Complete genome dynamics of a dominant-lineage strain of Xanthomonas oryzae pv. oryzae harbouring a novel plasmid encoding a type IV secretion system

Amandeep Kaur¹, Kanika Bansal¹, Sanjeet Kumar¹†, Ramesh V. Sonti²,³ and Prabhu B. Patil¹,*

INTRODUCTION

Xanthomonas is a yellow-pigmented, Gram-negative bacteria that causes serious diseases in a wide range of economically important plants, such as rice, sugarcane, banana, citrus, pepper, cabbage, etc. Among Xanthomonas species, Xanthomonas oryzae pv. oryzae (Xoo), which causes bacterial blight disease in rice, is a major threat to rice cultivation. Xoo strains display tremendous variation in their pathogenicity towards different resistance genes that are deployed for the development of resistant rice cultivars [1]. This is also reflected in the large-scale inter-strain variation found at a locus encoding lipopolysaccharide biosynthetic gene clusters in the genome of Xoo isolates [2]. Population genomic studies of more than 100 isolates revealed that inter-strain variation is more conspicuous than inter-lineage variation in Xoo species [3]. Hence, it is important to carry out systematic and genomic characterizations of Xoo strains to comprehend the mechanisms underlying the rapid evolution and success of this important pathogen infecting a major staple crop.

Third-generation long-read sequencers such as PACBIO and ONT generate longer reads for assembling the complete genomes of many species, allowing us to understand the dynamic variations at inter-strain levels, which occur through plasmids, repetitive elements such as transcriptional activator-like effectors (TALE), integrons, conjugative and IS elements, etc. [4]. The presence of plasmids can provide ecological fitness to the pathogen to survive in every environment or to withstand adverse conditions [5]. Plasmids are also known to carry virulence genes or toxins contributing to the pathogenicity of a bacterium [6, 7]. Unfortunately,
the vast majority of available genomic sequences are drafts, making it difficult to study inter-strain variations mediated by plasmids in a systematic manner. However, with the advent of cost-effective long-read technologies, it is possible to obtain gap-free genome sequences that provide insights into the plasmid content of an isolate and facilitate comparative genomic analyses [8].

In the present study, we carried out a genome-based investigation of Xoo strain BXO1, which belongs to a dominant lineage of this pathogen in India [3]. We also report here that this strain harbours two novel plasmids, pBXO1-1 and pBXO1-2, which have not been reported in the complete genome of any Xoo strain published to date. In a recent study, another strain from the dominant Indian Xoo lineage was found to harbour a novel plasmid with a size of 43 Kb [9], but this plasmid is not related to the plasmids of the BXO1 strain. While pBXO1-1 encodes a cluster of hypothetical genes with atypical GC content, pBXO1-2 encodes a T4SS cluster that has not yet been reported in Xoo on either chromosomes or plasmids. Further, its presence in another species of Xanthomonas i.e. Xanthomonas albilineans, which infects sugarcane (monocot), suggests that it provides an adaptive advantage to Xanthomonas spp. infecting different monocot plants. Our study has provided novel insights into plasmid-mediated variation in Xoo strains and emphasizes the importance of complete genome studies to address inter-strain variations in the Xoo population.

RESULTS

Complete genome sequencing of Xoo strain BXO1

Complete genome sequencing of the BXO1 strain was performed using the Oxford Nanopore MinION platform. The assembled sequence of BXO1 contains a circular chromosome of 4 991 257 bp with two plasmids (pBXO1-1 and pBXO1-2) with a size of 66 752 and 25 634 bp, respectively. The basic features of BXO1’s complete genome are given in Table 1. The genome coverage obtained was 127x with average G+C content of 63.7 %. A total of 4 857 coding sequences (CDSs), 53 tRNAs and 2 copies of the rrr operon (5 s, 16 s, 23 s) were found on the genome. The average G+C content of pBXO1-1 and pBXO1-2 was 61 and 57.3 %, with 71 and 29 CDSs, respectively. Further, completeness and contamination was found to be 100 and 0.36 %, respectively. We also looked for TAL effectors in the complete genome and found 19 TAL genes on the chromosomal DNA, whereas both plasmids lack TALE genes. A schematic representation of BXO1’s genome is shown in Fig. 1.

Complete genome-based investigation reveals novel plasmids in Xoo

We identified two plasmids in the BXO1 strain, pBXO1-1 and pBXO1-2, with sizes of 66.7 and 25.6 kb, respectively. The pBXO1-1 plasmid encodes 71 genes, out of which 38 genes were found to be hypothetical. Hypothetical genes were further checked for homology via a BLASTP search of the National Center for Biotechnology Information (NCBI) database. However, these hypothetical genes possess atypical GC content, pointing towards their acquisition through horizontal gene transfer. Although both plasmids lack type III effector genes, annotation of the plasmid pBXO1-1 revealed the presence of different conjugative transfer genes and toxins/antitoxins (relE/parE family toxin, phd antitoxin) (Fig. 2). The different conjugative transfer genes present on pBXO1-1 are traY, traW, traU, traQ, traO, traN, traM, traH, traL, traK, trwB and trbA, trbB, trbN and mobD, which enable self-transfer of plasmids. In addition to the conjugative transfer genes, the other core genes present on pBXO1-1 include repA (replication initiator protein A), parA (chromosome partitioning protein) and genes that encode DNA topoisomerase, restriction endonuclease and XRE family transcriptional regulator. Further, pBXO1-1 showed 93.61 % homology to Paraburkholderia strain DSM 17164 plasmid pEMT1 (CP026110.1) and 93.64 % homology to Paraburkholderia aromaticivorans strain BN5 plasmid pBN4 (CP022994.1).

pBXO1-2 encodes 29 genes, out of which 10 genes encode the T4SS cluster (Vir components) (Fig. 2). pBXO1-2-encoded vir genes, including virD2, virD4, virB2, virB3, virB4, virB5, virB6, virB8, virB9, virB10 and virB11. The other genes present on pBXO1-2 include parA and parB, traA, traC, spoT and kfrA, which are involved in core functions, i.e. maintaining plasmid replication, partitioning etc.

Inter-species movement of a novel plasmid in Xanthomonas pathogens infecting monocots

pBXO1-2 plasmid showed 96 % similarity to the X. albilineans str. GPE PC73 plasmid (FP340277.1), suggesting its movement at the inter-species level. A comparison between the pBXO1-2 and GPE PC73 plasmids is shown in Fig. 3. X. albilineans is another member of the genus Xanthomonas, which is a xylem-invading pathogen causing leaf scald disease in sugarcane. X. albilineans is considered to be an evolutionary intermediate between several Xanthomonas species that has undergone reductive genome evolution [10].

DISCUSSION

Bacterial populations display astonishing inter-strain variations mediated by horizontal gene transfer involving the movement of mobile genetic elements (MGEs) such as plasmids, integrative conjugative and IS elements, etc.

| Table 1. General features of Xanthomonas oryzae pv. oryzae BXO1 strain |
|--------------|---------|---------|---------|
| Chromosome   | Plasmid 1 (pBXO1-1) | Plasmid 2 (pBXO1-2) |
| Size (bp)    | 4 991 257 | 66 752 | 25 634 |
| GC content   | 63.7     | 61     | 57.3   |
| CDS          | 4857     | 71     | 29     |
| tRNA         | 53       | 0      | 0      |
| Ribosomal RNA operon | 2     | 0      | 0      |
| IS elements (complete) | 301  | 0      | 0      |
| TAL effector genes | 19    | 0      | 0      |
[11]. Hence, a systematic study of the genome dynamics of a successful strain is important, particularly in pathogenic bacteria. Presently, the vast majority of genome sequences are drafts, making it difficult to undertake a detailed and systematic study of mobile DNA. Our complete genome study allowed us to identify two novel plasmids in a Xoo strain, indicating that plasmid-mediated variation(s) may be playing an important role in the evolution of this rice pathogen. The fact that the plasmid is present in another Xanthomonas species (X. albilineans) suggests that plasmids may be actively moving between different species in this genus, thereby contributing to inter-strain and interspecies dynamics. However, the possibility that Xoo strain BXO1 and X. albilineans str. GPE PC73 might have independently obtained this plasmid from a non-Xanthomonas source cannot be ruled out. Interestingly, both Xoo and X. albilineans infect the vascular tissues of monocot plants, suggesting that this plasmid may have adaptive value for growth within this tissue.

Xanthomonas T4SS seems to be a highly diverse and versatile system [12]. It has been shown that Xanthomonas citri deploys the T4SS to secrete toxins that contain conserved C-terminal domains to kill other bacterial species in a contact-dependent manner [13]. Similarly, Burkholderia cepacia contains a 92 kb plasmid that encodes T4SS, which is responsible for providing the plant tissue water-soaking (PTW) phenotype on onion tissue [14]. Intriguingly, pBXO1-2-encoded T4SS did not show any similarity with Xanthomonas axonopodis pv. citri and Xanthomonas campestris pv. vesicatoria (strain 85–10) T4SS, pointing towards functional diversification of T4SS among Xanthomonas [15]. Thus, further studies are required to prove its role in fitness, virulence or pathogenicity in the hosts. However, the distribution of the plasmid with T4SS indicates that the plasmid is playing a role in the ecology and evolution of Xanthomonas species.

Short-read sequencing technology platforms have enabled researchers to generate draft genomes of more than 100 strains in a population-based study [3]. This has allowed an understanding of relationship and population structure to a level of detail that was not possible using a few housekeeping genes [16–18]. However, a complete genome sequence is required to systematically understand inter-strain variation originating from repetitive and mobile elements. The advent of long-read technologies such as Oxford Nanopore is allowing us to obtain complete genomes of bacteria rapidly and in a cost-effective manner. We were successful in using both short-read sequence data from Illumina and long-read data from Oxford Nanopore to carry out assembly. This enabled us to sequence the complete genome and detect two plasmids in this strain. The Xoo genome is particularly rich in repetitive elements such as TAL genes and IS elements.

Fig. 1. Circular representation of the BXO1 genome. The rings represent (from outside to inside): (i) protein-coding genes on the forward strand (red), (ii) protein-coding genes on the reverse strand (blue), (iii) tRNA genes (purple), (iv) transcriptional activator-like effectors (TALEs) (green), (v) GC content and (vi) GC skew (distribution of guanine and cytosine nucleotides in the genome).
To the best of our knowledge, the present study is the first report on the use of Oxford Nanopore Technology (MinION) in successfully obtaining the complete genome of a Xoo strain. Our study has unravelled an important facet of inter-strain variation mediated by novel plasmids harbouring novel gene clusters of importance in host–microbe interactions, including the T4SS. The presence of two novel plasmids in one strain highlights the importance of strain-specific genome dynamics mediated by such elements. In one of the plasmids, a large number of clustered hypothetical genes are present and this suggests the importance of plasmids in providing Xoo strains with novel functions. The use of such cost-effective long-read sequencing platforms can enable researchers to provide comprehensive insights into inter-strain and inter-lineage genome dynamics.

METHODS
DNA extraction and nanopore sequencing
BXO1 culture was grown in peptone sucrose (PS) broth media at 28°C for 48 h. The cells were harvested and genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen). The quantity and quality of DNA was assessed using Nanodrop and the Qubit 2.0 Fluorometer. Library preparation was performed using the Ligation Sequencing Kit 1D (SQK-LSK108). Initially, 4μg of genomic DNA was sheared using Covaris g-TUBE (Covaris, Brighton, UK). Then, the DNA end prep step was performed using NEBNext Ultra II End-repair/dA-Tailing modules. All bead washing steps were performed using AMPure beads (Beckman Coulter). Further, native barcoding and adaptor ligation steps were performed as per the protocols given by Oxford Nanopore Technologies. Finally, 12μl of prepared library DNA was used and sequenced using the MinION (FLO-MIN-106 vR9.4) flow cell with MinKNOW software (v1.13.1) (http://community.nanoporetech.com; Oxford Nanopore Technologies) for 48 h. Nanopore raw FAST5 reads were base-called and converted to the FASTQ format using Albacore v2.3.1 software (http://community.nanoporetech.com).

Genome assembly and annotation
The reads obtained after demultiplexing were assembled using Unicycler v0.4.4 [19] with the conservative mode. The assembled genomes were then error-corrected for multiple rounds with short reads generated by Illumina using Pilon v1.22 [20]. The assembled genome was then checked for completeness and the presence of contamination using CheckM v1.0.11 [21]. Genome coverage was checked using the BBMap v38.20 tool [22]. The genome was submitted to the NCBI Whole Genome Shotgun (WGS) portal with the accession numbers CP033201, CP033202 and CP033203 and annotated using the NCBI’s Prokaryotic Genome Annotation Pipeline (PGAP; https://www.ncbi.nlm.nih.gov/genome/annotation_prok/).
Identification of mobile elements

The genes were represented in circular form and the GC content of the genes in chromosome and plasmids was visualized using DNA Plotter [23]. Hypothetical genes were further validated using blastp [24]. IS elements were identified using ISSaga [25]. In order to look for TAL effectors, TAL DNA sequences were retrieved from http://www.xanthomonas.org/experts.html and a BLASTN search was performed against the sequenced genome. We also validated our TAL gene sequences using AnnoTALE software [26]. Comparison of pBXO1-2 and GPE PC73 plasmids was generated using blast ring image generator.

Funding information
A. K. is supported by a DST-INSPIRE fellowship. This work is supported by a project entitled ‘High throughput and integrative genomic approaches to understand adaptation of probiotic and pathogenic bacterium’ (OLP148) of the Council of Scientific and Industrial Research (CSIR) to P.B.P. R.V. S. was also supported by a J. C. Bose fellowship from the Science and Engineering Research Board, Government of India.

Author contributions
A. K., K. B. and S. K. carried out the complete genome sequencing and NCBI submission. A. K. performed downstream analysis with the help of S. K. A. K. and K. B. drafted the manuscript with input from R. V. S. and P. B. P. P. P. participated in its design and the interpretation of data with all the authors and applied for funding. All authors have read and approved the manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References
1. Mishra D, Vishnupriya MR, Anil MG, Konda K, Raj Y et al. Pathotype and genetic diversity amongst Indian isolates of Xanthomonas oryzae pv. oryzae. PLoS One 2013;8:e81996.
2. Patil PB, Bogdanove AJ, Sonti RV. The role of horizontal transfer in the evolution of a highly variable lipopolysaccharide biosynthesis locus in xanthomonads that infect rice, citrus and crucifers. BMC Evol Biol 2007;7:243.
3. Midha S, Bansal K, Kumar S, Girija AM, Mishra D et al. Population genomic insights into variation and evolution of Xanthomonas oryzae pv. oryzae. Sci Rep 2017;7:40694.
4. Gochez AM, Huguet-Tapia JC, Minsavage GV, Shantaraj D, Jalan N et al. Pacbio sequencing of copper-tolerant Xanthomonas citri reveals presence of a chimeric plasmid structure and provides insights into reassembly and shuffling of transcription activator-like effectors among X. citri strains. BMC Genomics 2018;19:16.
5. Álvarez-Alvarez R, Rodriguez-Garcia A, Martinez-Burgo Y, Robles-Reglero V, Santamarta I et al. A 1.8-Mb-reduced Streptomyces clavuligerus genome: relevance for secondary metabolism and differentiation. Appl Microbiol Biotechnol 2014;98:2183–2195.
6. Bansal K, Kumar S, Patil PB. Complete genome sequence reveals evolutionary dynamics of an emerging and variant pathovar of Xanthomonas euvesicatoria. Genome Biol Evol 2018;10:3104–3109.
7. Nakanoaka K, Ogura Y, Toyoda A, Yoshida M, Fukano H et al. Naturally occurring a loss of a giant plasmid from Mycobacterium ulcerans subsp. shinshuense makes it non-pathogenic. Sci Rep 2018;8:2818.
8. Li R, Xie M, Dong N, Lin D, Yang X et al. Efficient generation of complete sequences of MDR-encoding plasmids by rapid assembly of MinION barcoding sequencing data. Gigascience 2018;7:gix132.
9. Carpenter SCD, Mishra P, Ghoshal C, Dash PK, Wang L et al. A Strain of an Emerging Indian Xanthomonas oryzae pv. oryzae Pathotype Defeats the Rice Bacterial Blight Resistance Gene xa13 Without Inducing a Clade III SWEET Gene and Is Nearly Identical to a Recent Thai Isolate. Front Microbiol 2018;9:2703.
10. Pieretti I, Royer M, Barbe V, Carrere S, Koebnik R et al. The complete genome sequence of Xanthomonas albilineans provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. BMC Genomics 2009;10:616.
11. MacLean RC, San Millan A. Microbial evolution: towards resolving the plasmid paradox. Curr Biol 2015;25:R764–R767.
12. Grohmann E, Christie PJ, Waksman G, Backert S. Type IV secretion in gram-negative and Gram-positive bacteria. Mol Microbiol 2018;107:455–471.
13. Souza DP, Oka GU, Alvarez-Martinez CE, Bisson-Filho AW, Danger G et al. Bacterial killing via a type IV secretion system. Nat Commun 2015;6:6453.
14. Engledow AS, Medrano EG, Mahenthiralingam E, LiPuma JJ, Gonzalez CF et al. Involvement of a plasmid-encoded type IV secretion system in the plant tissue water soaking phenotype of Burkholderia cenocepacia. J Bacteriol 2004;186:6015–6024.
15. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR et al. Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. Nature 2002;417:459–463.
16. Hauben Let al. Comparison of 16S ribosomal DNA sequences of all Xanthomonas species. International Journal of Systematic and Evolutionary Microbiology 1997;47:328–335.
17. Rademaker JL, Hoste B, Louws FJ, Kersters K, Swings J et al. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: Xanthomonas as a model system. Int J Syst Evol Microbiol 2000;50:665–677.
18. Parkinson N, Cowie C, Heeney J, Stead D. Phylogenetic structure of Xanthomonas determined by comparison of gyrB sequences. Int J Syst Evol Microbiol 2009;59:264–274.
19. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017;13:e1005595.
20. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 2014;9:e112963.
21. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW et al. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015;25:1043–1055.
22. Bushnell B. BBMap: a fast, accurate, splice-aware aligner 2014.
23. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J et al. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–120.
24. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S et al. Ncbi blast: a better web interface. Nucleic Acids Res 2008;36:W5–W9.
25. Varani AM, Siguer P, Gourbeyre E, Charneau V, Chandler M et al. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. Genome Biol 2011;12:R30.
26. Grau J, Reschke M, Erkes A, Streubel J, Morgan RD et al. AnnoTALE: bioinformatics tools for identification, annotation, and nomenclature of tales from Xanthomonas genomic sequences. Sci Rep 2016;6:21077.

Five reasons to publish your next article with a Microbiology Society journal
1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.