

Development of a Nested-MultiLocus Sequence Typing approach for a highly sensitive and specific identification of *Xylella fastidiosa* subspecies directly from plant samples

Sophie Cesbron, Enora Dupas, Quentin Beaupère, Martial Briand, Miguel Montes Borrego, Maria del Pilar Velasco Amo, Blanca Landa, Marie-Agnès Jacques

► **To cite this version:**

Sophie Cesbron, Enora Dupas, Quentin Beaupère, Martial Briand, Miguel Montes Borrego, et al.. Development of a Nested-MultiLocus Sequence Typing approach for a highly sensitive and specific identification of *Xylella fastidiosa* subspecies directly from plant samples. 2020. hal-02886841

HAL Id: hal-02886841

<https://hal.inrae.fr/hal-02886841>

Preprint submitted on 1 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Development of a Nested-MultiLocus Sequence Typing approach for a highly sensitive and specific identification of *Xylella* *fastidiosa* subspecies directly from plant samples

1 Sophie Cesbron¹, Enora Dupas^{1,3}, Quentin Beaupère¹, Martial Briand¹, Miguel
2 Montes Borrego², Maria del Pilar Velasco Amo², Blanca B. Landa², Marie-Agnès
3 Jacques¹

4 ¹ IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207 QuaSaV, 49071,
5 Beaucouzé, France

6 ² Institute for Sustainable Agriculture, Consejo Superior de Investigaciones Científicas (IAS-
7 CSIC), 14004, Córdoba, Spain

8 ³ French Agency for Food, Environmental and Occupational Health & Safety, Plant Health
9 Laboratory, Angers, France

10 Correspondence:

11 Sophie Cesbron

12 sophie.cesbron@inrae.fr

13 **Keywords :** *Xylella fastidiosa*, direct typing technique, MLST, nested PCR

14 Abstract

15 Different sequence types (ST) of *Xylella fastidiosa* were already identified in France
16 and Spain based on direct MultiLocus Sequence Typing (MLST) of plant DNA samples.
17 However, direct typing of plant DNA is partly efficient. In order to improve the sensitivity of

18 *X. fastidiosa* identification, we developed a direct nested-MLST assay on plant extracted
19 DNA. This method was performed based on a largely used scheme targeting seven
20 housekeeping gene (HKG) loci (*cysG*, *gltT*, *holC*, *leuA*, *malF*, *nuoL*, *petC*). Nested primers
21 were designed from multi-sequence alignments of 38 genomes representing all subspecies and
22 one genome of *Xylella taiwanensis*. Sequences obtained were long enough to be used for
23 BLAST comparison in PubMLST database. No nonspecific amplification products were
24 observed in these samples. Efficiency of the nested-MLST was tested on extracted DNA from
25 106 samples proven positive ($Cq < 35$) or equivocal ($35 \leq Cq \leq 40$) using the Harper's qPCR test.
26 Samples analyzed included 49 plant species and two insect species (*Philaenus spumarius*,
27 *Neophilaenus campestris*) that were collected in 2017 (106 plant samples in France), in 2018
28 (162 plant samples in France, 40 plant samples and 26 insect samples in Spain), and in 2019
29 (30 plant samples in Spain). With the conventional-MLST assay, no complete MLST profile
30 was obtained for any of the samples from France and for most samples (59/66) from Spain.
31 Conversely, with the nested approach, complete profiles were obtained for six French plant
32 samples, 55 Spanish plant samples and nine Spanish insect samples. The threshold was
33 improved by 100 to 1000 times compared to conventional PCR and was between 22 pg.mL^{-1}
34 to 2.2 pg.mL^{-1} depending on the HKG. Using nested-MLST assay, plants that were not yet
35 considered hosts tested positive and revealed novel alleles in France, whereas for Spanish
36 samples it was possible to assign the subspecies or ST to samples considered as new hosts in
37 Europe. Direct typing by nested-MLST from plant material has an increased sensitivity and
38 may be useful for epidemiological purposes.

39 **Introduction**

40 *Xylella fastidiosa* (*Xf*) is the causal agent of several devastating diseases of plants in the
41 Americas and this pathogen was recently detected in Europe, where it causes a severe disease
42 in olive tree in Italy and is present in several other regions. This species encompasses three

43 well recognized subspecies, namely *fastidiosa*, *multiplex*, and *pauca* (Marcelletti and
44 Scortichini, 2016; Denancé et al. 2019) but other subspecies are currently described (EFSA,
45 2018). The subspecies *fastidiosa* occurs in North and Central America and was recently
46 detected in Spain (<https://gd.eppo.int/taxon/XYLEFA/>). It infects a wide range of trees,
47 ornamentals, and other perennials and includes strains responsible for the well-known
48 Pierce’s disease on grapevine (Janse & Obradovic 2010; EFSA 2018). The subspecies
49 *multiplex* is present in North and South America and in Europe
50 (<https://gd.eppo.int/taxon/XYLEFA/>) and is associated with scorches and dieback of a wide
51 range of trees and ornamentals (EFSA 2018). The subspecies *pauca* is mostly found in South
52 and Central America on *Citrus* spp. and *Coffea* spp. (Almeida et al. 2008), but has been
53 recently detected also in olive trees in Spain (Landa, 2017), Brazil (Della Coletta-Filho et al.
54 2016), Argentina (Haelterman et al. 2015), and Italy (Saponari et al. 2013). Its host range
55 includes also ornamentals and other trees (EFSA 2018). Altogether more than 560 plant
56 species are hosts of *Xf* (EFSA 2018). This member of the *Xanthomonadaceae* family inhabits
57 the xylem of its host plants (Wells et al. 1987) and is naturally transmitted by insects from
58 plants to plants.

59 A range of detection tests has been proposed for *Xf* (EPPO 2019). Several immunological
60 methods are available (EPPO 2019). However, such methods have high limits of detection
61 (LoDs) that are close to 10^4 to 10^5 cells.mL⁻¹ (EPPO 2019). End point and also quantitative
62 PCR (qPCR) are nowadays widely used, with a better sensitivity as the LoD is around 10^2
63 cells.mL⁻¹ for several qPCR tests (Ouyang et al. 2013, Harper et al. 2010, Bonants et al. 2019,
64 Waliullah et al. 2019). The Harper’s qPCR test is often used in Europe for its high sensitivity,
65 its target is located in the gene coding for the 16S rRNA-processing RimM protein. Several
66 tests based on isothermal amplification have also been reported (Harper et al. 2010, Yaseen et
67 al. 2015, Li et al 2016, Burbank & Ortega 2018, Waliullah et al. 2019). The Harper’s test has

68 also been successfully transferred to be used in digital PCR (Dupas et al. 2019a). Some of
69 these tests were designed to detect only one subspecies. This is the case of the nested PCR test
70 proposed by Ciapina et al (2004) for detecting CVC strains (subspecies *pauca*) in
71 sharpshooters and citrus plants and also of the qPCR test targeting oleander leaf scorch strains
72 (that are included in the subspecies *fastidiosa*) (Guan et al. 2013). Other tests were designed
73 to detect and discriminate two or more subspecies (Burbank & Ortega 2018; Dupas et al.
74 2019b).

75 Precise identification of *Xf* at an infraspecific level is essential for epidemiological and
76 surveillance analyses, and to allow a proper description of the population structure and their
77 dynamics. The widely used multilocus sequence typing (MLST) scheme designed for *Xf*
78 (Sally et al. 2005, Yuan et al. 2010) is based on amplification by conventional PCR and
79 sequencing of seven HKG fragments (loci), either from strains or from plant samples
80 (Denancé et al., 2017). For each locus, the different sequence variants are considered as
81 distinct alleles. The combination of allele numbers defines the sequence type (ST). The
82 MLST-*Xf* data are stored in a public database (<https://pubmlst.org/xfastidiosa/>) that can be
83 used to automatically identify and assign new allele variants, and provide tools to analyze the
84 potential origin of the strains. The association of the different subspecies with their host plants
85 is useful to better understand *Xf* epidemiology.

86 A reliable and enough informative typing method is particularly relevant in cases of new
87 outbreaks or for the description of new host. Due to the large number of host plants to be
88 analyzed, various types of inhibitors can interfere with reagents of PCR and low bacterial
89 loads compromising PCR efficiency and hence typing. Improving DNA extraction methods
90 can, at least partly, solve the problem of PCR inhibitors, and nested PCR appears a solution to
91 allow the detection of low bacterial population sizes. A nested-MLST was already
92 successfully developed to detect and type *Xf* in vectors (Cruaud et al. 2018). Primers were

93 designed inside the gene fragments used in the conventional-MLST scheme and hence some
94 informative sites are lost. MLST with nested PCRs has also been developed in medical field
95 to enable the direct typing of samples infected by *Leptospira* or *Trichomonas*, for example
96 (Weiss et al. 2016; Van der Veer et al. 2016).

97 The objective of this study was to develop a *Xf* detection assay based on the largely used
98 MLST scheme (Yuan et al. 2010) that lowers the limit of detection (LoD) to enable at least
99 the identification of *Xf* subspecies and, if possible, provide larger sets of typing data directly
100 from plant samples. We used genomic sequences to improve each PCR efficiency and showed
101 a drastic increase in the sensitivity as compared to that of the conventional-MLST approach.

102 **Materials and Methods**

103 **Strains and media**

104 A collection of target and non-target bacterial strains was used to test *in vitro* the specificity
105 of the newly designed primers and the nested PCR assays. This set was made of five *X.*
106 *fastidiosa* strains from different subspecies and 33 strains representing bacteria
107 phylogenetically close to *Xf*, i.e. various *Xanthomonas*, as well as strains of other plant
108 pathogenic bacteria and endosymbionts potentially inhabiting the same niches as *Xf* (Table 1),
109 available at the French Collection of Plant-Associated Bacteria (CIRM-CFBP;
110 https://www6.inra.fr/cirm_eng/CFBP-Plant-Associated-Bacteria). The *Xf* strains were grown
111 on modified PWG media (agar 12 g.L⁻¹; soytone 4 g.L⁻¹; bacto tryptone 1 g.L⁻¹
112 ¹; MgSO₄.7H₂O 0.4 g.L⁻¹; K₂HPO₄ 1.2 g.L⁻¹; KH₂PO₄ 1 g.L⁻¹; hemin chloride (0.1% in NaOH
113 0.05 M) 10 ml.L⁻¹; BSA (7.5%) 24 ml.L⁻¹ ; L-glutamine 4 g.L⁻¹) at 28°C for one week.
114 *Agrobacterium* and *Rhizobium* were grown at 25°C for one to two days on MG medium
115 (Mougel et al. 2001) ; TSA was used (tryptone soybroth 30 g.L⁻¹; agar 15 g.L⁻¹) for
116 *Clavibacter*, *Ensifer*, *Stenotrophomonas*, *Xanthomonas* and *Xylophilus*; and King's medium B

117 (King et al. 1954) for *Dickeya*, *Erwinia*, *Pantoea* and *Pseudomonas*. For PCRs, bacterial
118 suspensions were prepared from fresh cultures in sterile distilled water, adjusted at
119 $OD_{600\text{ nm}} = 0.1$ and used as templates for amplification after boiling for 20 minutes, thermal
120 shock on ice and centrifugation 10 000g, 10 min.

121 **DNA extraction**

122 Genomic DNA from *Xf* strain CFBP 8070 was extracted with the Wizard genomic DNA
123 Purification Kit (Promega, France) and used to prepare a 10-fold serial dilutions from 220
124 ng.mL^{-1} (corresponding to 0.8×10^8 copies. mL^{-1} of genomic DNA) to 22 fg.mL^{-1} (8 copies. mL^{-1}
125 ¹) to evaluate the LoD of the nested-MLST. Copies number were calculated using an
126 estimated genome size of 2 903 976 bp, knowing that $1 \text{ pg} = 9.78 \times 10^8 \text{ bp}$ (Doležel et al.
127 2003). A total of 268 plant samples were collected in Corsica, France, based on symptoms
128 compatible with those caused by *X. fastidiosa*; 106 samples were collected in June 2017 and
129 162 in September 2018. For each French sample, DNA was extracted as described in PM7/24
130 (EPPO 2019) using two methods in order to optimize the chances of detection. CTAB-based
131 extraction and robotic QuickpickTM SML kit from Bio-Nobile were used with the following
132 modification: a sonication step (1 min, 42 KHz) was added after the samples (petioles, twigs)
133 were finely cut, and was followed by a 15-min incubation period at room temperature. For
134 initial laboratory diagnosis MLST results were compared with the Harper's qPCR test (Harper
135 et al. 2010) as in EPPO (2019) with following modifications: primers Xf-F and Xf-R, and
136 probe Xf-P (Harper et al. 2010) were used at a final concentration of $0.6 \mu\text{M}$ and $0.2 \mu\text{M}$
137 respectively, non-acetylated BSA was used at final concentration of $1.5 \mu\text{g}.\mu\text{L}^{-1}$, and $2 \mu\text{l}$ of
138 DNA were used in $10 \mu\text{l}$ reaction volume. Each DNA sample were tested in triplicates. To
139 validate the nested PCR, DNA samples were provided by the National Reference Laboratory
140 for Phytopathogenic Bacteria, Valencia, Spain, and from the Official Phytosanitary
141 Laboratory of the Balearic Islands for determining *Xf* subspecies. Those DNA samples

142 correspond to DNA extractions made from symptomatic plants sampled during official
143 monitoring surveys. A total of 70 *Xf*-infected samples were analyzed from Balearic Islands
144 and mainland Spain during 2018 (40 samples) and 2019 (30 samples), as well as 26 insect
145 samples from both regions. DNA was extracted from petioles of symptomatic leaves as
146 described in PM7/24 (EPPO 2019) using a CTAB-based extraction method for plant samples
147 from Alicante and insect samples from Alicante and Balearic Islands. A Mericon DNeasy
148 Food kit from Qiagen was used for plant samples from Balearic Islands. All DNA extraction
149 methods have been validated; validation data is available in the EPPO Database on Diagnostic
150 Expertise (EPPO, 2019).

151

152 **Nested-MLST primers and reactions**

153 The seven HKG sequences (*cysG*, *gltT*, *holC*, *leuA*, *malF*, *nuoL*, *petC*) were extracted from 39
154 *Xf* genome sequences (S1 Table) (Denancé et al. 2019) to design the nested primers.
155 Alignments were performed with BioEdit sequence alignment editor. The primers designed
156 by Yuan *et al* (2010) were destined to be used as inner primers (PCR2) (Table 2) in our nested
157 assay. We checked their characteristics with Primer3 V4.1.0 software (<http://primer3.ut.ee/>).
158 Because of high T_m differences between forward and reverse primers for some primer pairs
159 (*gltT*) (S2 Table), or high hairpin T_m values (*holC* forward primer), some primers from Yuan
160 et al. (2010) were redesigned nearly at the same positions to improve their efficiency.
161 Moreover, as primer sequences were already near the locus sequence ends, we also had to
162 relocate some of them to design nested primers inside the sequence alignments without loss of
163 informative sites. Outer primers (PCR1) were designed with Primer3 V4.1.0 software
164 (<http://primer3.ut.ee/>) in flanking regions targeted by the inner primers. Outer and inner
165 primers were tested *in silico* using a primer search tool available in the galaxy toolbox of
166 CIRM-CFBP (<https://iris.angers.inra.fr/galaxypub-cfbp>) on 194438 bacterial Whole Genome

167 Shotgun (WGS) sequences available in the NCBI database (as on March, 2019) including 58
168 *Xylella* and 1292 *Xanthomonas*, and *in vitro* on target and non-target bacterial strains (Table
169 1).

170 PCRs were performed in 25 μ l reaction buffer (Promega) with $MgCl_2$ at 1.5mM final, 200 μ M
171 dNTP, 300 μ M each of the forward and reverse primers, 0.6 U GoTaq G2 (Promega) and 2 μ l
172 of sample DNA. The first-round PCR program consisted of an initial denaturation step of 3
173 min at 95°C followed by 35 cycles of 30 s denaturation at 95°C, 30 s annealing at the relevant
174 temperature according to each gene (determined by gradient PCR) and 60s elongation at 72°C
175 followed by a final extension step of 10 min at 72°C (Table 2). The second round was
176 performed with 30 cycles under same conditions and same concentrations but with a final
177 volume of 50 μ l for sequencing purposes and with 4 μ l of first-round PCR product. The primer
178 pairs of the second round of each nested PCR were used for sequencing (by Genoscreen,
179 Lille, France for French samples and by Stabvida, Caparica, Portugal, for Spanish samples)
180 the corresponding PCR products after 1.8% agarose gel visualization. To avoid
181 contamination, one sample was opened at a time and stringent cleaning measures were
182 applied after each experiment.

183

184 **Statistical analysis**

185 The sensitivity of detection by conventional- and nested-MLST PCRs were compared in plant
186 and vector samples for the seven HKGs that were analyzed by both approaches, by using a
187 Chi square test using SAS (version 9.4, SAS Institute, Cary, NC, USA). Analysis was
188 performed for the Spanish samples only, as HKG-PCRs were not systematically carried out
189 on the French samples. Results were considered significantly different when $p \leq 0.05$.

190 **Sequence acquisition, alignment and analyses**

191 Forward and reverse nucleotide sequences were assembled, and aligned using Geneious 9.1.8
192 software (French samples) or Bionumerics V7.6.3 software (Spanish samples) to obtain high
193 quality sequences. ST or loci assignation was performed according to
194 <http://pubmlst.org/xfastidiosa/>. To reduce the costs of sequencing for French samples, only
195 PCR products obtained for samples showing the highest rate of successful HKG
196 amplifications were sequenced. On the other hand, all positive *holC* amplifications were
197 sequenced to obtain a larger view of alleles present in Corsica.

198

199 **RESULTS**

200

201 **Nested-MLST proved to be specific**

202 The specificity of the outer and inner primer pairs was tested *in silico* and *in vitro*. *In silico*,
203 all primers pairs showed the best scores of alignment with *Xf* genomic sequences. Some non-
204 target organisms showed sequences nearly identical at outer primer locations with only one
205 mismatch and a similar expected fragment size, but sequences of inner primers were more
206 different indicating that there will be no amplification. This was the case for various
207 *Xanthomonas* strains that contained one mismatch at position 15 of the *petC* forward outer
208 primer and an identical sequence for the outer reverse primer. *X. taiwanensis holC* sequence
209 corresponding to inner primers contained also only one mismatch. The fragment size
210 predicted was as expected for *Xf*. Other predictions with one mismatch located in primers did
211 not end in fragment amplifications of the same expected size. Then, the specificity of the
212 outer and inner primer pairs (Table 2) was validated *in vitro* on five target strains and 33 non-
213 target strains (Table 1). Specificity of the nested-MLST assay could not been tested *in vitro*
214 on *X. taiwanensis* as no strain was available. Amplifications were obtained for all *Xf* strains.

215 No amplification was detected on the non-target strains except for strain CFBP 2532
216 (*Xanthomonas oryzae* pv *oryzae*) and CFBP 2533 (*Xanthomonas hortorum* pv. *pelargonii*) in
217 the first round of the nested PCR for the *petC* outer primers, providing a product of the
218 expected size. However, these products were not amplified in the second round of the nested
219 PCR and no false positive signal was finally obtained.

220 **Nested-MLST Limit of Detection is comparable to that of qPCR**

221 The sensitivity of each primer combination was evaluated on serial dilutions of a genomic
222 DNA solution calibrated (Qubit fluorimeter, Invitrogen) at 220 ng.mL⁻¹ (Figure 1). First round
223 PCRs gave a signal more or less intense for concentrations up to 2.2 ng.mL⁻¹ (0.8 x 10⁶
224 copies.mL⁻¹) for all HKG except *malF* and *cysG* (220 pg.mL⁻¹). The second round of PCRs
225 allowed a sufficiently strong signal for sequencing for concentrations up to 22 pg.mL⁻¹ (0.8 x
226 10⁴ copies.mL⁻¹) for *gltT*, *holC*, *petC*, *leuA*, *cysG*, and up to 2.2 pg.mL⁻¹ (0.8 x 10³ copies.mL⁻¹)
227 for *nuoL* and *malF*. The same range of genomic DNA solutions was tested with the
228 Harper's qPCR test to compare sensitivity of these two tests (S3 Table). The latest signal
229 (LoD) for the Harper's qPCR test (Cq =37.64) was obtained with the concentration of 0.8.10³
230 copies.mL⁻¹ and no amplification was obtained for lower concentrations.

231 Previously, we evaluated the LoD of the conventional PCRs for *cysG* and *malF* of the initial
232 MLST scheme (Yuan et al, 2010) on a range of dilutions of CFBP 8070 genomic DNA with
233 the Platinum Taq polymerase (Invitrogen) and tested the effect of adding BSA (final
234 concentration at 0.3µg. µL⁻¹) on the efficiency of the conventional PCRs. No improvement
235 was obtained as all signals remained around 0.8 x 10⁶ bacteria.mL⁻¹ (S1 Figure).

236 **Analysis of naturally infected samples**

237 Using qPCR Harper's test, 22 samples from 2017 and eight samples from 2018 collected in
238 France were positive (C_q values < 35) with one or both DNA extraction methods; 70 samples
239 from 2017 and 36 samples from 2018 were equivocal ($35 \leq C_q \leq 40$), 14 samples from 2017 and
240 118 from 2018 were negative ($C_q > 40$) (Table 3 and S4 Table). Positive and equivocal
241 samples were tested using the first round of PCR of the MLST assay: five samples from 2017
242 (one *Spartium junceum*, three *Polygala myrtifolia*, and one *Genista corsica*) gave a signal for
243 at least one gene, but no complete typing was obtained for any sample. No sample from 2018
244 gave a signal. Most of Spanish samples used to evaluate nested-MLST scheme were positive
245 using Harper's qPCR (only two out of 40 plant samples were equivocal in 2018 and eight out
246 of 26 vector samples).

247 **Nested-MLST improved successful HKG typing by increasing sensitivity level**

248 Using nested-MLST for French samples, full allelic profiles were obtained for five samples
249 from 2017 and one from 2018 corresponding to the lowest C_q in Harper's qPCR test (Table 4
250 and S4 Table) Among fully typed samples, four were *X. fastidiosa* subsp. *multiplex* ST7
251 (*Genista corsica*, *Polygala myrtifolia*, *Spartium junceum*), and two were *X. fastidiosa* subsp.
252 *multiplex* ST6 (*Polygala myrtifolia*).

253 Our scheme was also evaluated on Spanish samples already proved infected by *Xf*. These
254 samples from different outbreaks showed a wide range of C_q values ranging from 18.8 to 36.0
255 for plant samples and from 23.29 to 37.0 for insect samples (Table 3 and S4 Table). Samples
256 were first analyzed using the conventional-MLST assay (Yuan et al. 2010). Amplification
257 efficiency was variable and ranged from 10% for *glT* to 67% for *nuoL* with an average of
258 25% and 26% for the seven HKG in 2018 and 2019, respectively. The nested-MLST assay
259 improved the amplification efficiency that increased to 75% for *leuA* and up to 93% for *holC*
260 with an average of 81% and 91% in 2018 and 2019, respectively. In total, full allelic profiles

261 were obtained in seven plant samples using the conventional-MLST assay, whereas a total of
262 55 samples were fully typed with the improved nested-MLST assay (Table 4). For the 70
263 plant DNA samples that were tested by both protocols, for all the seven HKGs, conventional-
264 MLST showed a significant ($P < 0.0005$ for 2018 and $P < 0.0283$ for 2019) lower number of
265 samples amplified as compared to nested-MLST. Among fully typed plant samples using the
266 nested-MLST, we identified *X. fastidiosa* subsp. *fastidiosa* ST1 in *Ficus carica* and *Juglans*
267 *regia*, *X. fastidiosa* subsp. *multiplex* ST6 in *Helichrysum italicum*, *Olea europaea*, *Phagnalon*
268 *saxatile*, *Polygala myrtifolia*, *Prunus armeniaca*, *Prunus domestica*, *Prunus dulcis*, *Rhamnus*
269 *alaternus*, and *Rosmarinus officinalis*, *X. fastidiosa* subsp. *multiplex* ST7 in *Prunus dulcis*, *X.*
270 *fastidiosa* subsp. *multiplex* ST81 in *Lavandula angustifolia* and *Prunus dulcis*, and *X.*
271 *fastidiosa* subsp. *pauca* ST80 in *Cistus albidus*, *Prunus dulcis*, and *Rosmarinus officinalis*.

272 Not all insect samples could be tested by both protocols due to restrictions in DNA amount. In
273 samples tested only by the original MLST assay (Yuan et al., 2010), the percentages of
274 successful amplifications ranged from 8% (*gltT* and *malF*) to 65% (*cysG*). With the nested-
275 MLST assay, successful amplifications ranged from 54% (*malF*) to 81% (*cysG*), with an
276 average efficiency for the seven HKG of 22% to 67% for conventional and nested approach,
277 respectively (Table 3, S4 Table). Nine insect samples were fully typed using a combination of
278 both protocols (Table 4). *X. fastidiosa* subsp. *fastidiosa* ST1 was identified in insects from
279 Mallorca (Balearic Islands), *X. fastidiosa* subsp. *multiplex* ST6 in insects from Alicante
280 (mainland Spain) and *X. fastidiosa* subsp. *multiplex* ST81 in insects from Balearic Islands. For
281 the nine insect samples that were tested by both protocols, conventional-MLST showed a
282 significant ($P < 0.0247$) lower number of samples amplified as compared to nested-MLST for
283 six of the seven HKGs (excluding *cysG*). These results indicate that for insect samples it is
284 also better to use directly the improved nested-MLST assay.

285 No nonspecific amplicons were observed in any of the samples. Negative controls (water)
286 were run in the first and the second PCR and were always negative. The negative control
287 coming from the first reaction always tested negative in the second one, confirming the
288 absence of contamination during the entire process.. Positive control was a suspension of
289 strain CFBP 8084 (ST29) from the subspecies *morus* or strain CO33 (ST72) as this STs were
290 not previously found in Corsica, France or Spain, respectively.

291

292 **Nested-MLST allowed identification of new alleles among French samples**

293 Incomplete profiles were obtained for various French samples due to variable amplification
294 efficiencies varying according to the HKG. From 9% (with *gltT*) to 55% (with *holC*) of
295 French samples gave a signal applying the nested-MLST assay. Alleles that were not yet
296 described in plant samples in France were detected in 2017. This was the case for *holC_1* and
297 *holC_2* alleles known to occur in ST from ST1 to ST5 and ST75 that cluster in the subspecies
298 *fastidiosa* (<https://pubmlst.org/xfastidiosa/>). These alleles were sequenced in samples of
299 *Asparagus acutifolius*, *Eleagnus*, *Cistus monspeliensis* and *C. creticus*, *Quercus ilex*, *Myrtus*
300 *myrtifolia*, *Olea europea*, *Platanus*, *Arbutus unedo* (S4 Table). Other *holC* alleles already
301 described in STs clustering in the subspecies *fastidiosa* (*holC_24*) were also sequenced from
302 *Cistus monspeliensis* and *Pistaccia lentiscus*. *HolC_10* alleles described in STs clustering in
303 the subspecies *pauca* were sequenced from *Cistus monspeliensis* and *C. salicifolius*,
304 *Cypressus*, *Metrosideros excelsa*, *Myrtus communis*, *Pistaccia lentiscus*, *Quercus ilex*, *Rubia*
305 *peregrina*, *Smilax aspera* samples. Similarly, *holC_3* (known in ST6, ST7, ST25, ST34,
306 ST35, ST79, ST81 and ST87 clustering in the subspecies *multiplex*) were obtained from
307 samples of *Acer monspeliensis*, *Arbutus unedo*, *Calicotome spinosa*, *Cistus monspeliensis*,
308 *Genista corsica*, *Myrtus communis*, *Olea europea*, *Phyllirea angustifolia*, *Polygala myrtifolia*,
309 *Quercus ilex* and *Q. pubescens*, *Spartium junceum*. Among samples from 2018, only *holC_1*

310 allele was detected in *Olea europaea*, *Quercus ilex*, and *Platanus sp.* samples, and *holC_3*
311 allele in *Cistus monspeliensis*, *Acer monspeliensis*, *Myrtus communis*, and *Polygala myrtifolia*
312 samples.

313

314 **Recombinants or mixed infections were identified by nested-MLST**

315 Some French samples were further sequenced for several loci and these sequencing confirmed
316 the presence of alleles occurring in the subspecies *fastidiosa*, *multiplex* and *pauca* (S4 Table).
317 All alleles were previously described but were detected in combinations that were not
318 previously described, suggesting the presence of recombinants or of mix infections (S4
319 Table). This is the case for *Cistus monspeliensis* 7 showing an unknown combination of
320 *cysG_2/ petC_2/ nuoL_2/ gltT_2* (known in ST5) with *malF_4* (known in ST2), both from
321 subspecies *fastidiosa* ; *Helichrysum italicum* 1 showing *leuA_1* (known in subspecies
322 *fastidiosa*) with *petC_3/ holC_3* known in subspecies *multiplex*; *Myrtus communis* 4 with
323 *leuA_3/holC_2* respectively known in subspecies *multiplex* and *fastidiosa*; *Myrtus communis*
324 8 and *Platanus* presenting form 1 alleles for five HKG mixed with *malF_4* (all known in
325 subspecies *fastidiosa*) and *Q. ilex* 10 presenting form 1 alleles for two HKG mixed with
326 *malF4*); *Olea europaea* 2 with four *multiplex* alleles combined with *nuoL_1* (subspecies
327 *fastidiosa*); *Olea europaea* 5 with four *pauca* alleles combined with *malF_15* (known in ST72
328 and ST76, subspecies *fastidiosa*). Two samples gave a double sequence for *holC* that were
329 impossible to analyze (S4 Table). Some sequences were ambiguous with superimposed peaks
330 at some locations in otherwise good quality chromatograms revealing mixed infections. In
331 those 12 samples, the number of potential combinations was too high to detect one probable
332 allelic form, excepted for *Prunus dulcis* where the superimposed chromatograms
333 corresponded to only two allelic forms (*holC_3* or *holC_6* which are found in subspecies

334 *multiplex*). The *holC_6* allelic form and the *leuA_5* allele obtained for this sample are found
335 in ST10, ST26, ST36, ST46, and ST63.

336

337 **DISCUSSION**

338 A two-step nested procedure for MLST was developed to improve the typing of samples
339 infected with low *Xf* population sizes that cannot be typed using the conventional protocol. In
340 order not to affect the comparability of the results with the databases, the widely used MLST
341 scheme developed for *Xf* that is supported by the pubMLST public website (Yuan et al. 2010)
342 was re-used.

343 The nested-MLST approach proved to be specific and efficient. No nonspecific amplifications
344 were observed in any of the samples. Moreover, the sensitivities of the Harper's qPCR
345 detection test and the nested-MLST were similar with a LoD ranging from 10^3 bacteria.mL⁻¹
346 to 10^4 bacteria.mL⁻¹. These LoDs are similar to other nested-MLST approaches such as those
347 developed for *Burkholderia cepacia* (Drevinek et al., 2010) but higher than for the one
348 developed for *Neisseria meningitidis* (10 copies mL⁻¹) (Diggle et al, 2003). Consequently, in
349 resource-limited settings where qPCR facilities are not available, the assay may be used as a
350 useful diagnostic tool if applied with all necessary precautions to avoid cross-contamination
351 between samples. The sequencing, which is costly, can be done as a consecutive but separate
352 step to provide information on subspecies present in the sample. Higher bacterial loads (as
353 indicated by lower Cq values) were observed in Spanish samples than in French samples, for
354 which low amplification efficiency and partial profiles were observed. Full allelic profiles
355 (ST6 and ST7 from *multiplex* subspecies) were obtained for *Polygala myrtifolia*, *Spartium*
356 *junceum* and *Genista corsica* samples from France probably because they carried a higher
357 bacterial load as shown by the low Cq obtained with the Harper's qPCR test: five of the six

358 typed samples had a C_q value between 23.4 and 26.5. The use of the nested-MLST assay to
359 type plant Spanish samples allowed a higher number of successful complete typing (55
360 samples versus seven samples with the conventional approach). Spanish samples generally
361 showed higher *Xf* titer (i.e, lower C_q values in Harper's qPCR test) than the French samples
362 but also concerned different plant species.

363 In our nested-MLST assay as well as in the original MLST assay, the amplification
364 efficiencies were variable among genes, while all primers were designed using the same
365 parameters from the software. For example, the *holC* gene for French samples tested with the
366 nested-MLST assay was successfully amplified in 55% of samples collected in 2017 while the
367 *gltT* and *nuoL* genes gave the lowest rates (around 26%). For samples collected in Spain
368 tested with the original MLST assay, amplification rates among the seven HKGs ranged
369 from 10 to 67%. Success rate variations were also observed in medical research using MLST
370 between samples and between loci (Weiss et al. 2016). When conducted on strains, no
371 differences about amplification rates are observed because of DNA excess. Robustness of a
372 PCR reaction is determined by appropriate primers and it is not always obvious why some
373 primer combinations do not amplify well, even if some parameters such as DNA folding can
374 interfere in PCR efficiency (Bustin & Huggett 2017). In this study, even if primer annealing
375 temperature was adjusted, design of primers was limited by their arbitrary localization.

376 Typing results of French samples were concordant with previously published results (Denancé
377 et al. 2017) but also revealed the presence of alleles not yet described in France. It should be
378 noticed that no unknown sequence was obtained, refraining from evoking contaminations as
379 the origin of these yet undescribed alleles in France. Thanks to the high rate of amplification
380 of *holC* in nested PCR, it was also possible to obtain sequences for equivocal samples (C_q
381 with the Harper's qPCR test above 35) to confirm the presence of the bacterium in these
382 samples. Surprisingly, these amplifications led to alleles that correspond to subspecies other

383 than the *multiplex* subspecies. Thereby, alleles from subspecies *pauca* (*holC_10*) and
384 *fastidiosa* (*holC_1*, *holC_2*, *holC_24*) were sequenced. *HolC_10* was already reported in
385 *Polygala myrtifolia* in the south of France in 2015 (Denancé et al., 2017). *HolC_1* finding is
386 in agreement with Cruaud et al (2018), who also reported *holC_1* in insects in Corsica. Up to
387 now, no *holC_2* was reported in France but it is known in the USA. *HolC_24* was also
388 reported in *Polygala myrtifolia* in Corsica in 2015 (Denancé et al. 2017). Further plant
389 sampling efforts are needed to confirm the establishment of those strains in the environment
390 or to document further the dynamics of alleles revealing sporadic infections.

391 For French samples only, several samples could not be typed since the chromatograms
392 showed an overlap of two peaks precisely on the polymorphic sites (mainly with *leuA* and
393 *holC* genes). This has already been reported by Denancé et al (2017), it suggests the
394 simultaneous presence of several strains in the same sample since only one copy of these
395 genes are known in *Xf* (Yuan et al, 2010). Moreover, the report of previously unknown
396 combination of alleles belonging to different subspecies can also results from the presence of
397 co-infection or of recombinants. Recombination events are reported in *Xf* (Denancé et al.
398 2017; Jacques et al. 2016, Nunney et al. 2014a , Saponari et al. 2019) and could have led to
399 host shift (Nunney et al. 2014b). In this study, eight samples presented unknown
400 combinations of alleles from the same or different subspecies which could be explained by
401 intrasubspecies or intersubspecies recombination events. As reported in Potnis et al. (2019),
402 such events may exist and occur but not with the same frequency. Moreover, natural
403 competence can be variable among *Xf* strains (Kandel et al., 2017). These events could also
404 reflect a mechanism of adaptation (Kandel et al., 2016). Five samples among these eight
405 samples were collected in 2017 and three in 2018, and were different between years. In 2018
406 the three cases were a similar combination of alleles and were found in three different plants.
407 Future surveys will be necessary to know if some of these recombinants strains are indeed

408 present in Corsica or are the consequence of mixed infections and if they have adapted and
409 survived on different hosts.

410 The objective of this study was to improve the published MLST scheme supported by a public
411 website (<https://pubmlst.org/xfastidiosa/>) by designing nested primers to lower the limit of
412 detection and help in *Xf* diagnosis and typing. Thus, this improved MLST assay enables a
413 higher sensitivity and specific typing of *Xf* directly from plant and insects samples without the
414 need of isolating the strain and at an affordable cost.

415 **Conflict of Interest**

416 The authors declare that the research was conducted in the absence of any commercial or
417 financial relationships that could be construed as a potential conflict of interest. The present
418 work reflects only the authors' view and no analysis has been made in the French Reference
419 Lab; in particular ED is not authorized to perform any official tests at Anses.

420 **Author Contributions**

421 SC, QB and MMB, MPVA performed the experiments, ED took part to primer design, MB
422 helped with bioinformatic tools, SC conceived the study, MAJ and BL applied for funding,
423 SC, ED, MAJ, BL wrote the manuscript. All authors read and approved the final version of
424 the manuscript.

425 **Funding**

426 ED salary was funded by INRA SPE division and Anses. This work received support from the
427 European Union's Horizon 2020 research and innovation program under grant agreement
428 727987 XF_ACTORS (*Xylella fastidiosa* Active Containment Through a multidisciplinary-
429 Oriented Research Strategy), and from Projects E-RTA2017-00004-C06-02 from 'Programa
430 Estatal de I+D Orientada a los Retos de la Sociedad' from Spanish State

431 Research Agency, CSIC Intramural Project 2018 40E111, and from "Conselleria de
432 Agricultura, Desarrollo Rural, Emergencia Climática y Transición
433 Ecológica" from Valencia region, and the Ministry of Agriculture,
434 Fisheries and Food of Spain. The present work reflects only the authors' view and the EU
435 funding agency is not responsible for any use that may be made of the information it contains.

436 **Acknowledgments**

437 We thank Muriel Bahut (ANAN technical facility, SFR QUASAV, Angers, FR) for DNA
438 extraction automatization, CIRM-CFBP (Beaucouzé, INRA, France;
439 [http://www6.inra.fr/cirm_eng/CFBP-Plant Associated-Bacteria](http://www6.inra.fr/cirm_eng/CFBP-Plant%20Associated-Bacteria)) for strain preservation and
440 supply. We thank Ester Marco-Noales from the National Reference Laboratory for
441 Phytopathogenic Bacteria (IVIA), and Diego Olmo from the Official Phytosanitary
442 Laboratory of the Balearic Islands for providing DNA samples for MLST typing.

443

444 **Nomenclature**

445 BLAST: Basic Local Alignment Search Tool

446 C_q : quantification cycle

447 HKG : housekeeping gene

448 INRA: French National Institute for Agricultural Research

449 IRHS: Research Institute of Horticulture and Seeds

450 LoD : Limit of Detection

451 MLST: Multilocus Sequence Typing

452 NCBI: National Center for Biotechnology Information

453 ST: Sequence Type

454 *Xf. Xylella fastidiosa*

455 WGS: Whole Genome Shotgun

456

457 **References**

458

459 Almeida, R.P.P., Nascimento, F.E., Chau, J., Prado, S.S., Tsai, C.-W., Lopes, S.A., Lopes,
460 J.R.S., 2008. Genetic Structure and Biology of *Xylella fastidiosa* Strains Causing Disease in
461 Citrus and Coffee in Brazil. *Appl. Environ. Microbiol.* 74, 3690–3701.

462 <https://doi.org/10.1128/AEM.02388-07>

463 Bonants, P., Griekspoor, Y., Houwers, I., Krijger, M., van der Zouwen, P., van der Lee, T. A.,
464 & van der Wolf, J., 2019. Development and Evaluation of a Triplex TaqMan Assay and Next-
465 Generation Sequence Analysis for Improved Detection of *Xylella* in Plant Material. *Plant*
466 *disease* 103, 645–655. <https://doi.org/10.1094/PDIS-08-18-1433-RE>

467 Burbank, L. P., & Ortega, B. C., 2018. Novel amplification targets for rapid detection and
468 differentiation of *Xylella fastidiosa* subspecies *fastidiosa* and *multiplex* in plant and insect
469 tissues. *Journal of microbiological methods* 155, 8–18.

470 <https://doi.org/10.1016/j.mimet.2018.11.002>

471 Bustin, S., & Huggett, J., 2017. qPCR primer design revisited. *Biomolecular detection and*
472 *quantification* 14, 19–28. <https://doi.org/10.1016/j.bdq.2017.11.001>

- 473 Ciapina, L. P., Carareto Alves, L. M., & Lemos, E. G. M., 2004. A nested-PCR assay for
474 detection of *Xylella fastidiosa* in citrus plants and sharpshooter leafhoppers. Journal of applied
475 microbiology 96, 546–551. <https://doi.org/10.1111/j.1365-2672.2004.02176.x>
- 476 Cruaud, A., Gonzalez, A. A., Godefroid, M., Nidelet, S., Streito, J. C., Thuillier, J. M., ... &
477 Rasplus, J. Y., 2018. Using insects to detect, monitor and predict the distribution of *Xylella*
478 *fastidiosa*: a case study in Corsica. Scientific reports 8,
479 15628. <https://doi.org/10.1038/s41598-018-33957-z>
- 480 Della Coletta-Filho, H., Francisco, C. S., Lopes, J. R. S., De Oliveira, A. F., & Da Silva, L. F.
481 D. O., 2016. First report of olive leaf scorch in Brazil, associated with *Xylella fastidiosa*
482 subsp. *pauca*. Phytopathologia mediterranea 55(1).
483 https://doi.org/10.14601/Phytopathol_Mediterr-17259
- 484 Denancé, N., Legendre, B., Briand, M., Olivier, V., De Boissesson, C., Poliakoff, F., &
485 Jacques, M. A., 2017. Several subspecies and sequence types are associated with the
486 emergence of *Xylella fastidiosa* in natural settings in France. Plant Pathology 66, 1054–
487 1064. <https://doi.org/10.1111/ppa.12695>
- 488 Denancé, N., Briand, M., Gaborieau, R., Gaillard, S., & Jacques, M. A., 2019. Identification
489 of genetic relationships and subspecies signatures in *Xylella fastidiosa*. BMC genomics 20,
490 239. <https://doi.org/10.1186/s12864-019-5565-9>
- 491 Diggle, M. A., Bell, C. M., & Clarke, S. C., 2003. Nucleotide sequence-based typing of
492 meningococci directly from clinical samples. Journal of medical microbiology 52, 505–
493 508. <https://doi.org/10.1099/jmm.0.05078-0>

- 494 Doležel, J., Bartos, J., Voglmayr, H., & Greilhuber, J., 2003. Nuclear DNA content and
495 genome size of trout and human. *Cytometry. Part A: the journal of the International Society*
496 *for Analytical Cytology* 51A, 127–128. <https://doi.org/10.1002/cyto.a.10013>
- 497 Drevinek, P., Vosahlikova, S., Dedeckova, K., Cinek, O., & Mahenthiralingam, E., 2010.
498 Direct culture-independent strain typing of *Burkholderia cepacia* complex in sputum samples
499 from patients with cystic fibrosis. *Journal of clinical microbiology* 48, 1888–
500 1891. <https://doi.org/10.1128/JCM.02359-09>
- 501 Dupas, E., Legendre, B., Olivier, V., Poliakoff, F., Manceau, C., & Cunty, A., 2019a.
502 Comparison of real-time PCR and droplet digital PCR for the detection of *Xylella fastidiosa*
503 in plants. *Journal of microbiological methods* 162, 86–
504 95. <https://doi.org/10.1016/j.mimet.2019.05.010>
- 505 Dupas, E., Briand, M., Jacques, M. A., & Cesbron, S., 2019b. Novel tetraplex quantitative
506 PCR assays for simultaneous detection and identification of *Xylella fastidiosa* subspecies in
507 plant tissues. *Frontiers in Plant Science*, 10, 1732. <https://doi.org/10.3389/fpls.2019.01732>
- 508 EFSA (European Food Safety Authority), 2018. Scientific report on the update of the *Xylella*
509 spp. host plant database. *EFSA Journal* 16(9):5408, 87 pp.
- 510 EPPO Bulletin. 2019. PM 7/24 (3) *Xylella fastidiosa*. *EPPO Bulletin* 49 (2), 175–227
511 <https://doi.org/10.1111/epp.12575>
- 512 Guan, W., Shao, J., Singh, R., Davis, R. E., Zhao, T., & Huang, Q., 2013. A TaqMan-based
513 real time PCR assay for specific detection and quantification of *Xylella fastidiosa* strains
514 causing bacterial leaf scorch in oleander. *Journal of microbiological methods* 92, 108–
515 112. <https://doi.org/10.1016/j.mimet.2012.11.008>

- 516 Haelterman, R. M., Tolocka, P. A., Roca, M. E., Guzmán, F. A., Fernández, F. D., & Otero,
517 M. L., 2015. First presumptive diagnosis of *Xylella fastidiosa* causing olive scorch in
518 Argentina. *Journal of Plant Pathology* 97(2).
- 519 Harper, S. J., Ward, L. I., & Clover, G. R. G., 2010. Development of LAMP and real-time
520 PCR methods for the rapid detection of *Xylella fastidiosa* for quarantine and field
521 applications. *Phytopathology* 100, 1282–1288. <https://doi.org/10.1094/PHYTO-06-10-0168>
- 522 Jacques, M. A., Denancé, N., Legendre, B., Morel, E., Briand, M., Mississippi, S., ... &
523 Crouzillat, D., 2016. New coffee plant-infecting *Xylella fastidiosa* variants derived via
524 homologous recombination. *Appl. Environ. Microbiol.* 82, 1556–
525 1568. <https://doi.org/10.1128/AEM.03299-15>
- 526 Janse, J. D., & Obradovic, A., 2010. *Xylella fastidiosa*: its biology, diagnosis, control and
527 risks. *Journal of Plant Pathology* S35-S48.
- 528 Kandel, P. P., Lopez, S. M., Almeida, R. P., & De La Fuente, L. (2016). Natural competence
529 of *Xylella fastidiosa* occurs at a high frequency inside microfluidic chambers mimicking the
530 bacterium's natural habitats. *Appl. Environ. Microbiol.*, 82(17), 5269-5277.
- 531 Kandel, P. P., Almeida, R. P., Cobine, P. A., & De La Fuente, L. (2017). Natural competence
532 rates are variable among *Xylella fastidiosa* strains and homologous recombination occurs in
533 vitro between subspecies *fastidiosa* and *multiplex*. *Molecular Plant-Microbe*
534 *Interactions*, 30(7), 589-600.
- 535 King, E. O., Ward, M. K., & Raney, D. E., 1954. Two simple media for the demonstration of
536 pyocyanin and fluorescein. *The Journal of laboratory and clinical medicine* 44(2), 301-307.

- 537 Landa, B. B., 2017. Emergence of *Xylella fastidiosa* in Spain: current situation. Presentation
538 made at the European Conference on *Xylella* 2017. Available
539 at: <https://www.efsa.europa.eu/en/events/event/171113> [Accessed January 17, 2019].
- 540 Li, R., Russell, P., Mcowen, N., Davenport, B., & Zhang, S., 2016. Development of a rapid
541 and reliable isothermal AmplifyRP diagnostic assay for specific detection of *Xylella*
542 *fastidiosa*. *Phytopathology* 106 (12), 109-109
- 543 Marcelletti, S., & Scortichini, M. (2016). Genome-wide comparison and taxonomic
544 relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and
545 a new *Xylella* species. *Archives of microbiology*, 198(8), 803-812.
- 546 Mougel, C., Cournoyer, B., & Nesme, X., 2001. Novel tellurite-amended media and specific
547 chromosomal and Ti plasmid probes for direct analysis of soil populations of *Agrobacterium*
548 biovars 1 and 2. *Appl. Environ. Microbiol.* 67, 65–74. [https://doi.org/10.1128/AEM.67.1.65-](https://doi.org/10.1128/AEM.67.1.65-74.2001)
549 [74.2001](https://doi.org/10.1128/AEM.67.1.65-74.2001)
- 550 Nunney L, Hopkins DL, Morano LD, Russell SE, Stouthamer R., 2014a. Intersubspecific
551 recombination in *Xylella fastidiosa* native to the United States: infection of novel hosts
552 associated with an unsuccessful invasion. *Appl. Environ. Microbiol.* 80, 1159–1169.
553 <https://doi.org/10.1128/AEM.02920-13>
- 554 Nunney, L., Schuenzel, E. L., Scally, M., Bromley, R. E., & Stouthamer, R., 2014b. Large-
555 scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is
556 associated with the host shift to mulberry. *Appl. Environ. Microbiol.* 80, 3025–3033.
557 <https://doi.org/10.1128/AEM.04112-13>

- 558 Nunney L, Yuan XL, Bromley RE, Stouthamer R., 2012. Detecting genetic introgression:
559 high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl.
560 Environ. Microbiol. 78, 4702–4714. <https://doi.org/10.1128/AEM.01126-12>
- 561 Ouyang, P., Arif, M., Fletcher, J., Melcher, U., & Corona, F. M. O., 2013. Enhanced
562 reliability and accuracy for field deployable bioforensic detection and discrimination of
563 *Xylella fastidiosa* subsp. *pauca*, causal agent of citrus variegated chlorosis using Razor Ex
564 technology and TaqMan quantitative PCR. PLoS One. 8,
565 e81647. <https://doi.org/10.1371/journal.pone.0081647>
- 566 Potnis, N., Kandel, P. P., Merfa, M. V., Retchless, A. C., Parker, J. K., Stenger, D. C., ... &
567 De La Fuente, L. (2019). Patterns of inter-and intrasubspecific homologous recombination
568 inform eco-evolutionary dynamics of *Xylella fastidiosa*. *The ISME journal*, 13(9), 2319-2333.
569 <https://doi.org/10.1038/s41396-019-0423-y>
- 570 Saponari, M., Boscia, D., Nigro, F., & Martelli, G. P., 2013. Identification of DNA sequences
571 related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch
572 symptoms in Apulia (Southern Italy). *Journal of Plant Pathology* 95
573 (3). <https://doi.org/10.4454/JPP.V95I3.035>
- 574 Saponari, M., D'Attoma, G., Kubaa, R. A., Loconsole, G., Altamura, G., Zicca, S., ... &
575 Boscia, D., 2019. A new variant of *Xylella fastidiosa* subspecies multiplex detected in
576 different host plants in the recently emerged outbreak in the region of Tuscany,
577 Italy. *European Journal of Plant Pathology* 154, 1195–1200. [https://doi.org/10.1007/s10658-](https://doi.org/10.1007/s10658-019-01736-9)
578 [019-01736-9](https://doi.org/10.1007/s10658-019-01736-9)
- 579 Scally, M., Schuenzel, E. L., Stouthamer, R., & Nunney, L., 2005. Multilocus sequence type
580 system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination

581 and point mutation to clonal diversity. *Appl. Environ. Microbiol.* 71, 8491–
582 8499. <https://doi.org/10.1128/AEM.71.12.8491-8499.2005>

583 Van der Veer, C., Himschoot, M., & Bruisten, S. M., 2016. Multilocus sequence typing of
584 *Trichomonas vaginalis* clinical samples from Amsterdam, the Netherlands. *BMJ open.* 6,
585 e013997. <https://doi.org/10.1136/bmjopen-2016-013997>

586 Waliullah, S., Hudson, O., Oliver, J. E., Brannen, P. M., Ji, P., & Ali, M. E., 2019.
587 Comparative analysis of different molecular and serological methods for detection of *Xylella*
588 *fastidiosa* in blueberry. *PloS one.* 14, e0221903. <https://doi.org/10.1371/journal.pone.0221903>

589 Weiss, S., Menezes, A., Woods, K., Chanthongthip, A., Dittrich, S., Opoku-Boateng, A., ... &
590 Chalker, V., 2016. An extended multilocus sequence typing (MLST) scheme for rapid direct
591 typing of *Leptospira* from clinical samples. *PLoS neglected tropical diseases.* 10,
592 e0004996. <https://doi.org/10.1371/journal.pntd.0004996>

593 Wells, J. M., Raju, B. C., Hung, H. Y., Weisburg, W. G., Mandelco-Paul, L., & Brenner, D.
594 J., 1987. *Xylella fastidiosa* gen. nov., sp. nov: gram-negative, xylem-limited, fastidious plant
595 bacteria related to *Xanthomonas* spp. *International Journal of Systematic and Evolutionary*
596 *Microbiology* 37, 136–143. <https://doi.org/10.1099/00207713-37-2-136>

597 Yaseen, T., Drago, S., Valentini, F., Elbeaino, T., Stampone, G., Digiario, M., & D'ONGHIA,
598 A. M., 2015. On-site detection of *Xylella fastidiosa* in host plants and in “spy insects” using
599 the real-time loop-mediated isothermal amplification method. *Phytopathologia Mediterranea*
600 488-496. https://doi.org/10.14601/Phytopathol_Mediterr-15250

601 Yuan, X., Morano, L., Bromley, R., Spring-Pearson, S., Stouthamer, R., & Nunney, L., 2010.
602 Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf

603 scorch in the United States. *Phytopathology* 100, 601–611. <https://doi.org/10.1094/PHYTO->

604 [100-6-0601](https://doi.org/10.1094/PHYTO-100-6-0601)

Figure 1 : detection threshold of conventional-MLST (a) and nested-MLST (b) for seven HKGs using genomic DNA dilution range (1: 220 ng.mL⁻¹; 2: 22 ng.mL⁻¹; 3 : 2.2 ng.mL⁻¹; 4: 220 pg.mL⁻¹ ; 5: 22 pg.mL⁻¹ ; 6: 2.2 pg.mL⁻¹ ; 7 : 220 fg.mL⁻¹; 8: 22 fg.mL⁻¹)

Table 1: List of target and non-target strains used to verify the specificity of nested-MLST primers

CFBP code	Bacterial species	Host plant	Origin
6448	<i>Agrobacterium rubi</i>	<i>Rubus ursinus</i> var. <i>loganobaccus</i>	USA (1942)
2413	<i>Agrobacterium tumefaciens</i>	<i>Malus</i> sp.	NA (1935)
5523	<i>Agrobacterium vitis</i>	<i>Vitis vinifera</i>	Australia (1977)
2404	<i>Clavibacter insidiosus</i>	<i>Medicago sativa</i>	USA (1955)
4999	<i>Clavibacter michiganensis</i>	<i>Lycopersicon esculentum</i>	Hungary (1957)
3418	<i>Curtobacterium flaccumfaciens</i> pv. <i>flaccumfaciens</i>	<i>Phaseolus vulgaris</i>	Hungary (1957)
1200	<i>Dickeya dianthicola</i>	<i>Dianthus caryophyllus</i>	United Kingdom (1956)
5561	<i>Ensifer meliloti</i>	<i>Medicago sativa</i>	VA, USA (1984)
1232	<i>Erwinia amylovora</i>	<i>Pyrus communis</i>	United Kingdom (1959)
3845	<i>Pantoea agglomerans</i>	<i>Knee laceration</i>	Zimbabwe (1956)
3167	<i>Pantoea stewartii</i> pv. <i>stewartii</i>	<i>Zea mays</i> var. <i>rugosa</i>	USA (1970)
3205	<i>Pseudomonas amygdali</i>	<i>Prunus amygdalus</i>	Greece (1967)
8305	<i>Pseudomonas cerasi</i>	<i>Prunus cerasus</i>	Poland (2007)
7019	<i>Pseudomonas congelans</i>	na ¹	Germany (1994)

1573	<i>Pseudomonas syringae</i> pv. <i>persicae</i>	<i>Prunus persica</i>	France (1974)
1392	<i>Pseudomonas syringae</i> pv. <i>syringae</i>	<i>Syringa vulgaris</i>	United Kingdom (1950)
7436	<i>Rhizobium nepotum</i>	<i>Prunus cerasifera myrobolan</i>	Hungary (1989)
13100	<i>Stenotrophomas maltophilia</i>	<i>Phaseolus vulgaris</i>	Cameroon (2009)
3371	<i>Xanthomonas euvesicatoria</i> pv. <i>citrumelonis</i>	<i>Citrus</i> sp.	USA (1989)
2528	<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	<i>Juglans regia</i>	New Zealand (1956)
2535	<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	<i>Prunus salicina</i>	New Zealand (1953)
4924	<i>Xanthomonas axonopodis</i> pv. <i>axonopodis</i>	<i>Axonopus scoparius</i>	Colombia (1949)
5241	<i>Xanthomonas campestris</i> pv. <i>campestris</i>	<i>Brassica oleracea</i> var. <i>gemmifera</i>	United Kingdom (1957)
2901	<i>Xanthomonas citri</i> pv. <i>aurantifolii</i>	<i>Citrus limon</i>	Argentina (1988)
2525	<i>Xanthomonas citri</i> pv. <i>citri</i>	<i>Citrus limon</i>	New Zealand (1956)
7660	<i>Xanthomonas citri</i> pv. <i>viticola</i>	<i>Vitis vinifera</i>	India (1969)
2625	<i>Xanthomonas gardneri</i>	<i>Medicago sativa</i>	Reunion Island (1986)
4925	<i>Xanthomonas hortorum</i> pv. <i>hederae</i>	<i>Hedera helix</i>	USA (1944)
2533	<i>Xanthomonas hortorum</i> pv. <i>pelargonii</i>	<i>Pelargonium peltatum</i>	New Zealand (1974)
1156	<i>Xanthomonas hyacinthi</i>	<i>Hyacinthus orientalis</i>	Netherlands (1958)
2532	<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	<i>Oryza sativa</i>	India (1965)
2054	<i>Xanthomonas translucens</i>	<i>Hordeum vulgare</i>	USA (1933)
2543	<i>Xanthomonas vasicola</i> pv. <i>holcicola</i>	<i>Sorghum vulgare</i>	New Zealand (1969)
7970	<i>Xylella fastidiosa</i> subsp. <i>fastidiosa</i>	<i>Vitis vinifera</i>	USA (1987)
8416	<i>Xylella fastidiosa</i> subsp. <i>multiplex</i>	<i>Polygala myrtifolia</i>	France (2015)
8084	<i>Xylella fastidiosa</i> subsp. <i>morus</i>	<i>Morus alba</i>	USA (na ¹)

8070	<i>Xylella fastidiosa</i> subsp. <i>multiplex</i>	<i>Prunus</i> spp.	USA (2004)
8402	<i>Xylella fastidiosa</i> subsp. <i>pauca</i>	<i>Olea europea</i>	Italy (2014)

¹: not available

Table 2. Primer sequences used in the *X. fastidiosa* nested-MLST scheme.

locus	PCR round	Forward primer	Reverse primer	Position on Xf M12 genome (CP000941.1)	Annealing temperature (°C)	Size (pb) of reaction product
<i>cysG</i>	1	ccaacatagaagcacgccg	gcgagtgtttcagcgttcc	2111116-2111891	64	776
	2	gccgaagcagtgtggaag ¹	gccatttcgatcagtgcaaaag ¹	2111203-2111844	56	642
<i>gltT</i>	1	ggtgccatccaatccgtttt	tcaggatgtccaattccaacg	1731589-1732504	60	916
	2	tcatgatccaaatcactcgctt ¹	ttactggacgctgcctcg	1731783-1732482	56	700
<i>holC</i>	1	ccgatggtgaagaacagtagaca	gctcgagaaactsgattaatgg	133166-133714	62	549
	2	ggtcacatgctgtttgttc	cacgcgccgacttctattt	133269-133692	59	424
<i>leuA</i>	1	cgaaggtgcaaaacaagtga	cgcactggcttcgataatgtct	1271664-1272549	58	886
	2	ggtgcacgcaaatcgaatg ¹	actggtcctgtaccttctgt	1271752-1272525	60	774
<i>malF</i>	1	aacgtcgtcacccaagaa	atgaggcgggcttctttgg	1680264-1681108	56	845
	2	agcagaagcacgtcccagat	ctggctctgcggtgttgg	1680308-1681074	60	767
<i>nuoL</i>	1	ttggtacgttggtttggtg	gacaaaaccagattgcgtgc	325347-326191	60	845
	2	gcgacttacggttactgggc	accaccgatccacaacgcatt ¹	325454-326050	54	597
<i>petC</i>	1	tcaatgcacgtcctccaat	ggctgccattcgttgaagta	2020498-2021079	60	582
	2	acgtcctccaataagcct	cgttattcacgtatcgtgc	2020505-2021055	56	551

¹ : primers from Yuan et al. (2010).

Table 3: Number of samples, positive and equivocal in qPCR Harper. Percentage of successful amplifications obtained for each locus in conventional and nested PCR

Sample type	country	year	number of samples	qPCR Harper number of samples		Percentage of successful amplifications obtained for each locus in conventional and nested MLST-PCR ^a																average per year	
				Cq<35	Cq≥35	<i>cysG</i>		<i>gltT</i>		<i>holC</i>		<i>leuA</i>		<i>malF</i>		<i>nuoL</i>		<i>petC</i>		conv	nest		
				conv	nest	conv	nest	conv	nest	conv	nest	conv	nest	conv	nest	conv	nest	conv	nest	conv	nest	conv	nest
Plant	France	2017	106	22	70	1.1	28.3	2.2	26.1	4.3	55.4	4.3	34.8	1.1	35.9	0	26.1	1.1	46.7	2	36.5		
Plant	France	2018	162	8	36	0	11.4	0	9.1	0	27.3	0	27.3	0	15.9	0	27.3	0	25	0	20.4		
Plant	Spain	2018	40	38	2	55	90*	10	77.5*	15	80*	12.5	75*	30	75*	40	85*	15	85*	25.4	81.5		
Plant	Spain	2019	30	30	0	30	90*	13.3	90*	16.7	93.3*	16.7	90*	20	90*	66.7	90*	20	90*	26.2	90.5		
Insect	Spain	2018	26	18	8	65.4	80.8	7.7	73.1*	19.2	69.2*	11.5	57.7*	7.7	53.8*	26.9	57.7*	15.4	73.1*	22	66.2		

^a(*) Asterisk indicates a significant ($P<0.05$) higher number of successful amplifications for nested-MLST as compared to conventional-MLST (Yuang et al., 2010) according to a Chi-square test. The test was conducted only for the Spanish samples on the number of samples, even if frequencies are indicated in the table for MLST-PCR.

Table 4: Allele numbers and STs obtained for fully typed samples in France and Spain for plant and insect samples. The numbers correspond to the names of the samples.

Country	Sample names	<i>cysG</i>	<i>gltT</i>	<i>holC</i>	<i>leuA</i>	<i>malF</i>	<i>nuoL</i>	<i>petC</i>	sequence type (ST)
France	<i>Spartium junceum</i> 2	7	3	3	3	3	3	3	ST7
France	<i>Polygala myrtifolia</i> 3, 4	3	3	3	3	3	3	3	ST6
France	<i>Genista corsica</i> 1	7	3	3	3	3	3	3	ST7
France	<i>Polygala myrtifolia</i> 5, 6	7	3	3	3	3	3	3	ST7
Spain	<i>Cistus albidus</i> 2	31	15	10	7	17	16	6	ST80
Spain	<i>Ficus carica</i> 1	1	1	1	1	1	1	1	ST1
Spain	<i>Helichrysum italicum</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Juglans regia</i> 1	1	1	1	1	1	1	1	ST1
Spain	<i>Lavandula angustifolia</i> 1	32	3	3	3	3	3	3	ST81
Spain	<i>Olea europaea</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Phagnalon saxatile</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Polygala myrtifolia</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Prunus armeniaca</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Prunus domestica</i> 1	32	3	3	3	3	3	3	ST81
Spain	<i>Prunus domestica</i> 2	3	3	3	3	3	3	3	ST6
Spain	<i>Prunus dulcis</i> 4-8,10,11,15,18-26,30-47	3	3	3	3	3	3	3	ST6
Spain	<i>Prunus dulcis</i> 9	31	15	10	7	17	16	6	ST80
Spain	<i>Prunus dulcis</i> 1,2	32	3	3	3	3	3	3	ST81
Spain	<i>Prunus dulcis</i> 3	7	3	3	3	3	3	3	ST7

Spain	<i>Rhamnus alaternus</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Rosmarinus officinalis</i> 4	3	3	3	3	3	3	3	ST6
Spain	<i>Rosmarinus officinalis</i> 1,3	31	15	10	7	17	16	6	ST80
Spain	<i>Prunus domestica</i> 3	3	3	3	3	3	3	3	ST6
Spain	<i>Philaenus spumarius</i> 6,7,8,10,11	1	1	1	1	1	1	1	ST1
Spain	<i>Philaenus spumarius</i> 1	3	3	3	3	3	3	3	ST6
Spain	<i>Philaenus spumarius</i> 22	32	3	3	3	3	3	3	ST81
Spain	<i>Neophilaenus campestris</i> 1,2	3	3	3	3	3	3	3	ST6

1

2 **Supplementary Material**

3 S1 Table: List of *X. fastidiosa* genome sequences used in this study for primer and probe design
4 (Denancé et al. 2019)

5 S2 Table : Primers properties

6 S3 Table: detection threshold for Harper's qPCR test using genomic DNA dilution range (1: 220
7 ng.mL⁻¹; 2: 22 ng.mL⁻¹; 3 : 2.2 ng.mL⁻¹; 4: 220 pg.mL⁻¹; 5: 22 pg.mL⁻¹; 6: 2.2 pg.mL⁻¹; 7 : 220 fg.mL⁻¹
8 ¹; 8: 22 fg.mL⁻¹)

9 S4 Table: results obtained with qPCR and nested-MLST. (+) means that a signal has been obtained in
10 PCR but the PCR product has not been sequenced.

11

12 S1 Figure: detection threshold of conventional PCR for *cysG* and *malF* loci (Yuan et al., 2010) with
13 and without BSA (final concentration at 0.3 µg. µL⁻¹) using genomic DNA dilution range (1: 220
14 ng.mL⁻¹; 2: 22 ng.mL⁻¹; 3 : 2.2 ng.mL⁻¹; 4: 220 pg.mL⁻¹; 5: 22 pg.mL⁻¹; 6: 2.2 pg.mL⁻¹; 7 : 220
15 fg.mL⁻¹; 8: 22 fg.mL⁻¹). (+) positive control; (-) negative control.

16

