SCIENTIFIC OPINION

EFSA Journal

ADOPTED: 29 June 2016

doi: 10.2903/j.efsa.2016.4542

Statement on diversity of *Xylella fastidiosa* subsp. *pauca* in Apulia

EFSA Panel on Plant Health (PLH)

Abstract

This opinion addresses a request from the European Commission to evaluate whether heterogeneous populations of *Xylella fastidiosa* subsp. *pauca* have been found in Apulia (Italy) in addition to the strain named CoDiRO. After reviewing the most recent scientific literature and conducting further sequence analysis of the housekeeping genes used to genotype *X. fastidiosa* from Apulia, the EFSA Panel on Plant Health concluded that the currently available scientific evidence does not support the notion of the existence of heterogeneous populations of *X. fastidiosa* in Apulia. To reach this conclusion, several lines of arguments have been considered: (i) the currently accepted multilocus sequence analysis (MLST) approach provides a robust and sensitive framework to estimate *X. fastidiosa* diversity, which may be further improved by whole genome sequence analysis; (ii) all scientific papers evaluated on *X. fastidiosa* in Apulia come to the conclusion that – notwithstanding host and location – *X. fastidiosa* in Apulia is ST53; and (iii) a single article by Elbaino et al. (2014) provides contradictory statements. After assessment of the methodology and re-analysis of the sequence information underlying the conflicting statement, the Panel considered that the data presented in that article do not support firm conclusions on the existing diversity of isolates. The Panel further considered that MLST data are currently available for only 18 Apulian isolates, and thus more data are needed to study further *X. fastidiosa* diversity, strain evolution and route(s) of introduction. From all evidence currently available, the Panel concludes that *X. fastidiosa* isolates involved in the current epidemic in Apulia belong to a single sequence type, ST53. There is no other information at this moment supporting the notion that, in addition to the strain named CoDiRO, other diversity exists.

Keywords: CoDiRO strain, housekeeping genes, isolates, MLSA, MLST, olive disease, ST53

Requestor: European Commission

Question number: EFSA-Q-2016-00182

Correspondence: alpha@efsa.europa.eu

Panel members: Claude Bragard, David Caffier, Thierry Candresse, Elisavet Chatzivassiliou, Katharina Dehnen-Schmutz, Gianni Gilioli, Jean-Claude Grégoire, Josep Anton Jaques Miret, Michael Jeger, Alan MacLeod, Maria Navajas Navarro, Bjoern Niere, Stephen Parnell, Roel Potting, Trond Rafoss, Vittorio Rossi, Gregor Urek, Ariena Van Bruggen, Wopke Van Der Werf, Jonathan West and Stephan Winter

Acknowledgements: The Panel wishes to thank the members of the Working Group on urgent review of statements on *Xylella fastidiosa* and its control Claude Bragard, Leonardo De La Fuente, Marie-Agnès Jacques and Stephan Winter for the preparatory work on this scientific opinion and EFSA staff members Miren Andueza and Sara Tramontini for the support provided to this scientific opinion.

Suggested citation: EFSA PLH Panel (EFSA Panel on Plant Health), 2016. Statement on diversity of *Xylella fastidiosa* subsp. *pauca* in Apulia. EFSA Journal 2016;14(8):4542, 19 pp. doi:10.2903/j.efsa.2016.4542

ISSN: 1831-4732

The EFSA Journal is a publication of the European Food Safety Authority, an agency of the European Union.

Table of contents

1. Introduction

1.1. Background and Terms of Reference as provided by the requestor[1]

The purpose of this mandate is to request, pursuant to Article 29 of Regulation (EC) No 178/2002[2], scientific advice in the field of plant health as regards the regulated harmful organism *Xylella fastidiosa* (Wells et al.).

Specifically, the Commission has recently been confronted with a number of statements which are questioning the overall European Union (EU) control strategy against *X. fastidiosa* and some relevant legal provisions laid down under Decision (EU) 2015/789[3]. Such statements are the grounds for several appeals to the European Court of Justice which are pending for final ruling. Those statements and the related questions on which the Commission requests the European Food Safety Authority (EFSA) scientific advice are presented below:

1) It is considered that *the population of* Xylella fastidiosa *subsp.* pauca*, in Apulia (Italy) is heterogeneous as several different strains are present in the infected area, on top of the unique strain (referred to Co.Di.RO) reported so far*.
    - Is there any scientific conclusive evidence for such a statement?
2) The expression of the so-called 'quick declining symptoms in olive plants' (CoDiRO) seems *to be correlated, not only with the presence of* Xylella fastidiosa *or other fungi present within the xylematic vessels within the plant, but also with a number of other factors* which have not been fully taken into account in the EU Decision*.* Such factors are*: the degree of compactness of the soil, quantity of organic matter in the soil, presence of biodiversity between the micro-fauna of the soil, degree of salinisation of the soil, concentration of glyphosate (or other chemical toxic agents), nutrient concentration, as well as any pruning activities carried out, including ploughing of the soil and other agricultural practices*.
    - Is this statement in agreement with current scientific knowledge? Please advise whether this would affect the risk of *Xylella fastidiosa* for the rest of the Union.
3) *The causing link between* Xylella fastidiosa *and the quick declining symptoms of olive trees is still not established and Koch's postulates have not yet been fulfilled. Therefore, it is not sure that* Xylella fastidiosa *is the only and confirmed causing agent of the plant death*.
    - Can EFSA provide an update on the current scientific knowledge about this topic? In case Koch's postulates have not yet been fulfilled for the 'quick declining symptoms' in olives, please advise whether this would affect the risk of *Xylella fastidiosa* for the rest of the Union compared to that reported in the Pest Risk Assessment of January 2015 (EFSA Journal 2015, 13(1):3989)?
4) *Removal of infected trees is not considered to be a feasible option to contain or eradicate the bacterium, nor to prevent the further spread of the quick declining symptoms of olive plants, as also experienced in USA, Brazil and Taiwan.* Even more*, the removal of host plants, regardless of their health status, within a radius of 100 m around the infected plants as requested by Decision (EU) 2015/789* for any outbreak identified outside the province of Lecce*,* where the bacterium is not yet established*, is considered to be not scientifically validated*.
    - Can EFSA review such a statement on the basis of current scientific knowledge with regard to the level of prevention of further spread of *Xylella fastidiosa* in areas not yet infected? In particular:
        - In a system-based approach, as proposed in Article 7 of Decision (EU) 2015/789, can EFSA advise about the efficacy of removing infected plants located within an area where the bacterium is considered to be established (so-called

[1] Submitted by European Commission, ref. SANTE/G1/PDR/mm(2016)1031036.
[2] Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety, OJ L 31, 1.2.2002, p. 1–24, as last amended.
[3] Commission Implementing Decision (EU) 2015/789 of 18 May 2015 as regards measures to prevent the introduction into and the spread within the Union of *Xylella fastidiosa* (Wells et al.). OJ L 125, 21.5.2015, p. 36–53.

'containment area'), and particularly located in the proximity of the buffer zone, where the bacterium is not yet present with the aim to prevent further spread?
- In a system-based approach, as proposed in Article 6 of Decision EU 2015/789, can EFSA advise about the efficacy of removing host plants, regardless of their health status, located in the proximity of recently detected infected plants, located in areas where the bacterium was not known to occur before that detection (e.g. buffer zone or outside the 'containment area') with the aim to prevent further spread?

5) *It is alleged solutions to treat diseased plants in open field would be currently available. In this respect, it is often referred to experiments carried out by Prof. Marco Scortichini of CREA (Caserta, Italy) and the ones carried out by Prof. Francesco Lops and Dr. Antonia Carlucci from the University of Foggia (Italy).*
   - Can EFSA contact these researchers and assess the outcome, if provided, of these on-field experiments aiming at curing diseased plants?
   - Can EFSA also provide an update about recent treatment solutions, scientifically validated, if any, to cure diseased plants?

6) From the Pest Risk Assessment of EFSA (EFSA Journal 2015, 13(1):3989), reference is made to Section 3.5.2 '*The intensive use of insecticide treatment to limit the disease transmission and control the insect vector may have direct and indirect consequences for the environment by modifying whole food webs with cascading consequences, and hence affecting various trophic levels. For example, the indirect impact of pesticides on pollination is currently a matter of serious concern (EFSA, 2013b). In addition, large-scale insecticide treatments also represent risks for human and animal health*'; Section 4.3.2.2. '*large-scale application of insecticides may lead to the development of insecticide resistance as well as to environmental and human health issues*'; Section 4.3.3.4. '*Similarly, insecticide treatments could have a negative result by modifying insect population dynamics and favouring insect vectors, e.g. by placing proportionally higher pressure on the insects' natural enemies*'.
   - Can EFSA provide clarification on this matter in relation to the phytosanitary treatments required by Decision (EU) 2015/789 to be carried out prior to the removal of plants referred to in paragraph 2 of Article 6 and Article 7 against the vectors of *Xylella fastidiosa* and plants that may host those vectors? It is to be noted that those treatments, as appropriate, may include as well the removal of herbs where insect vectors lay down their eggs.

In view of a quick reaction expected by the Commission as part of the ongoing appeals to the European Court of Justice, EFSA is requested to prepare an urgent opinion by 18 March 2016. As regards specifically point 1 and point 5 above, an extended deadline could be set for 30 June and 31 March 2016, respectively.

1.2. Interpretation of the Terms of Reference

The Terms of Reference (ToR) are organised in six points, each of which refers to a different aspect of risks connected to *Xylella fastidiosa* presence in the EU. Under each point, the European Commission addresses one or more questions to EFSA.

In the current opinion, the PLH Panel replies to point 1 and related question as required by the indicated deadline.

The current opinion was prepared in the light of the Italian outbreaks of CoDiRO (Complesso del Disseccamento Rapido dell'Olivo, whose English equivalent is OQDS, from Olive Quick Decline Syndrome). Therefore, unless specified otherwise, the focus of the opinion is on *X. fastidiosa* subsp. *pauca* populations found in the infected area of Apulia.

The question whether there is scientific evidence for the existence of heterogeneous populations of *X. fastidiosa* subsp. *pauca* in the infected area is interpreted by the Panel as a request for confirmation of the presence of further strains associated with the Apulian outbreaks, in addition to the *X. fastidiosa* subsp. *pauca* strain associated with the CoDiRO disease of olives. The Panel considers that the terms 'heterogeneity' and 'strain' in the articulation of the question are used with reference to entities/ organisms that are distinguishable by genetic and/or biological characteristics.

2. Data and methodologies

2.1. Data

To revise the statement and reply to the connected question, targeted extensive literature searches were conducted. Searches were carried out on the research platform ISI Web of Science. The references retrieved were reviewed together with those cited in the EFSA risk assessment on *X. fastidiosa* produced earlier (EFSA PLH Panel, 2015). Further references and information were obtained from citations within the reviewed references and from experts in the field. In addition to the review of recent publications, evidence was collected from sequence data available at the pubMLST and NCBI GenBank databases and by considering this information with reference to (i) the CoDiRO disease on olives and (ii) the origin of the isolate(s).

2.2. Methodologies

The assessment was conducted in line with the principles described in the EFSA Guidance on transparency in the scientific aspects of risk assessment (EFSA, 2009). The present document is structured according to the Guidance on the structure and content of EFSA's scientific opinions and statements (EFSA Scientific Committee, 2014).

Sequences of Apulian isolate/strain genes that are highly conserved in *X. fastidiosa* and other bacteria and involved in basic cell maintenance function (housekeeping genes) and thus commonly used for *X. fastidiosa* subsp. comparison were downloaded from NCBI with accession numbers provided in Elbeaino et al. (2014), Bleve et al. (2016) and Loconsole et al. (2016). *X. fastidiosa* allele sequences were downloaded from the pubMLST database. The allele sequences of the CoDiRO strain were retrieved from its genome sequence (Giampetruzzi et al., 2015a). When necessary, sequences were trimmed to fit the fragments used in multilocus sequence typing (MLST) as described in http://pubmlst.org/xfastidiosa/. Sequences were analysed using Geneious software as described in Jacques et al. (2016).

3. Assessment

3.1. Point 1 – Heterogeneity of *Xylella fastidiosa* populations in Apulia

This request refers to PCR and DNA sequencing tests conducted to identify and to discriminate *X. fastidiosa* subsp. *pauca* isolates from olives and other infected plants in the contaminated area of the Apulia region, and to the possibility that the results published in recent manuscripts indicate or not the existence of genetic diversity among the analysed isolates, highlighting the existence of multiple strains.

3.1.1. Clarification on taxonomy of *Xylella fastidiosa* populations in Apulia

The Panel provides clarifications for some of the most common terms used in bacterial taxonomy and classification to clarify ambiguities regarding denominations on the taxonomical affiliation of *X. fastidiosa* isolates found in the recent literature on the Italian outbreak (Cariddi et al., 2014; Elbeaino et al., 2014; Loconsole et al., 2014, 2016; Giampetruzzi et al., 2015a; Bleve et al., 2016; Mang et al., 2016).

In bacteriology, the term 'isolate' refers to a colony derived from a single cell grown in pure culture, obtained from samples of environmental (e.g. insect, plant, soil, etc.), clinical (e.g. infected animal or human specimen) or any other origin. This first colony forming unit (CFU) obtained from a particular sample may originate from a single cell or not and is macroscopically identified on an agar plate as a distinct mass of bacterial growth. To ensure purity, single colonies are re-streaked and then each independent bacterial growth in the form of a colony is identified as a new CFU. These re-streaked CFUs are considered to be clonal populations of bacteria growing on an agar plate that originate from a single bacterial cell.

For the term 'strain', the Bergey's Manual of Systematic Bacteriology (Staley and Krieg, 1984) provides the following definition: *A strain is made up of the descendants of a single isolation in pure culture, and is usually made up of a succession of cultures ultimately derived from an initial single colony*. Strains are identities with unique phenotypic and/or genotypic characteristics and by based on further characterisation (e.g. by whole genome sequencing, virulence assessment, identification of

genetic or phenotypic markers, etc.), attributes can be added to further define the strain. Because a clear definition of 'isolate' and 'strain' is not given, in the scientific literature the two terms are sometimes used interchangeably, as they both refer to unique, homogeneous, clonal bacterial populations.

As an example, when isolating bacteria from an infected plant, all of the CFUs grown on agar media can be considered distinct isolates. *A priori* it could be considered that all CFUs cultured from a single plant reflect the same strain. However, because co-infection with different strains of the same pathogen can occur, caution has to be taken on the possible diversity of the isolates. Various plants from which bacterial isolates are taken can have a very different composition of 'strains', highly similar or identical ones as well as diverse strains (López-Villavicencio et al., 2007; Tollenaere et al., 2012; Darrasse et al., 2013; Susi et al., 2015; Wei et al., 2016). In other words, two isolates from the same plant may or may not belong to the same strain. Therefore, prior to further characterisation, a denomination of such isolates as 'strains' is not justified. When DNA sequences are obtained from plant material without isolation of the bacteria, these sequences should be considered as belonging to a 'sample', and should never be considered as a single isolate or a strain, since in this situation the purity of the bacterial sample is not confirmed.

In the recent literature on the Apulian outbreak the two terms (isolate and strain) have been applied inconsistently and sometimes incorrectly. Some publications on this topic mention that all 'isolates' obtained in a specific geographic area or from the same plants are the same 'strain' just because they cause the same symptoms and are supposed to be genetically identical only based on a limited set of gene fragments (Cariddi et al., 2014; Elbeaino et al., 2014; Giampetruzzi et al., 2015a; Bleve et al., 2016; Mang et al., 2016). Another problem contributing to the confusion is providing a name to an 'isolate' when amplification of a few bacterial genes from infected plant material is done without further development of pure cultures (Elbeaino et al., 2014; Loconsole et al., 2016; Mang et al., 2016), therefore not being able to discriminate if the plant was infected by a single or multiple strains. In particular, this confusing use of the terminology applies to the definition of the 'CoDiRO strain'. Saponari et al. (2014) used for the first time the term 'CoDiRO strain' to name all the isolates (PW1, OLDR-1) or DNA sequences (OL-G2) obtained from different host plants found infected in the outbreak area reported by Cariddi et al. (2014). Contributing to the confusion is the fact that the latter manuscript (Cariddi et al., 2014) uses the term 'Salento strain' and not 'CoDiRO strain'. Giampetruzzi et al. (2015a) sequenced the whole genome of the so-called 'CoDiRO strain' from infected periwinkle, although no information is provided about the origin of the sample/isolate. In a previous publication by the same group (Cariddi et al., 2014), an isolate from periwinkle, obtained by infection with insect vectors collected in the field, is denominated PW-1, but it is not clear if this isolate is the one whose whole genome sequence is presented by Giampetruzzi et al. (2015a). Thus, the ambiguous term 'CoDiRO strain' seems to be generically applied in the literature to refer to the *X. fastidiosa* subsp. *pauca* causing the CoDiRO disease in olives and present in different hosts in Apulia (Elbeaino et al., 2014; Loconsole et al., 2016; Saponari et al., 2016), although sequence data from different Apulian 'strains' with different accession numbers are available from the NCBI and pubMLST database (Appendix A). The fact of naming this 'strain' with the same disease name without a name/number for specific identification (in contrast with, for example, the 'Salento-1' 'strain' in Bleve et al., 2016) contributed to the confusion by not allowing to discriminate among the different isolates or 'strains' present in the infected area. In conclusion, the inconsistent application of the term 'strain' in the scientific literature about the *X. fastidiosa* outbreak in the Apulia region lead to uncertainties regarding the heterogeneity of the population of *X. fastidiosa* in the area.

3.1.2. Analysis of genetic diversity of *Xylella fastidiosa*

Classification of bacteria based on DNA sequence data is very useful since DNA sequence-based techniques are portable and highly reproducible across laboratories. Sequence analysis of DNA avoids variability inherent in classical experimental techniques used for bacterial classification, such as DNA–DNA hybridisation or techniques that depend on phenotypical traits (e.g. colony morphology, biochemical or metabolic characteristics) and PCR fragment analysis (Vinatzer and Bull, 2009; Perez-Losada et al., 2013; Glaeser and Kampfer, 2015). For bacterial classification and taxonomy, full genome analyses is becoming the reference tool, as exemplified by average nucleotide identity (ANI), a technique being proposed as a useful approach for species classification that is rapidly gaining popularity (Konstantinidis and Tiedje, 2005; Bull and Koike, 2015). However, until now, bacterial classification is sometimes uniquely based on sequencing of a limited set of genes, among which the

16S rRNA is the most common. The 16S rRNA gene sequence is currently the most popular approach for species identification because of its high sequence conservation and ubiquitous presence in Bacteria and Archaebacteria (Glaeser and Kampfer, 2015). Sequence information of the 16S rRNA gene across bacteria is deposited in the Ribosomal Data Project (http://rdp.cme.msu.edu/index.jsp), which currently comprises more than 1.5 million sequences of culturable and non-culturable bacteria. The major limitation of 16S rRNA analysis is the inadequate resolution to discriminate beyond the genus/species level (Bull and Koike, 2015; Glaeser and Kampfer, 2015).

The MLST approach first proposed by Maiden et al. (1998) for the classification of isolates within populations of pathogenic microorganisms takes nucleotide sequence data into consideration. The technique has been increasingly used in molecular epidemiology (Gevers et al., 2005; Glaeser and Kampfer, 2015). MLST comprises the analysis of partial sequences of several (typically six to eight) housekeeping genes dispersed across the genome of bacteria (or other microorganisms). For each of the selected genes, the different sequences are assigned to different allele numbers that are arbitrarily linked to that specific sequence. Gene sequences that differ by a single nucleotide are given different allelic numbers. The combination of the different allelic numbers assigned to each of the selected genes provides altogether an allelic profile or sequence type (ST). For MLST, the number of nucleotide differences between alleles is not taken into account. The multilocus sequence analysis (MLSA) uses the sequence information for the downstream analysis process (Gevers et al., 2005) and provides a framework for species definition (Bishop et al., 2009) as well as a tool for classification of isolates within a species (Hanage et al., 2006; Bishop et al., 2009). While MLST assigns a particular ST to an isolate, MLSA, using the same sequence information, appraises the molecular diversity and provides an indication of the relatedness of isolates within a particular bacterial species (Glaeser and Kampfer, 2015).

Both methods are particularly useful when isolates from the same bacterial species obtained from different geographical regions and/or hosts are to be compared. From contrasting and correlating sequence data, hypotheses can be drawn on the origins and progression of disease outbreaks (Perez-Losada et al., 2013).

For *X. fastidiosa*, the MLST/MLSA scheme established in 2005 (Scally et al., 2005) comprises the analysis of informative partial sequences of seven housekeeping genes (*leuA*, *petC*, *malF*, *cysG*, *holC*, *nuoL*, *gltT*). The MLST database established for *X. fastidiosa* (http://pubmlst.org/xfastidiosa/) currently contains the sequences of 293 *X. fastidiosa* isolates, including 249 from the United States, 33 from Costa Rica, 6 from Italy, 4 from Brazil and 1 from Mexico that taken together define a total of 76 STs. The database provides the opportunity for typing isolates and strains and allows categorising and cataloguing the diversity within bacterial species.

MLST showed that a coffee-intercepted *X. fastidiosa* subsp. *pauca* was classified as ST73, sharing only four of seven loci with ST53, highlighting the power of MLST to determine isolate diversity and resolving even closely related but not identical genotypes (Loconsole et al., 2016). Similarly, *X. fastidiosa* subsp. *fastidiosa* ST72 from Honduras was identical in six alleles with an ST76 isolate from Costa Rica with the only difference in *petC* (Loconsole et al., 2016). The recently published full genome sequence of the coffee *X. fastidiosa* isolate CO33, classified as ST72 (Giampetruzzi et al., 2015b), then finally revealed the entire complexity of the strain CO33 which is genetically related to isolates of both the subsp. *sandyi* and *fastidiosa* (Loconsole et al., 2016).

Only complete genome sequencing can fully resolve the complexity of *X. fastidiosa* as MLST/MLSA does not have resolution power at the level of the strain, and therefore two isolates cannot be guaranteed to belong to the same strain because they could differ significantly in other regions of the genome that are not covered by the MLST/MLSA analysis. However, MLST can still provide critical information and thus is a useful and robust tool to discriminate between isolates within a population and allows comparative studies. It is currently widely applied to *X. fastidiosa* (Scally et al., 2005; Nunney et al., 2014).

MLST/MLSA is usually done with pure bacterial isolates or with strains purified and proliferated from single colonies. This is to assure that single genetic entities are analysed and to prevent amplification of contaminants or mixed sequences from diverse isolates. MLST analysis based on DNA extracted from infected trees is not a common practice in bacteriology unless uncultivable agents cannot be analysed. This is because samples, and particularly plant samples, are colonised by a diversified microbiota, which could interfere during gene fragment amplification and final allele designation. While it is often not possible to culture bacteria, the validity of MLST/MLSA analysis directly from plant DNA bypassing culturing has to be assessed prior to its use. Similarly, the approaches followed to analyse the sequences (the quality/purity of DNA, the polymerase enzymes used for PCR amplification of the genes, whether direct sequencing of amplicons or sequencing of PCR clones was pursued, the single

or bidirectional orientation of sequencing and the handling of the sequence data to ensure high quality of reads and exclude ambiguous base calls) have significant influences on the quality of the sequencing result. Considering the intrinsic error rate of the method and that single nucleotide differences – also introduced by sequencing errors – can lead to false assignment, the highest prudence has to be applied at all steps of the MLSA/MLST analyses.

3.1.3. Genetic diversity and biological characteristics of *Xylella fastidiosa* bacterial isolates

MLST/MLSA is used to characterise and typify *X. fastidiosa* isolates around the world. While it is a very useful and robust tool for classification of bacterial isolates, the molecular characters studied allow grouping and comparing isolates are not associated with biological features such as virulence, host specificity or host adaptation (Vinatzer and Bull, 2009; Almeida et al., 2010; Perez-Losada et al., 2013; Glaeser and Kampfer, 2015). For example, the *X. fastidiosa* isolates Temecula1 and EB92-1 are both classified as *X. fastidiosa* subsp. *fastidiosa* ST1 (Jacques et al., 2016). However, while Temecula 1, the type strain of this subspecies (Van Sluys et al., 2003) was isolated in California during an epidemics of Pierce's disease in grapes, the isolate EB92-1 from elderberry colonises grapevines asymptomatically and was even proposed as biocontrol against Pierce's disease (Hopkins, 2005; Zhang et al., 2011, 2015).

At this time, there is only very limited information available from studies on virulence and host specificity of *X. fastidiosa* isolates (Almeida and Purcell, 2003; Lopes et al., 2010; Oliver et al., 2014, 2015). The study by Lopes et al. (2010) with *X. fastidiosa*-infecting alfalfa shows that the origin (host) of the strain is driving symptom expression and disease severity. Thus pending experimental evidence, the correlation between a particular ST and biological characteristics of the respective bacterial isolate cannot be drawn. Although the ST53 isolates identified in Apulia on different host plants can be considered genetically highly similar, only when experimental information from bioassays (including olives) becomes available, it will be possible to assess the diversity of biological features inherent in ST53 isolates. This would require transmission assays of ST53 isolates from various hosts to olives to monitor incidence of the disease and severity of symptoms. As shown in Zhang et al. (2015) comparing Temecula1 and EB92-1 isolates, a concurrent analysis of entire bacterial genomes may provide the information necessary to link phenotype (observable biological features) with genotype.

3.1.4. Genetic diversity of *Xylella fastidiosa* from interceptions of plant materials

The reported discoveries of *X. fastidiosa* in Europe so far have originated from: (1) interceptions of imported plant materials and (2) recent outbreaks in France and Italy.

It has to be emphasised that interceptions of *X. fastidiosa* in plant materials from consignments or plants for planting grown in nurseries are not automatically connected to a disease occurring in open fields and natural environments in Europe. A key plant species that represents a risk for introduction of *X. fastidiosa* into Europe is coffee (*Coffea arabica* L., *C. canephora* Pierre ex Froehner) as evidenced by interceptions of *X. fastidiosa* in France, Germany, Italy and the Netherlands on contaminated coffee plants from various countries of origin (Bergsma-Vlami et al., 2015; EFSA 2015; Jacques et al., 2016; Loconsole et al., 2016) and its diversity. The origins of the coffee interceptions reflect the diversity of the bacteria. *X. fastidiosa* subsp. *sandyi* ST72 and ST76 were detected in coffee from Honduras and Costa Rica, respectively, *X. fastidiosa* subsp. *fastidiosa/sandyi* ST75 was isolated from a coffee plant from Mexico, and *X. fastidiosa* subsp. *pauca* ST74 and ST53 were found in coffee plants from Ecuador and in a coffee plant of unknown origin, respectively (Jacques, 2016; Jacques et al., 2016). MLST showed diversity among the isolates, even adding new ST designations. In several cases, identical STs were assigned to isolates intercepted in France and in Italy, indicating potentially similar pathways of introduction and further spread of the pathogen (Jacques, 2016). The ST53 involved in the Apulian outbreak was also intercepted in France (Jacques, 2016). The genetic diversity documented for the intercepted plants infected by *X. fastidiosa* indicates that *X. fastidiosa* may arrive in Europe from different origins. Bacterial isolates that were identified by STs assigned to different subspecies and even the identification of atypical, recombinant isolates (CO33) reveal the high genetic diversity of *X. fastidiosa* in plant materials originating from various countries.

However, in northern Italy, where *X. fastidiosa* interceptions were made, there are no records of spread to open fields and the environment. Only in Apulia the outbreak of *X. fastidiosa* took place and the epidemic spread is connected to the sequence type ST53.

3.1.5. Genetic diversity of *Xylella fastidiosa* associated with the disease outbreak in Apulia

The most recent diversity studies with samples from olives and other plants naturally infected with *X. fastidiosa* over a large area in the Apulia region indicated that all *X. fastidiosa* isolates collected and analysed by MLST belong to a single sequence type, ST53 (Loconsole et al., 2016). These results confirm earlier results from studies where the genetic homogeneity conclusion was based on fewer samples and sites (Elbeaino et al., 2014; Loconsole et al., 2014; Saponari et al., 2014) (Table 1).

ST53 was defined during the MLST analysis of 10 *X. fastidiosa* isolates from Costa Rica (Nunney et al., 2014). Six of the ten DNA sequences, one from coffee and five from oleander (collected from three different sites), were classified as isolates of *X. fastidiosa* subsp. *pauca*, a subspecies previously described based on Brazilian isolates from citrus and coffee (Almeida et al., 2008; Nunney et al., 2012). The *cysG* and *nuoL* DNA sequences of these isolates differed from other alleles found in the database (http://pubmlst.org/xfastidiosa/), and this lead to a description of these six isolates as a novel Central American form of *X. fastidiosa* (Nunney et al., 2014).

Table 1: Studies of genetic diversity of *Xylella fastidiosa* isolates from the Italian (Apulia) outbreak[(a)]

Host plant	Number of isolates[(b)]	Number of sites	MLST classification	Reference
Catharanthus roseus (L.) G. Don	1[(c)]	1	ST53	Elbeaino et al. (2014)[(d)] Loconsole et al. (2016)
Nerium oleander L.	1[(c)]	1	ST53	Elbeaino et al. (2014)[(d)] Loconsole et al. (2016)
Olea europaea L.	2	NA[(e)]	ST53	Elbeaino et al. (2014)[(d)]
Olea europaea L.	2[(c)]	NA[(e)]	ST53	Elbeaino et al. (2014)[(d)] Loconsole et al. (2016)
Olea europaea L.	7	≥ 5	ST53	Loconsole et al. (2016)
Olea europaea L.	1	1	ST53	Bleve et al. (2016)
Polygala myrtifolia L.	1	1	ST53	Loconsole et al. (2016)
Prunus amygdalus Batsch	1[(c)]	1	ST53	Elbeaino et al. (2014)[(d)] Loconsole et al. (2016)
Prunus avium (L.) L.	1	1	ST53	Loconsole et al. (2016)
Westringia fructicosa (Willd.) Druce	1	1	ST53	Loconsole et al. (2016)

(a): Only manuscripts using MLST for classification were considered.
(b): Isolates are considered here according to the respective manuscript where the information was retrieved, that means not only bona fide isolates, but also DNA extracted directly from plants (samples).
(c): The same MLST sequences from five isolates were used in both studies of Elbeaino et al. (2014) and Loconsole et al. (2016): ALM-1 from *Prunus amygdalus*, OLDR-1 from *Nerium oleander*, OLG-2 and KM13 from *Olea europaea*, and PW-1 from *Catharanthus roseus*.
(d): From the manuscript (Elbeaino et al., 2014), it appears that MLST was performed on DNA extracted from olive tree material, and on 11 colonies that were grown from one olive tree sample. No information was provided concerning sampling from other plants.
(e): NA: information not available in the corresponding manuscript.

The first MLST analysis that identified the Apulian *X. fastidiosa* as ST53 was published by Elbeaino et al. (2014). Sequences of the five isolates/samples that were analysed in the study (OLG2, KM13, OLDR-1, PW-1, ALM1) were also subjected to the MLST analysis conducted by Loconsole et al. (2016) confirming their classification as ST53. In Elbeaino et al. (2014), seven isolates from olive (KM13, OLG2, Xf6, Xf9), oleander (OLDR-1), periwinkle (PW1) and almond (ALM1) were assigned to the profile ST53 (table 3 in Elbeaino et al., 2014). This conclusion should result in identical concatenated 4,161 base pair (bp) sequences for all analysed isolates. In contrast, the authors indicated that '*the sequences obtained from three different olive trees showed slight nucleotide divergence* (0.2%) *compared to "Xf9" and accordingly one ST, denoted "Xf6", was considered for analysis*'. Despite these differences, the Xf6 genotype and Xf9 were still assigned as ST53 (table 3 in Elbeaino et al., 2014). The phylogram (figure 2a in Elbeaino et al., 2014) based on the Neighbour Joining clustering method also suggests diversity among the six *X. fastidiosa* subsp. *pauca* strain CoDiRO isolates from olive, almond and periwinkle. The nucleotide diversity, which can be interpreted from the genetic distance

indicated in the figure, would involve between two and eight nucleotides on the total 4,161 bp (e.g. from 0.05% to 0.20%). Elbeaino et al. (2014) did not reflect on the conflicting evidence provided in their MLST and MLSA analyses and also did not discuss the diversity of *X. fastidiosa* subsp. *pauca* isolates highlighted in MLSA.

To tackle the conflicting evidence provided by Elbeaino et al., 2014 and to shed more light on the diversity of the genes analysed the sequences of Apulian isolates and of the DNA sample Xf6 that were deposited in the MLST databases were re-analysed for this report (Appendix A). The isolates KM13, OLG2, PW1, ALM1 and CoDiRO present the expected sequences corresponding to ST53. In contrast, only Xf6 and Xf9 from olive present divergent sequences at some of the seven loci. Xf9 presents a total of six SNPs (single nucleotide polymorphisms) on loci *malF*, *cysG*, *holC* and *gltT* in comparison with the expected sequences for ST53, while Xf6 shows differences at six loci (*leuA*, *petC*, *malF*, *cysG*, *holC* and *gltT*) and a total of 13 SNPs (not counting the 72 missing nucleotides) were recorded in comparison with ST53. Not considering the missing nucleotides, the two isolates would define novel STs, since loci differing by a single nucleotide are attributed to different allelic numbers. However, the MLST/MLSA analysis provided by Elbeaino et al. (2014) is contradictory because the conclusion to assign ST53 to the two isolates is not supported by the MLSA results. Divergence between Xf9 and ST53 is 0.14%, and somewhat higher between Xf6 and ST53 (0.31%, not taking into account missing data). Among the two alleles that are typical of ST53 (i.e. *cysG* and *nuoL*, Nunney et al., 2014), it should be noticed that both Xf6 and Xf9 present the expected *nuoL* allele. In contrast, the *cysG* allele is identical in Xf6 and Xf9 and differs by two SNPs from the expected ST53 allele (a T instead of an A at position 326, and a G instead of a C at position 485). These differences were observed for Xf9 and Xf6 only and this is somehow coherent with the diversity highlighted among Apulian isolates in the phylogenetic tree provided by Elbeaino et al. (2014). However, it is not conceivable how identical sequences for KM13, OLG2, PW1 and ALM1 can result in a phylogram indicating diversity (figure 2a in Elbeaino et al., 2014).

As stated above, the sequences of Xf9 and Xf6 in the pubMLST database represent diverging STs notwithstanding the erroneous assignation to ST53 in the Elbeaino et al. (2014) paper. This would qualify for two distinct strains different from KM13, OLG2, PW1, ALM1 and CoDiRO to which ST53 is assigned. However, the Panel identified several weaknesses in the methods used as well as their inaccurate presentation in the paper that question the validity of the sequence data obtained and would warrant resequencing of the isolates/strains if they would exist. The arguments essentially indicating a lack of prudence and confusing analysis are summarised:

1) Attempts to isolate the bacterium were made from petioles or midribs of 58 infected olive trees. *Xylella fastidiosa* was then cultured from only one of the 58 PCR-positive samples processed resulting in one isolate of which 11 colonies chosen randomly were tested by PCR. [...] *all locus sequences generated from PCR amplifications were identical to those of* X. fastidiosa *present in infected olive plant from which the bacterium was isolated*.

It is not stated whether the DNA of 11 colonies was combined and subjected to MLST/MLSA or separate analyses were made on DNA isolated from single colonies. Since the 11 colonies could essentially represent 11 'different' isolates, the combined DNA analysis could result in ambiguous sequences. There is no further information and since raw data are also not available, statements on identity with original plant DNA are not supported. More significantly, the sequences of the bacterial colonies were neither assigned to a *X. fastidiosa* accession nor isolate or strain.

2) 'In addition, sequences were also obtained from other 10 *X. fastidiosa*-infected plants. Accordingly, one representative ST, denoted Xf9, was analysed. Although the *X. fastidiosa* genome sequences obtained were significantly homogeneous...'.

No indication is provided whether the DNA sequences from the plant samples were combined with the sequences of the colonies and assigned to Xf9 or only *X. fastidiosa* DNA from infected plants comprises Xf9. This would mean that an environmental sample was combined with an isolate or multiple isolated bacterial colonies.

3) PCR amplicons were cloned and four independent clones were sequenced and a consensus sequence was determined based on each from four independent clones.

No statement is made on the orientation of sequencing, no further statement on how the consensus sequences were obtained. As independent DNA clones can also represent distinct sequences, a deriving consensus sequence of four independent clones can also result in sequence

ambiguities. It is unclear how this analysis was made. Notwithstanding the missing nucleotides at the ends of the DNA sequences, the loose description concerning the analyses, the number of samples singly analysed or in combination raises concern on the quality of the sequence analyses. It also indicates that these sequences were not obtained following the classical protocol (Scally et al., 2005; Yuan et al., 2010).

Thus it is not evident how the sequences of Xf9 and Xf6 were obtained. It is obvious that Xf6 is neither an isolate nor a strain as it represents a DNA sample from infected olive trees. To confirm its existence as an isolate/ strain and to identify diversity, isolation attempts would need to be made from the original tree(s). The association of the Xf9 sequence with a bacterial entity also is unclear as it is not evident whether a consensus sequence was derived from amplicons obtained from 11 bacterial colonies either analysed separately or with sequences obtained from an environmental sample. Finally, from the description of the sequence analysis, it is not evident that strategies were followed to identify single nucleotide exchanges and to eliminate false signals. Hence, considering the major weaknesses highlighted here, the validity of the evidence on Xf9 and Xf6 diversity presented by Elbeaino et al. (2014) has to be questioned. Because of the uncertainty that both Xf9 and Xf6 might be derived from sequencing errors and because they might not exist as bacterial entities, pending further investigation, these sequences should not be considered in diversity studies on *X. fastidiosa*.

Further approaches to reveal potential diversity among isolates of *X. fastidiosa* associated with the Apulian outbreak were conducted by different research groups in Italy. In the research report by Bleve et al. (2016), a strain, 'Salento-1' identified by MLST as ST53, was described. In addition to the sequences generally used for MLST, the authors used the sequences of two further genes, the polymerase sigma-70 factor (*rpoD*) and the chromosomal replication initiation protein DnaA (*dnaA*). The *rpoD* sequence analysis highlighted a repetitive sequence motif characteristic of subsp. *pauca* isolates from citrus and coffee (Chen et al., 2005) not present in the Apulian *X. fastidiosa* subsp. *pauca* isolates. However, the analysis did not reveal any further diversity within the subspecies. In the diversity study of Mang et al. (2016), instead of applying MLST, the sequence of gyrase B (*gyrB*, 384 bp), the polymerase sigma-70 factor (*rpoD*, 733 bp) and that of a fragment of the hypothetical protein HL gene (216 bp) were analysed. While a number of SNPs were identified in the entire data set, there were no differences in the sequences amplified from DNA isolated from Apulian olive trees. Since MLST/MLSA was not done for ST designation and all sequences essentially were identical, this study did not bring any additional information regarding a possible diversity of *X. fastidiosa* sequences in Apulian olives.

In summary, the currently available information from scientific publications does not support the notion that further diverse populations exist in Apulia in addition to the isolates/strains which were assigned to ST53. The Panel could observe that there is no evidence for a wide genetic diversity of *X. fastidiosa* in Apulia as all isolates tested and reported so far belong to ST53. The most recent publications (Jacques et al., 2016; Loconsole et al., 2016), while not providing definite answers, show that the currently accepted MLST approach (Almeida and Nunney, 2015) provides a framework robust enough to appraise/evaluate *X. fastidiosa* diversity, which then can be subjected to complete genome analysis to resolve differences. A large and robust evaluation of *X. fastidiosa* diversity in Apulia should be made to evaluate the current occurrence of these potential variants or to rule out any variation possibly linked to the sequencing strategy adopted.

Recent interception data (Loconsole et al., 2016) emphasise the serious threat from introduction of new isolates of *X. fastidiosa*. Considering that only the spatial and temporary interference of the pathogen with a highly efficient insect vector is needed for an outbreak with dramatic consequences, new diversity from further introductions presents a serious and unpredictable threat.

4. Conclusions

The PLH Panel was requested to provide a scientific opinion on the presence of heterogeneous populations of *X. fastidiosa* subsp. *pauca* associated with the Apulian outbreaks, in addition to the *X. fastidiosa* subsp. *pauca* 'strain' associated with the CoDiRO disease of olives. The Panel evaluated the evidence provided in publically available scientific papers and conducted a further analysis on sequence data available at the MLST database (Appendix A).

The Panel recognised that the analyses of *X. fastidiosa* based on MLST/MLSA were conducted with isolates/strains, as well as with DNA directly obtained from infected plants. Allelic profiles obtained for *X. fastidiosa* subsp. *pauca* identified a single ST53 present in Apulia. One publication (Elbeaino et al., 2014) while categorising Xf9 (isolates) and Xf6 (DNA from infected olives) as ST53

still highlighted diversity of *X. fastidiosa* sequences using MLSA on the same data set. The Panel identified several major shortcomings in this publication, contradictory evidence and interpretation, questioning the validity of the data obtained. Therefore, the Panel concludes that from data provided by MLST/MLSA, there is no evidence to corroborate further heterogeneous populations of *X. fastidiosa* subsp. *pauca* in Apulia in addition to those represented by ST53.

The Panel also wishes to highlight that to date, only a limited number of DNA samples, bacterial isolates or strains of *X. fastidiosa* were analysed by MLST/MLSA. Sequences of the seven housekeeping genes used in MLST are currently available for 18 Apulian isolates only. The meta-data associated with the available data are not sufficient to unequivocally associate an isolate to a host and to a ST. More data are needed for a robust testing of hypothesis concerning strain diversity, strain evolution and routes of introduction. To allow the spatio-temporal analysis of the epidemics, the data would need to include: (i) location of the samples analysed; (ii) data on host (species and cultivar); (iii) isolation, identification and culturing of bacterial isolates, and their availability in a reference collection; and (iv) MLST analysis. The limitations of the data found in this opinion deserve further studies on larger sample sets, including the analyses of complete *X. fastidiosa* genomes, to provide more comprehensive answers on the diversity of the pathogen.

References

Almeida RPP and Nunney L, 2015. How do plant diseases caused by *Xylella fastidiosa* emerge?. Plant Disease, 99, 1457–1467.

Almeida RPP and Purcell AH, 2003. Biological traits of *Xylella fastidiosa* strains from grapes and almonds. Applied and Environmental Microbiology, 12, 7447–7452.

Almeida RPP, Nascimento FE, Chau J, Prado SS, Tsai C-W, Lopes SA and Lopes JRS, 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Applied and Environmental Microbiology, 74, 3690–3701.

Almeida NF, Yan S, Cai R, Clarke CR, Morris CE, Schaad NW, Schuenzel EL, Lacy GH, Sun X, Jones JB, Castillo JA, Bull CT, Leman S, Guttman DS, Setubal JC and Vinatzer BA, 2010. PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. Phytopathology, 100, 208–215.

Bergsma-Vlami M, van de Bilt JLJ, Tjou-Tam-Sin NNA, van de Vossenberg BTLH and Westenberg M, 2015. *Xylella fastidiosa* in *Coffea arabica* ornamental plants imported from Costa Rica and Honduras in the Netherlands. Journal of Plant Pathology, 97, 391–403.

Bishop CJ, Aanensen DM, Jordan GE, Kilian M, Hanage WP and Spratt BG, 2009. Assigning strains to bacterial species via the internet. BMC Biology, 7, 20.

Bleve G, Marchi G, Ranaldi F, Gallo A, Cimaglia F, Logrieco AF, Mita G, Ristori J and Surico G, 2016. Molecular characteristics of a strain 9salento-1) of *Xylella fastidiosa* isolated in Apulia (Italy) from an olive plant with the quick decline syndrome. Phytopathologia Mediterranea, 55, 139–146.

Bull CT and Koike ST, 2015. Practical benefits of knowing the enemy: modern molecular tools for diagnosing the etiology of bacterial diseases and understanding the taxonomy and diversity of plant-pathogenic bacteria. Annual Review of Phytopathology, 53, 157–180.

Cariddi C, Saponari M, Boscia D, De Stradis A, Loconsole G, Nigro F, Porcelli F, Potere O and Martelli GP, 2014. Isolation of *Xylella fastidiosa* strain infecting olive and oleander in Apulia, Italy. Journal of Plant Pathology, 96, 425–429.

Chen J, Groves R, Civerolo EL, Viveros M, Freeman M and Zheng Y, 2005. Two *Xylella fastidiosa* genotypes associated with almond leaf scorch disease on the same location in California. Phytopathology, 95, 708–714.

Darrasse A, Carrère S, Barbe V, Boureau T, Arrieta-Ortiz ML, Bonneau S, Briand M, Brin C, Cociancich S, Durand K, Fouteau S, Gagnevin L, Guérin F, Guy E, Indiana A, Koebnik R, Lauber E, Munoz A, Noël LD, Pieretti I, Poussier S, Pruvost O, Robène-Soustrade I, Rott P, Royer M, Serres-Giardi L, Szurek B, Van Sluys M-A, Verdier V, Vernière C, Arlat M, Manceau C and Jacques M-A, 2013. Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. BMC Genomics, 14, 761.

EFSA (European Food Safety Authority), 2009. Guidance of the Scientific Committee on transparency in the scientific aspects of risk assessments carried out by EFSA. Part 2: General Principles. EFSA Journal 2009; 7(5):1051, 22 pp. doi:10.2903/j.efsa.2009.1051

EFSA (European Food Safety Authority), 2015. Categorization of plants for planting, excluding seeds, according to the risk of introduction of *Xylella fastidiosa*. EFSA Journal 2015;13(3):4061, 31 pp. doi:10.2903/j.efsa.2015.4061

EFSA PLH Panel (EFSA Panel on Plant Health), 2015. Scientific Opinion on the risk to plant health posed by *Xylella fastidiosa* in the EU territory, with the identification and evaluation of risk reduction options. EFSA Journal 2015;13(1):3989, 262 pp. doi:10.2903/j.efsa.2015.3989

EFSA Scientific Committee, 2014. Guidance on the structure and content of EFSA's scientific opinions and statements. EFSA Journal 2014;12(9):3808, 10 pp. doi:10.2903/j.efsa.2014.3808

Elbeaino T, Valentini F, Abou Kubaa R, Moubarak P, Yaseen T and Digiaro M, 2014. Multilocus sequence typing of *Xylella fastidiosa* isolated from olive affected by "olive quick decline syndrome" in Italy. Phytopathologia Mediterranea, 53, 533–542.

Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL and Swings J, 2005. Reevaluating prokaryotic species. Nature Reviews, 3, 733–739.

Giampetruzzi A, Chiumenti M, Saponari M, Donvito G, Italiano A, Loconsole G, Boscia D, Cariddi C, Martelli GP and Saldarelli P, 2015a. Draft genome sequence of the *Xylella fastidiosa* CoDiRO strain. Genome Announcements, 3, e01538–14.

Giampetruzzi A, Loconsole G, Boscia D, Calzolari A, Chiumenti M, Martelli GP, Saldarelli P, Almeida RPP and Saponari M, 2015b. Draft genome sequence of CO33, a coffee-infecting isolate of *Xylella fastidiosa*. Genome Announcements, 3, e01472–15.

Glaeser SP and Kampfer P, 2015. Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. Systematic and Applied Microbiology, 38, 237–245.

Hanage WP, Fraser C and Spratt BG, 2006. Sequences, sequence clusters and bacterial species. Philosophical Transactions of the Royal Society B, 361, 1917–1927.

Hopkins DL, 2005. Biological control of Pierce's disease in the vineyard with the strains of *Xylella fastidiosa* benign to grapevine. Plant Disease, 89, 1348–1352.

Jacques M-A, 2016. A sap alien: breaking news of *Xylella fastidiosa*, an emerging plant pathogen in France. International Training Workshop on *Xylella fastidiosa*, Bari, Italy, 19–22 April 2016. Available from: http://www.neppo.org/wp-content/uploads/2016/05/5-MA-Jacques.pdf

Jacques M-A, Denance N, Legendre B, Morel E, Briand M, Mississipi S, Durand K, Olivier V, Portier P, Poliakoff F and Crouzillat D, 2016. New variants of coffee-infecting *Xylella fastidiosa* issued from homologous recombination. Applied and Environmental Microbiology, 82, 1556–1568.

Konstantinidis KT and Tiedje JM, 2005. Genomic insights that advance the species definition for prokaryotes. Proceedings of the National Academy of Sciences of the United States of America, 120, 2567–2572.

Loconsole G, Potere O, Boscia D, Altamura G, Djelouah K, Elbeaino T, Frasheri D, Lorusso D, Palmisano F, Pollastro P, Silletti MR, Trisciuzzi N, Valentini F, Savino V and Saponari M, 2014. Detection of *Xylella fastidiosa* in olive trees by molecular and serological methods. Journal of Plant Pathology, 96, 7–14.

Loconsole G, Saponari M, Boscia D, D'Attoma G, Morelli M, Martelli GP and Almeida RPP, 2016. Intercepted isolates of *Xylella fastidiosa* in Europe reveal novel genetic diversity. European Journal of Plant Pathology, doi:10.1007/s10658-016-0894-x

Lopes JRS, Daugherty MP and Almeida RPP, 2010. Strain origin drives virulence and persistence of *Xylella fastidiosa* in alfalfa. Plant Pathology, 59, 963–971.

López-Villavicencio M, Jonot O, Coantic A, Hood ME, Enjalbert J and Giraud T, 2007. Multiple infections by the anther smut pathogen are frequent and involve related strains. PLoS Pathogens, 3, e176.

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IA, Achtman M and Spratt BG, 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America, 95, 3140–3145.

Mang AM, Frisullo S, Elshafie HS and Camele I, 2016. Diversity evaluation of *Xylella fastidiosa* from infected olive trees in Apulia (southern Italy). Plant Pathology Journal, 32, 1–10.

Nunney l, Yuan X, Bromley RE and Stouthamer R, 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Applied and Environmental Microbiology, 78, 4702–4714.

Nunney L, Ortiz B, Russell SA, Ruiz Sanchez R and Stouthamer R, 2014. The complex biogeography of the plant pathogen *Xylella fastidiosa*: genetic evidence of introductions and subspecific introgression in Central America. PlosOne, 9, e112463.

Oliver JE, Sefick SA, Parker JK, Arnold T, Cobine PA and De La Fuente L, 2014. Ionome changes in *Xylella fastidiosa*-infected *Nicotiana tabacum* correlate with virulence and discriminate between subspecies of bacterial isolates. Molecular Plant-Microbe Interactions, 27, 1048–1058.

Oliver JE, Cobine PA and De La Fuente L, 2015. *Xylella fastidiosa* isolates from both subsp. *multiplex* and *fastidiosa* cause disease on southern highbush blueberry (*Vaccinium* sp.) under greenhouse conditions. Phytopathology, 105, 855–862.

Perez-Losada M, Cabezas P, Castro-Nallar E and Crandall KA, 2013. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. Infection, Genetics and Evolution, 16, 38–53.

Saponari M, Boscia D, Loconsole G, Palmisano F, Savino V, Potere O and Martelli GP, 2014. New hosts of *Xylella fastidiosa* strain CoDiRO in Apulia. Journal of Plant Pathology, 96, 611.

Saponari M, Boscia D, Altamura G, D'Attoma G, Cavalieri V, Loconsole G, Zicca S, Dongiovanni C, Palmisano F, Susca L, Morelli M, Potere O, Saponari A, Fumarola G, Di Carolo M, Tavano D, Savino V and MArtelli P, 2016. Pilot project on *Xylella fastidiosa* to reduce risk assessment uncertainties. EFSA supporting publication 2016: EN-1013. 60 pp.

Scally M, Schuenzel EL, Stouthamer R and Nunney L, 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Applied and Environmental Microbiology, 71, 8491–8499.

Staley JT and Krieg NR, 1984. Classification of prokaryotic organisms: an overview. In: Krieg NR and Holt JG (eds.). *Bergey's Manual of Systematic Bacteriology*. Vol. 1. Williams & Wilkins, Baltimore, MD, USA. pp. 1–4.

Susi H, Barres B, Vale PF and Laine AL, 2015. Co-infection alters population dynamics of infectious disease. Nature Communications, 6. doi:10.1038/ncomms6975

Tollenaere C, Susi H, Nokso-Koivisto J, Koskinen P, Tack A, Auvinen P, Paulin L, Frilander MJ, Lehtonen R and Laine A-L, 2012. SNP design from 454 sequencing of *Podosphaera plantaginis* transcriptome reveals a genetically diverse pathogen metapopulation with high levels of mixed-genotype infection. PLoS ONE, 7, e52492.

Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LEA, da Silva ACR, Moon DH, Takita MA, Lemos EGM, Machado MA, Ferro MIT, da Silve FR, Goldman MHS, Goldman GH, Lemos MVF, El-Dorry H, Tsai SM, Carrer H, Carrao DM, de Oliveira RC, Nunes LR, Siqueira WJ, Coutinho LL, Kimura ET, Ferro ES, Haravaka R, Kuramae EE, Marino CL, Giglioti E, Abreu IL, Alves LMC, do Amaral AM, Baia GS, Blanco SR, Brito MS, Cannavan FS, Celestino AV, da Cunha AF, Fenille RC, Ferro JA, Formighieri EF, Kishi LT, Leoni SG, Oliveira AR, Rosa VE Jr, Sassaki FT, Sena JAD, de Souza AA, Truffi D, Tsukumo F, Yanai GM, Zaros LG, Civerolo EL, Simpson AJG, Almeida NF Jr, Setubal JC and Kitajima JP, 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. Journal of Bacteriology, 185, 1018–1026.

Vinatzer BA and Bull CT, 2009. The impact of genomic approaches on our understanding of diversity and taxonomy of plant pathogenic bacteria. In: Jackson RW (ed.). *Plant Pathogenic Bacteria: Genomics and Molecular Biology*. Caister Academic Press, School of Biological Sciences, University of Reading, Whiteknights, Reading, UK. pp 37–62.

Wei W, Davis W, Davis RE, Lee IM and Zhao Y, 2016. Development of molecular markers and a diagnostic tool for investigation of coinfections by and interactions between potato purple top and potato witches'-broom phytoplasmas in tomato. Annals of Applied Biology, 168, 133–141.

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

Zhang S, Flores-Cruz Z, Kumar D, Chakrabarty P, Hopkins DL and Gabriel DW, 2011. The *Xylella fastidiosa* biocontrol strain EB92-1 genome is very similar and syntenic to Pierce's disease strains. Journal of Bacteriology, 193, 5576–5577.

Zhang S, Chakrabarty PK, Fleites LA, Rayside PA, Hopkins DL and Gabriel DW, 2015. Three new Pierce's disease pathogenicity effectors identified using *Xylella fastidiosa* biocontrol strain EB92-1. PLoS ONE, 10, e0133796.

Abbreviations

ANI	average nucleotide identity
CFU	colony forming unit
CoDiRO	Complesso del Disseccamento Rapido dell'Olivo
MLSA	multilocus sequence analysis
MLST	multilocus sequence typing
MS	Member State
OQDS	Olive Quick Decline Syndrome
PCR	polymerase chain reaction
PLH Panel or the Panel	EFSA Panel on Plant Health
ST	sequence type
SNP	single nucleotide polymorphisms
ToR	Terms of Reference

Appendix A – Comparison of sequences of Apulian isolate-strain housekeeping genes

This appendix describes the sequences analysis of the seven housekeeping genes used for MLST/MLSA of *Xylella fastidiosa* conducted by Marie-Agnès Jacques, a member of the working group on urgent review of statements on *X. fastidiosa* and its control.

A.1. Methodology

Sequences of *X. fastidiosa* housekeeping genes were downloaded from NCBI with accession numbers provided in Elbeaino et al. (2014), Bleve et al. (2016) and Loconsole et al. (2016). *X. fastidiosa* allele sequences were downloaded from pubMLST database (http://pubmlst.org/xfastidiosa/) (Table A.1). The allele sequences of the CoDiRO strain were retrieved from its genome sequence (Giampetruzzi et al., 2015a). To fit the expected sizes and positions, alleles of *holC*, *leuA*, *nuoL* and *petC* of Xf6 strain were trimmed. 'N' was added at 3' and 5' positions of *malF* and *holC* sequences, respectively, to replace missing sequences. Sequences were analysed using Geneious Pro 4.8.5 software as described in Jacques et al. (2016).

Table A.1: Accession numbers for the seven housekeeping gene fragments available in pubMLST database for Apulian strains

Strain/ genotype	*leuA*	*petC*	*malF*	*cysG*	*holC*	*nuoL*	*gltT*
ALM1	KJ406220	KJ406226	KJ406232	KJ406256	KJ406238	KJ406244	KJ406250
KM13	KJ406218	KJ406224	KJ406230	KJ406254	KJ406236	KJ406242	KJ406248
OLG2	KJ406216	KJ406222	KJ406228	KJ406252	KJ406234	KJ406240	KJ406246
PW1	KJ406221	KJ406227	KJ406233	KJ406257	KJ406239	KJ406245	KJ406251
Salento-1	KU214453	KU214456	KU214454	KU214450	KU214452	KU214455	KU214451
Xf6	HG939499	HG939497	HG939503	LM999935	HG939495	HG939502	HG939500
Xf9	LM999929	LM999927	LM999933	LM999934	LM999925	LM999932	LM999930
CoDiRO	Accessed from genome sequence at JUJW01000000						

A.2. Results

After alignment and comparison of sequences for strains ALM1, KM13, OLG2, PW1, Salento-1, Xf6, Xf9 and CoDiRO with the sequences of alleles given in pubMLST, ALM1, KM13, OLG2, PW1, Salento-1 and CoDiRO presented alleles that correspond to ST53.

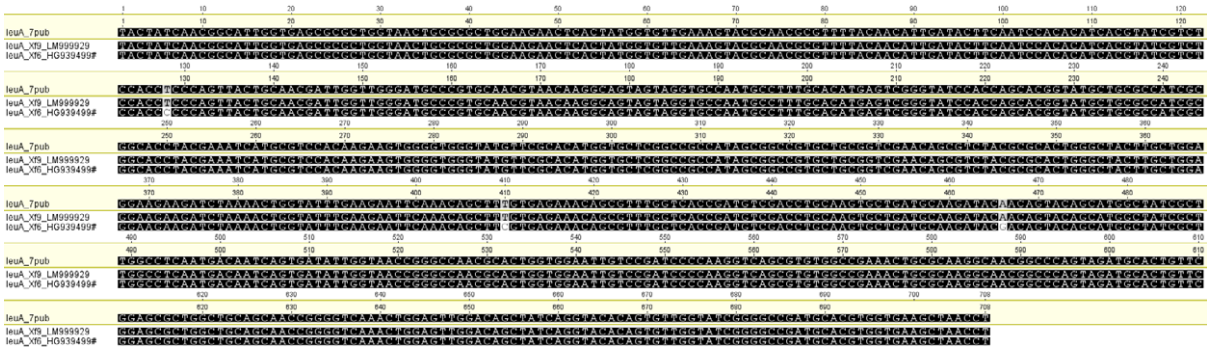
Xf6 and Xf9 did not present all expected alleles for ST53 (Figure A.1). The set of alleles of these two strains were different and differed from ST53. Xf9 presented unexpected sequences at loci *malF*, *cysG*, *holC* and *gltT* with a total of six SNPs, while Xf6 showed unexpected sequences at six out of the seven loci with a total of 13 SNPs, not taking into account the missing bases. The *nuoL* allele was allele #16, as expected in ST53.

According to Nunney et al. (2014), alleles *cysG* #24 and *nuoL* #16 are characteristic of ST53. The other alleles (*leuA* #7, *petC* #6, *malF* #16 (1 bp from #8), *holC* #10 and *gltT* #14 (1 bp from #8)) are typical or slightly different from alleles of known *pauca* strains from Brazil.

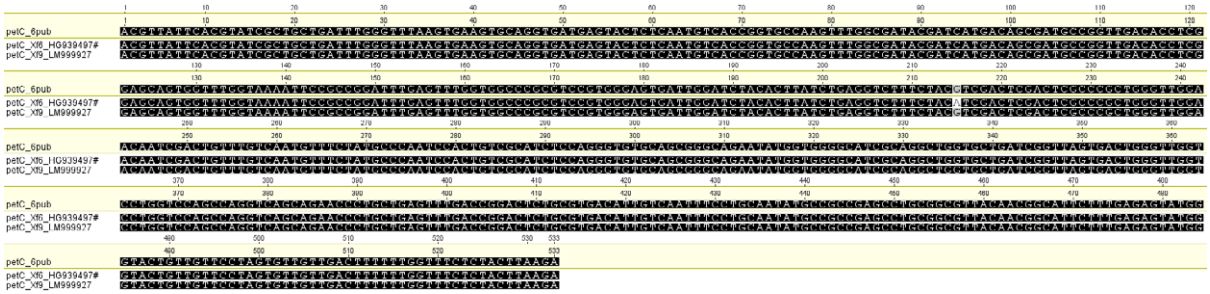
All these strains appeared to cluster in a divergent lineage from the Brazilian strains 6c, 32, 9a5c and CVC0018 (Figure A.2). This is coherent with previously reported positions for these *X. fastidiosa* subsp. *pauca* strains (Jacques et al., 2016; Loconsole et al., 2016). The strains 9a5c and CVC0018 share ST13 and appeared on the tree as identical but different from 32 and 6c. As expected, all three strains of *X. fastidiosa* subsp. *fastidiosa* (GB514, M13 and Temecula1) grouped together as they share the same ST, i.e. ST1.

In conclusion, the sequences deposited for Xf6 and Xf9 in pubMLST are not ST53. They define two novel STs that differ from ST53 at six loci (*cysG*, *gltT*, *holC*, *leuA*, *malF*, and *petC*) by at least 13 SNPS for novel Xf6 and at four loci (*malF*, *cysG*, *holC* and *gltT*) for novel ST Xf9.

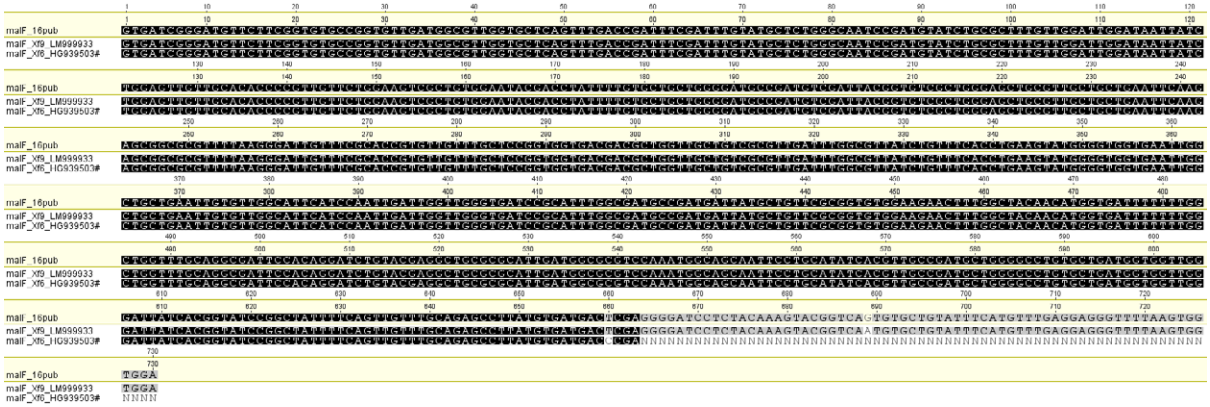
leuA alleles



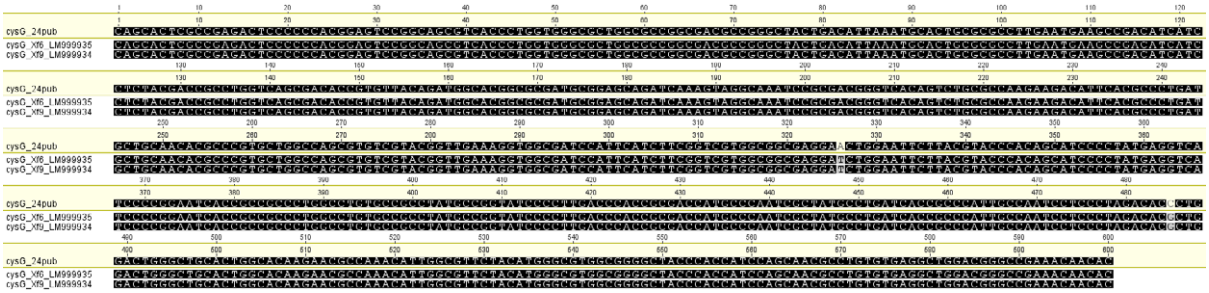
petC alleles



malF alleles

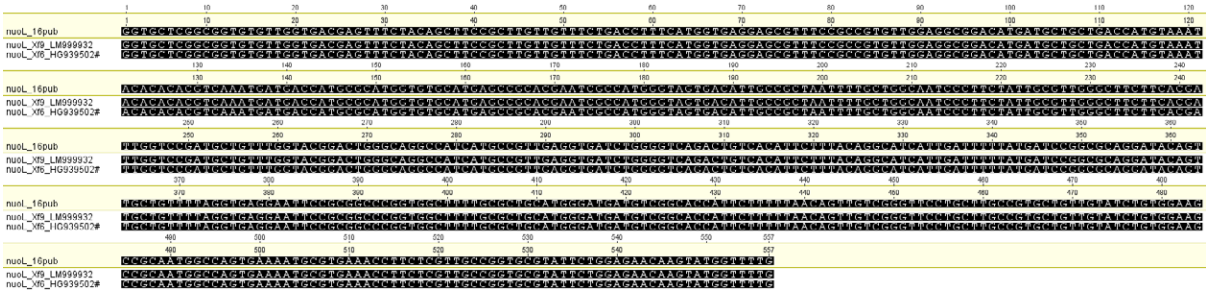


cysG alleles



holC alleles

nuoL alleles



gltT alleles

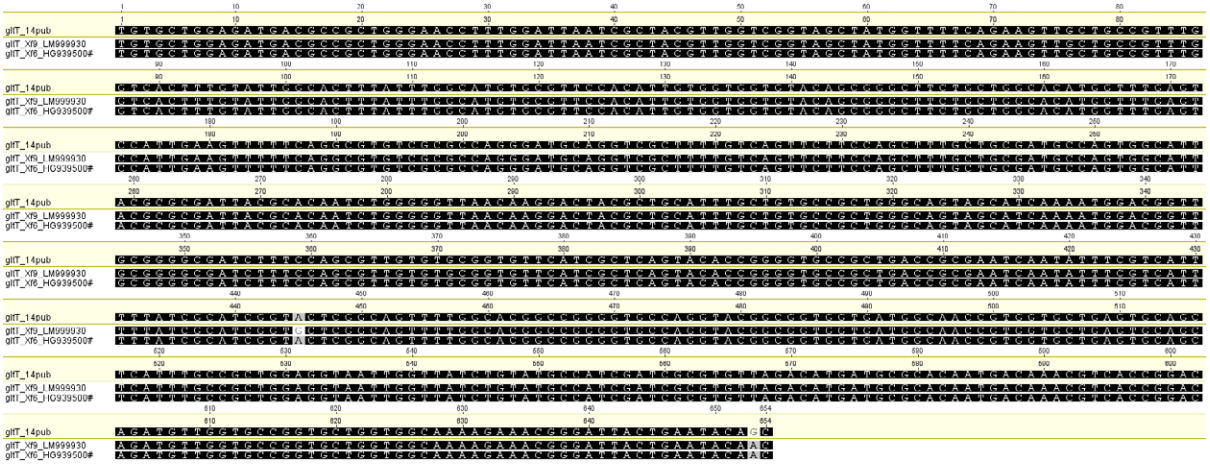


Figure A.1: Alignments of alleles from Xf6, Xf9 and the expected alleles for a ST53. Data were retrieved from NCBI database and pubMLST

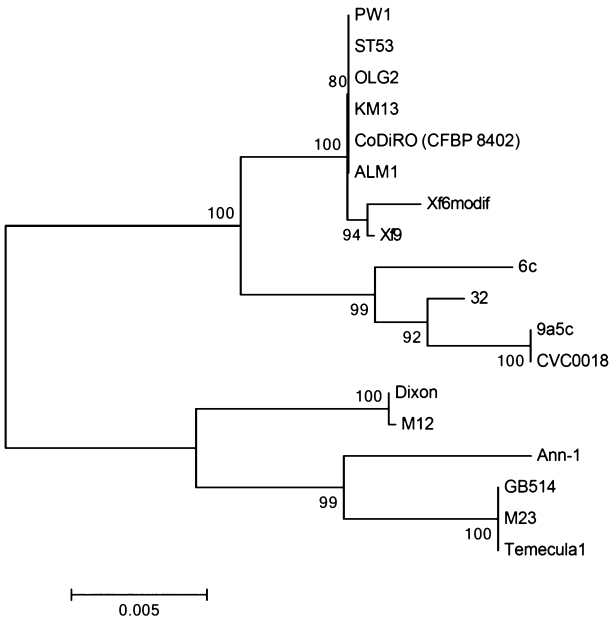


Figure A.2: Maximum likelihood tree based on the concatenated partial sequences of *cysG*, *glT*, *holC*, *leuA*, *malF*, *nuoL*, and *petC*. Bootstrap scores (1,000 replicates) are displayed at each node. ST53 corresponds to the sequences of the seven housekeeping genes for ST53, as they appear in pubMLST. Xf6 is named Xf6modif as sequences were trimmed