Arch Microbiol
DOI 10.1007/s00203-016-1245-1

ORIGINAL PAPER

Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new *Xylella* species

Simone Marcelletti[1] · **Marco Scortichini**[1,2]

Received: 4 March 2016 / Revised: 13 April 2016 / Accepted: 16 May 2016

Abstract A total of 21 *Xylella fastidiosa* strains were assessed by comparing their genomes to infer their taxonomic relationships. The whole-genome-based average nucleotide identity and tetranucleotide frequency correlation coefficient analyses were performed. In addition, a consensus tree based on comparisons of 956 core gene families, and a genome-wide phylogenetic tree and a Neighbor-net network were constructed with 820,088 nucleotides (i.e., approximately 30–33 % of the entire *X. fastidiosa* genome). All approaches revealed the occurrence of three well-demarcated genetic clusters that represent *X. fastidiosa* subspecies *fastidiosa*, *multiplex* and *pauca*, with the latter appeared to diverge. We suggest that the proposed but never formally described subspecies '*sandyi*' and '*morus*' are instead members of the subspecies *fastidiosa*. These analyses support the view that the *Xylella* strain isolated from *Pyrus pyrifolia* in Taiwan is likely to be a new species. A widely used multilocus sequence typing analysis yielded conflicting results.

Communicated by Erko Stackebrandt.

Electronic supplementary material The online version of this article (doi:10.1007/s00203-016-1245-1) contains supplementary material, which is available to authorized users.

✉ Marco Scortichini
marco.scortichini@crea.gov.it

[1] Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria – Centro di ricerca per le Colture Arboree, Via di Fioranello, 52, 00134 Rome, Italy

[2] Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria – Centro di ricerca per le Colture Arboree, Via Torrino, 3, 81100 Caserta, Italy

Keywords Average nucleotide identity (ANI) · Genome-wide phylogenetic tree · MLST · Subspecies · Xylem-restricted pathogens

Introduction

Xylella fastidiosa Wells et al. (1987) is a xylem-limited, fastidious phytopathogenic bacterium of the *Xanthomonadaceae* family which colonizes a very large number of hosts, including both cultivated and wild plant species. A recent review reported 309 plant species belonging to 63 families as hosts of this species group (EFSA 2015). *Xylella fastidiosa* is transmitted by several xylem sap-feeding insect vectors that primarily belong to the *Hemiptera* order and *Cicadellidae*, *Cercopidae* and *Machaerotidae* families (Purcell 2008). The European Plant Protection Organization (EPPO) includes *Xylella fastidiosa* in the A1 list of regulated pest pathogens. Based on a DNA–DNA hybridization (DDH) relatedness assay and the sequencing of 16S-23S rRNA intergenic transcribed spacers (ITS), Schaad et al. (2004a) described and formally proposed the occurrence of three subspecies: *X. fastidiosa* subsp. *piercei*, which was later emended into subsp. *fastidiosa* (Schaad et al. 2004b); *X. fastidiosa* subsp. *multiplex*; and *X. fastidiosa* subsp. *pauca*. However, it should be noted that for the subspecies *pauca* the strain that is claimed to be representative of the taxon, namely ICMP 15198, has not been deposited in two publicly accessible services collections in two countries (Bull et al. 2010, 2012) as required by rule 30 of the International Code of Nomenclature of Bacteria (Lapage et al. 1992). Subsequently, based on multilocus sequence typing (MLST) analyses of seven housekeeping genes, two additional subspecies that are restricted to host plants from which the pathogen was

Published online: 21 May 2016

originally isolated, i.e., *X. f.* subsp. '*sandyi*' isolated from *Nerium oleander* and *X. f.* subsp. '*morus*' isolated from *Morus alba*, have also been proposed but not formally described (Scally et al. 2005; Nunney et al. 2014; Almeida and Nunney 2015).

Recently, a *X. fastidiosa* strain belonging to the subsp. *pauca*, namely CoDiRO, was found to be associated with the 'olive quick decline syndrome' and is causing severe damage (i.e., twig and branch wilting and plant death) in cultivated olive (*Olea europea*) trees in the province of Lecce (Apulia region, Southern Italy) (Loconsole et al. 2014; Cariddi et al. 2014). Additionally, *X. fastidiosa* subsp. *multiplex* has recently been reported to be associated with the decline of the ornamental shrub *Polygala myrtifolia* in some locations of southern France (i.e., Southern Corsica and the Maritime Alps) (EPPO 2015). The rapid emergence of this regulated phytopathogen in Southern Europe has not previously been reported, and this imposes the application of rapid and reliable detection techniques to prevent the further introduction of latently infected plant material. Within this framework, knowledge of the basic taxonomy and population structure of the pathogen is fundamental for the development of an efficient detection and prevention protocol.

The incorporation of genomic data into bacterial taxonomies and systematic procedures has recently greatly contributed to the advancement of such disciplines (Thompson et al. 2013, 2015; Chun and Rainey 2014; Kim et al. 2014; Ramasamy et al. 2014). Regarding phytopathogenic bacteria, this approach has contributed to the clarification of the taxonomic relationships among phytopathogenic *Pseudomonas* species (Scortichini et al. 2013; Marcelletti and Scortichini 2014, 2015). In the present study, we retrieved complete and draft genomes of 21 *X. fastidiosa* strains isolated from 13 plant species that are cultivated on all the continents in which the pathogen has been reported. To infer the taxonomic relationships of these strains using these data, we performed the following analyses: a) a genome-wide phylogenetic tree (Lerat et al. 2003) and Neighbor-net network analysis (Huson and Bryant 2006) using 820,088 nucleotides (i.e., approximately 30–33 % of the entire genome); b) a consensus tree was created with 956 gene sequence alignments of the *X. fastidiosa* core genome (Huson and Bryant 2006); c) genome-wide-based average nucleotide identity (ANI) and tetranucleotide frequency correlation coefficient (TETRA) analyses (Konstantinidis and Tiedje 2005; Goris et al. 2007); and d) an MLST analysis (Maiden et al. 1998) with seven housekeeping genes that are currently used to type *X. fastidiosa* strains.

Materials and methods

Genome-wide phylogenetic and consensus trees and split Neighbor-net network

A data set containing ortholog alignment was prepared using a multistep procedure based on several ad hoc PERL scripts. First, the predicted protein sequences of all genomes were analyzed for the identification superfamilies of homologs by a procedure based on reciprocal smallest distance algorithm (Wall et al. 2003). Subsequent application of the branch clustering algorithm BranchClust (Poptsova and Gogarten 2007) allowed delineation of families of orthologs within superfamilies containing one or more paralogous gene families. The gene families selection of the *X. fastidiosa* core genome was obtained as described by Marcelletti and Scortichini (2015). For the *X. fastidiosa* strains CFBP 8072 and 8073 (Jacques et al. 2015), annotation was performed by the RAST server (http://rast.nmpdr.org/) (Aziz et al. 2008), which utilized GeneMark, Glimmer and tRNAscan-SE searches. In total, 956 gene sequence alignments, spanning 820,088 nucleotide sites, were selected for building the phylogenetic and the consensus trees. For the latter, the trees from each individual DNA sequence alignment were obtained by recursively running PhyML using LC as a substitution model and nearest neighbor interchange (NNI) for the tree topology estimate. From the 956 gene sequence alignment ML trees, a consensus tree was obtained with Splits Tree4 software (http//www.splitstree.org), using a mean network construction (Huson and Bryant 2006). In addition, with the 820,088 nucleotides, a Neighbor-net network was obtained with Splits Tree4 software (Huson and Bryant 2006). The gene sequences were concatenated to obtain a single large alignment that was submitted to Neighbor-net network analysis with Splits Tree4 using the neighbor-joining (NJ) algorithm with the Hamming distance method. Bootstrap analysis with 1.000 replications was performed by using the same software. *X. o.* pv. *oryzae* KACC 10331 was used as outgroup.

Average nucleotide identity (ANI) and tetranucleotide frequency correlation coefficients (TETRA)

To further evaluate the taxonomic relationships of 21 strains, previously classified as *Xylella fastidiosa* (Table 1), the whole-genome-based average nucleotide identity (ANI) and the tetranucleotide frequency correlation coefficients (TETRA) analyses were performed (Konstantinidis and Tiedje 2005; Goris et al. 2007). The genome data from *X. fastidiosa* strains were downloaded from the

Table 1 List of *Xylella fastidiosa* strains used in this study, their strain code, genome accession number and G+C percentage content

Strain	Host plant	Geographical origin	Accession number	%GC	Genome length
ATCC 35879* =DSM 10026 =LMG 17159 =ICMP 15177	Vitis vinifera	USA (Florida)	JQAP01000000 (WGS Project)	51.8	2,522,328
Temecula 1	Vitis vinifera	USA (California)	GCA_000007245.1 (GenBank)	51.8	2,521,148
GB514	Vitis vinifera	USA (Texas)	GCA_000148405.1 (GenBank)	51.8	2,517,383
Dixon =ATCC 700965	Prunus amygdalus	USA (California)	AAAL02000000 (WGS project)	52.0	2,622,287
M12	Prunus amygdalus	USA (California)	GCA_000019325.1 (GenBank)	51.9	2,475,130
M23	Prunus amygdalus	USA (California)	GCA_000019765.1 (GenBank)	51.7	2,573,987
MUL0034	Morus alba	USA (California)	GCA_000698825.1 (GenBank)	51.9	2,666,577
Mul-MD	Morus alba	USA (Maryland)	AXDP00000000 (WGS Project)	51.6	2,520,555
Griffin-1	Quercus rubra	USA (Georgia)	AVGA00000000 (WGS Project)	51.7	2,387,314
EB92-1	Sambucus canadensis	USA (Florida)	AFDJ01000000 (WGS Project)	51.5	2,475,426
Sy-Va	Platanus occidentalis	USA (Virginia)	JMHP00000000 (WGS Project)	51.6	2,475,880
Ann_1	Nerium oleander	USA (California)	GCA_000698805.1 (GenBank)	52.1	2,780,908
ATCC35871* =LMG 9063 =ICMP 15199	Prunus salicina	USA (Georgia)	AUAJ01000000 (WGS Project)	51.7	2,413,091
9a5c	Citrus sinensis	Brazil (Sao Paulo)	GCA_000006725.1 (GenBank)	52.6	2,731,748
32	Coffea arabica	Brazil (Sao Paulo)	AWYH00000000 (WGS Project)	52.5	2,607,546
6c	Coffea arabica	Brazil (Sao Paulo)	AXBS00000000 (WGS Project)	52.4	2,606,002
CO33	Coffea arabica	Costa Rica—Italy	LJZW00000000 (WGS Project)	51.7	2,681,926
CFBP 8072	Coffea arabica	Ecuador—France	LKDK00000000 (WGS Project)	51.9	2,496,662
CFBP 8073	Coffea canephora	Mexico—France	LKES00000000 (WGS Project)	51.6	2,582,150
CoDiRO	Olea europea	Italy (Apulia)	JUJW01000000 (WGS Project)	52.0	2,542,932
PLS 229	Pyrus pyrifolia	Taiwan	JDSQ00000000 (WGS Project)	53.0	2,733,013

The strains CO33, CFBP 8072 and CFBP 8073 have been intercepted in Italy and France, respectively, from *Coffea* spp. cultivated in Central America

* Subspecies type strain

National Center of Biotechnology Information (NCBI) (see Table 1). The analyses of sequences for the determination of their relatedness according to ANI and TETRA were performed with the software JSpecies (Richter and Rosselló-Móra 2009). ANI was calculated using algorithms obtained with the data structure named suffix tree and the BLASTN software (Altschul et al. 1997). TETRA was used as an alignment-free genomic similarity index as oligonucleotide frequencies carry a species-specific signal (Richter and Rosselló-Móra 2009). *Xanthomonas oryzae* pv. *oryzae* KACC 10331 was used as outgroup.

Multilocus sequence typing (MLST) analysis and recombination assessment

To compare the results obtained by means of the genome-wide assessments with a commonly used approach to type *X. fastidiosa* strains, a multilocus sequence typing (MLST) analysis was also carried out. The maximum likelihood (ML) analysis was performed with nucleotide sequences using seven housekeeping genes (*gltT, holC, lacF, leuA, nuoL, petC, rfbD*) commonly used for the *X. fastidiosa* typing (Scally et al. 2005), for a total of 6021 nucleotides. ML analysis was inferred with PhyML version 3.0 (Guindon and Gascuel 2003), with 1000 bootstrap replicates. To select the best-fit model for ML, we used a phymltest procedure implemented in the R package APE (Paradis et al. 2004), and JC69 was used as best nucleotide model for nucleotide. The tree was visualized using FigTree software, version 1.1.2 (http://tree.bio.ed.ac.uk/software/figtree/). *X. o.* pv. *oryzae* KACC 10331 was used as outgroup. To investigate homologous recombination within *X. fastidiosa* strains, with the seven housekeeping genes assessed by means of MLST, a recombination network analysis was performed by using Splits Tree4 software. In such evolutionary networks, reticulation indicates possible events of recombination among strains (Huson and Bryant 2006).

Results

Genome-wide phylogenetic and consensus trees, and split Neighbor-net network

The genome-wide phylogenetic tree (Fig. 1), the Neighbor-net network (Fig. 2), both constructed with 820,088 nucleotides, and the consensus tree with a cutoff value of 0.30 and based on 956 trees retrieved from the core genome of *X. fastidiosa* (Fig. 3) clearly revealed the robust occurrence of three distinct genetic clusters within *X. fastidiosa* that corresponded to subspecies *fastidiosa*, *multiplex* and *pauca.* Additionally, an evident divergence of the strain PLS 229, which was isolated in Taiwan from *P. pyrifolia*, was observed. Within the subspecies *fastidiosa*, the three strains isolated from *Vitis vinifera* and the strains obtained from *Sambucus canadensis* and *Prunus amygdalus* exhibited lower genetic divergence compared with the other strains of this subspecies. Interestingly, both the trees and the Neighbor-net network suggest that the five strains belonging to the subspecies *pauca* appeared to have diverged from the *X. fastidiosa* strains of the other two subspecies (Figs. 1, 2, 3).

Average nucleotide identity (ANI) and tetranucleotide frequency correlation coefficients (TETRA), GC content

The 21 *X. fastidiosa* strains were also cross-compared to determine their sequence similarities. The ANI value calculations and the TETRA analyses are shown in Fig. 4 and Tables S1 and S2. With the exception of PLS 229, the other 20 strains exhibited ANI values that ranged from 95.69 to 99.98 %. The five strains classified as *pauca* as inferred by the genome-wide phylogenetic and consensus tree approaches, exhibited ANI values between 95.69 and 96.63 % when compared with subspecies *fastidiosa* and *multiplex*. The ANI values for the subspecies *fastidiosa* and *multiplex* ranged from 97.82 to 99.93 %, and from 99.29 to 99.95 %, respectively. The GC content (i.e., percentage) of the three *X. fastidiosa* subspecies showed different ranges: from 51.5 to 52.1 % in subsp. *fastidiosa*; from 51.6 to 52.0 % in subsp. *multiplex*; and from 51.9 to 52.6 % in subsp. *pauca*. Notably, the ANI values for PLS 229 relative to the other strains ranged from 83.56 to 83.96 %. This strain also exhibited a GC content that was higher (i.e., 53.0 %) than those of the other *X. fastidiosa* strains (Table 1).

MLST and recombination analyses

The MLST analysis revealed the occurrence of genetic clusters that partially corresponded to the three subspecies, also highlighting the distinct nature of PLS 222 (Fig. S1). Two strains isolated from *Morus alba* that were included in the *fastidiosa* subspecies according to the genome-wide analyses were placed in the *multiplex* subspecies as diverging cluster based on the MLST analysis. The recombination networks performed with seven housekeeping genes

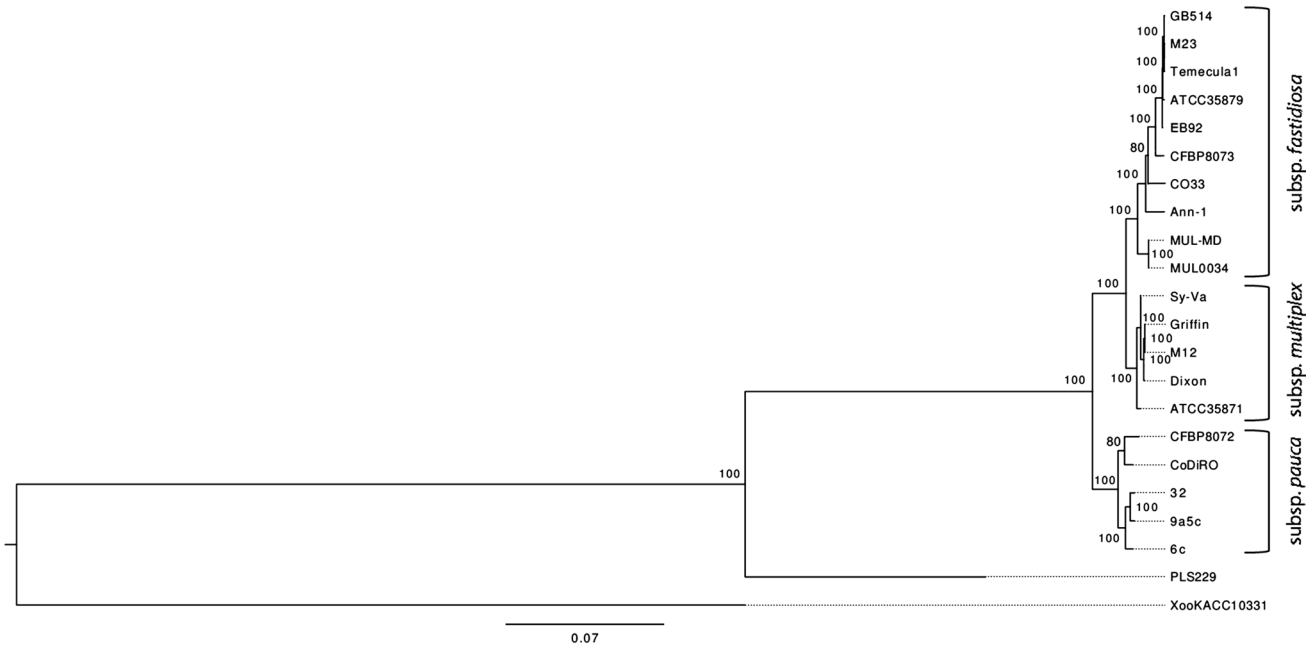

Fig. 1 Genome-wide phylogenetic tree as obtained from 820.088 nucleotides (i.e., about 30–33 % of the *Xylella fastidiosa* genome) showing the phylogenetic relationships among 21 *Xylella fastidiosa* strains. Bootstrap values are shown at the nodes. Strain legend is shown in Table 1. *Xanthomonas oryzae* pv. *oryzae* KACC 10331 was included as outgroup

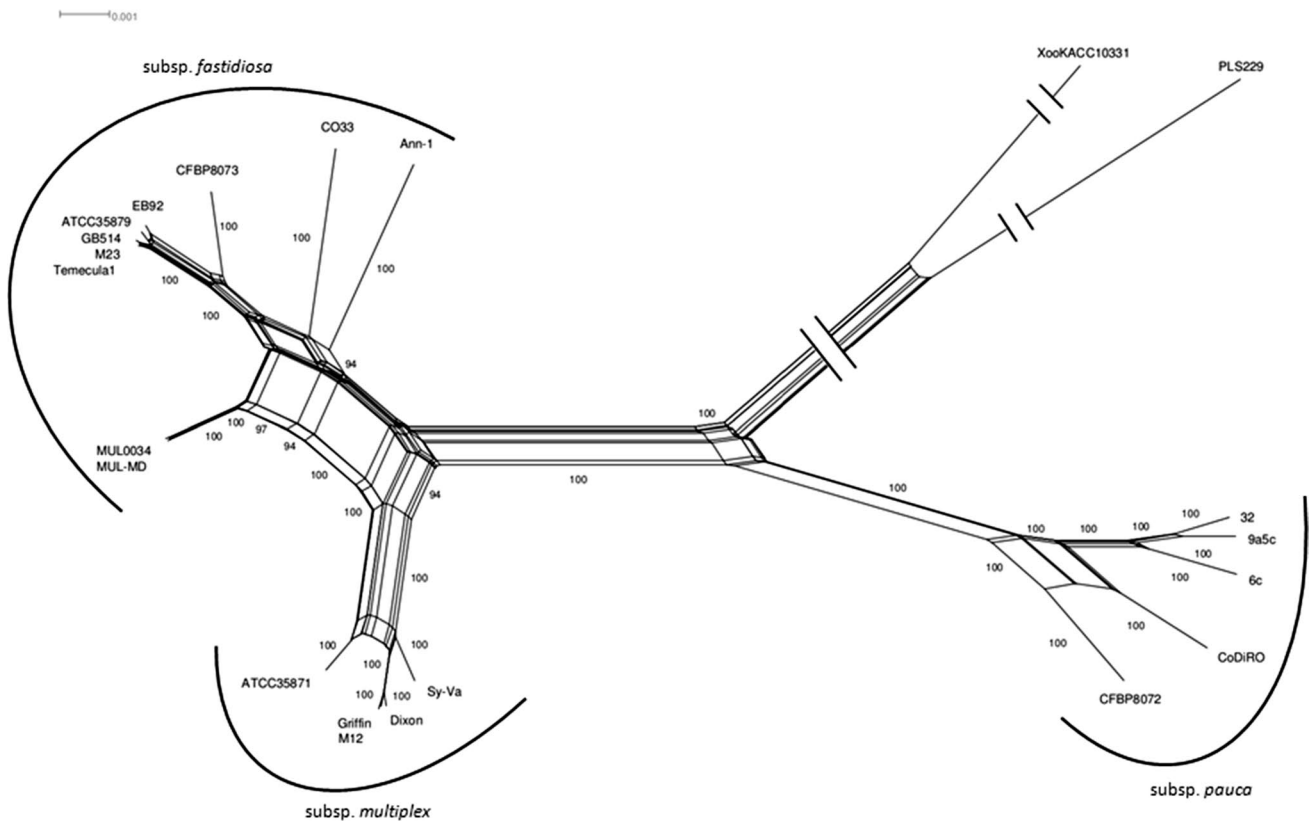

Fig. 2 Neighbor-net network as obtained from 820.088 nucleotides (i.e., about 30–33 % of the *Xylella fastidiosa* genome) and regarding 21 *Xylella fastidiosa* strains. Bootstrap values are shown at the nodes. Strain legend is shown in Table 1. *Xanthomonas oryzae* pv. *oryzae* KACC 10331 was included as outgroup

identified strong signs of recombination events within the *X. fastidiosa* strains in each of the genes assessed here (Figs. S2–S8). Note that each gene showed a different grouping of the *X. fastidiosa* strains.

Discussion

By applying different taxonomic approaches to 21 *X. fastidiosa* strains, this study clearly demonstrated the occurrence of three distinct genetic clusters within this species. Additionally, our analyses also highlighted the presence of one clearly distinct strain, i.e., PLS 229, which was isolated in Taiwan from *Pyrus pyrifolia* (Leu and Su 1993; Su et al. 2014). The three clusters matched with the three subspecies proposed and described by Schaad et al. (2004a, b), namely *fastidiosa*, *multiplex* and *pauca*. In the present study, we included only two of the three type strains of the subspecies used by these authors, i.e., ATCC 35879 and ATCC 35871 for *fastidiosa* and *multiplex*, respectively, because the *pauca* type strain ICMP 15198 has not yet been sequenced.

According to Staley and Krieg (1984), a subspecies is 'based on minor but consistent phenotypic variations within the species or on genetically determined clusters of strains within the species.'. Additionally, Wayne et al. (1987) stated that 'subspecies designation can be used for genetically close organism that diverge in phenotype.' Concerning two pathogenic species such as *Pasteurella multocida* and *Actinobacillus equuli*, Christensen et al. (2007) stressed that 'phenotypic cohesion between subspecies and related species must be fully documented as well as association with particular disease and host.' In addition to different DDH values Schaad et al. (2004a) found some differences between the three *X. fastidiosa* subspecies in terms their growth on a bacterial culture medium and in their sensitivity to penicillin and carbencillin, and host ranges. For these reasons, the designation of the three subspecies, which is based on current opinion and was confirmed in the present study, appears to be correct.

Schaad et al. (2004a) circumscribed the three subspecies by applying DDH and 16S-23S ITS sequencing assays. ANI and TETRA analyses have been found to correlate well with DDH, and ANI values of 95–96 % have been demonstrated to be equivalent to the 70 % DDH threshold, i.e., the value of the bacterial species boundary (Goris et al. 2007; Richter and Rosselló-Móra 2009). Therefore, this

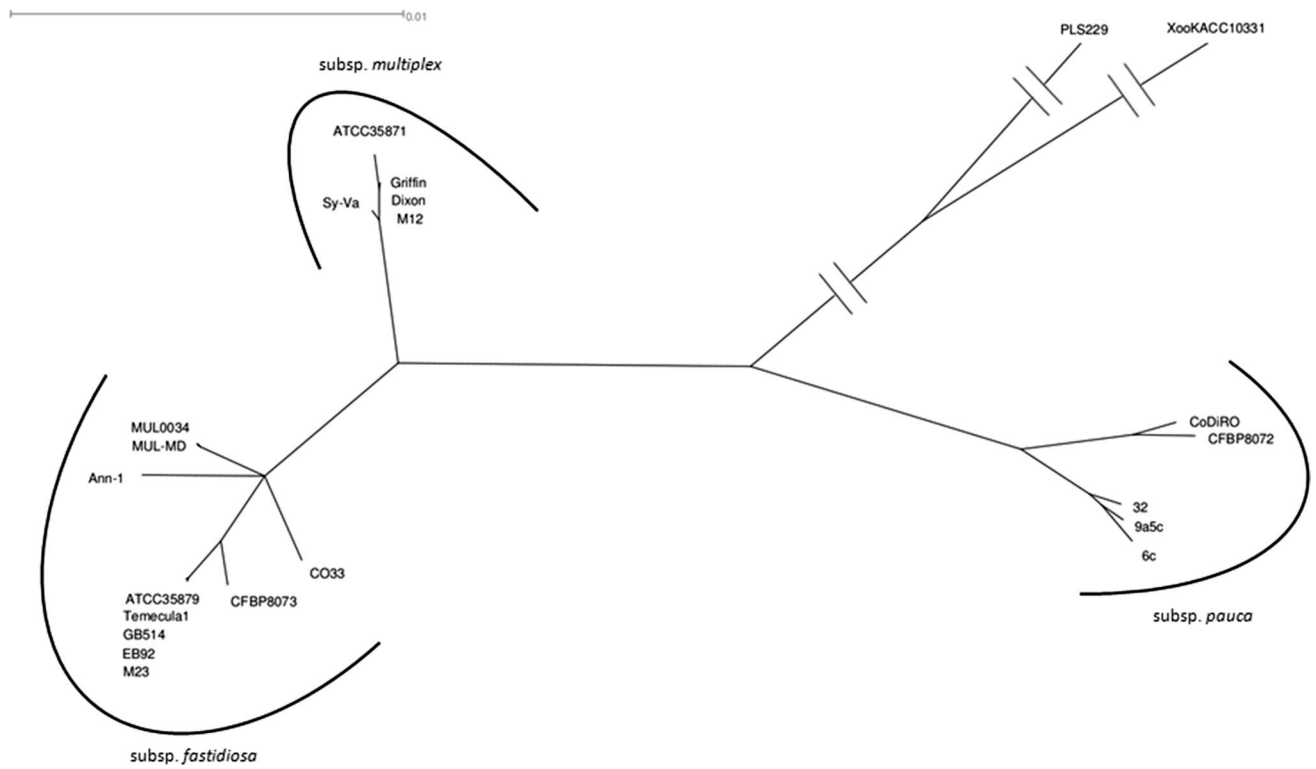

Fig. 3 Consensus tree as obtained from 956 gene sequence alignment of 21 *Xylella fastidiosa* strains. The *scale bar* indicates the number of substitution per nucleotide position. Strain legend is shown in Table 1. *Xanthomonas oryzae* pv. *oryzae* KACC 10331 was included as outgroup

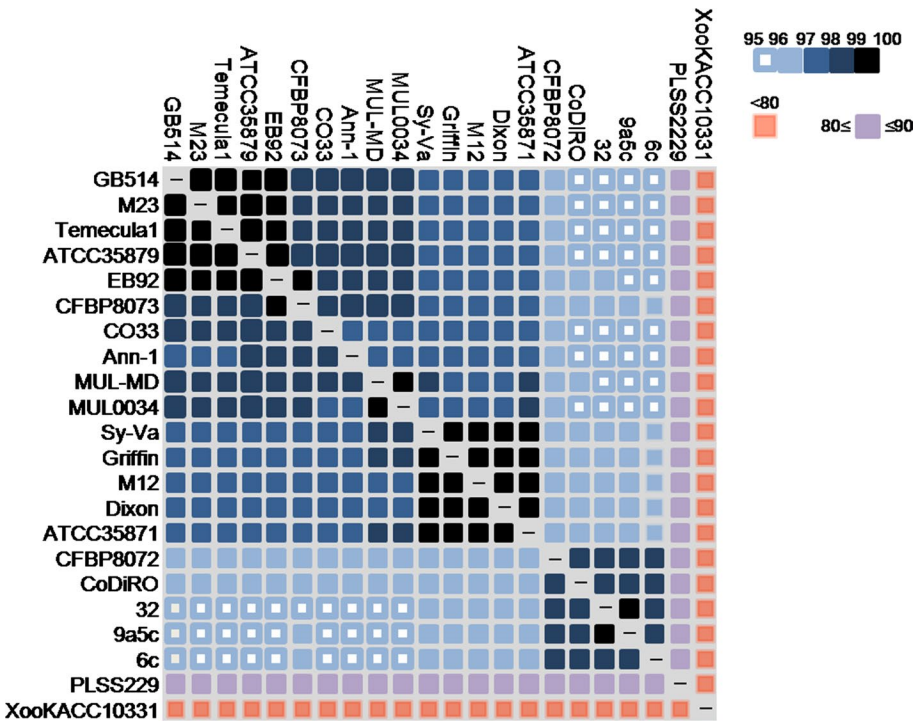

Fig. 4 Comparison of the average nucleotide identity (ANI) (i.e., percentage) between each genome of 20 *Xylella fastidiosa* strains and *Xylella* PLS 229 isolated in Taiwan from *Pyrus pyrifolia*, and the outgroup *Xanthomonas oryzae* pv. *oryzae* KACC 10331. The comparison shows higher similarities as the colors get darker

new approach has been proposed as a substitute for DDH in species demarcation (Richter and Rosselló-Móra 2009). When compared with the other 20 *X. fastidiosa* strains, the PLS 229 strain exhibited ANI values ranging from 83.56 to 83.95 %. This difference was confirmed by other taxonomic approaches used here. Indeed, in all of the dendrograms, this strain was found to be strongly segregated from the 20 genuine *X. fastidiosa* strains. The combination of the present and previous results (Chen et al. 2014) strongly supports the notion that PLS 229 represents a novel *Xylella* species. The name *X. taiwanensis* has been proposed for this strain (Chen et al. 2014). However, a formal description is necessary to properly name and describe this new taxonomic entity. This finding also indicates that phytopathogenic xylem-restricted bacteria that are transmitted by xylem sap-feeding insects are present also on continents other than North and South America.

Interestingly, we also found that the five *X. fastidiosa* strains classified as *pauca* that were assessed here exhibited ANI values in between 95.69 and 96.63 % relative to the other strains, suggesting that this subspecies has diverged from the other two subspecies. This finding was also found by Barbosa et al. (2015). The present study does not support the proposal for subspecies '*sandyi*' and '*morus*' in contrast to the studies of Scally et al. (2005) and Nunney et al. (2014). Additionally, the MLST analyses involving seven housekeeping genes that are commonly used to type *X. fastidiosa* strains provided partially discrepant results that suggested the inclusion of the strains obtained from *M. alba* into the subspecies *multiplex*. These results strongly support the importance of delineating taxonomic relationships between bacterial strains using genome-wide approaches.

These findings do not exclude the possibility that *X. fastidiosa* strains of the three subspecies can exchange portions of their genomes that are involved in the colonization of and multiplication in new host plants. Indeed, it has been demonstrated that *X. fastidiosa* is naturally competent (Kung and Almeida 2011) and that prophage-like regions in the genome can exert relevant influences on genome differentiation (De Mello Varani et al. 2008). Additionally, the acquisition of large mobile genetic elements can contribute to the augmentation of the variability within this species as demonstrated for the *pauca* strains associated with citrus variegated chlorosis in Brazil (Da Silva et al. 2007). Strong signals of gene recombination have been found also in the present study by assessing the seven MLST housekeeping genes by the recombination networks. For these reasons, we applied a more robust phylogenetic assay that has taken into consideration 956 gene families retrieved from the core genome of *X. fastidiosa*. Recently, such an approach has been used to elucidate γ-Proteobacteria (Lerat et al. 2003) and *Xanthomonas* taxonomy (Rodriguez-R et al. 2012).

The introduction of new *X. fastidiosa* strains to new geographic areas and subsequent intrasubspecific genomic introgressions and recombination with endemic strains have been demonstrated to play a relevant roles in shifts in the host ranges of the pathogen and consequent occurrences of new crop diseases in those areas (Nunney et al. 2012, 2014). A relevant order of recombination between *X. fastidiosa* strains has been also detected in the present study. The *X. fastidiosa* subsp. *pauca* CoDiRO strain which is currently involved in the 'olive quick decline syndrome' in Apulia region appears to be different from the other *pauca* strains that have been isolated in Central and South America from *Citrus sinensis* and *Coffea arabica*. Genome sequencing revealed that this strain exhibited the greatest similarities with a *pauca* strain that was isolated in Central America (Giampetruzzi et al. 2015) and is capable of infecting *N. oleander*, *C. arabica* and *Mangifera indica* (Digiaro and Valentini 2015). Other *X. fastidiosa* subsp. *pauca* strains infecting olive trees have been also detected in Argentina (Haelterman et al. 2015) and isolated in Brazil, and the latter strain differs from the CoDiRO strain based on MLST analysis (Coletta-Filho et al. 2016).

In the *pauca* strains isolated from *C. sinensis* in Brazil, Da Silva et al. (2007) found some molecular markers that were previously unique to the *X. fastidiosa* strains isolated in North America from other host plants. The utilization of specific molecular marker(s) is fundamental for the procedures leading to *X. fastidiosa* strain identification (Guan et al. 2015). However, regarding the diagnostic procedures for quarantine protocols to be adopted, especially in the areas where the pathogen is not yet present, we suggest to develop detection techniques that are capable of discriminating *X. fastidiosa* strains at the subspecies level. This approach is required to precisely intercept a known subset of the genetic variability of the pathogen. Indeed, the utilization of a strain-specific marker or markers that are targeted to identify particular strains infecting just one plant host (e.g., *M. alba*) can lead to the unexpected contemporary interception of strains of other subspecies which infect another crop (e.g., *O. europea*) (Guan et al. 2015). This issue is putatively related to other, still unknown, interrelationships between other diagnostic markers that are common to taxonomically different strains of this pathogen, and this possibility can cause doubts (i.e., those related to false positive results) especially for asymptomatic plant material from new, unexpected, *X. fastidiosa* hosts (Jacques et al. 2015), as has been recently occurred in Southern Europe.

Subspecies description: The description of *X. fastidiosa* subspecies is based on the features provided by Schaad et al. (2004a) and on the data obtained from the genome-wide assessment of the present study. For the composition of bacterial culture media, see EPPO (2004). The list

of major *X. fastidiosa* host plants has been updated even though new hosts could be found in the future.

Xylella fastidiosa subsp. *fastidiosa* (fas.tid.i.o'sa. N.L. m. adj. fastidiosus, highly critical, referring to the nutritional fastidiousness of the organism). Strains of this subspecies grow on the following media: PD2, PW, BCYE and CS20. They are less resistant to carbencillin than *X. fastidiosa* subsp. *multiplex* and *pauca* strains. Serology differentiates *X. fastidiosa* subsp. *fastidiosa* from *X. fastidiosa* subsp. *multiplex* and *X. fastidiosa* subsp. *pauca*. This subspecies cause disease to grape (*Vitis vinifera* L.), almond (*Prunus amygdalus* L.), cherry (*Prunus avium* L.), alfalfa (*Medicago sativa* L.), coffee (*Coffea arabica* L. and *C. canephora* L.), oleander (*Nerium oleander* L.), maples (*Acer* spp.), American black elderberry (*Sambucus canadensis L.).* Due to a not restricted host specialization, other cultivated and wild plant species can be infected and/or colonized by strains of this subspecies. An average nucleotide identity (ANI) assessment indicates a genome similarity percentage within the subspecies *fastidiosa* strains ranging from 97.82 to 99.93 %. The GC content of the subspecies ranges from 51.5 to 52.1 %. The type strain of the subspecies *fastidiosa* is ATCC 35879 (=ICPB 50025 = DSM 10026 = ICMP 15197 = LMG 17159).

Xylella fastidiosa subsp. *multiplex* (mul' ti. plex. adj. multiplex, numerous, named to recognize the large number of host plants in which the bacterium causes disease). Strains of this subspecies grow much faster on PW medium than on PD2, BCYE and CS20. They are more resistant to carbencillin and more susceptible to penicillin than *X. fastidiosa* subsp. *fastidiosa* strains. Serology differentiates *X. fastidiosa* subsp. *multiplex* from *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *pauca*. This subspecies causes disease to plum (*Prunus salicina* Lindl.), peach (*Prunus persica* (L.) Batsch), almond (*Prunus amygdalus* L.), sycamore (*Platanus occidentalis* L.), mulberry (*Morus alba* L.), oaks (*Quercus* spp.), elms (*Ulmus* spp.), myrtle-leaf milkwort (*Polygala myrtifolia* L.), Spanish broom (*Spartium junceum* L.). Due to a not restricted host specialization, other cultivated and wild plant species can be infected and/or colonized by strains of this subspecies. An average nucleotide identity (ANI) assessment indicates a genome similarity percentage within the subspecies *multiplex* strains ranging from 99.29 to 99.95 %. The GC content of the subspecies ranges from 51.6 to 52.0 %. The type strain of the subspecies *multiplex* is ATCC 35871 (=ICPB 50039 = ICMP 15199 = LMG 9063).

Xylella fastidiosa subsp. *pauca* (pau' ca L. fem. adj. pauca, few, named to recognize the narrow host range of this bacterium). Strains of this subspecies grow more slowly than *X. fastidiosa* subsp. *fastidiosa* strains on PD2, PW, BCYE and CS20. They are susceptible to penicillin and more resistant to carbencillin than *X. fastidiosa* subsp. *fastidiosa* strains. Serology differentiates *X. fastidiosa* subsp. *pauca* from *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex*. This subspecies causes disease to citrus (*Citrus* spp.) and coffee (*C. arabica*). Strains of this subspecies have been found associated with oleander, almond, cherry and with olive (*Olea europea* L.) trees showing extensive leaf scorching/wilting and twig die-back in Apulia (Southern Italy), Argentina and Brazil. Due to a not restricted host specialization, other cultivated and wild plant species can be infected and/or colonized by strains of this subspecies. An average nucleotide identity (ANI) assessment indicates a genome similarity percentage within the subspecies *pauca* strains ranging from 98.01 to 99.15 %. The ANI values related to the other two *X. fastidiosa* subspecies ranged from 95.69 to 96.63 %. The GC content of the subspecies ranges from 51.9 to 52.6 %. The type strain of the subspecies *pauca* is ICMP 15198.

Acknowledgments The study was financed by the ordinary funds of the Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria.

References

Almeida RPP, Nunney L (2015) How do plant diseases caused by *Xylella fastidiosa* emerge ? Plant Dis 99:1457–1467. doi:10.1094/PDIS-02-15-0159-FE

Altschul SF, Madden TL, Schäffer AA, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389–3402. doi:10.1093/nar/25.17.3389

Aziz RK, Bartels D, Best AA, De Jongh M, Disz T, Edwards RA, Formsma K, Gerdes L, Glass EM et al (2008) The RAST server: rapid annotation using subsystem technology. BMC Genom 9:75. doi:10.1371/journal.pone.0038596

Barbosa D, Alencar VC, Santos DS, De Freitas Oliveira AC, De Souza AA, Coletta-Filho HD, De Oliveira RS, Nunes LR (2015) Comparative genomic analysis of coffee-infecting *Xylella fastidiosa* strains isolated from Brazil. Microbiology 161:1018–1033. doi:10.1099/mic.0.000068

Bull CT, De Boer SH, Denny TP, Firrao G, Fischer-Le Saux M, Saddler GS, Scortichini M, Stead DE, Takikawa Y (2010) Comprehensive list of names of plant pathogenic bacteria, 1980-2007. J Plant Pathol 92:551–592. doi:10.4454/jpp.v92i3.302

Bull CT, De Boer SH, Denny TP, Firrao G, Fischer-Le Saux M, Saddler GS, Scortichini M, Stead DE, Takikawa Y (2012) List of new names of plant pathogenic bacteria (2008-2010). J Plant Pathol 94:21–27. doi:10.4454/jpp.fa.2011.003

Cariddi C, Saponari Boscia D, De Stradis A, Loconsole G, Nigro F, Porcelli F, Potere O, Martelli GP (2014) Isolation of a *Xylella fastidiosa* strain infecting olive and oleander in Apulia, Italy. J Plant Pathol 96:1–5. doi:10.4454/JPP.V96I2.024

Chen J, Su C-C, Deng W-L, Jan F-J, Chang C-J, Huang H (2014) Analyses of *Xylella* whole genome sequences and proposal of *Xylella taiwanensis* sp. nov. In: Proceedings of the 2014 Pierce's disease research Symposium, December 15–17, 2014, Sacramento, CA, USA

Christensen H, Kuhnert P, Busse H-J, Frederiksen WC, Bisgaard M (2007) Proposed minimal standards for the description of genera, species and subspecies of the *Pasteurellaceae*. Int J Syst Evol Microbiol 57:166–178. doi:10.1099/ijs.0.64838-0

Chun J, Rainey FA (2014) Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. Int J Syst Evol Microbiol 64:316–324. doi:10.1099/ijs.0.054171-0

Coletta-Filho HD, Francisco CS, Lopes JRS, De Oliveira AF, Da Silva LFO (2016) First report of olive leaf scorch in Brazil, associated with *Xylella fastidiosa* subsp. *pauca*. Phytopathol Medit. doi:10.14601/Phytopathol_Mediterr-17259

Da Silva VS, Shida CS, Rodrigues FB, Ribeiro DCD, De Souza AA, Coletta-Filho ED, Machado MA, Nunes LR et al (2007) Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. BMC Genom 8:474. doi:10.1186/1471-2164-8-474

De Mello Varani A, Souza RC, Nakaya HI, De Lima WC, De Almeida LGP, Kitajima EW, Chen J, Civerolo E, Vasconcelos ATR et al (2008) Origins of the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. PLoS One 3:e4059. doi:10.1371/journal.pone.0004059

Digiaro M, Valentini F (2015) The presence of *Xylella fastidiosa* in Apulia region (Southern Italy) poses a serious threat to the whole Euro-Mediterranean region. Centre International de Hautes Etudes Agronomiques Méditerranéennes Watch Letter, no 33

EFSA Panel on Plant Health (2015) Scientific opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory, with the identification and evaluation of risk reduction options. EFSA J 13:3989

European Plant Protection Organization (2004) Diagnostic protocol for regulated pests. *Xylella fastidiosa*. Bull OEPP/Eppo Bull 24:187–192. doi:10.1111/j.1365-2338.2004.00718.x

European Plant Protection Organization (2015) *Xylella fastidiosa* detected in Alpes-Maritimes, mainland France. Reporting Service no 10, 180

Giampetruzzi A, Chiumenti M, Saponari M, Donvito G, Italiano A, Loconsole G, Boscia D, Cariddi C, Martelli GP et al (2015) Draft genome sequence of the *Xylella fastidiosa* CoDiRO strain. Genome Announc 3:e01538-14. doi:10.1128/genomeA.01538-14

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007) DNA–DNA hybridization values and their relation to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. doi:10.1099/ijs.0.64483-0

Guan W, Shao J, Elbeaino T, Davis RE, Zhao T, Huang Q (2015) Specific detection and identification of American mulberry-infecting and Italian olive-associated strains of *Xylella fastidiosa* by polymerase chain reaction. PLoS One 10:e0129330. doi:10.1371/journal.pone.0129330

Guindon S, Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704. doi:10.1080/10635150390235520

Haelterman RM, Tolocka PA, Roca ME, Guzman FA, Fernandez FD, Otero ML (2015) First presumptive diagnosis of *Xylella fastidiosa* causing olive scorch in Argentina. J Plant Pathol 97:393. doi:10.4454/JPP.V97I2.023

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267. doi:10.1093/molbev/msj030

Jacques M-A, Denancé N, Legendre B, Morel E, Briand M, Mississipi S, Durand K, Olivier V, Portier P et al (2015) New variants of coffee-infecting *Xylella fastidiosa* issued from homologous recombination. Appl Environ Microbiol. doi:10.1128/AEM.03299-15

Kim M, Oh H-S, Park S-C, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. doi:10.1099/ijs.0.059774-0

Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci USA 102:2567–2572. doi:10.1073/pnas.0409727102

Kung SH, Almeida RPP (2011) Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278–5524. doi:10.1128/AEM.00730-11

Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, Clark VA (1992) International code of nomenclature of bacteria 1990. Revision bacteriological code. American Society for Microbiology, Washington

Lerat E, Daubin V, Moran NA (2003) From gene trees to organismal phylogeny in Prokaryotes: the case of the γ–Proteobacteria. PLoS Biol 1:e19. doi:10.1371/journal.pbio.0000019

Leu LS, Su CC (1993) Isolation, cultivation and pathogenicity of *Xylella fastidiosa*, the causal bacterium of pear leaf scorch disease in Taiwan. Plant Dis 77:642–646

Loconsole G, Potere O, Boscia D, Altamura G, Djelouah K, Elbeaino T, Frasheri D, Lorusso D, Palmisano F et al (2014) Detection of *Xylella fastidiosa* in olive trees by molecular and serological methods. J Plant Pathol 96:7–14. doi:10.4454/JPP.V96I1.041

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K et al (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA 95:3140–3145

Marcelletti S, Scortichini M (2014) Definition of plant-pathogenic *Pseudomonas* genomospecies of the *P. syringae* complex through multiple comparative approaches. Phytopathology 104:1274–1282. doi:10.1094/PHYTO-12-13-0344-R

Marcelletti S, Scortichini M (2015) Comparative genomic analyses of multiple *Pseudomonas* strains infecting *Corylus avellana* trees reveal the occurrence of two genetic clusters with both common and distinctive virulence and fitness traits. PLoS One 10:e0131112. doi:10.1371/journal.pone.0131112

Nunney LR, Yuan X, Bromley RE, Stouthamer R (2012) Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702–4714. doi:10.1128/AEM.01126-12

Nunney LR, Schuenzel EL, Scally M, Bromley RE, Stouthamer R (2014) Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025–3033. doi:10.1128/AEM.04112-13

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290. doi:10.1093/bioinformatics/btg412

Poptsova MS, Gogarten JP (2007) BranchClust: a phylogenetic algorithm for selecting gene families. BMC Bioinform 8:120. doi:10.1186/1471-2105-8-120

Purcell AH (2008) Transmission of *Xylella fastidiosa* bacteria by xylem-feeding insects. In: Capinera JL (ed) Encyclopedia of entomology. Springer, Dordrecht, pp 3885–3886

Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, Raoult D, Fournier PE (2014) A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 64:384–391. doi:10.1099/ijs.0.057091-0

Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 106:19126–19131. doi:10.1073/pnas.0906412106

Rodriguez-R LM, Grajales A, Arrieta-Ortiz ML, Salazar C, Restrepo S, Bernal A (2012) Genome-based phylogeny of the genus *Xanthomonas*. BMC Microbiol 12:43. doi:10.1186/1471-2180-12-43

Scally M, Schuenzel EL, Stouthamer R, Nunney L (2005) Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contribution of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491–8499. doi:10.1128/AEM.71.12.8491-8499.2005

Schaad NW, Postnikova E, Lacy G, Fatmi M, Chang CJ (2004a) *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp. *piercei*, subsp.

nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst Appl Microbiol 27:290–300. doi:10.1078/0723-2020/04/2369-848

Schaad NW, Postnikova E, Lacy G, Fatmi M, Chang CJ (2004b) *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp. *piercei*, subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Erratum Syst Appl Microbiol 27:763. doi:10.1078/0723-2020/04/27/06-763

Scortichini M, Marcelletti S, Ferrante P, Firrao G (2013) A genomic redefinition of *Pseudomonas avellanae* species. PLoS One 8:e75794. doi:10.1371/journal.pone.0075794

Staley JT, Krieg NR (1984) Classification of prokaryotic organisms: an overview. In: Krieg NR, Holt JG (eds) Bergey's manual of systematic bacteriology, vol 1. Williams & Wilkins, Baltimore, pp 1–7

Su C-C, Deng W-L, Jan F-J, Chang C-J, Huang H, Chen J (2014) Draft genome sequence of *Xylella fastidiosa* pear leaf scorch strain in Taiwan. Genome Announc 2:e00166-14. doi:10.1128/genomeA.00166-14

Thompson CC, Chimetto L, Edwards RA, Swings J, Stackebrandt E, Thompson FL (2013) Microbial genomic taxonomy. BMC Genom 14:913. doi:10.1186/1471-2164-14-913

Wall DP, Fraser HB, Hirsch AE (2003) Detecting putative orthologs. Bioinform Appl Note 19:1710–1711. doi:10.1093/bioinformatics/btg213

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevski MI, Moore LH, Moore WEC, Murray RGE et al (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematic. Int J Syst Bacteriol 37:463–464. doi:10.1099/00207713-37-4-463

Wells JM, Raju BC, Hung H-Y, Weisburg WG, Mandelco-Paul L, Beemer DJ (1987) *Xylella fastidiosa* gen. nov., sp. nov.: gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. Int J Syst Bacteriol 37:136–143. doi:10.1099/00207713-37-2-136