

RESEARCH

Open Access



High genomic plasticity and unique features of *Xanthomonas translucens* pv. *graminis* revealed through comparative analysis of complete genome sequences

Florian Goettelmann¹, Ralf Koebnik², Veronica Roman-Reyna³, Bruno Studer¹ and Roland Kölliker^{1*}

Abstract

Background *Xanthomonas translucens* pv. *graminis* (*Xtg*) is a major bacterial pathogen of economically important forage grasses, causing severe yield losses. So far, genomic resources for this pathovar consisted mostly of draft genome sequences, and only one complete genome sequence was available, preventing comprehensive comparative genomic analyses. Such comparative analyses are essential in understanding the mechanisms involved in the virulence of pathogens and to identify virulence factors involved in pathogenicity.

Results In this study, we produced high-quality, complete genome sequences of four strains of *Xtg*, complementing the recently obtained complete genome sequence of the *Xtg* pathotype strain. These genomic resources allowed for a comprehensive comparative analysis, which revealed a high genomic plasticity with many chromosomal rearrangements, although the strains were highly related. A high number of transposases were exclusively found in *Xtg* and corresponded to 413 to 457 insertion/excision transposable elements per strain. These mobile genetic elements are likely to be involved in the observed genomic plasticity and may play an important role in the adaptation of *Xtg*. The pathovar was found to lack a type IV secretion system, and it possessed the smallest set of type III effectors in the species. However, three XopE and XopX family effectors were found, while in the other pathovars of the species two or less were present. Additional genes that were specific to the pathovar were identified, including a unique set of minor pilins of the type IV pilus, 17 TonB-dependent receptors (TBDRs), and 11 plant cell wall degradative enzymes.

Conclusion These results suggest a high adaptability of *Xtg*, conferred by the abundance of mobile genetic elements, which could play a crucial role in pathogen adaptation. The large amount of such elements in *Xtg* compared to other pathovars of the species could, at least partially, explain its high virulence and broad host range. Conserved features that were specific to *Xtg* were identified, and further investigation will help to determine genes that are essential to pathogenicity and host adaptation of *Xtg*.

*Correspondence:
Roland Kölliker
roland.koelliker@usys.ethz.ch

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



© The Author(s) 2023. **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/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords *Xanthomonas translucens* pv. *graminis*, Complete genomes, Phylogeny, Host adaptation, Comparative genomics, Virulence factors

Introduction

Bacterial wilt of forage grasses, caused by *Xanthomonas translucens* pv. *graminis* (*Xtg*), is one of the main diseases of forage grasses in temperate grasslands [1]. *Xtg* is a member of the recently-defined clade Xt-III of *X. translucens*, which also contains four other forage grass-infecting pathovars (pv. *arrhenatheri*, *phlei*, *phleipratensis* and *poae*) and is distinguished from clades Xt-I and Xt-II, which contain mainly cereal-infecting pathovars such as pv. *translucens* and *cerealis*, respectively [2, 3]. While most *X. translucens* pathovars are generally restricted to a small host range, *Xtg* is able to infect many grass genera, including *Agrostis*, *Alopecurus*, *Dactylis*, *Deschampsia*, *Festuca*, *Lolium*, *Phalaris*, *Phleum*, *Poa*, and *Trisetum* [1, 4]. It is also the most widespread, being found in most of Europe, America, and New Zealand [4–6].

In contrast to cereal crops such as wheat, where *X. translucens* is considered a seedborne pathogen, no seed or soil transmission of *Xtg* has been observed on forage grasses so far [7, 8]. Here, the bacteria are mainly spread by contaminated mowing tools and penetrate via plant wounds. While some *X. translucens* pathovars such as pv. *undulosa* are restricted to the non-vascular tissue, all forage grass infecting *X. translucens* pathovars spread in the xylem, resulting in a disruption of the flow of water and nutrients [7, 9]. The disease affects mainly Italian ryegrass (*Lolium multiflorum* Lam.), which is cut several times during the growing season for harvesting. After penetration, the bacteria spread in the xylem, causing a disruption of water and nutrient absorption. The symptoms manifest as curling and wilting of the leaves and tillers, and eventually the death of the plant [9, 10]. Infection by the pathogen can lead to serious yield losses, estimated to reach up to 40% under unfavourable conditions [11]. To control the disease in an efficient manner, it is essential to understand the mechanisms involved in the pathogenicity of *Xtg*.

In plant pathogenic bacteria, successful infection of the plant host generally involves virulence factors that allow the bacteria to adhere to the plant surface, invade the host tissue, acquire nutrients, and suppress plant defence mechanisms. In xanthomonads, some of the most important virulence factors are effectors, in particular type III effectors (T3E), that are secreted into the plant host cells by the type III secretion system (T3SS). These T3E play a major role in pathogenicity by specifically targeting different pathways in the host cell. A total of 53 classes of T3E have been defined, generally referred to as “*Xanthomonas* outer proteins” (Xop) [12]. Other T3E can cause a hypersensitive response in their

plant host and are referred to by this characteristic, like AvrBs1 and AvrBs2. Additionally, transcription activator-like effectors (TALEs) are specific T3E that consist of repetitive sequences that allow them to bind to specific nucleotide sequences in the host genome and manipulate the expression of specific genes [13, 14]. However, while TALEs are characteristic to the *Xanthomonas* genus, no TALEs have ever been identified in *Xtg* [3, 15].

Other secretion systems also play an important role in the infection of *Xanthomonas* spp., including the type II, IV and VI secretion systems (T2SS, T4SS and T6SS, respectively). The T2SS is mostly involved in the secretion of plant cell wall degradative enzymes in the apoplast, allowing the bacteria to invade the plant host, as well as acquire the resulting degradation products as a source of nutrients [16]. Two types of T2SS can be found in *Xanthomonas* spp., the *xps* T2SS, which is conserved across the genus, and the *xcs* T2SS, which is only found in some species [17]. To date, no *xcs* T2SS was found in *X. translucens* [3].

The T4SS and T6SS have a similar function and are involved in defence against microbial predators and competition with other microorganisms [18, 19]. Three subtypes of the T6SS have been found in *Xanthomonas* spp., subtypes 1, 3 and 4, and subtype 3 is further divided into subgroups 3*, 3** and 3***. Previously, no T4SS has been identified in *Xtg*, and T6SS genes were identified only in strains Xtg2, Xtg9, Xtg10 and NCPPB 3709 [3, 20].

To reliably identify virulence factors in plant pathogenic bacteria, complete and high-quality genome sequences are essential, as they provide an accurate representation of the genes present in each strain. Such genome sequences enable comparative genomics, which can reveal conserved or divergent genes of a pathovar and help define genes that are crucial to pathogenicity. Previously, draft genome sequences of *Xtg* strains CFBP 2053, ICMP 6431, NCPPB 3709, Xtg2, Xtg9 and Xtg29 allowed a first comparative analysis, which revealed specific features of the pathovar, such as the lack of a flagellum and a distinct type IV pilus [20]. Hypothetical proteins specific to *Xtg* were also identified, including genes with predicted functions in nutrient acquisition, regulatory mechanisms, virulence, adhesion and motility. Effector proteins that were specific to *Xtg* included a YopT-like cysteine protease and a XopJ class effector protein, which were tested for their role in virulence. However, single and double knock-out mutants for these genes showed no significant difference in virulence with the wild type [20].

At that time, genomic resources for *X. translucens* pathovars consisted of highly fragmented and incomplete draft genome sequences based on short read sequencing, which might have prevented the identification of important genes. Furthermore, these genome assemblies fail to assemble complex repetitive sequences such as TALEs, which may have hindered their identification. Long read sequencing technologies now allow to assemble complete bacterial genome sequences, allowing for a comprehensive overview of the gene content within these genomes, including TALEs [21]. Recently, high-quality, complete genome sequences were produced for all pathotype strains of *X. translucens*, which allowed for a more comprehensive comparative analysis [3]. This revealed that the *Xtg* pathotype strain LMG 726 lacked many virulence features, such as a T4SS and T6SS, and had the smallest set of T3E. However, the inclusion of only one strain of *Xtg* did not allow for an in-depth comparative analysis. In this study, we aimed at producing high-quality complete genome sequences for four additional strains of *Xtg*, perform a comprehensive comparative genomic analysis of the pathovar, and identify unique features that could play a role in its virulence and host specificity.

Materials and methods

Bacterial strains, growth conditions and DNA extraction

The *Xtg* strains NCPPB 3709, *Xtg*2, *Xtg*9 and *Xtg*29 were selected to be sequenced in this study (Table 1). These strains were grown at 28 °C on YDC agar medium (2% dextrose, 1% yeast extract, 2% CaCO₃, 1.5% agar) for 48 h. Bacteria were then dissolved in 10 mL washing buffer (50 mM TRIS-HCl pH 8.0, 50 mM EDTA pH 8.0, 150 mM NaCl). Genomic DNA was then extracted with the NucleoSpin® Microbial DNA kit (Macherey Nagel, Duren, Germany), according to the manufacturer's recommendations.

Sequencing, genome assembly and annotation

Library preparation and DNA sequencing was performed at the Functional Genomics Center Zurich. Libraries were prepared and multiplexed with the PacBio SMRTbell® Express Template Prep Kit 2.0 (PacBio, Menlo Park, CA, USA) according to the published protocol [22]. After shearing and size selection to 7–10 kb, 1 µL of the non-size-selected library was spiked in the final solution, for a final volume of 10 µL, to allow for sequencing of potential plasmids. Libraries were then sequenced on the PacBio Sequel II platform.

Genomic sequences were *de novo* assembled with flye 2.7 using the "--plasmids --iterations 2" parameters [23]. Since the assembly of strain NCPPB 3709 did not result in a complete circular chromosome, we used canu with default parameters for this strain [24]. The assembled circular chromosomes were rotated to start with the

dnaA gene sequence for comparability. Genomes were functionally annotated with Prokka 1.14.6 with default parameters [25].

Phylogeny and comparative analysis

To determine phylogeny within clade Xt-III, we used all available complete genome sequences of the clade, as well as the genome sequences of pv. *translucens* strain DSM 18974 and pv. *cerealis* strain CFBP 2541 as representatives of clades Xt-I and Xt-II, respectively (Table 1). Average nucleotide identity (ANI) and alignment fraction (AF) were calculated using the anib subcommand of pyani 0.2.12 with default parameters [26], and a phylogeny dendrogram was constructed using Ward's hierarchical clustering method in R version 4.2.0 [27]. To identify genomic rearrangements and conserved genomic regions within *Xtg*, the genomic structure of the five *Xtg* genome sequences were compared with Mauve v20150226 [28].

For a comparative analysis of the gene content of *Xtg* compared to other *X. translucens* strains, all currently available complete genome sequences of the *X. translucens* species were retrieved and annotated with Prokka (Table 1, Table S1). The gene content within the *X. translucens* species was compared with Roary 3.13.0 using default parameters [29]. This allowed to predict singletons, genes that are exclusive to *Xtg*, by selecting genes that were only found in *Xtg* strains in our dataset. The pangenome and the predicted singleton contents of *Xtg* strains were visualized in Venn diagrams with jvenn [30]. Clusters of orthologous groups (COG)-based functional characterization of the gene content of *Xtg* was performed using eggNOG-mapper v2 [31, 32]. As the COG analysis revealed a high number of transposable elements in *Xtg*, these were further characterized using the mobileOG-db tool in the Proksee online software, with default parameters [33, 34]. Additionally, putative horizontal transfer events were identified using the Alien Hunter tool in Proksee [35]. The presence of signal peptides or non-classical secretion were predicted using SignalP 6.0 and SecretomeP 2.0, respectively [36, 37].

Minor pilins identified in the singleton analysis were investigated with InterProScan by identifying the characteristic pilin N-terminal transmembrane domain, as well as with the PilFind online tool [38, 39]. To compare the two filamentous hemagglutinin genes identified with similar previously-identified genes, the nucleotide sequence of these genes in the *Xtg*29 draft genome sequence in which they were identified was extracted [20]. These sequences were then used as query in a BLASTn analysis using the nucleotide sequences of the genes identified in this study or the whole genome sequences of *Xtg* strains as subject. The DNA sequences of these gene clusters were compared with Clinker v0.0.27 [40].

Presence of T2SS, T3SS, T4SS and type III effectors was determined as detailed by Goettelmann et al. [3]. Presence of a T6SS was determined using SecReT6 3.0 with default settings [41]. The presence of TALEs was determined using AnnoTALE 1.5 [42].

Results

Features of the *Xtg* genome sequences

Each strain was assembled to one circular complete chromosome. The chromosome sequences were of comparable size, ranging from 4,673,077 bp for Xtg29 to 4,734,182 bp for Xtg9. The sequence coverage ranged from 64- to 89-fold for all the sequenced strains (Table 1). Additionally, a circular contig sequence of 44,971 bp was assembled in strain NCPPB 3709. A comparison by BLASTn to the NCBI non-redundant nucleotide sequences database showed homology to plasmid sequences. This contig was thus considered to represent a plasmid of this strain.

Xtg displays a high genomic plasticity

Phylogeny based on ANI revealed that all strains of *Xtg* were very similar, with 99.9–100% ANI (Fig. 1). The

alignment fraction was >95% between *Xtg* strains, further supporting their high similarity (Figure S1). Comparison of the genomic structure within *Xtg* showed 65 to 124 locally collinear blocks (LCB), genomic regions that show no rearrangements between the compared strains (Fig. 2). This showed that although *Xtg* strains are very close genetically, a high level of rearrangements can be seen within the pathovar.

The number of genes found in each strain of *Xtg* was similar, ranging from 4,408 in Xtg9 to 4,445 in NCPPB 3709 (Table 2). The pangenome, i.e., the total number of genes, of the five *Xtg* strains used in this study consisted of 4,931 genes, while the core genome, i.e., the set of genes shared by all strains, consisted of 3,739 genes (Fig. 3A). Moreover, a total of 1,409 predicted singletons, genes that were not found in any other pathovar of the *X. translucens* species, were identified, with 536 being shared by all *Xtg* strains (Fig. 3B). A total of 433 predicted singletons were found to be strain-specific, while 131 were shared by two strains, 164 were shared by three strains, and 145 were shared by four strains. The total number of predicted singletons per strain ranged from 856 in Xtg9 to 922 in Xtg29 (Table 2).

Table 1 *Xanthomonas translucens* strains used for comparative genomic analysis and characteristics of the genome assemblies. The pv. *translucens* and pv. *cerealis* pathotype strains DSM 18974 and CFBP 2541 were included as representatives of clades Xt-I and Xt-II, respectively

Pathovar	Strain ¹	Country	Isolated from	Origin ²	Coverage	Chromosome length (bp)	Plasmid length (bp)	Accession	Reference
<i>graminis</i>	NCPPB 3709	Norway	<i>Lolium multiflorum</i>	NCPPB	89	4,732,427	44,971	CP118161-CP118162	This study
	Xtg2	Switzerland	<i>Lolium multiflorum</i>	[43]	78	4,722,726	NA	CP076253	This study
	Xtg9	Switzerland	<i>Lolium multiflorum</i>	[43]	84	4,734,182	NA	CP076252	This study
	Xtg29	Switzerland	<i>Lolium multiflorum</i>	[43]	64	4,673,077	NA	CP076257	This study
	LMG 726	Switzerland	<i>Lolium multiflorum</i>	BCCM	81	4,678,781	NA	CP076254	[3]
<i>arrhenatheri</i>	LMG 727	Switzerland	<i>Arrhenatherum elatius</i>	BCCM	103	4,843,101	NA	CP086333	[3]
<i>poae</i>	LMG 728	Switzerland	<i>Poa trivialis</i>	BCCM	86	4,792,655	NA	CP076250	[3]
<i>phlei</i>	LMG 730	Norway	<i>Phleum pratense</i>	BCCM	87	4,569,024	NA	CP076251	[3]
<i>phleipratensis</i>	LMG 843	USA	<i>Phleum pratense</i>	BCCM	121	4,902,099	NA	CP086332	[3]
<i>cerealis</i>	CFBP 2541	USA	<i>Bromus inermis</i>	CFBP	404	4,504,942	NA	CP074364	[3]
<i>translucens</i>	DSM 18974	USA	<i>Hordeum vulgare</i>	DSMZ	223	4,715,357	NA	LT604072	[44]

¹Pathotype strains are marked in bold

²NCPPB: National Collection of Plant Pathogenic Bacteria (Fera Science Ltd., Sand Hutton, United Kingdom); BCCM: Belgian Coordinated Collections of Microorganisms; CFBP: Collection of Plant Pathogenic Bacteria; DSMZ: German Collection of Microorganisms and Cell Cultures

³TBDRs: TonB-dependent receptors

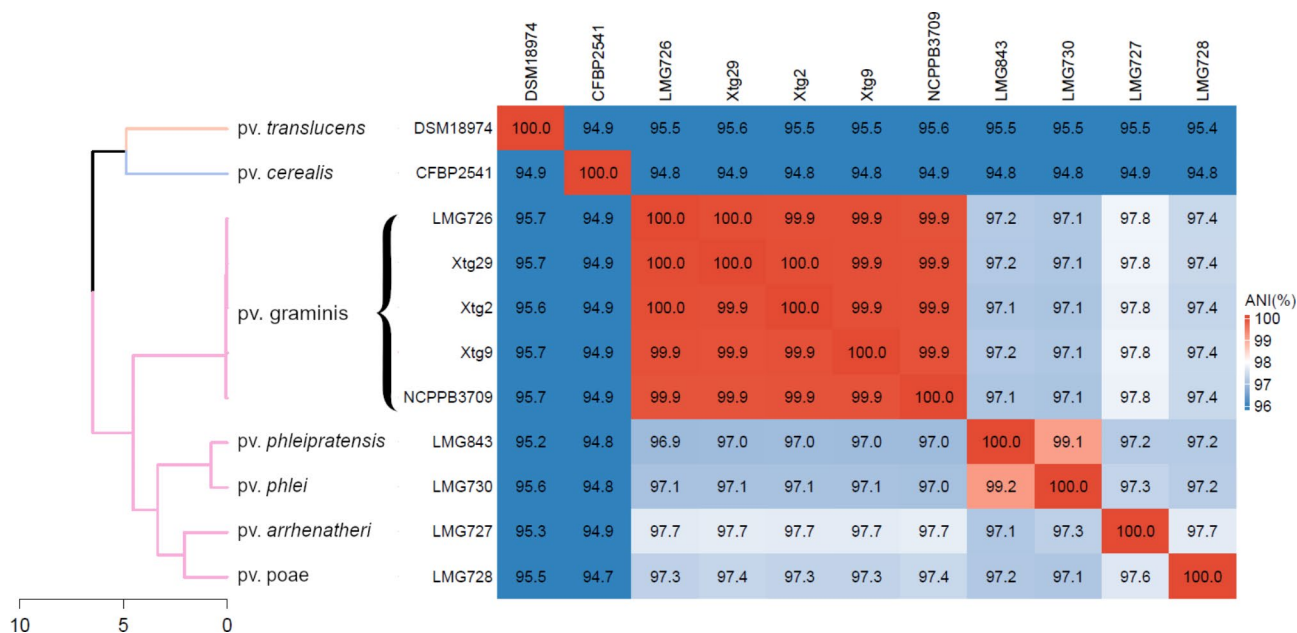


Fig. 1 Average nucleotide identity (ANI)-based phylogeny of *Xanthomonas translucens* clade Xt-III, including all complete genome sequences currently available for the clade, as well as DSM 18974 and CFBP 2541 as representatives of clades Xt-I and Xt-II, respectively. Phylogeny was constructed with Ward's hierarchical clustering method. ANI is depicted as a gradient from blue (<96%) to white (98%) to red (100%). Distance in the dendrogram represents dissimilarity between nodes. Orange: clade Xt-I, blue: clade Xt-II, pink: clade Xt-III

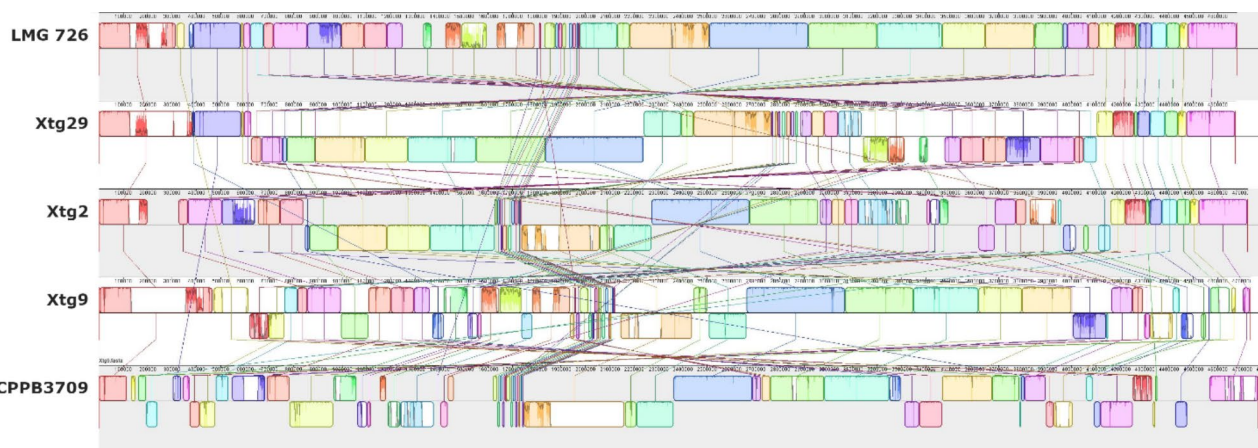


Fig. 2 Pairwise comparisons of the chromosome structure within *Xanthomonas translucens* pv. *graminis* using all five currently available complete genome sequences of the pathovar. The LMG 726 genome sequence at the top was used as a reference. Colours represent locally colinear blocks (LCB), regions that show no rearrangement across all the compared genome sequences. Synteny is displayed by a line that links orthologous LCBs. LCBs found on the bottom part represent regions that are found in reverse orientation compared to the reference

The functional characterization based on COG (Fig. 4) showed that many genes within *Xtg* could not be functionally annotated (13–15%) or were of unknown function (15–16%). The largest category with a known function was replication, recombination, and repair, with 733 to 781 genes of this category per strain (16–18%). Furthermore, this category was also largely represented in predicted singletons (48–53%), ranging from 424 genes of this category in NCPPB 3709 to 453 in *Xtg9*. Investigation of these genes revealed that most of them were annotated as transposases, suggesting a high transposase

activity in *Xtg*. Further investigation of mobile genetic elements in the *Xtg* genome sequences showed an overrepresentation of integration/excision elements, corresponding with the previously identified transposases, with 413 to 457 of such elements found per strain, compared to 24 to 123 in other strains of the species (Table 2). The number of other types of genetic mobile elements was similar between *Xtg* and other pathovars. Additionally, *Xtg* strains displayed a higher number of putative horizontal gene transfer events, with 99 to 124 predicted

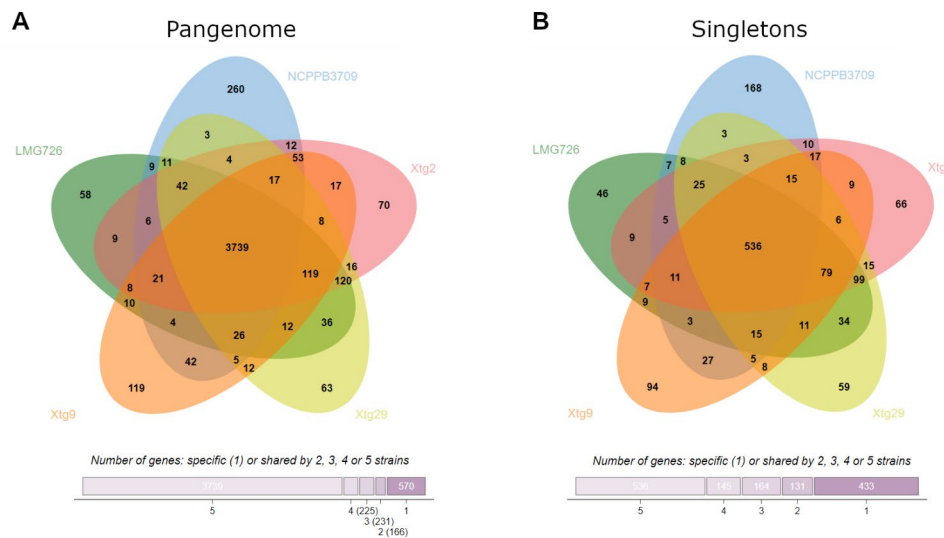


Fig. 3 **A**) Pangenome (i.e., total gene content) of the five *Xanthomonas translucens* pv. *graminis* strains investigated, and **B**) content in predicted singletons (i.e., genes that were exclusive to *X.t.* pv. *graminis* strains). Numbers indicate the number of genes shared by specific sets of strains. Bottom bars indicate the number of genes that were shared by five, four, three or two strains or that were specific to one strain

putative events, compared to 49 to 62 in the rest of the species (Table 2).

Unique features of *Xtg*

Investigation of secretion systems showed that all *Xtg* strains have the *xps* T2SS and the T3SS but not the *xcs* T2SS and T4SS (Fig. 5). Furthermore, a T6SS-i4 was found in NCPPB 3709, *Xtg2* and *Xtg9*, but no T6SS could be found in LMG 726 and *Xtg29*. The type III effector (T3E) repertoire of *Xtg* was comparable among the strains of the pathovar, though only one effector of the XopP family was found in LMG 726 and *Xtg29*, compared to two in the other strains, and a XopL family effector was identified in LMG 726 only. When comparing to other pathovars, differences could be seen in the number of effectors of some families, such as AvrBs2 and XopF effectors, with only one found in *Xtg*, compared to two in most other pathovars, though only one XopF effector was found in LMG 730 and LMG 843 as well (Fig. 5). No XopG effector could be identified in *Xtg*, while all other pathovars possessed one. Moreover, aside from one gene found in LMG 726, no XopL effector was found in the other four *Xtg* strains, when other pathovars possessed one to three effectors of this family. While all other pathovars of clade Xt-III possessed exactly two XopX effectors, three could be found in *Xtg*, as was the case for pv. *translucens* (Xt-I) and pv. *cerealis* (Xt-II). Finally, three genes of the XopE family were identified in all *Xtg* strains, while only two were found in CFBP 2541, one in LMG 730, LMG 843 as well as DSM 18974, and none were found in LMG 728 and LMG 727. Overall, the *Xtg* strains displayed the smallest T3E set, ranging from 20 to 21 per strain, compared to 24 to 26 in other pathovars of clade

Xt-III and 31 in pv. *translucens* and pv. *cerealis*. No TAL effector could be identified in any *Xtg* strain.

Among the 856 to 922 predicted singletons observed in *Xtg* strains, 627 to 647 were functionally annotated (Table 2). After excluding transposases, 181 to 222 annotated predicted singletons could be identified, of which 148 were shared by all strains (Table S2). Among these predicted singletons, genes that could have a role in the pathogenicity of *Xtg* were further investigated. As secreted or cell surface proteins are often involved in interaction with other organisms, we focused on genes that harbored a predicted secretion signal or a putative non-classical secretion. Among these, genes encoding for a filamentous hemagglutinin family outer membrane protein, a YadA family autotransporter adhesin, ten TonB-dependent receptors (TBDRs), and 11 plant cell wall degradative enzymes, including two subtilases, two esterases, a serine hydrolase, a glycerophosphodiester phosphodiesterase, a pectate lyase, a phospholipase, an alpha-galactosidase, a beta-galactosidase, and an aminopeptidase, were found. Seven additional TBDRs that were not shared by all strains were also found. Of the 17 identified TBDRs, five were annotated as diverse iron-siderophore receptors, five as oar-like proteins, three as vitamin B12 transporters, and one as a colicin receptor. However, while TBDRs were overrepresented in *Xtg*, ranging from 74 to 77 per strain, compared to 51 to 67 in other pathotype strains of the species (Table 2), many genes with similar annotations were also found among the TBDR genes that were conserved between pathovars, suggesting a functional redundancy of these TBDRs.

Additionally, a second gene predicted as encoding a filamentous hemagglutinin gene was found neighbouring

Table 2 Gene content of the *Xanthomonas translucens* strains compared in this study

Pathovar	Strain ¹	Total number of genes	Predicted singletons	Functionally annotated singletons	Functionally annotated singletons, excluding transposases	Integration/excision elements	Replication/recombination/repair elements	Phage elements	Stability/transfer/defense elements	Putative horizontal gene transfer events	TBDRs ³
<i>graminis</i>	NCPBP 3709	4,445	859	647	222	413	37	16	17	108	76
	Xtg2	4,443	915	631	189	433	34	16	17	124	76
	Xtg9	4,408	856	641	184	457	34	16	17	103	76
	Xtg29	4,422	922	627	182	427	33	15	16	99	74
<i>arrhenatheri</i>	LMG 726	4,415	904	627	181	428	34	15	17	111	77
	LMG 727	3,998	NA	NA	NA	57	35	15	19	49	63
<i>poae</i>	LMG 728	4,067	NA	NA	NA	118	34	15	12	58	62
<i>phlei</i>	LMG 730	3,957	NA	NA	NA	123	33	16	18	49	64
	LMG 843	4,158	NA	NA	NA	108	35	37	17	62	67
<i>cerealis</i>	CFBP 2541	3,886	NA	NA	NA	24	35	20	22	51	51
<i>translucens</i>	DSM 18974	4,022	NA	NA	NA	103	36	18	26	57	62

¹Pathovtype strains are marked in bold²NA: Not applicable³TBDRs: TonB-dependent receptors

the first one, being shared by all strains except NCPBP 3709. Further investigation of the locus containing the two genes revealed the presence of transposases in between them, suggesting it may be a single gene that was truncated by the insertion of these transposable elements. A BLASTn comparison with similar partial filamentous hemagglutinin genes that were previously identified in a draft genome sequence of Xtg29 revealed they corresponded to the same locus. One of the previously identified genes corresponded to the 5'-most gene that was identified in the singleton analysis, while the second gene corresponded to an additional smaller gene downstream of the two other genes we identified, which was present in all strains except NCPBP 3709 (Fig. 6). A BLASTx comparison of the hypothetical full-length gene, consisting of the combined partial sequences, revealed a high identity to other filamentous hemagglutinin genes from *Xanthomonas* spp., further confirming this locus corresponds to a single gene that was truncated.

Furthermore, a cluster containing five minor pilins of the type IV pilus (T4P) *fimT*, *pilV*, *pilW*, *pilX*, and *pilE*, as well as the anti-retraction factor *pilY1* was also found. This cluster is also present in other *X. translucens* pathovars, but the cluster found in Xtg is highly divergent from the one found in other pathovars, and these genes were thus predicted as singletons by the roary pipeline (Fig. 7A). An in-depth analysis of the minor pilins revealed that the *pilX* gene lacked the characteristic N-terminal methylation domain and class III signal peptidase cleavage site in any *X. translucens* strains investigated, suggesting it may not be functional. In contrast, these domains were present in *pilE*, *pilV* and *fimT* for all strains. When investigating *pilW*, the gene was predicted as encoding a putative pilin only in Xtg, pv. *phleipratensis* strain LMG 843 and pv. *poae* strain LMG 728 (Fig. 7B). However, the conserved N-terminal phenylalanine residue was present only in LMG 843 and was replaced by a leucine in LMG 728 and all Xtg strains.

Discussion

Comparative analysis of five Xtg strains revealed a very high genome plasticity within the pathovar compared to all other pathovars of *X. translucens*. While the strains showed 99.9 to 100% nucleotide identity, ANI compares conserved sequences, thus the genome sequences may still differ in lost or acquired sequences, as well as in genome structure. Indeed, the analysis of conserved genomic regions showed a high level of rearrangement between the five strains with 65 to 125 LCBs. Additionally, regions without any link to a similar block can be seen, which may be due to the presence of duplicated sequences, which are not considered by the MAUVE software.

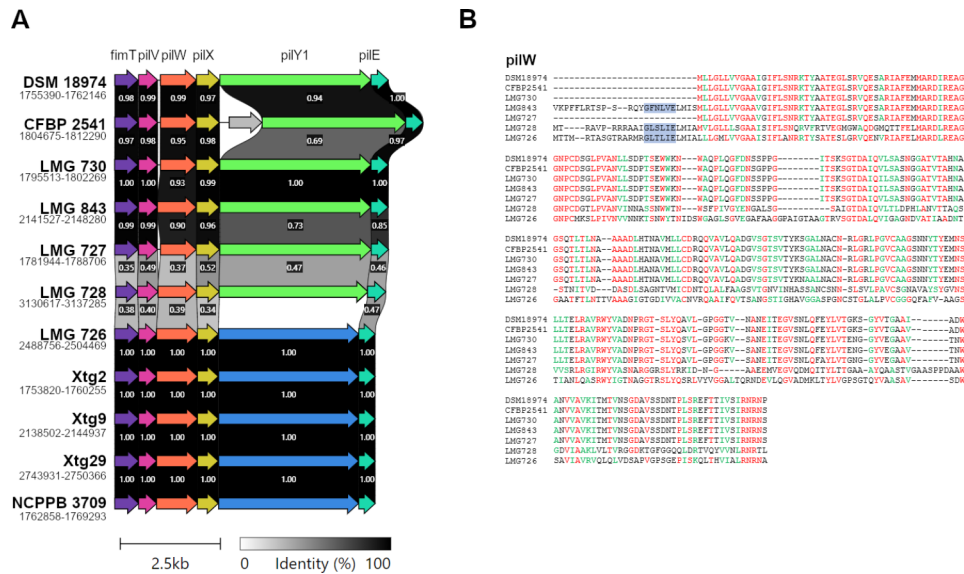


Fig. 7 **A)** Pairwise comparisons of the minor pilins gene cluster between strains of *Xanthomonas translucens* (i.e. *Xt. pv. translucens* DSM 18974, *pv. ce-real* CFBP 2541, *pv. phlei* LMG 730, *pv. phleipratensis* LMG 843, *pv. arhenatheri* LMG 727, *pv. poae* LMG 728, and *pv. graminis* LMG 726, Xtg2, Xtg9, Xtg29 and NCCPB 3709). Colors denote similar genes. The *pilY1* gene is represented in different colours in *pv. graminis* strains compared to other strains as its sequence was too divergent and they were considered as different genes by the software. Numbers below strain names refer to the position of the gene cluster in this genome sequence. **B)** Alignment of the amino acid sequences of pilW. Only the *pv. graminis* strain LMG 726 is represented as the sequences were identical in all five *pv. graminis* strains. Colours represent amino acids that were conserved among at least six out of seven strains (red) or amino acids with similar properties that were found in at least six out of seven strains (green). Highlighted in blue is the class III signal peptidase cleavage site that is characteristic to pilins

This is the first report of such a high level of rearrangements within *Xtg*, which could be observed thanks to the generation of complete genome sequences that allowed a clear observation of the genome organization in each strain. In comparison, a similar analysis with data published by Shah et al. [45] and using complete genome sequences of four *pv. translucens* and eight *pv. undulosa* strains with ANI values as low as 98.9 and 99.5%, respectively, showed a considerably lower amount of rearrangements, with 53 to 67 and 39 to 67 LCB (38, data not shown). This is in line with a previous comparative analysis within the *X. translucens* species that showed that clade Xt-III, which contains *Xtg*, was the clade showing the most rearrangements [3]. Additional genome sequences for other pathovars of clade Xt-III would help to determine whether this high level of plasticity is exclusive to *Xtg* or whether it is a characteristic of the entire clade.

The singleton analysis revealed 1,409 genes that were exclusive to the *Xtg* strains, with 536 of them being found in all five strains. Functional characterization revealed that most of these predicted singletons were annotated as transposases, and further analysis of mobile genetic elements also revealed a high amount of insertion/excision elements in the genomes of *Xtg* strains, corresponding to the identified transposases. Such a high number of transposable elements in *Xtg* could explain the high genomic plasticity that was observed, as they are genetic elements

that are mobile within the genome and can facilitate gene duplications, chromosomal recombinations and horizontal transfer [46, 47]. A high genome plasticity has previously been found in many *Xanthomonas* spp., and was suggested to play an important role in host adaptation [48]. A comparably large number of insertion sequences was reported in *X. oryzae pv. oryzae*, where these numerous mobile elements were suggested to play a role in race differentiation [49]. Insertion elements can also cause the loss of gene function, and virulence can emerge from the inactivation of avirulence genes [50, 51]. Furthermore, transposable elements have been suggested to play a role in the movement of “passenger genes” in *Xanthomonas* spp., including virulence genes, contributing to their diversification and spread [52, 53]. As such, mobile genetic elements can play a crucial role in pathogen adaptation, and the large amount of such elements in *Xtg* compared to other pathovars of the species could, at least partially, explain its high virulence and broad host range. Investigating all the identified transposable elements, their site of insertion, and the genes flanking them would be crucial in understanding their role in pathogenicity and adaptation of *Xtg* and should be tackled in the future.

The total number of genes found in each strain of the pathovar was comparable, confirming the high number of genes already identified in the *Xtg* pathotype strain LMG 726. When compared with other pathotype strains, LMG 726 exhibited the highest number of genes of the species,

as well as the highest number of genes that were exclusive to the strain [3]. While in this previous comparative analysis, as much as 950 genes were found to be exclusive to LMG 726, this number was influenced by the strains included in the comparison. The finding of 1,409 genes exclusive to *Xtg*, with 536 shared by all strains, using all available complete genome sequences of *X. translucens* strains is most likely a better estimation of the number of genes specific to *Xtg*. However, this unusually high amount seems to mainly be explained by the abundance of transposases in the pathovar. Nonetheless, many predicted singletons that could play a role in the pathogenicity and host range of *Xtg* have been uncovered.

First, a cluster containing five minor pilins and the anti-retraction factor *pilY1* was found to greatly differ between *Xtg* and other pathovars of the species. These pilins are components of the T4P, which mediates surface adhesion, biofilm formation and twitching motility [54, 55]. The T4P consists of a filamentous polymer of the major pilin PilA, and minor pilins may be incorporated into the pilus. Further investigation of minor pilins of this cluster revealed that only *pilE*, *pilV* and *fimT* may be functional in all pathovars of the species. For *pilW*, the N-terminal cleavage and methylation domain that is characteristic to pilins and required for their processing was present only in pvs. *graminis*, *poae*, and *phleipratensis*, though the conserved phenylalanine that undergoes methylation after cleavage was only present in pv. *phleipratensis* and was substituted by a leucine in the other two pathovars. However, it was previously shown that although the phenylalanine residue is highly conserved in pilins, its substitution by another amino acid did not affect secretion, cleavage and methylation of the pilin and subsequent T4P assembly [56]. It is thus plausible that *pilW* is still functional in pvs. *graminis* and *poae*. The six identified genes, as well as the major pilin PilA were previously shown to differ significantly in *Xtg*29 compared to strains of other *X. translucens* pathovars, while the genes encoding the main components of the T4P machinery were highly conserved [20]. That this divergent cluster is conserved within *Xtg* could indicate a role in the virulence of the pathovar. Indeed, the T4P is an extracellular feature of the bacteria that may be detected by plant defense mechanisms. The high sequence variation in the minor pilins, as well as the major pilin PilA may allow *Xtg* to avoid plant defenses while mediating surface adhesion and biofilm formation.

Among other predicted singleton genes, two filamentous hemagglutinin genes were identified. However, further investigation revealed that they likely correspond to a single gene in which multiple transposases were inserted. Moreover, these genes corresponded to the two partial filamentous hemagglutinin genes that were previously reported as predicted singletons of *Xtg* [20].

The N-terminus of filamentous hemagglutinins possesses a signal peptide, as well as a two partner system domain that are involved in their secretion to the periplasm and the apoplasm, respectively [57, 58]. Whether the gene encoding for the N-terminal part of the protein is still functional remains to be determined, though it is unlikely, as essential domains may be found in the C-terminal part of the gene. In *X. axonopodis* pv. *citri*, filamentous hemagglutinins were shown to be crucial to virulence by promoting plant surface adhesion and biofilm formation but acted as elicitors of plant defense in *X. campestris* pv. *vesicatoria* [59, 60]. Thus, the loss or truncation of this gene in *Xtg* may help it evade recognition by the plant host.

Additionally, 17 predicted singleton genes coding for TBDRs were found, with ten being shared by all *Xtg* strains. TBDRs are known for their role in iron and vitamin B12 uptake and are mostly regulated by the ferric uptake regulator gene [61]. Mutants of this regulator were unable to induce disease symptoms in *X. o. pv. oryzae* which was hypothesized to be due to an inability to cope with the oxidative stress conditions encountered during the infection, emphasizing the role of TBDRs in virulence [62]. Moreover, in *X. campestris* pv. *campestris*, some TBDRs were found to be activated by HrpG and HrpX, regulatory proteins of the type III secretion system, and the SuxA TBDR was found to play a major role in pathogenicity by acting on sucrose import [63]. A comparative analysis across 226 eubacterial genome sequences showed that TBDRs were overrepresented in *Xanthomonas* spp. compared to other bacteria, ranging from 36 TBDRs in *X. oryzae* pv. *oryzae* to 68 in *X. axonopodis* pv. *citri* [63]. In this study, similar numbers of TBDRs were found in all strains of *X. translucens*, with an overrepresentation in *Xtg*. The identified predicted singletons were mainly annotated with functions related to iron and vitamin B12 uptake, and similar genes were found to be shared across the species. It is thus unclear if these 17 predicted singletons serve a specific function in the pathogenicity of *Xtg*, or whether they are functionally redundant with other TBDRs.

A total of 11 plant cell wall degradative enzymes were also found in singletons that were shared by all *Xtg* strains and were predicted to be secreted. Such enzymes are often secreted by the T2SS and involved in the degradation of the plant cell wall [16, 18]. These include cellulases, xylanases, pectate lyases, and polygalacturonases, which degrade the main polysaccharide cell wall constituents, as well as various proteases and lipases. *X. oryzae* pv. *oryzae* mutants lacking a T2SS were shown to be virulence deficient and multiple enzymes secreted by the T2SS were shown to be necessary for virulence [64–66]. As *Xtg* mutants lacking a T3SS showed a drastic reduction in symptoms, but still survived in the plant, it was

hypothesized that other virulence factors may be important in *Xtg* [15]. Therefore, the T2SS and the enzymes it secretes may play a crucial role in the *in planta* growth of *Xtg*.

Finally, a gene encoding for a YadA family autotransporter adhesin was predicted as a singleton shared by all *Xtg* strains. Such adhesins showing a high similarity to YadA from *Yersinia* spp. have been shown to mediate cell to cell aggregation, biofilm formation, and adhesion to the host cell surface [67, 68]. In *Xanthomonas oryzae* pv. *oryzae*, mutations in the genes encoding for the YadA-like adhesins XadA and XadB were shown to result in reduced surface attachment and virulence after surface inoculation [69, 70]. In *Xtg*, this adhesin may therefore play a similar role in attachment to the host.

An in-depth analysis of the secretion systems within *Xtg* showed that a T6SS-i4 was present in strains Xtg2, Xtg9 and NCPPB 3709 and not in LMG 726 and Xtg29, confirming previous reports investigating these strains [20]. It was previously hypothesized that the T4SS and T6SS could act as elicitors of plant defense, and that their absence in LMG 726 could help it evade plant defenses [3]. However, the presence of a T6SS in Xtg2, Xtg9 and NCPPB 3709 may disprove this hypothesis. Nonetheless, its absence in Xtg29 and LMG 726 indicates that it is not crucial to the pathogen's survival and virulence. Conversely, the *xps* T2SS and the T3SS were found to be conserved, suggesting they may be crucial to the pathovar.

The T3E repertoire within the pathovar was largely conserved, except for only one XopP family effector being present in strains Xtg2, Xtg9 and NCPPB 3709, compared to two in LMG 726 and Xtg29, as well as one XopL family effector being present only in the LMG 726 strain. This XopL effector could have been conserved in LMG 726, or it might have been recently acquired through horizontal transfer. Nonetheless, as it is unique to LMG 726, it is likely not crucial to the lifecycle of *Xtg*. Rather, the absence of XopL effectors may have been beneficial to other strains of *Xtg*, as they all lacked such effectors, while all other pathovars of the species possessed at least one and up to three. Similarly, several effectors appear to have been lost in *Xtg*, such as AvrBs2, XopF, or XopG, as the pathovar possesses fewer of these effectors than other pathovars. Furthermore, a comparable loss of a XopP effector may have occurred in strains LMG 726 and Xtg29. Interestingly, *Xtg* possesses the smallest set of T3E in the species, which may be explained by such a loss of effectors during evolution. Finally, no TALEs were identified, further confirming the absence of such effectors in the pathovar [3, 15]. This suggests that these T3E were either unnecessary or even deleterious to the pathovar. As these effectors could trigger plant defence, their loss may have contributed to the success of *Xtg* as a pathogen of many grass species. On the other hand, three effectors

of the XopE family were present in *Xtg*, while only two, one, or none were found in other pathovars. Additionally, three XopX effectors were found in *Xtg*, while other pathovars of clade Xt-III possessed only two. XopE effectors belong to the HopX family, which contains transglutaminases and possess a catalytic site that may be involved in proteolysis [71]. The XopE2 effector was shown to suppress hypersensitive response in *Nicotiana* spp. and to inhibit pattern-triggered immunity [72, 73]. XopX effectors were found to be required for virulence in *X. euvesicatoria* on tomato and pepper by interfering with plant immune responses [74, 75]. This suggests that XopE and XopX effectors may play a role in the pathogenicity of *Xtg* by suppressing plant defences.

Phylogeny based on ANI showed that all five available *Xtg* strains were highly identical, with 99.9 to 100% ANI. A previous analysis of the genetic diversity of *Xtg* using AFLP markers also showed that multiple strains from Switzerland, France and Belgium displayed a very low diversity [43]. In other pathovars of the species, a higher diversity was observed, with only 99% ANI within pv *undulosa* and pv. *translucens* [45]. Although these strains were from a more diverse origin than in *Xtg*, with strains from North and South America as well as Iran, there was still a high diversity when comparing strains of a similar origin. This further highlights the particularity of the low diversity observed within *Xtg*. Still, as aside from one Norwegian strain, all *Xtg* strains compared in this study originate from Switzerland, obtaining additional complete genome sequences from strains of more diverse origins would help to better determine the genomic diversity existing within the pathovar.

Overall, many common features of *X. translucens* were lacking in *Xtg* and seem to have been lost over the course of evolution, as they were present in the rest of the species. No T4SS was found in the pathovar and a T6SS was found only in strains NCPPB 3709, Xtg2 and Xtg9. It had the smallest set of T3E and no TALEs were identified. Previous research also showed that although a nearly complete flagellar gene cluster is found in strain NCPPB 3709, such a cluster was absent in all other strains investigated [20]. In this regard, the high mobile genetic element activity observed in *Xtg* could have played an important role in this loss of virulence features through transposable element insertion within a gene, or by facilitating chromosomal region deletions [50, 51, 76]. Furthermore, a highly divergent T4P was found, both in terms of the amino acid sequence of the pilins, as well as in the presence or absence of functional minor pilins, compared to the rest of the species. This loss or divergence of main virulence features in *Xtg* may be responsible for its ability to infect a large range of plant hosts, by allowing the bacteria to evade plant defense mechanisms.

Conclusion

Our study provided new high-quality complete genome sequences for *X. translucens* *Xtg*, allowing for an in-depth comparative genome analysis within the pathovar. This revealed that *Xtg* exhibits a remarkable genome plasticity, likely related to an unusually high amount of transposases, which could explain its success as a critical pathogen of a large range of forage grasses. A set of potential virulence factors of the pathovar were identified, and a comprehensive analysis of these genes will contribute to better understand its virulence and host range, providing a basis for the development of new cultivars of forage grasses with increased resistance to the disease.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09855-8>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

This article is based upon work from COST Action CA16107 EuroXanth, supported by COST (European Cooperation in Science and Technology). Library preparation and sequencing were done by the Functional Genomics Center Zurich (Zurich, Switzerland). RaK and RoK are grateful for an international mobility EXPLORE grant from the Montpellier University of Excellence (MUSE). This manuscript was made available as a preprint on bioRxiv (<https://doi.org/10.1101/2023.06.29.547023>).

Author contributions

Design of the study: FG and RoK, with the help of BS. Genome sequencing, assembly and data analysis: FG with the help of VR and RaK. Data interpretation, writing and reviewing: FG, RoK, VR, RaK and BS. All authors contributed to the article and approved the submitted version.

Funding

Open access funding provided by Swiss Federal Institute of Technology Zurich. Funding was provided by the Swiss National Science Foundation (Grant No: IZCOZO_177062). Open access funding provided by Swiss Federal Institute of Technology Zurich

Data Availability

The genome sequences generated for this study can be found in the NCBI GenBank repository under the accession numbers specified in Table 1. The Prokka annotations can be found at <https://zenodo.org/records/10077657>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland

²Plant Health Institute of Montpellier, University of Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

³Department of Plant Pathology and Environmental Microbiology, The Pennsylvania State University, University Park, PA, USA

Received: 5 July 2023 / Accepted: 28 November 2023

Published online: 05 December 2023

References

1. Egli T, Goto M, Schmidt D. Bacterial wilt, a new forage grass disease. *J Phytopathol.* 1975;82:111–21.
2. Sapkota S, Mergoum M, Liu Z. The translucens group of *Xanthomonas translucens*: complicated and important pathogens causing bacterial leaf streak on cereals. *Mol Plant Pathol.* 2020;21:291–302.
3. Goettelmann F, Roman-Reyna V, Cunnac S, Jacobs JM, Bragard C, Studer B, et al. Complete genome assemblies of all *Xanthomonas translucens* pathotype strains reveal three genetically distinct clades. *Front Microbiol.* 2022;12:817815.
4. Egli T, Schmidt D. Pathogenic variation among the causal agents of bacterial wilt of forage grasses. *J Phytopathol.* 1982;104:138–50.
5. Roberts DL, Vargas JM Jr, Detweiler R. Occurrence of bacterial wilt on *Poa annua* and other turfgrasses. *Phytopathology.* 1985;75:1289.
6. Falloon RE, Hume DE. Productivity and persistence of prairie grass (*Bromus willdenowii* Kunth) 1. Effects of the head smut fungus *Ustilago bullata* Berk. *Grass Forage Sci.* 1988;43:179–84.
7. Duveiller E, Bragard C, Maraite H. Bacterial leaf streak and black chaff caused by *Xanthomonas translucens*. In: Duveiller E, Fucikovskiy L, Rudolph K, editors. *The bacterial Diseases of wheat: concepts and methods of disease management.* Mexico City, Mexico: CIMMYT; 1997. pp. 25–47.
8. Leyns F, De Cleene M, Van Bogaert G, Van de Wiele A, De Ley J. Preliminary investigations about the mode of transmission and spread of *Xanthomonas campestris* pv. *graminis* on forage grasses. *J Phytopathol.* 1988;122:76–88.
9. Rudolph K. Infection of the plant by *Xanthomonas*. In: Swings JG, Civerolo EL, editors. *Xanthomonas.* Dordrecht: Springer Netherlands; 1993. pp. 193–264.
10. Leyns F. *Xanthomonas campestris* Pv *graminis*: cause of bacterial wilt of forage grasses. In: Swings JG, Civerolo EL, editors. *Xanthomonas.* Dordrecht: Springer Netherlands; 1993. pp. 55–7.
11. Schmidt D, Nuesch B. Resistance to bacterial wilt (*Xanthomonas graminis*) increases yield and persistency of *Lolium multiflorum*. *EPPO Bull.* 1980;10:335–9.
12. White FF, Potnis N, Jones JB, Koebnik R. The type III effectors of *Xanthomonas*. *Mol Plant Pathol.* 2009;10:749–66.
13. Boch J, Bonas U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol.* 2010;48:419–36.
14. Zhang B, Han X, Yuan W, Zhang H. TALEs as double-edged swords in plant-pathogen interactions: Progress, challenges, and perspectives. *Plant Commun.* 2022;3:100318.
15. Wichmann F, Vorhölter F-J, Hersemann L, Widmer F, Blom J, Niehaus K, et al. The noncanonical type III secretion system of *Xanthomonas translucens* pv. *graminis* is essential for forage grass infection. *Mol Plant Pathol.* 2013;14:576–88.
16. Jha G, Rajeshwari R, Sonti RV. Bacterial type two secretion system secreted proteins: double-edged swords for plant pathogens. *Mol Plant Microbe Interact.* 2005;18:891–8.
17. Szczesny R, Jordan M, Schramm C, Schulz S, Coge V, Bonas U et al. Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. *New Phytologist.* 2010;187:983–1002.
18. Büttner D, Bonas U. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiol Rev.* 2010;34:107–33.
19. Alvarez-Martinez CE, Sgro GG, Araujo GG, Paiva MRN, Matsuyama BY, Guzzo CR, et al. Secrete or perish: the role of secretion systems in *Xanthomonas* biology. *Comput Struct Biotechnol J.* 2021;19:279–302.
20. Hersemann L, Wibberg D, Blom J, Goesmann A, Widmer F, Vorhölter F-J et al. Comparative genomics of host adaptive traits in *Xanthomonas translucens* pv. *graminis*. *BMC Genomics.* 2017;18:35.

21. Peng Z, Hu Y, Xie J, Potnis N, Akhunova A, Jones J, et al. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics*. 2016;17:21.
22. PacBio. Procedure & checklist – preparing multiplexed microbial libraries using SMRTbell express template prep kit 2.0. 2021.
23. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 2019;37:540–6.
24. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res*. 2017;27:722–36.
25. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
26. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods*. 2016;8:12–24.
27. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022.
28. Darling ACE. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 2004;14:1394–403.
29. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31:3691–3.
30. Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C. Jvarkit: an interactive Venn diagram viewer. *BMC Bioinformatics*. 2014;15:293.
31. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47:D309–14.
32. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825–9.
33. Brown CL, Mullet J, Hindi F, Stoll JE, Gupta S, Choi M, et al. mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. *Appl Environ Microbiol*. 2022;88:e00991–22.
34. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C et al. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res*. 2023;gkad326.
35. Vernikos GS, Parkhill J. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. *Bioinformatics*. 2006;22:2196–203.
36. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol*. 2022;40:1023–5.
37. Bendtsen JD, Kiemer L, Fausbøll A, Brunak S. Non-classical protein secretion in bacteria. *BMC Microbiol*. 2005;5:58.
38. Imam S, Chen Z, Roos DS, Pohlschröder M. Identification of surprisingly diverse type IV pili, across a broad range of Gram-positive bacteria. *PLoS ONE*. 2011;6:e28919.
39. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterPro-Scan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–40.
40. Gilchrist CLM, Chooi Y-H. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics*. 2021;37:2473–5.
41. Li J, Yao Y, Xu HH, Hao L, Deng Z, Rajakumar K, et al. SecReT6: a web-based resource for type VI secretion systems found in bacteria: Genomics update. *Environ Microbiol*. 2015;17:2196–202.
42. Grau J, Reschke M, Erkes A, Streubel J, Morgan RD, Wilson GG, et al. AnnoTALE: bioinformatics tools for identification, annotation and nomenclature of TALEs from *Xanthomonas* genomic sequences. *Sci Rep*. 2016;6:21077.
43. Kölliker R, Kraehenbuehl R, Boller B, Widmer F. Genetic diversity and pathogenicity of the grass pathogen *Xanthomonas translucens* pv. *raminis*. *Syst Appl Microbiol*. 2006;29:109–19.
44. Jaenicke S, Bunk B, Wibberg D, Spröer C, Hersemann L, Blom J, et al. Complete genome sequence of the barley pathogen *Xanthomonas translucens* pv. *translucens* DSM 18974^T (ATCC 19319^T). *Genome Announc*. 2016;4:e01334–16.
45. Shah SMA, Khojasteh M, Wang Q, Taghavi SM, Xu Z, Khodaygan P, et al. Genomics-enabled novel insight into the pathovar-specific population structure of the bacterial leaf streak pathogen *Xanthomonas translucens* in small grain cereals. *Front Microbiol*. 2021;12:674952.
46. Gray YHM. It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. *Trends Genet*. 2000;16:461–8.
47. Vandecraen J, Chandler M, Aertsen A, Van Houdt R. The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit Rev Microbiol*. 2017;43:709–30.
48. Timilsina S, Potnis N, Newberry EA, Liyanapathirana P, Iruegas-Bocardo F, White FF, et al. *Xanthomonas* diversity, virulence and plant–pathogen interactions. *Nat Rev Microbiol*. 2020;18:415–27. <https://doi.org/10.1038/s41579-020-0361-8>.
49. Ochiai H, Inoue Y, Takeya M, Sasaki A, Kaku H. Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. *JARQ*. 2005;39:275–87.
50. Kang S, Lebrun MH, Farrall L, Valent B. Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. *Mol Plant Microbe Interact*. 2001;14:671–4.
51. Inami K, Yoshioka-Akiyama C, Morita Y, Yamasaki M, Teraoka T, Arie T. A genetic mechanism for emergence of races in *Fusarium oxysporum* f. sp. *lycopersici*: inactivation of avirulence gene AVR1 by transposon insertion. *PLoS ONE*. 2012;7:e44101.
52. Ferreira RM, de Oliveira ACP, Moreira LM, Belasque J, Gourbeyre E, Siguiet P, et al. A TALE of transposition: Tn3-like transposons play a major role in the spread of pathogenicity determinants of *Xanthomonas citri* and other xanthomonads. *mBio*. 2015;6:e02505–14.
53. Gochez AM, Huguet-Tapia JC, Minsavage GV, Shantaraj D, Jalan N, Strauß A, et al. Pacbio sequencing of copper-tolerant *Xanthomonas citri* reveals presence of a chimeric plasmid structure and provides insights into reassortment and shuffling of transcription activator-like effectors among *X. citri* strains. *BMC Genomics*. 2018;19:16.
54. Dunger G, Guzzo CR, Andrade MO, Jones JB, Farah CS. *Xanthomonas citri* subsp. *citri* type IV pilus is required for twitching motility, biofilm development, and adherence. *Mol Plant Microbe Interact*. 2014;27:1132–47.
55. Dunger G, Llontop E, Guzzo CR, Farah CS. The *Xanthomonas* type IV pilus. *Curr Opin Microbiol*. 2016;30:88–97.
56. Strom MS, Lory S. Amino acid substitutions in pilin of *Pseudomonas aeruginosa*. Effect on leader peptide cleavage, amino-terminal methylation, and pilus assembly. *J Biol Chem*. 1991;266:1656–64.
57. Villarino Romero R, Osicka R, Sebo P. Filamentous hemagglutinin of *Bordetella pertussis*: a key adhesin with immunomodulatory properties? *Future Microbiol*. 2014;9:1339–60.
58. Guérin J, Bigot S, Schneider R, Buchanan SK, Jacob-Dubuisson F. Two-partner secretion: combining efficiency and simplicity in the secretion of large proteins for bacteria–host and bacteria–bacteria interactions. *Front Cell Infect Microbiol*. 2017;7:148.
59. Gottig N, Garavaglia BS, Garofalo CG, Orellano EG, Ottado J. A filamentous hemagglutinin-like protein of *Xanthomonas axonopodis* pv. *citri*, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS ONE*. 2009;4:e4358.
60. Choi HW, Kim DS, Kim NH, Jung HW, Ham JH, Hwang BK. *Xanthomonas* filamentous hemagglutinin-like protein Fha1 interacts with pepper hypersensitive-induced reaction protein CaHIR1 and functions as a virulence factor in host plants. *Mol Plant Microbe Interact*. 2013;26:1441–54.
61. Postle K, Kadner RJ. Touch and go: tying TonB to transport. *Mol Microbiol*. 2003;49:869–82.
62. Subramoni S, Sonti RV. Growth deficiency of a *Xanthomonas oryzae* pv. *oryzae* *fur* mutant in rice leaves is rescued by ascorbic acid supplementation. *Mol Plant Microbe Interact*. 2005;18:644–51.
63. Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, et al. Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE*. 2007;2:e224.
64. Ray SK, Rajeshwari R, Sonti RV. Mutants of *Xanthomonas oryzae* pv. *oryzae* deficient in general secretory pathway are virulence deficient and unable to secrete xylanase. *Mol Plant Microbe Interact*. 2000;13:394–401.
65. Hu J, Qian W, He C. The *Xanthomonas oryzae* pv. *oryzae* *eglXoB* endoglucanase gene is required for virulence to rice. *FEMS Microbiol Lett*. 2007;269:273–9.
66. Rajeshwari R, Jha G, Sonti RV. Role of an in planta-expressed xylanase of *Xanthomonas oryzae* pv. *oryzae* in promoting virulence on rice. *Mol Plant Microbe Interact*. 2005;18:830–7.
67. Feil H, Feil WS, Lindow SE. Contribution of fimbrial and afimbrial adhesins of *Xylella fastidiosa* to attachment to surfaces and virulence to grape. *Phytopathology*. 2007;97:318–24.
68. Totsika M, Wells TJ, Beloin C, Valle J, Allsopp LP, King NP, et al. Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic *Escherichia coli*. *Appl Environ Microbiol*. 2012;78:2179–89.

69. Ray SK, Rajeshwari R, Sharma Y, Sonti RV. A high-molecular-weight outer membrane protein of *Xanthomonas oryzae* pv. *oryzae* exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. *Mol Microbiol.* 2002;46:637–47.
70. Das A, Rangaraj N, Sonti RV. Multiple adhesin-like functions of *Xanthomonas oryzae* pv. *oryzae* are involved in promoting leaf attachment, entry, and virulence on rice. *Mol Plant Microbe Interact.* 2009;22:73–85.
71. Nimchuk ZL, Fisher EJ, Desveaux D, Chang JH, Dangl JL. The HopX (AvrPphE) family of *Pseudomonas syringae* type III effectors require a catalytic triad and a novel N-terminal domain for function. *Mol Plant Microbe Interact.* 2007;20:346–57.
72. Lin R-H, Peng C-W, Lin Y-C, Peng H-L. The XopE2 effector protein of *Xanthomonas campestris* pv. *vesicatoria* is involved in virulence and in the suppression of the hypersensitive response. *Bot Stud.* 2011;52:55-72.
73. Popov G, Fraiture M, Brunner F, Sessa G. Multiple *Xanthomonas euvesicatoria* type III effectors inhibit flg22-triggered immunity. *Mol Plant Microbe Interact.* 2016;29:651–60.
74. Metz M, Dahlbeck D, Morales CQ, Sady BA, Clark ET, Staskawicz BJ. The conserved *Xanthomonas campestris* pv. *vesicatoria* effector protein XopX is a virulence factor and suppresses host defense in *Nicotiana Benthamiana*. *Plant J.* 2005;41:801–14.
75. Stork W, Kim J-G, Mudgett MB. Functional analysis of plant defense suppression and activation by the *Xanthomonas* core type III effector XopX. *Mol Plant Microbe Interact.* 2015;28:180–94.
76. Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev.* 2014;38:865–91.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.