocus Sequence Type (MLST) analysis (Yuan et al. 2010). To further enhance sensitivity, a nested-MLST protocol was developed based on the same loci as conventional MLST but offering improved detection thresholds (Cesbron et al. 2020).

Additionally, several quantitative real-time quantitative PCR (qPCR) tests have been developed, enabling not only the detection of *Xf* but also its quantification. In Europe, the most commonly used qPCR tests for *Xf* diagnosis since the first outbreak in Italy include the Francis-qPCR (Francis et al. 2006), Harper-qPCR (Harper et al. 2010), Li-qPCR (Li et al. 2013), and Ouyang-qPCR (Ouyang et al. 2013). Additionally, recently, a multiplex TaqMan qPCR test was developed that combines the primers and probes of Harper-qPCR and Ouyang-qPCR with an internal control, providing analytical sensitivity comparable to the original tests (Bonants et al. 2019). In addition, specific qPCR tests have also been developed to simultaneously detect and assign *Xf* at the subspecies level (Burbank and Ortega 2018; Dupas et al. 2019a; Hodgetts 2022).

Due to their higher analytical sensitivity compared to other molecular tests, qPCR tests are primarily recommended for testing asymptomatic plant material, insect vectors, and samples in areas where *Xf* has not been detected yet or in buffer zones of demarcated areas. The extensive application of some of the qPCR tests mentioned above has been exemplified in various studies, including those investigating the distribution and host range of *Xf* in different outbreaks across Europe (Jacques et al. 2016; Moralejo et al. 2020; Marco-Noales et al. 2021; Olmo et al. 2021; Saponari et al. 2013), characterizing potential insect vectors of *Xf* (Cunty et al. 2020; Elbeaino et al. 2014), supporting pathogen surveillance efforts (Cornara et al. 2024), or providing complementary data for remote sensing studies (Zarco-Tejada et al. 2018, 2021).

Time and portability are also important factors in the diagnosis of quarantine plant pathogens (Aglietti et al. 2019). In light of the need for simplified and rapid detection assays that reduce nucleic acid amplification time compared to conventional methods, various isothermal nucleic acid amplification techniques have been developed for *Xf* such as loop-mediated isothermal amplification (LAMP) (Aglietti et al. 2019; Amoia et al. 2023a; Notomi et al. 2000; Waliullah et al. 2019) and the recombinase polymerase amplification (RPA) AmplifyRP XRT+ kit (Kersting et al. 2014). This latter test is user-friendly, requiring minimal laboratory staff training, and has been adapted for field use, enabling preliminary on-site screening (Cesbron et al. 2023).

Detection of phytopathogenic bacteria through droplet digital PCR (ddPCR) has demonstrated high detection efficiency, particularly at low pathogen concentrations, and offers improved tolerance to PCR inhibitors commonly found in plant tissues (Dreo et al. 2014). A ddPCR protocol for *Xf*, based on the Harper-qPCR test, was developed and proved to have comparable, superior, or inferior detection limits than Harper-qPCR test depending on the host plants tested (Dupas et al. 2019b). In a separate study, this ddPCR protocol was also applied to detect *Xf* in insect vectors and economically important crops, demonstrating higher sensitivity than the Harper-qPCR test (Amoia et al. 2023b).

To date, diagnostic and epidemiological monitoring in Europe have relied mainly on quantitative PCR (qPCR) and multilocus sequence typing (MLST). While qPCR remains the method of choice for rapid detection due to its sensitivity and specificity, it cannot provide


Funded by the European Union

information on genetic diversity beyond presence/absence. Conversely, MLST, based on seven housekeeping genes, allows strain-level classification but has limited resolution for inferring phylogenetic relationships and tracking evolutionary processes. New sequencing-based methods are therefore needed to complement and enhance existing tools, providing both higher sensitivity in complex host matrices and greater discriminatory power for epidemiological applications.

This Deliverable presents advances in the development and application of novel molecular tools aimed at improving the detection, characterization, and phylogenetic resolution of *Xf*. The document comprises five main components: i) Detection of *Xf* in host plants and insect vectors by droplet digital PCR (ddPCR); ii) Detection of *Xf* in almond orchards by colorimetric LAMP; iii) Modeling the accuracy of *Xf* molecular diagnostic tests in naturally infected almond tree samples; iv) Development of a High-Resolution diagnostic approach based on targeted next-generation sequencing enrichment (*Xf*-TSCE); and v) Development of a FLOW-FISH method for detection and quantification of *Xf*.

Together, these sections highlight the contribution of the Deliverable to the objectives of BeXyl project, demonstrating how the generated knowledge and molecular diagnostic methodologies can strengthen both research capacities and applied management strategies for *Xf*.



Funded by the European Union

# 2. DETECTION OF *XYLELLA FASTIDIOSA* IN HOST PLANTS AND INSECT VECTORS BY DROPLET DIGITAL PCR

## 2.1 Background

Droplet digital polymerase chain reaction (ddPCR) is an advanced and powerful molecular technology that allows for accurate detection and absolute quantification of the nucleic acid present in samples, even when the target is present at a very low level. The principle of digital PCR was first described in the 1990s (Sykes et al. 1992; Vogelstein et al. 1999) and, later on, was further optimized for different diagnostic purposes. Differently from the qPCR, ddPCR does not require a calibration curve nor the use of housekeeping genes to normalize the results, thus reducing the reaction requirements and facilitating the data analysis and comparability (Hindson et al. 2011; Pinheiro et al. 2012). The ddPCR technology massively distributes the sample into thousands of independent nanoliter-sized droplets in a random approach. Each droplet acts as an individual PCR test tube in which amplification can take place (Huggett et al. 2013; Whale et al. 2020). Partitions containing amplified fluorescent products are considered positive, whereas those without fluorescence are considered negative. The absolute number of target DNA molecules in a sample can be calculated directly from the ratio of positive droplets to total partitions using Poisson's law of small numbers (Hindson et al. 2011; Pinheiro et al. 2012). Although this molecular technique is more expensive and time-consuming than qPCR, the high accuracy and sensitivity support its application in the diagnosis of microorganisms in the early stages of infection. Indeed, timely detection of new outbreaks of quarantine pests is fundamental to avoid the establishment of exotic pathogens in new areas. Digital droplet PCR has been shown to achieve more precise detection results than qPCR, and this molecular method has now gained interest in different biological applications for research and diagnostic purposes. Our work aimed to implement the use of ddPCR for the detection and quantification of *Xf* in different susceptible plant species, especially in insect vectors, as the latter are very rich in PCR inhibitors and contaminants. The sensitivity and efficiency were compared to those obtained by qPCR using the same DNA preparations.

## 2.2 Materials and Methods

### 2.2.1 Bacterial Strain

The reference bacterial strain 'De Donno', belonging to the subsp. *pauca* (*Xfp*) and harboring the sequence type (ST) 53, was isolated in the Apulia region (Italy) in 2016 from an olive tree affected by olive quick decline syndrome (OQDS) (Giampetruzzi et al. 2017). It was used as target strain to prepare artificially contaminated plant sap containing known concentrations of bacterial cells/copies of the target DNA. Bacterial colonies, grown on PD3 solid medium at 28 °C for 7 days (Davis 1980), were scraped from plates and dispersed in sterile demineralized water to prepare a bacterial suspension with an optical density (OD) at 600 nm of 0.5 OD. This suspension was diluted 1:4 (approx. $10^8$ CFU/mL) and used to contaminate plant saps recovered from all the matrices selected in this study. The concentration of the bacterial suspension (CFU/mL) was also determined by plate counting.



Funded by the European Union

### 2.2.2 Plant Sources and Insects

Plant matrices used to evaluate ddPCR assay were collected from *Xf*-free sources of *Olea europea*, *Vitis vinifera*, *Citrus sinensis*, *Prunus dulcis*, and *Nerium oleander*, grown in the greenhouse. For the insect vectors, adult specimens of *Philaenus spumarius* (Hemiptera: Aphrophoridae) were collected in fields located in an *Xf*-free area of the Apulia region. The accuracy of ddPCR assays was also evaluated on naturally infected olive and insect samples collected in the framework of the regional monitoring program carried out in Apulia in 2021–2022. The panel of naturally infected olive samples was selected in order to obtain a full representation of samples with high, medium, and low bacterial concentrations.

### 2.2.3 Sample Preparation and DNA Extraction

Plant tissues used in our tests consisted of leaf tissues (either petioles, midribs, or basal leaf parts) for oleander, grapevine, and citrus; stem/twig portions of 1–1.5 cm long for olive; and scraped xylem tissue from hardwood cutting for almond. Insect samples were prepared by removing the head from the body (EPPO 2023). Plant and insect tissues were homogenized in extraction buffer (CTAB buffer) following the standard ratio (w:v) (EPPO 2023). Artificially contaminated samples were prepared by spiking both the plant sap and the macerated insect with the *Xf* suspension previously described in order to obtain 3 replicates/matrices of 10-fold serial dilutions containing from $10^7$ to 10 CFU/mL. Both the spiked and the naturally infected samples were processed using the Maxwell® RSC PureFood GMO and Authentication Kit (Promega), following the manufacturer's instructions to extract DNA from plants and insects. For insects, individual excised heads were homogenized in 500 µL of CTAB buffer (EPPO 2023), and DNA was purified using the same kit described for plant tissues, with the only difference being that the total DNA was eluted in 50 µL of elution buffer (EB, Promega). Negative and positive internal controls (NIC, PIC) for each matrix were prepared and processed as described previously. The PIC was contaminated with a bacterial suspension $10^7$ CFU/mL before the DNA extraction.

### 2.2.4 Quantitative PCR Assay

For qPCR assays, the procedure described in Annex 5—EPPO PM 7/24 (4) (EPPO 2023)—was based on the set of primers/probe designed by Harper et al. (2010), but was slightly modified. Briefly, the reaction mix was prepared in a final volume of 11 µL containing: 1x TaqMan™ Fast Advanced Master Mix (Applied biosystem), 300 nM of each primer (*XF*-F/*XF*-R), 200 nM of 6-FAM/BHQ-1 labeled probe (*XF*-P), and 1 µL of DNA extract. Positive and negative amplification controls (i.e., PAC, NAC) were included in each test. The amplification conditions were 95 °C for 5 min, followed by 40 cycles of 94 °C for 10 s and 62 °C for 40 s. Tests were performed on the thermal cycler CFX96 Touch Real-Time PCR Detection System (Bio-rad, Hercules, CA, USA) using 96-well plates (Bio-rad). For artificially contaminated samples, each sample/replicate was amplified in triplicate and each replicate/matrix (including NIC) was tested in at least three independent assays. DNA recovered from the naturally infected samples was tested in duplicate wells. Data acquisition and analysis were performed using the CFX Maestro 1.1 version N. 4.1 software (Bio-rad, Hercules, CA, USA). The qPCR amplification efficiency was


Funded by the European Union

estimated for each matrix from the slopes of the standard curves generated by the 10-fold serial dilutions, using the equation E = 10−1/slope ± 1. The limit of quantification (LOQ) of the assay was defined based on the lowest dilution yielding positive results for the three replicates tested when these were in the linear range of the standard calibration curve (i.e., when the ΔCq, among the dilution considered and the previous one tested, was close to 3). The limit of detection (LOD) corresponded to the lowest serial dilution yielding a positive result in more than 50% of the overall replicates of each matrix tested in all independent assays. A recombinant plasmid containing a region of the rimM gene targeted by the primers/probe reported by Harper et al. (2010), was used to set up PCR reactions with known concentrations of target DNA (ng/µL) and, from there, to calculate the copy number of *Xf* bacterial DNA.

### 2.2.5 Droplet Digital PCR Assay

Digital droplet PCR was performed on QX200 Droplet Digital PCR System (Bio-Rad) according to the manufacturer's instructions (Whale et al. 2020), using the primers/probes of Harper et al. (2010). To implement the performance of the ddPCR assay reported by Dupas et al. (2019b), i.e., to achieve a better separation of positive and negative droplet clusters without 'droplet rain' in naturally infected olive plant and insect samples, different volumes of plant/insect extracts (2 µL, 4 µL, 6 µL, and 8 µL) were tested using ddPCR conditions reported by Dupas et al., (2019b). Once the optimal amount of purified DNA of plants and insects was determined, 3 different primers/probe concentrations (900 nM/300 nM; 600 nM/300 nM; and 300 nM/300 nM) were tested in 20 µL of the reaction mix (2x ddPCR™ Supermix for Probes (No dUTP)) using both spiked and naturally infected olive and insect samples as templates. The reaction mix (20 µL) and 70 µL of droplet-generating oil (Bio-rad) were added to a droplet-generating DG8 cartridge and loaded onto Bio-rad Automated Droplet Generator (Bio-rad). The water-in-oil droplets (40 µL) were carefully transferred to a 96-well PCR plate, which had been heat-sealed at 175 °C with a pierceable foil using a PX1TM PCR Plate Sealer (Bio-rad) and then placed in a C1000 Thermal Cycler (Bio-rad) for end-point PCR. Amplification reactions were performed with the cycling parameters optimized by Dupas et al. (2019b), except for increasing the number of cycles to 45 with a temperature ramp rate of 2 °C/s. After amplification, the PCR plate was directly transferred to the droplet reader (QX200TMDroplet DigitalTM System (Bio-rad)) set in the absolute quantification (ABS) modality. When the assay was carried out on spiked matrices, each replicate/matrix of the 10-fold serial dilution from 106 to 10 CFU/mL was amplified in duplicate in 3 different runs. Then, for each replicate, samples with concentrations close to the LOQ, identified by qPCR, were amplified in duplicate in several separated runs to evaluate the reproducibility of the results between different amplification reactions. The LOD was calculated as previously described. Finally, field plant and insect samples were checked by ddPCR in duplicate in a single run.

### 2.2.6 Evaluation of ddPCR Performance in Naturally Infected Plant and Insect Samples

The performance efficiency of both ddPCR and qPCR assays to detect *Xf* in naturally infected samples from olive trees and insects was assessed using 34 olive samples from the field and



Funded by the European Union

27 specimens of *P. spumarius* collected by sweep net in the demarcated infected area of Apulia in July 2022. DNA from plants and insects was extracted as previously described. Each sample was then amplified in duplicate in the same amplification run. In addition, negative controls (samples collected in an *Xf*-free area) were included in each test to evaluate the presence of PCR contaminants.

### 2.2.7 Data Analysis

The linear regression of the standard curves generated by qPCR assay was calculated by CFX™ Maestro Software version 2.2 (Bio-Rad). For ddPCR, fluorescent signals of droplets were acquired, expressed in target copies/µL of the reaction, and analyzed by the QX-Manager-V.1.2-STD software (Bio-Rad). For each ddPCR experiment, positive droplets, with higher fluorescent signals, and negative droplets, with lower fluorescent signals (considered as background), were divided by applying an amplitude threshold line; ddPCR reactions with fewer than 10.000 generated droplets were excluded from the analysis, and a reaction was considered positive if at least three positive droplets were counted (Raguseo et al. 2021). Poisson's statistic was used to calculate the absolute concentration in each sample as a copy number/µL of the target DNA. The error reported for every single well was represented as the Poisson assessed at a 95% confidence level. The linear range of the ddPCR assay was determined by plotting the number of expected target copies/µL vs. the number of target copies/µL generated by ddPCR using both serial dilutions of the plasmid DNA and the bacterial suspensions. Pearson's correlations and linear regression were also used to evaluate the relationship between the results of ddPCR and qPCR assays.

## 2.3 Results

### 2.3.1 Optimization of the ddPCR Assay

Preliminary experiments included the use of different amounts of DNA preparations from plants and insects (Supplementary Figure S1). Once the optimal number of samples was defined, different primers/probe concentrations were tested in an attempt to improve the detection of the bacterium in spiked and naturally infected olive and insect samples. By increasing the number of cycles to 45, with respect to the study reported by Dupas et al. (2019b), the optimal primer/probe concentrations and DNA quantities were found to correspond to 600/300 nM (primers/probe) and 4 and 6 µL of purified plant/insect DNA, respectively. These reaction conditions proved to be suitable for detecting *Xf* in all six matrices analyzed, with an optimal difference in the fluorescence signals between positive and negative droplets and generating a high number of amplification products (number of positive droplets) (Figure 1A–F). When compared to qPCR, ddPCR yielded a narrower linearity range, from $10^6$ to $10^1$ CFU/mL, since the high concentration of positive droplets saturated the fluorescence signal at concentrations above $10^6$ CFU/mL, making the Poisson algorithm invalid. Negative and positive controls correctly produced the expected results, with no fluorescence signal detected in NIC or NAC.


Funded by the European Union

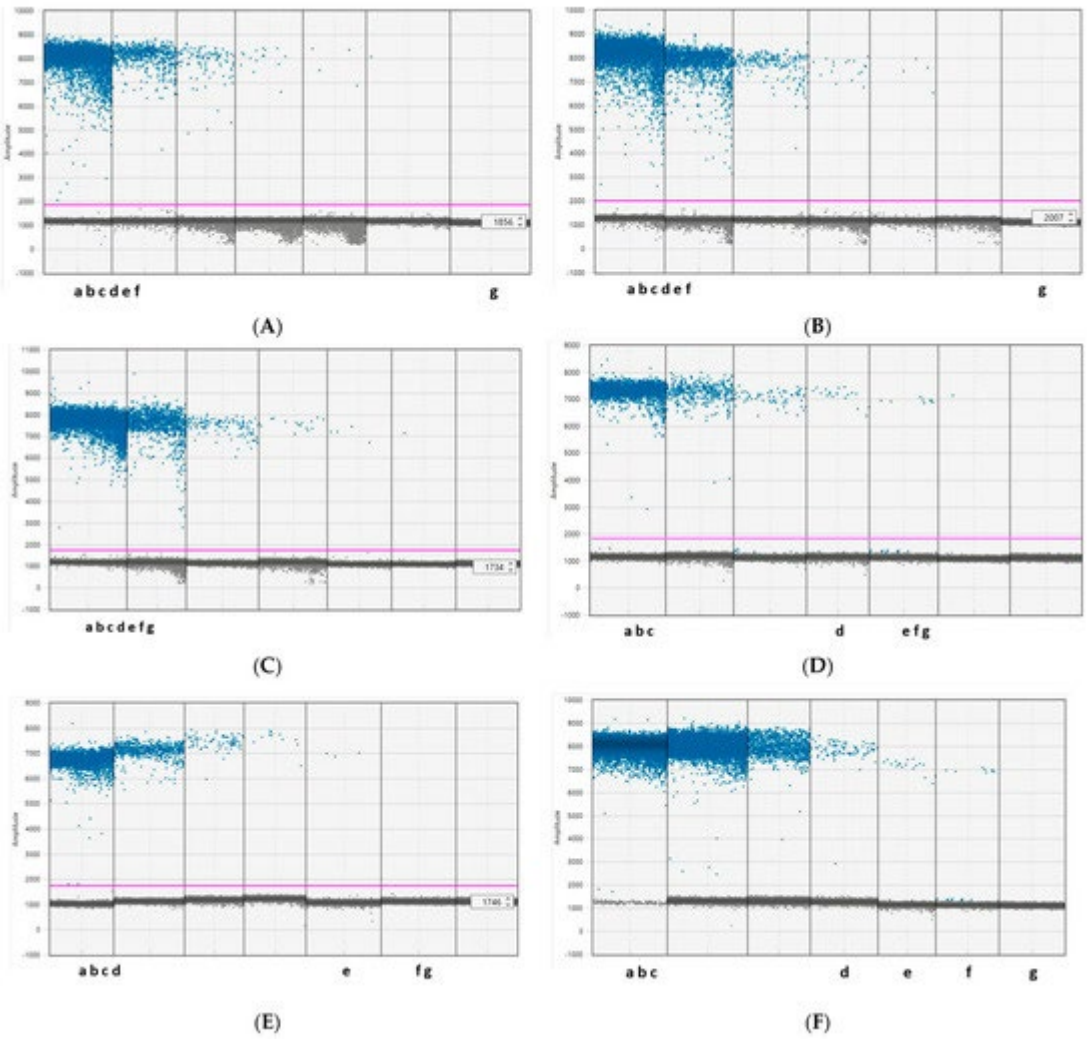



**Figure 1.** Comparison of the different limits of detection of *Xylella fastidiosa* (*Xf*) obtained by ddPCR in the five plant spiked matrices and in the insect vector at the optimized ddPCR parameters (600/300 nM primer/probe concentration and 45 end-point PCR cycles). Blue dots represent positive droplets above the pink horizontal threshold line. Gray dots represent the negative droplet background, with no amplification. (A): *O. europaea*; (B): *V. vinifera*; (C): *P. dulcis*; (D): *N. oleander*; (E): *C. sinensis*; and (F): *P. spumarius*. On x-axis: ten-fold dilution of *Xf* suspension reported in wells a to f; a: $10^6$ CFU/mL; b: $10^5$ CFU/mL; c: $10^4$ CFU/mL; d: $10^3$ CFU/mL; e: $10^2$ CFU/mL; and f: 10 CFU/mL. Well g: NIC (negative internal control), specific for each matrix; on the y-axis: amplitude value.

### 2.3.2 Linearity and Analytical Sensitivity of qPCR Assay

Ten-fold serial dilutions of both plasmid DNA and *Xfp* bacterial suspension were used to generate the standard curves for the qPCR tests (Figure 2A,B). When testing artificially contaminated samples, containing either plasmid DNA (50 pg/µL–50 ag/µL) or bacterial suspension ($1.00 \times 10^7$–$1.00 \times 10^1$), the recovered curves showed good linearity, with R2 values of 0.9943 and 0.9973 for the rimM-recombinant DNA plasmid and the bacterial suspension, respectively. The qPCR efficiency values retrieved from these standard curves were 100.72% and 116.96% for plasmid DNA and bacterial suspension, respectively.



Funded by the European Union

According to the standard curves, the analytical sensitivity of the qPCR test was 50 ag/µL, corresponding to 9.45 copies/µL, and 1.00 × $10^2$ CFU/mL, corresponding to 0.1 copies/µL, for plasmid DNA and bacterial cells, respectively.

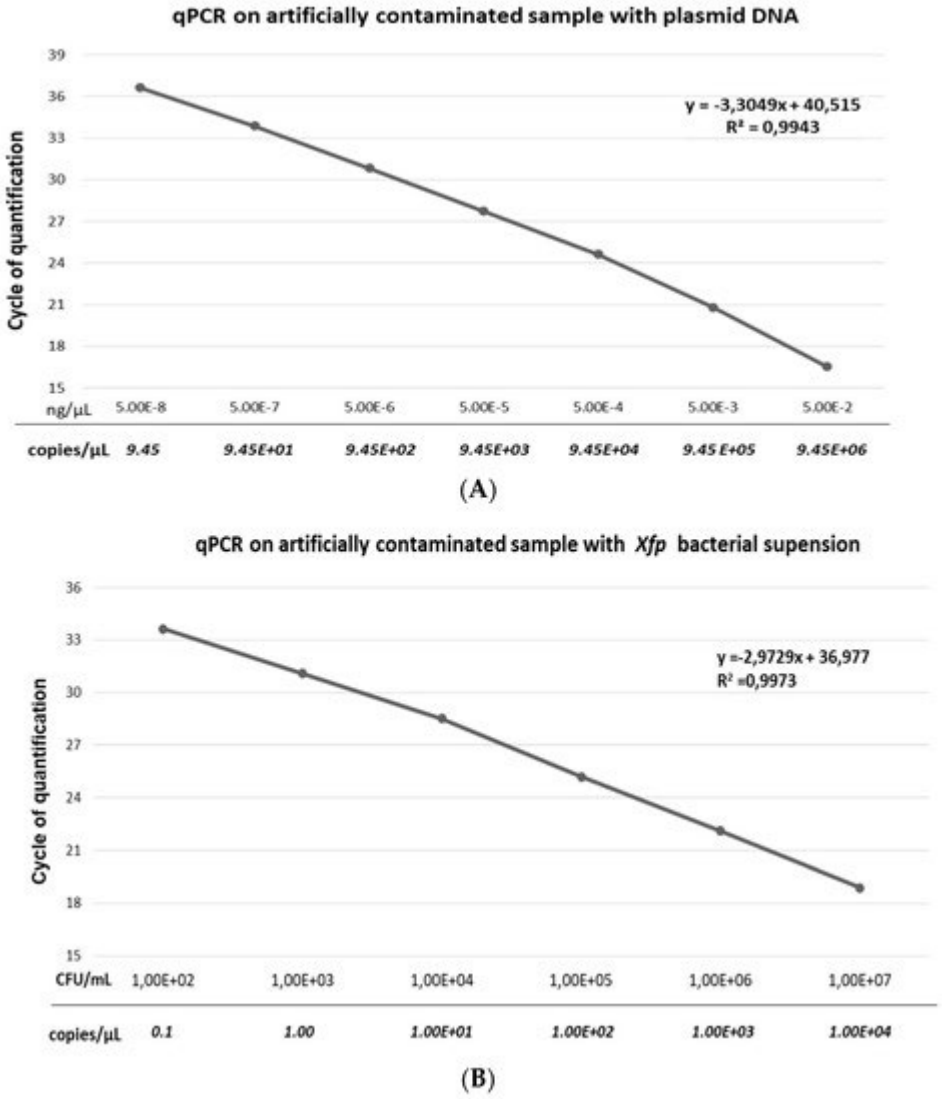



**Figure 2.** Standard curves generated by qPCR assay on *rim*M plasmid DNA (A) and bacterial suspension (B). Plasmid DNA was ten-fold serially diluted from 50 pg/µL to 50 ag/µL (9.45 × $10^6$–9.45 copies/µL). The bacterial suspension was ten-fold serially diluted from 1.0 × $10^7$ to 1.0 × $10^1$ CFU/mL (1.0 × $10^4$–0.1 copies/µL). On the y-axis: the cycle of quantification. (A) On the x-axis, ng/µL and copies/µL (italic and bold) are indicated. (B) On the x-axis, CFU/mL and copies/µL (italic and bold) are reported.

### 2.3.3 Linearity and Analytical Sensitivity of the ddPCR Assay

Ten-fold serial dilutions of rimM-recombinant plasmid DNA and *Xfp* bacterial suspension were used to construct the regression curves for the ddPCR assay. As shown in Figure 3A,B, the trend line of copies/µL obtained by ddPCR compared to our expectation indicated a high level of linearity in the evaluated range of concentrations, with an R2 value of 0.9981 for the


Funded by the European Union

bacterial suspension and 0.9928 for the plasmid DNA. The detection limits of the assay were determined to be $4.30 \times 10^{-1}$ copies/µL and $5.06 \times 10^{-1}$ copies/µL for the bacterial suspension and the plasmid, respectively, corresponding to $8.60 \times 10^{0}$ and $1.01 \times 10^{1}$ copies per reaction in a 20 µL volume. The correlation values between the expected and measured concentrations were significant (p-value < 0.00001 at $p < 0.05$), with $r = 1$ for both bacterial suspension and rimM plasmid DNA dilutions.

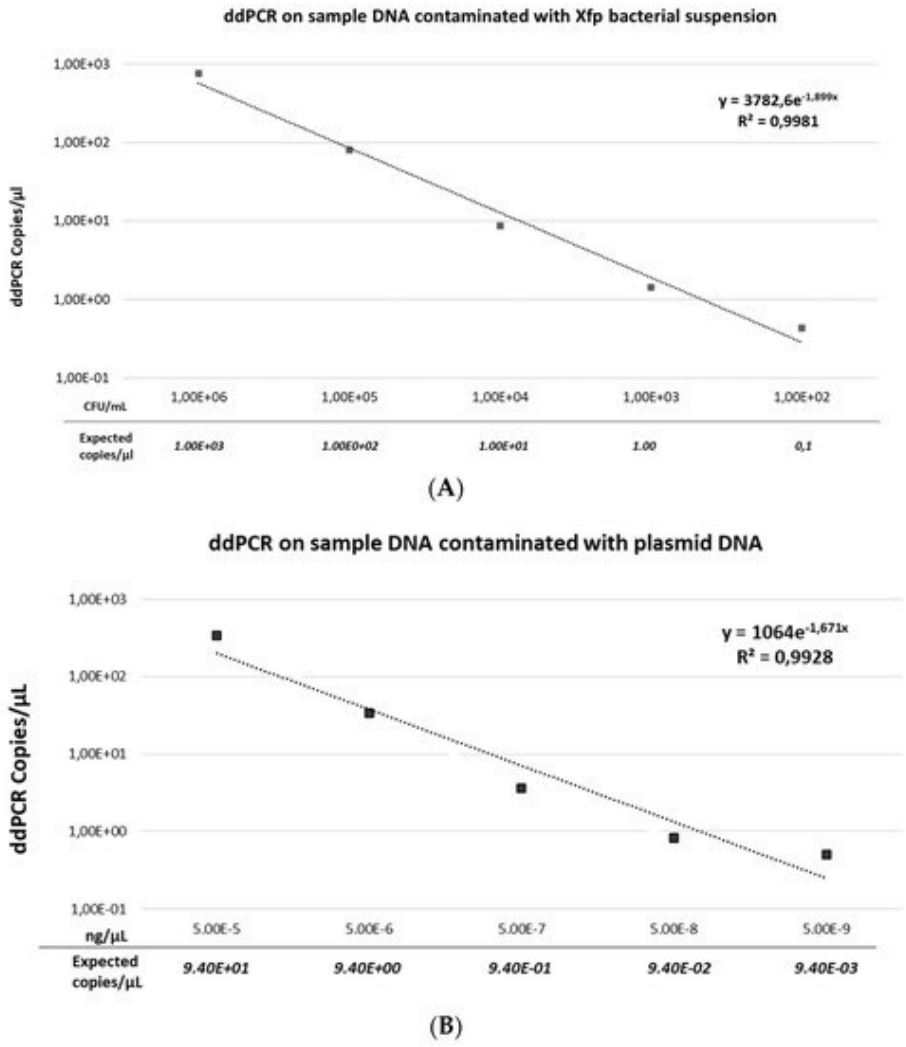



**Figure 3.** Linear regression of the ddPCR assays generated using the same 10-fold dilution series of bacterial suspension (A) and rimM plasmid DNA (B) tested with the qPCR assay. Number of copies/µL measured by ddPCR on the y-axis are correlated to (A) dilutions of bacterial suspension ranging from $1.0 \times 10^{6}$ to $1.0 \times 10^{1}$ CFU/mL, and (B) dilutions of plasmid ranging from 50 fg/µL to 5 ag/µL. Both figures indicated the corresponding expected copies/µL on the x-axis.



Funded by the European Union

### 2.3.4 Comparison between ddPCR and qPCR Assays in Artificially Contaminated Samples

The quantitative PCR and ddPCR assays carried out under the conditions implemented in this work exhibited optimal performance values, with high determination coefficients (Table 1 and Table 2). Based on the standard and regression curves generated by the qPCR and ddPCR assays (Supplementary Figures S2 and S3), the linearity, LOQ, and LOD of both assays were determined. As reported in Table 1 and Table 2, the qPCR and ddPCR assays exhibited good linearity for all the analyzed matrices, with (R2) ranging from 0.9905 (for almond) to 0.9995 (for insects) and from 0.9743 (for insect) to 0.9985 (for almond), respectively. The slope value of the standard curve generated by qPCR for each matrix ranged from −2.94 for oleander to −3.1993 for olive, corresponding to efficiency values of 118.84% and 105.38%, respectively. *Xfp* was correctly detected by qPCR in all three replicates/matrix, containing the lowest bacterial concentration, with LOQ and LOD being equal. The LOQ and LOD values recorded for ddPCR were identical for all plant matrices, while they differed in the case of the ddPCR assays on insects. More specifically, for the insects, LOQ corresponded to 10 CFU/mL (i.e., $9.22 \times 10^{-1}$ copies/µL), but at this concentration, the bacterium was detected in 4 out of 8 replicates. As such, $10^2$ CFU/mL (corresponding to 2.02 copies/µL) was considered to be the LOD, given that at this concentration, all 8 replicates tested positive. As shown in Table 3 and Table 4, the ddPCR assay showed generally higher analytical sensitivity compared to the qPCR test in the range of one order of magnitude, detecting *Xfp* in the replicates containing lower bacterial concentrations than qPCR for olive (up to 10 CFU/mL, corresponding to 1.2 copies/µL in 4 of 7 replicates), citrus (up to $10^2$ CFU/mL corresponding to 0.24 copies/µL in 9 of 10 matrices), and oleander (up to $10^2$ CFU/mL corresponding to 0.922 copies/µL in 7 of 8 replicates). Equivalent LOQ and LOD levels (102 CFU/mL) were recorded in *V. vinifera* and *P. dulcis*, corresponding in ddPCR to 4.74 × 10−1 and $3.92 \times 10^{-1}$ copies/µL, respectively. Indeed, *Xfp* was correctly detected in all replicates containing 102 CFU/mL in V. vinifera by qPCR (3/3) and ddPCR (12/12). However, in P. dulcis, the bacterium was detected by ddPCR in 6 of the 10 replicates containing 102 CFU/mL, and by qPCR in all 3 replicates containing $10^2$ CFU/mL. No positive droplets and no amplification curves were produced in the NTC or NIC for any of the matrices considered. The accuracy of the ddPCR assay when testing the spiked matrices reached 100%.

**Table 1.** Equation of regression curves and R2 values generated by ddPCR for the six matrices spiked with *Xfp* bacterial suspension.

| ddPCR | Curve Equation | $R^2$ |
|---|---|---|
| *Olea europea* | $y = 1635.4 \times 10^{-1.278x}$ | $R^2 = 0.9867$ |
| *Vitis vinifera* | $y = 6607.7 \times 10^{-1.922x}$ | $R^2 = 0.9885$ |
| *Citrus sinensis* | $y = 2220.3 \times 10^{-1.86x}$ | $R^2 = 0.9949$ |
| *Prunus dulcis* | $y = 4432.9 \times 10^{-1.854x}$ | $R^2 = 0.9985$ |
| *Nerium oleander* | $y = 653.8 \times 10^{-1.392x}$ | $R^2 = 0.9949$ |
| *Philaenus spumarius* | $y = 25{,}107 \times 10^{-1.814x}$ | $R^2 = 0.9743$ |



Funded by the European Union

**Table 2**. Equation of standard curves, R2 value, and efficiency generated by qPCR for the six matrices spiked with *Xfp* bacterial suspension.

| qPCR | Curve Equation | $R^2$ | Efficiency |
|---|---|---|---|
| *Olea europea* | y = −3.1993x + 38.963 | 0.9962 | 105.38% |
| *Vitis vinifera* | y = −3.1095x + 38.341 | 0.9987 | 109.70% |
| *Citrus sinensis* | y = −3.1376x + 38.847 | 0.9993 | 108.31% |
| *Prunus dulcis* | y = −3.0538x + 38.553 | 0.9905 | 112.55% |
| *Nerium oleander* | y = −2.94x + 37.563 | 0.9933 | 118.84% |
| *Philaenus spumarius* | y = −3.0427x + 38.005 | 0.9995 | 113.14% |

**Table 3.** Mean concentrations were estimated in copies/µL, as measured by ddPCR in each serial dilution of the spiked plant and insect matrices.

| Dilution Range | *Olea europea* | | *Vitis vinifera* | | *Citrus sinensis* | | *Prunus dulcis* | | *Nerium oleander* | | *Philaenus spumarius* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CFU/mL | copies/µL | Replicates * | copies/µL | Replicates * | copies/µL | Replicates * | copies/µL | Replicates * | copies/µL | Replicates * | copies/µL | Replicates * |
| $1.00 \times 10^6$ | $6.09 \times 10^2$ | 2/2 | $7.97 \times 10^2$ | 2/2 | $3.74 \times 10^2$ | 2/2 | $6.62 \times 10^2$ | 2/2 | $2.21 \times 10^2$ | 2/2 | $4.99 \times 10^3$ | 2/2 |
| $1.00 \times 10^5$ | $1.08 \times 10^2$ | 5/5 | $2.00 \times 10^2$ | 2/2 | $5.43 \times 10^1$ | 2/2 | $1.28 \times 10^2$ | 2/2 | $3.90 \times 10^1$ | 2/2 | $7.73 \times 10^2$ | 2/2 |
| $1.00 \times 10^4$ | $4.36 \times 10^1$ | 5/5 | $1.98 \times 10^1$ | 2/2 | $7.79 \times 10^0$ | 2/2 | $1.33 \times 10^1$ | 2/2 | $6.37 \times 10^0$ | 2/2 | $1.20 \times 10^2$ | 2/2 |
| $1.00 \times 10^3$ | $5.10 \times 10^0$ | 11/11 | $2.55 \times 10^0$ | 12/12 | $1.08 \times 10^0$ | 10/10 | $3.24 \times 10^0$ | 10/10 | $2.02 \times 10^0$ | 8/8 | $8.33 \times 10^0$ | 8/8 |
| $1.00 \times 10^2$ | $2.40 \times 10^0$ | 14/14 | $4.74 \times 10^{-1}$ | 12/12 | $2.43 \times 10^{-1}$ | 9/10 | $3.92 \times 10^{-1}$ | 6/10 | $9.22 \times 10^{-1}$ | 7/8 | $2.02 \times 10^0$ | 8/8 |
| $1.00 \times 10^1$ | $1.20 \times 10^0$ | 4/7 | n.d. ** | | n.d. ** | | n.d. ** | | n.d. ** | | $9.22 \times 10^{-1}$ | 4/8 |

* Replicates: number of positive replicates/number of replicates analyzed. ** n.d.: not detectable.

**Table 4.** Mean Cq values of the three replicates of 10-fold serial dilutions (CFU/mL) produced by qPCR Harper et al. (2010), modified for the spiked plant and insect matrices.

| Dilution Range CFU/mL | *Olea europea* | *Vitis vinifera* | *Citrus sinensis* | *Prunus dulcis* | *Nerium oleander* | *Philaenus spumarius* |
|---|---|---|---|---|---|---|
| $1.00 \times 10^6$ | 22.67 | 22.97 | 26.20 | 23.02 | 26.02 | 22.92 |
| $1.00 \times 10^5$ | 26.24 | 25.61 | 29.53 | 26.19 | 28.60 | 25.69 |
| $1.00 \times 10^4$ | 29.74 | 29.04 | 32.67 | 30.23 | 31.31 | 28.84 |
| $1.00 \times 10^3$ | 32.76 | 32.23 | 35.61 | 32.25 | 34.92 | 31.90 |
| $1.00 \times 10^2$ | 35.41 | 35.21 | n.d | 35.26 | n.d | 35.03 |
| $1.00 \times 10^1$ | n.d. | n.d | n.d | n.d | n.d | n.d |



Funded by the European Union

### 2.3.5 Evaluation of the ddPCR Assay in Field Samples

In this study, a comparison was carried out between the diagnostic performances of ddPCR and qPCR assays on naturally infected olive and insect samples. The DNA recovered from 34 olive and 27 insect samples was simultaneously tested by ddPCR and qPCR. Based on the Cq values generated in the qPCR assay, samples were clustered in three and two groups, respectively, for olives and insects. The olive samples were grouped as follows: (i) Cq from 27 to 29; (ii) Cq from 30 to 32; and (iii) Cq from 33 to 34. On the other hand, the insect samples were subdivided as follows: (i) Cq from 23 to 32 and (ii) Cq from 33 to 36. As shown in Table 5, all the olive samples belonging to the first and the second group were correctly identified as positive by ddPCR (19/19); while 13 out of 15 samples belonging to the third group (Cq 33–34), which had been classified as undermined by qPCR, were identified as positive by ddPCR. All insect samples included in the first group (11 specimens), yielding positive qPCR reactions, were also identified as positive by ddPCR (Table 6). Eight of the thirteen samples of the second group, classified as undetermined by qPCR, tested positive with ddPCR, two remained as undetermined samples, and three tested negative (Table 6). Moreover, all insects and olive samples collected in *Xf*-free areas, used as NICs, tested negative with both assays. NTCs included in all runs gave no amplification and no fluorescence signal. Therefore, the accuracy level reached by ddPCR was 100% in comparison with that of qPCR, with an increment level of approx. 38 and 30% to detect *Xfp* in olive and insect samples, respectively.

**Table 5.** Number of naturally infected olive samples assessed by qPCR assay as positive and undetermined (grouped in different ranges of the Cq values) compared with number of positive and negative samples detected by ddPCR assay.

| Olive Samples Grouped Based on the Cq Values Obtained by qPCR | | | | ddPCR | | |
|---|---|---|---|---|---|---|
| Group | Cq Values | N. of Samples | qPCR Result | Positive | Negative | Total |
| (i) | 27–29 | N. 10 | Positive | 10 | 0 | 10 |
| (ii) | 30–32 | N. 9 | Positive | 9 | 0 | 9 |
| (iii) | 33–34 | N. 15 | Undetermined | 13 | 2 | 15 |
| | | | Total | 32 | 2 | 34 |

**Table 6.** Performance of ddPCR and qPCR assays for detection of *Xfp* in naturally infected insects. Insect samples were divided into two groups according to the Cq value obtained by qPCR.

| Insect Samples Grouped Based on the Cq Values Obtained by qPCR | | | | ddPCR | | | |
|---|---|---|---|---|---|---|---|
| Group | Cq Values | N. of Samples | qPCR Result | Positive | Negative | Undetermined | Total |
| (i) | 23–32 | N. 11 | Positive. | 11 | 0 | 0 | 11 |
| (ii) | 33–36 | N. 13 | Undetermined | 8 | 3 | 2 | 13 |
| (iii) | >36 | N. 3 | Negative | 0 | 3 | 0 | 3 |
| | | | Total | 19 | 6 | 2 | 27 |



Funded by the European Union

# 3. NAKED-EYE MOLECULAR TESTING FOR THE DETECTION OF *XYLELLA FASTIDIOSA* IN ALMOND ORCHARDS BY COLORIMETRIC LAMP

## 3.1 Background

The loop-mediated isothermal amplification (Notomi et al. 2000) technique, already established for an extensive variety of pathogens (De Boer et al. 2012; Donoso et al. 2018), is extremely efficient and highly specific. Due to the auto-cycling strand displacement DNA synthesis and the tendency of the Bst polymerase enzyme to synthesize new DNA more quickly than conventional PCR at a constant temperature, reactions can be performed outdoors using only a portable, battery operated heat block or any other device able to maintain a single temperature for the relatively short time needed. The stability and abundance of the produced amplicons make optical readout options, such as colour changes or turbidimetric measurements, possible. Therefore, LAMP is an attractive alternative tool in resource-limited areas and in situations where time plays a crucial role in controlling the disease spread (Baldi et al. 2020). The present work describes an open-field colorimetric LAMP (cLAMP) application for *Xf* detection on trees in naturally *Xf*-infected almond orchards in Mallorca, employing only a stabilized reaction mix, unpurified buffered plant extracts as templates, and a simple portable microtube incubator. In addition, to evaluate the sensitivity and reliability of the method, the cLAMP results were compared with those obtained in laboratories using real-time quantitative PCR (qPCR assay) with the same crude extracts. This study aimed to demonstrate that the bulk detection of a reasonable percentage of infection in woody plants, like almond trees, up to the border of a pathogen titer where trees do not show yet any symptoms, could enable the cLAMP workflow as a preliminary screening tool in *Xf*-infected almond orchards for the timely support of management and eradication strategies.

## 3.2 Materials and Methods

In the framework of the European Union Horizon 2020 BIOVEXO project, during the summer of 2023, when the temperature ranged between 29 and 34 °C, a field survey was conducted in two orchards located on Mallorca Island, where *Xfm* (harboring the genotypes ST7 and ST81 (Moralejo et al. 2020; Greco et al. 2021) was detected. Figure 4A–C shows the most representative symptoms induced by *Xf* on almond trees.

During the field survey in Mallorca (Balearic Islands) in July 2023, a total of 53 almond trees were randomly sampled from the two orchards having different planting ages and plantation plots (Table 7). All the samples used in the experiment consisted of chips from semi-lignified shoots. In detail, five young shoots per tree were collected from the basal branches, sometimes close to ALS symptoms when these were evident, but not yet senescent or desiccated. Plant material was stored in plastic bags and put in a refrigerated box with melting ice until processing.



Funded by the European Union

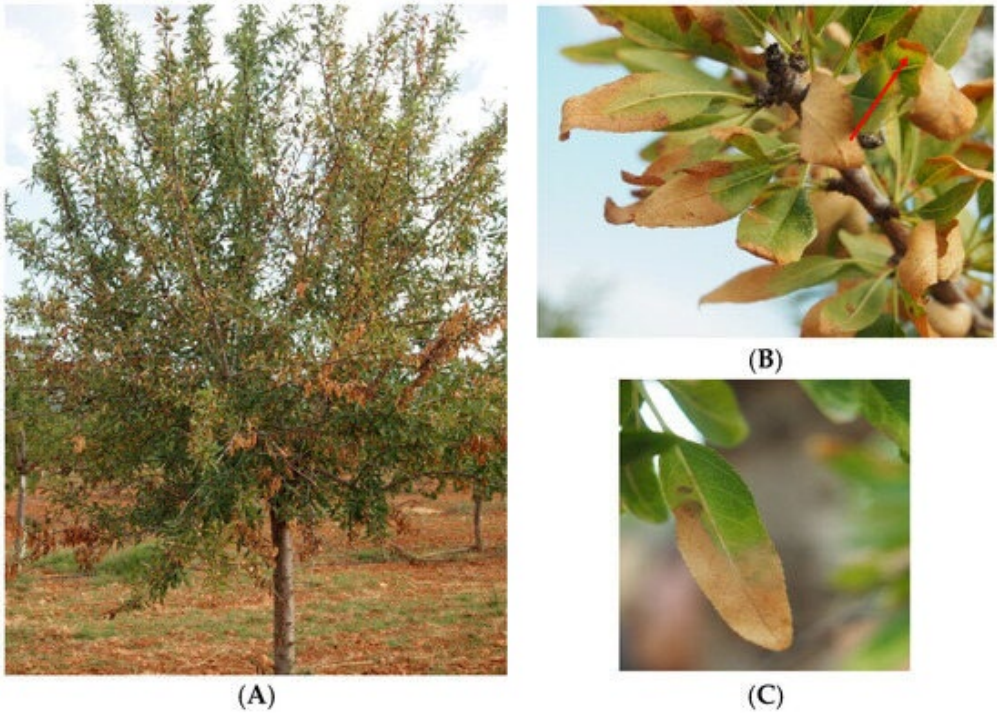



**Figure 4.** (A) Infected almond tree in Mallorca orchard with 'golden death' symptoms; (B) scorch symptoms on almonds with the typical yellow band, marked with a red arrow; and (C) on the leaf tissue between the green and dead areas.

Small slices (0.2 cm thickness) of cortical tissues (up to 250 mg in total) were cut with a razor blade, disinfected after chopping each sample with diluted commercial bleach (ca. 1% active chlorine), and placed inside a mesh plastic bag (Agdia, Elkhart, IN, USA) in the presence of 1.5 mL sodium chloride–Tris–EDTA extraction buffer (STE 1x), containing also 2% DIECA (diethyldithiocarbamate), 1% PVP-40k (polyvinylpyrrolidone) and Triton 0.1%. Bags were manually crumbled with a hammer to facilitate the breaking of plant tissues and the release of bacterial cells in the extraction buffer (Schaad et al. 2002). The spilled liquid was recovered in a 1.5 mL centrifuge tube and stored on ice. One hundred µL of the decanted suspension was retrieved in a new labelled tube, avoiding the plant debris at the bottom. This crude sap aliquot was denatured by a 2 min incubation with an equal volume of 0.4 M NaOH, then subsequently neutralized by mixing for a short time with the same volume of HCl 0.2 M. The neutralized suspension was further diluted to 1:10 in distilled water and finally subjected to the cLAMP assay according to Amoia et al. (2023a). Crude sap from healthy almond tissues recovered from greenhouse-grown seedlings, previously tested by qPCR as being *Xf*-free, and nuclease-free water served as negative internal and no-template controls (NIC and NTC), respectively, for each cLAMP round. A standard 30 min incubation was performed at 65 °C in a portable BIOPIX DNA Technology battery-powered heat block (Papadakis et al. 2022). Moreover, to compare the field-performed cLAMP (Harper et al. 2010) test, the presence of *Xf* was then assessed by testing, in duplicate, the same almond crude extracts by qPCR in two different laboratories (IPSP-CNR, Bari and AIT, Vienna) involved with the project, following the TaqMan protocol described by Harper et al. (2010). One µL of the alkaline-denatured, neutralized and 1:10 water-diluted sap was added to the cLAMP and qPCR reactions. The primer sequences used in both tests that targeted the rimM gene, highly conserved among


Funded by the European Union

all *Xf* subspecies' genomes and widely validated for *Xf* detection (EPPO 2019), are reported in 7.

**Table 7.** List of collected almond samples analyzed by dual molecular assays. Comparative analysis of cLAMP and qPCR (Cq) results for the detection of *Xf* in the almond samples used in the study. The Cq values reported are the average obtained by the two laboratories.

| Sample No. | Average Cq Values | cLAMP Result |
|---|---|---|
| 1 | 20.43 | + |
| 2 | 24.71 | + |
| 3 | 24.27 | + |
| 4 | 24.64 | + |
| 5 | 25.74 | + |
| 6 | 25.96 | + |
| 7 | 25.65 | + |
| 8 | 25.89 | + |
| 9 | 26.49 | + |
| 10 | 26.30 | – |
| 11 | 26.89 | – |
| 12 | 26.54 | + |
| 13 | 26.03 | + |
| 14 | 26.33 | – |
| 15 | 26.58 | + |
| 16 | 26.44 | – |
| 17 | 26.80 | + |
| 18 | 27.14 | – |
| 19 | 27.20 | + |
| 20 | 27.21 | + |
| 21 | 27.16 | – |
| 22 | 27.49 | – |
| 23 | 28.53 | + |
| 24 | 28.97 | – |
| 25 | 28.99 | – |
| 26 | 28.14 | – |
| 27 | 29.62 | – |
| 28 | 29.44 | – |
| 29 | 29.90 | – |
| 30 | 29.96 | + |
| 31 | 29.18 | – |
| 32 | 30.17 | + |
| 33 | 30.30 | – |
| 34 | 31.39 | – |
| 35 | 31.88 | – |
| 36 | 32.99 | – |
| 37 | 32.07 | – |
| 38 | 32.01 | – |
| 39 | 32.01 | – |
| 40 | 32.62 | + |
| 41 | 32.60 | – |
| 42 | 33.87 | – |
| 43 | 33.75 | – |
| 44 | 34.13 | – |
| 45 | 35.92 | – |
| 46 | 35.83 | – |
| 47 | N/A | – |
| 48 | N/A | – |
| 49 | N/A | – |
| 50 | N/A | – |
| 51 | N/A | + |
| 52 | N/A | – |
| 53 | N/A | – |

Cq value > 33 is considered negative. N/A: no absorbance. Successful amplification in cLAMP is indicated with "+", whereas "–" indicates a negative result.



Funded by the European Union

To evaluate the limit of detection (LoD; i.e., the lowest copy number reliably detected of the *Xf* genome) of the cLAMP assay for the amplification of the *Xf* target DNA in almond crude extracts, 6-days-grown bacterial cells of *Xfm* strain ESVL ST6 on PD3 agar plates (Giampetruzzi, A. et al. 2019; Marco-Noales, E.et al. 2021) were employed. The cell suspension was 10-fold serially diluted up to 10 CFU/mL (starting from an $OD_{600}$ value = 0.5, equivalent to 4 × $10^8$ CFU/mL; (D'Attoma, G.et al. 2019)) in healthy crude almond sap and then subjected to alkaline and neutralization treatments before the test, as specified above. NIC and NTC were also included. The formation of amplicon DNA was assessed by observing the color shift from pink (negative) to yellow (positive) after 30 min of incubation at 65 °C, due to the presence of phenol red, a pH indicator present in the master mix. The cLAMP assay was performed with a final volume of 25 µL, including, as follows: 12.5 µL Bst 2.0 WarmStart® DNA Polymerase 2x Master Mix (New England Biolabs, Ipswich, MA, USA); 2.5 µL LAMP primer mix (0.2, 1.6 and 0.4 µM of F3/B3, FIP/BIP and LF/LB primers, respectively); 1 µL of neutralized and diluted crude DNA extract and double-distilled water to a final volume; and 25 µL of mineral oil, added to each tube to cover the LAMP mix to avoid possible evaporation and contaminations. The consistency of the reaction for some tubes was also evaluated using 1.2% agarose gel electrophoresis run in Tris–acetate–EDTA (TAE 1x) buffer and stained with GelRed. The performance criteria, such as the diagnostic sensitivity (proportion of infected samples testing positive compared with the number of positive results from an alternative test, considered to be the gold standard) and specificity (calculated as the ratio of samples testing negative compared with the results from the alternative test) of the assay, were calculated according to EPPO PM 7/76 (EPPO 2018). A statistical comparison of correspondence between the results of positive and negative samples in both tests was performed by Fisher's exact test on a contingency table. For all the tested samples (positive and negative) in cLAMP and qPCR, respectively, receiver operating characteristic curve (ROC) analyses were calculated by plotting the true positive rate (sensitivity) and the false positive rate (1-specificity) with a 95% confidence interval (CI). The area-under-the-curve (AUC) values were also calculated to evaluate the performance of both assays.

## 3.3 Results

A rough extraction of nucleic acids was adopted by using a simple, homemade buffer. The preliminary homogenization step in STE extraction buffer, before the NaOH-based denaturation, differently from what was reported in (Amoia et al. 2023a) for the olive tissues, was introduced to strongly reduce the quick polyphenols oxidation occurring in almond sap, when directly extracted by a hefty alkaline buffer. The sensitivity of the cLAMP assay showed that the LoD of the assay was 100 CFU/mL (Figure 5A, reaction tube 5), as was previously demonstrated in crude olive sap for *Xfp* ST53 using the same amplification conditions (Amoia et al. 2023a). The qPCR detection limit on the same crude extracts was 10 CFU/mL, thus demonstrating a sensitivity that was an order of magnitude higher than that of the cLAMP assay. The cLAMP results were verified by observation of the amplification products in agarose gel electrophoresis, where a 180 bp-band was more evident and a ladder of larger-sized faint bands were visible in cases of positive amplification (Figure 5B, lanes 1 to 5).



Funded by the European Union

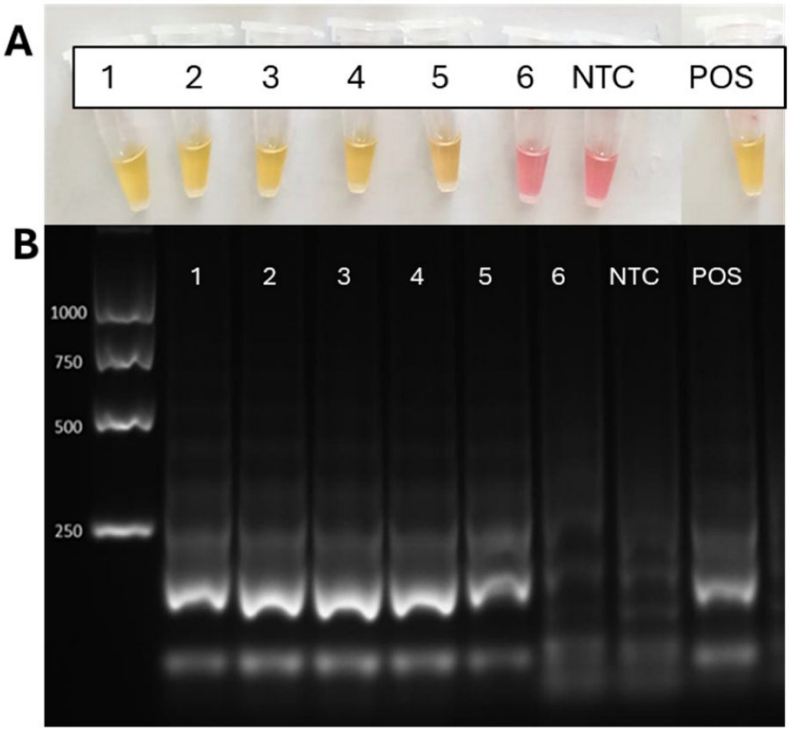



**Figure 5.** (A) cLAMP reactions from healthy almond sap spiked with 10-fold serial dilutions of *Xfm* ESVL cells, from $10^6$ to 10 CFU/mL (1–6; NTC: no template control; Pos: positive control); and (B) amplified products of each dilution analysed by gel electrophoresis.

Concerning specificity, the primer sequences used in this experiment were derived from a single-copy gene in the *Xf* genome and are well-known to be highly specific for all the *Xf* subspecies only EPPO PM 7/24 (EPPO 2023). The same amplification procedures did not show any cross-reactions with some tested non-target bacteria such as different *Xanthomonas* spp. or other common, non-correlated, epiphytic, and endophytic bacteria (Amoia et al. 2023a).

The crude sap extracted from field-grown almond samples was tested by either cLAMP or qPCR assays. Results obtained from the tested trees for both methods are presented in Table 7. The trees were classified according to their Cq value generated in qPCR. The average Cq produced by the analysis of the two laboratories, CNR-IPSP and AIT, was considered. As reported in Table 8, 41 out of 53 samples tested positive in qPCR (77.35%); the remaining 12 tested negative (22.64%). Notably, in both laboratories, a real-time qPCR result is considered negative or undetermined when Cq is > 33, as these are borderline values close to the detection limit of the technique, due to the uneven and erratic distribution of the bacterium in host tissues (Amoia et al. 2023b). Overall, the presence of the bacterium was properly determined by cLAMP in 19 out of 41 samples that tested positive using qPCR (i.e., 46.34% along with all the 'true positive' Cq values up to 33). For the samples showing negative results in qPCR, the cLAMP assay, correctly, did not exhibit any amplification signal in 11 out of 12 samples, except for sample no. 51, which behaved as a false positive sample. Therefore, the diagnostic specificity corresponds to 92.30%. The cLAMP AUC value (0.5938), illustrating the diagnostic ability and connection between diagnostic sensitivity and specificity, properly confirmed this concept with respect to the qPCR AUC number equal to 0.8750 (Supplementary Figure S5).



Funded by the European Union

**Table 8.** Comparative performance of cLAMP vs. qPCR assay for the detection of *Xf* in almond alkaline crude extracts. Positive samples in cLAMP are indicated for each average Cq group obtained from qPCR assay. Percentage of diagnostic sensitivity and specificity generated by cLAMP assay are also reported.

| qPCR Value | cLAMP Positives/Nr Analyzed Plants (Per Cq Value) | Diagnostic Sensitivity % | Diagnostic Specificity % |
|---|---|---|---|
| 18–24 | 4/4 | 100 | / |
| 25 | 4/4 | 100 | / |
| 26 | 5/9 | 55.55 | / |
| 27 | 2/5 | 40 | / |
| 28 | 1/4 | 25 | / |
| 29 | 1/5 | 20 | / |
| 30–33 | 2/10 | 20 | / |
| 33.01-NA | 1/12 | | 92.30 |
| Total | 20/53 | | |

# 4. MODELING THE ACCURACY OF *XYLELLA FASTIDIOSA* MOLECULAR DIAGNOSTIC TESTS IN NATURALLY INFECTED ALMOND TREE SAMPLES

## 4.1 Background

Many of the qPCR tests used for *Xf* diagnosis have been validated by different laboratories across Europe in the framework of several research initiatives, including the H2020 *XF*-ACTORS and Euphresco PROMODE projects. These initiatives have facilitated test performance studies (TPS) and proficiency tests (PT), which harmonized diagnostic procedures for the effective identification of *Xf* (Giovani et al. 2019). They also validated DNA extraction methods from host plants and insect vectors, followed by *Xf* detection via qPCR (Saponari and Loconsole 2020), and examined the influence of reagents in performing tetraplex real-time PCR (Pucci et al. 2023). However, these interlaboratory validation tests were performed using spiked samples (EPPO- Database on Diagnostic Expertise [https://dc.eppo.int/validation_data/validationlist] Roi et al. 2010). In contrast, only a few studies have addressed the performance of some of these tests in parallel using naturally infected samples where the DNA matrix may be more complex and *Xf* concentrations are unknown (e.g., Anguita-Maeso et al. 2022; Dupas et al. 2019a, 2019b; Loconsole et al. 2021).

The most common diagnostic parameters used to evaluate the performance of detection tests are sensitivity, specificity, accuracy, and positive and negative likelihood ratios (Vučurović et al. 2022). These diagnostic parameters can be calculated when the true infection status of the population under study is known. In cases where this is not possible, the suitability of the detection methods is often estimated by comparison with a gold standard technique (Olmos et al. 2016). However, the absence of a perfect gold standard test can lead to the misclassification of the true infection status (Olmos et al. 2016). When a gold standard is


Funded by the European Union

unavailable, multiple tests can be evaluated on a set of known samples using the same DNA extracts (Chikere et al. 2019). In such cases, statistical approaches like Bayesian Latent Class Models (LCMs), can be useful for assessing tests agreement (Olmos et al. 2016; Turechek et al. 2013). LCMs combine the results of multiple detection tests to estimate the prevalence, sensitivity and specificity of a method, even in the absence of a gold standard test (Olmos et al. 2016; Turechek et al. 2013).

In this study, we conducted a comprehensive evaluation of several diagnostic parameters for six molecular detection tests for *Xf*, including four qPCR tests (Francis-qPCR, Harper-qPCR, Li-qPCR, and Ouyang-qPCR), a ddPCR test (Dupas et al. 2019b), and an isothermal amplification test based on RPA (Cesbron et al. 2023), to detect *Xf* in naturally infected almond trees. Additionally, we assessed the impact of the plant sample source (leaf petioles versus woody chips from xylem tissues) on the performance of these tests. Finally, using a Bayesian LCM approach, we identified the most suitable and reliable molecular detection methods for *Xf* based on the sampled tissue. These findings are significant as they may contribute to the selection of the most effective diagnostic protocols for *Xf* in *Prunus dulcis* and may also serve as a model for other deciduous species, representing a crucial step in combating the spread of this harmful plant pathogen.

## 4.2 Materials and Methods

### 4.2.1 Plant material

A total of 88 naturally infected almond (*P. dulcis*) tree samples were used in this study. Plant samples were collected from various municipalities within the *Xf* DA of Alicante, Valencian Community, Spain, during a survey conducted in July 2018 (Tender for phytosanitary evaluation of trees in Alicante Ref. Ares (2018) 2552442-16/05/2018). A detailed description of the field plots and sampled trees is available in previous studies (Anguita-Maeso et al. 2022; Camino et al. 2022; Zarco-Tejada et al. 2021). For each tree, DNA was extracted from leaf petioles and bark, both derived from the same branches.

### 4.2.2 Visual assessment of ALS symptoms

Almond trees were visually inspected for the presence of ALS symptoms using a Disease Severity (DS) rating scale ranging from 0 to 5, based on the percentage of the tree crown showing ALS symptoms (where 0 refers to no visual symptoms (i.e., asymptomatic); 1= 1-25%; 2= 26–50%, 3= 51–75%; 4= 76–95%, and 5 corresponds to a tree with mostly dead branches) (Anguita-Maeso et al. 2022; Zarco-Tejada et al. 2021).

### 4.2.3 DNA extraction

In deciduous species such as *Prunus* spp., asymptomatic leaves collected early in the season may not contain detectable levels of *Xf*. However, by late summer, detectable levels are typically present in the leaf petioles of symptomatic plants (EPPO 2023; Zecharia et al. 2024). Therefore, in this study, DNA was extracted from both leaf petioles and woody chips. For the leaf petioles, small pieces of midribs and basal leaf parts were taken; while to prepare the


Funded by the European Union

woody chips, the stems were debarked, and the xylem woody chips were obtained by scraping most of the external layer using a sterile scalpel. A total of 0.5 g of woody chips or leaf petioles were used for DNA extraction using the cetyltrimethylammonium bromide (CTAB) method described in the EPPO protocol (2023). Additionally, leaf petioles and xylem tissues from a sample from a *Xf*-free P. dulcis tree were used to extract DNA and serve as the Negative isolation control (NIC). The NIC was also used as host DNA background in the standard curves used to determine the LoD. Extracted DNA was kept in 1.5-ml Eppendorf safe-lock tubes at -20°C until use. All molecular tests were performed within 1.5 year of DNA extraction.

### 4.2.4 DNA standard curves

To determine the LoD for each test for detecting *Xf* in plant DNA samples, two independent sets of DNA standard curves were obtained using 10-fold dilution series of *Xf* DNA. For that, *Xf* subsp. multiplex strain ESVL (Giampetruzzi et al. 2019) was cultured in PD2 for 7 days (Davis et al. 1980). DNA was extracted using the DNeasy Plant Mini Kit (Qiagen). Pathogen and host DNA concentrations were quantified in triplicate using a spectrofluorometric method with the Qubit dsDNA HS kit (Thermo Fisher Scientific, MA, USA) to ensure accuracy. Then, extracted DNA was adjusted and serially diluted in a fixed background of almond DNA (15 ng/µl), to obtain two independent DNA standard curves (DNA Series I-Ab and II-Ab) with final concentrations starting from 3 ng/µl *Xf*-DNA to 3 x$10^{-6}$ ng/µl (DNA Series I-Ab) or 2.8 ng/µl *Xf*-DNA to 2.8 x $10^{-6}$ ng/µl (DNA Series II-Ab). This DNA concentration corresponded to a concentration of approximately 1 × $10^{6}$ to 1 cells/µl based on the estimated genome size of the strain ESVL including its two plasmids (2,613,641 bp) and knowing that 1 pg = 9.78 × $10^{8}$ bp (Doležel et al. 2003). To evaluate the potential influence of PCR inhibitors in almond DNA, an additional DNA standard curve was obtained by spiking the same *Xf* DNA dilutions of DNA Series I into a water background (DNA Series I-Wb). Analytical sensitivity, linearity, and dynamic range of the different assays were assessed by testing two or three replicates of each DNA concentration simultaneously in the same run, depending on the molecular test used, with the only exception of the RPA test, for which a curve was tested for each background independently due to the equipment limitations that allowed amplification of only eight samples at a time. Each standard DNA curve always included an almond NIC and water as negative amplification control (NAC).

### 4.2.5 Quantitative real-time polymerase chain reaction assays (q-PCR)

All four qPCR tests were performed on a LightCycler® 480 thermal cycler (Roche Applied Science), using clear 96-well plates (White, Landgraaf, Netherlands). Each qPCR test was performed as described by EPPO (EPPO 2023) with the exception that the iTaq™ Universal Probes Supermix (BioRad) was used in all the tests, and BSA (Molecular Biology Grade, New England Biolabs) was added to a final concentration of 0.3 µg/µl. A total of 2 µl of DNA (from each dilution series of the DNA standard curves or from naturally infected almond plant samples) was used in each qPCR replicate.

In each qPCR plate, two dilutions (1/1 and 1/10) of genomic amplified DNA (GenomiPhiTM DNA Amplification Kit; Amersham Biosciences Co. UK) of *Xf* subsp. *sandyi* at a concentration of 3 ng/µl were used as positive amplification control (PAC), while sterile PCR-grade water and


Funded by the European Union

almond DNA known to be *Xf*-free were used as NACs. Each sample was amplified in duplicate on the same PCR run.

The data acquisition and data analyses were performed using LightCycler® 480 Software, v1.5. The determination of the quantification cycle (Cq) values for each reaction was calculated first by determining the PCR cycle number at which the fluorescence signal exceeded background using the Absolute Quantification-Fit Points method and the noise band was used as threshold line. Cq values up to 35 were considered as a positive result, while a Cq value above 38 was considered a negative result for *Xf*. A Cq between 35 and 38 was considered undetermined (EPPO, 2023). Those undetermined samples were repeated and considered positive if amplification was detected in more than 50% of all PCR wells.

### 4.2.6 Recombinase-polymerase amplification assay (RPA)

The RPA assay was performed using the AmplifyRP® XRT+ kit (Agdia® Inc. Cat# XCS 34501/0048) for *Xf*. The processing of plant samples from naturally infected almond trees was performed at the same time as the CTAB DNA extraction. A mix of 0.3 g of woody chips or leaf petioles was macerated in the extraction bag (*Xf* AmplifyRP® XRT+) and homogenized with 3.0 mL of AMP1 buffer. For maximum sensitivity, the plant extracts were incubated for 30 min at RT before testing. Then, 10 µl of each sample extract were mixed with 1 mL of PD1 solution in a microcentrifuge tube and mixed well. Next, 25 µl of this mix were transferred into the reaction pellet tubes and incubated for 20 min in the AmplifyRP® portable heat block (Agdia1 Inc., Cat#ACC 00150) at 39˚C, according to manufacturer's instructions for sample amplification. The probe and primer sequences of this test are protected by intellectual property (Li et al. 2016). To compare the LoD of the RPA with that of the other tests and if the presence of a background DNA influenced *Xf* detection, 2 µl of each dilution of the DNA standard curves in water or almond DNA background were added directly to the reaction pellet tubes and processed as described above. Results were considered positive or negative according to the output provided by the APF device software. Additionally, the amplification curve files obtained from each DNA dilution series were exported from the APD device to estimate, by interpolation, the Cq at which the signal exceed the fluorescence default threshold value set by the manufacturer.

### 4.2.7 Droplet digital polymerase chain reaction (ddPCR)

A ddPCR reaction mix was prepared using primers and probes from Harper's qPCR test (Harper et al. 2010) according to Dupas et al. (2019b) with minor modifications. Amplification reactions were performed in a final volume of 20 µl containing: 1× ddPCR™ Supermix for Probes (No dUTP) (Bio-Rad), 0.2 µM forward and reverse primers (*XF*-F and *XF*-R) and 0.6 µM 6'FAM/BHQ-1 labelled probe (*XF*-P), and 2 µl of DNA samples to be tested (from each dilution series of the DNA standard curves or from almond plant samples). Positive and negative amplification controls were similar to those used in qPCR assays. At the time we performed the ddPCR assays, the study by Amoia et al. (2023b) describing the optimization of the ddPCR conditions for *P. dulcis* was not available; so, it could not be implemented in our study.

Droplets were generated with the QX200™ Droplet Digital™ System (Bio-Rad) as recommended by the manufacturer and the ddPCRs were performed on a T-100TM Thermo



Funded by the European Union

cycler (Bio-Rad) with the following amplification conditions: DNA polymerase activation of 95°C for 10 min, then 45 cycles of two-steps of 94°C for 30 s for denaturation and 62°C for 60 s for hybridization and elongation, followed by a final step at 98 °C for 10 min for droplet stabilization. After amplification, the ddPCR plate was directly transferred to the QX200™ Droplet Digital™ System (Bio-Rad) reader and the QuantaSoft Analysis Pro 1.0.596 software was used for data acquisition and analysis. Each DNA sample was amplified in duplicate in the same run. The threshold was manually set to differentiate droplets, categorizing them as positive (high level of fluorescence) or negative (low and constant level of fluorescence) (Dupas et al. 2019b; Lievens et al. 2016). PCR reactions that generated fewer than 10,000 droplets were excluded from the analysis. A sample was classified as positive if at least three positive droplets were detected, following manufacturer's recommendations.

### 4.2.8 Statistical Analysis

All analyses were performed using R software v 4.3.1 (R Core Team, 2023) unless otherwise stated. To appropriately compare and establish relationships between the DNA dilution series generated from the different qPCR tests, linear regressions were performed, modelling the natural logarithm of known concentrations of the *Xf* target DNA (Log(DNA)) versus the Cq values (qPCR and RPA assays) or positive event counts (ddPCR assay). These regressions were calculated for each DNA dilution series across the different levels of the two main experimental factors (molecular diagnostic test and background DNA). Significant differences in slopes were determined by testing the interaction terms involving Log(DNA) and the experimental factors using an ANOVA F-test. To compare intercepts, estimated marginal means (EMMs) of Cq values for specific Log(DNA) were calculated using the 'emmeans' package in R v1.10.5 (Lenth, 2024). For the qPCR tests, amplification efficiency (AE) was calculated from the slopes of the standard regression curves using the equation AE = 10(1/slope) – 1 (Higuchi et al. 1993; Waliullah et al. 2019).

For naturally infected samples, the frequency of infected plants detected by each molecular diagnostic test was analysed using generalized linear models with binomial distribution and a logit link function using the GENMOD procedure in SAS software (version 9.4; SAS Institute, North Carolina, USA). A likelihood ratio test was performed to determine the significant effects ($P<0.05$) of the molecular diagnostic test, the type of plant tissue used for DNA extraction, and their interaction on the frequency of *Xf* detection. The statistical significance at $P<0.05$ of the likelihood ratio was determined by a chi-square test. Pairwise comparison of least squares (LS) means at $P<0.05$ were performed to test for differences among molecular diagnostic tests or plant tissue type. The contrast statement was used to identify significant differences ($P<0.05$). Additionally, differences in Cq among qPCR tests and between DNA sources (leaf petioles or woody chips) were determined by standard analysis of variance (ANOVA) with a factorial design, using only those Cq values obtained for all qPCR tests and both DNA sources. To further assess relationships among assays, Pearson's correlations and linear regression were used to evaluate the agreement between the results of RPA kit and qPCR assays.

To compare the performance and agreement among the different assayed tests, Harper-qPCR was initially selected as the gold standard, as it is the most widely used *Xf* detection test in European diagnostic laboratories. Next, comparisons were made using the ddPCR test as the gold standard, since it uses the same primers and probes as the Harper-qPCR test. Finally,


Funded by the European Union

results were also compared using the Li-qPCR test as the gold standard due to its similar sensitivity and robustness to the Harper-qPCR test.

Agreement between pairs of detection tests was estimated using binary outputs (i.e., *Xf*-positive (1) or *Xf*-negative (0)). The prevalence and bias-adjusted kappa (PABAK) index (Byrt et al. 1993) was used to account for disease prevalence (when one category occurs much more frequently than the other) and correct for any potential bias in the tests. PABAK values were interpreted based on the benchmarks suggested by Landis and Koch (1977), where 0 is poor agreement, >0 to 0.2 is slight agreement, >0.2 to 0.4 is fair agreement, >0.4 to 0.6 is moderate agreement, >0.6 to 0.8 is substantial agreement, and >0.8 is almost perfect agreement. McNemar's test (McNemar 1947) was used to compare the diagnostic performance (assessing the existence of significant differences) of two tests when using the same DNA sample. McNemar's $\chi 2$ test and PABAK indices were estimated using the R package epiR v2.0.77 (Stevenson 2024).

Sensitivity, specificity, false positive and false negative rates, and Youden's Index (J) (Youden 1950) were estimated using the R platform package epiR v2.0.77 (Stevenson 2024). Youden's Index (J = Sensitivity + Specificity − 1), which ranges from -1 to 1, indicates the balance between sensitivity and specificity, with values closest to 1 representing both perfect specificity and sensitivity compared to the reference tests. Tests with a Youden's Index below 0.5 are generally considered unreliable (Youden 1950). Agreement plots, showing the agreement between pairs of detection tests (Bangdiwala et al. 2008), were generated using the R package vcd-R v1.4-13 (Meyer et al. 2024).

Finally, a latent class model (LCM) analysis was conducted to evaluate the performance of the three diagnostic protocols for *Xf* that showed the highest sensitivity: Harper-qPCR, Li-qPCR, and ddPCR. The LCM analysis aimed to integrate and compare the data from these protocols to identify the most effective diagnostic test. The optimal number of latent classes was determined based on the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). The LCM analysis was performed using the R package poLCA v1.6.0.1 (Polytomous Variable Latent Class Analysis) (Linzer and Lewis 2011).

## 4.3 Results

### 4.3.1 Reproducibility, efficiency and sensitivity of qPCR, RPA and ddPCR tests

All qPCR tests produced positive and reproducible amplifications up to a concentration of 28 or 30 fg/µl of *Xf* DNA, depending on the DNA dilution series used. Additionally, for the Li-qPCR and Harper-qPCR tests, some positive amplifications were obtained up to 2.8 and 3 fg/µl, depending on the DNA dilution series assessed, although those amplifications were less reproducible. Linear regressions of Log (*Xf* DNA) versus Cq values were obtained for each qPCR test and DNA dilution series (Fig. 6). Standard regression lines obtained for all combinations showed highly reproducible amplifications, high efficiency (AE = 79.45–112.02%) and high coefficients of determination (R2 = 0.988–0.999), which fall within the ideal range for real-time qPCR assays (Table 9). However, statistical comparisons of the regression lines indicated significant differences among intercepts and slopes ($P < 0.0001$) of the four qPCR tests for a given background DNA. Similar results were obtained for the two independent DNA dilution



Funded by the European Union

series evaluated (Table 9). Overall, Harper-qPCR and Li-qPCR tests showed significantly (P<0.0219) lower intercept and slope values compared to the Ouyang-qPCR and Francis-qPCR tests, indicating better efficiency in the PCR amplification process, higher sensitivity -at least one order of magnitude- and earlier detection of *Xf* DNA of the former tests (Table 9). On the other hand, the efficiency and reproducibility of amplifications for all four qPCR tests were not influenced by the presence of host DNA as indicated by the lack of statistically significant (P>0.05) differences between slopes and intercepts for a given qPCR test across backgrounds (Table 9; Fig. 6). No false positive reactions (i.e., no detection in negative control samples) were obtained for any of the DNA dilution series evaluated.

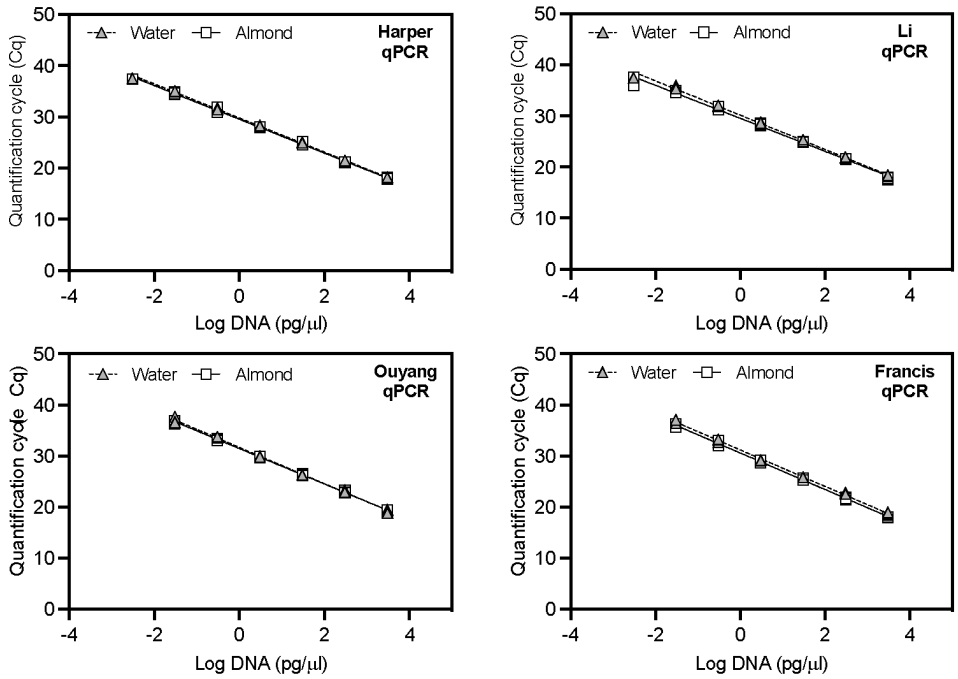



**Figure 6**. Quantification cycle (Cq) values obtained for Harper-qPCR, Li-qPCR, Ouyang-qPCR, and Francis-qPCR tests using a seven-point 10-fold serial dilution of Xylella fastidiosa subsp. multiplex strain ESVL DNA standard curve (3 ng/µl) diluted in almond DNA background (15 ng/ µl) (DNA Series I-Ab) or ultrapure sterile water (DNA Series I-Wb). Values from the linear regression analysis for each treatment are shown in Table 9.



Funded by the European Union

**Table 9.** Regression equation parameters from standard curves of *Xylella fastidiosa* subspecies multiplex strain ESVL DNA of four quantitative polymerase chain reaction (qPCR) assays evaluated for quantification of *X. fastidiosa* in ultrapure sterile water background or almond background.

| qPCR test | DNA standard series | DNA background | $R^2$ | Slope[a] | Intercept[b] | AE (%) |
|---|---|---|---|---|---|---|
| Harper | Series I-Wb | Water | 0.999 | -3.324 c | 29.80 c | 99.91 |
| Li | Series I-Wb | Water | 0.997 | -3.363 bc | 30.16 c | 98.31 |
| Ouyang | Series I-Wb | Water | 0.998 | -3.547 b | 31.68 a | 91.39 |
| Francis | Series I-Wb | Water | 0.998 | -3.578 a | 31.20 b | 90.32 |
| Harper | Series I-Ab | Almond | 0.998 | -3.296 bc | 29.54 c | 101.09 |
| Li | Series I-Ab | Almond | 0.997 | -3.240 c | 29.67 c | 100.50 |
| Ouyang | Series I-Ab | Almond | 0.999 | -3.470 ab | 31.45 a | 94.17 |
| Francis | Series I-Ab | Almond | 0.998 | -3.575 a | 30.58 b | 90.42 |
| Harper | Series II-Ab | Almond | 0.999 | -3.449 b | 30.97 c | 94.96 |
| Li | Series II-Ab | Almond | 0.997 | -3.490 b | 32.06 b | 93.43 |
| Ouyang | Series II-Ab | Almond | 0.997 | -3.938 a | 36.27 a | 79.45 |
| Francis | Series II-Ab | Almond | 0.988 | -3.064 c | 30.30 c | 112.02 |

[a] Significant differences between slopes within the same DNA standard series, as indicated by different letters were determined by testing the interaction terms involving Log(DNA) and the factors of the study using an ANOVA F-test at $P < 0.05$.

[b] Significant differences between intercepts within the same DNA standard series as indicated by different letters were estimated using the marginal means (EMMs) of Cq values for specific Log(DNA) at $P < 0.05$.

Results for the RPA tests indicated a lower sensitivity compared to qPCR, with a LoD that was at least one or two orders of magnitude higher. Additionally, the reproducibility of RPA results was lower, especially when almond DNA was included as background (Suppl. Fig. 6). Nevertheless, background DNA did not affect the AE or LoD, as indicated by the lack of significant differences (P>0.357) between slopes and intercepts of the regression lines for Log(*Xf* DNA) versus Cq values across DNA dilution series for the RPA tests. A strong correlation was observed between the RPA test and the four qPCR tests at *Xf* DNA concentrations ranging between 3 ng/µl and 3 pg/µl as indicated by the linear relationships between the Cq values of each qPCR test and the RPA Cq values (i.e., R2 > 0.965) (Suppl. Fig. 7).

The ddPCR test showed a strong linear response between the Log(*Xf* DNA) and Log(number of positive droplets +1) across DNA dilution series (Fig. 7). Positive and reproducible amplifications were obtained up to 300 fg/µl, with occasional amplification observed up to 30 fg/µl. These results indicate that the ddPCR test has lower sensitivity compared to the most sensitive qPCR tests (i.e., Harper-qPCR and Li-qPCR). The number of positive droplets amplified was not affected by the presence of almond DNA, as indicated by the lack of significant differences (P>0.357) between slopes and intercepts of the regression lines fitted for each background DNA.



Funded by the European Union

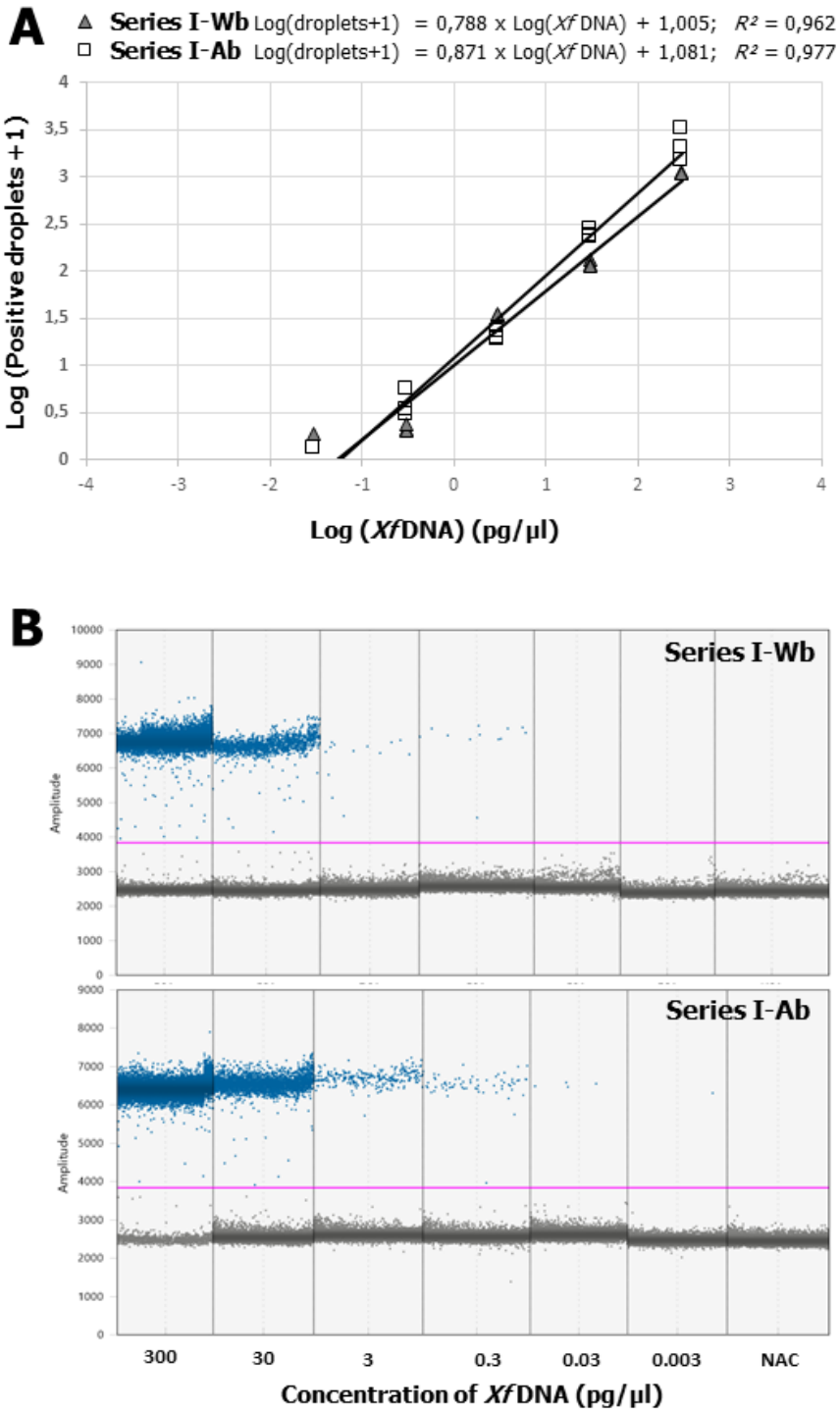



**Figure 7.** Limit of detection obtained by ddPCR of a six-point 10-fold serial dilution of *Xylella fastidiosa* (*Xf*) subsp. *multiplex* DNA standard curve (0.3 ng/µl) diluted in almond DNA background (15 ng/ µl) (DNA Series I-Ab) or ultrapure sterile water (DNA Series I-Wb). (A) Linear regression between the logarithm of the DNA concentration of *Xf* plotted against the logarithm of the number of positive droplets + 1 (y-axis). (B) Positive (in blue) and negative (in gray) droplets obtained by ddPCR in a representative assay. The pink line represents the threshold to differentiate between the negative and positive droplets. NAC corresponds to the negative amplification controls of almond DNA (Ab) or water (Wb).



Funded by the European Union

### 4.3.2 Detection of *Xf* in naturally infected almond samples

Of the 88 almond trees analysed, 46.6% were classified as *Xf*-positive by visual inspection, exhibiting typical ALS symptoms during the field sampling campaign (Zarco-Tejada et al. 2021) (Suppl. Table 9). The presence of *Xf* on the 88 trees was determined by four qPCR tests, RPA and ddPCR analysing DNA extracted from the same plant samples but from different plant tissues (leaf petioles or woody chips). The percentage of samples diagnosed as *Xf*-positive varied significantly depending on the origin of DNA (i.e., sample type: leaf petioles versus woody chips) (P<0.0001), the molecular test used (P<0.0001), and the interaction between these two factors (P<0.0001) (Fig. 8). All trees exhibiting ALS symptoms were diagnosed as *Xf*-positive by at least two molecular methods, except for one tree with an ALS severity =1 and another tree that was nearly dead (severity = 4). The number of asymptomatic trees that tested positive depended on the molecular test and tissue type. Across molecular tests, when DNA was extracted from leaf petioles, positive results ranged from 3 to 20 trees, whereas DNA extracted from woody chips yielded positive results for 4 to 44 trees (Suppl. Table 9).

Li-qPCR, Harper-qPCR and ddPCR detected a significantly higher (P<0.0122) number of *Xf*-positive samples when using woody chip tissues as compared to leaf petioles (Fig. 8). When using woody chip tissues, the number of *Xf*-infected trees detected was significantly higher (P<0.0122) for ddPCR (96.6%) and Harper-qPCR (93.2%), followed by Li-qPCR (79.6%). A similar trend was observed when using leaf petioles, although the number of infected trees detected was lower. On the other hand, Ouyang-qPCR, Francis-qPCR and the RPA tests were less efficient, detecting *Xf* in 46.6 to 53.4% of trees using leaf petioles and 51.1 to 54.6% when using woody chips. Interestingly, the number of trees diagnosed as *Xf*-positive by the Ouyang-qPCR, Francis-qPCR and RPA tests did not differ significantly (P>0.05) from the visual assessment irrespective of whether DNA was extracted from petioles or woody tissues (Fig. 8).

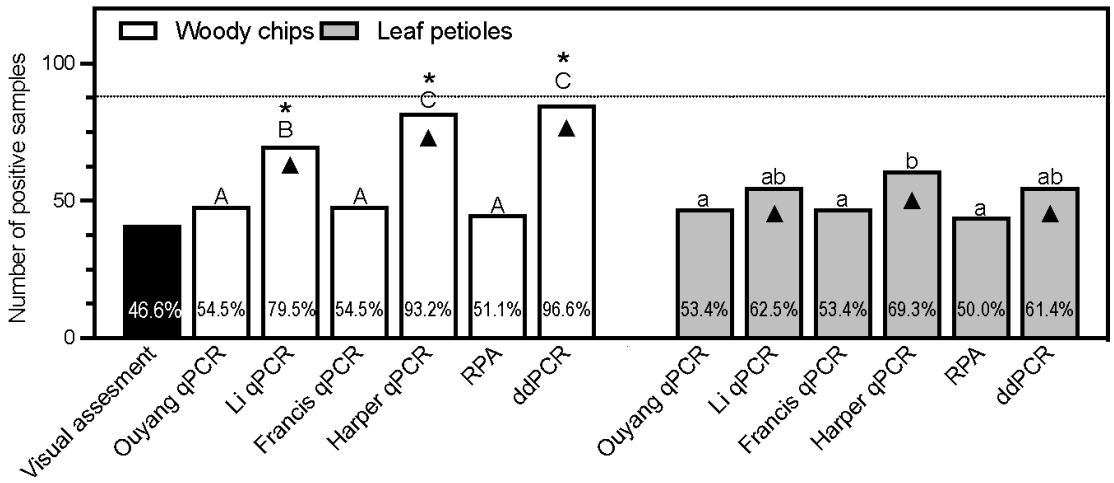



**Figure 8.** Number of *Xylella fastidiosa* (*Xf*) positive samples as determined by six molecular diagnostic tests when using different DNA sources (woody chips or leaf petioles). Values inside the bars indicate the percentage of positive samples (N=88). (*) Indicates significant (P<0.05) differences between the different DNA sources for a given molecular test. (▲) Indicates significant (P<0.05) differences in the number of trees considered infected by *Xf* as compared to the visual assessment. Protocols with same letters did not differ significantly (P<0.05) in the number of trees diagnosed as *Xf*-positive for the same DNA source.



Funded by the European Union

Cq values obtained from the four qPCR tests differed significantly depending on the molecular test used (P<0.0001), but not according to the DNA source (leaf petioles versus woody chips, P=0.1109). However, a significant interaction was detected between both factors (P<0.0001) (Fig. 9). Harper-qPCR and Li-qPCR exhibited the lowest Cq values for DNA extracted from woody chips. For DNA extracted from leaf petioles, Harper-qPCR showed significantly lower Cq values compared to the other three qPCR tests (Fig. 9).

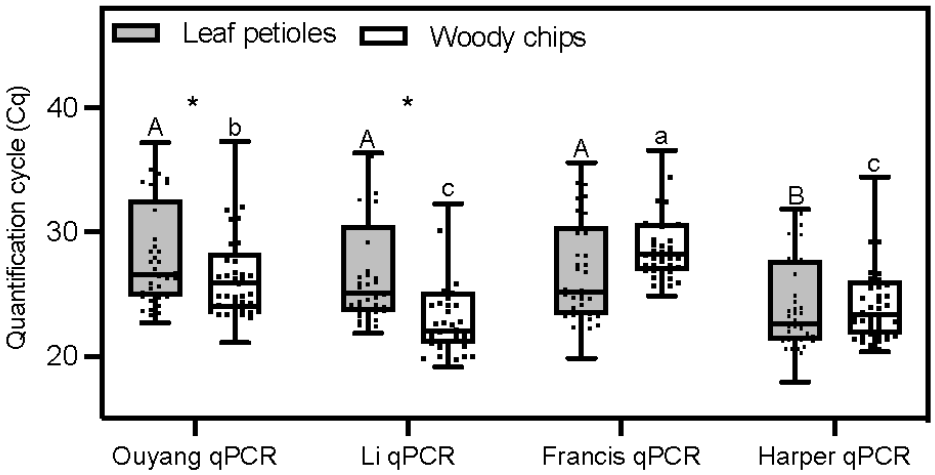



**Figure 9.** Quantification cycle (Cq) range of *Xylella fastidiosa* positives samples as determined by six molecular diagnostic tests when using different DNA sources (woody chips or leaf petioles). (*) Indicates significant (P<0.05) differences between the different DNA sources for a given molecular test. Protocols with same letters did not differ significantly (P<0.05) in the Cq values for the same DNA source. For statistical analysis, only almond samples amplifying with the four qPCR tests were used.

### 4.3.3 Performance and agreement between diagnostic tests

The diagnostic performance of the six molecular tests for detecting *Xf* infection in almond trees was evaluated using the three tests with the highest sensitivity (i.e., Li-qPCR, Harper qPCR, and ddPCR) as gold standards. Agreement plots for each pair of diagnostic tests are shown in Suppl. Figs. 8 and 9. The agreement between the diagnostic tests and the gold standards was assessed using various metrics, including the observed and expected proportion of agreement, PABAK and McNemar's $\chi^2$ tests (Suppl. Table 10). Overall, pairwise agreement reached better metrics when DNA was extracted from leaf petioles compared to DNA extracted from woody chips, with a few exceptions (Suppl. Table 10). McNemar's $\chi^2$ test was not significant (P>0.180) for the comparison between Harper-qPCR and ddPCR for both plant tissues, indicating no systematic disagreement between these tests, whereas this occurred only for Li-qPCR and ddPCR when testing leaf petioles. When comparing results obtained with Harper-qPCR with the other molecular tests analyzing DNA extracted from woody chip tissues, PABAK index indicated a slight agreement with RPA, fair agreement with Ouyang-qPCR and Francis-qPCR, substantial agreement with Li-qPCR and almost perfect agreement with ddPCR (Suppl. Table 10). On the other hand, substantial agreement was obtained between Harper-qPCR and all the tests when DNA was extracted from leaf petioles. For ddPCR, the PABAK index indicated high substantial agreement with Li-qPCR for both plant


Funded by the European Union

tissues. Finally, for Li-qPCR, PABAK index indicated a moderate agreement with Francis-qPCR, Ouyang-qPCR and RPA when using woody chips and substantial agreement, for leaf petioles (Suppl. Table 10).

**Table 10.** Diagnostic accuracy measures[a] and 95% confidence intervals of different molecular tests when compared to a reference standard for diagnosing *Xylella fastidiosa* in almond trees and when using DNA extracted from different plant sources

| Gold standard | DNA source | Diagnostic test | Sensitivity[a] | Specificity[a] | Youden's Index[a] | False T+ proportion for true D- | False T- proportion for true D+ |
|---|---|---|---|---|---|---|---|
| Harper | woody chips | Ouyang | 0.59 (0.47, 0.69) | 1.00 (0.54, 1.00) | 0.585 (0.01, 0.69) | 0.00 (0.00, 0.46) | 0.41 (0.31, 0.53) |
| | | Li | 0.83 (0.73, 0.90) | 0.67 (0.22, 0.96) | 0.496 (-0.05, 0.86) | 0.33 (0.04, 0.78) | 0.17 (0.10, 0.27) |
| | | Francis | 0.59 (0.47, 0.69) | 1.00 (0.54, 1.00) | 0.585 (0.012, 0.69) | 0.00 (0.00, 0.46) | 0.41 (0.31, 0.53) |
| | | RPA | 0.55 (0.43, 0.66) | 1.00 (0.54, 1.00) | 0.549 (-0.02, 0.66) | 0.00 (0.00, 0.46) | 0.45 (0.34, 0.57) |
| | | ddPCR | 0.99 (0.93, 1.00) | 0.33 (0.04, 0.78) | 0.321 (-0.002, 0.78) | 0.67 (0.22, 0.96) | 0.01 (0.00, 0.07) |
| | leaf petioles | Ouyang | 0.77 (0.65, 0.87) | 1.00 (0.87, 1.00) | 0.77 (0.52, 0.87) | 0.00 (0.00, 0.13) | 0.23 (0.13, 0.35) |
| | | Li | 0.90 (0.80, 0.96) | 1.00 (0.87, 1.00) | 0.901 (0.67, 0.96) | 0.00 (0.00, 0.13) | 0.10 (0.04, 0.20) |
| | | Francis | 0.77 (0.65, 0.87) | 1.00 (0.87, 1.00) | 0.77 (0.52, 0.87) | 0.00 (0.00, 0.13) | 0.23 (0.13, 0.35) |
| | | RPA | 0.72 (0.59, 0.83) | 1.00 (0.87, 1.00) | 0.77 (0.52, 0.87) | 0.00 (0.00, 0.13) | 0.28 (0.17, 0.41) |
| | | ddPCR | 0.77 (0.65, 0.87) | 0.70 (0.50, 0.86) | 0.474 (0.14, 0.73) | 0.30 (0.14, 0.50) | 0.23 (0.13, 0.35) |
| ddPCR | woody chips | Ouyang | 0.95 (0.88, 0.99) | 0.66 (0.09, 0.99) | 0.619 (-0.02, 0.98) | 0.33 (0.00, 0.90) | 0.05 (0.00, 0.12) |
| | | Harper | 0.97 (0.90, 0.99) | 0.22 (0.06, 0.47) | 0.194 (-0.04, 0.47) | 0.77 (0.52, 0.94) | 0.03 (0.03, 0.1) |
| | | Francis | 0.56 (0.45, 0.67) | 1.00 (0.29, 1.00) | 0.565 (-0.26, 0.67) | 0.00 (0.00, 0.71) | 0.44 (0.33, 0.55) |
| | | RPA | 0.53 (0.42, 0.64) | 1.00 (0.29, 1.00) | 0.529 (-0.29, 0.64) | 0.00 (0.00, 0.71) | 0.47 (0.36, 0.58) |
| | | Li | 0.82 (0.73, 0.89) | 1.00 (0.05, 0.39) | 0.823 (0.02, 0.90) | 0.00 (0.00, 0.71) | 0.17 (0.10, 0.27) |



Funded by the European Union

| Gold standard | DNA source | Diagnostic test | Sensitivity[a] | Specificity[a] | Youden's Index[a] | False T+ proportion for true D- | False T- proportion for true D+ |
|---|---|---|---|---|---|---|---|
| ddPCR | leaf petioles | Ouyang | 0.85 (0.73, 0.93) | 0.58 (0.39, 0.74) | 0.43 (0.13, 0.68) | 0.42 (0.25, 0.61) | 0.15 (0.06, 0.27) |
| | | Harper | 0.85 (0.73, 0.93) | 0.57 (0.39, 0.74) | 0.43 (0.13, 0.68) | 0.42 (0.25, 0.61) | 0.15 (0.06, 0.27) |
| | | Francis | 0.8 (0.67, 0.89) | 0.9 (0.76, 0.98) | 0.71 (0.43, 0.88) | 0.09 (0.01, 0.24) | 0.2 (0.10, 0.33) |
| | | RPA | 0.74 (0.61, 0.85) | 0.91 (0.75, 0.98) | 0.65 (0.37, 0.83) | 0.09 (0.01, 0.24) | 0.25 (0.15, 0.39) |
| | | Li | 0.82 (0.69, 0.91) | 0.70 (0.51, 0.84) | 0.515 (0.20, 0.75) | 0.30 (0.16, 0.49) | 0.18 (0.09, 0.31) |
| Li | woody chips | Ouyang | 0.69 (0.56, 0.79) | 1.00 (0.81, 1.00) | 0.685 (0.38, 0.79) | 0.00 (0.00, 0.19) | 0.31 (0.21, 0.44) |
| | | Harper | 0.97 (0.90, 0.99) | 0.22 (0.06, 0.47) | 0.194 (-0.04, 0.47) | 0.77 (0.52, 0.94) | 0.03 (0.03, 0.1) |
| | | Francis | 0.69 (0.56, 0.79) | 1.00 (0.81, 1.00) | 0.685 (0.38, 0.79) | 0.00 (0.00, 0.19) | 0.31 (0.21, 0.44) |
| | | RPA | 0.64 (0.52, 0.75) | 1.00 (0.81, 1.00) | 0.643 (0.33, 0.75) | 0.00 (0.00, 0.19) | 0.35 (0.24, 0.48) |
| | | ddPCR | 1 (0.95, 1.00) | 0.17 (0.05, 0.39) | 0.167 (-0.001, 0.39) | 0.833 (0.61, 0.94) | 0.00 (0.00, 0.05) |
| | leaf petioles | Ouyang | 0.75 (0.67, 0.82) | 0.98 (0.89, 1.00) | 0.77 (0.52, 0.87) | 0.02 (0.00, 0.01) | 0.25 (0.18, 0.33) |
| | | Harper | 0.86 (0.79, 0.91) | 0.94 (0.8, 1.00) | 0.8 (0.59, 0.91) | 0.06 (0.01, 0.20) | 0.14 (0.08, 0.21) |
| | | Francis | 0.75 (0.67, 0.82) | 0.98 (0.89, 1.00) | 0.77 (0.52, 0.87) | 0.02 (0.00, 0.01) | 0.25 (0.18, 0.33) |
| | | RPA | 0.78 (0.65, 0.88) | 0.97 (0.85, 1.00) | 0.75 (0.49, 0.88) | 0.03 (0.00, 0.16) | 0.22 (0.18, 0.35) |
| | | ddPCR | 0.82 (0.69, 0.91) | 0.70 (0.51, 0.84) | 0.515 (0.20, 0.75) | 0.30 (0.16, 0.49) | 0.18 (0.09, 0.31) |

[a] Sensitivity = TP/(TP+FN), Specificity = TN/(TN+FP) and Youden's Index (J) = Sensitivity + Specificity − 1); where TP and TN represent the true positive and true negative, respectively, and FP and FN represent the false positive and false negative, respectively.

### 4.3.4 Diagnostic accuracy of diagnostic tests

The diagnostic sensitivity and specificity were used to determine the ability of each test to correctly detect the presence or absence of *Xf* using three different gold standard tests for



Funded by the European Union

comparison (Table 10). When Harper-qPCR was the reference test, only Ouyang-qPCR, Francis-qPCR and RPA gave an acceptable diagnostic performance for woody chips, with improved results for leaf petioles. On the other hand, ddPCR gave values of Youden's Index lower than 0.5 for woody chips (0.321) and leaf petioles (0.474). This could be attributed to its lower specificity and a higher number of false positives relative to Harper-qPCR. When selecting ddPCR as gold standard, Li-qPCR and Francis-qPCR showed the best performance with DNA from woody chips or leaf petioles, respectively. Finally, when using Li-qPCR as gold standard, Harper-qPCR and ddPCR exhibited the lowest diagnostic accuracy for woody chips, and ddPCR for leaf petioles. These results were likely due to their lower specificity and a higher number of false positives, as both tests identified a higher number of *Xf*-infected trees (Table 10).

**Table 11.** Observed and predicted (by the Latent Class Model) number of trees for each combination of diagnostic results for Li-qPCR, Harper-qPCR and ddPCR molecular tests

| Tissue | Li-qPCR | Harper-qPCR | ddPCR | Observed trees (%) | Predicted Class[a] |
|---|---|---|---|---|---|
| Petioles | 0 | 0 | 0 | 19 (21,6) | Healthy |
| | 0 | 0 | 1 | 8 (9,1) | Healthy |
| | 0 | 1 | 0 | 4 (4,5) | Healthy |
| | 0 | 1 | 1 | 2 (2,3) | Healthy |
| | 1 | 0 | 0 | 0 (0,0) | --- |
| | 1 | 0 | 1 | 0 (0,0) | --- |
| | 1 | 1 | 0 | 11 (12,5) | Diseased |
| | 1 | 1 | 1 | 44 (50,0) | Diseased |
| Woody chips | 0 | 0 | 0 | 2 (2,0) | Healthy |
| | 0 | 0 | 1 | 2 (2,0) | Healthy |
| | 0 | 1 | 0 | 1 (1,0) | Healthy |
| | 0 | 1 | 1 | 13 (13,0) | Diseased |
| | 1 | 0 | 0 | 0 (0,0) | --- |
| | 1 | 0 | 1 | 2 (2,0) | Diseased |
| | 1 | 1 | 0 | 0 (0,0) | --- |
| | 1 | 1 | 1 | 68 (68,0) | Diseased |

[a] Individuals were assigned to each latent class (healthy or diseased) based on the highest posterior probability using a maximum likelihood estimation. (---)= No observation for this combination of test results existed in the data set.

### 4.3.5 Latent Class Model (LCM) Analysis

Based on the previous results, in the absence of a perfect gold standard, a LCM analysis was performed to estimate the true disease status of each almond tree when using the three diagnostic tests with the highest sensitivity (i.e., Harper-qPCR, Li-qPCR and ddPCR). Among the eight possible response patterns that can arise from the three diagnostic tests, there were no observations for two patterns when using DNA extracted from either petioles or woody chips (Table 11). Based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion



Funded by the European Union

(BIC) values, a 2-class model provided the most parsimonious and effective fit for explaining the diagnostic outcomes of the three molecular tests for DNA samples extracted from both tissues (Table 11). This simplified LCM effectively distinguished the two latent groups across both tissue types, capturing cases with consistent positive results across protocols (diseased trees) and those with mixed or negative diagnostic pattern (healthy trees). Table 12 shows the class probabilities, indicating the proportion of the trees in each latent class for each protocol and DNA source. For DNA extracted from woody chips, the probability of detecting a true positive (i.e., sensitivity) was maximum for ddPCR (100%), followed by Harper-qPCR (97.1%) and Li-qPCR (84.8%). For DNA extracted from leaf petioles, Harper-qPCR and Li-qPCR exhibited the highest sensitivity (>99.3%), followed by ddPCR (80.0%). Regarding specificity, Li-qPCR achieved the maximum probability of detecting a true negative (100%) while ddPCR showed the lowest specificity (<70.3%) irrespective of the DNA source. The LCM also indicated a high concordance between protocols when analyzing DNA extracted from woody chips, with estimated and predicted population shares of 94%. However, concordance was much lower when using DNA extracted from leaf petioles with a population share of 63% (Table 12).

**Table 12.** Response probabilities' results indicating proportion of the trees in each latent class for each molecular diagnostic test and DNA source obtained by Latent Class Model analysis

| | | | Item response probabilities (conditional probabilities)[x] | |
|---|---|---|---|---|
| **DNA source** | **Diagnostic test** | **Latent class** | **Pr(−) healthy** | **Pr(+) diseased** |
| Woody chips | Li-qPCR | Healthy (class 1) | **1** | 0 |
| | | Diseased (class 2) | 0.1519 | **0.8481** |
| | Harper-qPCR | Healthy (class 1) | **0.6667** | 0.3333 |
| | | Diseased (class 2) | 0.0286 | **0.9714** |
| | Droplet digital PCR (ddPCR) | Healthy (class 1) | **0.5492** | 0.4508 |
| | | Diseased (class 2) | 0 | **1** |
| | Estimated class population shares[y] | 0.0621 | 0.9379 | |
| | Predicted class memberships[z] | 0.0568 | 0.9432 | |
| Leaf petioles | Li-qPCR | Healthy (class 1) | **1** | 0 |
| | | Diseased (class 2) | 0.0074 | **0.9926** |
| | Harper-qPCR | Healthy (class 1) | **0.8285** | 0.1715 |
| | | Diseased (class 2) | 0 | **1** |
| | ddPCR | Healthy (class 1) | **0.7033** | 0.2249 |
| | | Diseased (class 2) | 0.2000 | **0.8000** |
| | Estimated class population shares[y] | 0.3703 | 0.6297 | |
| | Predicted class memberships[z] | 0.3750 | 0.6250 | |

[a] Values indicate the probability of obtaining a negative Pr(−) or positive Pr(+) test result given membership in a specific latent class.
[b] Proportion of the samples predicted to belong to each latent class.
[c] Proportion of trees assigned to each latent class based on the highest posterior probability.


Funded by the European Union

# 5. HIGH-RESOLUTION DIAGNOSIS OF *XYLELLA FASTIDIOSA* USING TARGETED NGS ENRICHMENT

## 5.1 Background

*Xf* is a genetically diverse, polyphyletic bacterium that comprises multiple evolutionary lineages rather than sharing a recent common ancestor. This complex population structure—shaped by horizontal gene transfer, recombination events, and host-specific adaptations—poses significant challenges for phylogenetic classification and outbreak tracking using traditional molecular methods focused on a few loci (Dupas et al. 2023; Landa et al. 2019; Nunney et al., 2014; Parker et al. 2012; Potnis et al. 2019; O'Leary and Burbank, 2023).

Whole-genome sequencing (WGS) provides a higher resolution for identifying and characterizing *Xf* strains (Landa et al., 2019; Moralejo et al., 2020), but it typically requires isolation of pure cultures, a labour-intensive and time-consuming process that is often impractical during emergent outbreaks. To address these limitations, recent studies have explored high-throughput sequencing (HTS) directly from infected plant material. Bonants et al. (2019) demonstrated that deep metagenomic sequencing enables detection of *Xf* DNA even in low-infection samples (Cq ≈ 35), allowing in some cases subspecies-level identification and even full genome assembly. Similarly, Román-Reyna et al. (2021) proposed a shotgun metagenomic workflow using Illumina short reads to detect and assemble *Xf* sequences, achieving strain-level resolution at low bacterial concentrations (as little as 1 pg DNA per gram of tissue). However, these untargeted approaches require high sequencing depth to recover key genetic markers, such as those used in multilocus sequence typing (MLST), which limits their practical utility and cost-efficiency.

However, the effectiveness of these untargeted HTS approaches remains limited by the low abundance of bacterial DNA in plant or insect samples. To overcome this, hybridization-based targeted sequencing, commonly referred to as pathogen enrichment sequencing (PenSeq), has been successfully applied in other pathosystems. Using biotinylated baits, PenSeq enriches pathogen-specific sequences prior to WGS, improving detection sensitivity and reducing sequencing cost (Andermann et al. 2020; Grillova et al. 2023; Thilliez et al. 2019;). Recently, Boutigny et al. (2023) applied such an approach to *Xf*, demonstrating the feasibility of capturing nearly complete genomes from samples with moderate infection levels (Cq ≈ 31).

In the present study, we developed and validated a *Xylella fastidiosa*-Targeted Sequence Capture Enrichment (TSCE) method, coupled with Illumina-based HTS, to enable accurate diagnosis and high-resolution phylogenetic analysis directly from DNA extracted from infected plants or insect vectors. Capture probes were designed based on available RefSeq genomes, targeting 138 coding genes from both the chromosome and *Xf* subsp. *multiplex* plasmids. We also demonstrated the ability to recover nearly complete plasmid sequences when probes spanned their entire length. This study has provided with a TSCE method for *Xf* diagnosis, with resolution at the strain level, which is has cost-effectiveness similar to currently used approaches for subspecies assignment as required by European regulatory frameworks.



Funded by the European Union

## 5.2 Materials and Methods

Figure 10 illustrates the workflow of the TSCE-HTS methodology developed in this study, outlining the steps from sample reception to the output of the bioinformatic pipeline.

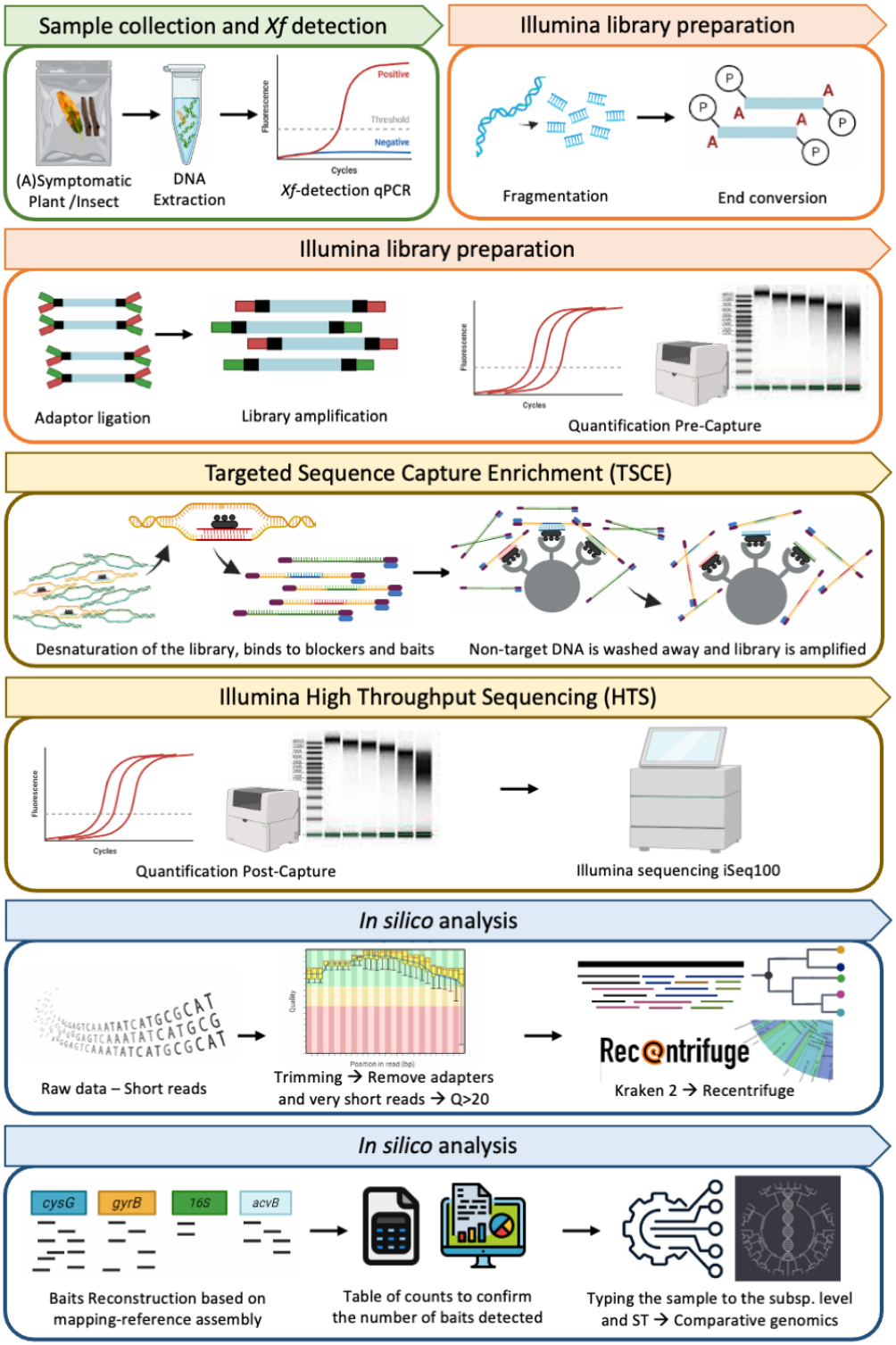



**Figure 10.** Workflow of the TSCE-HTS methodology developed in this study.



Funded by the European Union

### 5.2.1 *Xylella fastidiosa* strains and DNA isolation

The optimization of the protocol was carried out using mock inoculated samples with known DNA concentrations of host and of different *Xf* subspecies strains (Table 13). A total of four strains were used to prepare mock inoculated samples: Strain IVIA5235 of subspecies *fastidiosa* ST1 isolated from a cherry tree in Mallorca Island, Spain ( Arias-Giraldo et al. 2020), strain ESVL of subspecies *multiplex* ST6 isolated from an almond tree in Alicante, Spain (Giampetruzzi et al. 2019), strain De Donno of subspecies *pauca* ST53 isolated from an olive tree in Apulia, Italy (Saponari et al., 1993) provided by M. Saponari (CNR, Bari, Italy), and strain CFBP8074 of subsp. *pauca* ST74 intercepted in Europe in a coffee plant from Ecuador acquired from the CFBP (Collection for Plant-associated Bacteria, INRAE, France).

All strains are stored at -80ºC and belong to the *Xf* collection at the Instituto de Agricultura Sostenible (IAS-CSIC), Córdoba, Spain. Strains were cultured in BCYE (Wells et al. 1981) or PD2 (Davis et al. 1980) solid media at 28˚C in the dark for 7 to 12 days (depending on the strain). Cells were scrapped from the media and used for DNA extraction using the DNeasy Plant Mini Kit (Quiagen). Total DNA was quantified using a spectrofluorometer (Qubit; Thermo Fisher Scientific, Ciudad, Pais), the quality was measured by using a nanodrop spectrophotometer and the DNA integrity was verified by 1% agarose gel electrophoresis.

### 5.2.2 Mock inoculated samples of plants and insect vectors

Four spiked samples were prepared with a final concentration of 3 pg/µl of *Xf* DNA in an 80 ng/µl background of Olive (*Olea europaea*) DNA.

Additionally, a total of 10 *Xf*-spiked samples from the Interlaboratory test comparison EU-*XF*-IC-2020-03 were also included in the study. The samples consisted on heat-inactivated (incubation at 70˚C for 15 min) bacterial suspensions of *Xf* subsp. *pauca* strain De Donno (CFBP 8402) at different concentrations mixed in a *Polygala myrtifolia* and insect DNA background of *Philaenus spumarius*. Bacterial suspension was prepared by scraping 10-days old colonies of *Xf* subsp. *pauca* strain De Donno cultures growing at 28ºC on BCYE (Wells et al. 1981) plates. DNA was extracted using the a cetyltrimethylammonium bromide (CTAB) DNA-based extraction protocol. Olive and insect DNA were extracted using the CTAB-based extraction method according to EPPO (2023). Additionally, a blind sample of a *Xf* bacterial culture was provided by M. Román-Ecija (IAS-CSIC).

### 5.2.3 Development of *Xf* DNA standard curves

To establish the limit of detection of the *Xf*-TSCE two sets of DNA standard curves were obtained using 10-fold dilution series of DNA of *Xf* subsp. *multiplex* strain ESVL. For this purpose, *Xf* from this strain was serially diluted in a fixed background of *Olea europaea* (25 ng/µl) extracted from leaf petioles using the CTAB-based DNA extraction procedure (EPPO 2023) to obtain final concentrations starting from 3 ng/µl to 3 fg/µl. Total DNA was quantified twice using a spectrofluorometer (Qubit; Thermo Fisher Scientific) and the Cq of the sample was estimated by using the qPCR-Harper.



Funded by the European Union

**Table 13.** *Xylella fastidiosa* mock-inoculated samples used for the optimization of the Targeted Sequence Capture Enrichment protocol and for comparing the effect of library multiplexing

| DNA ID [a] | DNA background | *Xf* subsp. | ST[b] | Origin | Cq[c] | Cq[c] [1ng/µl] | Capture type[d] |
|---|---|---|---|---|---|---|---|
| 19C-PmDD | *Polygala myrtilofia* | *pauca* | 53 | Interlaboratory | 21 | 26 | $t_2$ |
| 9C-PmDD | *Polygala myrtilofia* | *pauca* | 53 | Interlaboratory | 24 | 30 | s and $t_5$ |
| 4C-PmDD | *Polygala myrtilofia* | *pauca* | 53 | Interlaboratory | 28 | 33 | s and $t_5$ |
| 20C-PmDD | *Polygala myrtilofia* | *pauca* | 53 | Interlaboratory | 31 | na | $t_5$ |
| 32C-PsDD | *Philaenus spumarius* | *pauca* | 53 | Interlaboratory | 26 | 28 | s and $t_3$ |
| 36C-PsDD | *Philaenus spumarius* | *pauca* | 53 | Interlaboratory | 29 | 31 | wc and $t_3$ |
| 37C-PsDD | *Philaenus spumarius* | *pauca* | 53 | Interlaboratory | 33 | na | wc and $t_4$ |
| C4d4-OeESVL* | *Olea europaea* | *multiplex* | 6 | Alicante, Spain | 26 | 31 | s |
| C4d5-OeESVL* | *Olea europaea* | *multiplex* | 6 | Alicante, Spain | 30 | 32 | s |
| C4d6-OeESVL* | *Olea europaea* | *multiplex* | 6 | Alicante, Spain | 33 | na | s |
| C4d7-OeESVL* | *Olea europaea* | *multiplex* | 6 | Alicante, Spain | 36 | na | s |
| Oe-CFBP8074** | *Olea europaea* | *pauca* | 74 | Ecuador* | 24 | 28 | $t_1$ |
| Oe-DD | *Olea europaea* | *pauca* | 53 | Lecce, Italy | 24 | 27 | s, $d_1$, $t_1$ and wc |
| Oe-ESVL | *Olea europaea* | *multiplex* | 6 | Alicante, Spain | 24 | 28 | s, $d_1$, $t_1$ and wc |
| Oe-IVIA5235 | *Olea europaea* | *fastidiosa* | 1 | Mallorca, Spain | 24 | 28 | $t_1$ |

[a] (*)= DNA samples used for stardard diltution curves. (**)= Blind sample provided by M. Román-Écija (IAS-CSIC, Spain).

[b] ST= Sequence type as determined by MLST analysis.

[c] Cq= Threshold quantitative cycle as determined by Harper-qPCR (Harper et al. 2010) of DNA after extraction or after normalization to 1 ng/ µl. na= no amplification.

[d] Type of capture performed: s = singleplex, t= tetraplex, wc= without capture. Same subscript numbers indicate samples processed within the same pooling run of captures (i.e., the libraries were captured simultaneously or with other samples included in Table 14).

### 5.2.4 Naturally-infected plants and insect vectors samples

A total of 30 plant and insect samples naturally infected by different *Xf* subspecies and STs were used in this study (Table 14). DNA samples were provided by different official laboratories in Europe and were obtained from a total of 12 plant hosts and *P. spumarius*. DNA was extracted as described in the EPPO protocol (EPPO 2023) from leaf petioles or wood chip samples using the CTAB-based extraction or the EZNA HP Plant Mini kit (Omega-Biotek, Norcross, Georgia, USA), depending on the laboratory providing the samples.



Funded by the European Union

**Table 14.** List of *Xylella fastidiosa* naturally-infected samples used in this study

| Plant/Insect DNA sample[a] | Host background | *Xf* subsp. | ST[b] | origin | Cq (1ng/ul) | Capture type[c] | Illumina platform |
|---|---|---|---|---|---|---|---|
| *XF*4358/18 bis2 | *Rhamnus alaternus* | *fastidiosa* | 1 | Mallorca, Spain | 26 | $t_2$ | MiSeq |
| XYL5773 | *Vitis vinifera* | *fastidiosa* | 1 | Mallorca, Spain | 29 | $t_2$ | MiSeq |
| Xm1215 | *Philaenus spumarius* | *fastidiosa+ multiplex** | 1+ 81* | Mallorca, Spain | 33 | $t_6$ | MiSeq |
| XYL2774/23 | *Prunus dulcis* | *fastidiosa+ multiplex** | 1+ 81* | Mallorca, Spain | 26 | s | iSeq 100 |
| IVIA8926 | *Lavandula angustifolia* | *multiplex* | 6 | Alicante, Spain | 31 | $t_8$ | MiSeq |
| IVIA8938 | *Prunus dulcis* | *multiplex* | 6 | Alicante, Spain | 32 | $t_4$ | MiSeq |
| IVIA8922 | *Prunus dulcis* | *multiplex* | 6 | Alicante, Spain | 36 | wc and $t_5$ | MiSeq |
| XYL583/16bis | *Prunus dulcis* | *multiplex* | 7 | Mallorca, Spain | 34 | $t_8$ | MiSeq |
| XYL1374/17 | *Prunus dulcis* | *multiplex* | 7 | Mallorca, Spain | 24 | $t_6$ | MiSeq |
| 2531 (2021) | *Quercus suber* | *multiplex* | 7 | Portugal | 31 | $t_7$ | MiSeq |
| 2017 (2021) | *Gazania* sp. | *multiplex* | 7 | Portugal | 35 | $t_{98}$ | MiSeq |
| XYL610 | *Polygala myrtifolia* | *multiplex* | 7 | Mallorca, Spain | 28 | s | MiSeq |
| XYL0895/20 | *Calicotome spinosa* | *multiplex* | 81 | Mallorca, Spain | 26 | $t_6$ | MiSeq |
| Xm1550 | *Philaenus spumarius* | *multiplex* | 81 | Menorca, Spain | 30 | $t_7$ | MiSeq |
| XYL1050/19 | *Prunus dulcis* | *fastidiosa* | 81 | Mallorca, Spain | 27 | $t_3$ | MiSeq |
| L3/0300 | *Olea europaea* | *multiplex* | 81 | Mallorca, Spain | 31 | wc and $t_4$ | MiSeq |
| CF5GrapSI2* | *Olea europaea* | *pauca* | *unk* | Brazil | 28 | $t_6$ | MiSeq |
| PUG18-193 | *Olea europaea* | *pauca* | 53 | Italy | 26 | $t_2$ | MiSeq |
| IVIA30/24 | *Olea europaea* | *pauca* | 53 | Mallorca, Spain | 33 | s | iSeq 100 |
| IVIA33/24 | *Nerium oleander* | *pauca* | 53 | Mallorca, Spain | 31 | s | iSeq 100 |
| IVIA35/24 | *Olea europaea* | *pauca* | 53 | Mallorca, Spain | 29 | s | iSeq 100 |
| XYL0078/24 | *Olea europaea* | *pauca* | 53 | Mallorca, Spain | 26 | wc and t | iSeq 100 |
| Xm1489 | *Philaenus spumaius* | *pauca* | 80 | Ibiza, Spain | 31 | $t_7$ | MiSeq |
| IVIA11394 | *Ulex parvifolius* | *pauca* | 80 | Ibiza, Spain | 31 | $t_8$ | MiSeq |
| XYL4933/20 | *Olea europaea* | *pauca* | 80 | Ibiza, Spain | 27 | s and $t_3$ | MiSeq |
| XYL0090/19 | *Salvia rosmarinus* | *pauca* | 80 | Ibiza, Spain | 33 | wc and $t_4$ | MiSeq |
| D* | *Vitis vinifera* | *unknown* | *unk* | INRAE, France | 29 | $t_7$ | MiSeq |
| M9-2 | *Olea europaea* | *unknown* | 90 | La Rioja, Argentina | 23 | s | iSeq 100 |
| F9P4 | *Olea europaea* | *unknown* | 90 | Córdoba, Argentina | 29 | s | iSeq 100 |
| F8P13 | *Olea europaea* | *unknown* | 90 | Córdoba, Argentina | 23 | s | iSeq 100 |

[a] (*)= 'CF5GrapSI2' is a blind sample provided by H. Coletta (APTA, Brazil) and 'D' is a blind sample provided by M.A. Jacques (INRAE, France).

[b] ST= Sequence type. (*) The qPCR of Dupas et al. (2019a) indicated the coinfection by both subspecies. MLST analysis indicated a mixed infection by two *Xylella fastidiosa* subspecies (*fastidiosa* ST1 and *multiplex* ST81).

[b] Type of capture performed: s = singleplex, t= tetraplex, wc= without capture. Same subscript numbers indicate samples processed within the same pooling run of captures (i.e., the libraries were captured simultaneously or with other samples included in Table 13).



Funded by the European Union

At IAS-CSIC the Harper's qPCR test (Harper et al. 2010) was used to quantify the amount of *Xf* present in all DNA samples; qPCR assays were performed as described on section 4.2.5 of this deliverable. The MLST (Yuan et al. 2010) and Nested-MLST (Cesbron et al. 2020) assays were used to assign the strain to the subspecies and Sequence type (ST) level before performing the TSCE. All samples were run in duplicate for qPCR analysis with the exception of samples used in the DNA standard curves that were run in triplicate on the same PCR plate. In each PCR plate, two concentrations (1/1 and 1/10) of boiled *Xf* subsp. *sandyi* strain CFBP8478 were used as positive amplification control (PAC), while sterile PCR-grade water and olive DNA known to be *Xf*-free were used as negative amplification controls (NAC). All samples for the conventional or Nested MLST were analyzed using two aliquots of DNA.

In total, 22 samples were obtained from Spain, including Mallorca Island (subspp. *fastidiosa* ST1, *multiplex* ST7 and ST81, and *pauca* ST53), Ibiza Island (subsp. *pauca* ST80), Menorca Island (subsp. *multiplex* ST81), and from the Demarcated area of Alicante, Valencian Community (subsp. *multiplex* ST6). Additionally, six samples including two from Portugal (subsp. *multiplex* ST7), one from Italy (subsp. *pauca* ST53), and three from Argentina, two belonging to the ST78 and one from an unknown subspecies but belonging to a new ST (ST90) were included in the study (Table 14). Finally, a blind sample was provided by INRAE (France) consisting of DNA extracted from a grapevine plant artificially inoculated with an unknown *Xf* strain, and a second natural blind sample provided by APTA (Brazil) from a *Xf*-infected olive tree.

### 5.2.5 Selection of *Xylella fastidiosa* target sequences and baits design

We selected 138 target sequences (Supplementary Table 11) covering coding conserved and non-conserved and non-coding regions of the different subspecies and ST of known *Xf* strains sequenced which genome was present at Genbank up to of 2021. The selected target sequences are shown in Supplementary Table S11 and correspond to genes used for *Xf* diagnostics assays, house-keeping genes used in MLST analysis differentiating subspecies and STs (Yuan et al. 2010), specific sequences to distinguish certain subspecies and strains (Citrus variegated chlorosis and mulberry strains), genes used in the MLSA of environmentally mediated genes (Parker et al. 2012), genes from the pangenome of *Xf* which are crucial for understanding *Xf* evolution and adaptation to different ecological niches, horizontal gene transfer, etc.

Additionally, to analyze the possibility of being able to capture entire plasmids, specific targets were designed for the two plasmids found in the ESVL strain (Giampetruzzi et al. 2019).

For the development of the sequence capture target enrichment approach, biotinylated RNA baits were custom-designed by Arbor Biosciences (Ann Arbor, MI, USA) to target the selected *Xf* specific loci of interest (Supplementary Table S11). Similar hybridization-based enrichment strategies using Arbor Biosciences baits have been successfully implemented for plant pathogen surveillance and phylogenomic studies, including the recovery of *Xf* genomes from infected plant material (Boutigny et al. 2023), and the enrichment of other obligate plant pathogens from complex environmental samples (Rodriguez et al. 2021). These approaches enable the efficient retrieval of targeted genomic regions even when pathogen DNA occurs at low abundance, yielding higher sequencing depth at the chosen loci compared to shotgun metagenomics. This improved coverage enhances downstream phylogenetic resolution, strain discrimination, and epidemiological tracing in agricultural and natural ecosystems.


Funded by the European Union

### 5.2.6 Illumina library preparation of studied samples

Total DNA of samples was quantified using the Qubit spectrofluorometer (Thermo Fisher Scientific) and the TapeStation 4150 system with the DNA D1000 kit and cartridge (Agilent Technologies, Santa Clara, California, USA). This later system was also used to check DNA integrity. DNA was diluted to start from an initial total mass of 250-500 ng for library preparation. Next, DNA was fragmented to approximately 350 bp using the NEBNext Ultra II FS DNA Prep Kit. Fragmentation time was adjusted empirically according to the DNA integrity.

Libraries were prepared using the NEBNext Multiplex Oligos (Dual Index Primers) from New England Biolabs (Ipswich, Massachusetts, USA) for both Illumina sequencers following the manufacturer's protocol with some small modifications. In the PCR enrichment step of adaptor-ligated DNA, the volume used of P5/P7 indexes was reduced to 2.5 µl, and a volume of 20 µl of the adaptor-ligated DNA fragment. The number of cycles of enrichment were adapted as indicated in the protocol according to the DNA input for enrichment between 750 ng and 1 µg. After library preparation, an additional cleaning step with 0.9X AMPure XP beads (Beckman Coulter, High Wycombe, UK) was performed in the case that primer dimers (120-125 bp) or primer remnants (60 bp) were observed after quantifying the library concentration on the TapeStation 4150 system.

For some selected DNA samples two libraries were prepared simultaneously using different dual index primers, one library was sequenced directly after this step and the second one followed the Targeted Sequence Capture Enrichment (TSCE) described below.

### 5.2.7 *Xylella fastidiosa*-TSCE and sequencing

Hybridization with *Xf* baits was performed using the myBaits kit (Arbor Biosciences, Ann Arbor, MI, USA) following the manufacturer's protocol. For environmental samples with low numbers of *Xf* target copies, more than 200 ng of library DNA was used as input, and hybridization was extended to at least 20 hours, as recommended. The number of post-capture PCR cycles for each sample was adjusted according to its known *Xf* concentration, using the minimum number of cycles needed to generate sufficient library material for sequencing. After post-capture amplification, the quality and concentration of enriched libraries were assessed using a TapeStation 4150 with the DNA D1000 kit (Agilent Technologies, Santa Clara, CA, USA).

To improve cost-effectiveness, we evaluated the possibility of pooling DNA libraries during sequence capture. Libraries were combined either in pairs (duplex) or in groups of four (tetraplex) and compared with individually processed libraries (singleplex) or with non-captured libraries. Pooling was performed by adjusting the proportion of each library based on the Ct values of the original samples, ensuring an equimolar number of *Xf* molecules across pooled libraries prior to hybridization.

Enriched captured libraries were pooled equimolarly and then they were sequenced on an Illumina MiSeq instrument and the v2 kit (2 x 250 bp; 300-cycles) kit at the Genomics Unit at "Fundación Parque Científico de Madrid" or using the Illumina iSeq100 sequencer and the iSeq 100 i1 Reagent v2 (2 x 150 bp; 300-cycles) kit at the IAS-CSIC facilities.



Funded by the European Union

### 5.2.8 *In silico* bioinformatic pipeline

Raw DNA reads in FASTQ format were processed using *fastp* v0.23.2 (Chen et al. 2018) for trimming and initial quality control. Adapters and low-quality reads (Q<20) were removed. A detailed quality report was generated using MultiQC v.1.13. Initially, potential contamination and non-targeting DNA sequences were detected and filtered out with Kraken2 v2.1.2 (Wood et al. 2019) and KrakenTools v.1.2 with a precompiled reference database (PlusPFP v.6-5-2023). The results of the taxonomic classification performed with Kraken were visualized and analyzed using Recentrifuge v1.9.1 (Martí 2019).

For the reconstruction of selected marker genes, reads classified as *Xf* (TAXID: 2370) were then extracted and aligned to the reference *arbor* bait sequences using *BWA-MEM* v0.7.17 (Li, 2013). Consensus regions were subsequently extracted with *bcftools* v1.20, and the resulting FASTQ file was converted to FASTA format using *seqtk* v1.3, filtering out contigs with 75% or more ambiguous bases (Ns). Gene identification of the assembled contigs was performed using *BLASTn*, which generated a report of coverage and gene counts against a database of the sequences used to design the baits. Additionally, a MultiLocus Sequence Typing (MLST) profiling was performed from the assembled housekeeping gene sequences in FASTA format using *mlst* tool v2.19.0 to identify the corresponding ST of the *Xf* present in the sample. The MLST profile for each sample was then exported to a CSV file, containing the typing results for each assembled sequence.

A dataset of 204 complete *Xf* genomes, representing all subspecies: *fastidiosa*, *morus*, *multiplex*, *pauca*, and *sandyi* was created using all genomes available at the Reference genome database of GenBank as a date of December 2024.

Genome annotation of the captured genes was performed using Prokka v1.14.6 (Seemann 2014) with default settings. The coding sequences (CDSs) obtained were utilized to estimate the core genome through CoreCruncher, (Harris et al. 2021) which employed the MAFFT algorithm (Katoh and Standley 2013) for constructing the core genome alignment. Initially, the frequency parameter in CoreCruncher was set at 100%, but this threshold was reduced to 90% to retain important genes, thus assembled genes present in less than 90% samples were removed. Subsequently, poorly aligned regions and ambiguous sequences were removed from the multiple sequence alignment using ClipKIT v1.3. A maximum likelihood (ML) phylogenetic tree was generated with IQ-TREE v2.2.0, applying the GTR+I+G4 substitution model as determined by ModelTest-NG v0.1.6 (Darriba et al. 2020; Nguyen et al. 2015). Different ML trees were created depending of the samples to analyze, the purpose of the experiment and the number of captured genes. The tree was visualized using ggtree v3.6.2 (Yu 2020). The tree topology was rooted using the midpoint method, and branch support was assessed through 1,000 bootstrap replicates.



Funded by the European Union

## 5.3 Results

### 5.3.1 Limit of detection of the *Xf*-Targeted Sequence Capture Enrichment

To assess the sensitivity of the *Xf*-TSCE, two replicates of each concentration were tested using independent dilutions which correspond from 3000 fg/µl to 3 fg/µl of *Xf* DNA by using three diagnostic protocols. In Figure 11, the Cq values from Harper-qPCR and the amplification of the seven genes through both MLST and nested-MLST are shown. This comparison highlights the differences in amplification efficiency and signal intensity across the different diagnostic tests and genes. Results were very reproducible for both serial dilutions of *Xf* DNA.

In MLST analysis, the genes that produced the strongest signal were *leuA*, *petC*, *nuoL*, and *gltT*, amplifying up to 30 fg/µl. Though, PCR products from these genes do not always yielded high-quality sequencing results (+/-) which are necessary to determine the exact allele sequence type. *malF* and *holC* genes amplify up to 300 fg/µl although with a lower signal intensity. The sensitivity of the nested-MLST tests increased significantly with amplification reached up to 3 fg/µl with equal signal intensity across all the seven genes (Figure 11).

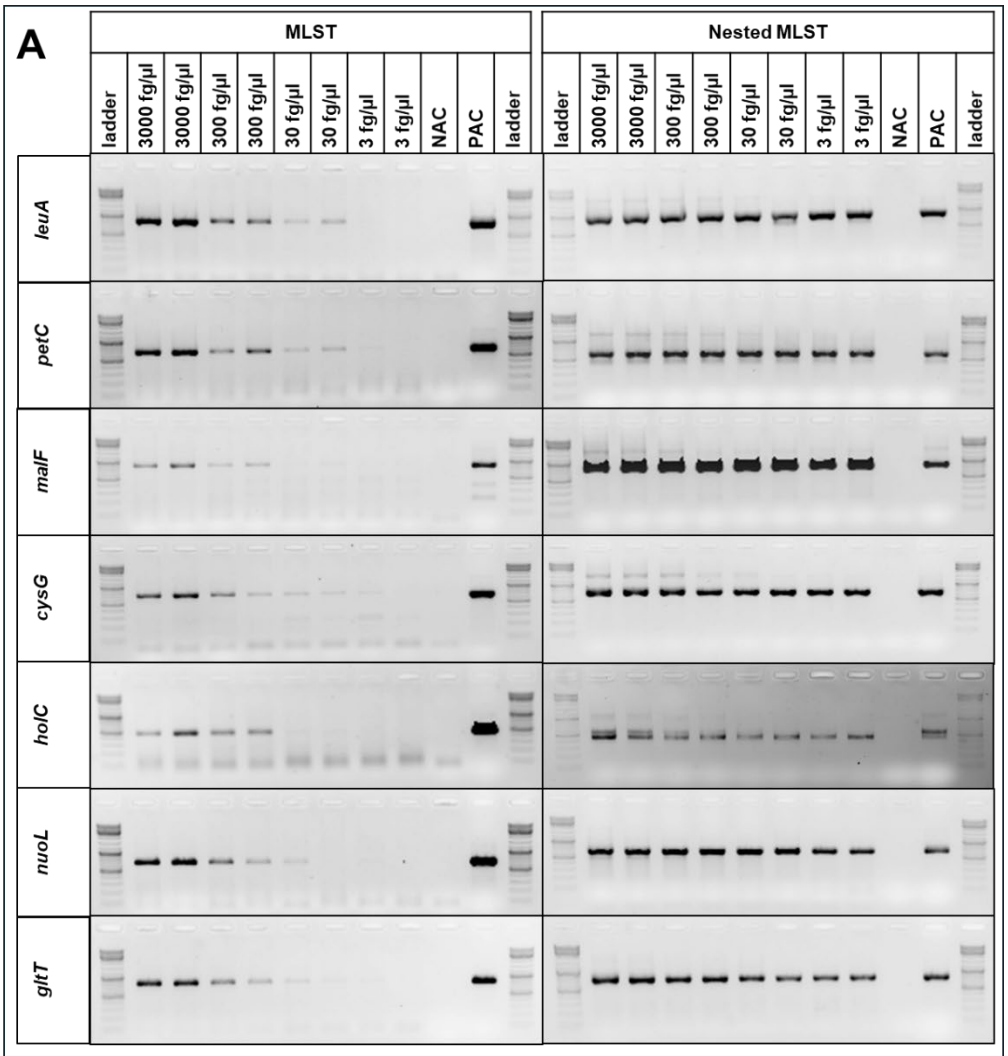
In the case of the *Xf*-TSCE developed in this study, 132 out of 138 genes were successfully captured at concentrations of 3000 fg/µl and 300 fg/µl, including the 7 MLST genes (Table 15). Additionally, sequences targeting both plasmids corresponding to the *multiplex* subspecies strain ESVL were reconstructed. For the sample with a concentration of 30 fg/µl and a Cq value of 33, a total of 54 genes were captured, showing a percentage identity between 96.8% and 99.4% and good coverage with the corresponding *Xf* genes. Finally, for the sample with a concentration of 3 fg/µL and a Cq value of 36, only 2 genes were captured (*queH, hfq*), but with a high identity (>96.3) and good coverage with *Xf* genes (Table 15). It should be noted that before starting the process of TSCE the olive DNA was normalized to a final concentration of 1 ng/µl which in turn reduced significantly the amount of *Xf* target DNA in about 1/25.

### 5.3.2 Effect of library pooling and concentration of Xylella fastidiosa on Xf-TSCE results

To reduce experimental costs and increase throughput, we evaluated whether multiplexing during the sequence capture of *Xf* DNA affected capture efficiency. The same two samples were processed under three capture configurations: individually (singleplex), in pairs (duplex), and in groups of four (tetraplex) and results were compared to the sample without performing TSCE (Table 15).

The impact of capture configuration on *Xf* enrichment was evaluated for plant and insect samples mock-inoculated with strains DD and ESVL (Table 15). In olive samples inoculated with the DD and ESVL strains, sequencing without capture yielded very few *Xf* reads and only partial gene recovery. For Oe-DD without capture, 1,212 reads (0.24% of total) were obtained, corresponding to six genes. By contrast, singleplex, duplex, and tetraplex captures produced between 436,694 and 546,294 reads (53.3–64.2% of total), with consistent recovery of 126 genes and >99% genome coverage. For Oe-ESVL, the improvement was similar: only 723 reads (0.15%) and two genes were retrieved without capture, whereas capture configurations yielded 515,953–755,332 reads (58–74%), consistently recovering 131 genes, including the plasmid sequence, with >99% coverage.


Funded by the European Union




**B**

| [DNA] (fg/µl) | Ct | MLST | | | | | | | Nested MLST | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *leuA* | *petC* | *malF* | *cysG* | *holC* | *nuoL* | *gltT* | *leuA* | *petC* | *malF* | *cysG* | *holC* | *nuoL* | *gltT* |
| 3000 | 26.0<br>26.1 | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ |
| 300 | 29.5<br>29.6 | ++ | ++ | +/– +/– | ++/– | +/– +/– | ++/– | ++/– | ++ | ++ | ++ | ++ | ++ | ++ | ++ |
| 30 | 32.8<br>33.1 | +/– +/– | +/– +/– | – – | +/– +/– | – – | +/– – | +/– +/– | ++ | ++ | ++ | ++ | +– | ++ | ++ |
| 3 | 35.6<br>36.3 | – – | – – | – – | +/– – | – – | – – | +/– – | ++ | ++ | ++ | ++ | ++ | ++ | ++ |

**Figure 11.** (**A).** Agarose gels showing the sensitivity of the seven house-keeping genes (*leuA, petC, malF, cysG, holC, nuoL, gltT*) MLST analysis for each gene (left panels) and of nested-MLST (right panels) using serial dilutions of *Xylella fastidiosa* (*Xf*) subspecies *multiplex* strain ESVL DNA (from 3000 to 3 fg/µL) mixed in an olive background DNA. PAC: Positive amplification control contained boiled cells of *Xf* subsp. *sandyi* strain CFBP8478. NAC: Negative amplification control contained DNA from a non *Xf* infected plant or sterile PCR-grade water. L: 10 kb GeneRuler™ DNA Ladder Mix (Thermo Scientific™) **B.** Data summarizing the results from two independent standard DNA curves for each protocol performed: (+) = Positive amplification, (+/-) = Positive but weak amplification, and (–) = No amplification. (Ct)= Cycle threshold values obtained for Harper's qPCR using the same DNA standard curves. For the *Xf*-TSCE results, only the results of one replicate are shown: (+) = captured gene; (-) = not captured gene.



Funded by the European Union

**Table 15.** Comparison of TSCE results for DNA plant and insect samples mock inoculated with different *Xylella fastidiosa* strains at different concentrations using different capture configurations

| ID sample | Cq[a] (1ng /µl) | Capture type[b] | Total number of reads | *Xf* reads | (% of total) | Number of *Xf* captured genes | Mean Identity (%) | Mean Coverage (%) | MLST genes (ST)[c] |
|---|---|---|---|---|---|---|---|---|---|
| Oe-DD | 27 | wc | 495,731 | 1,212 | (0.24) | 6 | 98.85 | 80.17 | 0C (ST-) |
| | | s | 869,738 | 546,294 | (62.81) | 126 | 97.89 | 99.98 | 7C (ST53) |
| | | $d_1$ | 688,075 | 441,544 | (64.17) | 126 | 97.90 | 99.95 | 7C (ST53) |
| | | $t_1$ | 819,222 | 436,694 | (53.31) | 126 | 97.89 | 99.96 | 7C (ST53) |
| Oe-ESVL | 28 | wc | 495,731 | 723 | (0.15) | 2 | 96.45 | 88.00 | 0C (ST-) |
| | | s | 1,022,617 | 755,332 | (73.86) | 131 | 98.77 | 99.97 | 7C (ST6) |
| | | $d_1$ | 757,021 | 515,953 | (68.26) | 131 | 98.78 | 99.98 | 7C (ST6) |
| | | $t_1$ | 962,134 | 557,731 | (57.97) | 131 | 98.78 | 99.95 | 7C (ST6) |
| 19C-PmDD | 26 | $t_2$ | 603,346 | 270,935 | (44.91) | 126 | 97.89 | 99.95 | 7C (ST53) |
| 9C-PmDD | 30 | s | 460,385 | 172,535 | (37.48) | 126 | 97.89 | 99.94 | 7C (ST53) |
| | | $t_5$ | 18,786 | 617 | (3.28) | 61 | 97.78 | 83.95 | 1C (ST-) |
| 4C-PmDD | 33 | s | 586,984 | 94,933 | (16.17) | 126 | 97.90 | 99.95 | 7C (ST53) |
| | | $t_5$ | 150,827 | 16,997 | (11.27) | 55 | 97.75 | 83.47 | 0C (ST-) |
| 20C-PmDD | na | $t_5$ | 2,213,426 | 792 | (0.04) | 69 | 98.07 | 82.28 | 0C (ST-) |
| 32C-PsDD | 28 | s | 644,808 | 94,933 | (14.72) | 126 | 97.90 | 97.89 | 7C (ST53) |
| | | $t_3$ | 714,981 | 16,997 | (2.38) | 126 | 97.90 | 99.96 | 7C (ST53) |
| 36C-PsDD | 29 | wc | 399,781 | 68 | (0.02) | 4 | 99.31 | 99.5 | 0C (ST-) |
| | | $t_3$ | 3,347,999 | 8,763 | (0.26) | 127 | 97.90 | 99.81 | 7C (ST53) |
| 37C-PsDD | 28 | wc | 369,416 | 43 | (0.01) | 2 | 97.27 | 87.5 | 0C (ST-) |
| | | $t_4$ | 1,329,284 | 8,895 | (0.26) | 126 | 97.85 | 98.63 | 5C,2P (ST53?) |
| C4d4- OeESVL | 26* | s | 973,236 | 218,120 | (22.41) | 132 | 98.75 | 99.95 | 7C (ST6) |
| C4d5- OeESVL | 30* | s | 1,096,631 | 31,327 | (2.86) | 132 | 98.77 | 99.42 | 7C (ST6) |
| C4d6- OeESVL | 33* | s | 1,156,653 | 27,777 | (2.40) | 54 | 98.56 | 82.33 | 0C (ST-) |
| C4d7- OeESVL | 36* | s | 484,231 | 112 | (0.02) | 2 | 98.87 | 93.5 | 0C (ST-) |
| Oe-CFBP8074 | 28 | $t_1$ | 819,791 | 354,241 | (34.21) | 124 | 98 | 99.96 | 7C (ST74) |
| Oe-IVIA5235 | 28 | $t_1$ | 1,032,098 | 369,668 | (35.82) | 131 | 99.87 | 99.99 | 7C (ST1) |

[a] Threshold quantitative cycle of the sample as determined by Harper-qPCR (Harper et al 2010). na= no amplification. (*) Cq values correspond to the library withouth normalization to 1 ng/µl.

[b] Capture configuration: (s)= singleplex, (d)= duplex, (t)= tetraplex, (wc)= without capture. Subscript numbers indicate samples processed within the same pooling run of captures (i.e., the libraries were captured simultaneously or with other samples included in Table 16).

[c] The number of housekeeping genes captured is indicated (C)= correctly assigned to an allele; (P)=Partially assigned to the closest allele. When the seven housekeeping genes are captured the ST is provided; when some genes arecaptured although with no perfect matching the closest ST is indicated with?.



Funded by the European Union

Consistent with the results obtained for the DD and ESVL strains, the Oe-CFBP8074 sample (Cq 28), tetraplex capture yielded 354,241 *Xf* reads (34.2% of the total), with 124 genes recovered at 98% identity and 99.9% genome coverage. The Oe-IVIA5235 sample (Cq 28) produced 369,668 reads (35.8%), with 131 genes recovered at 99.9% identity and nearly complete genome coverage (99.99%). In both cases, MLST typing was successful: strain IVIA5235 was assigned to ST1, while the blind CFBP8074 sample was identified as belonging to ST74.

The efficiency of capture in insect- and *P. myrtifolia* samples inoculated with the DD strain (PmDD and PsDD backgrounds, respectively) was more variable. For PmDD samples with Cq 30–33, singleplex runs recovered 94,933–172,535 *Xf* reads (16–38% of total), with 126 genes retrieved. Tetraplex runs of the same samples resulted in lower read numbers (e.g., 617 reads, 3.3%, and 16,997 reads, 11.3%) and fewer captured genes (61 and 55, respectively). In PsDD libraries, uncaptured sequencing produced only 43–68 reads (0.01–0.02% of total) and 2–4 genes, whereas tetraplex capture recovered 8,763–8,895 reads (0.26%), with 126–127 genes retrieved. Notably, in 32C-PsDD (Cq 28), both singleplex and tetraplex runs yielded similar gene recovery (126 genes), despite differences in read number (Table 15).

The amount of *Xf* present in the sample, estimated from the Harper-qPCR Cq values, was associated with the number of reads and genes captured across plant and insect matrices. In olive (Oe) samples with relatively low Cq values (27–28), sequencing after TSCE consistently yielded high proportions of *Xf* reads (53–79%) and complete recovery of 126–131 genes, regardless of capture configuration. In *P. myrtifolia* (Pm) and *P. spumarius* (Ps) insect samples, higher Cq values (30–36) were generally accompanied by lower numbers of *Xf* reads. For example, the 9C-PmDD sample (Cq 30) produced 172,535 reads in singleplex capture (37.5%), while the 36C-PsDD sample (Cq 36) yielded only 8,763 reads in tetraplex mode (0.26%). Notably, however, the number of recovered genes was not always reduced in proportion to the decline in read counts. Samples with Cq values up to 32 frequently produced near-complete recovery (124–131 genes), and only samples with very high Cq values (>33) showed compromised gene recovery, as illustrated by 4C-PmDD and C4d6, which captured fewer than 70 genes (Table 15).

Most captured libraries allowed full or near-complete recovery of the seven MLST loci, with successful sequence type assignments (e.g., ST53 and ST6), only some tetraplexed captures at low concentration of *Xf* failed in retrieving the MLST. Without capture, MLST profiling was not possible due to the low number of recovered genes (Table 15).

### 5.3.3 *Xylella fastidiosa*-TSCE in naturally infected samples

TSCE demonstrated strong performance in naturally infected plant samples, enabling near-complete genome recovery even at higher Cq values (Table 16). In low-Cq samples, nearly all reads corresponded to *Xf*, resulting in >90% genome coverage and full MLST recovery. Even in more challenging samples (Cq ≥ 32), hundreds of genes were still retrieved, highlighting the robustness of capture-based enrichment. In contrast, sequencing without capture consistently failed, producing negligible reads and insufficient loci for meaningful analysis (Table 16). For instance, sample XYL2774/23 (Cq 26) produced 278,930 reads (92.1% of total), with 126 genes captured, while sample XYL1374/17 (Cq 24) recovered 1,078,000 reads (96.7%) and 128 genes, both with nearly complete genome coverage. In contrast, uncaptured libraries such as


Funded by the European Union

XYL0078/24 (Cq 26) and IVIA8922 (Cq 36) yielded negligible recovery (3,707 and 10 reads, respectively, corresponding to ≤9 genes). Even in samples with higher Cq values (≥31), tetraplex captures enabled substantial enrichment. For example, IVIA8938 (Cq 32) recovered 232,443 reads (34.2%) and 130 genes, while IVIA30/24 (Cq 33) produced only 5,626 reads (1.6%) but still yielded 123 genes.

Insect-derived libraries also benefited greatly from TSCE, though with greater variability than plants, and on those samples capture efficiency varied according to input concentration and capture configuration. For example, sample XF4358/18 bis2 (Cq 26) produced 281,962 *Xf* reads (51.1%) with 131 genes recovered, while Xm1550 (Cq 30) yielded 107,665 reads (36.3%) with 128 genes. In higher Cq samples, the proportion of *Xf* reads was lower but gene recovery remained relatively robust. For example, Xm1215 (Cq 33) produced only 342 reads (2.8%) but still yielded 67 genes, while Xm1489 (Cq 31) recovered 52,689 reads (21.3%) and 125 genes. This illustrates the capacity of TSCE to extract meaningful genomic data from samples where shotgun sequencing would be uninformative (Table 16).

Across both plant and insect samples, the number of reads generally declined with increasing Cq values, reflecting lower bacterial loads. Low-Cq samples (≤28) yielded abundant reads and complete or near-complete gene recovery, while mid-range Cq values (30–32) were associated with reduced read numbers but still permitted recovery of 120–131 genes. Only very high Cq samples (>33) consistently showed limited gene recovery (e.g., 67 genes in Xm1215, 123 genes in IVIA30/24). This indicates that capture probes are broadly effective across the genome, allowing comprehensive recovery from samples with low bacterial loads (Table 16).

Importantly, TSCE enabled consistent MLST profiling across both plant and insect samples. Full seven-locus profiles were recovered in most cases, allowing unambiguous sequence type assignments, while uncaptured sequencing was always unsuccessful. Samples with complete recovery of the seven housekeeping loci could be assigned to known STs (e.g., ST1, ST6, ST7, ST16, ST53, ST78, ST80, ST81). In contrast, uncaptured libraries always failed to recover any loci, preventing MLST typing (Table 16). MLST typing was successful for the two blind samples ‘D’ sample from grapevine was assigned to ST2, while the blind CF5GrapSI2 sample was identified as belonging to ST16.

### 5.3.4 Phylogenetic reconstruction using *Xf*-TSCE

Beyond improving genome recovery and strain typing, *Xf*-TSCE also provided a valuable framework for phylogenetic inference. Traditional MLST, which relies on seven housekeeping genes, is widely used for assigning STs but has limited resolution for reconstructing evolutionary relationships, particularly among closely related lineages. In contrast, the broader set of loci retrieved through TSCE allows the inclusion of dozens of genes in comparative analyses, thereby increasing phylogenetic resolution and reducing biases associated with single-gene or few-gene datasets.



Funded by the European Union

**Table 16.** Comparison of TSCE results for plant and insect samples naturally infected by *Xylella fastidiosa* using different capture configurations

| Plant/Insect DNA sample | Cq (1ng/ul)[a] | Capture type[b] | Total number of reads | reads | (% of total) | Number of *Xf* captured genes | Mean Identity (%) | Mean Coverage (%) | MLST genes (ST)[c] |
|---|---|---|---|---|---|---|---|---|---|
| XF4358/18 bis2 | 26 | $t_2$ | 552,158 | 281,962 | 51.07 | 131 | 99.76 | 99.95 | 7C (ST1) |
| XYL5773 | 29 | $t_2$ | 2,612,560 | 406,115 | 15.54 | 131 | 99.87 | 99.99 | 7C (ST1) |
| Xm1215 | 33 | $t_6$ | 12,194 | 342 | 2.80 | 67 | 99.56 | 87.25 | 0C (ST-) |
| XYL2774/23 | 26 | s | 302,923 | 278,930 | 92.08 | 126 | 99.91 | 100 | 7C (ST1) |
| IVIA8926 | 31 | $t_8$ | 186,948 | 62,682 | 33.53 | 130 | 98.75 | 99.49 | 7C (ST6) |
| IVIA8938 | 32 | $t_4$ | 679,058 | 232,443 | 34.23 | 130 | 98.75 | 99.48 | 7C (ST6) |
| IVIA8922 | 36 | wc | 523,610 | 10 | 0 | 0 | - | - | 0C (ST-) |
| | | $t_5$ | 1,258,951 | 4,002 | 0.32 | 77 | 98.47 | 85.98 | 0C (ST-) |
| XYL583/16bis | 34 | $t_8$ | 668,906 | 338,472 | 50.60 | 129 | 98.76 | 99.95 | 7C (ST7) |
| XYL1374/17 | 24 | $t_6$ | 1,114,859 | 1,078,000 | 96.69 | 128 | 98.75 | 99.95 | 7C (ST7) |
| 2531 (2021) | 31 | $t_7$ | 416,420 | 174,082 | 41.80 | 128 | 98.72 | 99.95 | 7C (ST7) |
| 2017 (2021) | 35 | $t_8$ | 973,396 | 34,734 | 3.57 | 128 | 98.74 | 99.35 | 7C (ST7) |
| XYL610 | 28 | s | 310,937 | 214,142 | 68.87 | 128 | 98.75 | 99.95 | 7C (ST7) |
| XYL0895/20 | 26 | $t_6$ | 620,520 | 584,390 | 94.18 | 128 | 98.76 | 99.95 | 7C (ST81) |
| Xm1550 | 30 | $t_7$ | 296,967 | 107,665 | 36.25 | 128 | 98.73 | 99.95 | 7C (ST81) |
| XYL1050/19 | 27 | $t_3$ | 653,469 | 425,603 | 65.13 | 131 | 99.89 | 99.95 | 7C (ST1) |
| L3/0300 | 31 | wc | 421,030 | 386 | 0.09 | 4 | 99.27 | 100 | 0C (ST-) |
| | | $t_4$ | 276,146 | 87,755 | 31.78 | 129 | 98.76 | 99.88 | 7C (ST81) |
| CF5GrapSI2 | 28 | $t_6$ | 240,020 | 207,933 | 86.63 | 124 | 97.83 | 99.58 | 7C (ST16) |
| PUG18-193 | 26 | $t_2$ | 376,348 | 156,774 | 41.66 | 125 | 97.89 | 99.95 | 7C (ST53) |
| IVIA30/24 | 33 | s | 346,634 | 5,626 | 1.62 | 123 | 97.88 | 97.07 | 3C,4P (ST53?) |
| IVIA33/24 | 31 | s | 3,275,101 | 179,251 | 5.47 | 124 | 97.89 | 99.94 | 7C (ST53) |
| IVIA35/24 | 29 | s | 118,682 | 52,886 | 44.56 | 120 | 98.27 | 99.96 | 3C,4P (ST53) |
| XYL0078/24 | 26 | wc | 2,207,437 | 3,707 | 0.17 | 9 | 98.92 | 94.44 | 0C (ST-) |
| | | t | 763,697 | 291,327 | 38.15 | 125 | 97.89 | 99.98 | 7C (ST53) |
| Xm1489 | 31 | $t_7$ | 247,353 | 52,689 | 21.30 | 125 | 97.99 | 99.94 | 7C (ST80) |
| IVIA11394 | 31 | $t_8$ | 280,398 | 23,336 | 8.32 | 124 | 97.99 | 99.95 | 7C (ST80) |
| XYL4933/20 | 27 | s | 464,442 | 172,535 | 37.15 | 125 | 97.98 | 99.95 | 7C (ST80) |
| | | $t_3$ | 270,125 | 22,203 | 8.22 | 127 | 98 | 99.65 | 7C (ST80) |
| XYL0090/19 | 33 | wc | 571,248 | 58 | 0.01 | 5 | 98.45 | 97.8 | 0C (ST-) |
| | | $t_4$ | 1,000,816 | 13,379 | 1.34 | 127 | 97.98 | 99.67 | 7C (ST80) |
| D* | 29 | $t_7$ | 873,374 | 725,684 | 83.09 | 130 | 99.85 | 99.99 | 7C (ST2) |
| M9-2 | 23 | s | 346,634 | 264,511 | 76.31 | 123 | 97.84 | 99.98 | 7C (ST78) |
| F9P4 | 29 | wc | 1,133,939 | 144 | 0.01 | 0 | - | - | 0C (ST-) |
| | | s | 1,060,698 | 185,514 | 17.49 | 126 | 98.77 | 99.94 | 7P (ST90?) |
| F8P13 | 23 | s | 258,732 | 117,161 | 45.28 | 123 | 97.81 | 99.97 | C (ST78) |



Funded by the European Union

[a] Threshold quantitative cycle of the sample as determined by Harper-qPCR (Harper et al 2010). na= no amplification.

[b] Type of capture performed: s = singleplex, t= tetraplex, wc= without capture. Same subscript numbers indicate samples processed within the same pooling run of captures (i.e., the libraries were captured simultaneously or with other samples included in Table 13).

[c] The number of housekeeping genes captured is indicated (C)= correctly assigned to an allele; (P)=Partially assigned to the closest allele. When the seven housekeeping genes are captured the ST is provided; when all genes are captured although with no perfect matching the closest ST is indicated with ?.

Phylogenetic relationships among all analysed samples and publicly available genomes were reconstructed using maximum likelihood (ML) (Figure 12). The resulting tree revealed a clear clustering of isolates according to *Xf* subspecies, as indicated by branch colors. Samples were distributed across three major clusters, reflecting the underlying genetic diversity and separating the different subspecies within the phylogeny, even when based on a reduced dataset of 51 genes that were captured across all samples. This clustering highlights strong genetic differentiation among most subspecies, with the notable exception of *fastidiosa*, *sandyi*, and *morus*, which were not as clearly resolved.

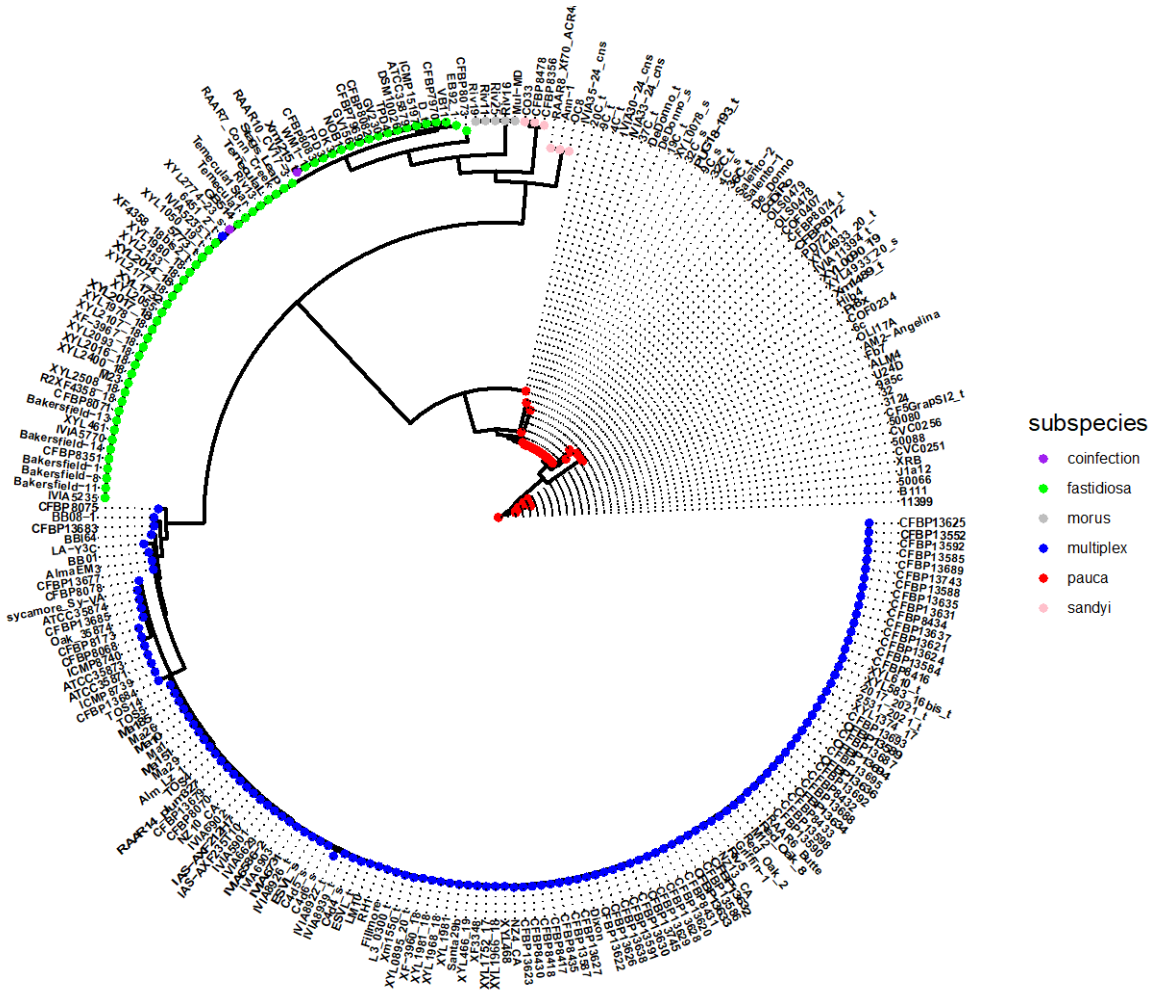



**Figure 11.** based on 51 core genes captured using *Xf*-TSCE across all samples in the study. Branch colors indicate *X. fastidiosa* subspecies.



Funded by the European Union

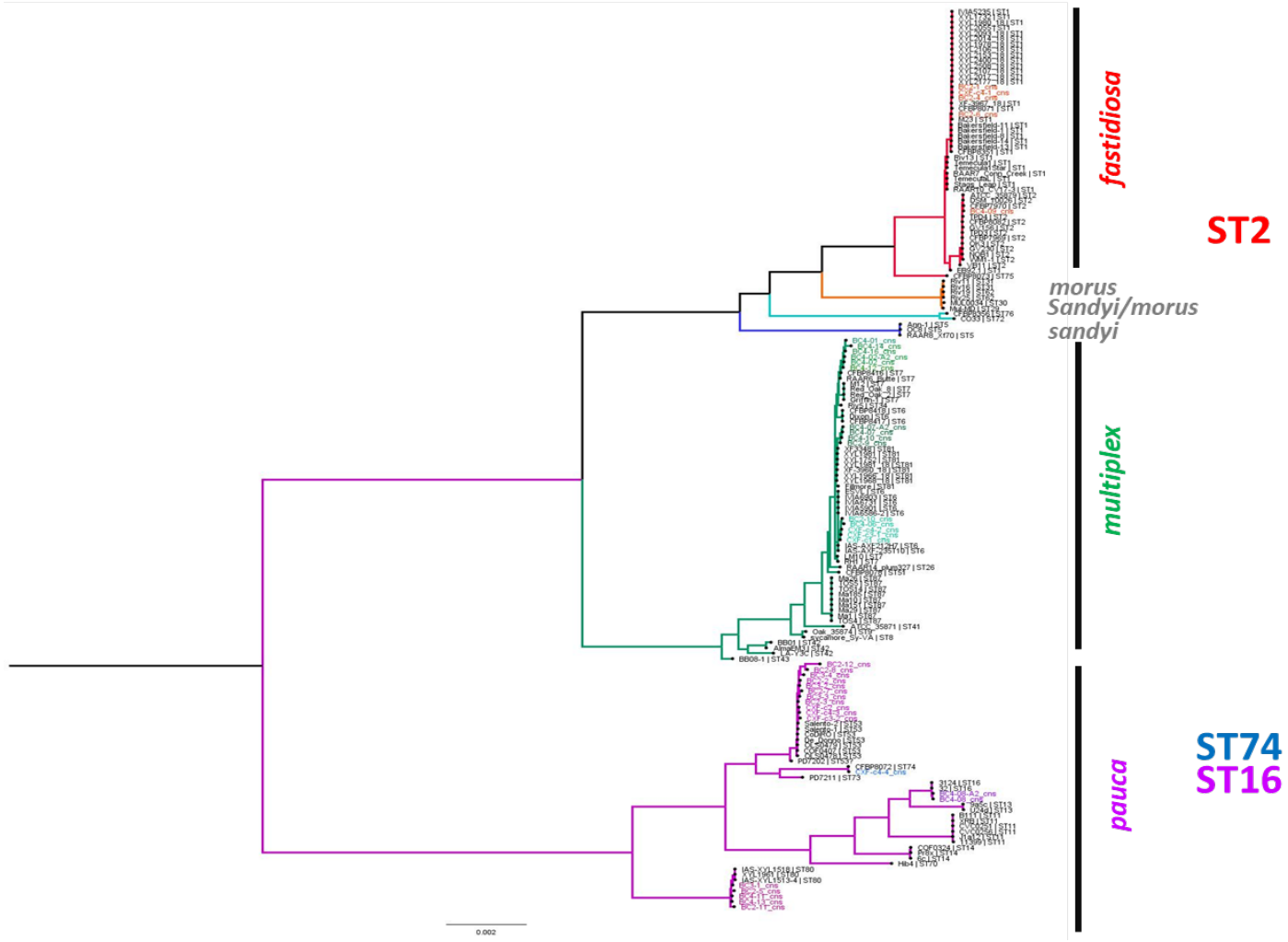



**Figure 12.** Maximum-likelihood phylogeny based on 122 core genes captured using *Xf*-TSCE from mock-inoculated, naturally infected plant samples and blind samples. Branch colors represent *Xylella fastidiosa* subspecies, and sequence types (STs) are indicated for blind samples.

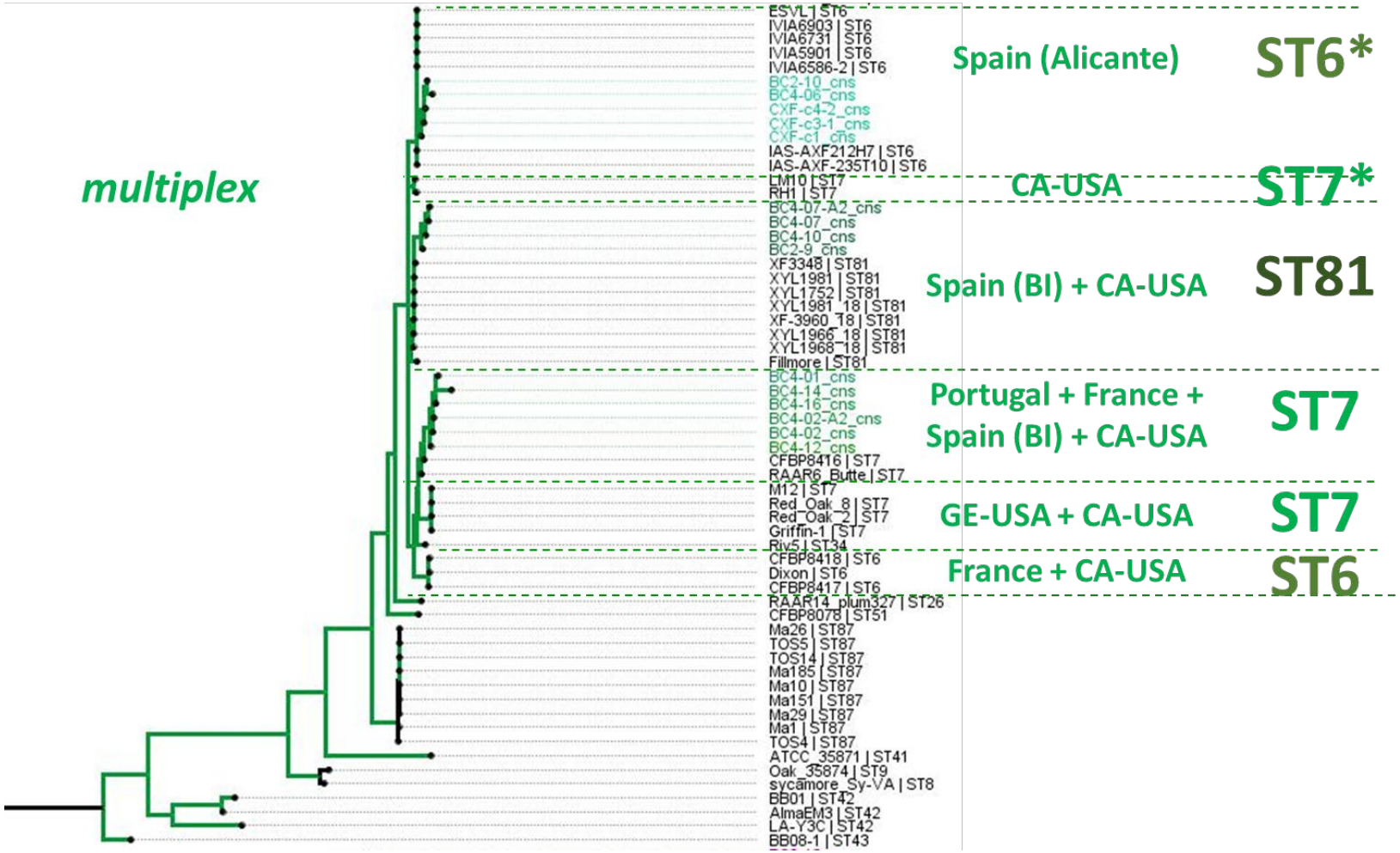



**Figure 13.** Partial zoom of the maximum-likelihood phylogeny shown in Figure 12 based on 122 core genes captured using *Xf*-TSCE from mock-inoculated and naturally infected plant samples belonging to subspecies *multiplex* (ST6, ST7, and ST81).

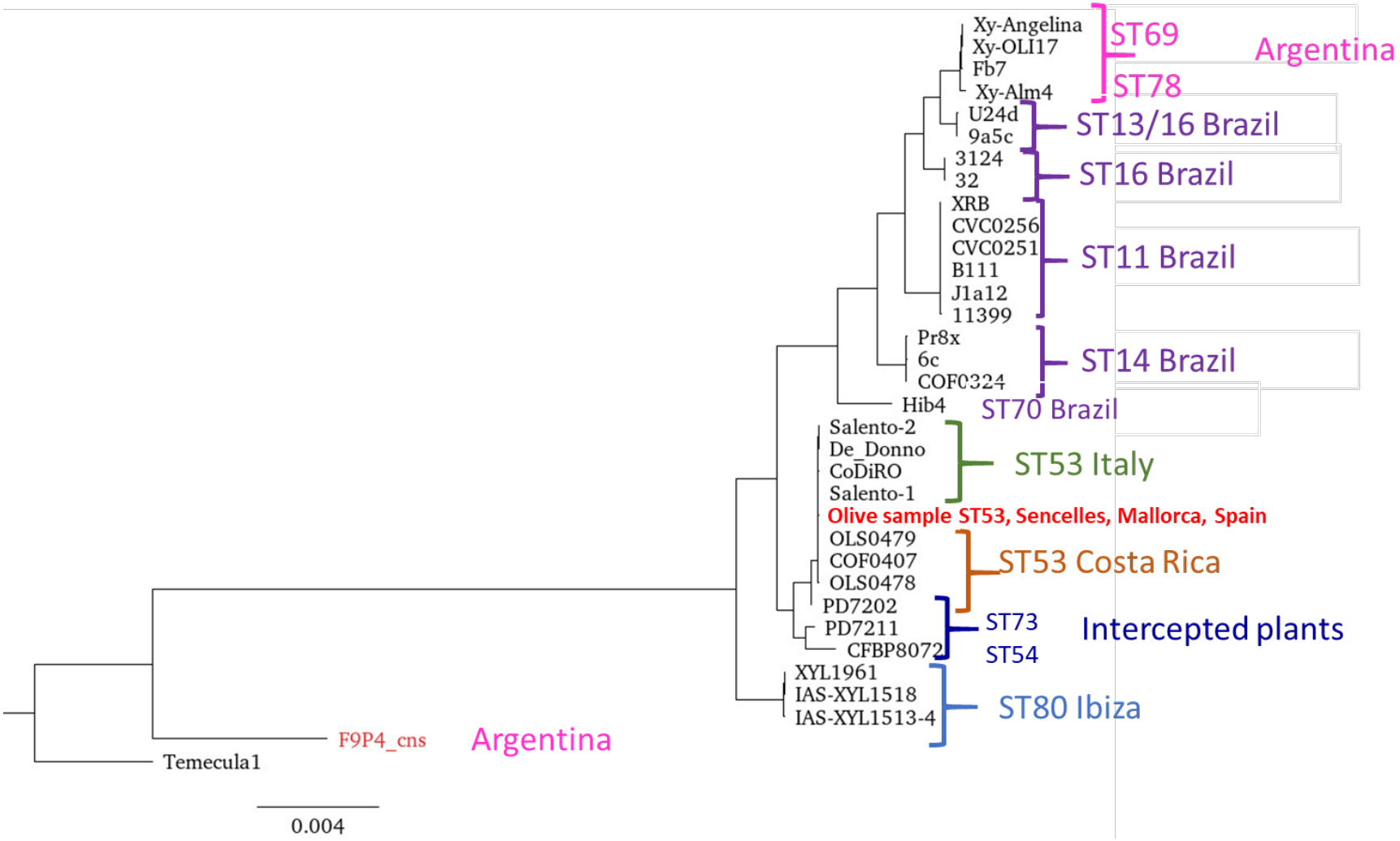



**Figure 13.** Maximum-likelihood phylogeny based on 133 core genes captured using *Xf*-TSCE from naturally infected plant samples infected by *Xf* isolates belonging to the ST13, ST14, ST16, ST53, ST69, ST73, ST74, ST80 and ST90 of subspecies *pauca*. Colors represent strains from same STs.

The maximum-likelihood tree reconstructed from 122 core genes captured by *Xf*-TSCE (Figure 12) showed clear clustering of isolates according to *Xf* subspecies. Mock-inoculated and naturally infected plant samples grouped with their expected subspecies clades and reference genomes, confirming the accuracy of the approach. Blind samples were consistently placed within their respective subspecies clusters, allowing robust sequence type inference (Figure 12). For example, sample D clustered within the *fastidiosa* group and was assigned to ST2, whereas the blind CFBP8074 sample grouped with ST74 and CF5GrapSI2 with strains from ST16.

Within the *multiplex* subspecies, distinct clusters corresponding to ST6, ST7, and ST81 were observed, illustrating the capacity of *Xf*-TSCE to resolve intra-subspecific lineages (Figure 13). Interestingly, all samples or strains from ST7 from France, Spain and Portugal clustered together and close to strains from same ST from Georgia and California, USA.

The maximum-likelihood phylogeny revealed that the Sencelles sample from Mallorca, isolated from olive previously assigned to ST53 by MLST, clustered almost identically with strains from Apulia, Italy, known to be responsible for the severe olive decline epidemic in that region. This strong phylogenetic similarity underscores the genetic identity between the Mallorcan and Italian ST53 populations and highlights the epidemiological relevance of this lineage for olive cultivation in the Mediterranean basin.

In contrast, the sample from olive in Argentina assigned to ST90 occupied an unusual position within the tree. Unlike other well-defined clades, ST90 appeared intermediate between the *fastidiosa* and *pauca* subspecies. Its placement suggests a history of extensive recombination, reflecting genetic exchange across subspecies boundaries. This atypical pattern supports previous evidence that recombination plays a major role in *Xf* evolution and may complicate subspecies-level classification.

Overall, the phylogeny demonstrates that the capture of over 122 genes by *Xf*-TSCE provides robust resolution across subspecies and sequence types, allowing not only confirmation of sample identity but also fine-scale phylogenetic placement, even for unknown or blind samples. This level of resolution extends beyond what is typically achievable with MLST, which relies on only seven housekeeping loci.

# 6. A FLOW-FISH METHOD FOR DETECTION AND QUANTIFICATION OF *XYLELLA FASTIDIOSA*

## 6.1 Background

Fluorescence *In Situ* Hybridization (FISH) is a molecular tool traditionally coupled with microscopy. It enables the direct identification of microorganisms within complex samples within short processing times. FISH has led to the development of numerous protocols for both fundamental and applied research. Within *Xf* pathosystems, has been used to localize the bacterium in planta by combining FISH-microscopy with histology in diseased olive plants (Cardinale et al, 2018), and to estimate its abundance in honeydew and plant sap (Rodrigues et al., 2006). Similar applications have been reported in other bacterial models (Freen-van-Heeren, 2021).



Funded by the European Union

Coupling FISH with flow cytometry (FLOW-FISH) provides a powerful molecular and cellular microbiology approach, allowing high-throughput single-cell analysis of large populations. Although initially applied to eukaryotic cells, FLOW-FISH applications for microorganisms have expanded considerably in recent years (Freen-van-Heeren, 2021). In the case of *Xf*, implementing FLOW-FISH offers a complementary strategy to current research and diagnostic methods for the study of *Xf*-plant and *Xf*-insect interactions. FLOW-FISH allows absolute specific quantitative detection of *Xf* from different plant matrices, absolute quantitative comparison of the presence of *Xf* versus other endophytic populations, in asymptomatic and diseased plants, complementing relative microbiome composition analysis.

In Costa Rica, *Xf* subsp. *fastidiosa* is endemic. Coffee plantations show high prevalence and bacterial load, but disease symptoms are rarely observed (Montero-Astúa et al, 2008). Here, we investigated the extent of bacterial infection in coffee and its insect vectors. Our aim was to establish a FLOW-FISH protocol as proof of concept for studying *Xf*–coffee interactions and for its direct detection, with potential diagnostic value in endemic areas. We optimized a FLOW-FISH procedure using an *Xf*-specific probe (KO210) and a universal bacterial probe (Eub338) *Xf*-specific probe (KO210) and a universal probe (Eub338). Probe working conditions were validated against FISH-microscopy using pure cultures and spiked coffee matrices, and detection limits were determined. The resulting protocol broadens the toolkit available for research and diagnostics of *Xf*, providing a high-resolution qualitative and quantitative tool.

## 6.2 Materials and Methods

### 6.2.1 Bacterial culture and plant material

A *Xf* subsp. *fastidiosa* endemic strain (*XF*-1093), was used in this study for optimization of FLOW-FISH. This strain was genomically characterized in a previous study (Castillo et al., 2020). *XF*-1093 was grown for seven days on BCYE solid medium before inoculum preparation for FISH protocols. A panel of bacteria was used for probe testing and specificity, including other endemic *Xf* subsp. *fastidiosa* strains (*XF*-68, *XF*-1090, *XF*-1094, and *XF*-1105) as well as other bacterial species including *Stenotrophomonas maltophilia* and *Xanthomonas campestris*, that are among the closest relatives to *Xf; Ralstonia solanacearum,* a pathogen that can colonize xylem vessels; *Methylobacterium extorquens,* a known endophytic bacterium present in coffee, and *Escherichia coli* and *Staphylococcus aureus*.

Bacterial cultures were grown to the exponential phase and were harvested by centrifugation at 5000 g for 10 min, washed with sterile phosphate-buffered saline (PBS), and fixed with 4% (wt/vol) paraformaldehyde in PBS for 20 min. After fixation, the bacteria were washed with PBS and stored in 50% ethanol–PBS until FISH procedure. Previous to fixation an aliquot of the bacterial suspension was used to determine the cell concentration (UFC/ml) for each bacterium used in the panel.

Coffee plants were grown from seed in greenhouse conditions, and plants were tested for absence of *Xf* by using the Francis-qPCR assay (Francis et al., 2006) before their use as a source of coffee plant matrix.



Funded by the European Union

### 6.2.2 Optimization of FISH and FLOW-FISH protocol using *Xylella fastidiosa*-specific and eubacteria probes

Figure 15 shows a schematic overview of the procedures used for FLOW-FISH. Bacteria were pelleted by centrifugation (microcentrifuge; 5000g, 5 min, 4°C), washed with PBS, and resuspended in 100 µL of 0.5% Triton-PBS (Sigma-Aldrich Corporation, Burlington, Massachusetts) and 100 µL of lysozyme (10mg/mL in PBS; Sigma L6876). The samples were incubated for 30 min at 37°C in a Thermo Mixer (Eppendorf). Pellets were washed in 1 mL of PBS and then resuspended in 50 µL of Hybridization Buffer (HB; 20mM Tris-HCl, 3M NaCl, 0.01% SDS, [pH=7.5] filtered with 0.2µm filter).

The *Xf*-specific probe KO210 (5´-GCCACCCATGGTATTACTACC-3´) (Franke-Whittle et al., 2005) labeled with A647 and a universal bacterial probe EUB338 (5´-GCTGCCTCCCGTAGGAGT-3´) (Amann et al., 1990) labeled with Cy3, were added at the appropriate amount to the HB. Samples were incubated for 3h at 42°C with or without probe (negative control) in dark conditions in constant shaking at 300 rpm. Non-bound probe was removed by adding 1 mL of HB to each tube for an additional 45 min incubation. After centrifugation, pellets were washed with 1 mL of PBS and resuspended in PBS for data acquisition.

Data acquisition was performed using a Cytoflex flow cytometer (Beckman Coulter) equipped with Violet-Blue-Red lasers or was analyzed using a Nikon Eclipse 80i fluorescence microscope (magnification: 60x oil iris). For FLOW-FISH assessment and detection limits in coffee matrix, coffee extractions were spiked with known concentrations of *Xf* before the start of the protocol.

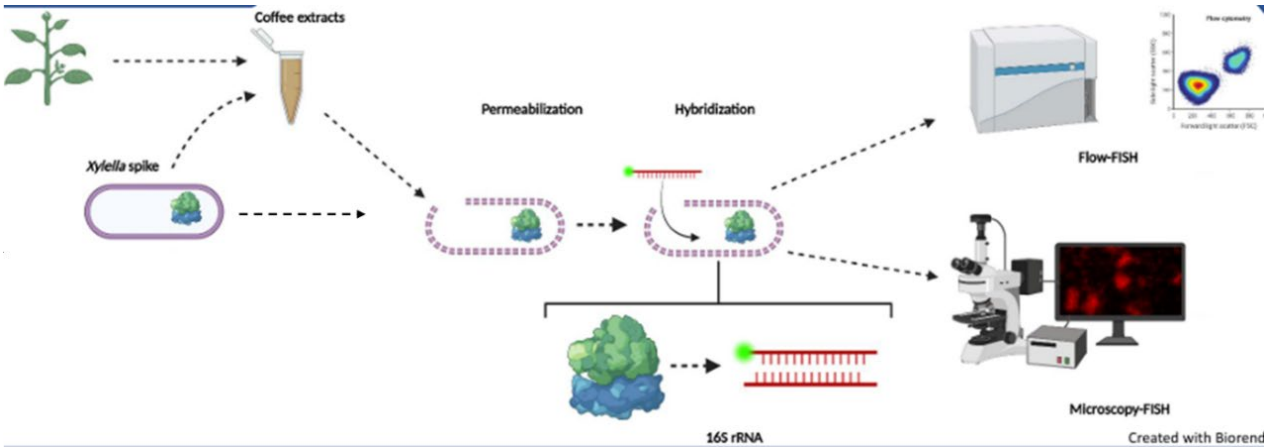



**Figure 15.** Schematic representation of the methodology performed for FLOW-FISH optimization.



Funded by the European Union

## 6.3 Results

### 6.3.1 Evaluation of probe specificity and cross-reactivity for FLOW-FISH optimization

The specificity and cross-reactivity of the FISH *Xf* (KO210) and eubacteria (Eu388) probes was tested against pure cultures of *Xf* (*XF*-68, *XF*-1090, *XF*-1093, *XF*-1094, and *XF*-1105) and a panel of different bacteria. *Xf* KO210 probe recognized all endemic *Xf* strains. For further analysis, we selected *XF*-1093 as a reference because it was isolated from coffee and is a well-characterized strain. The *Xf* KO210 probe did not cross-react with the other tested bacterial species (Figure 16A), ensuring that the probe is specific for *Xf* under the tested conditions. The EU338 probe, as expected, recognized all the bacteria assayed in the panel; *S. maltophilia* represents an example of a positive signal seen for all tested bacterial species (Figure 16B).

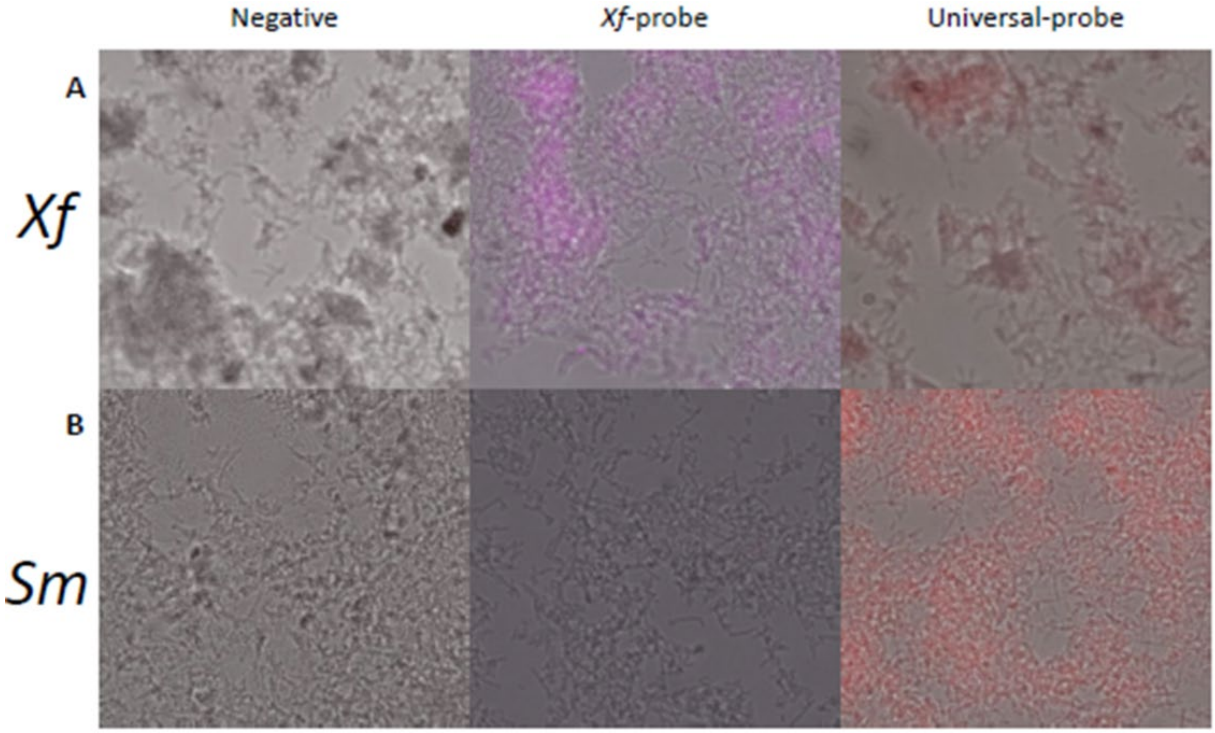



**Figure 16.** Microscopy-FISH probe specificity tested against *Xylella fastidiosa* (*Xf*) cells (**A**) and *S. maltophilia* (*Sm*) (**B**) cells. Evaluation of *Xf* K0210 coupled with Alexa647 (middle panel) and Eub338 coupled with Cy3 probes (right panel). Left panel: negative control. Preparations were analyzed using a Nikon Eclipse 80i fluorescence microscope (magnification: 60x oil immersion).



Funded by the European Union

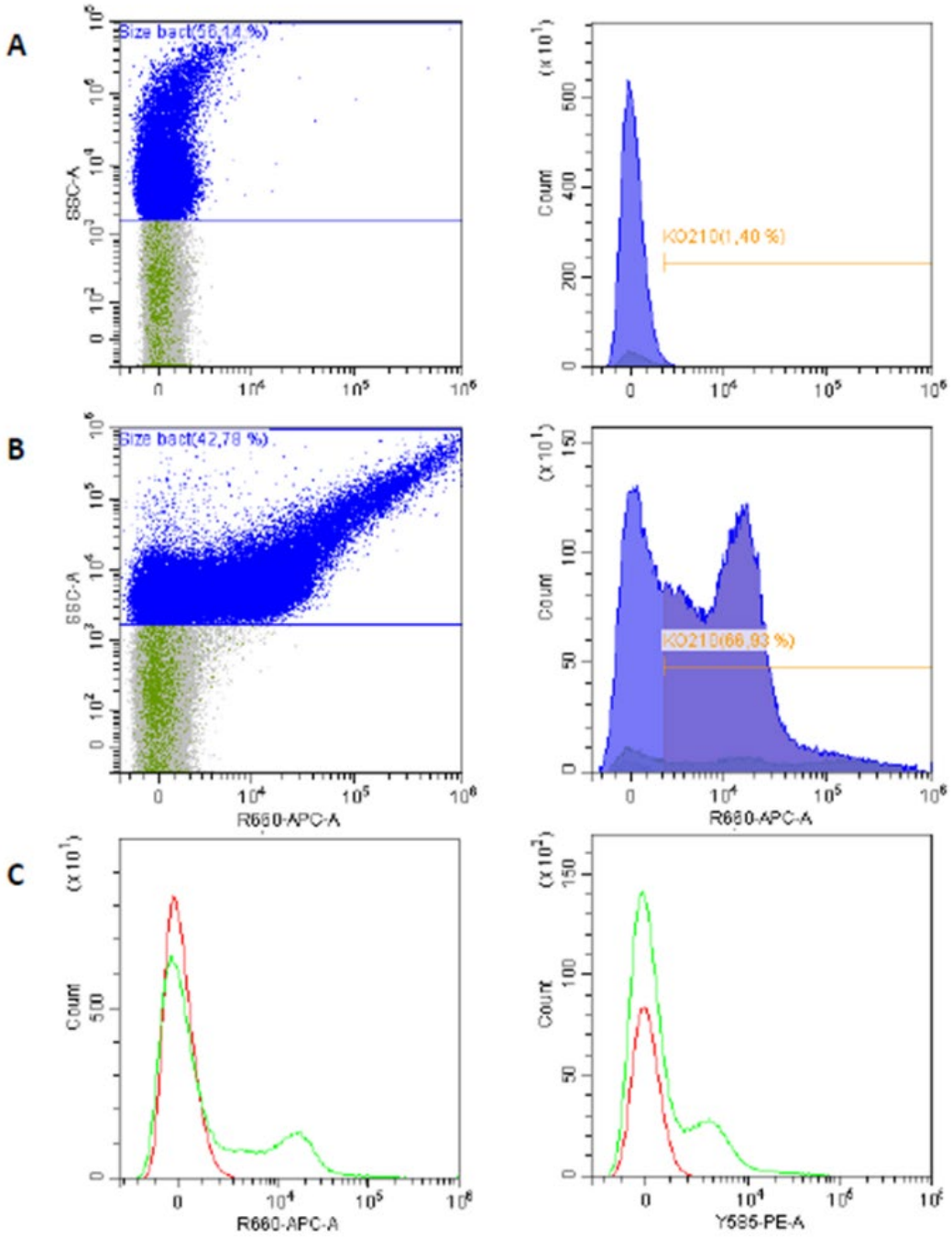



**Figure 17.** FLOW-FISH performance for K0210 and Eub338 probes against *Xylella fastidiosa* (*Xf*). **A.** Fixed *Xf XF*-1093 cells without probe. Left panel: scatterplots of SSC versus baseline fluorescence intensity of Alexa 647 (R660-APC channel). A size gate (blue) was defined to reduce the buffer background signal. Right panel: histogram showing baseline fluorescence intensity for Alexa 647 signal of the size-gated population. A *Xf* KO210 positive gate (orange) was defined considering basal autofluorescence. **B.** Fixed *Xf* incubated for 3h at 42°C with KO210-A647 probe. Left panel: scatter plots of SSC versus the fluorescence intensity of Alexa 647. As expected, the size gate (blue) encompasses most of the marked bacteria. Right panel: histogram for the fluorescence intensity of Alexa 647 of the size-gated population. Most bacteria (around 70%) were KO210 positive. **C.** Overlay histograms of fixed *Xf* incubated for 3h at 42°C with probe (green) and without probe (red). Left panel: using KO210-A647; and right panel: using Eub338-Cy3, fluorescence intensity detected in Y585-PE channel.



Funded by the European Union

After confirming the optimal performance of the *XF*-KO210 and Eub338 probes in FISH-microscopy, the procedure was adapted to flow cytometry to assess FLOW-FISH performance. Fluorescence intensity was analyzed in relation to side scatter (SSC). Bacterial *XF*-1093 without probe was used for the definition of a size gate (blue) (Figure 17A) to reduce background signal, as well as a positive gate (orange) (Figure 17A, left panel), to establish baseline autofluorescence. When marked with *Xf*-KO210 probe, the majority of *XF*-1093 were within the established size gate, representing 70% of the population according to the specific fluorescence signal (Figure 17B). Under these conditions, the KO210-specific probe outperformed the Eu338 probe in resolving positive populations (Figure 17C).

### 6.3.2 Optimization of XF-KO210 and Eub338 probe concentrations and interaction analysis

To evaluate if the excess of *Xf* K0210 probe could impact and induce unspecific signals *E. coli* cells were used with different concentrations of *Xf*-KO210 probe. An increase in basal fluorescence was observed at high concentrations (Figure 18A). Based on these results, probe titration was carried out to determine the concentrations that yielded a strong specific signal without increasing basal fluorescence. The optimal concentration was 0.5 µmol/L for the *XF*-KO210 probe and 0.375 µmol/L for the Eu388 probe (Figure 18B, upper and lower panel). Finally, FLOW-FISH results were compared and validated against FISH microscopy (Figure 18C and 18D).

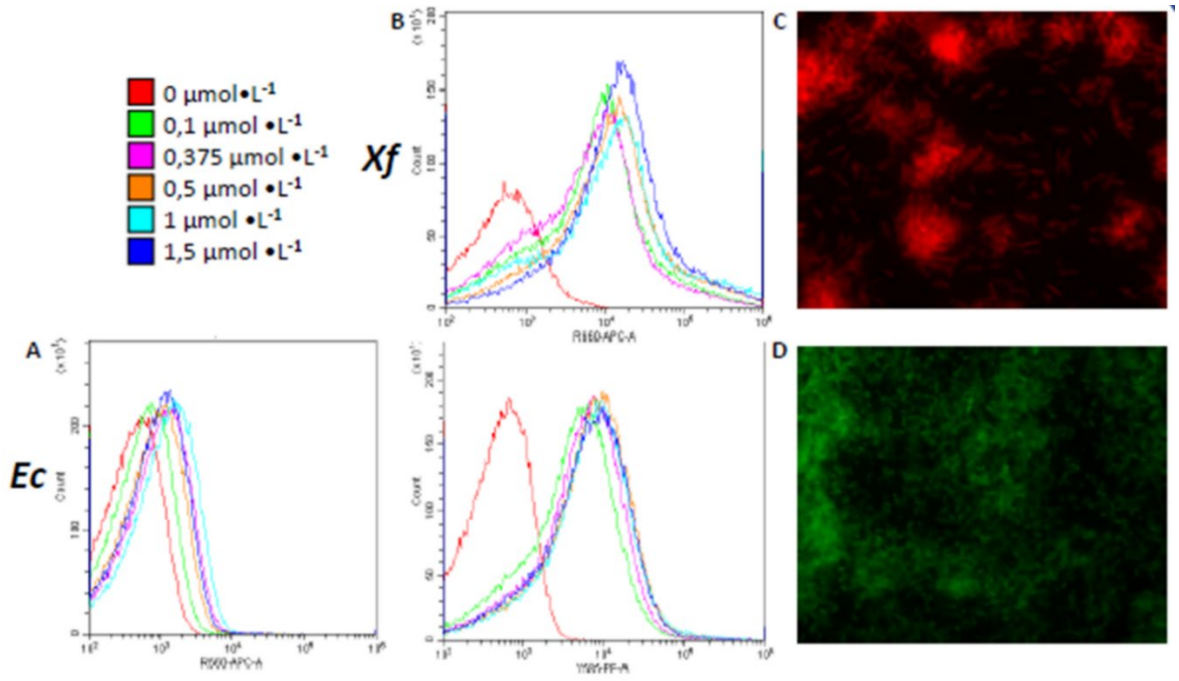



**Figure 18.** Probe titration for optimal FLOW-FISH performing concentrations using fixed exponential growth phase pure bacterial cultures incubated for 3h at 42°C with different probe concentrations. (**A**) Flow-FISH analysis of *Echerichia coli* with *Xf*-KO210 probe at different concentrations showing an increase in autofluorescence due to unspecific probe binding. (**B**) Flow-FISH analysis of *Xf*-KO210 probe titration with *Xylella fastidiosa* (*Xf*) (upper panel) and Eub338 probe titration with *E. coli* (Lower panel). Microscopy-FISH observation with the selected probe concentrations of *Xf*-KO210 probe against *Xf* (**C**), and Eub338 probe against *E. coli* (**D**).



Funded by the European Union

Additionally, we evaluated the interaction of using the specific *Xf* K0210 probe simultaneously with the universal eubacteria Eu388 probe. The probes did not interfere with each other; the signal for the *Xf* KO210 probe was the same as for the Eu388 probe (Figure 19A). There were no problems concerning fluorescent signal or competition; the merged image correlated with positive staining for both probes. This result was also seen for the FLOW-FISH double positive population (Figure 19B). This is an important result since one of the intended applications for *Xf* FLOW-FISH is to determine the absolute number of *Xf* cells versus the total bacterial population in healthy plants as well as in asymptomatic and symptomatic infected plants.

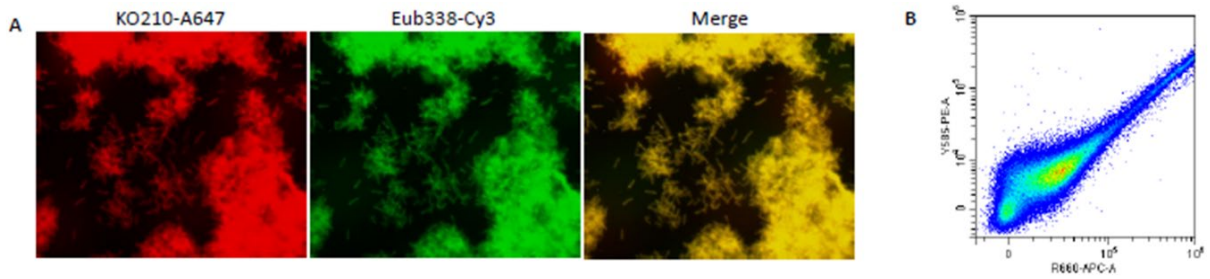



**Figure 19.** Probe interaction analysis. *Xylella fastidiosa* cells were incubated for 3h at 42°C with selected probe concentrations simultaneously. All bacterial cells were positive for both *Xf*-KO210 and Eub338 probes as demonstrated by Microscopy-FISH observation (**A**) and Flow-FISH (**B**).

### 6.3.3 Performance evaluation of *Xylella fastidiosa* FLOW-FISH in coffee matrices spiked with *Xylella fastidiosa* cells

The *Xf* FLOW-FISH optimized protocol was tested against *Xf* spiked coffee matrices. Coffee petioles were macerated and used immediately. High concentrations of the coffee matrix impacted *Xf* detection. In a 1:50 *Xf* cells:coffee matrix ratio, a discrete positive population was detected (Figure 20A, upper and lower panel), representing 32% of the positive population. When the ratio was 1:1 *Xf* cells:coffee matrix, the positive population was doubled to 60% (Figure 20B, upper and lower panel).

Employing crude matrix proportions diminished FISH and FLOW FISH power, due to clumping of cells within the excess matrix. The implementation of sonication and a low-speed centrifuge step to eliminate debris, at the beginning of the FLOW-FISH protocol, improved the yield of detected positive populations. FISH microscopy revealed the complex nature of the coffee matrix and potential interference in FLOW-FISH analysis.

### 6.3.4 Detection limit of *Xylella fastidiosa* FLOW-FISH in coffee extracts

To assess detection limits, coffee extracts were spiked with different concentrations of *Xf* cells ranging from $10^7$ to $10^2$. A time gate was established to ensure that data capture represents optimal flow conditions and, for accurate comparative data analyses (Figure 21A). We were able to determine that our FLOW-FISH protocol has a detection limit of $10^5$ bacteria (Figure 21B and 21C). Spiked coffee matrix with lower concentrations did not give a positive signal according to our gating parameters. This data is important to value its potential uses in different diagnostics or experimental applications.



Funded by the European Union

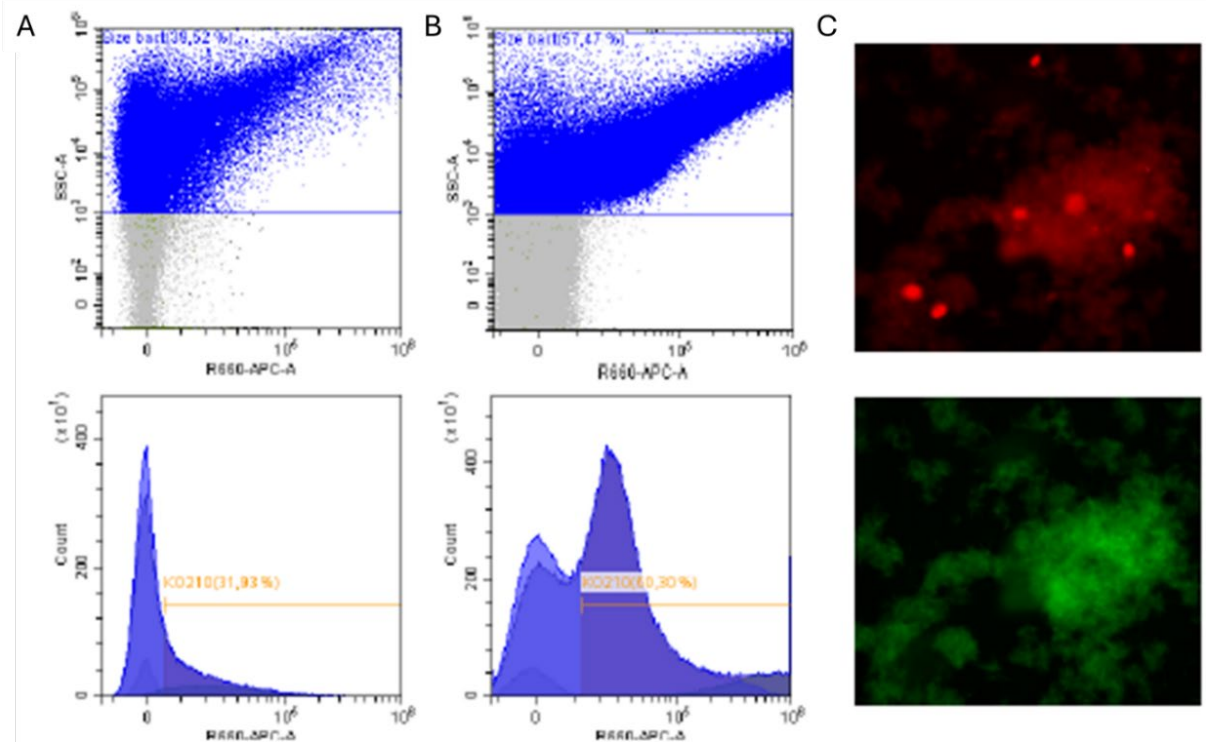



**Figure 20**. Performance of *Xylella fastidiosa* FLOW-FISH in spiked *Xf* coffee matrix. (**A**) Flow-FISH analysis at a 1:50 *Xf* cells:coffee matrix ratio. Upper panel: scatter plots showing the size-gated population(blue) with a reduced KO210-positive population. Lower panel: histogram of the same population, with only 32% *Xf* KO210 positive bacteria. (**B**) Coffee matrix at a 1:1 *Xf* cells:coffee matrix ratio. Upper panel: scatter plots showing the size-gated population(blue) with a significant KO210 positive population. Lower panel: histogram of the same population, with about 60% KO210-positive bacteria, doubling the signal obtained for the 1:50 ratio. (**C**) Microscopy-FISH observation of 1:1 *Xf* cells:coffee matrix ratio, simultaneously hybridized with both KO210 and Eub338. Upper panel: KO210-A647 fluorescence signal; and lower panel: Eub338-Cy3 fluorescence signal.

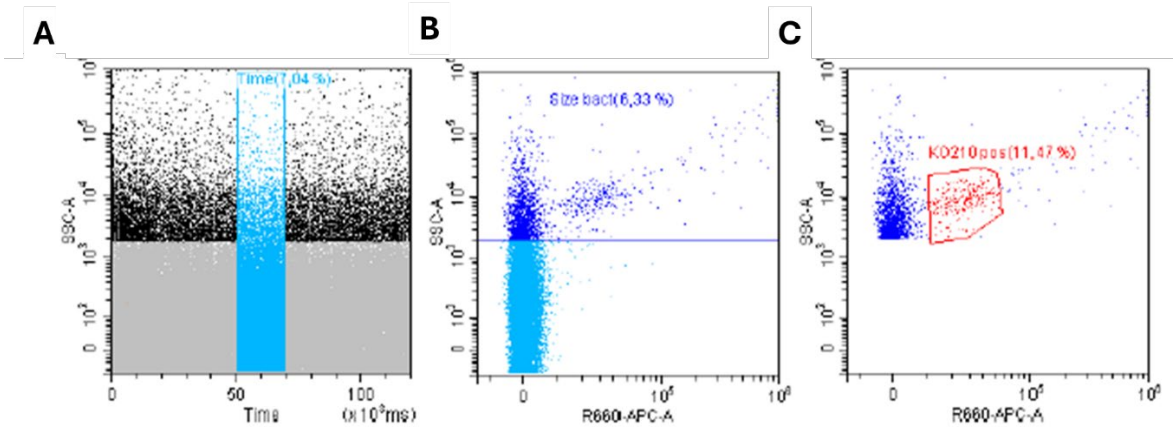



**Figure 21.** Detection limit of *Xylella fastidiosa* in coffee extracts by *Xf* FLOW-FISH. *Xf* detection limits were assessed using spiked coffee extracts, incubated for 3h at 42°C. Standardized acquisitions performed establishing a (**A**) time gate (light blue) and (**B**) sizegate. (**C**) *Xf*-KO210 positive populations (red) were quantified through concentrations ranging from $10^7$ to $10^3$, showing differences from background noise at concentrations over $10^5$. Images represent the results of three independent experiments.



Funded by the European Union

# 7. CONCLUSSIONS

- **ddPCR** proved to be a highly sensitive and precise tool for the detection of *Xf* in both host plants and insect vectors, allowing absolute quantification and improved diagnostic confidence, particularly in samples with low bacterial load.
- **Colorimetric LAMP** was successfully adapted for rapid field-level detection of *Xf* in almond orchards, offering a simple, low-cost, and portable alternative to laboratory-based methods, thus enhancing surveillance capacity in endemic areas.
- **Modeling approaches** applied to naturally infected almond samples provided robust estimates of diagnostic test performance, demonstrating the importance of statistical validation for optimizing sampling strategies and for improving the reliability of molecular detection in heterogeneous field conditions.
- The development of **targeted next-generation sequencing enrichment** (Xf-TSCE) enabled high-resolution genome recovery from complex plant and insect matrices, outperforming traditional MLST by providing improved phylogenetic resolution and accurate sequence typing, even from samples with low bacterial loads.
- The implementation of a **FLOW-FISH** assay expanded the diagnostic toolbox by enabling single-cell detection and quantification of *Xf* populations within plant and insect tissues, thereby complementing existing molecular methods with a technique that combines specificity, sensitivity, and cellular-level resolution.

# 8. PUBLICATIONS FROM THIS DELIVERABLE

The following publications have been compiled for this deliverable D2.2:

- Amoia, S.S.; Minafra, A.; Ligorio, A.; Cavalieri, V.; Boscia, D.; Saponari, M.; Loconsole, G. Detection of *Xylella fastidiosa* in Host Plants and Insect Vectors by Droplet Digital PCR. Agriculture 2023, 13, 716.
- Serena, A.S.; Falcón-Piñeiro, A.; Pastar, M.; Garcìa-Madero, J.M.; Contaldo, N.; Muegge, M.; Compant, S.; Saldarelli, P.; Minafra, A. Naked-Eye Molecular Testing for the Detection of *Xylella fastidiosa* in Mallorca (Balearic Island) Almond Orchards by Colorimetric LAMP. Appl. Sci. 2025, 15, 739.
- Velasco Amo, M.P.; Olivares-García, C.; Román-Écija, M.; Marco-Noales, E.; Navas-Cortés J.A.; Landa B.B. Modeling the accuracy of *Xylella fastidiosa* molecular diagnostic tests in naturally-infected almond tree samples. Plant Disease 2025 0:1-12.
- Arroyo-Portilla C, Chaves-González L, Zúñiga Pereira A, Garita L, Montero-Astúa M, Chacón-Díaz C. 2023. Fluorescence In situ hybridization coupled to Flow cytometry and cell imaging analyses as potential tools for *Xylella fastidiosa* research and diagnostics. 4th European Conference on *Xylella fastidiosa* August 19 and 20, 2023. Lyon, France". A manuscript with the final optimization and application of this tool is in preparation.
- Velasco Amo, M.P.; Arias-Giraldo, L.F., Landa B.B. Kigh-resolution diagnosis of *Xylella fastidiosa* using targeted NGS enrichment. A manuscript is in preparation.
-



Funded by the European Union

# 9. REFERENCES

Aglietti, C., Luchi, N., Pepori, A. L., Bartolini, P., Pecori, F., Raio, A., Capretti, P., and Santini, A. 2019. Real-time loop-mediated isothermal amplification: an early-warning tool for quarantine plant pathogen detection. AMB Express 9:50.

Almeida, R. P. P. 2016. Can Apulia's olive trees be saved? Science 353:346–348.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R., Stahl, D.A. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol. 56(6):1919-25.

Amoia, S. S., Loconsole, G., Ligorio, A., Pantazis, A. K., Papadakis, G., Gizeli, E., Minafra, A. 2023a. A colorimetric LAMP detection of *Xylella fastidiosa* in crude alkaline sap of olive trees in Apulia as a field-based tool for disease containment. Agriculture 13:448.

Amoia, S. S., Minafra, A., Ligorio, A., Cavalieri, V., Boscia, D., Saponari, M., and Loconsole, G. 2023b. Detection of *Xylella fastidiosa* in host plants and insect vectors by droplet digital PCR. Agriculture 13:716.

Anguita-Maeso, M., Ares-Yebra, A., Haro, C., Román-Écija, M., Olivares-García, C., Costa, J., Marco-Noales, E., Ferrer, A., Navas-Cortés, J. A., and Landa, B. B. 2022. *Xylella fastidiosa* infection reshapes microbial composition and network associations in the xylem of almond trees. Front. Microbiol. 13:866085.

Baldi, P., and La Porta, N. 2017. *Xylella fastidiosa*: Host range and advance in molecular identification techniques. Front. Plant Sci. 8:944.

Baldi, P.; La Porta, N. 2020. Molecular Approaches for Low-Cost Point-of-Care Pathogen Detection in Agriculture and Forestry. Front. Plant Sci. 11:570862.

Bangdiwala, S. I., Haedo, A. S., Natal, M. L., and Villaveces, A. 2008. The agreement chart as an alternative to the receiver-operating characteristic curve for diagnostic tests. J. Clin. Epidemiol. 61:866–874.

Bonants, P., Griekspoor, Y., Houwers, I., Krijger, M., van der Zouwen, P., van der Lee, T. A. J., and van der Wolf, J. 2019. Development and evaluation of a triplex TaqMan assay and next-generation sequence analysis for improved detection of *Xylella* in plant material. Plant Dis. 103:645–655.

Boutigny, A., Remenant, B., Legendre, B., Beven, V., Rolland, M., Blanchard, Y., and Cunty, A. 2023. Direct *Xylella fastidiosa* whole genome sequencing from various plant species using targeted enrichment. J. Microbiol. Methods 208:106719.

Burbank, L. P., and Ortega, B. C. 2018. Novel amplification targets for rapid detection and differentiation of *Xylella fastidiosa* subspecies *fastidiosa* and *multiplex* in plant and insect tissues. J. Microbiol. Meth. 155:8–18.

Byrt, T., Bishop, J., and Carlin, J. 1993. Bias, Prevalence and Kappa. J. Clin. Epidemiol. 46:423–429.

Camino, C., Araño, K., Berni, J. A., Dierkes, H., Trapero-Casas, J. L., León-Ropero, G., Montes-Borrego, M., Román-Écija, M., Velasco-Amo, M. P., Landa, B., Navas-Cortes, J. A., and Beck, P. S. A. 2022. Detecting *Xylella fastidiosa* in a machine learning framework using Vcmax



Funded by the European Union

and leaf biochemistry quantified with airborne hyperspectral imagery. Remote Sens. Environ. 282:113281.

Cardinale, M., Luvisi, A., Meyer, J.B., Sabella, E., De Bellis, L., Cruz, A.C., Ampatzidis, Y., Cherubini, P. 2018. Specific Fluorescence *in Situ* Hybridization (FISH) Test to Highlight Colonization of Xylem Vessels by *Xylella fastidiosa* in Naturally Infected Olive Trees (*Olea europaea* L.). Front. Plant Sci. 6;9:431.

Castillo, A.I., Chacón-Díaz, C., Rodríguez-Murillo, N., Coletta-Filho, H.D., Almeida, R.P.P. 2020. Impacts of local population history and ecology on the evolution of a globally dispersed pathogen. BMC Genomics. 21(1):369.

Cendoya, M., Lázaro, E., Navarro-Quiles, A., López-Quílez, A., Conesa, D., and Vicent, A. 2024. Performance of outbreak management plans for emerging plant diseases: the case of almond leaf scorch caused by *Xylella fastidiosa* in mainland Spain. Phytopathology 114:1566-1576.

Cesbron, S., Dupas, E., Beaurepère, Q., Briand, M., Montes-Borrego, M., Velasco-Amo, M. del P., Landa, B. B., and Jacques, M.-A. 2020. Development of a Nested-MultiLocus Sequence Typing approach for a highly sensitive and specific identification of *Xylella fastidiosa* subspecies directly from plant samples. Agronomy 10:1099.

Cesbron, S., Dupas, E., and Jacques, M.-A. 2023. Evaluation of the AmplifyRP® XRT+ kit for the detection of *Xylella fastidiosa* by Recombinase Polymerase Amplification (RPA). PhytoFrontiers 3:225-234.

Chikere, C. M. U., Wilson, K., Graziadio, S., Vale, L., and Allen, A. J. 2019. Diagnostic test evaluation methodology: A systematic review of methods employed to evaluate diagnostic tests in the absence of gold standard – An update. PLoS ONE 14:e0223832.

Cornara, D., Boscia, D., D'Attoma, G., Digiaro, M., Ligorio, A. M., Loconsole, G., Minutillo, S. A., Montilon, V., Palmisano, F., Ragone G., Silletti, M. R., Verrastro, V., Zaza, C., and Saponari, M. 2024. An integrated strategy for pathogen surveillance unveiled *Xylella fastidiosa* ST1 outbreak in hidden agricultural compartments in the Apulia region (Southern Italy). Eur. J. Plant Pathol. 1-9.

Cunty, A., Legendre, B., de Jerphanion, P., Juteau, V., Forveille, A., Germain, J.-F., Ramel, J.-M., Reynaud, P., Olivier, V., and Poliakoff, F. 2020. *Xylella fastidiosa* subspecies and sequence types detected in *Philaenus spumarius* and in infected plants in France share the same locations. Plant Pathol. 69:1798–1811.

Cunty, A., Legendre, B., de Jerphanion, P., Dousset, C., Forveille, A., Paillard, S., & Olivier, V. 2022. Update of the *Xylella fastidiosa* outbreak in France: two new variants detected and a new region affected. Eur. J. Plant Pathol. 163:505-510.

D'Attoma, G., Morelli, M., Saldarelli, P., Saponari, M., Giampetruzzi, A., Boscia, D., Savino, V.N., De La Fuente, L., Cobine, P.A. 2019. Ionomic differences between susceptible and resistant olive cultivars infected by *Xylella fastidiosa* in the outbreak area of Salento, Italy. Pathogens 8:272.

Davis, M. J., Purcell, A. H., and Thomson, S. V. 1980. Isolation media for the Pierce's Disease bacterium. Phytopathology 70:425–429.



Funded by the European Union

De Boer, S.H., López, M.M. 2012. New grower-friendly methods for plant pathogen monitoring. Annu. Rev. Phytopathol. 50:197–218.

De La Fuente, L., Navas-Cortés, J. A., and Landa, B. B. 2024. Ten challenges to understand and manage the insect-transmitted, xylem-limited bacterial pathogen *Xylella fastidiosa*. Phytopathology 114:869-884.

Doležel, J., Bartoš, J., Voglmayr, H., and Greilhuber, J. 2003. Nuclear DNA Content and Genome Size of Trout and Human. Cytometry Part A 51A:127–128.

Donoso, A., Valenzuela, S. 2018. In-field molecular diagnosis of plant pathogens: recent trends and future perspectives. Plant Pathol. 67:1451–1461.

Dreo, T., Pirc, M., Ramšak, Ž., Pavšič, J., Milavec, M., Žel, J., and Gruden, K. 2014. Optimising droplet digital PCR analysis approaches for detection and quantification of bacteria: a case study of fire blight and potato brown rot. Anal. Bioanal. Chem. 406:6513–6528.

Dupas, E., Briand, M., Jacques, M.-A., and Cesbron, S. 2019a. Novel tetraplex quantitative PCR assays for simultaneous detection and identification of *Xylella fastidiosa* subspecies in plant tissues. Front. Plant Sci. 10:1732.

Dupas, E., Legendre, B., Olivier, V., Poliakoff, F., Manceau, C., and Cunty, A. 2019b. Comparison of real-time PCR and droplet digital PCR for the detection of *Xylella fastidiosa* in plants. J. Microbiol. Meth. 162:86–95.

EFSA, E. F. S., Gibin, D., Pasinato, L., and Delbianco, A. 2023. Update of the *Xylella* spp. host plant database – systematic literature search up to 31 December 2022. EFSA J. 21:e08061.

Elbeaino, T., Yaseen, T., Valentini, F., Moussa, I. E. B., Mazzoni, V., and D'onghia, A. M. 2014. Identification of three potential insect vectors of *Xylella fastidiosa* in Southern Italy. Phytopathol. Mediterr. 53:328–332.

European and Mediterranean Plant Protection Organization (EPPO). 2023. PM 7/24 (5) *Xylella fastidiosa*. EPPO Bull. 53:205–276.

European and Mediterranean Plant Protection Organization (EPPO). 2018. PM 7/76 (5) Use of EPPO Diagnostic Standards. EPPO Bull. 48, 373–377.

Francis, M., Lin, H., Cabrera-La Rosa, J., Doddapaneni, H., and Civerolo, E. 2006. Genome-based PCR primers for specific and sensitive detection and quantification of *Xylella fastidiosa*. Eur. J. Plant Pathol. 115:203-213.

Freen-van Heeren, J.J. Flow-FISH as a Tool for Studying Bacteria, Fungi and Viruses. 2021. BioTech (Basel). 10:21.

Generalitat Valenciana (GVA). 2024. Situación de *Xylella fastidiosa* en la Comunitat Valenciana. Servicio de Sanidad Vegetal, DG Agricultura, Ganadería y Pesca, Generalitat Valenciana. https://portalagrari.gva.es/es/agricultura/xylella-fastidiosa#h3_num_0 (accessed December 2, 2024).

Giampetruzzi, A., Saponari, M., Almeida, R.P.P., Essakhi, S., Boscia, D., Loconsole, G., Saldarelli, P. 2017. Complete genome sequence of the olive-infecting strain *Xylella fastidiosa* subsp. *pauca* De Donno. Genome Announc. 5:e00569-17.

Giampetruzzi, A., Velasco-Amo, M. P., Marco-Noales, E., Montes-Borrego, M., Román-Écija, M., Navarro, I., Monterde, A., Barbé, S., Almeida, R. P. P., Saldarelli, P., Saponari, M.,



Funded by the European Union

Montilon, V., Savino, V. N., Boscia, D., and Landa, B. B. 2019. Draft genome resources of two strains ("ESVL" and "IVIA5901") of *Xylella fastidiosa* associated with Almond Leaf Scorch Disease in Alicante, Spain. Phytopathology 109:219–221.

Giovani, B., Cellier, G., McMullen, M., Saponari, M., Stefani, E., and Petter, F. 2019. From transnational research collaboration to regional standards. Biotechnol. Agron. Soc. Environ. 23:30–35.

Greco, D., Aprile, A., De Bellis, L., Luvisi, A. 2021. Diseases caused by *Xylella fastidiosa* in *Prunus* Genus: an overview of the research on an increasingly widespread pathogen. Front. Plant Sci. 12:712452.

Harper, S. J., Ward, L. I., and Clover, G. R. G. 2010. Development of LAMP and real-time PCR methods for the rapid detection of *Xylella fastidiosa* for quarantine and field applications. Phytopathology 100:1282–1288.

Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. 1993. Kinetic PCR analysis: Real-time monitoring of DNA amplification reactions. Nat. Biotechnol. 11:1026–1030.

Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., Legler, T.C., et al. 2011. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal. Chem. 83:8604–8610.

Hodgetts, J. 2022. A panel of Real-Time PCR assays for the direct detection of all of the *Xylella fastidiosa* subspecies. In plant pathology: method and protocols, methods in molecular biology. New York, NY, Springer US, 201–230.

Huggett, J.F., Whale, A. 2013. Digital PCR as a novel technology and its potential implications for molecular diagnostics. Clin. Chem. 59:1691–1693.

Jacques, M.-A., Denancé, N., Legendre, B., Morel, E., Briand, M., Mississipi, S., Durand, K., Olivier, V., Portier, P., Poliakoff, F., and Crouzillat, D. 2016. New coffee plant-infecting *Xylella fastidiosa* variants derived via homologous recombination. App. Environ. Microbiol. 82:1556–1568.

Kersting, S., Rausch, V., Bier, F. F., and von Nickisch-Rosenegk, M. 2014. Rapid detection of *Plasmodium falciparum* with isothermal recombinase polymerase amplification and lateral flow analysis. Malar J. 13:99.

Landis, J. R., and Koch, G. G. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159.

Lenth, R. V., Banfai, B., Bolker, B., Buerkner, P., Giné-Vazquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., Singmann, H. 2024. Emmeans: estimated marginal means, aka least-squares means. R package version 1.10.5. Available at https://CRAN.R-project.org/package=emmeans.

Li, R., Russell, P., Mcowen, N., Davenport, B., and Zhang, S. 2016. Development of a rapid and reliable isothermal AmplifyRP diagnostic assay for specific detection of *Xylella fastidiosa*. Phytopathology 106:109–109.

Li, W., Teixeira, D. C., Hartung, J. S., Huang, Q., Duan, Y., Zhou, L., Chen, J., Lin, H., Lopes, S., Ayres, A. J., and Levy, L. 2013. Development and systematic validation of qPCR assays for



Funded by the European Union

rapid and reliable differentiation of *Xylella fastidiosa* strains causing citrus variegated chlorosis. J. Microbiol. Meth. 92:79–89.

Lievens, A., Jacchia, S., Kagkli, D., Savini, C., and Querci, M. 2016. Measuring digital PCR quality: Performance parameters and their optimization. PLoS ONE 11:e0153317.

Linzer, D. A., and Lewis, J. B. 2011. poLCA: An R package for polytomous variable latent class analysis. J. Stat. Softw. 42:1–29.

Loconsole, G., Zicca, S., Manco, L., El Hatib, O., Altamura, G., Potere, O., Elicio, V., Valentini, F., Boscia, D., and Saponari, M. 2021. Diagnostic procedures to detect *Xylella fastidiosa* in nursery stocks and consignments of plants for planting. Agriculture 11:922.

Marco-Noales, E., Barbé, S., Monterde, A., Navarro-Herrero, I., Ferrer, A., Dalmau, V., Aure, C. M., Domingo-Calap, M. L., Landa, B. B., and Roselló, M. 2021. Evidence that *Xylella fastidiosa* is the causal agent of Almond Leaf Scorch disease in Alicante, mainland Spain (Iberian Peninsula). Plant Dis. 105:3349–3352.

McNemar, Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12:153–157.

Meyer, D., Zeileis, A., Hornik, K., Friendly, M. 2024. vcd: Visualizing Categorical Data. R package version 1.4-13. Available at https://CRAN.R-project.org/package=vcd.

Minsavage, G. V., Thompson, C. M., Hopkins, D. L., Leite, R. M. V. B. C., and Stall, R. E. 1994. Development of a Polymerase Chain Reaction protocol for detection of *Xylella fastidiosa* in Plant Tissue. Phytopathology 84:456–461.

Montero-Astúa, M., Chacón-Díaz, C., Aguilar, E., Rodríguez, C.M., Garita, L., Villalobos, W., Moreira, L., Hartung, J.S., Rivera, C. 2008. Isolation and molecular characterization of *Xylella fastidiosa* from coffee plants in Costa Rica. J. Microbiol. 46:482-90.

Moralejo, E., Gomila, M., Montesinos, M., Borràs, D., Pascual, A., Nieto, A., Adrover, F., Gost, P. A., Seguí, G., Busquets, A., Jurado-Rivera, J. A., Quetglas, B., García, J. de D., Beidas, O., Juan, A., Velasco-Amo, M. P., Landa, B. B., and Olmo, D. 2020. Phylogenetic inference enables reconstruction of a long-overlooked outbreak of almond leaf scorch disease (*Xylella fastidiosa*) in Europe. Commun. Biol. 3:1–13.

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., and Hase, T. 2000. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 28:E63.

Olmo, D., Nieto, A., Borràs, D., Montesinos, M., Adrover, F., Pascual, A., Gost, P. A., Quetglas, B., Urbano, A., García, J. de D., Velasco-Amo, M. P., Olivares-García, C., Beidas, O., Juan, A., Marco-Noales, E., Gomila, M., Rita, J., Moralejo, E., and Landa, B. B. 2021. Landscape epidemiology of *Xylella fastidiosa* in the Balearic Islands. Agronomy 11:473.

Olmos, A., Bertolini, E., Ruiz-García, A. B., Martínez, C., Peiró, R., and Vidal, E. 2016. Modeling the accuracy of three detection methods of grapevine leafroll-associated virus 3 during the dormant period using a Bayesian approach. Phytopathology 106:510–518.

Ouyang, P., Arif, M., Fletcher, J., Melcher, U., and Ochoa Corona, F. M. 2013. Enhanced reliability and accuracy for field deployable bioforensic detection and discrimination of *Xylella fastidiosa* subsp. pauca, causal agent of citrus variegated chlorosis using razor ex technology and TaqMan quantitative PCR. PLoS ONE 8:e81647.



Funded by the European Union

Papadakis, G., Pantazis, A.K., Fikas, N., Chatziioannidou, S., Tsiakalou, V., Michaelidou, K., Pogka, V., Megariti, M., Vardaki, M., Giarentis, K., et al. 2022. Portable real-time colorimetric LAMP-device for rapid quantitative detection of nucleic acids in crude samples. Sci. Rep. 12:3775.

Pinheiro, L.B., Coleman, V.A., Hindson, C.M., Herrmann, J., Hindson, B.J., Bhat, S., Emslie, K.R. 2012. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. Anal. Chem. 84:1003–1011.

Pucci, N., Scala, V., Cesari, E., Crosara, V., Fiorani, R., L'Aurora, A., Lucchesi, S., Tatulli, G., Barra, E., Ciarroni, S., De Amicis, F., Fascella, S., Giacobbi, F., Gaffuri, F., Gualandri, V., Landi, L., Loconsole, G., Molinatto, G., Pollastro, S., Raimondo, M. L., Rizzo, D., Roggia, C., Silletti, M. R., Talevi, S., Testa, M., and Loreti, S. 2023. An inter-laboratory comparative study on the influence of reagents to perform the identification of the *Xylella fastidiosa* subspecies using tetraplex Real Time PCR. Horticulturae 9:1053.

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.r-project.org

Raguseo, C., Gerin, D., Pollastro, S., Rotolo, C., Rotondo, P.R., Faretra, F., Angelini, R.M.D.M. 2021. A duplex-droplet digital PCR assay for simultaneous quantitative detection of *Monilinia fructicola* and *Monilinia laxa* on stone fruits. Front. Microbiol. 12:747560.

Rodrigues, J. L. M., Silva-Stenico, M. E., Gomes, J. E., Lopes, J. R. S., and Tsai, S. M. 2003. Detection and diversity assessment of *Xylella fastidiosa* in field-collected plant and insect samples by using 16S rRNA and gyrB Sequences. Appl. Environ. Microbiol. 69:4249–4255.

Rodrigues, J.L., Silva-Stenico, M.E., de Souza, A.N., Lopes, J.R., Tsai, S.M. 2006. *In situ* probing of *Xylella fastidiosa* in honeydew of a xylem sap-feeding insect using 16S rRNA-targeted fluorescent oligonucleotides. Environ. Microbiol. 8:747-54.

Rodriguez, J., Hausbeck, M., and Sakalidis, M. 2021. Application of target enrichment sequencing for population genetic analyses of the obligate plant pathogens *Pseudoperonospora cubensis* and *P. humuli*. MPMI 34:1103–1118

Roy, A. S., Petter, F., Griessinger, D. 2010.- EPPO database on diagnostic expertise. Available at http://dc.eppo.org. Bull. OEPP/EPPO Bull. 40:127-130

Sanchez B., Barreiro J., Soto I., Rodriguez E., 2019. The impact indicator for priority pests (I2P2): a tool for ranking pests according to regulation (EU) 2016/2031. Maastricht: European Union.

Saponari, and Loconsole, G. 2020. Interlaboratory comparison EU-*XF*-IC-2020-03: Evaluation of molecular methods for the detection of *Xylella fastidiosa*. https://www.*Xf*actorsproject.eu/research-results-xylella-fastidiosa/.

Saponari, M., Boscia, D., Nigro, F., and Martelli, G. P. 2013. Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J. Plant Pathol. 95.

Schaad, N.W., Opgenorth, D., Gaush, P. 2002. Real-time polymerase chain reaction for one-hour on-site diagnosis of Pierce's Disease of grape in early season asymptomatic vines. Phytopathology 92:721–728.



Funded by the European Union

Stevenson, M., Sergeant, E., Heuer, C., Nunes, T., Marshall, J., Sanchez, J., Thornton, R., Reiczigel, J., Robison-Cox, J., Sebastiani, P., Solymos, P., Yoshida, K., Jones, G., Pirikahu, S., Firestone, S., Kyle, R., Popp, G., Jay, M., Cheung, A., Singanallur, N., Szabo, A., and Rabiee, A. 2024. epiR: Tools for the Analysis of Epidemiological Data. R package version 2.0.77. Available at https://cran.r-project.org/web/packages/epiR/index.html.

Sykes, P.J., Neoh, S.H., Brisco, M.J., Hughes, E., Condon, J., Morley, A.A. 1992. Quantitation of targets for PCR by use of limiting dilution. BioTechniques 13:444–449.

Turechek, W. W., and Wang, H. 2023. Evaluation of a Viable-Cell Detection Assay for *Xanthomonas fragariae* with Latent Class Analysis. PhytoFrontiers 3:214–224.

Turechek, W. W., Webster, C. G., Duan, J., Roberts, P. D., Kousik, C. S., and Adkins, S. 2013. The use of latent class analysis to estimate the sensitivities and specificities of diagnostic tests for Squash vein yellowing virus in cucurbit species when there is no gold standard. Phytopathology 103:1243–1251.

Vogelstein, B., Kinzler, K.W. 1999. Digital PCR. Proc. Natl. Acad. Sci. USA 96:9236–9241.

Vučurović, A., Mehle, N., Anthoine, G., Dreo, T., and Ravnikar, M., eds. 2022. Critical Points for the Organisation of Test Performance Studies in Microbiology: Plant Pathogens as a Case Study. Springer, Cham, Switzerland.

Waliullah, S., Hudson, O., Oliver, J. E., Brannen, P. M., Ji, P., and Ali, M. E. 2019. Comparative analysis of different molecular and serological methods for detection of *Xylella fastidiosa* in blueberry. PLoS ONE 14:e0221903.

Whale, A.S., De Spiegelaere, W., Trypsteen, W., Nour, A.A., Bae, Y.-K., Benes, V., Burke, D., Cleveland, M., Corbisier, P., Devonshire, A.S., et al. 2020. The digital MIQE guidelines update: minimum information for publication of quantitative digital PCR experiments for 2020. Clin. Chem. 66:1012–1029.

Youden, W. J. 1950. Index for rating diagnostic tests. Cancer 3:32–35.

Yuan, X., Morano, L., Bromley, R., Spring-Pearson, S., Stouthamer, R., and Nunney, L. 2010. Multilocus Sequence Typing of *Xylella fastidiosa* causing Pierce's Disease and oleander leaf scorch in the United States. Phytopathology 100:601–611.

Zarco-Tejada, P. J., Camino, C., Beck, P. S. A., Calderon, R., Hornero, A., Hernández-Clemente, R., Kattenborn, T., Montes-Borrego, M., Susca, L., Morelli, M., Gonzalez-Dugo, V., North, P. R. J., Landa, B. B., Boscia, D., Saponari, M., and Navas-Cortes, J. A. 2018. Previsual symptoms of *Xylella fastidiosa* infection revealed in spectral plant-trait alterations. Nat. Plants 4:432–439.

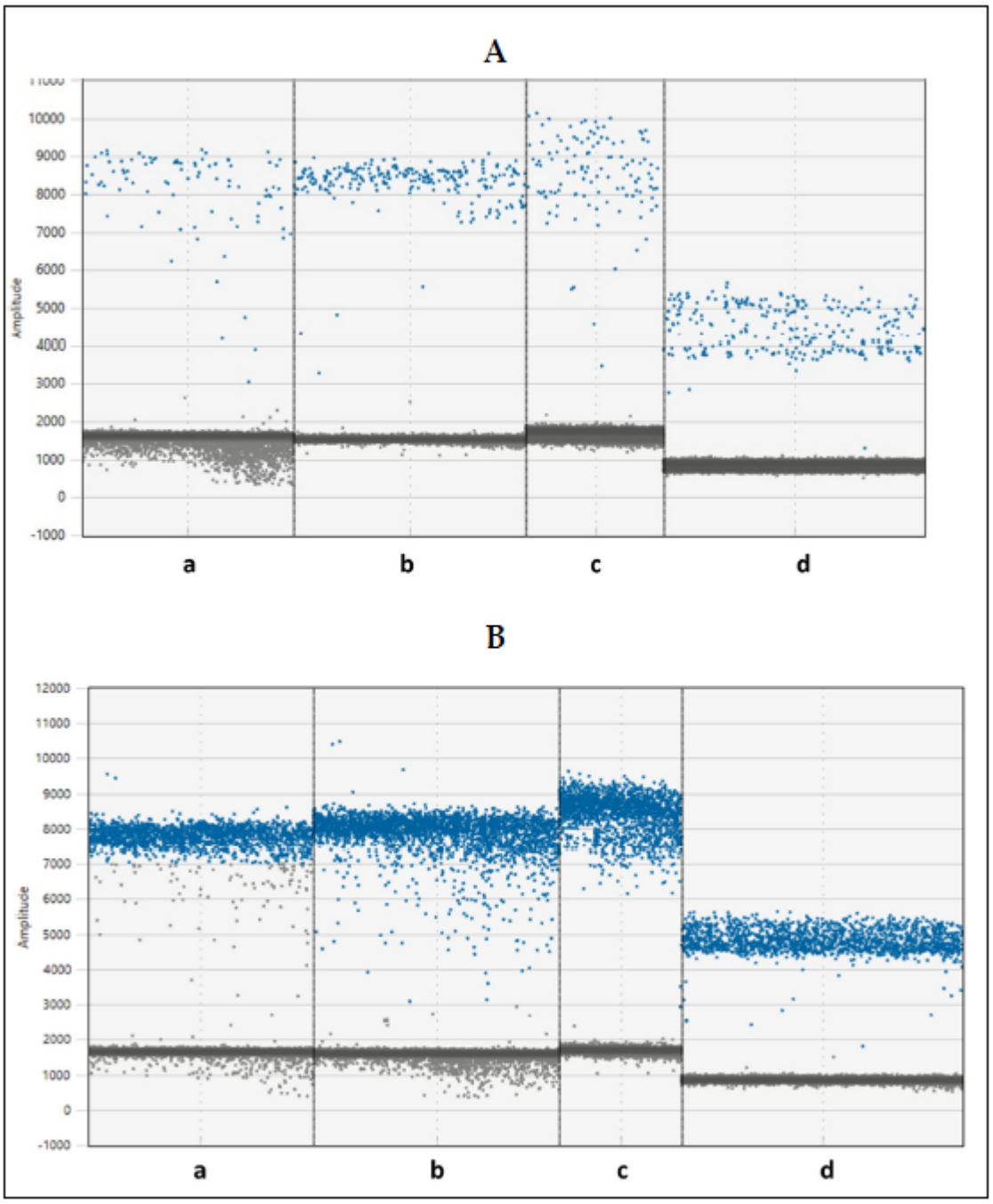
Zarco-Tejada, P. J., Poblete, T., Camino, C., Gonzalez-Dugo, V., Calderon, R., Hornero, A., Hernandez-Clemente, R., Román-Écija, M., Velasco-Amo, M. P., Landa, B. B., Beck, P. S. A., Saponari, M., Boscia, D., and Navas-Cortes, J. A. 2021. Divergent abiotic spectral pathways unravel pathogen stress signals across species. Nat. Commun. 12:6088.

Zecharia, N., Miri, V., Dror, O., Hatib, K., Holland, D., Dani, S., et al. 2024. Seasonal dynamics and distribution of *Xylella fastidiosa* in infected almond trees. Phytopathology 114:1186–1195.



Funded by the European Union

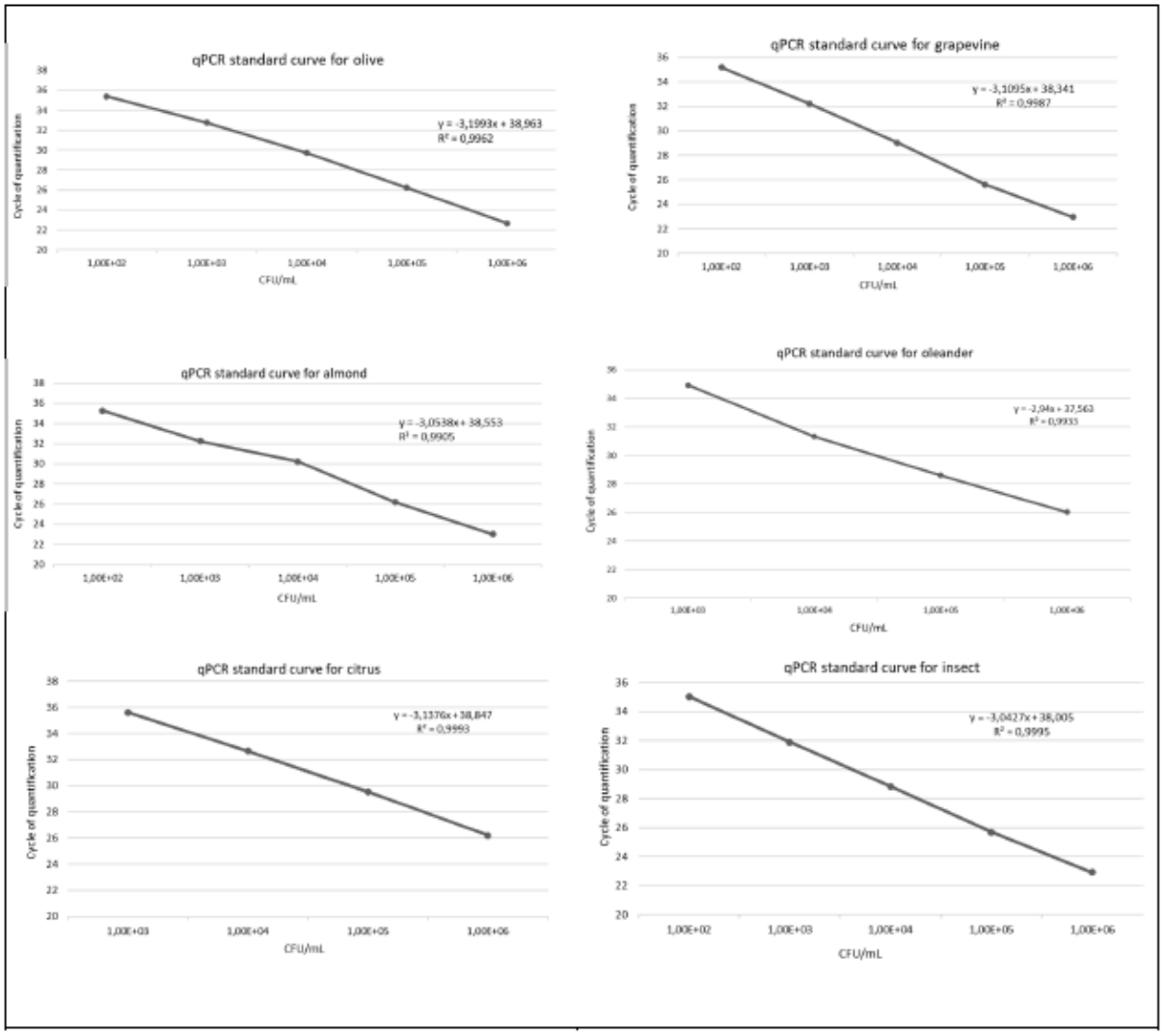
# ANNEX 1. SUPPLEMENTARY FIGURES




**Supplementary Figure S1**. Evaluation of the optimal amount of olive (A) and insect extracts (B) using the ddPCR conditions reported by Dupas et al., 2019. On x-axis: Amplitude value. On y-axis: DNA amounts tested; a: 2 µL; b : 4 µL;c: 6 µL; d: 8 µL. The best positive and negative droplet separation was achieved employing 4 µL and 6 µL of purified DNA for plant and insect, respectively.



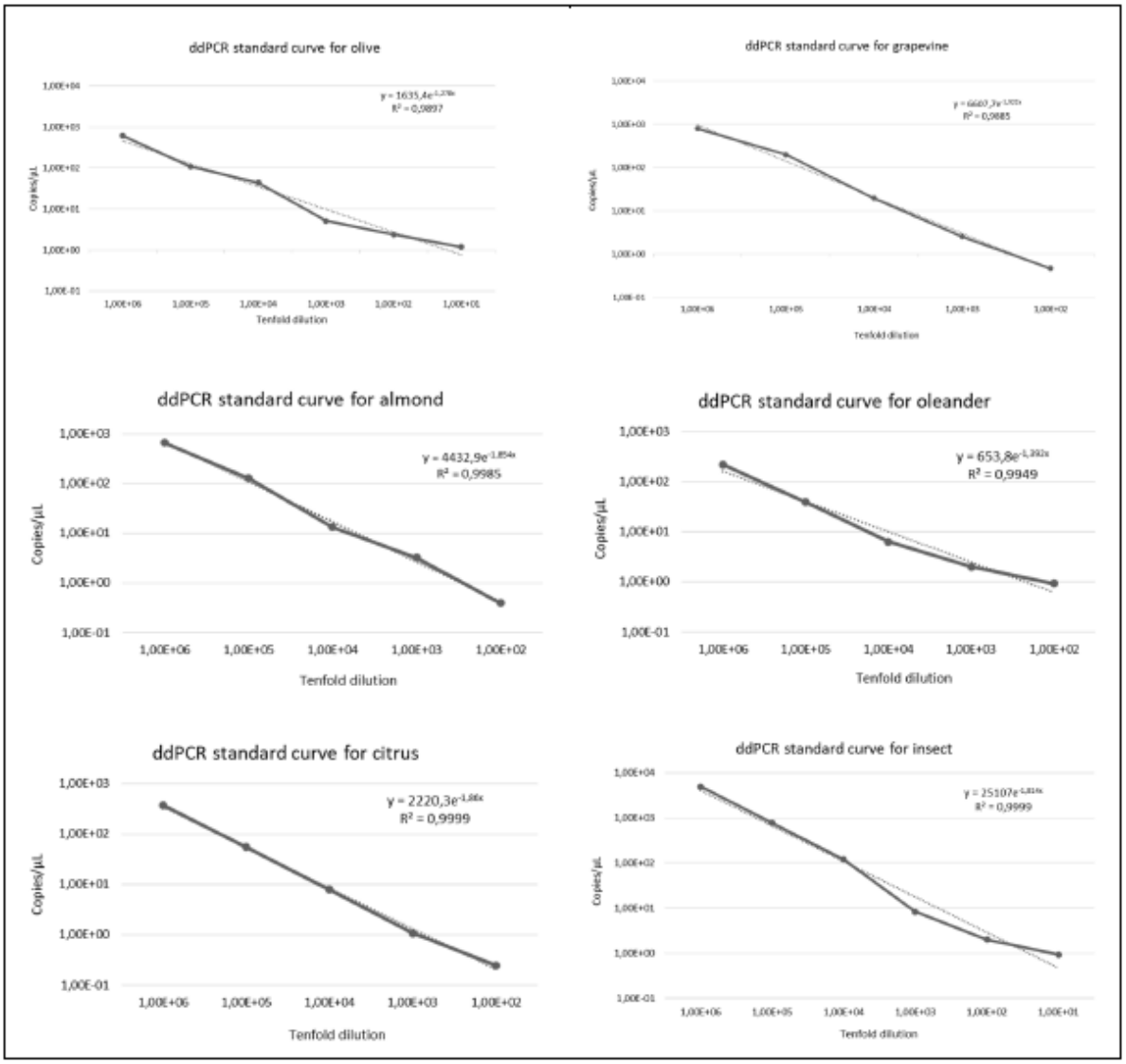
Funded by the European Union




**Supplementary Figure S2**. Standard curves of qPCR assays on purified DNA from ten-fold serial dilutions of bacterial cell suspension (from 106 to the limit of detection, LOD) spiked in plant and insect matrices. Slopes and determination coefficients (R2) are indicated in the figures.



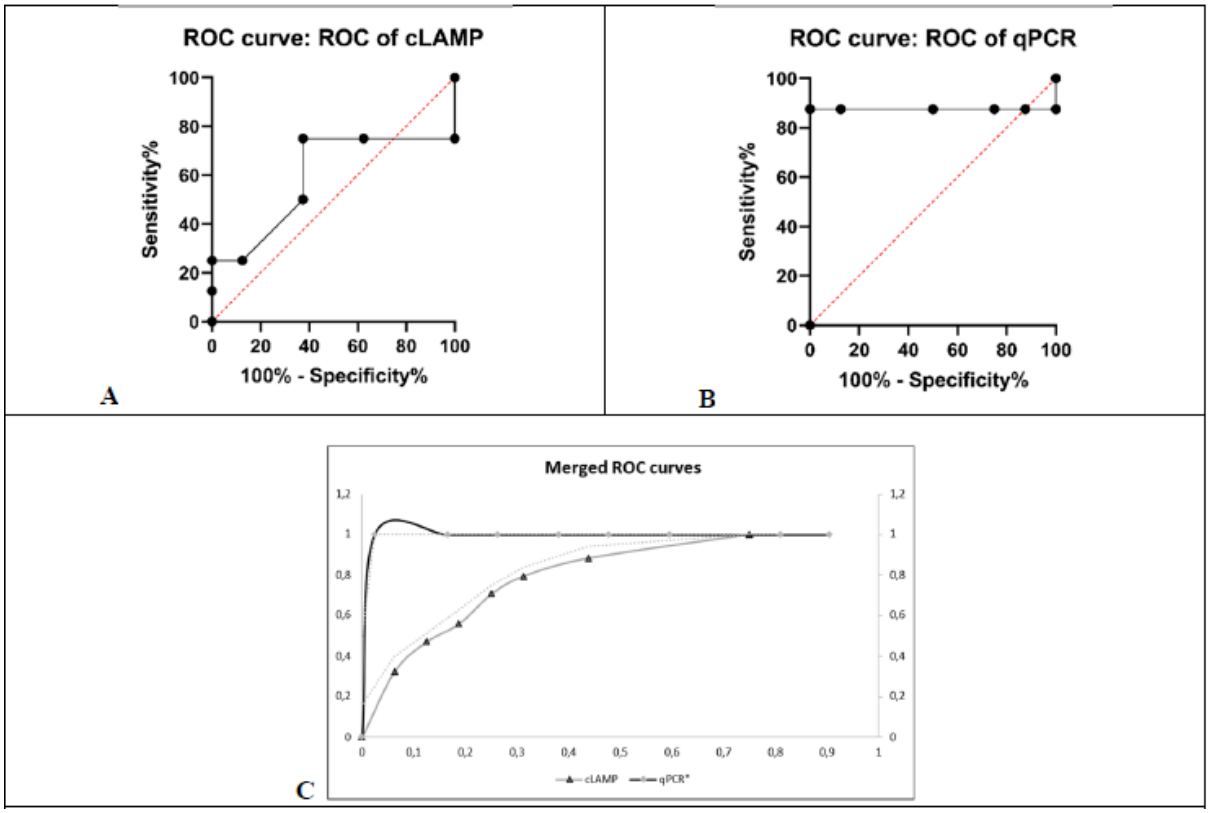
Funded by the European Union




**Supplementary Figure S3**. Linear regression curves of the ddPCR assay, for the plant and insect matrices evaluated, were constructed employing the same ten-fold dilution series tested with the qPCR assay. The vertical axis shows the measured copies/µL of the ddPCR reaction mixture.



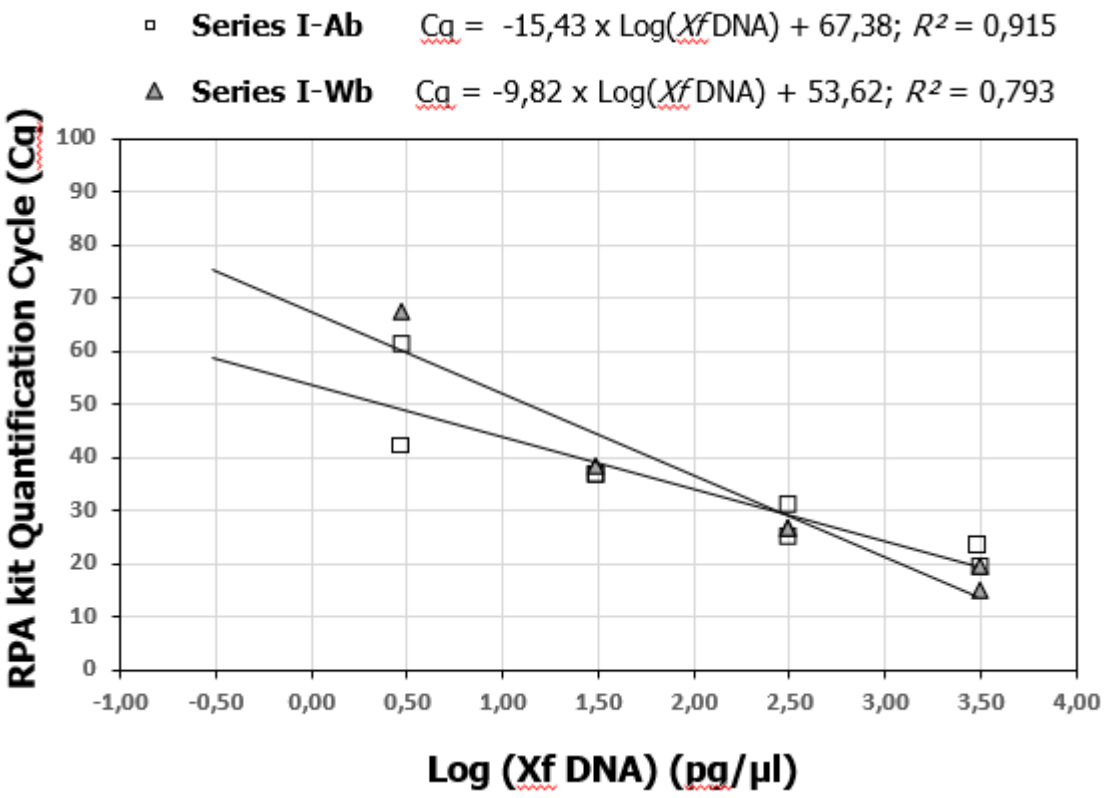
Funded by the European Union




**Supplementary Figure 5**. (A) Receiver Operating Characteristic (ROC) curve for cLAMP assay with an AUC of 0.5938 and 95% confidence interval 0.2933 to 0.8942 (p value = 0.5286; SE = 0.1533). (B) Receiver Operating Characteristic (ROC) curve for qPCR with AUC of 0.8750 and 95% confidence interval (CI) 0.6458 to 1.00 (p value = 0.0117; SE = 0.1169). In both cases, sensitivity (y axis) vs specificity (x axis) values, expressed in %, for each assay are plotted. (C) Overlapped ROC curves analysis performed to better evaluate the performance of the qPCR (black curve) and cLAMP (grey curve) assays for the detection of *Xf* in naturally infected almond trees; Moving average trendlines (light grey) of both curves are shown for a more clearly pattern.



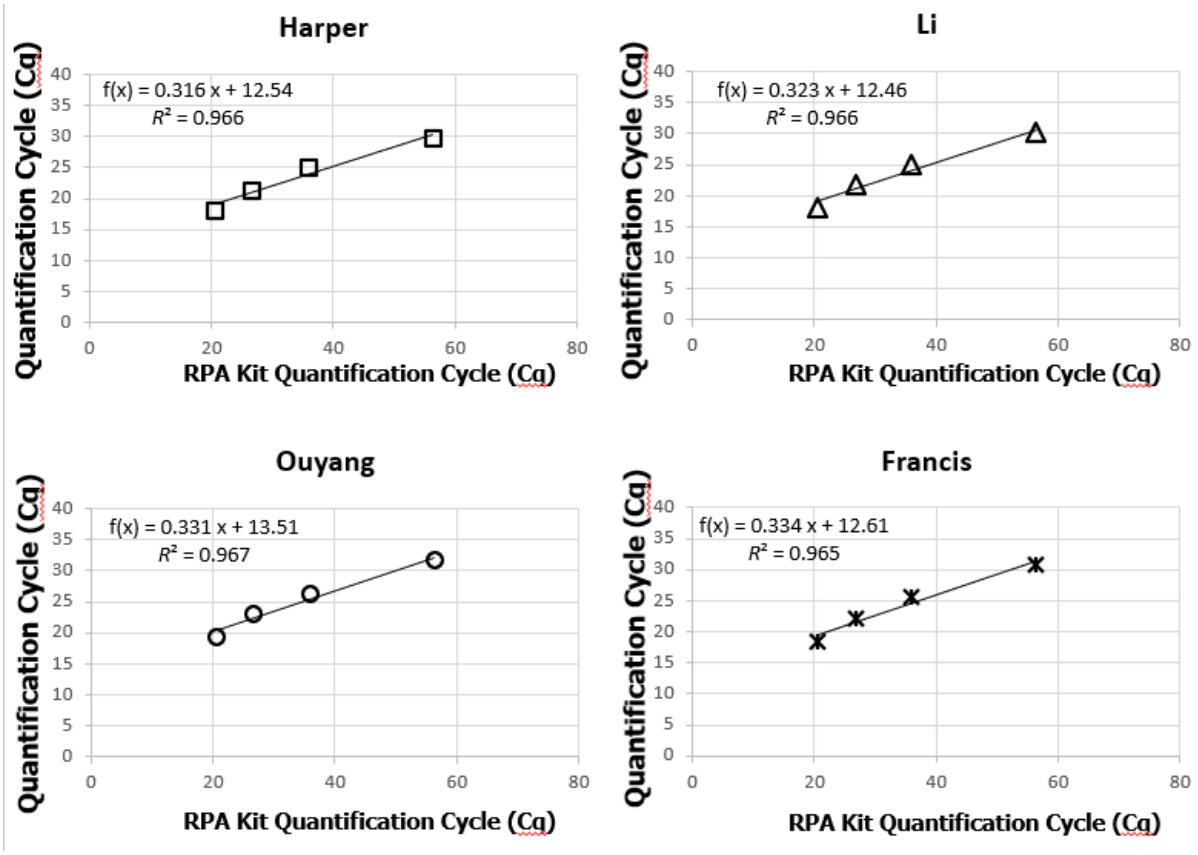
Funded by the European Union




**Supplementary Figure S6**. Quantification cycle (Cq) values obtained for the RPA kit using a seven-point 10-fold serial dilution of *Xylella fastidiosa subsp. multiplex* strain ESVL DNA standard curve diluted in almond DNA background (15 ng/ µl) (DNA Series I-Ab) or ultrapure sterile water (DNA Series I-Wb). Values from the linear regression analysis for each treatment are shown in the figure.



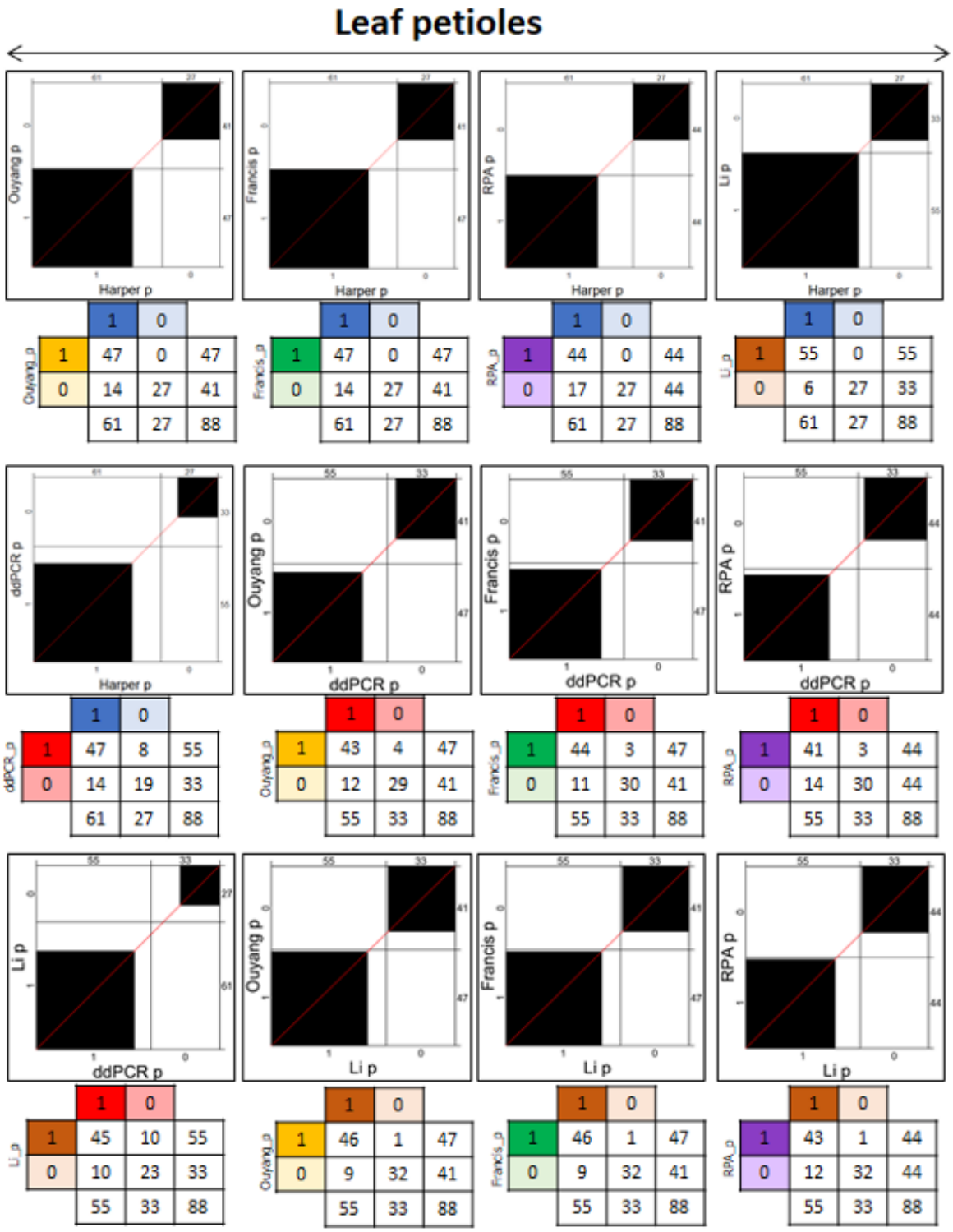
Funded by the European Union




**Supplementary Figure S7.** Comparison between the quantification cycle (Cq) values obtained with the RPA kit and those from Harper-, Li-, Ouyang-, and Francis-qPCR assays using serial dilutions of genomic DNA from *Xylella fastidiosa subsp. multiplex* strain ESVL. Linear regression analysis was performed on the RPA cycle threshold values (x-axis) against the qPCR cycle threshold values (y-axis).



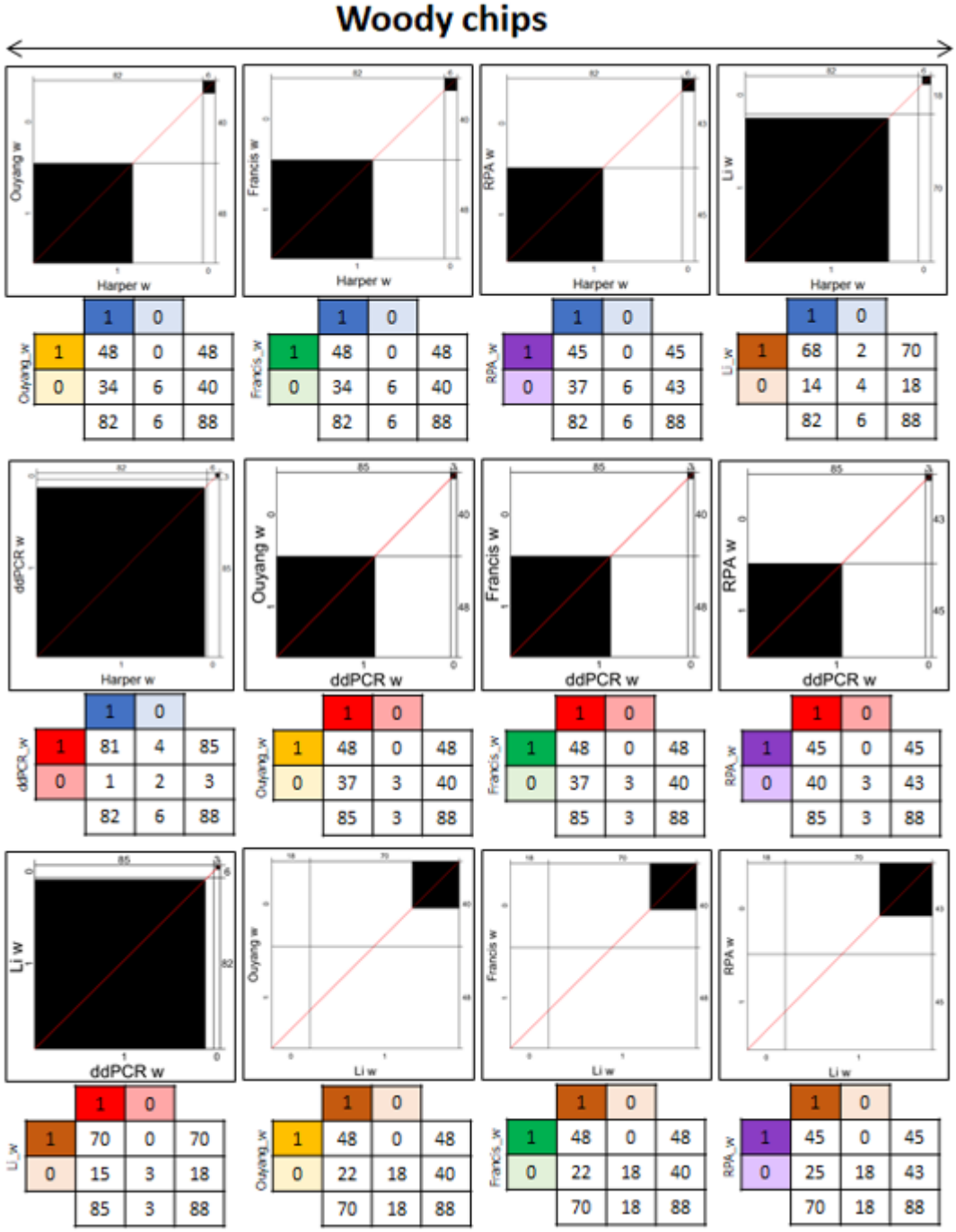
Funded by
the European Union




**Supplementary Figure S8**. Agreement plots for each pair of six *Xylella fastidiosa* molecular diagnostic tests obtained when using DNA extracted from leaf petioles. The bottom-left and top-right black rectangles represent the proportion of cases in which both diagnostic tests had the same result (negative and positive respectively), whereas the top- left and bottom-right white squares are representative of the cases in which the outcome of the tests was different. Individual numbers shown on the x- and y-axis indicate the number of positive and negative test results.



Funded by the European Union




**Supplementary Figure S9**. Agreement plots for each pair of six *Xylella fastidiosa* molecular diagnostic tests obtained when using DNA extracted from woody chips. The bottom-left and top-right black rectangles represent the proportion of cases in which both diagnostic tests had the same result (negative and positive respectively), whereas the top- left and bottom-right white squares are representative of the cases in which the outcome of the tests was different. Individual numbers shown on the x- and y-axis indicate the number of positive and negative test results.



Funded by the European Union

# ANNEX 2. SUPPLEMENTARY TABLES

**Supplementary Table 1**. List of primers used for qPCR and cLAMP assays (from Harper et al., 2010). The extra, non-sequence-specific T stretches, in the middle of the internal LAMP primers are indicated in blue

| | |
|---|---|
| **qPCR** | |
| XF-F | CACGGCTGGTAACGGAAGA |
| XF-R | GGGTTGCGTGGTGAAATCAAG |
| XF-P | TCGCATCCCGTGGCTCAGTCC |
| **cLAMP** | |
| Xf3 | CCGTTGGAAAACAGATGGGA |
| Xb3 | GAGACTGGCAAGCGTTTGA |
| Xfip | ACCCCGACGAGTATTACTGGG*TTTT*TCGCTACCGAGAACCACAC |
| Xbib | GCGCTGCGTGGCACATAGA*TTTT*TGCAACCTTTCCTGGCATCAA |
| XLf | TGCAAGTACACACCCTTGAAG |
| XLb | TTCCGTACCACAGATCGCT |



Funded by the European Union

**Supplementary Table 9.** Disease incidence and severity score of Almond Leaf Scorch (ALS), and diagnostic results of the six molecular tests obtained when using DNA samples extracted from leaf petioles or woody chips for a total of 88 Prunus dulcis trees analyzed in the study

| ID Tree | Incidence-Visual assessment | Severity ALS | | Ouyang - Woody chips | Li-Woody chips | Francis-Woody chips | Harper-Woody chips | RPA-Woody chips | ddPCR-Woody chips | | Ouyang - Leaf petioles | Li-Leaf petioles | Francis-Leaf petioles | Harper-Leaf petioles | RPA-Leaf petioles | ddPCR-Leaf petioles |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AXF_262 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_263 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_264 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_265 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_266 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_267 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_268 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_269 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_270 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_271 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 1 | 0 | 1 |
| AXF_272 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_273 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_274 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_275 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_276 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_277 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_278 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_279 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_280 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 0 | | 1 | 1 | 0 | 1 | 0 | 1 |
| AXF_281 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 1 | 1 | 1 | 1 | 0 | 1 |
| AXF_282 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_283 | 1 | 0,5 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 1 | 0 | 1 | 0 | 0 |
| AXF_284 | 1 | 0,5 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 1 |
| AXF_285 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_286 | 1 | 3 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_287 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 0 | 1 |
| AXF_288 | 1 | 1 | | 0 | 1 | 0 | 0 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_289 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_290 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 1 | 0 | 1 | 0 | 0 |
| AXF_291 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_292 | 1 | 1,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_293 | 1 | 1,5 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 1 | 0 | 1 |
| AXF_294 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 1 | 0 | 1 | 0 | 0 |
| AXF_295 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |



Funded by the European Union

| ID Tree | Incidence-Visual assessment | Severity ALS | | Ouyang - Woody chips | Li-Woody chips | Francis-Woody chips | Harper-Woody chips | RPA-Woody chips | ddPCR-Woody chips | | Ouyang - Leaf petioles | Li-Leaf petioles | Francis-Leaf petioles | Harper-Leaf petioles | RPA-Leaf petioles | ddPCR-Leaf petioles |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AXF_296 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_297 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_298 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 0 | 1 |
| AXF_299 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 0 | 1 | 0 | 1 | 1 | 0 |
| AXF_300 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_301 | 1 | 0,2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 0 |
| AXF_302 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_303 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 0 | 1 | 1 | 1 | 1 | 1 |
| AXF_304 | 1 | 0,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_305 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_306 | 0 | 0 | | 1 | 1 | 1 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_307 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_308 | 0 | 0 | | 1 | 1 | 1 | 1 | 0 | 1 | | 0 | 0 | 0 | 1 | 0 | 0 |
| AXF_309 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_310 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_311 | 1 | 2,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_312 | 1 | 3 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_313 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 1 | | 0 | 1 | 1 | 1 | 0 | 0 |
| AXF_314 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 1 | 1 | 0 | 1 | 0 | 0 |
| AXF_315 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_316 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 0 |
| AXF_317 | 0 | 0 | | 0 | 1 | 0 | 0 | 0 | 1 | | 0 | 1 | 0 | 1 | 0 | 0 |
| AXF_318 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_319 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_320 | 1 | 1,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_321 | 1 | 1,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_322 | 1 | 3 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_323 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_324 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 1 | 0 | 0 |
| AXF_325 | 1 | 4 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| AXF_326 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_327 | 1 | 1,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_328 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |



Funded by the European Union

| ID Tree | Incidence-Visual assessment | Severity ALS | | Ouyang - Woody chips | Li-Woody chips | Francis-Woody chips | Harper-Woody chips | RPA-Woody chips | ddPCR-Woody chips | | Ouyang - Leaf petioles | Li-Leaf petioles | Francis-Leaf petioles | Harper-Leaf petioles | RPA-Leaf petioles | ddPCR-Leaf petioles |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AXF_329 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_330 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 0 | 1 | 1 | 1 | 1 | 1 |
| AXF_331 | 1 | 2,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_332 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_333 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_334 | 1 | 3 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_335 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_336 | 1 | 2,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_337 | 0 | 0 | | 1 | 1 | 1 | 1 | 0 | 1 | | 1 | 1 | 0 | 1 | 0 | 0 |
| AXF_338 | 1 | 0,2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 0 | 1 | 1 | 1 | 0 |
| AXF_339 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_340 | 0 | 0 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 0 | 0 | 1 | 0 | 0 |
| AXF_341 | 1 | 1,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_342 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_343 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_344 | 1 | 3 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 0 |
| AXF_345 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_346 | 1 | 2,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_347 | 1 | 3,5 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_348 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| AXF_349 | 1 | 2 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| Number | 41 | | | 48 | 70 | 48 | 82 | 45 | 85 | | 47 | 55 | 47 | 61 | 44 | 54 |
| Percentage (%) | 46,6 | | | 54,5 | 79,5 | 54,5 | 93,2 | 51,1 | 96,6 | | 53,4 | 62,5 | 53,4 | 69,3 | 50,0 | 61,4 |



Funded by the European Union

**Supplementary Table 10.** Comparison of diagnostic test performance agreement between six molecular diagnostic tests for Xylella fastidiosa when using two DNA sources and three tests selected as the gold standard.

| Gold standard | DNA source | Diagnostic test | Proportion of agreement observed | Proportion of agreement expected | PABAK(1) | McNemar's $\chi^2$ (2) |
|---|---|---|---|---|---|---|
| **Li** | woody chips | Ouyang | 0,750 | 0,527 | 0.5 (0.293, 0.672) | 22 *P* = 2.727e-06 |
| | | Francis | 0,750 | 0,527 | 0.5 (0.293, 0.672) | 22 *P* = 2.727e-06 |
| | | RPA | 0,716 | 0,507 | 0.432 (0.22, 0.614) | 25 *P* = 5.733e-07 |
| | leaf petioles | Ouyang | 0,8863636 | 0,509 | 0.773 (0.602, 0.888) | 6.4 *P* = 0.01141 |
| | | Francis | 0,8863636 | 0,509 | 0.773 (0.602, 0.888) | 6.4 *P* = 0.01141 |
| | | RPA | 0,8522727 | 0,500 | 0.614 (0.418, 0.766) | 9.3077 *P* = 0.002282 |

(1) The prevalence and bias-adjusted kappa (PABAK) index (Byrt et al. 1993) was used to account for disease prevalence and correct for any potential bias in the tests, and was estimated as PABAK= 2Po - 1; where Po = Observed Agreement (proportion of cases where both rates agree).
(2) McNemar's Test was used to determine the existence of significant differences in proportions between two tests that are dependent (i.e., the same DNA sample was used for each pair of tests).



Funded by the European Union

**Supplementary Table 11.** List of target sequences used to design the baits for the process of targeted next-generation sequencing enrichment (*Xf*-TSCE)

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **MLST** | *holC* | DNA polymerase III holoenzyme chis subunit | NC_004556 | 133973 | 134394 |
| **MLST** | *nuoL* | NADH-ubiquinone oxidoreductase NQO12 subunit | NC_004556 | 328700 | 329298 |
| **MLST** | *gltT* | Glutamate symport protein | NC_004556 | 1763125 | 1763822 |
| **MLST** | *cysG* | cysG Siroheme synthase | NC_004556 | 2156444 | 2157085 |
| **MLST** | *petC* | Ubiquinol cytochrome c oxidoreductase | NC_004556 | 2066205 | 2066780 |
| **MLST** | *leuA* | 2-Isopropylmalate synthase | NC_004556 | 1250098 | 1250846 |
| **MLST** | *malF* | ABC transporter sugar permease | NC_004556 | 1707484 | 1708256 |
| **MLST** | *pilU* | Twitching motility protein | NC_004556 | 1345718 | 1346306 |
| **MLSA-E** | *acvB* | Virulence protein: suggested to regulate pathogenicity | NC_004556 | 2231493 | 2232235 |
| **MLSA-E** | *copB* | Copper resistance protein B precursor | NC_004556 | 129489 | 130155 |
| **MLSA-E** | *cvaCa* | Colicin V precursor | NC_004556 | 272372 | 272701 |
| **MLSA-E** | *fimA* | Fimbrial subunit precursor | NC_004556 | 82627 | 83183 |
| **MLSA-E** | *gaa* | Glutaryl-7-aminocephalosporanic acid acylase precursor | NC_004556 | 398233 | 399361 |
| **MLSA-E** | *pglAa* | Polygalacturonase precursor | NC_004556 | 1728368 | 1729172 |
| **MLSA-E** | *pilA* | Fimbrial protein | NC_004556 | 2256827 | 2257208 |
| **MLSA-E** | *rpfF* | Regulator of pathogenicity factors | NC_004556 | 499366 | 500167 |
| **MLSA-E** | *xadAa* | Outer membrane afimbrial adhesin | NC_004556 | 897803 | 898866 |
| ***Xf* Diagnosis** | *rpoD*-short | RNA polymerase sigma-70 factor, | NC_004556 | 726358 | 727078 |
| **Diagnosis** | *rpoD*-Long | RNA polymerase sigma-70 factor | NC_004556 | 725501 | 726488 |
| **Diagnosis** | 16S *rrs* | 16S rDNA | NC_004556 | 171190 | 171982 |



Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **Diagnosis** | *gyrB* | DNA gyrase-β-subunit polypeptide-coding gene16S rRNA processing RimM protein | NC_002488 | 5261 | 5689 |
| **Diagnosis** | *RimM* | 16S rRNA processing RimM proteinhypothetical protein HL gene | NC_004556 | 104919 | 104989 |
| **Genome** | | methyl transferase | NC_002488 | 1878143 | 1878780 |
| **Genome** | | fimbrial protein | NC_002488 | 2414734 | 2415145 |
| **Genome** | | Intergenic region of GCVA | NC_010513 | 188931 | 189450 |
| **Genome** | | DNA gyrase-β-subunit polypeptide-coding gene | | | |
| **Plasmid** | *VapD* | Present in plasmid p*XF*51, a 17.6 kDa virulence-associated protein D (*VapD*) predicted to belong to the CRISPR/Cas2 family (Pfam: PF09827). Refer.: https://doi.org/10.1371/journal.pone.0145765 | NC_002490 | 41316 | 41771 |
| ***Xf* Diagnosis** | ORF-mulberry | Hypothetical protein. PCR Amplicon specific of mulberry strains and in some subsp. *pauca* strains (ST53 and ST80) | NZ_CP020870 | 873202 | 873489 |
| ***Xf* Diagnosis*** | 272-1 int / 272-2 int | Specific detection *Xylella fastidiosa* subsp. *pauca* | NC_002488 | 1051239 | 1051710 |
| ***Xf* Diagnosis*** | CVC-1 / 272-2 int | General detection all *Xylella fastidiosa* strains | NC_002488 | 1051239 | 1051745 |
| **MLSA-E** | *fts*Z | Cell division protein; forms a Z-ring guiding septum formation. | NC_004556 | 2187566 | 2188085 |
| **MLSA-E** | *grpE* | Co-chaperone in protein folding; works with DnaK/DnaJ in heat-shock response. | NC_004556 | 1594853 | 1595113 |
| **MLSA-E** | *infC* | Translation initiation factor IF-3; ensures proper start codon selection. | NC_004556 | 2244925 | 2245389 |
| **MLSA-E** | *L10* | Ribosomal protein L10; part of 50S ribosomal subunit, crucial for protein synthesis. | NC_004556 | 2365033 | 2365449 |



Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **MLSA-E** | *recA* | DNA recombination and repair protein; maintains genomic integrity. | NC_004556 | 118352 | 118875 |
| **MLSA-E** | *rimP* | Ribosome maturation factor; aids in assembly of the 30S subunit. | NC_004556 | 242477 | 242737 |
| **MLSA-E** | *rpoA* | RNA polymerase alpha subunit; essential for transcription initiation and enzyme assembly. | NC_004556 | 549106 | 549361 |
| **MLSA-E** | *rpoA* | RNA polymerase alpha subunit; essential for transcription initiation and enzyme assembly. | NC_004556 | 549872 | 550104 |
| **MLSA-E** | *ruvA* | Holliday junction DNA helicase; important in DNA repair and recombination. | NC_004556 | 1097816 | 1098076 |
| **MLSA-E** | *secG* | Part of the Sec translocase complex; involved in protein export across the cytoplasmic membrane. | NC_004556 | 317051 | 317449 |
| **MLSA-E** | *truB* | tRNA pseudouridine synthase; modifies tRNA for stability and function. | NC_004556 | 247597 | 248097 |
| **MLSA-E** | exbD1 | Biopolymer transport exbD1 protein | NC_004556 | 12265 | 12684 |
| **MLSA-E** | *pilJ* | Chemotaxis protein (MCP) | NC_004556 | 1048576 | 1050618 |
| **MLSA-E** | *pyrE* | Orotate phosphoribosyl transferase | NC_004556 | 155324 | 155983 |
| **Pangenome** | *acpP* | Acyl carrier protein | NC_004556 | 1744038 | 1743799 |
| **Pangenome** | *mreB* | Rod shape-determining protein MreB | NC_004556 | 684701 | 685744 |
| **Pangenome** | *rpmB* | 50S ribosomal protein L28 | NC_004556 | 584054 | 584290 |
| **Pangenome** | *speD* | S-adenosylmethionine decarboxylase proenzyme | NC_004556 | 928617 | 927823 |
| **Pangenome** | *metE* | 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase | NC_004556 | 1534676 | 1536952 |
| **Pangenome** | *queH* | Epoxyqueuosine reductase QueH | NC_004556 | 1537019 | 1537729 |



Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **Pangenome** | *rne* | Ribonuclease E | NC_004556 | 2330801 | 2334199 |
| **Pangenome** | group_2814 | hypothetical protein | NC_004556 | 2206239 | 2204377 |
| **Pangenome** | *hupB* | DNA-binding protein HU-beta | NC_004556 | 569629 | 569913 |
| **Pangenome** | *clsA* | Cardiolipin synthase A | NC_004556 | 584655 | 586058 |
| **Pangenome** | *apaH* | Bis(5'-nucleosyl)-tetraphosphatase, symmetrical | NC_004556 | 1411376 | 1412263 |
| **Pangenome** | *sodB* | Superoxide dismutase [Mn/Fe] | NC_004556 | 2338762 | 2339373 |
| **Pangenome** | *rpsL* | 30S ribosomal protein S12 | NC_004556 | 2355257 | 2355631 |
| **Pangenome** | *pfp* | Pyrophosphate--fructose 6-phosphate 1-phosphotransferase | NC_004556 | 283945 | 282692 |
| **Pangenome** | *mraY* | Phospho-N-acetylmuramoyl-pentapeptide-transferase | NC_004556 | 2196607 | 2195522 |
| **Pangenome** | *rplW* | 50S ribosomal protein L23 | NC_004556 | 538248 | 538550 |
| **Pangenome** | *rpsK* | 30S ribosomal protein S11 | NC_004556 | 548018 | 548410 |
| **Pangenome** | *cyoC* | Cytochrome bo(3) ubiquinol oxidase subunit 3 | NC_004556 | 768096 | 768704 |
| **Pangenome** | *rph* | Ribonuclease PH | NC_004556 | 885458 | 886183 |
| **Pangenome** | group_3514 | S-methyl-5'-thioinosine phosphorylase | NC_004556 | 1604269 | 1605015 |
| **Pangenome** | *rplK* | 50S ribosomal protein L11 | NC_004556 | 2366767 | 2367195 |
| **Pangenome** | group_3660 | hypothetical protein | NC_004556 | 1911392 | 1910085 |
| **Pangenome** | *prfC* | Peptide chain release factor 3 | NC_004556 | 182199 | 183803 |
| **Pangenome** | *qmcA* | Protein QmcA | NC_004556 | 194156 | 193200 |
| **Pangenome** | group_3901 | NADH-quinone oxidoreductase subunit 11 | NC_004556 | 327057 | 327362 |
| **Pangenome** | *phaB* | Acetoacetyl-CoA reductase | NC_004556 | 333231 | 333971 |
| **Pangenome** | *rpsU* | 30S ribosomal protein S21 | NC_004556 | 1903305 | 1903520 |
| **Pangenome** | *ihfA* | Integration host factor subunit alpha | NC_004556 | 2240030 | 2240329 |



Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **Pangenome** | *gcvR* | Glycine cleavage system transcriptional repressor | NC_004556 | 2025403 | 2025978 |
| **Pangenome** | *prfB* | Peptide chain release factor RF2 | NC_004556 | 493039 | 494164 |
| **Pangenome** | *atpC* | ATP synthase epsilon chain | NC_004556 | 528547 | 528125 |
| **Pangenome** | *rplB* | 50S ribosomal protein L2 | NC_004556 | 538562 | 539389 |
| **Pangenome** | *clpP* | ATP-dependent Clp protease proteolytic subunit | NC_004556 | 564751 | 565377 |
| **Pangenome** | *kdsA* | 2-dehydro-3-deoxyphosphooctonate aldolase | NC_004556 | 666942 | 667772 |
| **Pangenome** | group_4230 | hypothetical protein | NC_004556 | 997348 | 997899 |
| **Pangenome** | *lepA* | Elongation factor 4 | NC_004556 | 1505205 | 2161315 |
| **Pangenome** | *scoF* | Cold shock protein ScoF | NC_004556 | 1604179 | 1603976 |
| **Pangenome** | *pglF* | UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase | NC_004556 | 1693449 | 1691542 |
| **Pangenome** | *sucC* | Succinate--CoA ligase [ADP-forming] subunit beta | NC_004556 | 1691542 | 1693449 |
| **Pangenome** | *rpsR* | 30S ribosomal protein S18 | NC_004556 | 2288237 | 2288007 |
| **Pangenome** | *asnS* | Asparagine--tRNA ligase | NC_004556 | 2289479 | 2290879 |
| **Pangenome** | *icd_1* | Isocitrate dehydrogenase [NADP] | NC_004556 | 2320637 | 2321644 |
| **Pangenome** | *rpoC* | DNA-directed RNA polymerase subunit beta' | NC_004556 | 2360104 | 2355881 |
| **Pangenome** | *dnaA* | Chromosomal replication initiator protein DnaA | NC_004556 | 146 | 1465 |
| **Pangenome** | *accC* | Biotin carboxylase | NC_004556 | 45068 | 46435 |
| **Pangenome** | *nusA* | Transcription termination/antitermination protein NusA | NC_004556 | 242734 | 244245 |
| **Pangenome** | *rpsO* | 30S ribosomal protein S15 | NC_004556 | 248684 | 248944 |
| **Pangenome** | *nuoB* | NADH-quinone oxidoreductase subunit B | NC_004556 | 318051 | 318605 |
| **Pangenome** | *nuoD* | NADH-quinone oxidoreductase subunit D | NC_004556 | 319406 | 320713 |
| **Pangenome** | *nuoI* | NADH-quinone oxidoreductase subunit I | NC_004556 | 325899 | 326387 |


Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **Pangenome** | *phoP_1* | Virulence transcriptional regulatory protein PhoP | NC_004556 | 1949318 | 1948635 |
| **Pangenome** | *sasA_2* | Adaptive-response sensory-kinase SasA | NC_004556 | 1948619 | 1947171 |
| **Pangenome** | *btuD_1* | Vitamin B12 import ATP-binding protein BtuD | NC_004556 | 1928285 | 1929586 |
| **Pangenome** | *mlaE* | putative phospholipid ABC transporter permease protein MlaE | NC_004556 | 1922396 | 1923145 |
| **Pangenome** | *rpmI* | 50S ribosomal protein L35 | NC_004556 | 2244482 | 2244679 |
| **Pangenome** | *ftsL* | Cell division protein FtsL | NC_004556 | 2201585 | 2201322 |
| **Pangenome** | *ftsA* | Cell division protein FtsA | NC_004556 | 2189757 | 2188522 |
| **Pangenome** | *fda* | Fructose-bisphosphate aldolase class 1 | NC_004556 | 2162319 | 2161315 |
| **Pangenome** | *ettA* | Energy-dependent translational throttle protein EttA | NC_004556 | 2038245 | 2039909 |
| **Pangenome** | *sdhA* | Succinate dehydrogenase flavoprotein subunit | NC_004556 | 441564 | 443354 |
| **Pangenome** | *lolC* | Lipoprotein-releasing system transmembrane protein LolC | NC_004556 | 444958 | 446199 |
| **Pangenome** | *atpD1* | ATP synthase subunit beta 1 | NC_004556 | 528626 | 530026 |
| **Pangenome** | *atpH_1* | ATP synthase subunit c | NC_004556 | 534087 | 533782 |
| **Pangenome** | *rplN* | 50S ribosomal protein L14 | NC_004556 | 541658 | 542026 |
| **Pangenome** | *clpX* | ATP-dependent Clp protease ATP-binding subunit ClpX | NC_004556 | 565505 | 566785 |
| **Pangenome** | *pyrG* | CTP synthase | NC_004556 | 665117 | 666781 |
| **Pangenome** | *hprK* | HPr kinase/phosphorylase | NC_004556 | 784843 | 783893 |
| **Pangenome** | *lptB* | Lipopolysaccharide export system ATP-binding protein LptB | NC_004556 | 787669 | 786947 |
| **Pangenome** | *clpS* | ATP-dependent Clp protease adapter protein ClpS | NC_004556 | 813335 | 813655 |
| **Pangenome** | *clpA* | ATP-dependent Clp protease ATP-binding subunit ClpA | NC_004556 | 813788 | 816064 |
| **Pangenome** | *infA* | Translation initiation factor IF-1 | NC_004556 | 816631 | 816413 |
| **Pangenome** | *sufB* | FeS cluster assembly protein SufB | NC_004556 | 850414 | 848957 |



Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **Pangenome** | *hslV* | ATP-dependent protease subunit HslV | NC_004556 | 857829 | 858380 |
| **Pangenome** | *xcpT* | Type II secretion system protein G | NC_004556 | 902880 | 903314 |
| **Pangenome** | *clp* | CRP-like protein Clp | NC_004556 | 928760 | 929449 |
| **Pangenome** | *ybaB* | Nucleoid-associated protein YbaB | NC_004556 | 1261433 | 1261113 |
| **Pangenome** | *guaB* | Inosine-5'-monophosphate dehydrogenase | NC_004556 | 1689427 | 1687970 |
| **Pangenome** | *rpsA* | 30S ribosomal protein S1 | NC_004556 | 1698067 | 1696391 |
| **Pangenome** | *efp* | Elongation factor P | NC_004556 | 1734038 | 1733472 |
| **Pangenome** | *sufA* | Protein SufA | NC_004556 | 2289258 | 2288920 |
| **Pangenome** | *fusA* | Elongation factor G | NC_004556 | 2354638 | 2352521 |
| **Pangenome** | *purF* | Amidophosphoribosyltransferase | NC_004556 | 1059528 | 1058071 |
| **Pangenome** | *maeB* | NADP-dependent malic enzyme | NC_004556 | 346574 | 348868 |
| **Pangenome** | *nrdB* | Ribonucleoside-diphosphate reductase subunit beta | NC_004556 | 578511 | 579551 |
| **Pangenome** | *typA* | GTP-binding protein TypA/BipA | NC_004556 | 589082 | 590911 |
| **Pangenome** | *nuoF* | NADH-quinone oxidoreductase subunit F | NC_004556 | 321242 | 322576 |
| **Pangenome** | *sasA_4* | Adaptive-response sensory-kinase SasA | NC_004556 | 2252003 | 2253340 |
| **Pangenome** | *hfq* | RNA-binding protein Hfq | NC_004556 | 88377 | 88099 |
| **Pangenome** | *sucD* | Succinate--CoA ligase [ADP-forming] subunit alpha | NC_004556 | 2266275 | 2267147 |
| **Pangenome** | *rpmH* | 50S ribosomal protein L34 | NC_004556 | 2519722 | 2519582 |
| **Pangenome** | *rpsJ* | 30S ribosomal protein S10 | NC_004556 | 536660 | 536971 |
| **Pangenome** | *minE* | Cell division topological specificity factor | NC_004556 | 697554 | 697297 |
| **Pangenome** | *nusG* | Transcription termination/antitermination protein NusG | NC_004556 | 2367950 | 2367393 |



Funded by the European Union

| Type | Gene/target name | Gene/target function | NCBI reference sequence[a] | Forward position | Reverse position |
|---|---|---|---|---|---|
| **Plasmid**** | p*XF*64-HB_ESVL | p*XF*64-HB_ESVL plasmid, entire sequence | NZ_CM013392.1 | 58914 | 13 |
| **Plasmid**** | p*XF*64-HB_ESVL | p*XF*64-HB_ESVL, partial sequence | NZ_CM013392.1 | 40189 | 41429 |
| **Plasmid** | p*XF*PD1.3 | p*XF*PD1.3 plasmid, partial sequence | NC_019302 | 457 | 1171 |

[a] NCBI reference sequences correspond to: NZ_CM013392= *Xylella fastidiosa* subsp. *multiplex* strain ESVL genome; NC_019302= *X. fastidiosa* plasmid pUCLAc; NC_002488= *X. fastidiosa* subsp. *pauca* strain 9a5c genome; NC_010513= *X. fastidiosa* subsp. *multiplex* M12 genome; and NZ_CP0208= *X. fastidiosa* subsp. *pauca* strain De Donno genome.



Funded by the European Union