

Figure S1. Illustration of symptoms observed in coffee plants, bacterial cells and colonies isolated from these symptoms. Coffee leaf scorch on (A) a *Coffea arabica* plant originating from Ecuador, from which Strain CFBP 8072 was isolated and (B) on a *C. canephora* plants originating from Mexico, from which Strain CFBP 8073 was isolated. Cells of CFBP 8072 (C) and CFBP 8073 (D) visualized by immunofluorescence using a *Xylella fastidiosa*-specific antibody. Colony of CFBP 8072 (E) and CFBP 8073 (F) growing on PWG and B-CYE media, respectively.

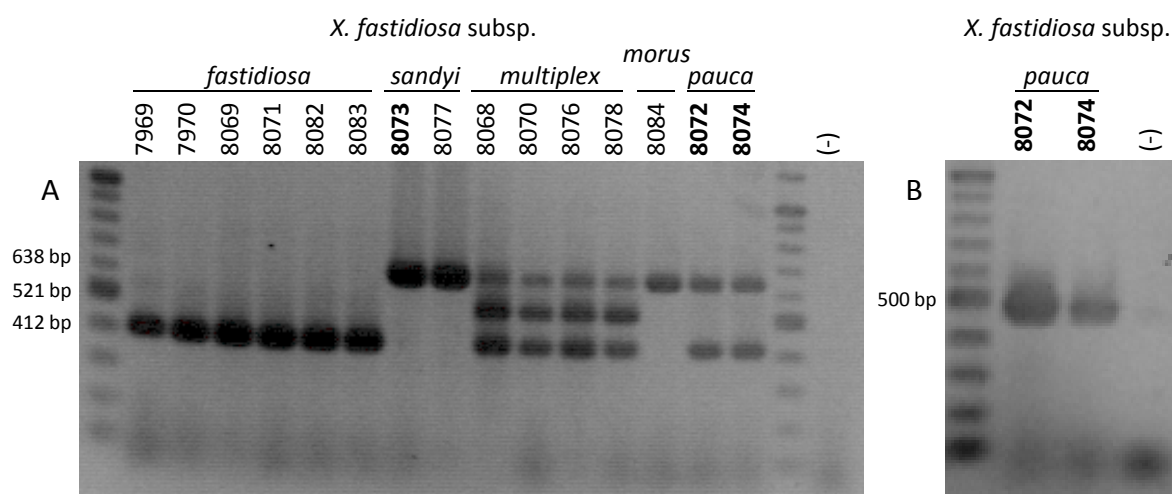


Figure S2. Gel photographs showing the signals obtained for the three *Xylella* strains (CFBP 8072, CFBP 8073, and CFBP 8074) isolated from coffee cuttings in comparison to strains representing the subspecies *fastidiosa*, *sandyi*, *multiplex*, and *morus* of *X. fastidiosa* with the identification tests described by Hernandez-Martinez et al. (1) (a) and Pooler and Hartung (2)(b). The CFBP code of each strain is indicated above each lane. (-) indicates negative control, i.e. tests with water.

1. Hernandez-Martinez, R, Costa HS, Dumenyo CK, Cooksey DA. 2006. Differentiation of strains of *Xylella fastidiosa* infecting grape, almonds and oleander using a multiprimer PCR assay. *Plant Dis.* **90**:1382-1388.

2. Pooler MR, Hartung JS. 1995. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. *Curr Microbiol* **31**:377-381.

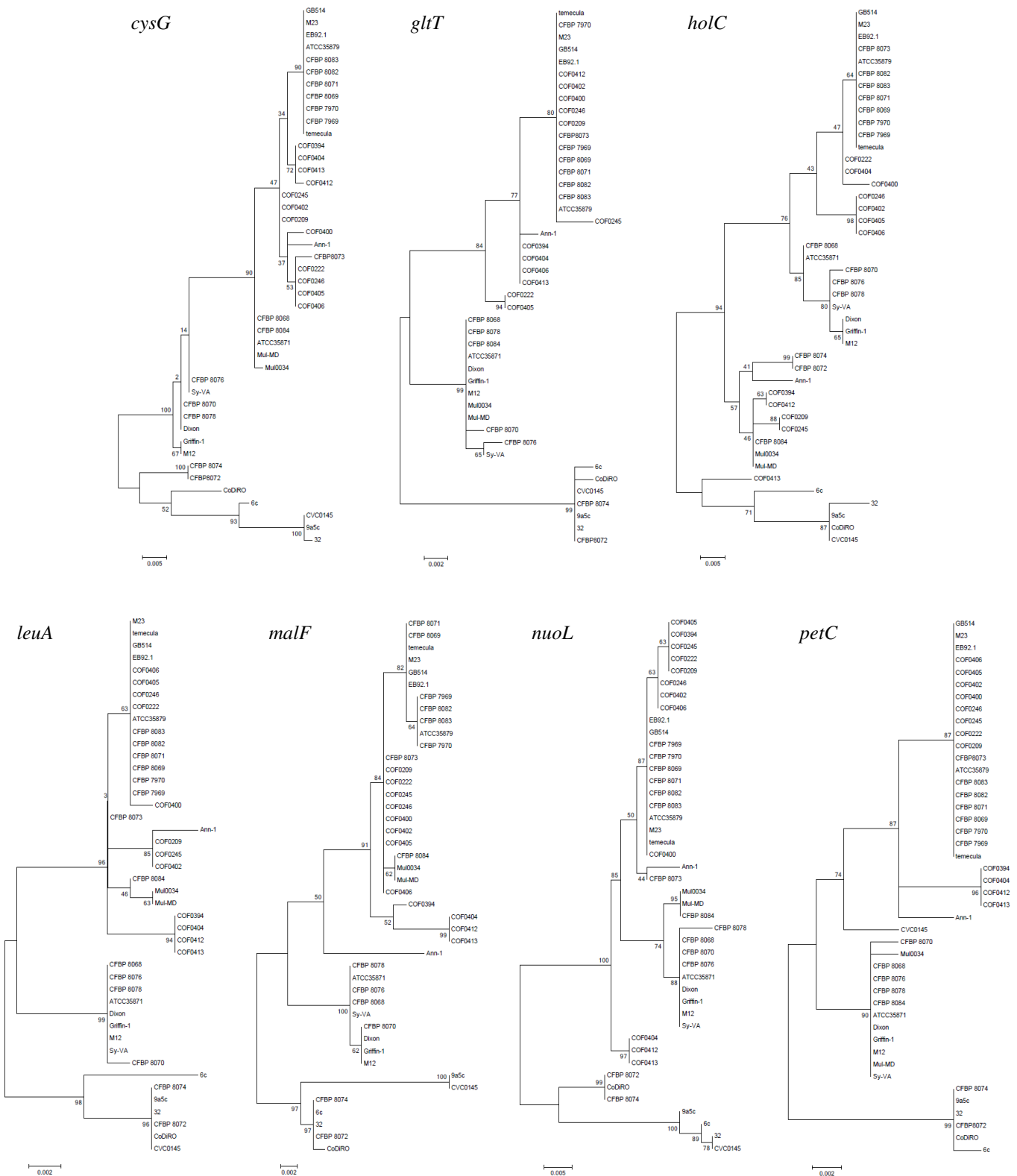


Figure S3. Maximum likelihood trees based on the partial sequences of *cysG*, *gltT*, *holC*, *leuA*, *malF*, *nuoL*, *petC*. Bootstrap scores (1,000 replicates) are displayed at each node.

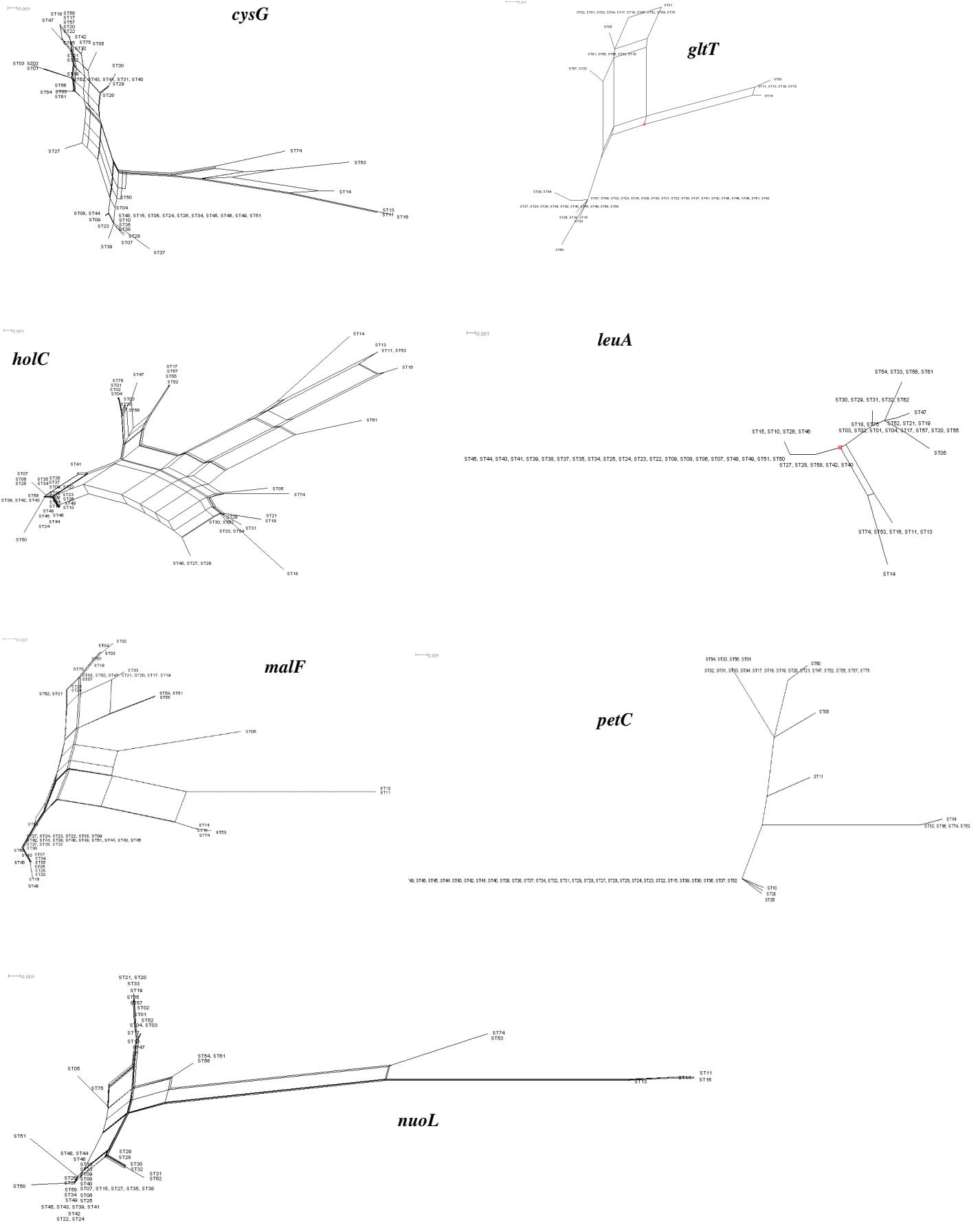


Figure S4. Split graphs of multilocus sequence analysis of the *X. fastidiosa* strains of each sequence type (ST) for the 7 loci (*cysG*, *gltT*, *holC*, *leuA*, *malF*, *nuoL* and *petC*). The designation at each of the leaves indicates the ST number. See Tables 1 and 4 for strain designations and ST correspondences, respectively.

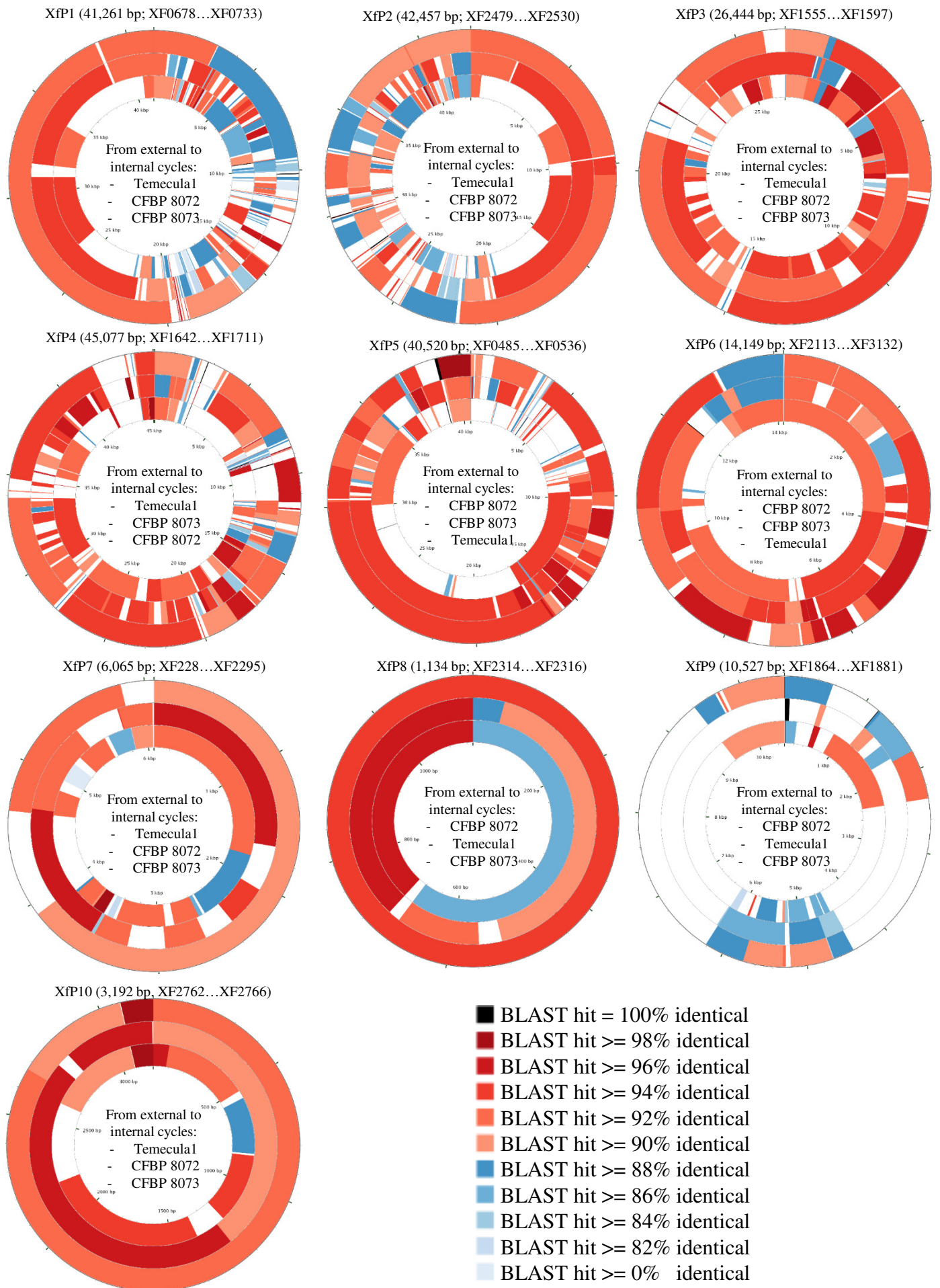


Figure S5. Conservation of Xfp1-10 phages from *X. fastidiosa* Strain 9a5c in genome sequences of CFBP 8072, CFBP 8073 and Temecula-1.



Figure S6. Conservation of Xfp1-8 phages from *X. fastidiosa* Strain Temecila-1 in genome sequences of CFBP 8072, CFBP 8073 and 9a5c.

Methionyl-t-RNA-synthetase (*metG*)

ATCC35871 (F820_RS0104705) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Griffin-1 (M233_00365) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

M12 (XfasM12_1753) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Mul0034 (P303_07735) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Mul-MD (P910_001364) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Sy-VA (DF22_001966) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Dixon (XFASADRAFT_RS00950) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Ann-1 (D934_RS11260) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

EB92.1 (XFEB_00975) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

GB514 (XFLM_02185) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

M23 (XfasM23_1678) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

Temecula1 (PD1590) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

9a5c (XF0549) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

CoDiRO (RA12_02050) **AACCGCATCCCATGTCAAACACAGCACCACCA**
 N A S M S N T A P P

CFBP 8072 (XYFPCFBP8072_i00120) **AACCGC** **TCCCATGTCAAACACAGCACCACCA** (2x)
 N A S M S N T A P P

CFBP 8073 (XYFPCFBP8073_aa00390) **AACCGC** **TCCCATGTCAAACACAGCACCACCA** (1x)
 N A S M S N T A P P

6c (B375_RS16320) **AACCGC** **TCCCATGTCAAACACAGCACCACCA** (5x)
 N A S M S N T A P P

32 (B398_RS16135) **AACCGC** **TCCCATGTCAAACACAGCACCACCA** (1x)
 N A S M S N T A P P

Hypothetical protein

Mul0034 (P303_05870) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q M E M F R

Mul-MD (P910_RS09730) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q M E M F R

Ann-1 (D934_RS00755) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

EB92.1 (XFEB_00416) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

GB514 (XFLM_05125) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E L V E M F R

M23 (XfasM23_1318) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

Temecula1 (PD1239) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

9a5c (XF2190-2191) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

CoDiRO (RA12_11140) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E L V E M F R

CFBP 8072 (XYFPCFBP8072_cc00010) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

CFBP 8073 (XYFPCFBP8073_dx00010) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

6c (B375_RS18005) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

32 (B398_RS17775) **AGGAACAGCATGGAGATGTTTCAGAAA**
 E Q V E M F R

Von Willebrand factor type A

ATCC35871 (F820_RS0101115) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

Griffin-1 (M233_11120) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

M12 (XfasM12_0914) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

(XfasM12_0653) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

Mul0034 (P303_03805) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

(P303_07930) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

Mul-MD (P910_002361) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

Sy-VA (DF22_RS10795) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

Dixon (XFASADRAFT_RS01420) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

(XFASADRAFT_RS02665) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

(XFASADRAFT_RS11395) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

Ann-1 (D934_RS07330) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

(D934_RS08480) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

(D934_RS09595) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

EB92.1 (XFEB_00076) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

GB514 (XFLM_09455) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

M23 (XfasM23_0811) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

Temecula1 (PD0767) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

9a5c (XF0526) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D A I L

(XF2115) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D A I L

CoDiRO (RA12_09015) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D A I L

CFBP 8072 (XYFPCFBP8072_bs00010) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

CFBP 8073 (XYFPCFBP8073_fd00010) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

6c (B375_RS18275) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

(B375_RS15870) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

32 (B398_RS18210) **CCGATCCTCACCAGCAGGATGCCATCCTGC**
 P I L T D D D I L

Figure S7. Candidate genes with k-mer fragments specific to four coffee-infecting strains of *X. fastidiosa*. For each candidate gene, the region containing the k-mer specific to coffee strains (*X. fastidiosa* subsp. *pauca* Strains 6c, 32 and CFBP 8072 and *X. fastidiosa* subsp. *fastidiosa/sandyi* Strain CFBP 8073) was aligned with corresponding regions in orthologs from other *X. fastidiosa* strains. The strain code is indicated for each strain followed by the locus tag of the candidate gene in that genome sequence.

Cellobiosidase

ATCC35871 (F820_RS12200)	TC CAG C C C C G C T T C T A G C T T C S S S G S S
Griffin-1 (M233_RS19950)	TC CAG C C C C A C C C C C G C T T C S S S S S S G
M12 (Xfasm12_0598)	TC CAG C C C C A C C C C C G C T T C S S S S S S G
Mul0034 (P303_02185-P303_02190)	TC CAG C C C C G C T T C T A C T T C S S S G S T
Mul-MD (P910_RS04945)	TC CAG C C C C G C T T C T A C T T C S S S G S T
Sy-VA (DF22_RS00940)	TC CAG C C C C A C C C C C G C T T C S S S S S S G
Dixon (XFASADRAFT_RS02915)	TC CAG C C C C A C C C C C G C T T C S S S S S S G
Ann-1 (D934_RS03550)	TC A A C C C C C G C T T C T A G C T T C S S S G S S
EB92.1 (XFEB_01892)	TC T A G T T C T A C T T C T A G C T T C S S S S S S
GB514 (XFLM_08105-XFLM_08110)	TC T A G T T C T A C T T C T A G C T T C S S S S S S
M23 (XfasM23_0553)	TC T A G T T C T A C T T C T A G C T T C S S S S S S
Temecula1 (PD0529)	TC T A G T T C T A C T T C T A G C T T C S S S S S S
9a5c (XF1267)	TC CAG C C C C G C T T C T A G C T T C S S S G S S
CoDiRO (RA12_05855-RA12_05860)	TC CAG C C C C G C C W A A A A A C C S S S G * K
CFBP 8072 (XYFPCFBP8072_p00410-XYFPCFBP8072_p00420)	TC C G C C C C C G C T T C T A G C T T C S G S G S S
CFBP 8073 (XYFPCFBP8073_cm00010)	TC C G C C C C C G C T T C T A G C T T C S G S G S S
6c (B375_RS0113680)	TC C G C C C C C G C T T C T A G C T T C S G S G S S
32 (B398_RS0113725)	TC C G C C C C C G C T T C T A G C T T C S G S G S S

Phage antirepressor

ATCC35871 (F820_RS0109215)	G R R G T G R G G T A R C G T R G C C G G T A A R C G T V A S
(F820_RS0109470) (intergenic; 245bp-upstream)	G R R G T G R G G T A R C G T R G C C G G T A A
Griffin-1 (M233_RS22595) (intergenic; 244bp-upstream)	A T G R R G T G R G G T A R C G T R G C C G G T A A
M12 (Xfasm12_1162)	A T G R R G T G R G G T A R C G T R G C C G G T A A R C G T V A S
Mul0034 (P303_02835)	A T G R R G T G R G G T A R C G T R G C C G G T A A R C G T V A S
(P303_09510) (intergenic; 241bp-upstream)	A T G R R G T G R G G T A R C G T R G C C G G T A A
Mul-MD (P910_RS10720) (intergenic; 246bp-upstream)	G R R G T G R G G T A R C G T R G C C T A G T A A
(P910_RS10770)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
Sy-VA (DF22_RS10700)	G R R G T G R G G T A R C G T R G C C G G T A A R C G T V A S
Dixon (XFASADRAFT_RS05730) (intergenic; 241bp-upstream)	A A R G R T G R G G T A R C G T R G C C G G T A A
Ann-1 (D934_RS05835) (intergenic; 248bp-upstream)	G R R G T G R G G T A R C G T R G C C G G T A A
(D934_RS06120) (intergenic; 244bp-upstream)	G R R G T G R G G T A R C G T R G C C G G T A A
EB92.1 (XFEB_02046)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
(XFEB_02052)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
GB514 (XFLM_11365) (intergenic; 244bp-upstream)	G R R G T G R G G T A R C G T R G C C T A G T A A
M23 (XfasM23_1182)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
(XfasM23_1184)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
Temecula1 (PD1117)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
(PD1119)	G R R G T G R G G T A R C G T R G C C T A G T A A R C G T V A S
9a5c (XF2507) (intergenic; 145bp-upstream)	A R R G R T G R G G T A R C G T R G C C T G G T A A
CoDiRO (RA12_07555) (intergenic; 37bp-downstream)	G R R G T G R G G T A R C G T R G C C G G T A A
CFBP 8072 (XYFPCFBP8072_ap00120)	A R R G R T G R G G T A R C G T R G C C G G T A A R C G T V A S
CFBP 8073 (XYFPCFBP8073_aq00060)	A R R G R T G R G G T A R C G T R G C C G G T A A R C G T V A S
6c (B375_RS20650) (intergenic; 243bp-upstream)	A R R G R T G R G G T A R C G T R G C C G G T A A
32 (B398_RS0114070) (intergenic; 244bp-upstream)	A R R G R T G R G G T A R C G T R G C C G G T A A

Figure S7 (continued)

Table S1. *P* values determined using the Shimodaira-Hasegawa test of tree topologies run on each of the maximum likelihood trees based on the 7 loci and the data set of concatenated sequences.

	<i>cysG</i>	<i>glfT</i>	<i>holC</i>	<i>leuA</i>	<i>malF</i>	<i>nuoL</i>	<i>petC</i>	Conc ¹
<i>cysG</i>	-	0.005	0.000	0.000	0.000	0.018	0.027	0.000
<i>glfT</i>	0.000	-	0.000	0.000	0.000	0.000	0.000	0.000
<i>holC</i>	0.000	0.000	-	0.000	0.000	0.000	0.000	0.000
<i>leuA</i>	0.000	0.000	0.000	-	0.008	0.000	0.007	0.000
<i>malF</i>	0.000	0.000	0.000	0.066	-	0.000	0.005	0.000
<i>nuoL</i>	0.000	0.000	0.000	0.007	0.000	-	0.000	0.000
<i>petC</i>	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000
Conc	0.415	0.019	0.100	0.215	0.624	0.077	0.105	-

¹Concatenated data set

Table S2. Sequence variation at the seven loci among the 44 *X. fastidiosa* strains used in the MLSA.

Locus	GC %	Sites ^a	S ^b	Hap ^c	Hd ^d	θ_{π} ^e	θ_w ^f	Tajima's D ^g	Fu & Li's D ^g	Fu 's F ^g
<i>cysG</i>	61.8	600	54	18	0.915	0.01825	0.20690	-0.52740	0.17484	0.290
<i>gltT</i>	53.3	654	25	12	0.800	0.01212	0.00880	0.92589	0.37741	2.359
<i>holC</i>	61.6	379	37	17	0.906	0.02466	0.02250	0.13773	0.44541	0.074
<i>leuA</i>	52.9	708	30	12	0.833	0.00953	0.00974	-0.18139	-0.42531	1.527
<i>malF</i>	50.6	730	46	12	0.905	0.01441	0.01463	-0.19598	0.39165	4.108
<i>nuoL</i>	51.0	557	42	13	0.876	0.01932	0.01836	-0.06145	1.19143	3.076
<i>petC</i>	51.5	533	24	9	0.735	0.01095	0.01035	0.19324	-0.41645	3.094
Concat. ^h	54.2	4161	259	32	0.970	0.01493	0.01441	-0.03785	0.36260	1.172

^anumber of analyzed sites

^bnumber of polymorphic sites

^cnumber of haplotypes

^dhaplotype diversity (**Nei M.** 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.)

^enucleotide diversity (**Nei M.** 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.)

^fnucleotide diversity from S (**Watterson GA.** 1975. On the number of segregating sites in genetical models without recombination. Theoret. Pop. Biol. 7:256-276.)

^gneutrality tests of Tajima (**Tajima F.** 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.) and Fu & Li (**Fu YX, Li WH.** 1993. Statistical tests of neutrality of mutations. Genetics 133:693–709.) and associated P-value: (* $p < 0.05$)

^hdata set of concatenated sequences

Table S3. Average nucleic identity (ANIb) values derived from 19 *X. fastidiosa* genome sequence pairwise comparisons.

Strain code	PLS 229	ATCC 35871	Sy- VA	M12	Griffin	Dixon	Mul- MD	Mul 0034	Ann-1	CFBP 8073	Temecula1	M23	GB 514	EB 92.1	9a5c	6c	32	CoDi- RO	CFBP 8072
PLS229	---	83.64	83.56	83.64	83.65	83.59	83.47	83.41	83.44	83.49	83.41	83.40	83.46	83.51	83.77	83.77	83.80	83.77	83.77
ATCC 35871	83.89	---	99.45	99.39	99.40	99.43	98.26	98.25	97.60	97.72	97.74	97.73	97.76	97.72	96.30	96.23	96.31	96.30	96.55
Sy-VA	83.70	99.38	---	99.31	99.31	99.36	98.17	98.17	97.50	97.57	97.70	97.69	97.72	97.65	96.18	96.17	96.23	96.15	96.45
M12	83.74	99.31	99.35	---	99.95	99.68	97.85	97.79	97.51	97.47	97.55	97.53	97.52	97.50	96.15	96.22	96.23	96.21	96.48
Griffin	83.77	99.39	99.37	99.98	---	99.69	97.92	97.90	97.49	97.51	97.56	97.56	97.54	97.52	96.21	96.22	96.21	96.26	96.56
Dixon	83.59	99.33	99.27	99.58	99.57	---	97.86	97.82	97.44	97.40	97.46	97.55	97.44	97.40	96.03	96.05	96.10	96.13	96.39
Mul-MD	83.51	98.19	98.10	97.73	97.78	97.76	---	99.77	97.99	98.29	98.40	98.37	98.40	98.49	95.81	95.78	95.83	95.97	96.20
Mul0034	83.30	98.06	97.95	97.61	97.65	97.64	99.79	---	97.89	98.22	98.33	98.29	98.39	98.42	95.67	95.65	95.74	95.80	96.08
Ann-1	83.38	97.41	97.31	97.12	97.16	97.16	97.89	97.82	---	98.12	97.97	97.97	97.99	98.02	95.65	95.77	95.74	95.93	96.16
CFBP 8073	83.47	97.39	97.32	97.13	97.17	97.16	98.11	98.01	97.97	---	98.75	98.72	98.74	98.84	95.65	95.76	95.77	95.84	96.14
Temecula1	83.40	97.68	97.58	97.35	97.4	97.36	98.41	98.39	98.04	98.89	---	99.90	99.91	99.71	95.82	95.80	95.90	95.92	96.23
M23	83.33	97.69	97.56	97.35	97.36	97.42	98.32	98.30	98.00	98.83	99.87	---	99.83	99.64	95.79	95.78	95.87	95.93	96.19
GB514	83.44	97.65	97.58	97.40	97.43	97.42	98.43	98.39	98.09	98.82	99.93	99.89	---	99.72	95.83	95.83	95.90	95.91	96.25
EB92.1	83.48	97.63	97.59	97.33	97.33	97.26	98.45	98.44	98.09	98.78	99.64	99.62	99.67	---	95.74	95.73	95.78	95.97	96.15
9a5c	83.73	96.17	96.12	96.07	96.13	96.04	95.85	95.76	95.83	95.95	95.84	95.86	95.84	95.85	---	98.85	99.16	98.10	97.96
6c	83.91	96.24	96.16	96.17	96.18	96.13	95.83	95.82	95.90	95.95	95.85	95.82	95.84	95.86	98.86	---	98.88	98.19	97.96
32	83.92	96.09	96.00	95.89	95.93	95.89	95.68	95.64	95.68	95.82	95.8	95.77	95.75	95.80	99.06	98.77	---	97.98	97.86
CoDiRO	83.92	96.27	96.21	96.16	96.18	96.16	96.07	96.01	96.06	96.19	96.08	96.06	96.05	96.11	98.21	98.25	98.28	---	98.66
CFBP 8072	83.72	96.19	96.06	96.09	96.08	96.06	95.96	95.92	96.01	96.02	96.02	95.99	96.09	96.07	97.74	97.67	97.86	98.29	---