*Diversity* **2010**, *2*, 1205-1222; doi:10.3390/d2111205

OPEN ACCESS

*diversity*

**ISSN 1424-2818**

www.mdpi.com/journal/diversity

*Article*

# Multilocus Sequence Typing Reveals Relevant Genetic Variation and Different Evolutionary Dynamics among Strains of *Xanthomonas arboricola* pv. *juglandis*

**Simone Marcelletti, Patrizia Ferrante and Marco Scortichini ***

Department of Bacteriology, C.R.A.—Research Center for Fruit Trees, Via di Fioranello, 52, 00134 Roma, Italy; E-Mails: simone._82@hotmail.it (S.M.): patferrante@tiscali.it (P.F.)

* Author to whom correspondence should be addressed; E-Mail: marco.scortichini@entecra.it; Tel: +39-0679348102; Fax: +39-0679340158.

*Received: 14 October 2010; in revised form: 11 November 2010 / Accepted: 16 November 2010 / Published: 22 November 2010*

---

**Abstract:** Forty-five *Xanthomonas arboricola* pv. *juglandis* (*Xaj*) strains originating from *Juglans regia* cultivation in different countries were molecularly typed by means of MultiLocus Sequence Typing (MLST), using *acnB*, *gapA*, *gyrB* and *rpoD* gene fragments. A total of 2.5 kilobases was used to infer the phylogenetic relationship among the strains and possible recombination events. Haplotype diversity, linkage disequilibrium analysis, selection tests, gene flow estimates and codon adaptation index were also assessed. The dendrograms built by maximum likelihood with concatenated nucleotide and amino acid sequences revealed two major and two minor phylotypes. The same haplotype was found in strains originating from different continents, and different haplotypes were found in strains isolated in the same year from the same location. A recombination breakpoint was detected within the *rpoD* gene fragment. At the pathovar level, the *Xaj* populations studied here are clonal and under neutral selection. However, four *Xaj* strains isolated from walnut fruits with apical necrosis are under diversifying selection, suggesting a possible new adaptation. Gene flow estimates do not support the hypothesis of geographic isolation of the strains, even though the genetic diversity between the strains increases as the geographic distance between them increases. A triplet deletion, causing the absence of valine, was found in the *rpoD* fragment of all 45 *Xaj* strains when compared with *X. axonopodis* pv. *citri* strain 306. The codon adaptation index was high in all four genes studied, indicating a relevant metabolic activity.

**Keywords:** MLST; *Juglans regia*; haplotype; walnut apical necrosis; *rpoD*; valine; CAI

## 1. Introduction

*Xanthomonas arboricola* pv. *juglandis* (*Xaj*) is the causal agent of Persian (English) walnut blight, a disease causing severe economic losses to *Juglans regia* worldwide. Common symptoms include stem, fruit and leaf spotting (Figure 1) and lesions, catkin necrosis as well as fruit drop.

**Figure 1.** Leaf spotting caused by *Xanthomonas arboricola* pv. *juglandis* on *Juglans regia.*

During the last 15–20 years, an additional severe syndrome, the so–called "apical necrosis" of fruit, has been more frequently observed on *J. regia* cultivars for which *Xaj* plays a fundamental role in causing such a symptom [1]. So far, this syndrome has been reported in Mediterranean countries and California [2,3] (Figure 2).

Previous studies on *Xaj* mainly concerned its cycle of disease and control [3]. Investigations concerning the assessment of *Xaj* genetic variability have been performed using repetitive-sequence PCR (rep-PCR) [4], amplified fragment length polymorphism (AFLP) [5] and integron gene cassette array screening [6] All of these studies showed a high level of genetic diversity within this pathovar, which is partly explained by the geographic origin of the strains. However, strains with different fingerprint patterns, which were isolated in the same country, were also found, as well as strains showing the same fingerprint obtained from different sites [4,5].

Multilocus Sequence Typing (MLST) is a powerful technique for inferring phylogenetic relationships at the interspecific and intraspecific levels [7]. Analysis is based on the sequencing of portions of at least four genes to compare at least 2.0–2.5 kb of the bacterial genome. MLST can target the "core" genome when slowly evolving housekeeping genes encoding for proteins essential for microorganism survival are analyzed. It can also be effectively used to address taxonomic issues, and

when representative strains of a genus are assessed, the technique is commonly called multilocus sequence analysis (MLSA) [8]. Recently, the genus *Xanthomonas* was revised, and new species and pathovars were also described using MLSA [9-11]. Moreover, when MLST is coupled with the appropriate mathematical-statistical tests and algorithms, it can help unravel the evolutionary dynamics (*i.e.*, different strain evolution) that the species, a group of strains or even a single gene can show. In addition, MLST analysis allows investigation of whether recombination or mutation have a major role in determining genetic variability [12]. Establishing different kinds of strain evolution (*i.e.*, diversifying, directional, purifying) is a fundamental prerequisite for better understanding the epidemiology and the cycle of disease of the pathogen [12].

**Figure 2.** Apical necrosis incited by *Xanthomonas arboricola* pv. *juglandis* to *Juglams regia* fruit.

So far, the studies concerning MLST analysis and plant pathogenic bacteria revealed that *Xylella fastidiosa*, the *Pseudomonas syringae* complex and the "core" genome of *Ralstonia solanacearum* have a clonal structure [12-14], whereas in the pathogenicity islands, in *gyrB* and in four other gene fragments of *P. viridiflava*, in *gyrB* and *rpoD* of *P. syringae* pv. *Tomato*, as well as in the virulence-related *egI* and *hrpB* genes of *R. solanacearum*, recombination contributed more than mutation to the genetic variability among strains [12,15,16]. In addition, genes essential for *R. solanacearum* survival were under purifying selection, whereas those involved in pathogenicity as well as the *hrpA* gene encoding the type III system of *P. syringae*, were under diversifying selection [12,17].

By contrast, no study has been performed on strains belonging to the genus *Xanthomonas* using MLST coupled with recombination and selection analysis tests, to unravel the structure and the evolutionary dynamics of populations. We investigated four housekeeping genes of *Xaj* already used to infer phylogenetic relationships within the genus *Xanthomonas*: *acnB*, coding for second aconitase, *gapA* encoding glyceraldeide-3-phosphate dihydrogenase, *gyrB* encoding DNA gyrase, and *rpoD*

encoding the sigma factor 70 [9,10]. The choice of such genes is believed to be fundamental for inferring phylogenetic relationships and evolutionary dynamics because these genes belong to the "core" genome and, consequently, are less subject to horizontal gene transfer.

The aims of this study were to assess (a) the phylogenetic structure of the pathogen; (b) the evolutionary forces that are currently acting upon this microorganism; (c) the possible presence of recombination events within the genes; and (d) the genetic divergence among groups of strains. For these aims, 45 *Xaj* strains obtained from different geographic areas, as well as from different *J. regia* cultivars grown in the same area and isolated at the same time, were used. *Xaj* strains isolated from walnut fruit showing apical necrosis were also included into the analysis.

## 2. Materials and Methods

### *2.1. Bacterial Strains and Growth Conditions*

The 45 *Xaj* strains used in this study are listed in Table 1. Four strains (*i.e.*, WB 252b, WB 301c, NA 0826, isolated in Spain, and CRA-FRU 30, isolated in Italy) were obtained from walnut fruits showing apical necrosis. The *Xaj* strains isolated in Italy in 2008 from different *J. regia* cultivars (*i.e.*, Rome and Caserta) were obtained from fields hosting collections of walnut germplasm. Strain identification was achieved using procedures described elsewhere [4]. Strains were cultured on glucose-yeast extract-calcium carbonate (GYCA) agar plates at 25–27 °C.

**Table 1.** List of *Xanthomonas arboricola* pv. *juglandis* strains used in this study.

| Strain name | Country | Year of isolation | Host-cultivar |
|---|---|---|---|
| CRA-FRU 16 | Italy | 2006 | *Juglans regia* |
| CRA-FRU 19 | Italy | 2006 | *Juglans regia* |
| CRA-FRU 1 | Italy-Caserta | 2008 | *Juglans regia* cv. Malizia |
| CRA-FRU 10 | Italy-Caserta | 2008 | *Juglans regia* cv. Malizia |
| CRA-FRU 44 | Italy-Caserta | 2008 | *Juglans regia* cv. Malizia |
| CRA-FRU 66 | Italy-Rome | 2008 | *Juglans regia* cv. Malizia |
| CRA-FRU 26 | Italy-Caserta | 2008 | *Juglans regia* cv. Sorrento |
| CRA-FRU 30 | Italy-Caserta | 2008 | *Juglans regia* cv. Sorrento |
| CRA-FRU 33 | Italy-Caserta | 2008 | *Juglans regia* cv. Sorrento |
| CRA-FRU 68 | Italy-Rome | 2008 | *Juglans regia* cv. Sorrento |
| CRA-FRU 63 | Italy-Rome | 2008 | *Juglans regia* cv. Srem |
| CRA-FRU 71 | Italy-Rome | 2008 | *Juglans regia* cv. Tisa |
| CRA-FRU 75 | Italy-Rome | 2008 | *Juglans regia* cv. Tisa |
| CRA-FRU 76 | Italy-Rome | 2008 | *Juglans regia* cv. Srem |
| CRA-FRU 77 | Italy-Rome | 2008 | *Juglans regia* cv. Srem |
| CRA-FRU 78 | Italy-Rome | 2008 | *Juglans regia* cv. Medvida |
| CRA-FRU 79 | Italy-Rome | 2008 | *Juglans regia* cv. Champion |
| CRA-FRU 80 | Italy-Rome | 2008 | *Juglans regia* cv. Champion |
| CRA-FRU 81 | Italy-Rome | 2008 | *Juglans regia* cv. Champion |
| CRA-FRU 83 | Italy-Rome | 2008 | *Juglans regia* cv. Baca |
| CRA-FRU 84 | Italy-Rome | 2008 | *Juglans regia* cv. Baca |
| CRA-FRU 85 | Italy-Rome | 2008 | *Juglans regia* cv. Mire |

**Table 1.** *Cont.*

| | | | |
|---|---|---|---|
| CRA-FRU 86 | Italy-Rome | 2008 | *Juglans regia* cv. Mire |
| CRA-FRU 87 | Italy-Rome | 2008 | *Juglans regia* cv. Chico |
| CRA-FRU 88 | Italy-Rome | 2008 | *Juglans regia* cv. Chico |
| CRA-FRU 60 | Turkey | 2008 | *Juglans regia* cv. Kaman 1 |
| CRA-FRU 61 | Turkey | 2008 | *Juglans regia* cv. Kaman 1 |
| NCPPB 411 [T] | New Zealand | 1956 | *Juglans regia* |
| NCPPB 412 | New Zealand | 1956 | *Juglans regia* |
| NCPPB 413 | New Zealand | 1957 | *Juglans regia* |
| NCPPB 362 | United Kingdom | 1955 | *Juglans regia* |
| NCPPB 1659 | United Kingdom | 1964 | *Juglans regia* |
| Aus 76583 | Australia | --- | *Juglans regia* |
| INIA 391 | Portugal | 1994 | *Juglans regia* |
| INIA 392 | Portugal | 1994 | *Juglans regia* |
| INIA 393 | Portugal | 1994 | *Juglans regia* |
| INIA 394 | Portugal | 1994 | *Juglans regia* |
| NCPPB 1447 | Romania | 1962 | *Juglans regia* |
| PD 157 | The Netherlands | 1987 | *Juglans regia* |
| WB 301c | Spain | 2006 | *Juglans regia* cv. Hartley |
| NA 0823 | Spain | 2008 | *Juglans regia* cv. Chandler |
| WB 252b | Spain | 2006 | *Juglans regia* |
| IVIA 884-4 | Spain | 1988 | *Juglans regia* |
| IVIA 884-6 | Spain | 1988 | *Juglans regia* |
| CFBP 7179 | France | 2002 | *Juglans regia* cv. Fernot |

CFBP: Collection Française de Bactéries Phytopathogénes, Angers, France
CRA-FRU: Culture Collection of C.R.A.-Centro di Ricerca per la Frutticoltura, Roma, Italy
DAR: Australian Collection of Plant Pathogenic Bacteria, Rydalmere, Australia
NCPPB: National Collection of Plant Pathogenic Bacteria, York, United Kingdom
INIA: Culture Collection of Instituto Nacional de Investigaciones Agrarias, Oeiras, Portugal
IVIA: Culture Collection of Instituto Valenciano de Investigaciones Agrarias, Moncada-Valencia, Spain
PD: Culture Collection of Plant Protection Service, Wageningen, The Netherlands
WB and NA: Culture Collection of University of Girona, Spain
[T]: type-strain of *Xaj*

### *2.2. PCR and MLST Sequencing*

Gene fragments of *acnB* (second aconitase), *gapA* (glyceraldehyde-3-phosphate dehydrogenase), *gyrB* (subunit B of DNA gyrase) and *rpoD* (sigma subunit of RNA polymerase) were amplified from genomic DNA of the 45 *Xaj* strains. The DNA gene fragments were extracted using the alkaline lysis method. Briefly, a loopful of pure colonies was suspended into Eppendorf tubes containing sterile saline (0.85% of NaCl in distilled water) and mixed with a Vortex. Subsequently, the tubes were centrifuged for 2 min at 10,000 g. Then, the pellet was resuspended in 100 µL of 0.05 M NaOH and heated at 95 °C for 15 min. After centrifugation for 2 min at 10,000 g, the supernatant was used as a DNA template or stored at −20 °C. Gene fragments were amplified and sequenced with primers described by Parkinson *et al*. [9] (*i.e.*, *acnB, gapA* and *rpoD*) and Young *et al*. [10] (*i.e.*, *gyrB*). PCR was carried out in a total volume of 25 µL containing 1 × PCR buffer (10 mM Tris-HCl, 50 mM KCl,

0.1% Triton X-100, pH 9); 50 pmol of each primer; 1.25 U GoTaq® DNA polymerase (Promega); 0.2 mM each dNTP (Promega); 2 mM $MgCl_2$ and 1 μL of 50 ng DNA. All PCR reactions were performed in a BioRad MJ Mini thermal cycler using the following conditions: denaturation at 95 °C for 5 min; 30 s of annealing at 58 °C for *acnB* and *gapA* and at 60 °C for *gyrB* and *rpoD* and extension at 72 °C for 1 min for 35 cycles; and 5 min at 72 °C for the final extension. The PCR reactions were then sent to Primm (Milano, Italy) for sequencing.

### *2.3. Phylogenetics, Amino Acid Deletion and Recombination Analysis*

Sequences were edited with Geneious 4.7.4 (http://www.geneious.com/) and aligned using ClustalW 1.83 (http://www.ebi.ac.uk/tools/clustalw2/). All ambiguous and terminal sequences were edited before data analysis. Each gene fragment was analyzed independently, per each phylotype and collectively at the pathovar level. Each *Xaj* strain was characterized by sequence typing (ST) and allele assignment using the non-redundant database (NRDB) program, available at http://www.mlst.net/. The program PhyML [18] was used to determine the best model of evolution for maximum likelihood (ML) analysis. Both the hierarchical likelihood ratio test (hLRT) and the standard Akaike Information Criterion (AIC) were used to evaluate the model scores. Phylogenetic trees and bootstrap values for the nucleotidic and amino acidic sequences of each gene fragment and for concatenated data, were obtained using the PhyML method by means of TOPALi program version 2.5 [19], available at http://www.topali.org/. A radial phylogenetic tree using the Neighbor-joining (NJ) algorithm [20] was built with Geneious 4.7.4 software. *X. axonopodis* pv. *citri* strain 306 was used as outgroup.

To point out possible amino acid deletions, the four gene fragment sequences were translated into amino acid sequences using the web service, SeWeR, available at http://www.bioinformatics.org/, which translates the nucleotidic sequences into all the six possible readings. All outputs of each gene fragment were analyzed using the Uniprot database (http://www.uniprot.org/) to find out the amino acid deletions. To identify possible conflicting signals within sequence data, which would suggest possible recombination events or acquisition of exogenous genetic material through lateral gene transfer, phylogenetic networks were generated in Splits-Tree 4.6 [21], which is available at http://www.splitstree.org/, using the Neighbor-Net algorithm. Subsequently, for each gene, the web-based service GARD (Genetic Algorithm for Recombination Detection) (http://datamonkey.org/GARD/) was employed to detect and locate recombination breakpoints [22]. In addition, to assess the clonality of phylotypes, the standardized index of association $I^s_A$ was estimated using START2 [23], a web tool available at http://www.pubmlst.org/, to test the null hypothesis of linkage equilibrium for multilocus data [24]. *P* values, calculated using both the parametric and the Monte Carlo methods, were used to discriminate the significance of $I^s_A$.

### *2.4. Genetic Diversity, Selection Analysis, Codon Adaptation Index and Gene Flow*

Both phylogenetic trees and phylogenetic networks indicated four main phylotypes within the *Xaj* strains analyzed. Consequently, we analyzed the genetic diversity and the selection indices between the phylotypes and, when indicated, for each gene fragment or for strains isolated from the same geographic area or plant organ. The number of haplotypes, number and percentage of polymorphic

sites, base frequency and haplotypic diversity (Hd) were calculated using the DnaSP software version 5.10.00 package (http://ub.edu/dnasp/) [25].

The codon adaptation index (CAI), which provides a quantitative measure to assess the synonymous codon bias of genes [26], was calculated using the CAI calculator software (http://genomes.urv.es/CAIcal/). In addition, we also calculated the expected value of CAI (E-CAI) to determine if the CAI values were statistically significant [27]. For each gene fragment and, when applicable, between phylotypes and for the four strains isolated from fruit with apical necrosis, the Tajima's D genetic diversity and population divergence test [28] and the Fu and Li's D tests for neutrality of mutations [29] were estimated using the DnaSP software. Both tests were also assessed for their statistical significance according to Kimura neutral theory [30]. The Mc Donald-Kreitman test between the *Xaj* strains obtained from walnut fruit showing "apical necrosis" symptoms and the other *Xaj* strains was calculated at http://bioinf3.uab.cat/mkt/MKT.asp. Such a test calculates the ratio between nonsynonymous and synonymous polymorphism within different populations of the species (*i.e.*, pathovar), thus indicating evidence for selection in case of non-neutral polymorphism [31,32]. The gene flow estimate between the *Xaj* strains isolated from the same geographic area was assessed by using the fixation index ($F_{ST}$) measuring the population differentiation among strains [33]. The $F_{ST}$ was determined by means of the Arlequin version 3.11 software [34], available at http://cmpg.unibe.ch/software/arlequin3/.

## 3. Results

### *3.1. Genetic and Haplotype Diversity*

The NRDB program was used to assign different allele profiles at each locus among the 45 *Xaj* strains to obtain unique profiles called sequence types (STs). The strains studied have a number of STs that varies from four, for *acnB*, *gapA* and *gyrB,* to seven, for *rpoD*. The highest level of polymorphism was found within the *acnB* and *rpoD* genes (Table 2). The content of guanine plus cytosine is within the range of *X. arboricola*. The number of sequence types (*i.e.*, 12 for 45 strains) found in *Xaj* strains is relevant when compared with those of two other *X. arboricola* pathovars, namely *corylina* and *pruni*, showing nine STs in 40 strains and solely five STs in 47 strains, respectively (data not shown).

**Table 2.** Number of sequence types, polymorphic sites and nucleotide frequencies found in four housekeeping genes of *Xanthomonas arboricola* pv. *Juglandis*.

| Gene | Lenght (bp) | Number of ST | No. of polymorphic sites | % of polymorphic sites | Frequence of A/C/G/T | % G+C |
|---|---|---|---|---|---|---|
| *acnB* | 589 | 4 | 38 | 6.45 | 0.160/0.310/0.356/0.174 | 66.64 |
| *gapA* | 598 | 4 | 16 | 2.67 | 0.214/0.320/0.323/0.143 | 64.30 |
| *gyrB* | 775 | 4 | 18 | 2.32 | 0.182/0.341/0.334/0.143 | 67.50 |
| *rpoD* | 597/594* | 7 | 35 | 5.36 | 0.220/0.316/0.330/0.134 | 64.61 |
| Concatenated | 2,559/2,556* | 12 | 104 | 4.10 | 0.194/0.322/0.335/0.149 | 65.70 |

ST : sequence type ; * three strains have a triplet deletion

### *3.2. Phylogenetic Reconstructions, Amino Acid Sequences and Recombination Analysis*

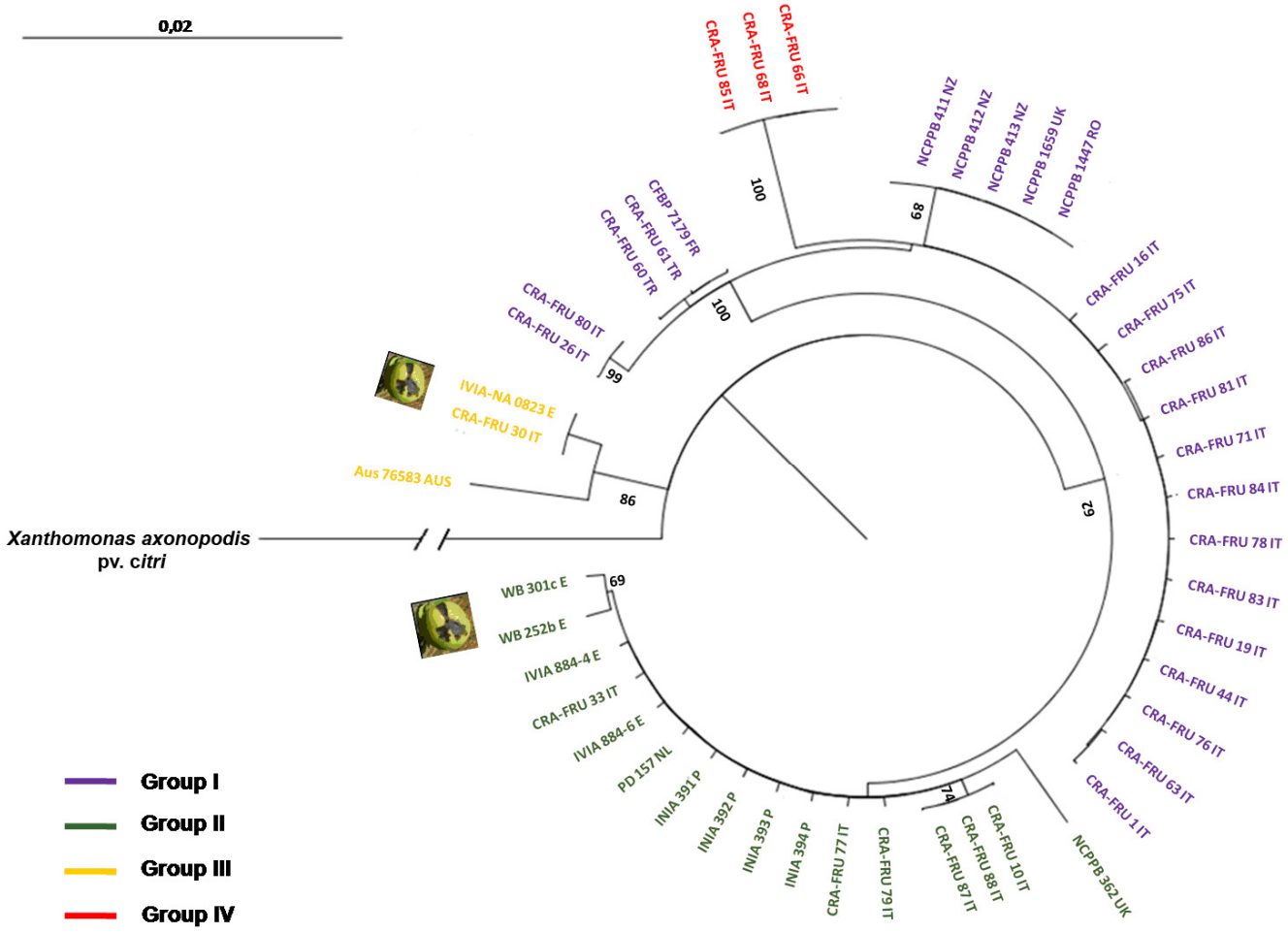
Forty-one *Xaj* strains obtained from different cultivars and geographic areas of walnut cultivation and four strains obtained from walnut fruit apical necrosis were typed at four loci, *acnB*, *gapA*, *gyrB* and *rpoD*. Each locus provided between 589 and 775 bp of common sequence resulting in a total sequence length of 2,556 bp. A ML dendrogram was constructed using either sequences of each locus or concatenated data of the four genes. The best-fit nucleotide substitution model used to infer the ML dendrogram was consistent among the four genes and fit in the GTR (general time-reversible) model. In each case, both the hLRT and AIC found the same model. Either the dendrogram built with ML or the radial dendrogram obtained with NJ algorithms showed the same phylotype distribution among the *Xaj* strains. The radial dendrogram using concatenated data and the NJ algorithm is shown in Figure 3. It reveals two main phylotypes within *Xaj* strains, termed groups I and II. In addition, two other minor phylotypes, named groups III and IV, were also found. Phylotype I contains a subgroup with the type-strain of the pathovar (*i.e.*, NCPPB 411), two others strains from New Zealand, one strain from the United Kingdom and one strain from Romania, all showing the same gene sequences. This phylotype also includes a subgroup of 15 homogenous *Xaj* strains isolated in 2006 and 2008 in Rome and Caserta, Italy, as well as a subgroup with two strains obtained from Turkey and one from France showing the same gene sequences. Phylotype II includes a main subgroup with strains isolated from Italy, Spain, Portugal and the Netherlands all showing the same sequences. This phylotype also contains two *Xaj* strains obtained in Spain from fruit apical necrosis. Phylotype III contains two strains isolated in Italy and Spain from walnut apical necrosis, which showed the same sequences, and one strain from Australia. Phylotype IV includes three *Xaj* strains, isolated in 2008 in Rome, Italy, which showed a triplet deletion in the *rpoD* gene fragment. The strains of this phylotype are the same as phylotype I, except for a mutation within the *rpoD* gene fragment. The sequences diversity of the four *Xaj* phylotypes was relatively high, varying from 0.567 (phylotype I) to 0.667 (phylotype III). The *Xaj* strains isolated in Caserta, Italy, showed the highest level of haplotypic diversity (Table 3).

**Table 3.** Genetic diversity observed within *Xanthomonas arboricola* pv. *juglandis* according to phylogenetic reconstruction (phylotypes) and geographic origin of the strains.

| Phylotype/Origin of strains | Number of sequences | Number of haplotypes | Number of polymorphic site | Haplotype diversity |
|---|---|---|---|---|
| I | 23 | 4 | 16 | 0.636 |
| II | 16 | 5 | 33 | 0.600 |
| III | 3 | 2 | 25 | 0.667 |
| IV | 3 | 1 | 0 | 0.000 |
| Italy | 25 | 6 | 57 | 0.707 |
| Italy (Caserta) | 6 | 5 | 37 | 0.933 |
| Italy (Rome) | 19 | 5 | 40 | 0.649 |
| Spain+Portugal | 9 | 4 | 36 | 0.583 |
| New Zealand | 4 | 1 | 0 | 0.000 |

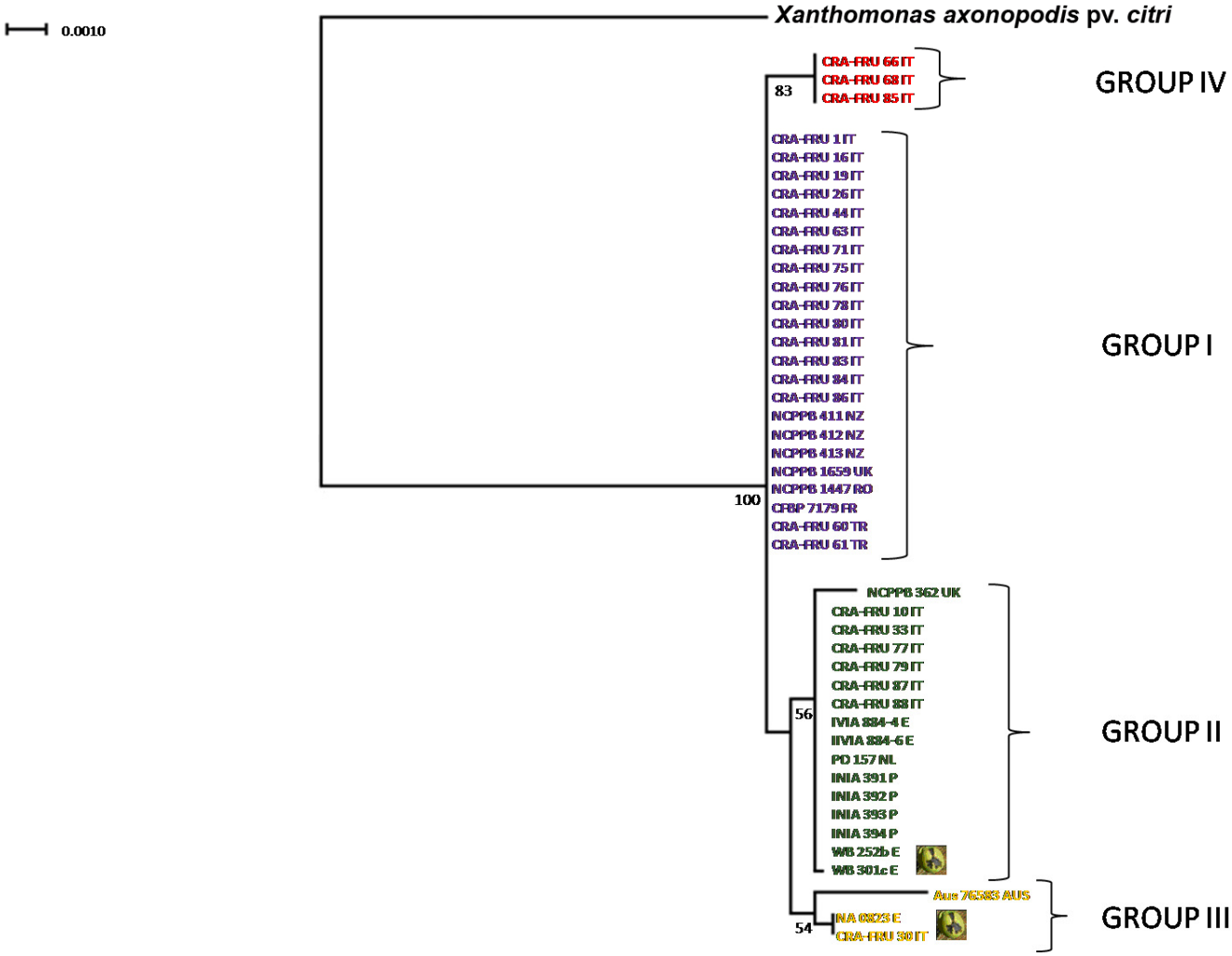
The assessment of the amino acid sequences revealed that all *Xaj* strains show a triplet deletion within the *rpoD* gene fragment, corresponding to valine, when compared with the corresponding *rpoD* sequence of *X. axonopodis* pv. *citri* strain 306. In addition, three *Xaj* strains all isolated in 2008 in Rome, Italy, from leaf lesions of three different walnut cultivars, namely CRA-FRU 66, 68 and 85 (*i.e.*, phylotype 4), show an additional triplet deletion corresponding to arginine.

**Figure 3.** Radial dendrogram using concatenated data and the Neighbor-joining (NJ) algorithm obtained with *acnB*, *gapA*, *gyrB* and *rpoD* nucleotidic gene sequences. Bootstrap values are reported at the main nodes. The *X. a.* pv. *juglandis* strains isolated from apical necrosis of walnut fruits are pointed out.



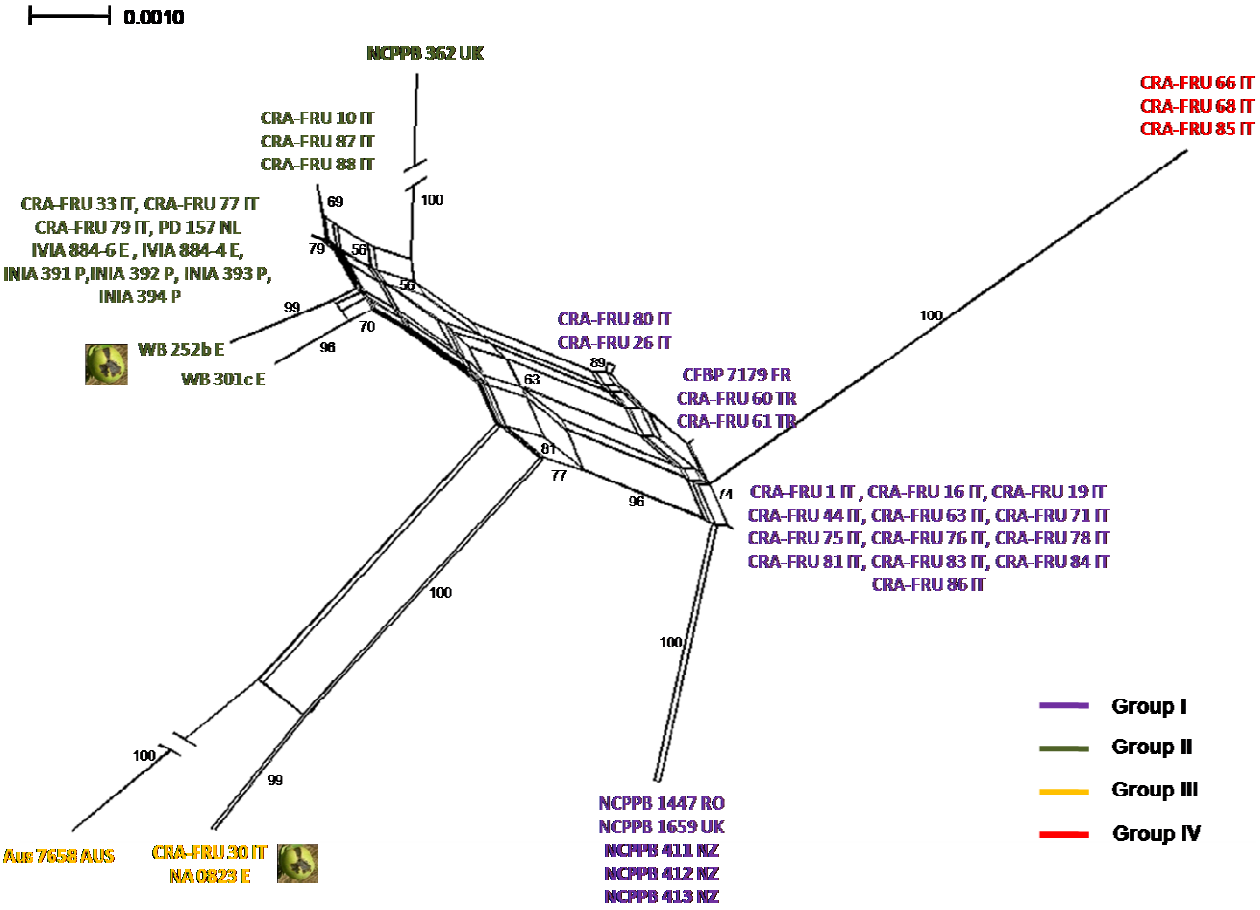

It is noteworthy that the dendrogram built with amino acid sequences identified the same four phylotypes found using the ML algorithm (Figure 4). For representative strains of each phylotype (*i.e.*, NCPPB 411, NCPPB 1659, DAR 76583, CRA-FRU 33, CRA-FRU 61, CRA-FRU 66, CRA-FRU 75, CRA-FRU 80, CRA-FRU 87, NA 0823, WR 252b), the sequences of the four gene fragments were deposited at the NCBI databank. Their accession numbers are the following: *acnB*: from FN667744 to FN667754; *gapA* from FN667755 to FN667765; *gyrB* from FN667766 to FN667776; *rpoD* from FN667777 to FN667787.

**Figure 4.** Dendrogram of relationships among *X. a*. pv. *juglandis* strains obtained by using the amino acid sequences of *acnB*, *gapA*, *gyrB* and *rpoD*. Bootstrap values are reported at the main nodes. The *X. a*. pv. *juglandis* strains isolated from apical necrosis of walnut fruits are pointed out.




Phylogenetic networks were constructed using the Neighbor-Net algorithm to highlight conflicting signals in the gene sequence data, which would suggest the exchange or acquisition of genetic material among *Xaj* strains. In a phylogenetic network, alternative phylogenies are represented by splits or reticulation. By definition, the more reticulation there is in a network, the more conflicting signals exist in the sample, possibly due to exchange of genetic material. The phylogenetic network obtained using the MLST concatenated data is shown in Figure 5. A relevant reticulation was found connecting strains of all phylotypes. This result suggests that possible recombination events can occur within *Xaj* populations.

**Figure 5.** NeighborNet analysis of the concatenated set of four housekeeping genes, *acnB*, *gapA*, *gyrB* and *rpoD* of *X. a.* pv. *juglandis* strains. Bootstrap confidence values are shown along each branch. The *X. a.* pv. *juglandis* strains isolated from apical necrosis of walnut fruits are pointed out.




To further investigate possible recombination events in *Xaj* strains, we subsequently applied the GARD program, a likelihood-based model selection procedure detecting recombination breakpoints in the gene fragments. A breakpoint was predicted within *rpoD* based on the AIC. This finding suggests that the reticulation pointed out with the NeighborNet analysis can be explained with the occurrence recombination events within *Xaj* strains. Table 4 shows the GARD output for all genes.

**Table 4.** Recombination breakpoints predicted in analyzed gene fragment using GARD (Genetic Algorithm for Recombination Detection) and the Akaike Information Criterion (AIC).

| Gene | Number of breakpoints | Lenght of fragment for each partition (nt) | NeighborNet results for partition | Number of detectable breakpoints/Total of breakpoints (P < 0.05) ** |
|---|---|---|---|---|
| *acnB* | 0 | NA | NA | NA |
| *gapA* | 0 | NA | NA | NA |
| *gyrB* | 0 | NA | NA | NA |
| *rpoD* | 1 * | 156* | reticulation | 1/1 |

NA: not applicable; * statistically significant at P = 0.05; ** value represent detectable breakpoints over the number of breakpoints tested

To test the possible role of homologous recombination in *Xaj* strains, we checked the level of linkage disequilibrium in our sample. If the loci are under linkage equilibrium, then high rates of recombination are operating in the population. The standardized index of association $I^S_A$ measures the extent of the linkage, and it is equal or very close to zero when the strains are experiencing free recombination. By contrast, if this value is significantly different from zero, the population is considered clonal (*i.e.*, linkage disequilibrium). The assessment of linkage disequilibrium at the pathovar level indicated that the *Xaj* strains tested here are clonal because the $I^S_A$ value, although relatively low (0.32) was statistically different from 0 (*i.e.*, $V_D$: 1.836 > $L_{MC}$: 1.047).

### *3.2. Selection Analysis and Gene Flow*

The results of the McDonald-Kreitman test are shown in Table 5. This analysis revealed that in *acnB* and *gapA* gene fragments of *Xaj* strains there is a very large amount of non-neutral polymorphism. This data confirm that a possible different selection is acting in the *Xaj* strains isolated from different walnut organs. With the aim of investigating the evolution model that best fit our sample, we also performed the Tajima's D and the Fu and Li's D tests and evaluated their statistical significance. Both tests indicated that all four genes and the *Xaj* strains isolated in Italy, Spain and Portugal are currently under neutral selection. In fact, in each case, the negative values obtained are not statistically significant. By contrast, the four *Xaj* strains obtained from fruit apical necrosis, showed high significant positive values for both the Tajima's and Fu and Li's D tests, indicating diversifying selection (Table 6).

**Table 5.** Neutral and non-neutral polymorphism between *Xanthomonas arboricola* pv. *juglandis* strains isolated from walnut "apical necrosis" of fruits and the other strains according to the Mc Donald-Kreitman test. The test provides the ratio between non-synonymous amino acids and synonymous amino acids.

| | **Neutral** | **Non-neutral** | **Total** |
|---|---|---|---|
| *acn B* | 2 | 52 | 54 |
| *gap A* | 0 | 19 | 19 |
| *gyt B* | 18 | 0 | 18 |
| *rpo D* | 34 | 2 | 36 |

We also investigated the CAI. High CAI values correlate with high levels of gene expression, and average or low CAI values correlate with low levels of gene expression. The CAI value for all of the four genes was high, varying from 0.805 to 0.855, and it was statistically significant, as shown by the corresponding E-CAI values always being lower than the CAI values (Table 6). The gene flow among *Xaj* strains isolated from different geographic areas was estimated using the $F_{ST}$ test (Table 7). This value indicated that the genetic differentiation between strains from different areas is rather low. This does not support the hypothesis of geographic isolation, even though the genetic differentiation between populations increases as the geographic distance between them increases with the strains from New Zealand showing the highest differentiation (Table 7).

**Table 6.** Selection tests, Codon Adaptation Index (CAI) and expected CAI (E-CAI) observed in four *Xanthomonas arboricola* pv. *juglandis* housekeeping gene fragments at gene and at strain geographic origin levels. AN: *X. a.* pv. *juglandis* strains obtained from walnut fruits showing "apical necrosis" symptoms.

| Gene/Origin of strains | Tajima's D | Fu and Li's D | CAI | E-CAI |
|---|---|---|---|---|
| *acnB* | −0.3630* | −0.2351* | 0.805 | 0.743 |
| *gapA* | −1.3869* | −3.9380* | 0.819 | 0.787 |
| *gyrB* | −1.2075* | −1.4734* | 0.855 | 0.789 |
| *rpoD* | −1.1436* | −0.0671* | 0.823 | 0.745 |
| Italy | −0.3346* | −0.1122* | NA | NA |
| Spain+Portugal | −1.6609* | −1.8527* | NA | NA |
| AN strains | 1.5571** | 1.6086** | NA | NA |

* not significant at P = 0.05; ** significant at P = 0.01; NA: Not Applicable.

**Table 7.** Gene flow estimates ($F_{ST}$) among *Xanthomona arboricola* pv. *juglandis* strains isolated from different geographic areas.

| | $F_{ST}$ |
|---|---|
| New Zealand *vs.* Italy | 0.47561 |
| New Zealand *vs.* Spain + Portugal | 0.61021 |
| Italy *vs* Spain + Portugal | 0.34149 |
| Italy (Rome) *vs.* Italy (Caserta) | 0.26806 |

## 4. Discussion

The MLST study performed with four housekeeping genes, namely *acnB*, *gapA*, *gyrB* and *rpoD* pointed out an overall relevant genetic variability within the 45 *Xaj* strains analyzed here, as revealed by the dendrograms built with nucleotidic and amino acidic sequences. By contrast, two other pathovars of *X. arboricola*, namely pvs, *corylina* and *pruni*, assessed using the same technique, appeared considerably more homogeneous (data not shown). This confirms previous studies on *Xaj* performed with other techniques, namely rep-PCR, AFLP and integron cassette arrays, also indicating variation among *Xaj* strains [4-6]. In this study, the phylogenetic reconstruction using ML algorithm indicated two main phylotypes, which also include some subgroups and two other minor phylotypes. In addition, for strains isolated from the same site from different cultivars, the haplotypic diversity can be high as observed for strains obtained in Caserta, Italy. However, some strains isolated either from the same site or from different geographic areas, showed the same gene sequences. These findings suggest that some *Xaj* haplotypes persist over many years and infect walnut trees in different continents, as observed for the phylotype I, which included strains from New Zealand and Europe. In fact, the three strains from New Zealand, including the type strain, isolated in 1956–1957, have the same gene sequences as two other strains isolated in Romania in 1962 and in the United Kingdom in 1964.

However, it is difficult to ascertain the origin of this pathogen because *Xaj* strains isolated from the center of origin of *Juglans regia* (*i.e.*, Central Asia to Eastern China) [35] are difficult to obtain. Our data on gene flow estimate would not support the hypothesis of geographic isolation for *Xaj* strains, even though an increase of strain genetic differentiation associated with an increase in geographic

distance was observed. This feature was also observed by using AFLP [5]. More strain assessment is required to test if *Xaj* is structured on a geographic basis. This hypothesis cannot be ruled out also because in many areas of walnut cultivation in Southern Europe, Minor Asia and Central Asia the production is still largely based on the cultivation of *J. regia* ecotypes, which are well adapted to a typical environment [36], and *Xaj* could adapt accordingly. A geographic structure for phytopathogenic bacteria has been demonstrated for *Pseudomonas avellanae* [37,38] and *Ralstonia solanacearum* [12].

Our study also pointed out possible recombination events within the *rpoD* gene fragment of *Xaj* strains as revealed by the NeighborNet dendrogram and GARD analysis. The occurrence of exchange of genetic material within *Xaj* strains might partly explain the relevant genetic diversity showed by this *X. arboricola* pathovar. The possibility of recombination of chromosomal genes within xanthomonads has already been shown for *X. axonopodis* pv. *vesicatoria* [39]. For this pathogen, the authors found evidence of *in planta* horizontal transfer for genes related to pathogenicity and pigmentation. Interestingly, Yan *et al*. [16] found evidence that in *P. syringae* pv. *tomato, rpoD* and *gyrB* recombination plays a relevant role in the evolution of the pathovar. Moreover, MLST analysis revealed recombination also within *Xylella fastidiosa* subspecies [40].

One interesting finding revealed by the present study is the different kind of selective pressure currently acting on the *Xaj* strains isolated from fruit apical necrosis. In fact, the four strains associated with this pathological syndrome (three from Spain and one from Italy) would appear under diversifying selection, whereas according to Tajima's D and Fu and Li's D tests all the *Xaj* strains isolated in Italy, Spain and Portugal are under neutral selection for all genes tested. In addition, the McDonald-Kreitman test pointed out different selection in gene fragments of the *Xaj* strains. This analysis revealed that a large number of non-neutral polymorphism occurs in *acnB* and *gapA* gene fragments. More detailed studies are necessary to establish which housekeeping gene(s) is relevant for the adaptation of *Xaj* strains. At the pathovar level, *Xaj* is under neutral selection, as expected for populations in equilibrium with their ecological niche. A similar complex evolutionary model was proposed for *X. arboricola* pv. *pruni* strains infecting stone fruit species in areas outside of North America, the center of origin of this pathogen [41]. In the case of the present study, the *X. arboricola* pv. *pruni* strains colonizing new areas in Western Europe were under diversifying selection, suggesting that those strains composed a new and recently introduced population when compared to the older North America populations. In our case, the reasons why the strains obtained from fruit apical necrosis are possibly under a diversifying selection could be explained by the fact that this pathological syndrome is rather new and mainly reported on new walnut cultivar of recent introduction (*i.e.*, not the traditional ecotypes). Consequently, different *Xaj* populations are still adapting to such a niche. Moreover, walnut fruit is chemically very different from walnut leaf. For example, the phenol content in walnut green husks is six times higher than in the leaf [42,43], and such different *milieu* might select a diverse population of the pathogen. However, the assessment of more strains isolated from walnut fruit apical necrosis is necessary to confirm this finding.

Despite the presence of recombination within the *rpoD* gene and the occurrence of strains from apical necrosis under diversifying selection, our study indicated that, at the pathovar level, *Xaj* is clonal and under neutral selection. In neutral evolution, many of the nucleotide changes are selectively neutral, but they contribute to the genetic diversity of the microorganism, as well. This would mean

that, at least for the *Xaj* populations not involved in the fruit apical necrosis, they have reached an equilibrium with the host plant, an indication for an ancient co-evolution with *J. regia*. Another xanthomonad, namely *X. axonopodis* pv. *vesicatoria*, showed a reduced genetic variation that is consistent with a population having undergone a recent bottleneck and/or a selective sweep followed by population expansion [44]. By coalescent analysis, the authors argued that this tomato pathogen is quite young (*i.e.*, 1.432–28 years) and expanded rapidly after the domestication of the tomato to modern agriculture.

In this study we also report a triplet deletion in the *rpoD* gene fragment. For all *Xaj* strains, one triplet deletion in this gene causes the absence of valine in the protein coded by *rpoD*. By contrast, this amino acid is present in *X. axonopodis* pv. *citri* 306. Interestingly, three *Xaj* strains have an additional triplet deletion in *rpoD* that causes the absence of arginine. We could not determine if these deletions alter the function of the sigma subunit of RNA polymerase of *Xaj*.

Finally, we found high CAI values for all the four housekeeping genes. High CAI values predispose the microorganism to be metabolically versatile and facilitate adaptation to new habitats. Therefore, a bacterium can rapidly explore its metabolic potential to adapt to changes in the supply of nutrients or other environmental factors. A high CAI value is ideal for ubiquitous microorganisms living in a niche with a low supply of nutrients that metabolize virtually any carbon source [45]

## 5. Conclusions

*X. arboricola* pv. *juglandis* shows relevant genetic variation as pointed out using MLST performed with four housekeeping genes. Possible recombination events detected within the *rpoD* gene fragment and deletion of a triplet coding for valine in some strains partially explain such a genetic variability. This worldwide spread phytopathogen has some haplotypes which persists over the years in different continents. The pathogen appears clonal at the pathovar level. However, different evolutionary dynamics were pointed out among strains isolated from different walnut tissues. The strains obtained from apical necrosis of fruits, a recently observed pathological syndrome, would seem under diversifying selection, whereas all the other are under neutral selection.

## Acknowledgements

The authors wish to thank Maria M. Lopez (Instituto Valenciano de Investigaciones Agrarias, Moncada-Valencia, Spain) and Concepciò Moragrega (University of Girona, Spain) for supplying the *Xaj* strains from Spain.

## References

1. Moragrega, C.; Ozaktan, H. Apical necrosis of Persian (English) walnut (*Juglans regia*): An update. *J. Pl. Pathol*. **2010**, *92*, 67-71.
2. Garcin, A.; Duchesne, D. Walnut blight and apical necrosis. *Acta Horticulturae* **2001**, *544*, 379-387.

3. Buchner, R.P.; Adaskaveg, J.E.; Olson, W.H.; Lindow, S.E. Walnut blight (*Xanthomonas campestris* pv. *juglandis*) control. Investigations in northern California, USA. *Acta Horticulturae* **2001**, *544*, 369-378.
4. Scortichini, M.; Marchesi, U.; Di Prospero, P. Genetic diversity of *Xanthomonas arboricola* pv. *juglandis* (synonyms: *X. campestris* pv. *juglandis*; *X. juglandis* pv. *juglandis*) strains from different geographical areas shown by repetitive polymerase chain reaction genomic fingerprinting. *J. Phytopathol*. **2001**, *149*, 325-332.
5. Loreti, S.; Gallelli, A.; Belisario, A.; Wajnberg, E.; Corazza, L. Investigation of genomic variability of *Xanthomonas arboricola* pv. *juglandis* by AFLP analysis. *Eur. J. Pl. Pathol*. **2001**, *107*, 583-591.
6. Barionovi, D.; Scortichini. M. Integron variability in *Xanthomonas arboricola* pv. *juglandis* and *Xanthomonas arboricola* pv. *pruni* strains. *FEMS Microbiol. Lett*. **2008**, *288*, 19-24.
7. Maiden, K.J.C.; Bygraves, J.A.; Feil, E.; Morelli, G.; Russell, J.E.; Unwin, R.; Zhou, J.; Zurth, K.; Caugant, D.A.; Feavers, I.M.; Achtman, M.; Spratt, B.G. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3140-3145.
8. Gevers, D.; Cohan, F.M.; Lawrence, J.G.; Spratt, B.J.; Coenye, T.; Feil, E.J.; Stackebrandt, E.; van de Peer, Y.; Vandamme, P.; Thompson, F.L.; Swings, J. Re-evaluating prokaryotes species. *Nat. Rev. Microbiol.* **2005**, *3*, 733-739.
9. Parkinson, N.; Aritua, V.; Heeney, J.; Cowie, C.; Bew, J.; Stead, D. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. *Int. J. Syst. Evol. Microbiol*. **2007**, *57*, 2881-2887.
10. Young, J.M.; Park, D-C.; Shearman, H.M.; Fargier, F. A multilocus sequence analysis of the genus *Xanthomonas*. *Syst. Appl. Microbiol*. **2008**, *31*, 366-377.
11. Young, J.M.; Wilkie, J.P.; Park, D-C.; Watson, D.R.W. New Zealand strains of plant pathogenic bacteria classified by multi-locus sequence analysis; proposal of *Xanthomonas dyei* sp. nov. *Plant Pathol*. **2009**, *59*, 270-281.
12. Castillo, J.A.; Greenberg, J.T. Evolutionary dynamics of *Ralstonia solanacearum*. *Appl. Environ. Microbiol.* **2007**, *73*, 1225-1238.
13. Sarkar, S.F.; Guttman, D.S. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl. Environ. Microbiol*. **2004**, *70*, 1999-2012.
14. Scally, M.; Schuenzel, E.L.; Stouthamer, R.; Nunney, L. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Appl. Environ. Microbiol*. **2005**, *71*, 8491-8499.
15. Goss, E.M.; Kreitman, M.; Bergelson, J. Genetic diversity, recombination and cryptic genes in *Pseudomonas viridiflava* infecting natural populations of *Arabidopsis thaliana*. *Genetics* **2005**, *139*, 21-35.
16. Yan, S.; Liu, H.; Mohr, T.J.; Jenrette, J.; Chiodini, R.; Zaccardelli, M.; Setubal, J.C., Vinatzer, B.A. The role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. tomato DC3000, a very atypical tomato strain. *Appl. Environ. Microbiol*. **2008**, *74*, 3171-3181.

17. Guttman, D.S.; Gropp, S.L.; Morgan, R.L.; Wang, P.W. Diversifying selection drives the evolution of the type III secretion system pilus of *Pseudomonas syringae. Mol. Biol. Evol.* **2006**, *23*, 2342-2354.
18. Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate phylogenies by maximum likelihood. *Syst. Biol*. **2003**, *52*, 696-704.
19. Milne, I.; Lindner, D.; Bayer, M.; Husmeier, D.; McGuire, G.; Marshall, D.F.; Wright, F. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* **2008**, *24*, 126-127.
20. Huson, D.H.; Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **2006**, *23*, 254-267.
21. Bryant, D.; Moulton, V. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol*. **2004**, *21*, 258-265.
22. Kosakovski Pond, S.L.; Posada, D.; Gravenor, M.B.; Wall, C.H.; Frost, S.D. Automated phylogenetic: detection of recombination using a genetic algorithm. *Mol. Biol. Evol*. **2006**, *23*, 1891-1901.
23. Jolley, K.A.; Feil, E.J.; Chan, M.S.; Maiden, M.C. Sequence type analysis and recombinational tests (START). *Bioinformatics* **2001**, *17*, 1230-1231.
24. Haubold, B.; Travisano, M.; Rainey, P.B.; Hudson, R.R. Detecting linkage disequilibrium in bacterial populations. *Genetics* **1998**, *150*, 1341-1348.
25. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451-1452.
26. Sharp, P.M.; Li, W-H. The codon adaptation index: a measure of directional synonymous codon usage bias, and its potential application. *Nucleic Acid Res*. **1987**, *15*, 1281-1295.
27. Puigbò, P.; Bravo, I.G.; Garcia-Vallvé, S. E-CAI: a novel server to estimate an expected value of codon adaptation index (eCAI). *BMC Bioinformatics* **2008**, *9*, 65.
28. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585-598.
29. Fu, Y.X.; Li, W.H. Statistical tests of neutrality of mutation. *Genetics* **1993**, *133*, 693-709.
30. Kimura, M. Retrospective of the last quarter century of the neutral theory. *Japan. J. Genetics* **1993**, *68*, 521-528.
31. Mc Donald, J.H.; Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila. Nature* **1991**, *351*, 652-654.
32. Nielsen, R. Statistical tests of selective neutrality in the age of genomics. *Heredity*, **2001**, *86*, 641-647.
33. Wright, S. The genetical structure of populations. *Ann. Eugenics* **1951**, *15*, 323-354.
34. Excoffier, L.; Laval, G.; Schneider, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **2005**, *1*, 47-50.
35. Aradhya, K.; Potter, D.; Gao, F.; Simon, C.J. Molecular phylogeny of *Juglans* (*Juglandaceae*): a biogeography perspective. *Tree Genetics & Genomes* **2007**, *3*, 363-378.

36. Fornari, B.; Cannata, F.; Spada, M.; Malvolti, M.E. Allozyme analysis of genetic diversity and differentiation in European and Asiatic walnut (*Juglans regia* L.) populations. *For. Genetics* **1999**, *6*, 115-127.
37. Scortichini, M.; Natalini, E.; Marchesi, U. Evidence for separate origins of the two *Pseudomonas avellanae* lineages. *Plant Pathol*. **2006**, *55*, 451-457.
38. Wang, P.W.; Morgan, R.L.; Scortichini M.; Guttman, D.S. Convergent evolution of phytopathogenic pseudomonads onto hazelnut. *Microbiol*. **2007**, *153*, 2067-2073.
39. Basim, H.; Stall, R.E.; Minsavage, G.V.; Jones, J.B. Chromosomal gene transfer by conjugation in the plant pathogen *Xanthomonas axonopodis* pv. *vesicatoria. Phytopathol.* **1999**, *89*, 1044-1049.
40. Yuan, X.; Morano, L.; Bromley, R.; Spring-Pearson, S.; Stouthamer, R.; Nunney, L. Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. *Phytopathol*. **2010**, *100*, 601-611.
41. Boudon, S.; Manceau, C.; Notteghem, J.L. Structure and origin of *Xanthomonas arboricola* pv. *pruni* populations causing bacterial spot of stone fruits in western Europe. *Phytopathol.* **2005**, *95*, 1081-1088.
42. Pereira, J.A.; Oliveira, I.; Sousa, A.; Valentao, P.; Andrade, P.B.; Ferreira, I.C.F.R.; Ferrerei, F.; Bento, A.; Seabra, R.; Estevinho, L. Walnut (*Juglans regia* L.) leaves: phenolic compounds, antibacterial activity and antioxidant potential of different cultivars. *Food Chem. Toxicol*. **2007**, *45*, 2287-2295.
43. Oliveira, I.; Sousa, A.; Ferreira, I.C.F.R.; Bento, A.; Estevinho, L.; Pereira, J.A. Total phenols, antioxidant potential and antimicrobial activity of walnut (*Juglans regia* L.) green husks. *Food Chem. Toxicol.* **2008**, *46*, 2326-2331.
44. Wichmann, G.; Ritchie, D.; Kousik, C.S.; Bergelson, J. Reduced genetic variation occurs among genes of the high clonal plant pathogen *Xanthomonas axonopodis* pv. vesicatoria, including the effector gene *avrBs2*. *Appl. Environ. Microbiol*. **2005**, *71*, 2418-2432.
45. Kiewitz, C.; Tummler, B. Sequence diversity of *Pseudomonas aeruginosa*: impact on population structure and genome evolution. *J. Bacteriol*. **2000**, *182*, 3125-3135.