

RESEARCH

Open Access



# Comparative genomics of *Xanthomonas fragariae* and *Xanthomonas arboricola* pv. *fragariae* reveals intra- and interspecies variations

Michael Gétaz<sup>1</sup>, Jochen Blom<sup>2</sup> , Theo H. M. Smits<sup>1\*</sup>  and Joël F. Pothier<sup>1</sup> 

## Abstract

The quarantine bacterium *Xanthomonas fragariae* causes angular leaf spots on strawberry. Its population structure was recently found to be divided into four (sub)groups resulting from two distinct main groups. *Xanthomonas arboricola* pv. *fragariae* causes bacterial leaf blight, but the bacterium has an unclear virulence status on strawberry. In this study, we use comparative genomics to provide an overview of the genomic variations of a set of 58 *X. fragariae* and five *X. arboricola* pv. *fragariae* genomes with a focus on virulence-related proteins. Structural differences within *X. fragariae* such as differential plasmid presence and large-scale genomic rearrangements were observed. On the other hand, the virulence-related protein repertoire was found to vary greatly at the interspecies level. In three out of five sequenced *X. arboricola* pv. *fragariae* strains, the major part of the Hrp type III secretion system was lacking. An inoculation test with strains from all four *X. fragariae* (sub)groups and *X. arboricola* pv. *fragariae* resulted in an interspecies difference in symptom induction since no symptoms were observed on the plants inoculated with *X. arboricola* pv. *fragariae*. Our analysis suggests that all *X. fragariae* (sub)groups are pathogenic on strawberry plants. On the other hand, the first genomic investigations of *X. arboricola* pv. *fragariae* revealed a potential lack of certain key virulence-related factors which may be related to the difficulties to reproduce symptoms on strawberry and could question the plant-host interaction of the pathovar.

**Keywords:** Type III effector, Virulence, Pathogenicity test, Secretion system

## Background

Strawberry is a small fruit crop of great economic importance in the world (Amil-Ruiz et al. 2011). The worldwide strawberry production increased from 2.4 Mt in 1990 to 9.1 Mt in 2016, representing a progress from 5.1 to over 10.9 billion US\$ yearly (FAOSTAT 2020). Strawberry became a part of the major fruit industry for several countries (Kim et al. 2016). Strawberry cultivars exhibit diverse susceptibilities to a large variety of harmful organisms, reducing fruit quality and plant yield

production (Simpson 1991; Maas 1998). Such diseases cause economic losses in strawberry fields and require to develop corresponding control measures (Amil-Ruiz et al. 2011). One of the main bacterial diseases affecting strawberry is caused by *Xanthomonas fragariae*, to which all commercial strawberry cultivars commercialized before 2003 were found to be susceptible (Hartung et al. 2003). Numerous strawberry cultivars including wild species and advanced breeding clones from breeding programs were assessed for resistance to *X. fragariae*, leading to the result that four resistant genotypes were detected (Maas et al. 2000; Hartung et al. 2003; Roach et al. 2016).

\* Correspondence: [theo.smits@zhaw.ch](mailto:theo.smits@zhaw.ch)

<sup>1</sup>Environmental Genomics and Systems Biology Research Group, Institute of Natural Resource Sciences, Zurich University of Applied Sciences (ZHAW), CH-8820 Wädenswil, Switzerland

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

*X. fragariae* is considered as quarantine organism by the European and Mediterranean Plant Protection Organization (OEPP/EPPO 1986), and the symptoms it caused are defined as angular leaf spots (ALS) affecting strawberry plant leaves. The bacterium was first described in 1960 from the USA (Kennedy and King 1960), and was subsequently found in most major strawberry producing regions worldwide (Zimmermann et al. 2004; OEPP/EPPO 2006). The disease begins with *X. fragariae* invading the plant through natural openings, such as stomata, hydathodes or wounds (Bestfleisch et al. 2015). The first symptoms occur as water-soaked bacterial lesions in the early stages and appear angular in shape. Then, the lesions spread over the foliage and form larger necrotic spots before the plants suffer from vascular collapse (Hildebrand et al. 1967). The artificial infection of *X. fragariae* displayed different disease incidences on strawberry cultivars indicated by variably severe symptoms on plant leaves (Bestfleisch et al. 2015).

In the last decade, genomic information of *X. fragariae* was made publicly available (Vandroemme et al. 2013a; Henry and Leveau 2016; Gétaz et al. 2017b). Four subgroups (*Xf*-CGr-IA, *Xf*-CGr-IB, *Xf*-CGr-IC and *Xf*-CGr-II) included in two major groups of strains (*Xf*-CGr-I and *Xf*-CGr-II) were defined using two types of molecular markers: variable numbers of tandem repeats (VNTRs) and clustered regularly interspaced short palindromic repeats (CRISPRs) (Gétaz et al. 2018b). When compared to other *Xanthomonas* genomes, the size of the genome of *X. fragariae* was smaller. Absent genes/regions are potentially involved in xylan degradation and metabolism, the  $\beta$ -ketoacid pathway, one of two type II secretion systems (T2SS) and the glyoxylate shunt pathway. The absence of these genes could possibly impact the plant-host interaction (Vandroemme et al. 2013a). However, a type III secretion system (T3SS), known to be essential for bacterial pathogenicity (Galán and Collmer 1999; Ghosh 2004), a distinct type III secretion system effector (T3E) repertoire, a type IV secretion system (T4SS) and a type VI secretion system (T6SS), all of which could play a role in specific and mostly endophytic association of *X. fragariae* with its plant host, were observed from the first draft genome (Vandroemme et al. 2013a).

*X. fragariae* was long considered as the only bacterial pathogen causing disease on strawberry (Kennedy and King 1962). However, in 1993, another causal agent, *Xanthomonas arboricola* pv. *fragariae* (Janse et al. 2001), was observed on strawberry plants in northern Italy (Scortichini 1996), and later in Turkey (Ustun et al. 2007). In contrast to the water-soaked regions on the leaves resulting from *X. fragariae*, the symptoms caused by *Xanthomonas arboricola* pv. *fragariae* were reddish-brown lesions on the leaf surface that enlarge and

become surrounded by a chlorotic halo (Ferrante and Scortichini 2018). *X. arboricola* pv. *fragariae* incites the leaf symptoms mainly in open-field cultivations and during mid-autumn weather conditions characterized by a very high relative air humidity (Scortichini and Rossi 2003). The pathogenicity of *X. arboricola* pv. *fragariae* upon artificial inoculation on strawberry plants was not always reproducible in glasshouse or in laboratory experiments and virulence among strains was variable (Scortichini and Rossi 2003; Vandroemme et al. 2013b; Merda et al. 2016). However, two studies could obtain symptoms like extensive vascular discoloration and wilting leaves after vein inoculations (Janse et al. 2001) as well as extensive necrotic lesions on the major leaf vein (Ferrante and Scortichini 2018). Genetic variability among Italian strains of *X. arboricola* pv. *fragariae* using repetitive PCR genomic fingerprinting revealed a high overall similarity of the pattern but with distinct genomic profiles (Scortichini and Rossi 2003). A multilocus sequence analysis (MLSA) showed that *X. arboricola* pv. *fragariae* strains are not in a monophyletic group, but were found to be spread within the *X. arboricola* clade (Vandroemme et al. 2013b).

In this study, we applied a comparative genomics workflow to assess the genomic variations within and between *X. fragariae* and *X. arboricola* pv. *fragariae* species. For *X. fragariae*, differential plasmid numbers and/or content, genomic rearrangements but a conserved virulence-related gene repertoire were revealed among strains. And meanwhile, all the *X. fragariae* strains tested were pathogenic on strawberry. For *X. arboricola* pv. *fragariae*, two groups of strains relative to its virulence-related gene repertoire were revealed, but none of the inoculated strains could incite symptoms on the tested strawberry cultivar.

## Results

### Genome sequences

As the genome sequences of *X. fragariae* and *X. arboricola* pv. *fragariae* strains included in this study were resulted from different sequencing technologies with different read lengths, the genomes vary in their total numbers of contigs and genome size (Table 1). Based on the genome data for *X. fragariae*, an average of 3960 genes was found through PacBio or Illumina MiSeq sequencing, whereas an average of 3510 genes was obtained when Illumina HiSeq sequencing was applied. The differential gene content of on average 450 CDS which were lacking in the Illumina HiSeq assemblies were examined by a comparison of the complete PacBio assembly of dually sequenced strains (PD 885<sup>T</sup>, NBC 2815, PD 5205) with their respective Illumina HiSeq assemblies. The lacking CDS in the gaps were identified as hypothetical proteins (51%), transposases (37%),

**Table 1** List of *Xanthomonas fragariae* and *Xanthomonas arboricola* pv. *fragariae* strains used in this study. The table provides geographic origin, host plant, year of isolation, CRISPR typing, genomic information such as genome length, contig number, GC content and number of CDS. The presence (+) or absence (-) of the pXf21 and pXf29 plasmids based on tBLASTn analysis of the origins of replication or based on BLASTn analysis of the complete plasmid sequences is also reported. Growth was assessed for each strain and the exponential time before stationary and generation time in hour are listed

Strain	Species	Geographic origin	Host plant	Year of isolation	(sub)group	Genome length	Contig number	GC content (%)	Sequencing technology	CDS	pXf21	pXf29	Exponential time before stationary (h)	Generation time (h)
PD 885 <sup>T</sup>	<i>X. fragariae</i>	US	<i>Fragaria chiloensis</i> var. <i>ananassa</i>	1960	Xf-CGr-IA	4,247,102	1 chr. + 2 plasmids	62.22	PacBio	4021	+	+	14	3.80
PD 4314	<i>X. fragariae</i>	US	<i>Fragaria chiloensis</i> var. <i>ananassa</i>	1960	Xf-CGr-IA	3,915,166	269	62.21	HiSeq	3526	+	+	28	4.52
PD 2659	<i>X. fragariae</i>	US	<i>Fragaria</i> sp.	1962	Xf-CGr-IA	3,906,466	307	62.22	HiSeq	3521	+	+	28	4.57
ICMP 659	<i>X. fragariae</i>	US	<i>Fragaria</i> sp.	1982	Xf-CGr-IA	4,261,751	430	62.22	MiSeq	4077	+	+	30	5.96
ICMP 661	<i>X. fragariae</i>	US	<i>Fragaria</i> sp.	1982	Xf-CGr-IA	4,237,456	439	62.25	MiSeq	4044	+	+	33	5.48
JVD-0051	<i>X. fragariae</i>	BE	NA	NA	Xf-CGr-IB	4,153,782	372	62.30	MiSeq	3931	+	-	25	4.37
Xf 61	<i>X. fragariae</i>	NL	var. Figaro	2010	Xf-CGr-IB	3,850,449	290	62.33	HiSeq	3428	+	-	26	4.57
PD 2905	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	NA	Xf-CGr-IB	3,884,146	309	62.22	HiSeq	3466	+	-	20	4.29
LMG 25863 <sup>a</sup>	<i>X. fragariae</i>	BE	<i>Fragaria</i> x <i>ananassa</i> Elsanta	2011	Xf-CGr-IC	4,183,685	95	62.18	HiSeq	3920	+	+	28	4.59
GBBC 2087	<i>X. fragariae</i>	BE	NA	NA	Xf-CGr-IC	4,193,468	414	62.25	MiSeq	4024	-	+	20	4.31
JVD-0046	<i>X. fragariae</i>	BE	NA	NA	Xf-CGr-IC	4,189,841	296	62.24	MiSeq	3987	+	+	25	4.21
JVD-0047	<i>X. fragariae</i>	BE	NA	NA	Xf-CGr-IC	4,151,553	413	62.31	MiSeq	3936	-	-	22	3.68
GBBC 2088	<i>X. fragariae</i>	BE	NA	NA	Xf-CGr-IC	4,186,171	414	62.31	MiSeq	3984	+	-	26	4.62
CFBP 5253	<i>X. fragariae</i>	FR	<i>Fragaria</i> sp.	1999	Xf-CGr-IC	4,144,074	282	62.30	MiSeq	3899	+	-	17	4.07
CFBP 5257	<i>X. fragariae</i>	FR	<i>Fragaria</i> sp.	1999	Xf-CGr-IC	4,168,345	436	62.27	MiSeq	3977	-	+	26	3.98
CFBP 5260	<i>X. fragariae</i>	FR	<i>Fragaria</i> sp.	1999	Xf-CGr-IC	4,182,173	413	62.27	MiSeq	3995	+	+	15	3.95
ICMP 13779	<i>X. fragariae</i>	IT	<i>Fragaria</i> x <i>ananassa</i> Duchesne	1996	Xf-CGr-IC	4,214,022	416	62.27	MiSeq	4017	+	+	24	4.69

**Table 1** List of *Xanthomonas fragariae* and *Xanthomonas arboricola* pv. *fragariae* strains used in this study. The table provides geographic origin, host plant, year of isolation, CRISPR typing, genomic information such as genome length, contig number, GC content and number of CDS. The presence (+) or absence (-) of the pXf21 and pXf29 plasmids based on tBLASTn analysis of the origins of replication or based on BLASTn analysis of the complete plasmid sequences is also reported. Growth was assessed for each strain and the exponential time before stationary and generation time in hour are listed (*Continued*)

Strain	Species	Geographic origin	Host plant	Year of isolation	(sub)group	Genome length	Contig number	GC content (%)	Sequencing technology	CDS	pXf21	pXf29	Exponential time before stationary (h)	Generation time (h)
ICMP 13780	<i>X. fragariae</i>	IT	<i>Fragaria x ananassa</i> Duchesne	1996	Xf-CGr-IC	4,177,334	395	62.28	MISeq	3978	+	+	24	4.13
PD 3160	<i>X. fragariae</i>	NL	NA	1996	Xf-CGr-IC	3,915,813	302	62.15	HiSeq	3533	+	+	18	4.28
PD 4634	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2003	Xf-CGr-IC	3,884,561	277	62.27	HiSeq	3493	+	+	24	4.41
PD 4932	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2004	Xf-CGr-IC	3,837,440	281	62.29	HiSeq	3540	+	+	28	5.21
PD 5205	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2005	Xf-CGr-IC	4,281,456	1 chr. + 2 plasmids	62.30	PacBio	4033	+	+	18	4.53
PD 5365	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2006	Xf-CGr-IC	3,877,123	298	62.27	HiSeq	3496	+	+	18	4.17
PD 5482	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2007	Xf-CGr-IC	3,888,298	286	62.27	HiSeq	3495	+	+	26	4.84
PD 6879	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2008	Xf-CGr-IC	3,869,542	296	62.27	HiSeq	3477	+	+	24	4.62
PD 6840	<i>X. fragariae</i>	NL	<i>Fragaria</i> sp.	2009	Xf-CGr-IC	3,870,901	288	62.28	HiSeq	3484	+	+	26	4.95
PD 6880	<i>X. fragariae</i>	NL	NA	2011	Xf-CGr-IC	3,868,438	279	62.28	HiSeq	3478	+	+	27	5.23
PD 6881	<i>X. fragariae</i>	NL	NA	2012	Xf-CGr-IC	3,883,071	287	62.27	HiSeq	3503	+	+	30	5.21
PD 6882	<i>X. fragariae</i>	NL	NA	2013	Xf-CGr-IC	3,884,412	274	62.28	HiSeq	3493	+	+	25	4.59
CFBP 4784	<i>X. fragariae</i>	PT	<i>Fragaria</i> sp. Thulda.	1995	Xf-CGr-IC	4,124,432	357	62.31	MISeq	3904	+	-	24	4.64
CFBP 4785	<i>X. fragariae</i>	PT	<i>Fragaria</i> sp.	1998	Xf-CGr-IC	4,125,383	283	62.29	MISeq	3868	+	-	18	4.43
CFBP 4786	<i>X. fragariae</i>	PT	<i>Fragaria</i> sp.	1998	Xf-CGr-IC	4,155,600	377	62.27	MISeq	3924	+	+	NA	NA
FoP29 <sup>b</sup>	<i>X. fragariae</i>	US	<i>Fragaria x ananassa</i> Portola	2011	Xf-CGr-IC	4,282,602	1 chr. + 2 plasmids	62.30	PacBio	3983	+	+	NA	NA
FoP21 <sup>b</sup>	<i>X. fragariae</i>	US	<i>Fragaria x ananassa</i> Portola	2011	Xf-CGr-IC	4,282,698	1 chr. + 2 plasmids	62.30	PacBio	3986	+	+	NA	NA

**Table 1** List of *Xanthomonas fragariae* and *Xanthomonas arboricola* pv. *fragariae* strains used in this study. The table provides geographic origin, host plant, year of isolation, CRISPR typing, genomic information such as genome length, contig number, GC content and number of CDS. The presence (+) or absence (-) of the pXf21 and pXf29 plasmids based on tBLASTn analysis of the origins of replication or based on BLASTn analysis of the complete plasmid sequences is also reported. Growth was assessed for each strain and the exponential time before stationary and generation time in hour are listed (*Continued*)

Strain	Species	Geographic origin	Host plant	Year of isolation	(sub)group	Genome length	Contig number	GC content (%)	Sequencing technology	CDS	pXf21	pXf29	Exponential time before stationary (h)	Generation time (h)
ICMP 20572	<i>X. fragariae</i>	US	<i>Fragaria x ananassa</i> Duchesne	2014	Xf-CGr-IC	4,248,096	455	62.22	MISeq	4119	+	+	27	5.21
ICMP 20573	<i>X. fragariae</i>	US	<i>Fragaria x ananassa</i> Duchesne	2014	Xf-CGr-IC	4,189,776	430	62.26	MISeq	4000	+	-	25	4.85
ICMP 20574	<i>X. fragariae</i>	US	<i>Fragaria x ananassa</i> Duchesne	2014	Xf-CGr-IC	4,226,734	426	62.24	MISeq	4068	+	+	23	4.52
ICMP 20575	<i>X. fragariae</i>	US	<i>Fragaria x ananassa</i> Duchesne	2014	Xf-CGr-IC	4,210,620	404	62.23	MISeq	4041	+	+	28	4.40
CFBP 6767	<i>X. fragariae</i>	AU	<i>Fragaria x ananassa</i> Duchesne	1975	Xf-CGr-II	4,129,734	293	62.31	MISeq	3870	-	-	13	3.91
ICMP 6646	<i>X. fragariae</i>	AU	<i>Fragaria x ananassa</i> Duchesne	1975	Xf-CGr-II	4,137,939	370	62.32	MISeq	3919	+	-	21	4.11
ICMP 6648	<i>X. fragariae</i>	AU	<i>Fragaria x ananassa</i> Duchesne	1975	Xf-CGr-II	4,153,233	396	62.31	MISeq	3950	+	-	22	4.37
ICMP 5813	<i>X. fragariae</i>	BR	<i>Fragaria x ananassa</i> Duchesne	1977	Xf-CGr-II	4,147,663	395	62.30	MISeq	3928	+	-	17	3.87
NCPPB 3743	<i>X. fragariae</i>	BR	<i>Fragaria</i> hybrid cv. Campinas	1990	Xf-CGr-II	4,121,221	257	62.30	MISeq	3861	+	-	19	3.87
F1-80	<i>X. fragariae</i>	CH	<i>Fragaria</i> cv. Polka	1993	Xf-CGr-II	4,107,680	264	62.31	MISeq	3829	-	-	22	3.61
F1-81	<i>X. fragariae</i>	CH	<i>Fragaria</i> sp.	1993	Xf-CGr-II	4,180,022	415	62.33	MISeq	3959	+	-	23	4.05
F2-02	<i>X. fragariae</i>	CH	<i>Fragaria</i> sp.	1994	Xf-CGr-II	4,122,612	273	62.31	MISeq	3875	+	-	27	4.70
CFBP 2510	<i>X. fragariae</i>	ES	<i>Fragaria</i> sp.	1982	Xf-CGr-II	4,136,197	280	62.29	MISeq	3886	+	-	18	3.91
PD 3145	<i>X. fragariae</i>	ES	<i>Fragaria</i> sp.	1997	Xf-CGr-II	3,838,331	315	62.32	HiSeq	3431	+	-	16	4.14
NBC 2815	<i>X. fragariae</i>	ES	<i>Fragaria</i> sp.	1997	Xf-CGr-II	4,268,155	1 chr. + 1 plasmid	62.31	PacBio	3939	+	-	18	5.10
CFBP 1558	<i>X. fragariae</i>	FR	<i>Fragaria</i> sp.	1974	Xf-CGr-II	4,155,692	284	62.30	MISeq	3909	+	-	18	3.80

**Table 1** List of *Xanthomonas fragariae* and *Xanthomonas arboricola* pv. *fragariae* strains used in this study. The table provides geographic origin, host plant, year of isolation, CRISPR typing, genomic information such as genome length, contig number, GC content and number of CDS. The presence (+) or absence (-) of the pXf21 and pXf29 plasmids based on tBLASTn analysis of the origins of replication or based on BLASTn analysis of the complete plasmid sequences is also reported. Growth was assessed for each strain and the exponential time before stationary and generation time in hour are listed (*Continued*)

Strain	Species	Geographic origin	Host plant	Year of isolation	(sub)group	Genome length	Contig number	GC content (%)	Sequencing technology	CDS	pXf21	pXf29	Exponential time before stationary (h)	Generation time (h)
PD 2662	<i>X. fragariae</i>	GR	<i>Fragaria</i> sp.	1979	Xf-CGr-II	3,825,417	276	62.31	HiSeq	3446	+	-	18	4.00
NCPBP 2473	<i>X. fragariae</i>	IT	<i>Fragaria vesca</i>	1972	Xf-CGr-II	4,169,971	319	62.30	MiSeq	3947	+	-	14	3.84
ICMP 6269	<i>X. fragariae</i>	NZ	<i>Fragaria</i> x <i>ananassa</i> Duchesne	1971	Xf-CGr-II	4,171,021	426	62.34	MiSeq	3954	+	-	17	4.61
ICMP 6277	<i>X. fragariae</i>	NZ	<i>Fragaria</i> x <i>ananassa</i> Duchesne	1972	Xf-CGr-II	4,178,618	426	62.33	MiSeq	3966	+	-	17	4.48
ICMP 6278	<i>X. fragariae</i>	NZ	<i>Fragaria</i> x <i>ananassa</i> Duchesne	1972	Xf-CGr-II	4,170,405	416	62.33	MiSeq	3984	-	-	19	4.14
PD 2660	<i>X. fragariae</i>	NZ	<i>Fragaria ananassa</i> var. <i>tioga</i>	1972	Xf-CGr-II	3,857,327	283	62.33	HiSeq	3453	+	-	20	4.21
ICMP 4019	<i>X. fragariae</i>	NZ	<i>Fragaria</i> x <i>ananassa</i> Duchesne	1974	Xf-CGr-II	4,160,354	405	62.31	MiSeq	3946	+	-	20	3.96
NCPBP 1822	<i>X. fragariae</i>	US	<i>Fragaria chiloensis</i> var. <i>ananassa</i>	1966	Xf-CGr-II	4,138,917	268	62.30	MiSeq	3877	-	-	24	3.92
LMG 19146	<i>X. arboricola</i> pv. <i>fragariae</i>	FR	NA	1986	NA	4,884,039	39	65.88	MiSeq	4020	NR	NR	8	1.56
CFBP 6773	<i>X. arboricola</i> pv. <i>fragariae</i>	NA	NA	NA	NA	4,692,498	39	65.95	MiSeq	3852	NR	NR	NA	NA
LMG 19145 <sup>PT</sup>	<i>X. arboricola</i> pv. <i>fragariae</i>	IT	<i>Fragaria</i> x <i>ananassa</i>	1993	NA	4,906,785	39	65.88	MiSeq	4019	NR	NR	8	1.56
LMG 19144	<i>X. arboricola</i> pv. <i>fragariae</i>	IT	<i>Fragaria</i> x <i>ananassa</i>	1993	NA	4,842,182	58	65.94	MiSeq	3954	NR	NR	7.4	1.50
CFBP 6762	<i>X. arboricola</i> pv. <i>fragariae</i>	IT	<i>Fragaria</i> x <i>ananassa</i>	1993	NA	4,889,284	38	65.75	MiSeq	4044	NR	NR	NA	NA

<sup>a</sup> Genome sequence data obtained from GenBank assembly accession GCA\_000376745.1 (Vandroemme et al. 2013a)

<sup>b</sup> Genome sequence data obtained from GenBank assembly accessions GCA\_001705565.1 and GCA\_001705545.1 (Henry and Leveau 2016)

NA: not applicable, NR: not relevant, AU: Australia, BE: Belgium, BR: Brazil, CH: Switzerland, ES: Spain, FR: France, GR: Greece, IT: Italy, NL: Netherlands, NZ: New Zealand, PT: Portugal, US: United States

integrase proteins (3%), and phage related proteins (3%). Some other annotated genes (6%) were expected to be located in observed gaps as they were surrounded on both sides by transposases leading to assembly ambiguities. The combination of highly repetitive regions and too short reads from HiSeq sequencing (120 bp) led to the result that more gaps remained in these genomes, but the low amount of CDS did not influence the current comparative genomics analysis.

#### Comparative genomics analysis based on whole-genome sequencing data

We have already reported the average nucleotide identities (ANI<sub>b</sub>) of the *X. fragariae* strains, and confirmed that they all belonged to the same species with identity values ranging between 99.48 and 99.97%, indicating close clonality of the isolates (Gétaz et al. 2018b). Here, we further evaluated the intraspecies and interspecies relatedness within and between these *X. fragariae* isolates and *X. arboricola* pv. *fragariae* isolates. For the analysis of the annotated genes, an amino acid identity (AAI) comparison between genome sequences of *X. fragariae* and *X. arboricola* pv. *fragariae* was computed with EDGAR (Additional file 1: Table S1). An overall AAI of 91.79 and 91.98% was found for interspecies relatedness. Intraspecies AAI within *X. fragariae* could discriminate both *Xf*-CGr-I and *Xf*-CGr-II groups with an average of 99.78% between groups, and 99.9–100% within groups.

The AAI comparison results within *X. arboricola* pv. *fragariae* isolates suggested that there were two groups, with AAI of 98.1–98.85% and greater than 99% between and within groups, respectively. The *X. arboricola* pv. *fragariae* strains CFBP 6773, LMG 19144 and LMG 19145<sup>PT</sup> grouped together whereas CFBP 6762 and LMG 19146 grouped separately. This corresponds well with the grouping of *X. arboricola* pv. *fragariae* strains in a partial *gyrB* sequence-based phylogeny, thus supporting the conclusion that *X. arboricola* is a polyphyletic group (Vandroemme et al. 2013b).

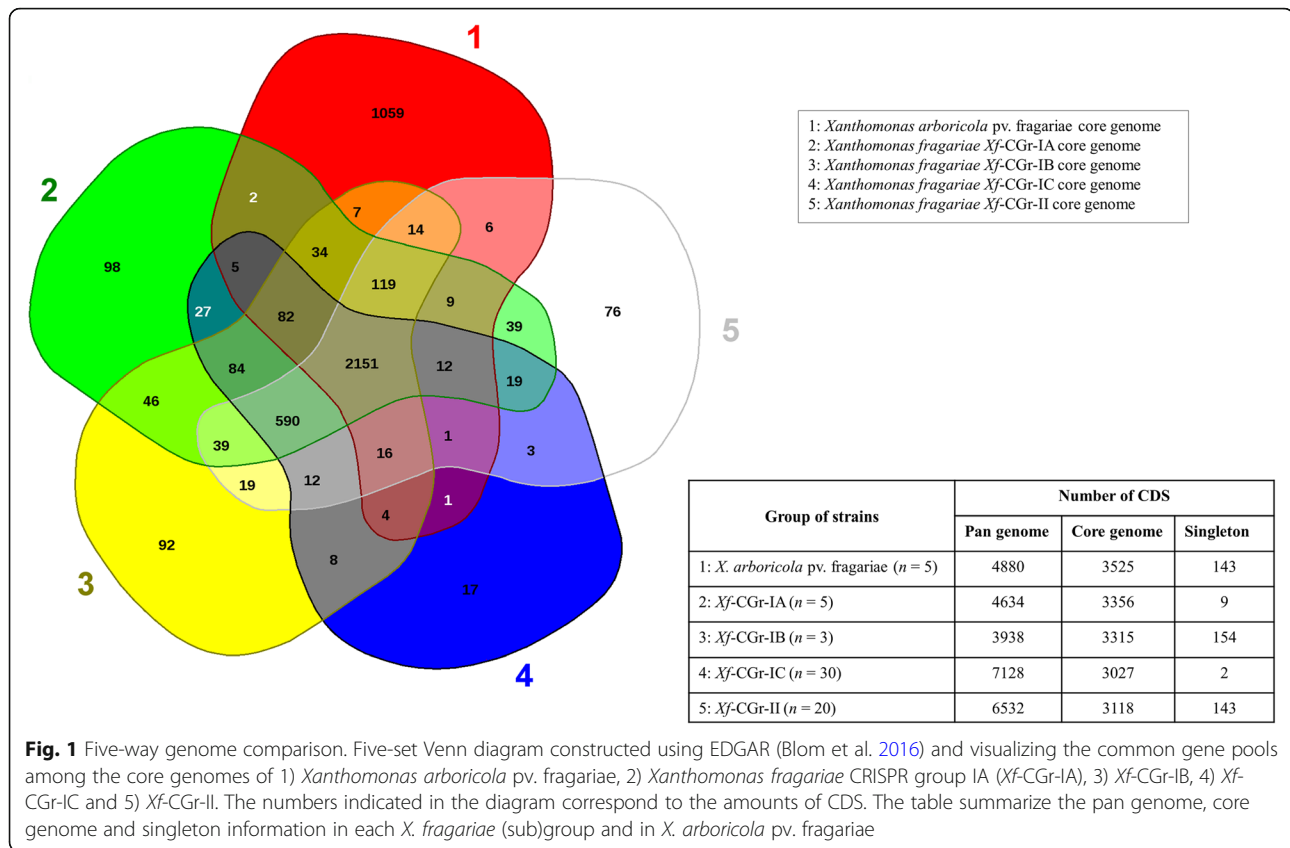
The alignment with MAUVE of the complete genomes of strains PD 5205, PD 885<sup>T</sup> and NBC 2815 (Gétaz et al. 2017b), belonging to three different *X. fragariae* (sub)groups (Gétaz et al. 2018b), showed an overall conservation of the genomes and illustrated that rearrangements of long syntenic regions between genomes has occurred (Additional file 2: Fig. S1). Two further complete genomes of strains *FaP21* and *FaP29* (Henry and Leveau 2016) were identical in structure to PD 5205, which belongs to the same (sub)group (i.e. *Xf*-CGr-IC; Additional file 2: Fig. S2). No large-scale genomic alterations such as insertions, deletions and duplications were observed between the aligned complete genomes, apart from the variable number of plasmids. Possibly, eight large-scale rearrangements have occurred between strains from *Xf*-

CGr-II and *Xf*-CGr-IA, whereas only two rearrangements were necessary to explain the differences between strains from *Xf*-CGr-IA and -IB. The synteny of genes flanking a recombination site indicated that these rearrangements were linked to the presence and/or activity of transposable elements.

EDGAR analysis comparing the CDS contents among *X. fragariae* (sub)groups and *X. arboricola* pv. *fragariae* genomes revealed a homogenous repartition of CDS between *X. fragariae* (sub)groups, while only 72% of the CDS from *X. arboricola* pv. *fragariae* were shared with *X. fragariae* (Fig. 1). The CDS that were found to be exclusively present in the *X. arboricola* pv. *fragariae* core genome ( $n = 1059$ , hypothetical = 270) represented 30% of its whole core genome and may reflect genes specific to the *X. arboricola* species level or even to the pathovar level. Among the CDS harbored uniquely in *X. arboricola* pv. *fragariae* strains, two clusters are potentially involved in degradation of lignin compounds. The presence of xylan degradation loci and a  $\beta$ -ketoacid pathway may indicate that *X. arboricola* pv. *fragariae* is able to degrade xylan and metabolize the phenolic monomeric components of lignin, two important elements of the secondary plant cell wall (Déjean et al. 2013). Here, 39 CDS suggested to function as plant cell wall degrading-enzymes (CWDE) were found uniquely in *X. arboricola* pv. *fragariae*. On the other hand, a smaller CWDE set was present in *X. fragariae* compared to other *Xanthomonas* (Vandroemme et al. 2013a).

In *X. arboricola* pv. *fragariae*, a distinct nitrogen assimilation cluster from nitrate form, possibly affecting cell metabolism and their cell growth (Snoeijers et al. 2000), was found. Additionally, the *gum*-associated genes *gumO*, and *gumP* were also found, but these genes were considered as unessential for xanthan biosynthesis and virulence (Lu et al. 2008). Subsequently, although one system was shared with *X. fragariae*, a second complete T2SS cluster containing proteins encoding both pilus (XpsG to XpsK) and membrane system (XpsC to XpsF, XpsL and XpsM) was found in *X. arboricola* pv. *fragariae* (Vandroemme et al. 2013a).

Common CDS to all *X. fragariae* (sub)groups (Fig. 1) represented a high percentage of core genomes in each (sub)group: 81.7% for *Xf*-CGr-IA, 82.6% for *Xf*-CGr-IB, 90% of *Xf*-CGr-IC and 87.7% of *Xf*-CGr-II. Of the remaining singletons, an overall proportion of 55–63% singleton CDS within each (sub)group was found to be hypothetical proteins. The number of CDS being common to all *X. fragariae* (sub)groups but absent from the *X. arboricola* pv. *fragariae* genomes was 590 CDS. Among them, a full T6SS consisting of 13 genes from both membrane complex (*tssA*, *tssK*, *tssL* and *tssM*) and phage-related complex (*tssA*, *tssB*, *tssC*, *hcp*, *tssE*, *tssF*, *tssG*, *clpV* and *vgrG*) required for an operational system



(Zoued et al. 2014) was present in all *X. fragariae* (sub)groups.

Overall, the *X. arboricola* pv. *fragariae* genomes did not contain phage-related genes. However, a cluster of phage genes integrated in the chromosome was found in the core genome of *X. fragariae*, and different clusters or remnants thereof with proteins annotated as phage-related were found in the singletons of *X. fragariae*.

In order to assess intra-(sub)group CDS variations, the generated pan genome was compared to all 58 *X. fragariae* strains. It confirmed an overall conservation of the genome content in *X. fragariae* (Fig. 2). However, in addition to CDS variations between (sub)groups, variations within the (sub)groups were observed suggesting that genetic variability also exists at a smaller scale. Overall, 13 out of 30 variable sites were found to be clusters of hypothetical proteins. Due to annotation limitations, no conclusion could be done about these variations.

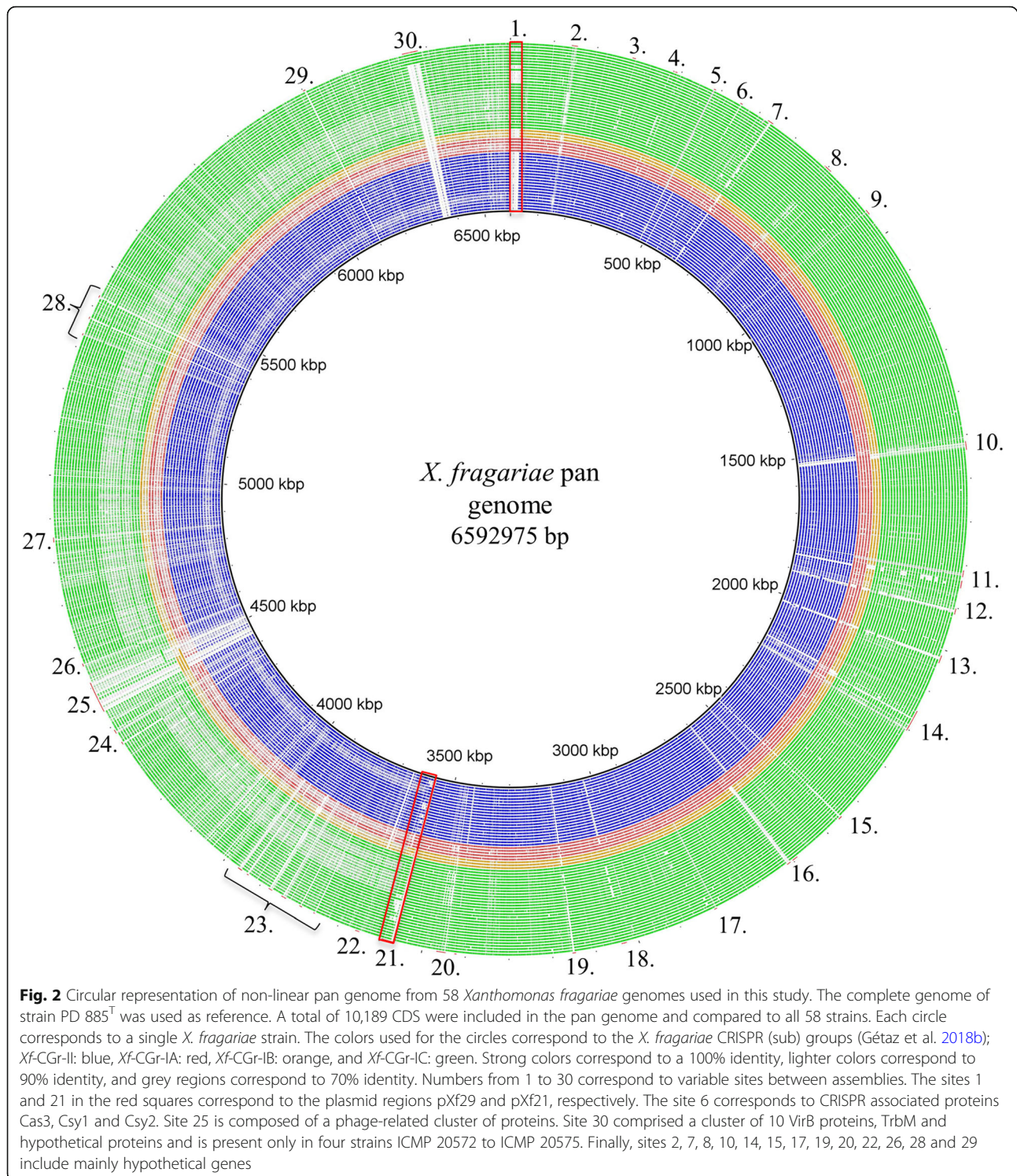
#### In silico screening of virulence-related genes

Of a collection of 163 virulence-related proteins, only 118 orthologous proteins were present in the strains included in this study (Fig. 3). The group of absent proteins included mostly T3E (Xop), and this repertoire was already reported to be smaller in *X. fragariae* compared

to other *Xanthomonas* (Vandroemme et al. 2013a). A total of 27 T3SS proteins, 31 flagellar T3SS, 27 T3E, 7 T4SS, 17 T6SS, 2 LPS and 7 EPS synthesis proteins were conserved in all *X. fragariae* (sub)groups (Fig. 3). The absence of XopE4 from twelve genomes of *X. fragariae* strains belonging to different (sub)groups can be explained by the presence of a gap at the corresponding position in the draft genome assemblies resulting from both Illumina MiSeq and HiSeq technologies. A cluster of T4SS proteins (VirB2, VirB3, VirB4, VirB6, VirB9 and VirB11) was found to be located on the chromosome and surrounded by transposases, suggesting being the remainder of an integrated plasmid. An additional tBLASTn search of these proteins in the five full-genome assemblies (Henry and Leveau 2016; Gétaz et al. 2017b) revealed another set of T4SS proteins with identity below 40% on plasmid regions. Paralogs of *virB11* and *virB6* were found in all five complete genome sequences, whereas the paralogs for *virB4*, *virB8* and *virD4* could only not be identified in strain NBC2815.

At the amino acid level, a limited number of variations were observed that may discriminate *X. fragariae* strains to the level of their (sub)group. Most of the virulence-related proteins with non-synonymous SNPs between (sub)groups had only 1 to 4 amino acid changes, which probably represent drift variations (between 0.18 to 1%





of total amino acid positions per protein, dependent on the size of the protein). These values correspond to the relative variation obtained by the intraspecies *X. fragariae* AAI comparison, approximating variations up to 0.25% (Additional file 1: Table S2).

If variation was observed, not more than two different alleles were observed per protein. Four proteins had a larger number of non-synonymous SNPs between *Xf-CGr-I* and *Xf-CGr-II*: GumE (2.5% amino acids substitutions), GumK (4.5%), XopAE (2.95%) and XopC (6.2%).



(See figure on previous page.)

**Fig. 3** Variable virulence-related genes within and between *X. fragariae* and *X. arboricola* pv. *fragariae* genomes. Comparison between genomes and protein sequences of virulence-related genes from type III secretion system (T3SS), flagellar-related T3SS, type III secretion effectors (T3E), type IV secretion system (T4SS), type VI secretion system (T6SS), lipopolysaccharide (LPS) and extracellular polysaccharides (EPS) synthesis proteins. Variations were observed between strains and reported with color-coded labels; red: absence, dark green: presence with 100% identical to reference, light green: truncated sequence due to end of contig, yellow: sequences variation between 1 and 14 amino acids, orange: amino acid variation above 14 different amino acids

The higher number of non-synonymous SNPs in these proteins may be indicative of a positive selection in these virulence-related genes. Transcription activator-like (TAL) effectors (Boch and Bonas 2010) were not found in the *X. fragariae* genomes with the exception of the non-TAL effector-like AvrBs2 (Kearney and Staskawicz 1990).

Compared to *X. fragariae*, *X. arboricola* pv. *fragariae* harbored a smaller virulence-related protein repertoire (Fig. 3). The genome sequences identified the same two groups of *X. arboricola* pv. *fragariae* as resulting from the AAI analysis (Additional file 1: Table S2) based on the differences in the virulence-related genes repertoire size. Strains CFBP 6762 and LMG 19146 encoded 25 and 26 T3SS proteins respectively, whereas only two of these, HrpG and HrpX, were found in the proteome of strains LMG 19144, LMG 19145<sup>PT</sup> and CFBP 6773. As a functional Hrp system is thus absent in strains LMG 19145<sup>PT</sup>, CFBP 6773 and LMG 19144, the absence of T3E in these strains may be evident. On the other hand, the strains CFBP 6762 and LMG 19146, which encode a functional Hrp system, have only four effectors: AvrBs2, XopF1, XopF2 and XopR (Fig. 3). We can hypothesize that the small T3SS and T3E repertoire in both groups of *X. arboricola* pv. *fragariae* may not elicit an HR in strawberry plants.

The screened proteins, which were present in both species, harbored interspecies amino acid variations between 1.15 and 39% (Additional file 1: Table S3). The variability of flagellar T3SS proteins, LPS and EPS synthesis protein was ranging between 1.15 and 11%, suggesting purification selection for highly interspecies conserved proteins and genetic drift for variability similar to the interspecies AAI values, thus suggesting a conservation of functional elements in both species. On the other hand, interspecies variations of most of the T3SS cluster with higher variability was observed, and this T3SS cluster, which was reported as conserved in *X. arboricola* strains (Cesbron et al. 2015), could have been independently acquired by horizontal transfer in both species. The four orthologs of T3E protein sequences in both species also greatly varied (between 13.20 and 32.70%), thus suggesting an independent acquisition of effectors. Most of the elements from the T3SS and T3E could therefore result from horizontal transfer as already reported for Gram-negative bacterial pathogens (Brown and Finlay 2011; Puhar and Sansonetti 2014). Similarly, most of the T4SS proteins were variable

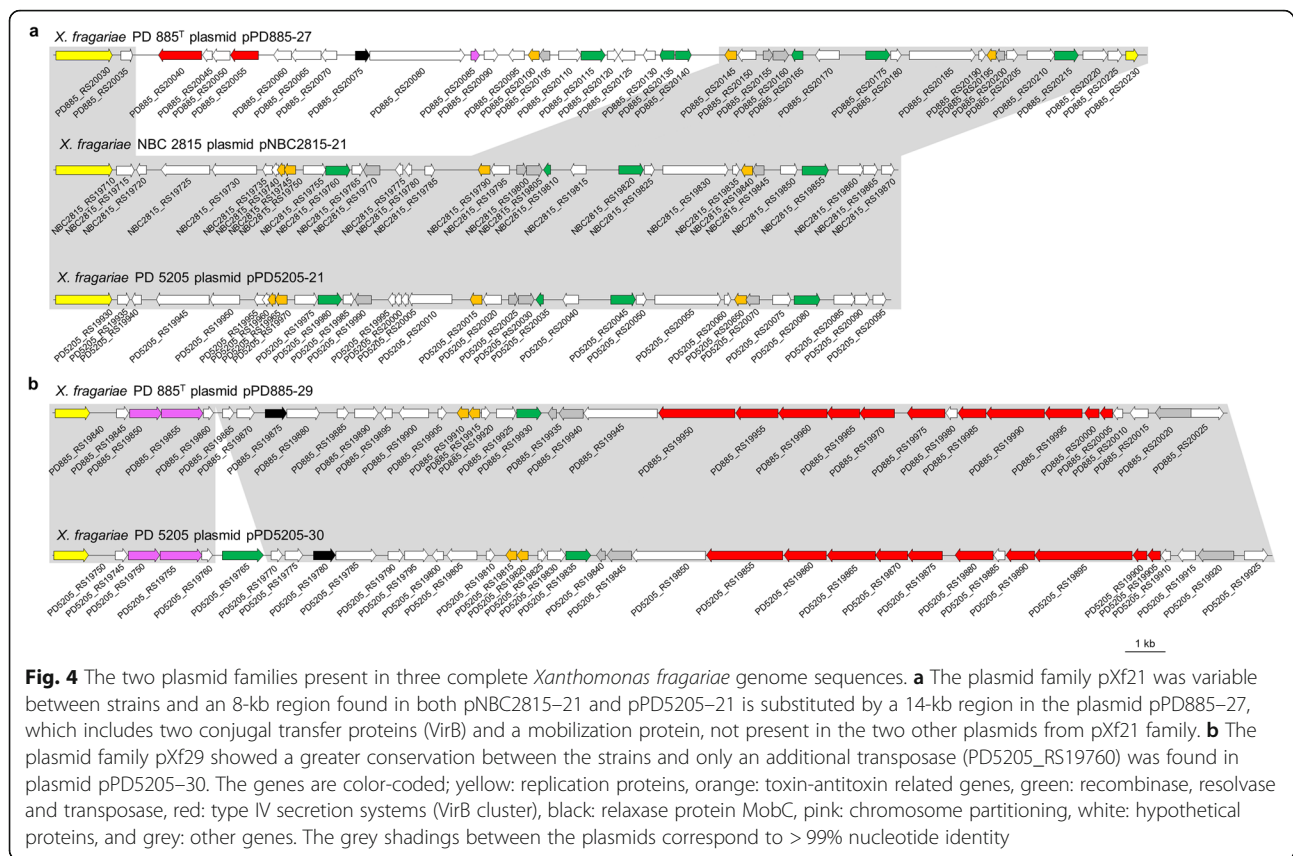
between 19.2 and 39% between species, suggesting similar acquisition patterns.

### Plasmid diversity

The five complete *X. fragariae* genomes (Henry and Leveau 2016; Gétaz et al. 2017b) contain either one or two plasmids that we can appoint to two distinct plasmid families (Fig. 4). The plasmids from strains FaP21 and FaP29 were identical to those from PD 5205 belonging to the same (sub)group (Additional file 2: Fig. S3).

The plasmid pXf21 family is represented by two different variants. The plasmid from strain PD 885<sup>T</sup> contains a 14-kb insert that encodes two T4SS proteins (VirB5 and VirB6), but this is not a complete T4SS. Additionally, the insert includes a relaxase from the MobC superfamily and more specifically in the MobC1 (Gammaproteobacteria) or MoBC\_CloDF13 family (Garcillán-Barcia et al. 2009), as well as a third RelE/ParE toxin/antitoxin element. The 14-kb region may have been inserted into the plasmid due to the concerted action of two transposable elements (PD885\_RS20135 and PD885\_RS20140) located at the border of the region. On the other hand, plasmids pNBC2815–21 and pPD5205–21 have a common 8-kb region instead. They include a RelE/Stb stabilization toxin family and a RelB/DinJ toxin/antitoxin element, in addition to two similar RelE/ParE toxin/antitoxin found as well in pPD885–27. Based on an incomplete T4SS in pPD885–27 and the lack of a conjugation system in the other plasmids from pXf21 family, this plasmid family may be non-transmissible. However, plasmid pPD885–27 is the only plasmid of the family harboring a relaxase and could therefore be mobilizable in the presence of an *oriT* (Smillie et al. 2010). The 14-kb insert was found to be common and unique to all strains from Xf-CGr-IA (Fig. 2), whereas the 8-kb region was found in the other groups.

The plasmid family pXf29 has a more conserved structure. Only one additional transposase (PD5205\_RS19765) was found in plasmid pPD5205–30 from *X. fragariae* PD5205. This protein is an IS3 family transposase found several times with high sequence identity in *X. fragariae* chromosomes. All plasmids from the pXf29 family contain a relaxase from the same MobC1 family (Garcillán-Barcia et al. 2009) as the one found in pPD885–27 and two toxin/antitoxin elements RelE/ParE and RelB/DinJ found as well in the pXf21 plasmids harboring the 8-kb region. Additionally, they all contain a complete T4SS. The pXf29



plasmid family could therefore be effectively mobilizable and conjugative (Garcillán-Barcia et al. 2009; Garcillán-Barcia et al. 2011).

The origin of replication of both plasmids was used to perform an in silico screening for the presence of the plasmid families within all genomes. Overall, a variation of plasmid number among strains screened in this study was observed as strains harbored either none, one or two plasmids (Table 1). Strains from subgroups *Xf*-CGr-IA and -IC had both plasmid families, whereas *Xf*-CGr-IB and -II strains usually had only the plasmid pXf21 family (Additional file 1: Table S4).

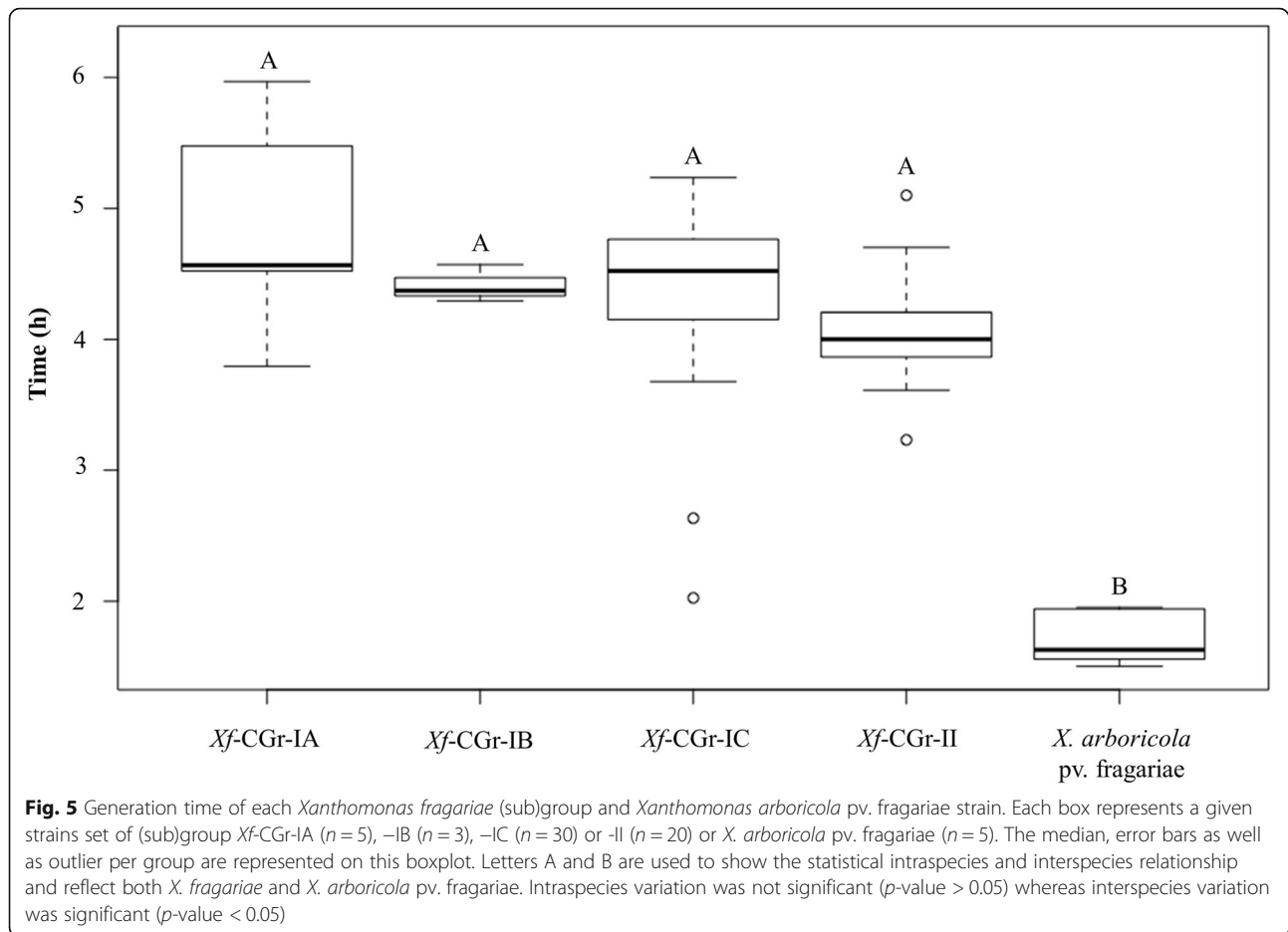
#### Growth behavior of *X. fragariae* strains

To investigate if the above described differences in the genome content and arrangement of the different *X. fragariae* (sub) groups would influence growth behavior, the strains were grown under the same culture conditions in liquid Willbrinks medium, of which sucrose is the carbon source. Average generation times for *X. fragariae* strains were ranging between 2.02 and 5.96 h (Table 1 and Fig. 5). No significant differences in generation time were observed between the *X. fragariae* (sub)-groups ( $p$ -value > 0.05). On the other hand, *X. arboricola* pv. *fragariae* strains grew significantly faster than *X. fragariae* strains with generation times between

1.5 and 1.95 h ( $p$ -value < 0.05) in the same growth medium.

#### Bacterial virulence to strawberry plants

To test whether there is a link between bacterial genotype and its phenotype of virulence on strawberry plants, plants were inoculated with a representative set of *X. fragariae* strains, covering all (sub)groups. *X. arboricola* pv. *fragariae* was also included in these assays. All strawberry plants inoculated with *X. fragariae* strains showed symptoms after 8 to 14 days post inoculation (dpi), dependent on the strain used (Fig. 6) indicating that the strains representing each of the *X. fragariae* (sub)groups were all pathogenic on strawberry. Symptoms caused by *X. fragariae* were not uniform at the plant level, as only a limited number of leaves with variable symptom intensities were observed. Approximately one third to half of the leaves per plant showed symptoms. Starting from the first appearance of the symptoms, plants were frequently evaluated for their symptom evolution. At each evaluation day, the symptom intensity corresponding to the leaf with the most advanced symptom reaction per plant was recorded (Fig. 6 and Fig. 7). Leaves, which were not yet present at the time of the inoculation, did not show any symptoms. A slight variation of intensity was observed between (sub)groups, but this could not be



statistically evaluated. On the other hand, in this study it was not possible to provoke symptoms by *X. arboricola* pv. *fragariae* strains on the Elsanta strawberry plants inoculated during the 30 days of incubation. The group of inoculated strains included strains with and without intact Hrp system and both types of strains resulted in absence of symptoms.

## Discussion

Our study provides a thorough comparative genomics analysis of the sequence data from 58 *X. fragariae* genomes obtained in previous studies (Vandroemme et al. 2013a; Henry and Leveau 2016; Gétaz et al. 2017b; Gétaz et al. 2018b). Based on the whole genome-analysis, we could confirm that all the tested *X. fragariae* were closely related and therefore belonged to the same species. The overall structure of the genomes was highly similar in all *X. fragariae* strains but some slight differences in genome organization were observed. The main difference resides especially in the variable number of plasmids in *X. fragariae*, a feature mainly revealed by long-read sequencing. The plasmid diversity analysis indeed revealed a differential plasmid presence in *X.*

*fragariae* strains that reflects the *X. fragariae* population structure previously reported using the same bacterial strains (Gétaz et al. 2018b). This population structure study suggested that both *Xf*-CGr-I and -II groups were separated before the description of the *X. fragariae* type strain in 1960 (Kennedy and King 1960). In this perception, strains from *Xf*-CGr-IA were considered as more ancestral due to their CRISPR spacer composition (Gétaz et al. 2018b). From this population structure results and following the principle of parsimony with the minimal evolutionary changes, we can hypothesize that all strains from *Xf*-CGr-I could have had two plasmids while the strains from *Xf*-CGr-IB subsequently lost the pXf29 plasmid. On the other hand, strains from *Xf*-CGr-II only harbored one plasmid, suggesting either the acquisition of pXf29 in *Xf*-CGr-I or its loss in *Xf*-CGr-II, after the separation of both groups.

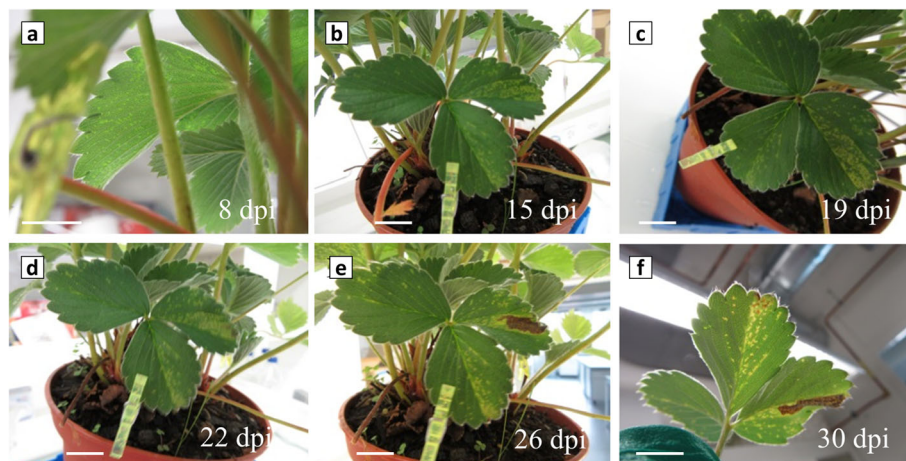
The comparative genomics approach focused mainly on the identification of known *Xanthomonas* virulence-related factors. The results revealed that the virulence-related gene repertoire was, with minor exceptions, identical among all *X. fragariae* genomes. During the thorough in silico screening of virulence-related genes,

Strain	Species	<i>X. fragariae</i> (sub)group	Geographic origin	Days post inoculation (dpi)																																		
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30					
PD 885 <sup>T</sup>	<i>X. fragariae</i>	Xf-CGr-IA	US	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
ICMP 661	<i>X. fragariae</i>	Xf-CGr-IA	US	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PD 4314	<i>X. fragariae</i>	Xf-CGr-IA	NL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XF 61	<i>X. fragariae</i>	Xf-CGr-IB	NL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PD 2905	<i>X. fragariae</i>	Xf-CGr-IB	NL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
JvD-0051	<i>X. fragariae</i>	Xf-CGr-IB	BE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CFBP 4784	<i>X. fragariae</i>	Xf-CGr-IC	PT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PD 3170	<i>X. fragariae</i>	Xf-CGr-IC	NL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LMG 25863	<i>X. fragariae</i>	Xf-CGr-IC	BE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CFBP 5257	<i>X. fragariae</i>	Xf-CGr-IC	FR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ICMP 13779	<i>X. fragariae</i>	Xf-CGr-IC	IT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ICMP 20575	<i>X. fragariae</i>	Xf-CGr-IC	US	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NCPPB 1822	<i>X. fragariae</i>	Xf-CGr-II	US	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CFBP 1558	<i>X. fragariae</i>	Xf-CGr-II	FR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NCPPB 3743	<i>X. fragariae</i>	Xf-CGr-II	BR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NBC 2815	<i>X. fragariae</i>	Xf-CGr-II	NL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
F1-80	<i>X. fragariae</i>	Xf-CGr-II	CH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ICMP 6648	<i>X. fragariae</i>	Xf-CGr-II	AU	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LMG 19145 <sup>PT</sup>	<i>X. arboricola</i> pv. <i>fragariae</i>	NA	IT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LMG 19146	<i>X. arboricola</i> pv. <i>fragariae</i>	NA	FR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NEG control	NA	NA	NA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

**Fig. 6** Outcome of the inoculation test of strawberry plant *Fragaria × ananassa* variety Elsanta. The inoculation tests were carried out with 18 *Xanthomonas fragariae* strains belonging to all *X. fragariae* (sub)groups. Additionally, inoculations with two *X. arboricola* pv. *fragariae* strains and a buffer-only control were carried out. The evolution of the symptoms was color-coded (red (0) = no symptoms; light green (1) = low number of spots, visible through the leaf with a light; middle dark green (2) = extensive spots visible on both sides of the leaf; dark green (3) = extensive spots coloring into yellow, tending to necrosis; brown (4) = necrotic regions on leaves)

GumE, GumK, XopAE and XopC were found to have a larger number of non-synonymous SNPs between Xf-CGr-I and Xf-CGr-II. The higher number of non-synonymous SNPs in these proteins may be indicative of a positive selection in these virulence-related genes. Our analysis also showed the absence of the TAL effectors in

*X. fragariae*, a result that was already reported before (Vandroemme et al. 2013a), but this conclusion was based on a genome that was sequenced with Illumina HiSeq technology yielding in short reads. However, due to the repetitive sequences of the TAL effectors, the absence might have been due to assembly issues, based on



**Fig. 7** Evolution of symptoms on strawberry leaves inoculated with the *Xanthomonas fragariae* strain JvD-0051. At 8 dpi (a), single spots are visible on the abaxial side of the leaf (category = 1). From 15 dpi (b) to 19 dpi (c), spots are visible on both sides of the leaf and aggregate (category = 2). At 22 dpi (d), spots turn to yellow color (category = 3), and then turn to extensive necrosis (category = 4) between 26 dpi (e) and 30 dpi (f), the last time point of the experiment. The white line represents 1 cm on the photography

the length of the used reads. Here, we confirm the absence of TAL effectors from the *X. fragariae* genomes based on the five complete genomes (Henry and Leveau 2016; Gétaz et al. 2017b). These structural variations, gene content and gene sequence differences were hypothesized to influence the ability of a strain to interact with its host. Therefore, plant inoculations using the susceptible strawberry cultivar Elsanta (Kastelein et al. 2014) were performed. These virulence tests revealed a slight variation of intensity between *X. fragariae* (sub)groups, but this could not be statistically evaluated. For this, additional inoculation tests, including more *X. fragariae* strains from all (sub)groups and more plant replicates per treatment would be required, eventually also considering the use of other strawberry cultivars with other sensitivity to the disease. Only then it can be determined whether the described genomic differences between (sub)groups are the reason for the influence on the symptom intensity on strawberry.

The cluster of phage genes and the different clusters with proteins annotated as phage remnant found in all the genomes of *X. fragariae* suggest a higher phage pressure on this species than on *X. arboricola* pv. *fragariae*. The phage pressure was already illustrated by the presence of CRISPR spacers targeted against different phage-related sequences between *X. fragariae* (sub) groups (Gétaz et al. 2018b). Recently, the pressure was also demonstrated by the isolation of the first phage that could infect seven of eight *X. fragariae* strains tested and none of the 14 other *Xanthomonas* species tested (Miller et al. 2020). More generally these phages also contribute to the diversification of the bacterial genome architecture, for instance by horizontal gene transfer and most of the current evidence for the involvement of phages in shaping bacterial genomes, bacterial fitness, and host-pathogen interactions deals with events at this lowest taxonomy level (Brüssow et al. 2004).

The comparative genomics analysis of *X. fragariae* genomes was supplemented by the analysis of five *X. arboricola* pv. *fragariae* strains (Gétaz et al. 2018a). For this species, pathogenicity upon artificial inoculation on strawberry plants was not always reproducible in glass-houses or in various laboratory experiments (Scortichini and Rossi 2003; Vandroemme et al. 2013b; Merda et al. 2016). The inoculation test performed in this study showed that, in contrast to all *X. fragariae* (sub)groups, symptoms were not obtained after inoculation of *X. arboricola* pv. *fragariae*. The lack of symptoms with *X. arboricola* pv. *fragariae* could be attributed to the presence of a smaller virulence gene repertoire in *X. arboricola* pv. *fragariae*, where especially the lack of T3SS and T3E in their genomes was hypothesized to influence its pathogenicity. However, the lack of most Hrp T3SS

proteins in LMG 19145<sup>PT</sup>, CFBP 6773 and LMG 19144 strains could not explain its non-pathogenic behavior since the strain LMG 19146 harboring the full Hrp cluster was not pathogenic either. A detailed study of the pathogenicity of this pathovar and its host range needs thus to be done, as strain LMG 19145<sup>PT</sup> was recently confirmed as causing symptoms on strawberry plants from cultivars Candonga, Sabrina, and Murano (Ferrante and Scortichini 2018) suggesting that another virulence factor may cause symptoms on some strawberry cultivars. In this study, the same strain, inoculated on the cultivar Elsanta, did not produce any symptoms. The lack of constancy in symptom reproduction on the various cultivars (Scortichini and Rossi 2003; Vandroemme et al. 2013b; Merda et al. 2016) may thus reflect that *X. arboricola* pv. *fragariae* could have a reduced host range possibly limited to certain strawberry cultivars only.

In previous studies on *X. arboricola*, strains from *X. arboricola* pv. *fragariae* were shown to be polyphyletic (Vandroemme et al. 2013b). Based on the genome sequences, we identified at least two groups with different gene repertoires in the five strains of this pathovar included in this study. Therefore, it is important that, when strains from *X. fragariae* pv. *fragariae* are tested, both groups are included in order to identify whether all strains of this pathovar are pathogenic on strawberry. Here, *X. arboricola* pv. *fragariae* was also shown to grow faster than *X. fragariae* in the same liquid medium under the tested conditions. We hypothesize that this difference in growth behavior may yield a positive detection of *X. arboricola* pv. *fragariae* in the process of isolation of the pathogen from symptomatic plants having both *X. arboricola* pv. *fragariae* and *X. fragariae* on the diagnosed plant material. Indeed, it was already reported that *X. arboricola* pv. *fragariae* and *X. fragariae* could be co-isolated from symptomatic strawberry leaves (Scortichini and Rossi 2003; Vandroemme et al. 2013b). In order to avoid enrichment and biased growth between both bacterial species, the recently designed *X. fragariae* loop-mediated isothermal amplification (LAMP) could be used to directly detect plant tissue (Gétaz et al. 2017a). For detection of *X. arboricola* pv. *fragariae*, a new assay would need to be designed using the available genome data.

Finally, the presence of a full T6SS in all *X. fragariae* (sub)groups is also reported in the present work. This T6SS cluster was recently described as having a particular genomic architecture in which structural genes are split in two clusters located around ~300 kb from each other while additional genes are present (Bayer-Santos et al. 2019). The relatively high number of putative T6SS effectors identified in the *X. fragariae* genomes (Bayer-Santos et al. 2019; Bosis 2019) could also point to an important ecological role in the life of this bacterium. These observations together with the wide distribution

of T6SS clusters in plant-associated bacteria could suggest that this system is crucial for optimal fitness during plant colonization, as the main role of T6SS in phyto-bacteria would be interbacterial competition rather than host manipulation (Bernal et al. 2018) as also recently demonstrated in the rice pathogen *X. oryzae* pv. *oryzicola* (Zhu et al. 2020). Since it was already reported that *X. fragariae* and *X. arboricola* pv. *fragariae* were co-isolated (Scortichini and Rossi 2003; Vandroemme et al. 2013b), the T6SS found only in *X. fragariae* could be an advantage as soon as both bacterial pathogens compete with each other in the plant.

## Conclusions

Overall, the comparative genomics approach on 58 *X. fragariae* strains and five *X. arboricola* pv. *fragariae* strains revealed both intraspecies and interspecies genomic variations. Within *X. fragariae*, this included large-scale genetic rearrangements at the full-genome level, the presence of none, one or two plasmids in individual strains, but also a conserved virulence-related gene repertoire. Although all tested *X. fragariae* strains were pathogenic to strawberry, none of the tested *X. arboricola* pv. *fragariae* provoked symptoms. The lack of many T3SS and T3E proteins in the genomes of the five *X. arboricola* pv. *fragariae* strains suggests that this pathovar may not be able to cause symptoms as the important genes required for virulence are missing.

The co-isolation of *X. fragariae* and *X. arboricola* pv. *fragariae* is complicating diagnostics from symptomatic plants, and wrong conclusions may be drawn when picking a single colony from plate. Tools for diagnostics directly from infected plant material like LAMP are already available for the specific detection of *X. fragariae* (Wang and Turechek 2016; Gétaz et al. 2017a). A similar test would need to be designed as well for detection of *X. arboricola* pv. *fragariae*. However, the polyphylous nature of the pathogen may prohibit finding *X. arboricola* pv. *fragariae*-specific genes easily within the *X. arboricola* clade. For optimal detection, the combination of both assays should be used in diagnostic settings in order to avoid isolation biases. This may then allow the clarification of the reported differences in the description of *X. arboricola* pv. *fragariae* pathogenicity to strawberries.

## Methods

### Bacterial strains

A total of 58 genomes of *X. fragariae* strains including five complete and 53 partial genome sequences obtained in a population structure analysis (Gétaz et al. 2018b) were used to assess their intraspecies variations (Table 1). Automatic genome annotation of the *X. fragariae* genomes was performed using the GenDB platform v.2.4 (Meyer et al. 2003) except for three of them that were directly obtained

from GenBank (Table 1). Additionally, five *X. arboricola* pv. *fragariae* strains (Gétaz et al. 2018a) were included in the analysis in order to assess variations within and between *X. arboricola* pv. *fragariae* and *X. fragariae* strains (Table 1). All genomes were added to EDGAR v.2.2 (Blom et al. 2016) for whole-genome comparisons.

### Virulence-related proteins for *Xanthomonas* spp.

A list of virulence-related proteins of xanthomonads was compiled from a list of proteins thought to play a role in bacterial virulence in the first *X. fragariae* draft genome LMG 25863 (Vandroemme et al. 2013a), including T3SS, T4SS, T6SS (Bayer-Santos et al. 2019), lipopolysaccharides (LPS) and extracellular polysaccharides (EPS) synthesis proteins and from an online list of T3SS effectors found in the genus *Xanthomonas* (Koebnik 2016). Overlaps between both lists were removed, resulting in a final list of 163 proteins (Additional file 1: Table S2) that was used as a reference in this study.

### Comparative genomics

Annotated complete genomes were aligned with MAUVE v.2.3.1 (Darling et al. 2004). Using tBLASTn v.2.3.0+ search (Camacho et al. 2009), all genomes were screened for the presence and sequence identity of virulence-related proteins (Additional file 1: Table S2). Sequences of present proteins were aligned using Clustal W (Thompson et al. 1994) on MEGA v.6.06 (Tamura et al. 2013) in order to manually screen intraspecies and interspecies amino acid variations, as well as hit length, which may be reduced due to contig edges, to late coding start or early stop codon. The sequence variation of virulence-related genes between strains was then based at amino acid sequence level, which does not consider synonymous mutations (Seo and Kishino 2008).

A non-linear pan genome of *X. fragariae*, containing all CDS of the 58 *X. fragariae* strains was generated using EDGAR v.2.2 (Blom et al. 2016) with the strain PD 885<sup>T</sup> as reference. In order to visualize variations of CDS at the strain level within *X. fragariae*, the generated pan genome was used as reference sequence to compare the 58 *X. fragariae* strains with BLAST ring image generator (BRIG) v.0.95 (Alikhan et al. 2011). The EDGAR platform v.2.2 (Blom et al. 2016) was used to generate specific subsets of CDS in order to compare core genomes of each *X. fragariae* (sub) group and *X. arboricola* pv. *fragariae*. Genes were considered orthologous when a reciprocal best BLAST hits was found between two genes, and when both BLAST hits were based on alignments exceeding 70% sequence identity spanning over at least 70% of the query gene length (Blom et al. 2009). Subsequently, the average amino acid identity (AAI) was computed for all strains from both bacterial species on the



EDGAR platform v.2.2 (Blom et al. 2016), to estimate their intraspecies and interspecies relatedness.

### In silico plasmid screening

The five complete genomes sequenced with PacBio technology harbored either one or two plasmids per genome (Henry and Leveau 2016; Gétaz et al. 2017b). The presence of these two plasmid families in all sequenced *X. fragariae* strains was examined in silico by searching for putative plasmid replication proteins of each plasmid family (GenBank locus tag SMR01212 in pNBC2815–21 and SMR01243 in pPD885–29) in all 53 *X. fragariae* draft genomes with tBLASTn (Camacho et al. 2009). The genome assemblies lacking both replication proteins were screened for full plasmid DNA sequences of both pNBC2815–21 and pPPD885–29 by performing a BLASTn v.2.3.0+ search (Camacho et al. 2009). Due to the large number of contigs in most of the draft genomes, plasmids may be fragmented into distinct contigs. These contigs were aligned on the used plasmid reference sequence and inspected for continuity on MEGA v.6.06 (Tamura et al. 2013). Whole-plasmid sequences of the five available complete genomes (Henry and Leveau 2016; Gétaz et al. 2017b) were screened for structural and gene content variation. The annotated plasmids were manually aligned and gene contents were compared using EDGAR (Blom et al. 2016). Sequence manipulations were completed using several subroutines of the Lasergene package v.12.1 (DNASTAR, Madison, USA).

### Growth experiments

*X. fragariae* and *X. arboricola* pv. *fragariae* strains (Table 1) were routinely grown at 28 °C in Wilbrinks-N medium (Koike 1965) for 48 h while shaking at 220 rpm. Cells were collected by centrifugation at 3000×g and washed twice with Ringer solution (Sigma-Aldrich, Buchs, Switzerland) and resuspended in liquid Wilbrinks-N medium (Koike 1965). Before the start of every growth experiment, the cell concentration was adjusted to  $A_{600}$  of 0.1 (Thermo Genesys 10 spectrophotometer; VWR International). All growth curves were obtained using an automated analysis system BioScreen C (Growth Curve AB Ltd., Helsinki, Finland), with attenuation recorded every 15 min. Bacteria were grown over 4 days, in order to ensure all bacterial strains to reach stationary phase. Each strain was tested in triplicate. Raw data of attenuation was used to calculate the specific growth rate (Maier and Pepper 2015). Two one-way ANOVA were performed using R v.3.4.3 (R Development Core Team 2008) in order to test if a significant difference of growth rate exists between the already described *Xf*-CGr (sub) groups, and also when *X. fragariae* was compared to *X. arboricola* pv. *fragariae*. Due to

multiple testing with same data, *p*-values were adapted with Bonferroni procedure (Hochberg 1988).

### Assessment of bacterial virulence in planta

Representative strains were selected for each *X. fragariae* (sub) group to be tested for plant inoculation. Following the protocol described by Kastelein (Kastelein et al. 2014), bacterial strains were grown overnight in Wilbrinks-N liquid culture medium (Koike 1965), washed and resuspended in Ringer solution (Sigma-Aldrich, Buchs, Switzerland). Strawberry plants were incubated in a plastic bag in order to keep high humidity. A total of 48 plants were inoculated by a spraying method in two 24-plant batches, which were tested in independent consecutive experiments. Each batch included ten *X. fragariae* strains, one *X. arboricola* pv. *fragariae* strain and a negative control using buffer only, all of which were inoculated onto two plants to have duplicates. Two of the *X. fragariae* strains (LMG 25863 and PD 885<sup>T</sup>) were used for both batches for between-batch comparison purposes. After inoculation, plants were kept for 30 days in a climate chamber (Fitotron, Weiss Technik, Leicestershire, United Kingdom). Controlled conditions were set to have a 16 h daylight period at 22 °C and 70% relative humidity and 8 h night period at 17 °C and 80% relative humidity. The appearance of symptoms was evaluated every 3 days. The number of leaves showing symptoms was recorded as well as the aspect of symptoms, in order to know the surface covered by the spots. A scale of 0 to 5 was applied to score the appearance and evolution of the symptoms on leaves as follows: 0 = no reaction, 1 = first few light spots visible on the abaxial side of a leaf, 2 = leaf spots visible on both sides of a leaf, 3 = extensive leaf spots turning to yellow color, 4 = extensive necrotic leaf spots, and 5 = dead leaf.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s42483-020-00061-y>.

**Additional file 1: Table S1.** Result of the pairwise amino acid identity (AAI) analysis represented as a matrix. **Table S2.** List of virulence-related proteins used as input for genome screening. **Table S3.** Amino acid comparison for the proteins found in both *Xanthomonas arboricola* pv. *fragariae* and *Xanthomonas fragariae*. **Table S4.** Plasmid representation of each *Xanthomonas fragariae* (sub)group.

**Additional file 2: Figure S1.** Alignment of three complete sequences of *Xanthomonas fragariae*. **Figure S2.** Alignment of chromosome from five complete sequences of *Xanthomonas fragariae*. **Figure S3.** The two plasmid families present in five complete *Xanthomonas fragariae* genome sequences.

### Abbreviations

AAI: Amino acid identity; CDS: Coding sequence; CRISPR: Clustered regularly interspaced short palindromic repeats; CWDE: Plant cell wall degrading enzyme; EPS: Extracellular polysaccharides; EPPO: European and Mediterranean Plant Protection Organization; HR: Hypersensitive reaction; LAMP: Loop-mediated isothermal amplification; LPS: Lipopolysaccharides;

MLSA: Multilocus sequence analysis; SNP: Single-nucleotide polymorphism; T2SS: Type II secretion system; T3E: Type III secretion system effector; T3SS: Type III secretion system; T4SS: Type IV secretion system; T6SS: Type VI secretion system; TAL: Transcription activator like; VNTR: Variable numbers of tandem repeats; Xf-CGR: *Xanthomonas fragariae* CRISPR group

#### Acknowledgements

The authors thank Marilena Palmisano (ZHAW, Wädenswil, CH) for technical assistance, Dr. Jan van der Wolf (Wageningen University & Research, Wageningen, the Netherlands), Dr. Maria Bergsma-Vlami (The Netherlands Food and Consumer Product Safety Authority) and Dr. Bart Cottyn (Flanders research institute for agriculture, fisheries and food – ILVO, Merelbeke, Belgium) for providing some of the *X. fragariae* strains used in this study. This article is based upon work from COST Action CA16107 EuroXanth, supported by COST (European Cooperation in Science and Technology).

#### Authors' contributions

MG, THMS, JFP conceived and designed the experiments; MG performed the experiments; MG, JB, THMS and JFP analyzed the data, MG, THMS and JFP interpreted the results. All authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

#### Funding

Funding was provided by the European Union Seventh Framework (FP7/2007–2013) under the grant agreement no. 613678 (DROPSA). The EDGAR platform is financially supported by the BMBF grant FKZ 031A533 within the de. NBI network.

#### Availability of data and materials

The genomic dataset supporting the conclusions of this article are deposited in the ENA repository under the study accession PRJEB25730.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare not to have any competing interests.

#### Author details

<sup>1</sup>Environmental Genomics and Systems Biology Research Group, Institute of Natural Resource Sciences, Zurich University of Applied Sciences (ZHAW), CH-8820 Wädenswil, Switzerland. <sup>2</sup>Bioinformatics and Systems Biology, Justus-Liebig-University Giessen, Giessen, Germany.

Received: 14 May 2020 Accepted: 10 July 2020

Published online: 24 July 2020

#### References

- Ali Khan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*. 2011;12:402.
- Amil-Ruiz F, Blanco-Portales R, Muñoz-Blanco J, Caballero JL. The strawberry plant defense mechanism: a molecular review. *Plant Cell Physiol*. 2011;52:1873–903.
- Bayer-Santos E, Ceseti LM, Farah CS, Alvarez-Martinez CE. Distribution, function and regulation of type 6 secretion systems of Xanthomonadales. *Front Microbiol*. 2019;10:1635.
- Bernal P, Llamas MA, Filloux A. Type VI secretion systems in plant-associated bacteria. *Environ Microbiol*. 2018;20:1–15.
- Bestfleisch M, Richter K, Wensing A, Wünsche JN, Hanke M-V, Höfer M, et al. Resistance and systemic dispersal of *Xanthomonas fragariae* in strawberry germplasm (*Fragaria L.*). *Plant Pathol*. 2015;64:71–80.
- Blom J, Albaum SP, Doppmeier D, Pühler A, Vorhölter FJ, Zakrzewski M, et al. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinf*. 2009;10:154.
- Blom J, Kreis J, Spänig S, Juhre T, Bertelli C, Ernst C, et al. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Res*. 2016;44:W22–8.
- Boch J, Bonas U. *Xanthomonas AvrBs3* family-type III effectors: discovery and function. *Annu Rev Phytopathol*. 2010;48:419–36.
- Bosis E. Computational analysis of type VI secretion systems in *Xanthomonadaceae*. In: Koebnik R, Baránek M, editors. *Proceeding of the 3rd Annual Conference of the EuroXanth COST Action*. Lednice: Mendel University, Brno; 2019. p. 77.
- Brown NF, Finlay B. Potential origins and horizontal transfer of type III secretion systems and effectors. *Mob Genet Elem*. 2011;1:118–21.
- Brüssow H, Canchaya C, Hardt W-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev*. 2004;68:560–602.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinf*. 2009;10:421.
- Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, et al. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveils molecular and evolutionary events linked to pathoadaptation. *Front Plant Sci*. 2015;6:1126.
- Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 2004;14:1394–403.
- Déjean G, Blanvillain-Baufumé S, Boulanger A, Darrasse A, Bernonville TD, Girard AL, et al. The xylan utilization system of the plant pathogen *Xanthomonas campestris* pv *campestris* controls epiphytic life and reveals common features with oligotrophic bacteria and animal gut symbionts. *New Phytol*. 2013;198:899–915.
- FAOSTAT. Food and agriculture data. <http://faostat.fao.org>. Accessed 6 May 2020.
- Ferrante P, Scortichini M. *Xanthomonas arboricola* pv. *fragariae*: a confirmation of the pathogenicity of the pathotype strain. *Eur J Plant Pathol*. 2018;150:825–9.
- Galán JE, Collmer A. Type III secretion machines: bacterial devices for protein delivery into host cells. *Science*. 1999;284:1322–8.
- Garcillán-Barcia MP, Alvarado A, de la Cruz F. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol Rev*. 2011;35:936–56.
- Garcillán-Barcia MP, Francia MV, de la Cruz F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev*. 2009;33:657–87.
- Gétaz M, Baeyen S, Blom J, Maes M, Cottyn B, Pothier JF. High-quality draft genome sequences of five *Xanthomonas arboricola* pv. *fragariae* isolates. *Genome Announc*. 2018a;6:e01585–17.
- Gétaz M, Bühlmann A, Schneeberger PHH, Van Malderghem C, Duffy B, Maes M, et al. A diagnostic tool for improved detection of *Xanthomonas fragariae* using a rapid and highly specific LAMP assay designed with comparative genomics. *Plant Pathol*. 2017a;66:1094–102.
- Gétaz M, Krijger M, Rezzonico F, Smits THM, van der Wolf JM, Pothier JF. Genome-based population structure analysis of the strawberry plant pathogen *Xanthomonas fragariae* reveals two distinct groups that evolved independently before its species description. *Microb Genomics*. 2018b;4:e000189.
- Gétaz M, Van der Wolf JM, Blom J, Pothier JF. Complete genome sequences of three isolates of *Xanthomonas fragariae*, the bacterium responsible for angular leaf spots on strawberry plants. *Genome Announc*. 2017b;5:e00642–17.
- Ghosh P. Process of protein transport by the type III secretion system. *Microbiol Mol Biol Rev*. 2004;68:771–95.
- Hartung JS, Gouin CC, Lewers KS, Maas JL, Hokanson S. Identification of sources of resistance to bacterial angular leafspot disease of strawberry. *Acta Horti*. 2003;626:155–9.
- Henry PM, Leveau JHJ. Finished genome sequences of *Xanthomonas fragariae*, the cause of bacterial angular leaf spot of strawberry. *Genome Announc*. 2016;4:e01271–16.
- Hildebrand DC, Schroth MN, Wilhelm S. Systemic invasion of strawberry by *Xanthomonas fragariae* causing vascular collapse. *Phytopathology*. 1967;57:1260–1.
- Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*. 1988;75:800–2.
- Janse JD, Rossi MP, Gorkink RFJ, Derks JHJ, Swings J, Janssens D, et al. Bacterial leaf blight of strawberry (*Fragaria (x) ananassa*) caused by a pathovar of *Xanthomonas arboricola*, not similar to *Xanthomonas fragariae* Kennedy & King. Description of the causal organism as *Xanthomonas arboricola* pv. *fragariae* (pv. Nov., comb. nov.). *Plant Pathol*. 2001;50:653–65.
- Kastelein P, Krijger M, Czajkowski R, van der Zouwen PS, van der Schoor R, Jalink H, et al. Development of *Xanthomonas fragariae* populations and disease progression in strawberry plants after spray-inoculation of leaves. *Plant Pathol*. 2014;63:255–63.

- Kearney B, Staskawicz BJ. Widespread distribution and fitness contribution of *Xanthomonas campestris* avirulence gene *avrBs2*. *Nature*. 1990;346:385–6.
- Kennedy BW, King TH. Angular leafspot, a new disease of strawberry. *Phytopathology*. 1960;50:641–2.
- Kennedy BW, King TH. Angular leaf spot of strawberry caused by *Xanthomonas fragariae* sp nov. *Phytopathology*. 1962;52:873–5.
- Kim D-R, Gang G-H, Jeon C-W, Kang NJ, Lee S-W, Kwak Y-S. Epidemiology and control of strawberry bacterial angular leaf spot disease caused by *Xanthomonas fragariae*. *Plant Pathol J*. 2016;32:290–9.
- Koebnik R. The *Xanthomonas* resource. <http://www.xanthomonas.org/t3e.html>. Accessed 11 July 2016.
- Koike H. The aluminium-cap method for testing sugarcane varieties against leaf scald disease. *Phytopathology*. 1965;55:317–9.
- Lu H, Patil P, Van Sluys MA, White FF, Ryan RP, Dow JM, et al. Acquisition and evolution of plant pathogenesis-associated gene clusters and candidate determinants of tissue-specificity in *Xanthomonas*. *PLoS One*. 2008;3:e3828.
- Maas JL. Angular leaf spot. *Compendium of strawberry diseases*; 1998. p. 16–7.
- Maas JL, Hartung JS, Gouin-Behe C, Hokanson SC. Sources of resistance for two differentially pathogenic strains of *Xanthomonas fragariae* in *Fragaria* genotypes. *HortScience*. 2000;35:128–31.
- Maier RM, Pepper IL. Chapter 3 - bacterial growth. In: Pepper IL, Gerba CP, Gentry TJ, editors. *Environmental Microbiology*. 3rd ed. San Diego: Academic Press; 2015. p. 37–56.
- Merda D, Bonneau S, Guimbaud J-F, Durand K, Brin C, Boureau T, et al. Recombination-prone bacterial strains form a reservoir from which epidemic clones emerge in agroecosystems. *Environ Microbiol Rep*. 2016; 8:572–81.
- Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, et al. GenDB - an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res*. 2003;31:2187–95.
- Miller M, Deulio A, Holland C, Douthitt C, McMahon J, Wiersma-Koch H, et al. Complete genome sequence of *Xanthomonas* phage RiverRider, a novel N4-like bacteriophage that infects the strawberry pathogen *Xanthomonas fragariae*. *Arch Virol*. 2020;165:1481–4.
- OEPP/EPPO. *Xanthomonas fragariae* Kennedy & King. *EPPO Bull*. 1986;16:17–20.
- OEPP/EPPO, et al. *Bull OEPP*. 2006;36:135–44.
- Puhar A, Sansonetti PJ. Type III secretion system. *Curr Biol*. 2014;24:R784–R91.
- R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2008.
- Roach JA, Verma S, Peres NA, Jamieson AR, van de Weg WE, Bink MCAM, et al. *FaRxf1*: a locus conferring resistance to angular leaf spot caused by *Xanthomonas fragariae* in octoploid strawberry. *Theor Appl Genet*. 2016;129:1191–201.
- Scortichini M. A new bacterial disease of strawberry caused by *Xanthomonas*. *Frutticoltura*. 1996;58:51–3.
- Scortichini M, Rossi MP. Genetic diversity of *Xanthomonas arboricola* pv. *fragariae* strains and comparison with some other *X. arboricola* pathovars using repetitive PCR genomic fingerprinting. *J. Phytopathol*. 2003;151: 113–9.
- Seo T-K, Kishino H. Synonymous substitutions substantially improve evolutionary inference from highly diverged proteins. *Syst Biol*. 2008;57:367–77.
- Simpson DW. Resistance to *Botrytis cinerea* in pistillate genotypes of the cultivated strawberry *Fragaria ananassa*. *J Horticult Sci*. 1991;66:719–23.
- Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. Mobility of plasmids. *Microbiol Mol Biol Rev*. 2010;74:434–52.
- Snoeiijers SS, Pérez-García A, Joosten MHJ, De Wit PJGM. The effect of nitrogen on disease development and gene expression in bacterial and fungal plant pathogens. *Eur J Plant Pathol*. 2000;106:493–506.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–9.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22:4673–80.
- Ustun N, Tjou-Tam-Sin NNA, Janse JD. First report on bacterial leaf blight of strawberry caused by *Xanthomonas arboricola* pv. *fragariae* Janse et al. in Turkey. *J Plant Pathol*. 2007;89:109–12.
- Vandroemme J, Cottyn B, Baeyen S, De Vos P, Maes M. Draft genome sequence of *Xanthomonas fragariae* reveals reductive evolution and distinct virulence-related gene content. *BMC Genomics*. 2013a;14:829.
- Vandroemme J, Cottyn B, Pothier JF, Pflüger V, Duffy B, Maes M. *Xanthomonas arboricola* pv. *fragariae*: what's in a name? *Plant Pathol*. 2013b;62:1123–31.
- Wang H, Turechek WW. A loop-mediated isothermal amplification assay and sample preparation procedure for sensitive detection of *Xanthomonas fragariae* in strawberry. *PLoS One*. 2016;11:e0147122.
- Zhu P-C, Li Y-M, Yang X, Zou H-F, Zhu X-L, Niu X-N, et al. Type VI secretion system is not required for virulence on rice but for inter-bacterial competition in *Xanthomonas oryzae* pv. *oryzicola*. *Res Microbiol*. 2020;171:64–73.
- Zimmermann C, Hinrichs-Berger J, Moltmann E, Buchenauer H. Nested PCR (polymerase chain reaction) for detection of *Xanthomonas fragariae* in symptomless strawberry plants. *J Plant Dis Prot*. 2004;111:39–51.
- Zoued A, Brunet YR, Durand E, Aschtgen M-S, Logger L, Douzi B, et al. Architecture and assembly of the type VI secretion system. *Biochim Biophys Acta*. 2014;1843:1664–73.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

