


BEXYL
BEYOND XYLELLA

# Development of genetic markers deployable for rapid screening of olive and almond genotypes

Deliverable 4.3

31 August 2026

Author(s): Soraya Mousavi, Roberto Mariotti, Sara Pinosio (P5-CNR); Ofir Bahar, Tamar Azoulay- Shemer, Rotem Beja, Hillel Brukental (P11-ARO-VOL).

Funded by the European Union

# DELIVERABLE INFORMATION

| | |
|---|---|
| **Start of the project:** | 01 Sep 2022 |
| **Duration:** | 48 months |
| **Coordinator:** | Blanca B. Landa |
| **Deliverable title:** | Development of genetic markers deployable for rapid screening of olive and almond genotypes |
| **Deliverable n°:** | D4.3 |
| **Nature of the deliverable:** | Report |
| **Dissemination level:** | Sensitive |
| **WP contributing:** | WP4 |
| **Lead beneficiary:** | P11-ARO-VOL |
| **Due date of deliverable:** | 31 August 2026 |
| **Actual submission date:** | 31 August 2026 |

## Deliverable status

| Ver. | Status | Date | Author(s) |
|---|---|---|---|
| 1.0 | Final | 31 August 2026 | Soraya Mousavi, Roberto Mariotti, Sara Pinosio (P5-CNR); Ofir Bahar, Tamar Azoulay- Shemer, Rotem Beja, Hillel Brukental (P11-ARO-VOL). |

**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

# Table of Contents

*1. INTRODUCTION* ..... *4*

*2. MATERIALS AND METHODS* ..... *5*

**2.1 Olive resistance marker development** ..... **5**
2.1.1 Plant material and DNA extraction ..... 5
2.1.2 Genetic identification and parentage validation via nuclear SSR markers ..... 5
2.1.3 SPET genotyping and bioinformatic processing ..... 6
2.1.4 Plant phenotyping ..... 6
2.1.5 Construction of genetic map for Leccino progeny ..... 7
2.1.6 QTL analysis ..... 7

**2.2 Almond resistance marker development** ..... **7**
2.2.1 Plant material and DNA extraction ..... 7
2.2.2 Plant inoculation and phenotyping ..... 8
2.2.3 SNP genotyping ..... 8
2.2.4 Genome-wide association study (GWAS) ..... 8
2.2.5 F2 population and planned Genotyping By Sequencing (GBS) genotyping ..... 9

*3. RESULTS* ..... *9*

**3.1. Olive** ..... **9**
3.1.1 SSR genetic identification and parentage analysis ..... 9
3.1.2 Categorization of progenies based on phenotypic response ..... 9
3.1.3 High-density genetic linkage map construction ..... 11
3.1.4 QTL mapping for *Xylella fastidiosa* resistance ..... 11

**3.2. Almond** ..... **15**
3.2.1 Almond germplasm phenotyping and SNP association ..... 15
3.2.2 ‘Lauranne’ self-pollinated F2 progeny phenotyping ..... 17

*4. DISCUSSION* ..... *18*

**4.1 Population curation and high-density linkage map collinearity** ..... **18**

**4.2 Parental segregation dynamics and the oligogenic architecture of *X. fastidiosa* resistance** ..... **18**

**4.3 Future directions, ongoing genetic mapping in diverse 'Leccino' progenies** ..... **19**

**4.4 Genetic variation and mapping of *Xylella fastidiosa* resistance in almond** ..... **20**

*5. CONCLUSIONS* ..... *20*

*6. REFERENCES* ..... *21*



Funded by the European Union

# 1. INTRODUCTION

Recognized globally as a severe agricultural threat, the xylem-dwelling bacterium *Xylella fastidiosa* (*Xf*) endangers a diverse array of vital perennial crops, such as olive, almond, citrus, and grapevine, alongside various forest and ornamental plants. Upon its arrival in Europe and the Mediterranean, *Xf* led to catastrophic agricultural and ecological consequences, particularly by inciting Olive Quick Decline Syndrome (OQDS) across southern Italy, alongside severe outbreaks in almonds and other susceptible Mediterranean species. Because there are currently no effective treatments to cure infected plants, deploying robust genetic resistance is widely considered the most viable long-term strategy to combat the disease (Morelli et al., 2021). Researchers have documented natural differences in how various plant cultivars handle the infection. For instance, the olive variety 'Leccino' demonstrates a much lower vulnerability to *Xf* than highly sensitive varieties, strongly pointing to the existence of inherited factors that allow the plant to recognize the threat and mount a defense (Walker et al., 2023). Similarly, the almond cultivar 'Lauranne' was found to be more resistant to *Xf* subsp. *fastidiosa* infection (Vanunu et al. 2026).

Genome-Wide Association Studies (GWAS) demonstrate that olive host resistance to *Xf* is governed by an oligogenic framework rather than a diffuse polygenic background. The primary layer of this defense relies on high-impact loss-of-function mutations in susceptibility factors such as EVM0013689 and EVM0017391, which deprive *Xf* of the biological pathways required for rapid multiplication and systemic damage. Beyond these essential knockouts, a pleiotropic locus on Chromosome 22 coordinates a secondary defense network (GWAS D4.2). In this network, Receptor-like protein kinase 2 detects the pathogen at the cell surface to restrict bacterial load (Rodrigues et al., 2013; Giampetruzzi et al., 2016), while E3 ubiquitin ligases and key transcription factors (NAC and ERF) dynamically regulate immune signaling and clear damaged proteins (Rodrigues et al., 2013; La Notte et al., 2024; Balan et al., 2025). To survive the vascular blockages caused by the infection, the plant also relies on Long-chain acyl-CoA synthetase 9 to fortify pit membranes through lipid metabolism and cuticular waxes (Balan et al., 2025), while the nearby Sodium/hydrogen exchanger 1-like gene maintains osmotic stability and cellular pH to combat severe physiological drought (Giampetruzzi et al., 2016; Balan et al., 2025).

Quantitative Trait Loci (QTL) mapping is a trusted technique for unraveling complex agricultural traits and locating disease resistance regions in perennial fruit trees (Khan & Korban, 2022). This method has successfully identified key resistance loci across several fruit species, including black spot resistance on linkage group 15 in Chinese sand pear (Hongyan et al., 2023), 26 Huanglongbing tolerance QTLs clustered across four linkage groups in citrus (Huang et al., 2023), European canker resistance segments in apples (Skytte af Sätra et al., 2023), and 26 distinct regions controlling grapevine heat tolerance (Pettenuzzo et al., 2025).

Although conventional olive breeding is severely constrained by a complex genome and long juvenile phases—ranging from 3 to 4 years in almond to up to 10–15 years in olive—modern QTL mapping combined with Whole Genome Sequencing (WGS) and transcriptomics (RNA-Seq) may assist in overcoming these hurdles. In olive, for instance, a multi-omics approach applied to a bi-parental F1 population successfully pinpointed a major QTL on chromosome 21 regulating verbascoside biosynthesis, a key phenolic defense metabolite (Pizzi et al., 2026). Similarly, in almond, high-density marker mapping and genome-wide association studies



Funded by the European Union

(GWAS) are accelerating the identification of candidate loci associated with physiological adaptations and pathogen responses (Brukental et al., 2021; Vanunu et al., 2026.

Extensive monitoring of wild olive plants in areas heavily plagued by *Xf* has brought to light many resistant individuals. Analysis of parentage and natural trait variation proves that 'Leccino' is an exceptional source of resistance, successfully transmitting defensive traits to a large majority (67%) of its offspring, even when mated with highly vulnerable cultivars (La Notte et al., 2024). Relying on these genetic findings, this report will investigate QTL mapping in the progeny of 'Leccino' (specifically the 'Leccino' x 'Oliana' cross) to evaluate their reaction to *Xf*. In parallel, systematic screening of almond genetic resources has identified 'Lauranne' as a premier source of host resistance against *Xf* subsp. *fastidiosa* (Vanunu et al., 2026). To dissect the genetic architecture of this resistance, a collection of 38 diverse almond germplasm accessions was phenotyped and subjected to a genome-wide association study (GWAS). In parallel, a self-pollinated F2 population of 'Lauranne' was generated, inoculated, and phenotyped, with DNA samples already extracted and pending to be submitted to a sequencing facility to undergo high-density genotyping-by-sequencing (GBS).

By identifying the exact genetic factors that cause these differing reactions in crossbred (or self) populations, we expect to accelerate breeding initiatives and forge novel tactics to keep the disease in check.

# 2. MATERIALS AND METHODS

## 2.1 Olive resistance marker development

### 2.1.1 Plant material and DNA extraction

Young leaf samples were collected in 2024 from 179 F1 individuals derived from the 'Leccino' × 'Oliana' controlled cross, along with both parent cultivars. This F1 progeny was obtained in 2020, subsequently in August 2021 the individual potted seedlings were exposed to infected spittlebugs (specimens of *Philaenus spumarius*) for 48-72h of inoculation access period in seedlings, prior to being transplanted in the field in October 2021. The experimental plot was located within the designated infected area for *Xf* subsp. *pauca* to further expose the plants to the natural pressure of inoculum.

Total genomic DNA was extracted using the Zymo Research Mini Plant/Seed DNA Extraction Kit (Zymo Research, Irvine, CA, USA) following the manufacturer's instructions. Plant tissues were homogenized in original extraction bags using a HOMEX XS homogenizer (Bioreba AG, Reinach, Switzerland). DNA concentration and purity were initially evaluated using a NanoPhotometer® NP80 spectrophotometer (IMPLEN, Munich, Germany).

### 2.1.2 Genetic identification and parentage validation via nuclear SSR markers

To confirm parentage and verify the genetic identity of the progeny, all individuals were genotyped using a panel of 10 highly polymorphic nuclear simple sequence repeat (SSR) loci: DCA3, DCA5, DCA9, DCA16, DCA18, EMO-90, GAPU71B, GAPU101, GAPU103A, and UDO99-043 (Trujillo et al., 2014; El Bakkali et al., 2019). Chromosomal locations for these loci were



Funded by the European Union

assigned according to Valeri et al. (2022). Forward primers were 5'-labeled with fluorescent dyes (VIC, PET, or NED). PCR amplification products were separated by capillary electrophoresis on a SeqStudio Genetic Analyzer using the GeneScan™ 500 LIZ Internal Size Standard, and allele sizing was performed using GeneMapper™ Software v.6 (Thermo Fisher Scientific, Waltham, MA, USA). Raw fractional fragment sizes were rounded to discrete integer values to generate a standardized SSR allele matrix. True-to-type hybrid validation and parentage exclusion analyses were performed using CERVUS software v.3.0.7 (Marshall et al. 1998) to identify selfing events or off-target cross-pollination. Non-hybrid individuals displaying non-parental alleles were excluded from downstream genetic mapping.

### 2.1.3 SPET genotyping and bioinformatic processing

DNA concentration was precisely quantified using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). High-density Single Nucleotide Polymorphism (SNP) genotyping of the validated 'Leccino' × 'Oliana' progeny was conducted using a customized Single Primer Enrichment Technology (SPET) probe panel previously designed and benchmarked for olive germplasm (Deliverable 4.2). Targeted enrichment libraries were constructed starting from 100 ng of input DNA using the Allegro Targeted Genotyping protocol (Tecan Genomics, Männedorf, Switzerland), incorporating 6-bp Unique Molecular Identifiers (UMIs) on original DNA fragments. Library concentration and fragment size distribution were verified on a Qubit 2.0 Fluorometer and an Agilent Bioanalyzer High Sensitivity DNA assay (Agilent Technologies, Santa Clara, CA, USA), respectively. Sequencing was performed at IGA Technology Services (IGATech, Udine, Italy) using an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA) in paired-end configuration (2 × 150 bp). Raw BCL files generated by the instrument were converted into FASTQ format using the standard Illumina pipeline. Short read mapping and SNP detection for SPET benchmark has been performed as reported in detail in the deliverable 4.2.

### 2.1.4 Plant phenotyping

Starting in 2024, the plants were monitored to assess the development of symptoms putatively induced by *Xf* subsp. *pauca*. Sampling was conducted over two consecutive years to determine infection frequency and quantify bacterial population size. Seedlings were also scored for their habitus, in particular if they had switched to the mature phase or still exhibited juvenile traits. Symptoms were scored on an empirical 0–5 scale, while the bacterial population was quantified using a standard qPCR protocol based on Harper et al. (2010), that included a 10-fold serial dilution standard curve with known bacterial concentrations. Diagnostic tests were performed using total DNA extracted from 1- to 2-year-old mature shoots.

Additionally, a panel of 10 F1-seedlings exhibiting a short juvenile phase was propagated by grafting. Six to eight months post-grafting, these plants were needle-inoculated with pure bacterial culture to validate the field data collected on the single individual seedling, with further experimental evidence. Visual inspections and leaf sampling were carried out 6 and 12 months after inoculation to confirm that bacterial colonization extended from the inoculation points.



Funded by the European Union

### 2.1.5 Construction of genetic map for 'Leccino' progeny

A high-density genetic linkage map was developed for an F1 full-sib olive population derived from the cross 'Leccino' × 'Oliana' using Lep-MAP3 and SPET-derived SNP markers. After parental genotype checking and filtering for missing data, minor allele frequency, and segregation distortion, the dataset was reduced from 235,210 informative SNPs to 73,488 high-quality markers. Linkage analysis recovered 23 linkage groups corresponding to the 23 olive chromosomes, with only two singleton markers remaining unassigned. Marker ordering was performed independently for each linkage group, producing male, female, and integrated maps.

Map quality was assessed by comparing genetic and physical marker positions. Most linkage groups displayed a highly monotonic relationship between genetic and physical coordinates, indicating accurate marker assignment and ordering. Only a small number of isolated markers deviated from the general trend and did not affect the overall map structure. Overall, the Lep-MAP3 map showed excellent collinearity with the reference genome and substantially improved marker ordering and map consistency compared with previous OneMap-based analyses.

### 2.1.6 QTL analysis

A total of 74,710 high-quality SNPs were tested for association with resistance to *Xf* subsp. *pauca* using two phenotypic datasets: (i) an ordinal resistance score ranging from 1 (highly susceptible) to 3 (highly resistant), and (ii) a quantitative resistance phenotype (CQ). Marker-trait associations were evaluated using a Kruskal-Wallis test for both phenotypic datasets, while the quantitative phenotype was additionally analysed using ANOVA. For each marker, nominal P-values, −log10(P), and Benjamini-Hochberg FDR values were calculated and integrated with the genetic map.

## 2.2 Almond resistance marker development

### 2.2.1 Plant material and DNA extraction

Evaluation of almond genotypes for resistance to *Xf* subsp. *fastidiosa* focused on two distinct resources: 1) a collection of 38 accessions from Israel's almond germplasm collection, representing diverse geographic origins; and 2) an F2 mapping population comprising approximately 200 progenies derived from self-pollinated resistant cultivar 'Lauranne' (termed F2). The 38 germplasm accessions were grafted onto GF677 rootstocks and maintained in 4-L pots, while the 'Lauranne' F2 progenies were sown directly from seed in identical pots. All potted plants were grown in a net-house at the Volcani Institute under an automated drip irrigation and fertigation system.

Young leaf tissue was collected from all evaluated plants. Genomic DNA was extracted from young leaves using the Plant/Fungi DNA Isolation Kit (Norgen Biotek Corp., Canada), following the manufacturer's instructions. Following concentration and purity assessments, samples meeting all quality criteria were submitted for SNP genotyping.



Funded by the European Union

### 2.2.2 Plant inoculation and phenotyping

Almond plants were mechanically inoculated with *Xf* subsp. *fastidiosa* SN17-4 strain, using the pin-prick method (Zecharia et al. 2022) during late spring (May–June). Briefly, a 10-µl drop of bacterial suspension ($10^8$ CFU/ml) was placed on green and tender stems (2-4 per plant) and pricked 5 times with a sterile 23G needle. Control plants were inoculated with PD3 medium.

Disease phenotyping was based on visual assessment of symptoms and molecular detection. For visual evaluation, the percentage of inoculated stems displaying clear almond leaf scorch (ALS) symptoms was recorded for each genotype. For molecular detection, quantitative PCR (qPCR) following Harper et al. (2010) protocol was carried out on leaf sample from each inoculated stem and the percentage of infected stems per genotype was determined. Disease phenotyping was done for two consecutive years.

### 2.2.3 SNP genotyping

Genome-wide SNP genotyping was performed using the targeted SNP-seq platform of LGC Genomics (Germany), following the targeted genotyping protocol previously described by Brukental et al. (2021). The assay was designed to genotype approximately 5,000 SNP markers distributed across the genome. SNP physical coordinates were mapped to the *Prunus dulcis* cv. Lauranne reference genome (NCBI accession GCA_008632915.2; https://www.ncbi.nlm.nih.gov/bioproject/553424).

Because samples were processed in two separate genotyping runs, technical variation in marker recovery between batches limited the shared panel to 3,300 SNPs that passed initial quality filtering. The two datasets were therefore merged using this common marker set.

Following additional filtering for read quality and a minor allele frequency (MAF) threshold of 0.05, a final panel of 2,050 high-quality polymorphic SNP markers was retained for downstream genome-wide association analysis.

### 2.2.4 Genome-wide association study (GWAS)

Genome-wide association analysis was conducted in R using the rrBLUP package. Prior to association testing, the filtered SNP dataset was used to construct a genomic kinship matrix characterizing the genetic relationships among accessions. Marker–trait associations were evaluated using the GWAS implementation in rrBLUP, with the kinship matrix included in the association model to control for genetic relatedness among accessions. GWAS was performed separately for two representations of the *Xf* subsp. *fastidiosa* resistance phenotype: arithmetic genotype means and estimated marginal means (EMMs). Both phenotypic formulations produced nearly identical association patterns, displaying only negligible differences. All analyses were performed across the final panel of 2,050 SNP markers that passed the genotyping quality filters and the MAF threshold of 0.05.



Funded by the European Union

### 2.2.5 F2 population and planned Genotyping By Sequencing (GBS) analysis

High-quality DNA samples from the phenotyped 'Lauranne' F2 population are being prepared for GBS analysis at a sequencing facility. Unlike the targeted SNP-seq approach applied to the germplasm collection, GBS is not restricted to the predefined 5,000-SNP panel and will yield a substantially higher genome-wide marker density. This high-density SNP dataset will enable high-resolution mapping of *X. fastidiosa* subsp. *fastidiosa* resistance loci within the segregating F2 population.

# 3. RESULTS

## 3.1. Olive

### 3.1.1 SSR genetic identification and parentage analysis

Molecular authentication of the 179 genotypes originating from the 'Leccino' × 'Oliana' mapping population was performed using a panel of SSR markers. Parentage exclusion analysis validated 166 individuals as true F1 hybrids resulting from the cross between 'Leccino' and 'Oliana'. In contrast, 13 genotypes exhibited unexpected allelic patterns inconsistent with the male parent ('Oliana'), likely representing contamination from foreign airborne pollen. To preserve the genetic integrity of the mapping population, these 13 non-conforming individuals were removed, leaving a refined population of 166 verified hybrids for genetic map construction and QTL analysis.

### 3.1.2 Categorization of progenies based on phenotypic response

Four years after planting the F1 progeny in the field, the majority of the plants were in the mature stage (~76%) and were bearing fruit. Consequently, they had overcome the juvenile stage, during which plants are generally more vigorous and less prone to symptom development, even when infected, which typically confounds the consistent classification of their response to *Xf* subsp. *pauca*.

Upon the first round of visual inspections (2024), 10% of the plants showed signs of shoot dieback/desiccation with symptom severity scores ranging from 1 to 3. However, some of the plants with the highest scores exhibited juvenile traits and tested negative for the bacterium, suggesting that their extensive decline was unrelated to *Xf* subsp. *pauca*. Consequently, these individuals were excluded from further analyses. Another 10% of the plants were scored for suspicious mild symptoms (score 0.5), among these only 35% tested positive, indicating that most likely these suspicious symptoms were not caused by the infections.

During the 2$^{nd}$ round of inspections (2025), 19% of the plants showed symptoms with symptom severity scores ranging from 1 to 3, confirming also during this 2$^{nd}$ assessment that some of them although highly symptomatic were not infected. Similarly to 2024, 13% of the plants showed suspicious mild symptoms (score 0.5), with 50% of them infected by *Xf*. With respect to the infections, 62% of the plants tested not infected with the remaining yielding for the majority Cq values >26 (75%) corresponding to a level of bacterial population below



Funded by the European Union

$10^4$CFU/ml; whereas few plants generated Cq values from 20 to 25 (25%) corresponding to a level of bacterial population of $10^5$-$10^6$ CFU/ml.

The symptom scores and the bacterial load detected in the plants were used to categorize the F1 individuals into six classes (Table 1).

**Table 1.** Criteria used to assign the 'Leccino' × 'Oliana' F1 progeny to six phenotypic response classes to *Xylella fastidiosa* subsp. *pauca*

| Class[a] | Qualitative assessment | Symtpom score | *Xf* subsp. *pauca* bacterial load (range) |
|---|---|---|---|
| 6 | Highly resistant | 0 | not detected or < $10^3$ CFU/ml |
| 5 | Resistant | 0 | $10^3$-$10^4$ CFU/ml |
| 4 | Medium resistant/tolerant | 0 | $10^5$-$10^6$ CFU/ml |
| 3 | Medium susceptible | 0.5-1 | $10^3$-$10^6$ CFU/ml |
| 2 | Suscpetible | 1.5-2 | $10^3$-$10^6$ CFU/ml |
| 1 | Highly susceptible | >2 | $10^3$-$10^6$ CFU/ml |

[a] Each class combines the field symptom severity score (empirical 0–5 scale, recorded in the 2024 and 2025 assessments) with the bacterial load estimated by qPCR and expressed as CFU/mL, derived from Cq values against a 10-fold serial dilution standard curve of known bacterial concentrations. Class 6 denotes the most resistant and class 1 the most susceptible phenotype; colour shading (green to red) illustrates the resistance-to-susceptibility gradient.

While 15 individuals were excluded from the classification due to the contrasting results gathered for the two parameters (i.e., highly symptomatic but negative for the bacterium), the remaining individuals were distributed among the different classes with the frequencies indicated in Table 2.

**Table 2.** Distribution of the 'Leccino' × 'Oliana' F1 progeny among the six phenotypic response classes defined in Table 1

| Class | Qualitative assessment | Frequency of individuals (%)[a] |
|---|---|---|
| 6 | Highly resistant | 64.24 |
| 5 | Resistant | 3.97 |
| 4 | Medium resistant/tolerant | 11.92 |
| 3 | Medium susceptible | 9.27 |
| 2 | Suscpetible | 3.31 |
| 1 | Highly susceptible | 7.28 |

[a] Frequencies are given as the percentage of classified individuals (n = 151); 15 of the 166 SSR-validated F1 hybrids were excluded because the two scoring parameters were contradictory (high symptom severity but negative qPCR detection). Class 6 denotes the most resistant and class 1 the most susceptible phenotype; colour shading (green to red) illustrates the resistance-to-susceptibility gradient.

This phenotypic classification served as the basis for QTL mapping using the SNP marker dataset, as described in paragraph 3.1.4. The occurrence of individuals distributed among all classes indicates that a segregation occurred in the progeny, even though their uneven



Funded by the European Union

balance between resistant and susceptible individuals may affect the efficiency of the genetic mapping. Furthermore, the high proportion of *Xf*-free individuals might stem from infection escape or erratic vector transmission rather than true genetic resistance.

### 3.1.3 High-density genetic linkage map construction

A high-density genetic linkage map was constructed for an F1 full-sib olive population derived from the cross 'Leccino' × 'Oliana' using Lep-MAP3 and SPET-derived SNP markers. Following parental genotype verification and filtering for missing data, minor allele frequency, and segregation distortion, the dataset was reduced from an initial 235,210 informative SNPs to 73,488 high-quality markers. Linkage analysis successfully recovered 23 linkage groups corresponding directly to the 23 olive chromosomes, leaving only two singleton markers unassigned. Male, female, and integrated maps were produced by performing marker ordering independently for each linkage group.
The quality of the map was evaluated by comparing the genetic coordinates (cM) against physical positions (Mb). Most linkage groups displayed a highly monotonic relationship between genetic and physical positions, demonstrating accurate marker assignment and ordering. This indicates excellent collinearity with the reference genome and represents a substantial improvement in map consistency compared to previous analysis methods (Fig. 1).

#### 3.1.4 QTL mapping for *Xylella fastidiosa* resistance

To identify genomic regions associated with resistance, 74,710 high-quality SNPs were tested against two distinct phenotypic datasets: an ordinal resistance score ranging from 1 (highly susceptible) to 6 (highly resistant) and a quantitative resistance phenotype based on bacterial load (CQ). Association testing was performed using the Kruskal-Wallis test for both datasets, and ANOVA was additionally utilized to evaluate the quantitative CQ phenotype.
Nominal P-values, and Benjamini-Hochberg FDR values were calculated and integrated with the genetic map. Under strict multiple-testing correction, no single marker achieved genome-wide significance (FDR values exceeded the 0.05 threshold across all tests). This suggests that host resistance to *Xf* in this pedigree is controlled by multiple loci of small-to-moderate individual effects rather than a single major-effect gene (Fig. 2). However, comparing the 50 most significant markers across the different statistical approaches revealed a remarkable overlap of recurrent association signals, concentrating on six primary genomic intervals:

- Chr7 (80–110 cM): This region was identified as a primary association zone. The top marker, physical position ~33.68 Mb, genetic position 109.26 cM, showed the highest overall statistical significance in the quantitative models with a $-\log_{10}(P)$ of 4.65 in the ANOVA model, 4.35 in the Kruskal-Wallis CQ analysis, and 3.53 in the Kruskal-Wallis ordinal score analysis.
- Chr 15 (94–129 cM): Anchored in physical position ~12.22 Mb, genetic position 94.14 cM, which yielded a $-\log_{10}(P)$ of 3.61 in ANOVA, 3.20 in Kruskal CQ, and 3.58 in Kruskal ordinal score. Other associated markers span up to 129.32 cM and 120.06 cM.
- Chr 16 (104–110 cM): This distal region was at 108.64 cM, exhibiting a $-\log_{10}(P)$ of 3.25 in ANOVA and 3.12 in Kruskal CQ. It also includes marker at 105.56 cM with a score of 3.10 in Kruskal.


Funded by the European Union

- Chr 17 (48–84 cM): Anchored at 47.53 cM, which had a $-\log_{10}(P)$ of 3.84 in ANOVA and 3.77 in Kruskal CQ. The association interval extends downstream to 84.26 cM with marker GWHAOPM00000017_24532674.
- Chr 21 (75–78 cM): This narrow genetic interval exhibited strong consistency, anchored at 77.78 cM, with a $-\log_{10}(P)$ of 3.60 in ANOVA, 3.21 in Kruskal CQ, and a peak of 4.39 in Kruskal ordinal score.
- Chr 22 (12–20 cM): Anchored at 20.37 cM, which showed a $-\log_{10}(P)$ of 3.10 in ANOVA and 3.00 in Kruskal ordinal score.

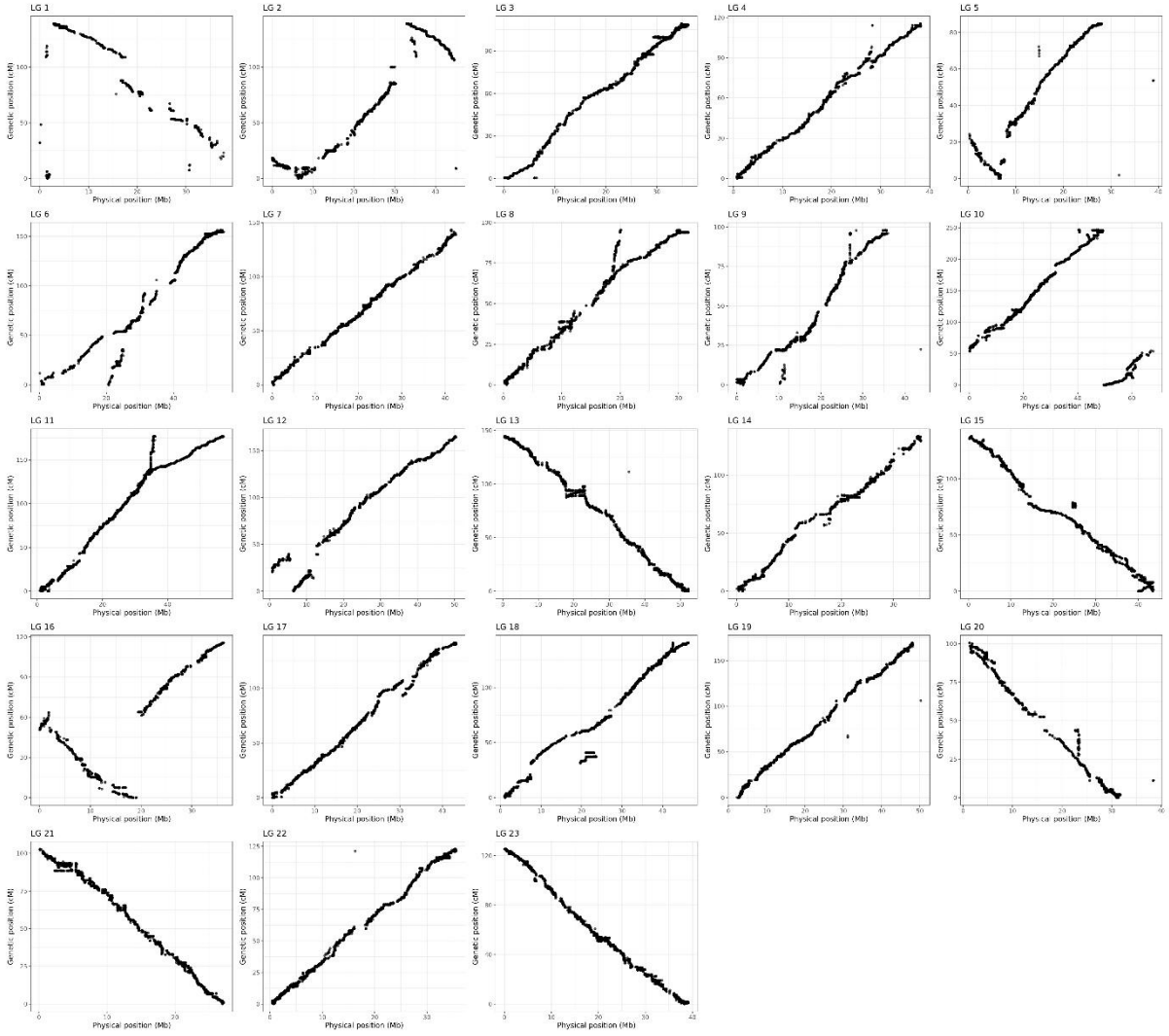



**Figure. 1.** Monotonic collinearity between genetic and physical coordinates across the 23 linkage groups of the 'Leccino' × 'Oliana' integrated map. Genetic positions (centimorgans, cM; y-axis) of the high-quality SNP markers are plotted against their corresponding physical positions (megabases, Mb; x-axis) on the reference genome. Each panel represents one of the 23 reconstructed linkage groups (LG 1 to LG 23). The highly linear and monotonic distribution of markers across the majority of chromosomes indicates accurate marker ordering and robust collinearity between the Lep-MAP3 genetic map and the physical reference assembly, providing high statistical confidence for downstream QTL mapping.



Funded by the European Union

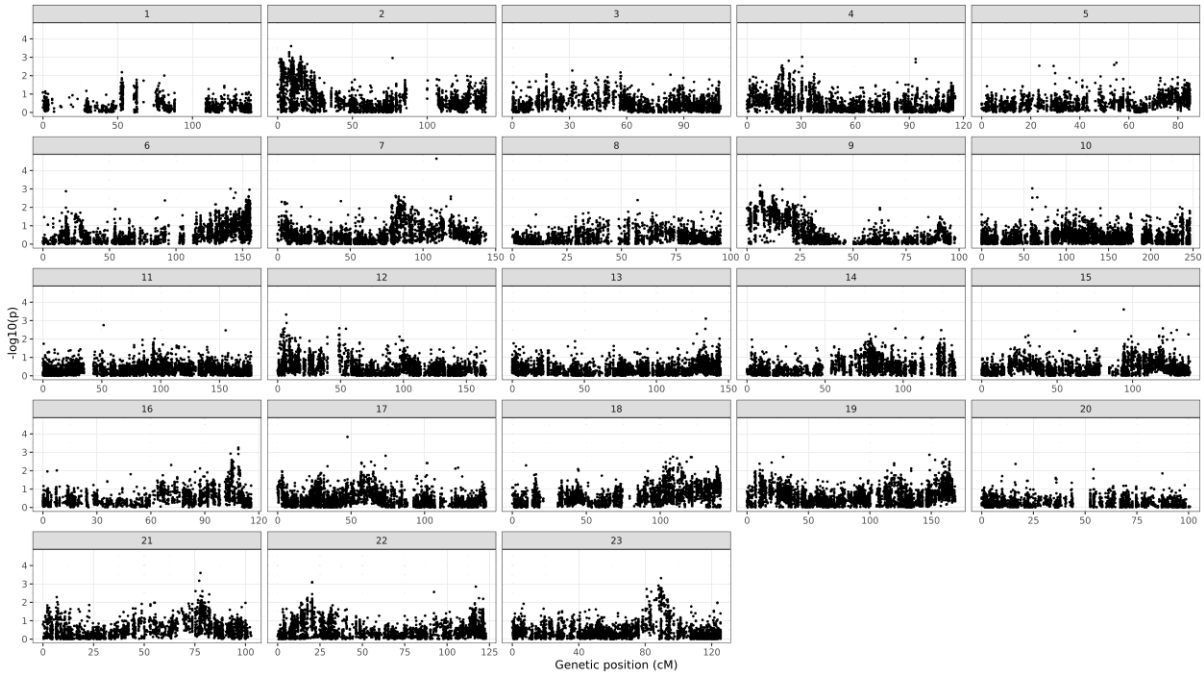



**Figure 2.** Manhattan plot of QTL mapping for *Xylella fastidiosa* quantitative resistance (bacterial load, Cq) in the 'Leccino' × 'Oliana' progeny using ANOVA. The statistical significance (y-axis) of the association tests is plotted against the genetic position in centimorgans (cM; x-axis) across the 23 reconstructed linkage groups (LG 1 to LG 23) of the olive genome. Individual panels represent the mapped chromosomes. The distribution of association signals highlights several localized, suggestive QTL regions, most notably the peak on Chromosome 7 (LG7) reaching a $-\log_{10}(P)$ of 4.65 near the distal end, alongside recurrent peaks on LG15, LG16, LG17, LG21, and LG22. The absence of any peaks exceeding a $-\log_{10}(P)$ value of 5 on the y-axis indicates that no single marker achieved strict genome-wide significance after false discovery rate (FDR) correction, supporting an oligogenic inheritance model governed by multiple moderate-effect loci.

Although no single marker achieved strict genome-wide significance after false discovery rate (FDR) correction, strong statistical consensus between independent parametric (ANOVA) and non-parametric (Kruskal-Wallis) models highlighted a set of candidate genes (Table 3).



Funded by the European Union

**Table 3**. The priority candidate genes associated with *Xylella fastidiosa* subsp. *pauca* olive response

| Gene ID | Linkage Group (LG) | Top Associated Progeny Marker(s) | Peak Significance (–log10 (P)) & Statistical Model | Functional Annotation (*sylvestris* Reference) | Functional Significance in *X. fastidiosa* Host Defense |
|---|---|---|---|---|---|
| EVM0026097 | LG7 | Chr 7_33678228 | 4.65 (ANOVA CQ)<br>4.35 (Kruskal CQ)<br>3.53 (Kruskal Class) | Proteasome subunit alpha type-1-A-like | Highest overall QTL peak. Controls targeted protein degradation and ubiquitin-mediated signalling cascades during active bacterial challenge |
| EVM0030218 | LG21 | Chr 21_8265144 | 4.39 (Kruskal Class)<br>3.60 (ANOVA CQ)<br>3.21 (Kruskal CQ) | Trihelix transcription factor ASR3-like | Primary symptom modifier. Coordinates stress-responsive transcriptional networks that regulate cellular survival and delay symptom expression |
| EVM0052690 | LG17 | Chr 17_15602733 | 3.84 (ANOVA CQ)<br>3.77 (Kruskal CQ)<br>2.92 (Kruskal Class) | ER membrane protein complex subunit 1-like | Endoplasmic Reticulum (ER) Homeostasis. Maintains proteostasis and coordinates secretory pathways under severe vascular and cellular stress |
| EVM0029114 | LG15 | Chr 15_12217367 | 3.61 (ANOVA CQ)<br>3.58 (Kruskal Class)<br>3.20 (Kruskal CQ) | Transmembrane 9 superfamily member 1 | Membrane-localized signalling. Facilitates the transport of proteins and monitors membrane dynamics during vascular occlusion |
| EVM0031945 | LG22 | Chr 22_6372959 | 3.10 (ANOVA CQ)<br>3.00 (Kruskal Class)<br>2.81 (Kruskal CQ) | EXPA3 (Expansin-A3) | Cell Wall Remodelling. Drives expansin-mediated wall loosening, which can be exploited by the path of systemic bacterial migration if not repressed |
| EVM0013689 | LG6 | Chr 6_51803622 | 2.94 (Kruskal Class) | Uncharacterized protein LOC111379872 | Candidate Susceptibility (S) Factor. Depriving the host of this functional allele blocks downstream pathways required for systemic symptom progression |
| EVM0026958 | LG9 | Chr 9_2418713 | 3.19 (ANOVA CQ)<br>3.09 (Kruskal CQ)<br>2.78 (Kruskal Class) | GDSL esterase/lipase CPRD49-like | Lipid-mediated defense. Catalyzes lipid metabolism and wax production to modify vascular hydraulic borders |
| EVM0043930 | LG6 | Chr 6_53741856 | 2.96 (ANOVA CQ) | Oxysterol-binding protein-related protein 1C-like | Intracellular transport. Manages lipid/sterol trafficking and membrane configuration under severe cellular stress |
| EVM0058261 | LG6 | Chr 6_52871733 | 2.71 (Kruskal Class) | Wall-associated receptor kinase 2-like | Cell Wall Integrity Monitoring. Wall-associated kinase (WAK) that senses degradation products from bacterial enzymes and initiates active defenses |



Funded by the European Union

## 3.2. Almond

### 3.2.1 Almond germplasm phenotyping and SNP association

Considerable phenotypic variation was observed among the tested genotypes for both symptoms and bacterial colonization (Fig. 3A,B). While several cultivars displayed high susceptibility, characterized by elevated symptoms incidence and bacterial colonization, other genotypes exhibited substantially lower disease levels. Consistent with previous experiments, the resistant cultivar 'Lauranne' was among the least affected genotypes, whereas UEF was among the most susceptible. Seven 'Lauranne' progenies which are part of the tested germplasm were scattered on the susceptibility scale but mainly on the resistance side (Fig. 3B).

**Table 4.** SNPs significantly associated ($P < 10^{-3}$) to %_Symptoms.

| Trait | Marker | *P* value | marker_$R^2$ |
|---|---|---|---|
| %_Symptom | AP019297.1_1251879 | 1.35E-05 | 0.42139* |
| %_Symptom | AP019297.1_4706870 | 4.09E-04 | 0.31653 |
| %_Symptom | AP019297.1_12430371 | 1.45E-04 | 0.35015 |
| %_Symptom | AP019297.1_13531181 | 4.30E-04 | 0.31483 |
| %_Symptom | AP019297.1_24329665 | 8.82E-04 | 0.24414 |
| %_Symptom | AP019298.1_2711160** | 7.63E-05 | 0.37753 |
| %_Symptom | AP019298.1_2875657 | 3.46E-04 | 0.32209 |
| %_Symptom | AP019298.1_4477456 | 8.07E-04 | 0.24728 |
| %_Symptom | AP019298.1_15369466 | 8.63E-04 | 0.29117 |
| %_Symptom | AP019298.1_17129101 | 3.78E-04 | 0.31913 |
| %_Symptom | AP019300.1_5674099** | 9.84E-04 | 0.23507 |
| %_Symptom | AP019302.1_994313 | 2.60E-04 | 0.3315 |
| %_Symptom | AP019304.1_16019613** | 2.55E-04 | 0.2754 |
| %_Symptom | AP019304.1_17798573 | 9.35E-04 | 0.2319 |
| %_Symptom | AP020417.1_63762 | 2.86E-04 | 0.32831 |

* Significant at the permutation test ($P < 0.05$)
** Associated to %_PCR_positive ($P < 10^{-2}$)

Marker-trait association analysis revealed 15 significant ($P < 10^{-3}$) SNPs associated with symptom severity (Table 4). One marker remained significant following permutation testing ($P < 0.05$) and therefore represented the strongest statistical association detected in the population. The significant associations were located to five chromosomes and one unassembled contig (Fig. 4). One of the regions at AP019298.1 includes two associated markers, AP019298.1_2711160 and AP019298.1_2875657 (164.5 kb; 2,711,160 to 2,875,657), supporting the presence of a resistance-associated locus. In contrast, no significant ($P < 10^{-3}$) associations were detected for the bacterial colonization metric.



Funded by the European Union

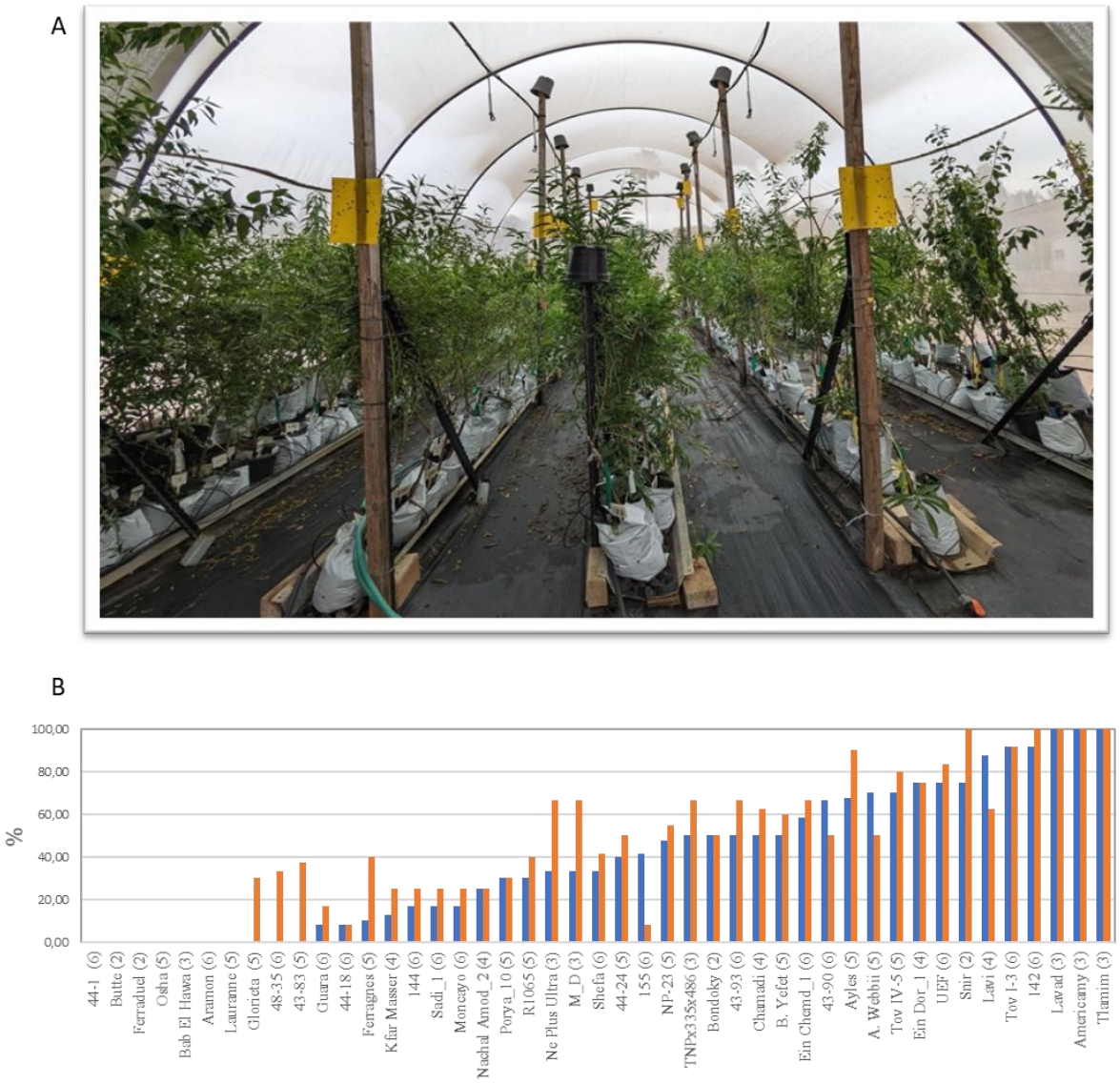



**Figure 3.** Symptoms and bacterial colonization in almond germplasm collection (two-year average). (**A**) Net-house plants. (**B**) Orange bars represent % symptomatic stems; blue bars represent % bacterial colonization. The number of replicates in for each genotype is shown in parenthesis.



Funded by the European Union

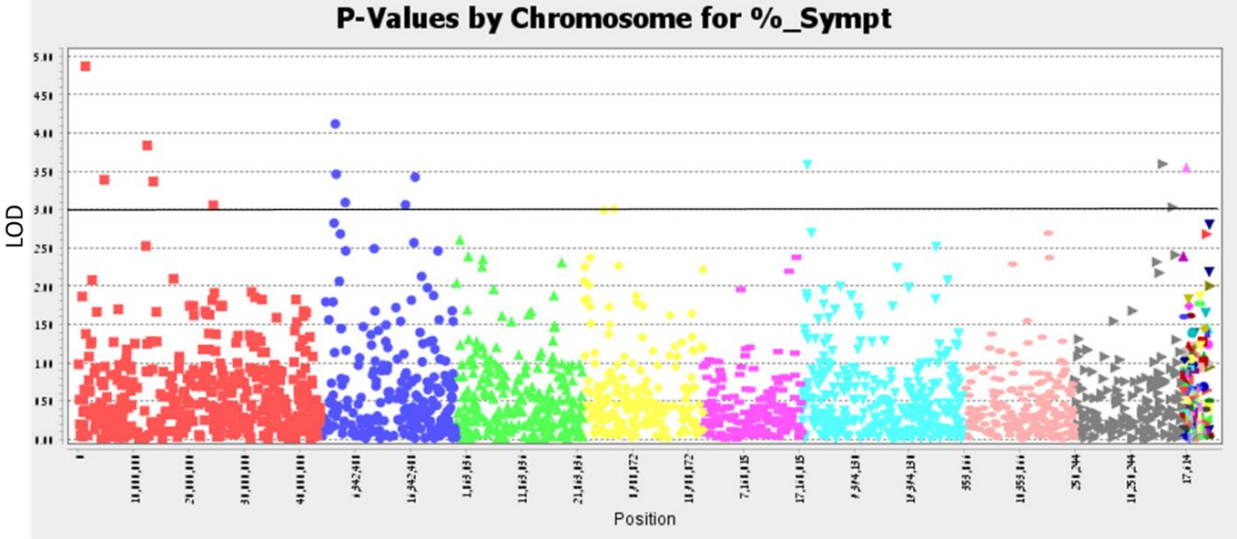



**Figure 4.** Manhattan plot for association to % symptomatic stems. Different colors represent the assembly of genome to eight chromosomes according to 'Lauranne' genome (ASM863291v2). X axis is the position along the genome (ASM863291v2). Dots are the LOD ($-\log_{10}(P)$) value of the association for each marker. The black line marks $-\log_{10}(P) = 3$, the significance threshold.

### 3.2.2 'Lauranne' self-pollinated F2 progeny phenotyping

Phenotypic data confirm a diverse range of responses, from high susceptibility to complete resistance against both *Xf* symptoms and bacterial colonization. When plants were described as either susceptible or resistant, 68% were resistant. This phenotypic segregation suggests that host resistance may be controlled by a single major gene, with resistance demonstrating dominance over susceptibility. The observed segregation did not deviate significantly from the 3:1 ratio expected for a single dominant resistance locus ($\chi^2$ test, $P = 0.289$). In addition, the parent of the segregating population – 'Lauranne', is resistant to *Xf*.

We have almost completed preparation of high-quality DNA for the F2 population and will soon send DNA samples for GBS analysis at an appropriate sequencing facility. When completed, the resulting reads will be aligned to the *P. dulcis* cv. 'Lauranne' reference genome to pinpoint the specific loci associated with the *Xf* resistance trait.



Funded by the European Union

# 4. DISCUSSION

## 4.1 Population curation and high-density linkage map collinearity

The molecular authentication of the 'Leccino' × 'Oliana' F1 mapping population establishes a highly reliable foundation for genetic marker discovery. In genetic studies of perennial crops, validating the pedigree of the progeny is a critical quality control step to prevent mapping errors caused by airborne pollen contamination. By utilizing a panel of simple sequence repeat (SSR) markers, foreign pollen contaminants can be identified and excluded, ensuring the integrity of the mapping panel. In this study, parentage exclusion analysis successfully identified and removed 13 off-type individuals, refining the population to 166 verified F1 hybrids.

The subsequently constructed high-density genetic map, ordered across 23 linkage groups using 73,488 high-quality SPET-derived SNPs, demonstrates excellent positional consistency with the physical olive reference genome. The highly linear and monotonic relationship observed between genetic distances (cM) and physical coordinates (Mb) across almost all chromosomes confirms accurate marker ordering. This robust collinearity provides strong genetic confidence when projecting cM QTL intervals to physical reference positions to search for candidate genes.

## 4.2 Parental segregation dynamics and the oligogenic architecture of *Xylella fastidiosa* resistance

The association analyses performed against quantitative bacterial loads (CQ) and physical symptom classes did not identify any genome-wide significant QTL zones or major single-marker peaks. In quantitative genetics, the lack of major-effect QTLs indicates that host resistance in this pedigree is governed by a complex, polygenic inheritance model composed of multiple loci of small-to-moderate individual effects. This polygenic trait architecture is highly consistent with the segregation patterns of the resistant parent, 'Leccino'.

Wide-scale natural surveys of olive offspring exposed to high inoculum pressure have documented that 'Leccino' acts as a highly effective genetic donor, transmitting traits of resistance (highly resistant, resistant, or tolerant phenotypes) to approximately 67% of its spontaneous progeny (La Notte et al., 2024). However, the inheritance and phenotypic distribution of this resistance are heavily shaped by the genetic background of the susceptible parental partner (La Notte et al., 2024). For instance, crosses between 'Leccino' and highly susceptible cultivars like 'Cellina di Nardò' yield offspring with exceptionally high rates of resistance. In contrast, when crossed with 'Ogliarola salentina', the frequency of resistant phenotypes in the progeny drops significantly to 45% (La Notte et al., 2024).

Because 'Leccino' transmits its resistance traits so highly to its offspring, the phenotypic segregation in the F1 mapping population is heavily skewed toward tolerance. This lack of a classic, balanced phenotypic segregation (susceptible vs. resistant) in the progeny significantly reduces the statistical power of standard QTL mapping, preventing the detection of major single-marker peaks under strict genome-wide corrections (Khan and Korban, 2022; Morales-Cruz et al. 2023; Pettenuzzo et al. 2025).



Funded by the European Union

Furthermore, transcriptomic evidence confirms that 'Leccino' and its resistant offspring do not rely on a single defense pathway; instead, they deploy a coordinated, multi-layered response that integrates:

- Active Cell-Surface Surveillance: Up-regulation of leucine-rich repeat receptor-like kinases (such as SOBIR1 and BAK1) to actively sense extracellular bacterial components (Giampetruzzi et al., 2016; La Notte et al., 2024; Balan et al., 2025).
- Vascular and Osmotic Homeostasis: Maintenance of cellular water relations and turgor pressure under negative physical tension through aquaporins and ion exchangers (Balan et al., 2025).
- Mechanical Wall Reinforcement: Physical sealing of xylem vessels via cinnamyl alcohol dehydrogenase (CAD)-driven lignification and secondary cell wall deposition (such as CESA4 and CESA8 cellulose synthases) to isolate the pathogen (Rodrigues et al., 2013; Giampetruzzi et al., 2016).

This distributed molecular network confirms that resistance is a highly complex, quantitative trait. Because these moderate-effect alleles segregate independently in the F1 hybrids, the resulting phenotypic variation is continuous, which limits the ability of any individual mapping marker to explain a major portion of the phenotypic variance.

## 4.3 Future directions, ongoing genetic mapping in diverse 'Leccino' progenies

To address these segregation limitations and resolve the genetic determinants of 'Leccino'-derived resistance with greater statistical clarity, further mapping initiatives are required. Within other ongoing research projects conducted by CNR, genetic map construction is currently being performed across four additional ‘Leccino’-derived progeny populations.
Evaluating 'Leccino' alleles against these four distinct parental genetic backgrounds will allow us to:

- Expose Altered Segregation Ratios: Crossing 'Leccino' with different partners will alter local allele frequencies and phenotypic segregation in the hybrids, potentially providing the balanced phenotypic variation needed to detect clear, statistically significant QTL regions.
- Identify Consensus QTLs: Overlapping mapping data across five independent crosses will help distinguish pedigree-specific background effects from universal, conserved QTL landmarks governing *Xylella* resistance in olives.

These ongoing multi-population analyses represent a vital step toward validating suggestive QTL intervals and developing high-utility diagnostic markers for marker-assisted breeding.



Funded by the European Union

## 4.4 Genetic variation and mapping of *Xylella fastidiosa* resistance in almond

The screening of almond germplasm revealed substantial variation in response to *Xf*, indicating that useful genetic variation for resistance exists within cultivated and wild almond material. The relatively low symptom development and bacterial colonization observed in 'Lauranne' were consistent with previous experiments and support its classification as a valuable source of *Xf* resistance. Moreover, the distribution of the seven 'Lauranne'-derived progenies, predominantly toward the resistant end of the phenotypic range, suggests that resistance can be transmitted to its offspring and may therefore be exploitable in almond breeding programs.

The association analysis identified several genomic regions potentially involved in resistance to *Xf*-induced symptoms. In particular, the detection of two associated SNPs within a relatively narrow 164.5-kb interval on AP019298.1 strengthens the evidence for a resistance-associated locus in this region. Nevertheless, only one SNP remained significant after permutation testing, indicating that these associations should be considered preliminary and require validation in independent populations. Interestingly, no significant associations were detected for bacterial colonization. This discrepancy may indicate that symptom development and bacterial population size are at least partially controlled by different host mechanisms, or alternatively may reflect greater experimental variability in quantification of bacterial colonization.

The segregation observed in the self-pollinated 'Lauranne' F2 population provides additional evidence that resistance may have a relatively simple genetic basis. Approximately 68% of the progeny were classified as resistant, and the segregation did not differ significantly from that expected for a 3:1 ratio, consistent with resistance being controlled predominantly by a single dominant locus. The occurrence of progeny ranging from highly susceptible to apparently completely resistant also makes this population particularly suitable for genetic mapping. High-density GBS analysis of the F2 population should enable this hypothesis to be tested directly and determine whether the resistance locus co-localizes with regions identified in the germplasm association analysis. Convergence between these complementary mapping approaches would provide strong support for candidate resistance loci and facilitate the subsequent development of molecular markers for *Xf* resistance breeding in almond.

# 5. CONCLUSIONS

Overall, this work highlights the substantial genetic variation for *Xf* resistance available in both olive and almond, while suggesting contrasting genetic architectures underlying resistance in the two crops. In olive, resistance appears to be a complex, quantitative trait requiring multi-population approaches to resolve its genetic determinants, whereas the almond results point toward a potentially simpler genetic basis involving a major resistance locus. The ongoing mapping efforts in both species will be essential for validating resistance-associated loci and translating these findings into robust molecular markers. Ultimately, these resources will support the development and selection of *Xf*-resistant cultivars as a sustainable strategy for mitigating the impact of *Xf* in perennial crops.



Funded by the European Union

# 6. REFERENCES

**Balan et al., 2025**. A systems biology framework integrating cross-species transcriptomics and PPI networks for *Xylella fastidiosa* resistance gene identification. BMC Plant Biology, 25(1), 1062.

**Brukental et al., 2021.** Revealing the genetic components responsible for the unique photosynthetic stem capability of the wild almond *Prunus arabica* (Olivier) Meikle. Frontiers in Plant Science, 12, 779970.

**El Bakkali et al., 2019** Away from the primary domestication centre, the far-west diversification area of the olive tree: towards the management of Moroccan natural resources. Annals of Botany, 138(1), 154-172.

**Giampetruzzi et al., 2016.** Transcriptome profiling of two olive cultivars in response to infection by the CoDiRO strain of *Xylella fastidiosa* subsp. *pauca*. BMC Genomics, 17(1), 475.

**Harper et al., 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.

**Hongyan et al., 2023.** SNP-based high-density linkage map construction and QTL mapping of black spot disease resistance in Chinese sand pear. Journal of Applied Genetics, 64(1), 23-36.

**Huang et al., 2023.** Mapping of QTLs and candidate genes associated with multiple phenotypic traits for Huanglongbing tolerance in citrus. Horticultural Plant Journal, 9(4), 705-719.

**Khan & Korban, 2022**. Breeding and genetics of disease resistance in temperate fruit trees: challenges and new opportunities. Theoretical and Applied Genetics, 135(11), 3961-3985.

**La Notte et al., 2024.** A survey in natural olive resources exposed to high inoculum pressure indicates the presence of traits of resistance to *Xylella fastidiosa* in Leccino offspring. Frontiers in Plant Science, 15, 1457831

**Marshall et al., 1998.** Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology, 7(5), 639-655.

**Morales-Cruz et al., 2023.** Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* spp.) and its implications within a changing climate. Communications Biology, 6(1), 580.

**Morelli et al., 2021.** *Xylella fastidiosa* in olive: A review of control attempts and current management. Microorganisms, 9(8), 1771.

**Pettenuzzo et al., 2025.** QTL mapping and underlying genes for heat tolerance in grapevine (*Rhine Riesling* × *Cabernet Sauvignon*) under field conditions. Theoretical and Applied Genetics, 138(8), 189.

**Pizzi et al., 2026.** Unveiling the genetic blueprint of Verbascoside biosynthesis in *Olea europaea*: An integrative metabolomics, genetics and transcriptomics approach. Horticulture Research, uhag154.

**Rodrigues et al., 2013.** RNA-Seq analysis of Citrus reticulata in the early stages of *Xylella fastidiosa* infection reveals auxin-related genes as a defense response. BMC Genomics, 14(1), 676.



Funded by the European Union

**Skytte af Sätra et al., 2023.** Parametric mapping of QTL for resistance to European canker in apple in 'Aroma'×'Discovery'. Tree Genetics & Genomes, 19(2), 12.

**Trujillo et al., 2014.** Identification of the Worldwide Olive Germplasm Bank of Córdoba (Spain) using SSR and morphological markers. Tree Genetics & Genomes, 10(1), 141-155.

**Valeri et al., 2022.** Exploring olive genetic diversity in the Maltese islands. Sustainability, 14(17), 10684.

**Vanunu et al., 2026.** Identification of an almond cultivar resistant to *Xylella fastidiosa* subsp. *fastidiosa* and insights into host-specific bacterial behavior. Plant Disease (First look). doi: 10.1094/PDIS-02-26-0369-RE.

**Walker et al., 2023.** The impact of xylem geometry on olive cultivar resistance to *Xylella fastidiosa*: an image-based study. Plant Pathology, 72(3), 521-535.

**Zecharia et al., 2022.** *Xylella fastidiosa* outbreak in Israel: population genetics, host range, and temporal and spatial distribution analysis. Phytopathology, 112(11), 2296-2309.



Funded by the European Union