MALAYSIAN JOURNAL OF BIOCHEMISTRY & MOLECULAR BIOLOGY

The Official Publication of The Malaysian Society For Biochemistry & Molecular Biology (MSBMB)
http://mjbmb.org

COMPARATIVE GENOMICS OF *Xanthomonas* BACTERIA: INSIGHTS INTO ITS PLANT PATHOGENICITY

Nik Yusnoraini Yusof[1], Muggunna Balasubramaniam[2], Anis Fathima Abdul Kareem[2] and Amira Suriaty binti Yaakop[2]*

[1]*Institute for Research in Molecular Medicine, Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia*
[2]*School of Biological Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia*

*Corresponding Author: amirasuriaty@usm.my

History
Received: 20 July 2022
Accepted: 8 August 2022

Keywords:

Xanthomonas; Virulence and pathogenicity; Bacterial secretion system; Secondary metabolites; Mobile genetic elements

Abstract

Xanthomonas is a group of Gram-negative bacteria from class Gammaproteobacteria that causes multiple diseases in many plant hosts. Bacterial secretion systems of *Xanthomonas* contribute to its pathogenicity and are one of the most important factors for symptom and disease development. Comparative genomic analysis between 20 *Xanthomonas* bacteria, well known for their pathogenicity was chosen to study the bacterial secretion system, specifically, the type III bacterial secretion system associated with bacterial pathogenesis. The analysis was performed using databases and software including NCBI, RAST, antiSMASH, KEGG and Islandviewer4 to compare the biological and taxonomical relationship between these genomes. The average nucleotide identity (ANI) similarity index ranges between 79% to 99.96%, indicating that the selected genomes were closely related. Screening of secondary metabolites using antiSMASH showed all 20 genomes produced secondary metabolites. Siderophore secondary metabolite gene cluster necessary for optimum virulence was present in all the genomes. Genomic island discovery using Islandviewer4 revealed the presence of mobile genetic elements such as phages and transposons, indicating horizontal gene transfer events. Genes and pathogenicity pathway mechanism identification of type III bacterial secretion system was done using SEED subsystem and KEGG database. A total of 34 genes associated with type III bacterial secretion system were identified. *HrpA*, *HrpB*, *HrpX* and *HrpG* were the conserved genes across the *Xanthomonas* species. The *HrpX* and *HrpG* genes are essential for symptom development, disease establishment and invasion into host cells.

INTRODUCTION

Plant pathogenic bacteria can be found mostly within the families of *Xanthomonadaceae*, *Pseudomonaceae*, and *Enterobacteriaceae* [1]. These bacterial families consist of several genera, namely *Dickeya, Liberibacter, Erwinia, Pectobacterium, Candidatus, Pantoea, Agrobacterium, Pseudomonas, Ralstonia, Burkholderia, Acidovorax, Xanthomonas, Clavibacter, Streptomyces, Xylella, Spiroplasma, Phytoplasma, Brenneria, Lonsdale*, and *Xylophilus* [1]. *Xanthomonas* is one of the most extensively studied genera and is known to infect and cause diseases in more than 400 plant hosts, classified as agriculturally essential food crops [2].

Xanthomonas or xanthomonads are short, straight-rod-shaped Gram-negative bacteria [3]. The growth of these species on a nutrient agar plate can be observed as distinct yellow-pigmented colonies [3]. Xanthomonadins are carotenoid-like, brominated, aryl-polyene esters associated with the outer membrane of the cell wall [4]. The presence of xanthomonadin contributes to yellow pigmentation, although not all strains can produce this pigment [2]. In *X. campestris* pv. *vesicatoria*, xanthomonadins protect phytopathogenic genus *Xanthomonas* against damage by

visible light in the presence of oxygen [5]. More than 35 species are currently classified in this genus and subdivided into subspecies or pathovars [2]. General symptoms of plants infected by *Xanthomonas* species include leaf spots, fruit spots, blights, vascular wilt and bacterial canker [6].

The advancement in bioinformatics, particularly Next Generation Sequencing (NGS) and functional genomics has led to the discovery and identification of novel genes that greatly contribute to understanding diversity, virulence, and plant-pathogen interactions [2]. Previous studies have reported that the bacterial secretion systems are one of the most important factors contributing to the pathogenicity of *Xanthomonas*. However, very few studies have been conducted that correlate pathogenicity factors to symptom and disease development. In this study, 20 *Xanthomonas* genomes known for pathogenic properties and to infect various plant hosts were chosen for comparative genomic analysis. Genomic analysis of *Xanthomonas* will help to expand our knowledge to correlate pathogenicity factors associated with bacterial secretion systems to host range, symptom development and evolution. Furthermore, identifying the underlying causes of pathogenicity provides insights into the characteristics and types of virulence genes and aids in developing novel infection management strategies for diseases caused by *Xanthomonas* species.

MATERIALS AND METHODS

Data Retrieval Information

A total of 20 complete sequences of *Xanthomonas* bacteria and one *Escherichia coli* (*E. coli*) genome as an outgroup were chosen from National Centre for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/genome/). The selected 20 genomes include pathovars and isolates of numerous strains from *Xanthomonas euvesicatoria*, *Xanthomonas vesicatoria*, *Xanthomonas perforans*, *Xanthomonas campestris*, *Xanthomonas citri*, *Xanthomonas vasicola*, *Xanthomonas oryzae*, *Xanthomonas hortorum*, *Xanthomonas fragariae*, *Xanthomonas cucurbitae*, *Xanthomonas euroxanthea*, and *Xanthomonas theicola* (Table 1). Only completely sequenced genomes were selected. The sequences of 20 genomes were retrieved in FASTA format file. Selected 20 *Xanthomonas* genomes are well-known for its pathogenicity.

Average Nucleotide Identity

Detailed information of each genome was obtained from IMG/ER and NCBI database. OrthoANIu (https://www.ezbiocloud.net/tools/ani), a web service within the EZbiocloud database was used to calculate average nucleotide identity (ANI) between pair of genome sequences [7]. The general features and ANI were summarised in respective Table 2. Obtained information was used to study the differences and similarities between the *Xanthomonas* genomes.

Genome Annotation

Genome annotation of the selected 20 genomes was done to identify the presence of different types of pathogenicity factors. The selected 20 Xanthomonas genomes were annotated using the Rapid Annotation using Subsystem Technology (RAST, https://rast.nmpdr.org/rast.cgi) [8]. In addition, SEED subsystem was used to retrieve genome annotation feature files in excel format for further screening of pathogenicity and virulence genes [9].

Screening of Secondary Metabolite

Online webserver of antibiotics and secondary metabolite analysis shell – antiSMASH (https://antismash.secondarymetabolites.org/), updated version 6 of antiSMASH, was used to screen for presence of secondary metabolites produced by the 20 genomes [10].

Metabolic and Pathogenicity Pathway Mechanism

General metabolic pathways of gene regulation and pathogenicity mechanism of the selected 20 Xanthomonas genomes were built by comparison to Kyoto Encyclopaedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/genome/) to identify genes and proteins involved in the contribution of pathogenicity, specifically bacterial secretion systems [11].

Genomic Island Discovery

Genomic islands of *Xanthomonas, E. coli* and *Lactobacillus acidophilus* were predicted using Islandviewer4 (https://www.pathogenomics.sfu.ca/islandviewer/) to study the presence of pathogenicity islands [12].

RESULTS AND DISCUSSION

Xanthomonas genomes were selected based on the availability of literature and how extensively a specific *Xanthomonas* species or strain is studied. For comparative genomic analysis, a total of 20 complete *Xanthomonas* genomes of different species and strains were selected. Complete genomes were chosen to minimise error during further analysis. Factors such as the number of contigs, sequencing technology, assembly level, literature availability, and research findings are considered during genome selection. Out of 20 genomes, four genomes, *X. theicola* CFBP 4691 [13], *X. euroxanthea* CPBF 426 [14], *X. cucurbitae* ATCC 23378 [15], and *X. hortorum* B007-007 [16] have fewer literature available compared to others (Table 1). In addition, ten genomes were categorised as reference genomes by NCBI. Complete bacterial sequences

assembled by more than one sequencing technology are considered during genome selection of the present study as they produce genomes of greater quality and fewer contigs [17].

The ANI values between the selected 20 *Xanthomonas* genomes are presented in Table 2. OrthoANIu from the EzGenome database was used to calculate ANI as it provides insights into the taxonomical relationship between species [18]. ANI is a computational analysis widely used to define both archaeal and bacterial species boundaries [19]. OrthoAniU by EzGenome database was chosen as the ANI calculator because it calculates ANI relatively faster than other known tools and it is efficient in large-scale comparative genomic analyses [19]. OrthoANIu was also chosen based on selected strains of the present study which belonged to the same genera. Findings from a study showed that OrthoANIu did not provide good results when compared with genomes from different genera [20]. Results indicate that ANI values between the 20 genomes ranged between 78.90% and 99.96%. The lowest ANI values ranged between 78.99% to 80.27% when comparing *X. theicola* CFBP 4691 with the rest of the *Xanthomonas* species strains indicating that it is the distantly related species among the selected strains. The results are supported by phylogenetic analyses of 16s rRNA gene shown in Figure 1.

Table 1. The 20 *Xanthomonas* genomes selected for comparative genomic analysis with the detail of accession, assembly level and method

Strain	Accession number	Assembly level	Assembly method	Reference
X. euvesicatoria pv. alfalfae strain CFBP3836	NZ_CP072268	Complete	HGAP PreAssembler Filter v. v1, Canu v. v1.5, Berokka v. v0.2.3, Circlator v. v1.5.1, variantCaller v. v2.2.2, Pilon v. v1.23	[45] [46]
X. campestris pv. vesicatoria str 85-10	NZ_CP017190	Complete	HGAP v. 2	
X. euvesicatoria strain LMG930	NZ_CP018467	Complete	SMRT Analysis HGAP v. 2.3	
X. vesicatoria ATCC35937 strain LMG911	NZ_CP018725	Complete	SMRT Analysis HGAP protocol v. 2.3	
X. vesicatoria strain LM159	NZ_CP018470	Complete	SMRT Analysis HGAP v. 2.3	
X. campestris pv. raphani strain MAFF106181	NZ_CP058243	Complete	Celera Assembler v. 8.3	
X. campestris pv. campestris str. ATCC 33913	NC_003902	Complete	NA	
X. campestris pv. campestris str. 8004	NC_007086	Complete	NA	
X. perforans 91-118	NZ_CP019725	Complete	NA	[47] [48]
X. perforans strain LH3	NZ_CP018475	Complete	SMRT Analysis HGAP v. 2.3	
X. citri strain UnB-Xtec2D	NZ_CP048044	Complete	Unicycler v. 0.4.7	[49] [50] [51] [52]
X. vasicola pv. vasculorum strain SAM119	NZ_CP028127	Complete	SPAdes v. 3.10.1 Pilon v. 1.22	[53]
X. vasicola strain NCPPB 1060	NZ_CP034649	Complete	HGAP v. 4 Pilon v. 1.22	
X. oryzae pv. oryzicola strain YM15	NZ_CP007810	Complete	FASTX-Toolkit Velvet sspace v. 2014.4	[45] [54] [55]
X. oryzae pv, oryzae strain DY89031 (J18)	NZ_CP064780	Complete	SPAdes v. 3.5.0	
X. hortorum strain B007 – 007	NZ_CP016878	Complete	CLC Genomics Workbench v.7.0.4 Celera v.CA 8	[16] [16]
X. fragariae strain PD885	NZ_LT853882	Complete	NA	[56]
X. cucurbitae strain ATCC 23378	NZ_CP033326	Complete	SPAdes v. June-2018	[15] [57]
X. euroxanthea isolate *Xanthomonas sp.* CPBF 426	NZ_LR824639	Complete	NA	[58] [14]
X. theicola strain CFBP 4691	NZ_CP049017	Complete	HGAP v. 4.0	[13]
E. coli K-12 sub strain MG1655	NC_000913	Complete	NA	
Lactobacillus acidophilus La-14	NC_021181	Complete	SeqMan NGen v. 4.1.2	

Table 2. ANI Similarity Index of *Xanthomonas* Strains

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
A	-	98.56	98.56	85.94	86.21	98.69	98.68	85.07	85.08	85.18	93.82	89.89	90.03	90.13	90.05	86.52	86.03	85.91	87.03	79.41
B	98.56	-	99.96	86.09	86.07	98.61	98.58	84.94	85.15	85.17	93.9	89.89	90.05	89.99	90.02	86.60	85.90	85.75	86.86	79.50
C	98.56	99.96	-	85.95	86.08	98.54	98.57	85.05	85.16	85.28	93.84	89.88	90.03	90.09	90.07	86.67	86.00	85.72	87.12	79.47
D	85.94	86.09	85.95	-	98.75	86.08	86.15	85.26	85.00	85.24	85.98	85.55	85.80	85.97	85.82	86.29	85.68	86.61	86.83	79.20
E	86.21	86.07	86.08	98.75	-	86.03	86.23	85.12	85.07	85.22	86.08	85.80	85.90	85.90	86.00	86.40	85.66	86.62	86.68	79.20
F	98.69	98.61	98.54	86.08	86.03	-	99.98	85.13	85.21	85.18	93.88	89.86	90.09	90.12	90.17	86.46	85.96	85.90	87.01	79.37
G	98.68	98.58	98.57	86.15	86.23	99.98	-	85.13	85.13	85.28	93.92	90.00	90.11	90.04	90.15	86.46	86.07	85.87	87.13	79.56
H	85.07	84.94	85.05	85.26	85.12	85.13	85.13	-	97.22	97.10	85.00	84.81	84.67	84.77	84.97	85.56	84.77	85.04	86.05	79.51
I	85.08	85.15	85.16	85	85.07	85.21	85.13	97.20	-	99.85	84.87	84.85	84.89	84.84	84.85	85.88	84.93	85.17	86.16	79.55
J	85.07	84.94	85.05	85.24	85.22	85.18	85.28	97.10	99.85	-	84.85	84.79	84.98	84.75	84.95	86.10	84.96	85.22	86.05	79.82
K	93.82	93.90	93.84	85.98	86.08	93.88	93.92	85.00	84.87	84.85	-	89.92	89.98	89.86	89.84	86.42	85.97	85.74	86.93	79.31
L	89.89	89.89	89.88	85.55	85.80	89.86	90.00	84.81	84.85	84.79	89.92	-	98.75	91.11	91.35	86.30	85.90	85.37	86.48	78.90
M	90.03	90.05	90.03	85.8	85.90	90.09	90.11	84.67	84.89	84.98	89.98	98.75	-	91.26	91.40	86.42	85.83	85.34	86.56	79.01
N	90.13	89.99	90.09	85.97	85.90	90.12	90.04	84.77	84.84	84.75	89.86	91.11	91.26	-.	97.84	86.26	86.03	85.43	86.56	78.99
O	90.05	90.02	90.07	85.89	86.00	90.17	90.15	84.97	84.85	84.95	89.84	91.35	91.40	97.84	-	86.43	86.11	85.46	86.73	79.11
P	86.52	86.60	86.67	86.29	86.40	86.46	86.46	85.56	85.88	86.10	86.42	86.3	86.42	86.26	86.43	-	87.35	85.60	88.97	79.41
Q	86.03	85.90	86.00	85.68	85.66	85.96	86.07	84.77	84.93	84.96	85.97	85.91	85.83	86.03	86.11	87.35	-	84.81	87.11	78.91
R	85.91	85.75	85.72	86.61	86.62	85.90	85.87	85.04	85.17	85.22	85.74	85.37	85.44	85.43	85.46	85.60	84.81	-	86.44	79.64
S	87.03	86.86	87.12	86.63	86.68	87.01	87.13	86.05	86.16	86.05	86.93	86.48	86.56	86.56	86.73	88.97	87.11	86.44	-	80.27
T	79.41	79.50	79.47	79.20	79.20	79.37	79.56	79.51	79.55	79.82	79.31	78.90	79.01	78.99	79.11	79.41	78.91	79.64	80.27	-

A = *X. euvesicatoria* pv. *alfalfae* strain CFBP3836, B = *X. campestris* pv. *vesicatoria* 85-10, C = *X. euvesicatoria* LMG930, D = *X. vesicatoria* ATCC 35937 strain LMG911, E = *X. vesicatoria* LM159, F = *X. perforans* 91-118, G = *X. perforans* LH3, H = *X. campestris* pv *raphani* MAFF 106181, I = *X. campestris* pv. *campestri*s ATCC 33913, J = *X. campestris* pv. *campestri*s 8004, K = = *X. citri* UnB-Xtec2D, L = *X. vasicola* pv. *vasculorum* SAM 119, M = *X. vasicola* NCPPB 1060, N = *X. oryzae* pv. *oryzicola* YM15, O = *X. oryzae* pv. *oryzae* DY89031 (J18), P = *X. hortorum* B007-007, Q = *X. fragariae* PD885, R= *X. cucurbitae* ATCC 23378, S = *X. euroxanthea* CPBF 426, T = *X. theicola* strain CFBP 4691

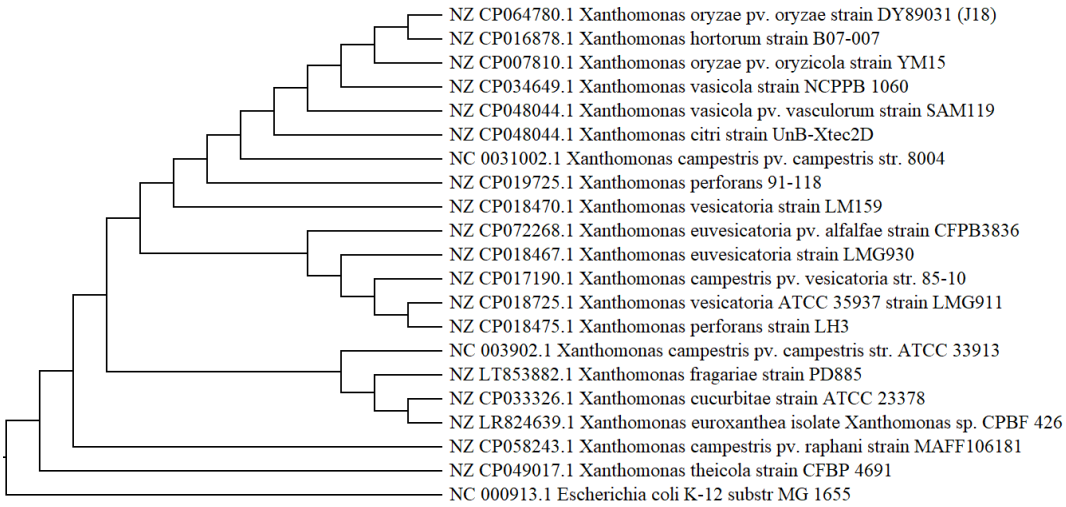

Figure 1. Phylogenetic Tree of 16s rRNA from Chosen *Xanthomonas* strains with *E. coli* K-12 sub strain MG1655

Genome Annotation

To further facilitate comparative genomic analysis, SEED subsystem by RAST server was used to predict functional genes. The gene distribution of *Xanthomonas* species based on the SEED annotation feature are is presented in Table 3. Based on the SEED subsystem, the largest functional group for all the species is amino acid derivatives with coding DNA sequences (CDS) count ranging from 263 to 284, followed by carbohydrates with CDS count ranging from 189 to 222, and membrane transport with CDS count ranging from 128 to 197. Moreover, functional genes clustered under protein metabolism and functional group consisting of cofactors, vitamins, and prosthetic groups also account for most of the CDS count across all the 20 *Xanthomonas* strains, ranging between 156 to 187 and 127 to 185, respectively.

The functional group of virulence, defence and disease is subdivided into several other groups. These groups include adhesions, toxins and superantigens, bacteriocins, resistance to antibiotics and toxic compounds, and lastly invasion and intracellular resistance. Among the subdivided groups, only two groups, namely antibiotics and toxic compounds, and invasion and intracellular resistance, were recorded to have gene clusters present in the SEED subsystem. The subcategories mercuric reductase and beta-lactamase are species-specific and categorised in subdivided groups of resistance to antibiotics and toxic compounds. The subcategory mercuric reductase was only found in *Xanthomonas campestris* pathovars while the subcategory beta lactamase was only found in *X. oryzae* pv. *oryzae* DY89031 (J18) and *X. theicola* CFBP 4691.

The second subdivided group in virulence, defence and disease functional group with gene clusters is the invasion and intracellular resistance which has four subcategories such as mycobacterium virulence operon involved in the protein synthesis (SSU ribosomal proteins and LSU ribosomal proteins), mycobacterium virulence operon involved in DNA transcription, and listeria surface proteins. Mycobacterium virulence operon of DNA transcription and protein synthesis involving LSU ribosomal proteins contained the same gene count in all the selected strains. All the species have 5 genes involving involved in mycobacterium protein synthesis virulence operon involving SSU ribosomal protein except for the *X. oryzae* pv. *oryzicola* YM15.

Secondary Metabolite of *Xanthomonas* Species

Microbial secondary metabolites are small organic molecules that play crucial roles in the microbial cell culture survival and ecological interaction with another organism [21]. They are involved in cellular activities such as signalling, regulation, nutrient scavenging and self-preservation [22]. Secondary metabolites produced by any bacteria, fungi, or plants are of great importance to humans as they are manipulated for the benefit of humankind. List of secondary metabolites identified in *Xanthomonas* genomes are summarised in Table 4.

Secondary metabolite gene clusters such as phosphonate and lassopeptide are species-specific. Phosphonate can only be found in *Xanthomonas vasicola* pathovars while lassopeptide are only found in *X. theicola* CFBP 4691, *X. euroxanthea* CPBF 426 and *X. hortorum* B007-007. Lankacidin C, xanthoferrin and xanthomonadin I are present in all the species while phosphonoacetic acid, cichopeptin, pseudopyronine A/pseudopyronine B, rhizomide A/ rhizomide B/ rhizomide C and dactlocycline A can only be found in Xanthomonas vasicola pathovars, *X. oryzae pv. oryzicola* YM 15, *X. oryzae pv. oryzae* DY89031 (J18), *X. cucurbitae* ATCC 23378 and *X. theicola* CFBP 4691, respectively.

Table 3. Gene Distribution of *Xanthomonas* species in different categories based on SEED subsystem

Subsystem Information	Respiration	Stress Responses	Carbohydrates	Photosynthesis	DNA Metabolism	RNA Metabolism	Miscellaneous	Potassium Metabolism	Sulphur Metabolism	Phosphorus Metabolism	Protein Metabolism	Secondary Metabolism	Cofactors, Vitamins, Prosthetic Groups, Pigments	Virulence, Disease, Defense	Phages, Prohages, Transposable Elements Plasmids
***Xanthomonas euvesicatoria* pv. *alfalfae* strain CFBP3836**	102	75	204	0	90	41	29	10	17	24	179	8	167	22	1
***Xanthomonas campestris* pv. *vesicatoria* strain 85-10**	101	78	206	0	92	41	29	11	18	22	184	8	167	24	8
***Xanthomonas euvesicatoria* strain LMG 930**	99	76	204	0	91	41	29	11	17	22	182	8	165	23	2
***Xanthomonas vesicatoria* ATCC 35937 strain LMG 911**	98	78	208	0	84	42	29	10	17	23	182	8	172	20	2
***Xanthomonas vesicatoria* strain LM 159**	100	80	209	0	90	41	29	10	17	23	181	8	168	20	1
***Xanthomonas perforans* 91-118**	100	75	204	0	84	41	29	10	17	22	180	8	167	22	2
***Xanthomonas perforans* strain LH3**	100	75	204	0	84	41	29	10	17	22	180	8	167	22	5
***Xanthomonas campestris* pv. *raphani* strain MAFF 106181**	99	77	215	0	93	41	30	10	17	22	179	8	175	22	0
***Xanthomonas campestris* pv. *campestris* ATCC 33913**	99	78	222	0	91	41	31	10	17	22	180	8	171	22	0
***Xanthomonas campestris* pv. *campestris* strain 8004**	99	78	224	0	97	42	30	10	17	22	156	8	171	22	0
***Xanthomonas citri* strain UnB-Xtec2D**	99	74	217	0	84	38	28	10	16	22	179	8	170	22	1
***Xanthomonas vasicola* pv. *vasculorum* strain SAM 119**	107	81	205	0	89	43	23	10	16	27	181	8	141	22	0
***Xanthomonas vasicola* strain NCPPB 1060**	101	66	200	0	86	43	21	10	17	27	179	8	139	21	3
***Xanthomonas oryzae* pv. *oryzicola* strain YM 15**	98	64	180	0	92	40	31	10	15	22	169	8	153	22	0
***Xanthomonas oryzae* pv. *oryzae* DY89031 (J18)**	98	70	185	0	86	39	28	10	18	23	176	8	160	24	1
***Xanthomonas hortorum* strain B07-007**	91	78	204	0	87	39	28	10	16	22	177	8	129	21	0
***Xanthomonas fragariae* strain PD885**	81	66	170	0	77	39	21	8	17	25	175	8	127	24	2
***Xanthomonas cucurbitae* strain ATCC 23378**	90	70	197	0	92	37	18	9	16	27	184	8	131	23	10
***Xanthomonas euroxanthea* CPBF 426**	93	84	208	0	81	37	29	10	17	23	187	8	132	21	0
***Xanthomonas theicola* strain CFBP 4691**	94	59	189	0	79	40	17	10	16	27	175	8	131	22	1

Table 3. Gene Distribution of *Xanthomonas* species based on SEED subsystem (continued)

Subsystem Information	Membrane Transport	Nitrogen Metabolism	Cell Wall And Capsules	Cell Division and Cell Cycle	Motility And Chemotaxis	Regulation and Cell Signalling	Fatty Acids, Lipids, and Isoprenodis	Dormance and Sporulation	Iron Acquistion and Metabolism	Nucleosides and Nucleotides	Metabolism In Aromatic Compounds	Amino Acid Derivatives
***Xanthomonas euvesicatoria* pv. *alfalfae* strain CFBP3836**	169	6	25	0	19	19	68	1	6	58	29	270
***Xanthomonas campestris* pv. *vesicatoria* strain 85-10**	157	6	24	0	19	20	71	1	6	57	30	274
***Xanthomonas euvesicatoria* strain LMG 930**	178	6	25	0	19	20	69	1	6	58	30	272
***Xanthomonas vesicatoria* ATCC 35937 strain LMG 911**	169	6	24	0	18	21	70	1	6	57	29	270
***Xanthomonas vesicatoria* strain LM 159**	177	6	24	0	19	21	72	1	6	57	29	271
***Xanthomonas perforans* 91-118**	171	6	24	0	19	19	68	1	6	56	29	272
***Xanthomonas perforans* strain LH3**	174	6	25	0	19	19	68	1	6	56	31	272
***Xanthomonas campestris* pv. *raphani* strain MAFF 106181**	175	6	29	0	18	21	67	1	6	56	32	276
***Xanthomonas campestris* pv. *campestris* ATCC 33913**	188	6	28	0	18	20	68	1	6	56	33	278
***Xanthomonas campestris* pv. *campestris* strain 8004**	190	6	28	0	18	20	68	1	6	56	31	277
***Xanthomonas citri* strain UnB-Xtec2D**	182	6	24	0	18	26	69	1	6	57	30	272
***Xanthomonas vasicola* pv. *vasculorum* strain SAM 119**	160	6	26	0	21	29	68	1	6	56	31	274
***Xanthomonas vasicola* strain NCPPB 1060**	161	6	26	0	22	26	70	1	6	56	29	274
***Xanthomonas oryzae* pv. *oryzicola* strain YM 15**	133	6	23	0	20	27	62	1	6	56	29	272
***Xanthomonas oryzae* pv. *oryzae* DY89031 (J18)**	132	6	23	0	20	26	48	1	6	54	28	284
***Xanthomonas hortorum* strain B07-007**	187	6	28	0	18	25	64	1	16	56	30	279
***Xanthomonas fragariae* strain PD885**	138	6	26	0	20	12	47	1	6	51	12	278
***Xanthomonas cucurbitae* strain ATCC 23378**	197	7	28	0	18	20	68	1	14	52	33	263
***Xanthomonas euroxanthea* CPBF 426**	149	6	29	0	18	23	68	1	6	55	38	276
***Xanthomonas theicola* CFBP 4691**	128	8	24	0	21	18	91	1	12	53	26	265

Table 4. List of Secondary Metabolites of Xanthomonas Genomes

***Xanthomonas* species**	**Type**								**Most Similar Known Cluster**									
	Redox-cofactor	**siderophore**	**NRPS**	**Arylpolyene**	**RiPP-like**	**RRE-containing**	**Lassopeptide**	**Phosphonate**	**Lankacidin C**	**Xanthoferrin**	**Xanthomonadin I**	**Cichopeptin**	**phosphonoacetic acid**	**Pseudopyronine A/ Pseudopyronine B**	**Rhizomide A/ Rhizomide B/ Rhizomide C**	**Xanthomonin I/ Xanthomonin II**	**Dactylocycline A**
Xanthomonas euvesciatoria pv. *alfalfae* CFBP 3836	/	/	/	/					/	/	/						
Xanthomonas campestris pv. *vesicatoria 85-10*	/	/	/	/	/				/	/	/						
Xanthomonas euvesciatoria LMG 930	/	/	/	/					/	/	/						
Xanthomonas vesicatoria LMG 911		/	/	/	/	/			/	/	/						
Xanthomonas vescicatoria LM 159	/	/	/	/	/				/	/	/						
Xanthomonas perforans 91-118	/	/	/	/					/	/	/						
Xanthomonas perforans LH3	/	/	/	/					/	/	/						
Xanthomonas campestris pv. *raphani* MAFF 106181	/	/	/	/			/		/	/	/						
Xanthomonas campestris pv. *campestris* ATCC 33913	/	/		/						/	/						
Xanthomonas campestris pv. *campestris strain 8004*	/	/		/						/	/						
Xanthomonas citri UnB-Xtec2D	/	/		/					/	/	/						
Xanthomonas vasicola pv. *vasculorum* SAM119	/	/	/	/	/			/	/	/	/		/				
Xanthomonas vasicola NCPPB 1060	/	/	/	/				/	/	/	/		/				
Xanthomonas oryzae pv. *oryzicola* YM15		/	/	/						/	/	/					
Xanthomonas oryzae pv. *oryzae* DY89031 (J18)	/	/		/					/	/				/			
Xanthomonas hortorum B007-007	/	/	/	/			/		/	/	/					/	
Xanthomonas fragariae PD 885	/	/	/						/	/							
Xanthomonas cucurbitae ATCC 23378	/	/	/	/	/				/	/	/				/		
Xanthomonas euroxanthea CPBF 426		/		/			/			/	/					/	
Xanthomonas theicola CFBP4691	/	/	/	/	/	/	/		/	/							/

Secondary metabolite gene cluster of redox cofactors which was only absent in *X. euvesicatoria* LMG 911, *X. euroxanthea* CPBF 426 and *X. oryzae pv. oryzae* DY89031 (J18) was always associated with lankacidin C with a similarity percentage of 13%. In addition, it was also observed that RRE-containing secondary metabolite in *X. vesicatoria* LMG 911 was also 13% similar to lankacidin C.

Siderophore was the only secondary metabolite gene cluster to be present in all the selected *Xanthomonas* genomes. It showed a 100% similarity with xanthoferrin in all the species except for *X. theicola* CFBP 4691 in which it only exhibited 85% similarity. Siderophore are characterized as low-molecular-weight metabolites with iron (III) chelating properties [23]. Siderophores secreted by bacteria aids in the acquisition of ferric ions during iron-limited conditions [24]. A study conducted in 2017 has reported that *X. campestris pv. campestris* which causes black rot in crucifers was found to secrete an α-hydroxy carboxylate type siderophore called xanthoferrin, necessary for optimum virulence and ferric iron uptake under low or restricted conditions [25]. It was also reported that mutation in tonB gene of *X. campestris pv. campestris* impaired ferric ion uptake, which resulted in elevated production of extracellular siderophore and reduced disease symptoms [25]. Moreover, identification of probable α-hydroxy carboxylate type siderophore (xanthoferrin) in *X. oryzae pv. oryzae* corresponds to the results obtained in the present study indicating the presence of xanthoferrin in *X. oryzae pv. oryzae* [24]. In addition, xanthomonads encode an xss (*Xanthomonas* siderophore synthesis) operon, homologous to the Vibrio parahaemolyticus siderophore (PVS) locus that produces a vibrioferrin-type siderophore under limited ferric ion conditions [25].

Conversely, subunits of non-ribosomal peptides synthetases (NRPS) was not related with any known similar cluster except for in *X. oryzae pv. oryzicola* YM15 in which NRPS was 46% similar to cichopeptin. NRPSs control the biosynthesis of nonribosomal peptides (NRP), the second major class of metabolites produced by Gram-negative bacteria [26]. Unfortunately, there was no evidence correlating the presence of cichopeptin with NRPS.

NRPS and RiPP-like secondary metabolite cluster finding in *X. cucurbitae* ATCC 23378 was 100% identical to rhizomide A/ rhizomide B/ rhizomide C gene cluster. RiPPs, an extensive group of secondary metabolites with wide structural diversity have been found in bacteria, fungi and archaea [26]. The statement of RiPPs is usually produced in Gram-negative bacteria correlates with the results obtained as *Xanthomonas* bacteria are Gram-negative bacteria. However, RiPPs are only found in six *Xanthomonas* genomes. The biosynthetic gene cluster of RiPP are composed of short precursor peptide of an N-terminal leader and a C-terminal core peptides and post-modification enzymes [27]. Currently, more than 13 representative types are classified. These classes include lanthipeptides, lasso peptides, linear azo containing peptides and thiopeptides [28].

The presence of lasso peptide, a class of RiPP secondary metabolite gene cluster, can be observed in four *Xanthomonas* genomes. Lasso peptides produced by bacteria are short bioactive compounds with a unique cyclic structure (lasso structure), hence the name lasso peptide [29]. Lassopeptide secondary metabolite found in *X. euroxanthea* CPBF 426 and *X. hortorum* B007-007 was 100% identical to xanthomonin I/ xanthomonin II. Xanthomonin I/ Xanthomonin II is a new class of lasso peptides characterized by macrolactam rings consisting of only seven amino acids [30]. Evidence states that lasso peptides produced by *Xanthomonas* are pathogenic, supporting results from antiSMASH, which shows that *Xanthomonas* species are indeed phytopathogenic. Moreover, it was also found in *X. theicola* CFBP 4691, but no similar known cluster was related to lassopeptide secondary metabolite.

Arylpolyene cluster found in almost of all of the genomes was mostly associated with xanthomonadin I with a similarity percentage of 64% and 71%. However, in X. theicola CFBP 4691, arylpolyene was associated with a distinct cluster called dactlocycline A which was only 5% similar to arylpolyene while in *X. oryzae pv. oryzae* DY89031 (J18) it was associated with pseudopyronine A/ pseudopyronine B with a similarity percentage of 12%. There was not much evidence associating arylpolyene secondary metabolite gene cluster to dactlocycline A that was present in *X. theicola* CFBP 4691. Similarly, findings correlating arylpolyene and pseudopyronine A/pseudopyronine B in *X. oryzae pv. oryzae* DY89031 (J18) were absent. However, pseudopyronine was isolated from *Pseudomonas* species [31].

Genomic Island of *Xanthomonas* Species

Genomics islands (GIs) are mobile DNAs or cluster of genes acquired by bacterial genomes by means horizontal gene transfer (HGT) [32]. HGT is the major driving force in the evolution of bacterial genomes as it enables bacteria to acquire foreign genetic material through conjugation, transduction, and transformation [33]. Predicted GIs of *Xanthomonas* genomes were compared with *E. coli* K-12 sub strain MG1655 and *L. acidophilus* La-14 genomes, as both are well studied. The current concept of GI, previously known as pathogenicity island,as initially studied by Hacker and colleagues using *E. coli* as the model organism [34]. On the other hand, *L. acidophilus* La-14 was chosen to differentiate genomic islands of non-pathogenic, Gram-positive bacteria with pathogenic, Gram-negative bacteria.

Based on the results obtained, it was observed that the GIs predicted in E. coli K-12 sub strain MG155 were prophages such as DLP12 phage, e14 phage, Rac phage, Qin phage, CP4-44 phage, CP4-57 phage, KpLE2 phage-like-element and its associated protein, insertions elements (IS) such as IS 2 and IS 5, and some putative proteins. Prediction

of GIs in *L. acidophilus* La-14 labelled in pink circular markers were homologs of resistance genes that include putative proteins, hypothetical proteins, ribosomal proteins, transfer RNA (tRNA) proteins, ATP-dependent proteins, and one transposase (Appendix A).

It was also observed that predicted GIs mostly included type IV secretion system (T4SS), ABC transporters, domain-containing proteins, different types of insertion sequences (IS), with IS3 and IS5 being the most common ones, toxins, pili and flagellum-associated proteins and hypothetical proteins. The GIs prediction of phages, phage-like proteins, insertion elements and transposases can be seen in both *Xanthomonas* and *E. coli* genomes. ABC transporters is one of the plant transporter families involved in membrane-trafficking and plant defence mechanism [35]. Additionally, the predicted transcription activator-like (TAL) effector proteins were related to T3SS in *X. euvesicatoria pv. alfalfae* CFBP 3836. TAL effectors serve as a host gene expression mediator either by binding to host resistance (R) genes which triggers resistance responses or susceptibility (S) genes which results in the induction of disease susceptibility [36].

Furthermore, hypothetical proteins were mostly predicted as pathogen-associated genes. However, in some instances, nucleoside hydrolases, pilus assembly protein were also predicted as pathogen-associated genes in *X. oryzae pv. oryzae* YM15. In the present study, *X. oryzae pv. oryzicola* DY89031 (J18) had the highest number of pathogen-associated GIs (17), followed by *X. campestris pv. campestris* strain 8004 with 16 pathogen-associated GIs and finally *X. campestris pv. campestris* strain ATCC 33913 with 11 pathogen-associated GIs. In both *X. campestris pathovars* of strain 8004 and ATCC 33913, domain containing proteins such as DUF 4189, DUF 3693, DUF 3060 and transfer protein car were grouped as genomic islands which were related to pathogenicity. *X. theicola* CFBP 4691, *X. cucurbitae* ATCC 23378 and *X. euroxanthea* CPBF 426 show absence of GIs even though these genomes are phytopathogenic. It can be also seen that most pathogen-associated GIs predicted were hypothetical protein.

Elongation factor Thermo-Unstable (Tu), aminoglycoside O-phosphotransferase, and efflux transporters were majorly recognised as resistance genes. Resistance genes categorized under resistance island supports the classification of GIs into distinct groups [37]. The presence of elongation factor Tu can be seen both in Gram-positive and Gram-negative bacteria whose function is to transport aminoacyl tRNA to the ribosome [38]. Moreover, it also has been associated with additional virulence functions, which involves adhesion to host extracellular matrix components [38].

Pathogenicity Factors and Mechanism of *Xanthomonas* Species

Presence of Genes in Type III Bacterial Secretion System

It is well known that Xanthomonas species employ different types of bacterial secretion systems to invade and colonise the host plant. Among the existing bacterial secretion system, the Type III bacterial secretion system (T3SS) is of interest as it is involved directly in the bacterial pathogenesis of Xanthomonas. Table 5 summarises the list of T3SS genes identified from the Xanthomonas genomes. Most phytopathogenic bacteria including Xanthomonas, employ the T3SS to cause leakage of nutrients to the apoplast of infected tissue through the secretion of virulence proteins [39]. T3SS, designated as Hypersensitive Response and Pathogenicity (hrp) genes are divided into two subgroups; hrp-conserved (*hrc*) and hrp-associated genes (*hpa*) [40].

It is believed that T3SS has evolved from flagellar apparatus and is currently grouped under the Hrp2 family, which is a part of seven distinct families of non-flagellar T3SSs [41]. The Hrp2 family of T3SS is the determining factor of virulence and is present in most xanthomonads, although *X. albilineans* do not possess T3SS [40]. Annotations from the SEED subsystem indicates the absence of T3SS and associated effectors in *X. euroxanthea* CBPF 426. The absence of T3SS is also observed in several *Xanthomonas* species such as *X. sacchari*, *X. cannabis*, *X. pseudoalbilineans*, and *X. maliensis*, which seem to lack T3SS genes and associated effectors [40].

T3SS encodes for 20 or more genes, located mostly on the chromosome and sometimes on the plasmid categorized as pathogenicity island [41]. T3SS not only encodes for core structural components but also encodes additional accessory and effectors genes located within the pathogenicity islands of T3SS, which may vary between species or pathovars [42].

From the data analysis, we found out that *hrpA*, *hrpB*, *hrpX* and *hrpG* genes are conserved across the selected *Xanthomonas* genomes (Figure 2). *hrpG* is a two-component signal transduction system belonging to the OmpR family, while *hrpX* is a regulator of the AraC family [43]. The function of hrpG of the OmpR family is to positively regulate the expression of *hrpX*, while the role of *hrpX* is to activate the transcription of *hrpB* to *hrpF* [43]. *hrpG* together with *hrpX* encodes for T3 effectors. *hrpX* regulates most genes of *hrpG* regulon [43]. The presence of two regulatory proteins HrpX and HrpG, is required for any Xanthomonas species to cause disease in host plants in a hrp-dependent manner [42]. Based on the previous study findings of Teper and colleagues, it support our current finding that the presence of regulatory proteins, HrpX and HrpG are vital for pathogenic

Table 5. List of Genes in Type III Bacterial Secretion System of Xanthomonas Genomes

Genes of T3SS	*Xanthomonas* Strains																			
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**
YopM	+	+	+	-	-	+	+	+	-	-	+	+	+	-	+	+	-	+	-	-
HrpF	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
HpaA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
HpaB	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrpE	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
HrpD6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
YscD	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
YscS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrcR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
YscQ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
YscP/HpaP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
HrcV	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrcU	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrpB1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
HrpB2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
HrcJ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
YscL	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
HrpB4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
HrcN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrpB7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
HrcT	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrcC	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
HrpG	+	+	+	+	+	+	+	+	-	+	+	+	-	+	+	+	+	+	+	+
HrpX	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
HrpA	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+
HrpB	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+
XopQ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-
YopJ	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
YopP	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HrpW	-	-	-	+	+	-	-	+	+	+	+	-	-	-	-	-	+	+	-	-

RMS methylation unit	+	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	+	-	+	-
Type III Hop protein	-	+	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-
Candidate effector (nucleoside Hydrolase)	+	-	-	-	+	+	+	+	-	+	+	+	-	-	-	+	+	-	-	-
Restriction endonuclease (putative)	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-

+ or – indicates presence or absence of T3SS genes

A = *X. euvesicatoria* pv. *alfalfae* strain CFBP3836, B = *X. campestris* pv. *vesicatoria* 85-10, C = *X. euvesicatoria* LMG930, D = *X. vesicatoria* ATCC 35937 strain LMG911, E = *X. vesicatoria* LM159, F = *X. perforans* 91-118, G = *X. perforans* LH3, H = *X. campestris* pv *raphani* MAFF 106181, I = *X. campestris* pv. *campestri*s ATCC 33913, J = *X. campestris* pv. *campestri*s 8004, K = = *X. citri* UnB-Xtec2D, L = *X. vasicola* pv. *vasculorum* SAM 119, M = *X. vasicola* NCPPB 1060, N = *X. oryzae* pv. *oryzicola* YM15, O = *X. oryzae* pv. *oryzae* DY89031 (J18), P = *X. hortorum* B007-007, P = *X. fragariae* PD885, Q = *X. cucurbitae* ATCC 23378, S = *X. euroxanthea* CPBF 426, T = *X. theicola* strain CFBP4691

HpaB, HrpE, HrpD6, HrpB4 = protein, YopM = possible injected virulence protein/internalin, putative, XopQ = type III effector, YscD, YscS, HrcR – inner membrane protein, YscP = HpaP protein, HrcV = Inner membrane channel protein, HrpB1, HrpB2 = secretion protein, HrcJ = Bridge Between inner and outer membrane lipoprotein, YscL = cytoplasmic protein, HrcN = cytoplasmic ATP synthase, HrcT = inner membrane protein, HrcC = outer membrane forming protein, HrpF = translocator of effector proteins, HrpW = hairpin with pectate lyase domain, candidate type III effector HolPtoQ = nucleoside hydrolase, Restriction Modification system methylation unit = RMS, YopP/YopJ = induces apoptosis, prevent cytokinin induction, inhibit NFkb activation, HrpA, HrpB = ATP-dependent helicas

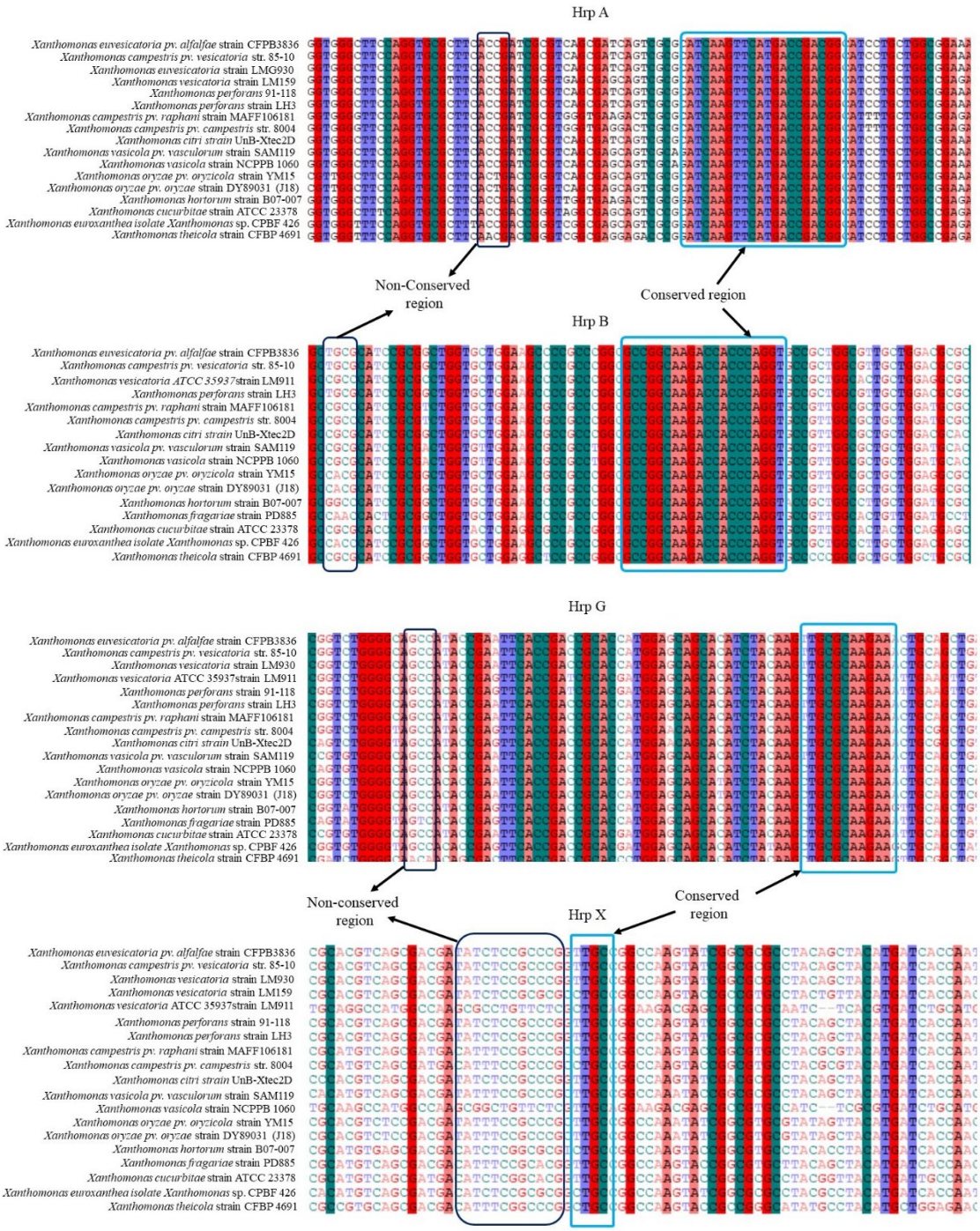

Figure 2. Multiple Sequence Alignment of Hrp genes in Xanthomonas genomes

properties of *Xanthomonas* species [42]. It was also reported that disruption or deletion of either *hrpX* or *hrpG* genes in *Xanthomonas* species possessing T3SS system resulted in complete loss of ability to causes disease, colonise the plant host or induce hypersensitive response (HR) in resistant or non-host plants [42].

T3SS contains Type III chaperones (T3Cs) which interact directly with the Type III effectors (T3Es) and delivers effectors to the host cells. T3Cs are divided into three classes. Class I is subdivided into class IA and class IB, which binds to one or more T3Es [44]. The second class includes chaperones of specialised translocator, while class III chaperones are-flagellar-associated T3SS chaperones [44]. This is well supported by the evidence from a study stating that HpaB belonging to class IB chaperone facilitates the secretion or translocation of effectors in *Xanthomonas*.

The result demonstrated that the HpaB is present in all the species except for *X. euroxanthea* CPBF 426 and *X. theicola* CFBP 4691.

Based on the annotation of the SEED subsystem, YopM is characterised as a part of the T3SS of *Xanthomonas*. The designation YopM usually indicates the Yersinia outer proteins in the species *Yersinia* [39]. In addition, the T3SS pathway mechanisms of *Xanthomonas* species displayed in the KEGG genome database also used the Yersinia gene designation, indicating that the Xanthomonas T3SS pathway mechanism is homologous to that of *Yersinia* species [11].

CONCLUSION

In conclusion, the average nucleotide identity indicates that the selected 20 *Xanthomonas* genomes are closely related. Pathogenic properties of the selected genomes could be observed by investigating the bacterial secretion system, especially the type III bacterial secretion system, which confers pathogenicity. From the present study, *hrpA*, *hrpB*, *hrpX* and *hrpG* of T3SS are conserved across the selected *Xanthomonas* species. Moreover, genomic islands in these genomes indicate that the *Xanthomonas* genomes are of high genome plasticity. Horizontal gene transfer is one of the main mechanisms used by these species to adapt quickly to a new environment. Finally, the presence of secondary metabolite gene clusters indicates these gene clusters' roles in virulence, pathogenesis, and environmental adaptability.

ACKNOWLEDGEMENTS

We would like to thank the School of Biological Sciences and Universiti Sains Malaysia for the facilities provided.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

REFERENCES

1. Aguilar-Marcelino, L., Mendoza-de-Gives, P., Al-Ani, L. K. T., López-Arellano, M. E., Gómez-Rodríguez, O., Villar-Luna, E., & Reyes-Guerrero, D. E. (2020) Chapter 26 - Using molecular techniques applied to beneficial microorganisms as biotechnological tools for controlling agricultural plant pathogens and pest. In V. Sharma, R. Salwan, & L. K. T. Al-Ani (Eds.). *Molecular Aspects of Plant Beneficial Microbes in* Agriculture, 333–349. Academic Press.

2. Timilsina, S., Potnis, N., Newberry, E. A., Liyanapathiranage, P., Iruegas-Bocardo, F., White, F. F., Goss, E. M., & Jones, J. B. (2020) *Xanthomonas* diversity, virulence and plant-pathogen interactions. *Nature Reviews. Microbiology*, ***18***(8), 415–427.

3. Sharma, A., Gautam, S., & Wadhawan, S. (2014) *Xanthomonas*. *Encyclopedia of Food Microbiology: Second Edition*, ***3***, 811–817

4. Goel, A. K., Rajagopal, L., Nagesh, N., & Sonti, R. V. (2002) Genetic locus encoding functions involved in biosynthesis and outer membrane localization of xanthomonadin in *Xanthomonas oryzae* pv. *oryzae*. *Journal of Bacteriology*, ***184***(13), 3539–3548.

5. Poplawsky, A. R., Urban, S. C., & Chun, W. (2000) Biological role of xanthomonadin pigments in *Xanthomonas campestris* pv. *campestris*. *Applied and Environmental Microbiology*, ***66***(12), 5123–5127

6. Agrios, G. N. (2005) Plant Diseases Caused by Prokaryotes: Bacteria and Mollicutes. *Plant Pathology*, 615–703.

7. Yoon, S.-H., Ha, S., Lim, J., Kwon, S., & Chun, J. (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek*, ***110***(10), 1281–1286.

8. Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., … Zagnitko, O. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, ***9***, 75.

9. Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., & Stevens, R. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, ***42***(Database issue), D206–D214.

10. Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, M. H., & Weber, T. (2021) antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Research*, ***49***(W1), W29–W35.

11. Yi, Y., Fang, Y., Wu, K., Liu, Y., & Zhang, W. (2020) Comprehensive gene and pathway analysis of cervical cancer progression. *Oncology Letters*, ***19***(4), 3316–3332.

12. Bertelli, C., Laird, M. R., Williams, K. P., Lau, B. Y., Hoad, G., Winsor, G. L., & Brinkman, F. S. L. (2017) IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research*, ***45***(W1), W30–W35.

13. Koebnik, R., Burokiene, D., Bragard, C., Chang, C., Saux, M. F. Le, Kölliker, R., Lang, J. M., Leach, J. E., Luna, E. K., Portier, P., Sagia, A., Ziegle, J., Cohen, S. P., & Jacobs, J. M. (2021) The complete genome sequence of *Xanthomonas theicola*, the causal agent of canker on tea plants, reveals novel secretion systems in clade-1 xanthomonads. *Phytopathology*, ***111***(4), 611–616

14. Fernandes, C., Martins, L., Teixeira, M., Blom, J., Pothier, J. F., Fonseca, N. A., & Tavares, F. (2021) Comparative genomics of X*anthomonas euroxanthea* and *Xanthomonas arboricola* pv. *juglandis* strains isolated from a single walnut host tree. *Microorganisms*, ***9***(3), 1–15.

15. Zhang, X., & Babadoost, M. (2018) Characteristics of *Xanthomonas cucurbitae* isolates from pumpkins and survival of the bacterium in pumpkin seeds. *Plant Disease*, ***102***(9), 1779–1784

16. Kimbrel, J. A., Givan, S. A., Temple, T. N., Johnson, K. B., & Chang, J. H. (2011) Genome sequencing and comparative analysis of the carrot bacterial blight pathogen, *Xanthomonas hortorum* pv. *carotae* M081, for insights into pathogenicity and applications in molecular diagnostics. *Molecular Plant Pathology*, ***12***(6), 580–594

17. Smits, T. H. M. (2019) The importance of genome sequence quality to microbial comparative genomics. *BMC Genomics*, ***20***(1), 662

18. Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, ***57***(Pt 1), 81–91.

19. Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., Brover, S., Schoch, C. L., Kimchi, A., & DiCuccio, M. (2018) Using

average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *International Journal of Systematic and Evolutionary Microbiology*, ***68***(7), 2386–2392.

20. Lee, I., Ouk Kim, Y., Park, S.-C., & Chun, J. (2016) OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology*, ***66***(2), 1100–1103.

21. Aiman, S., Shehroz, M., Munir, M., Gul, S., Shah, M., & Khan, A. (2018) Species-wide genome mining of *Pseudomonas putida* for potential secondary metabolites and drug-like natural products characterization. *Journal of Proteomics & Bioinformatics*, ***11***(1), 1–7.

22. Kenshole, E., Herisse, M., Michael, M., & Pidot, S. J. (2021) Natural product discovery through microbial genome mining. *Current Opinion in Chemical Biology*, ***60***, 47–54.

23. Etchegaray, A., Silva-Stenico, M. E., Moon, D. H., & Tsai, S. M. (2004) In silico analysis of nonribosomal peptide synthetases of *Xanthomonas axonopodis* pv. *citri*: Identification of putative siderophore and lipopeptide biosynthetic genes. *Microbiological Research*, ***159***(4), 425–437

24. Alok, P., & V., S. R. (2010) Role of the FeoB protein and siderophore in promoting virulence of *Xanthomonas oryzae* pv. *oryzae* on rice. *Journal of Bacteriology*, ***192***(12), 3187–3203.

25. Pandey, S. S., Patnana, P. K., Rai, R., & Chatterjee, S. (2017) Xanthoferrin, the α-hydroxycarboxylate-type siderophore of *Xanthomonas campestris* pv. *campestris*, is required for optimum virulence and growth inside cabbage. *Molecular Plant Pathology*, ***18***(7), 949–962

26. Masschelein, J., Jenner, M., & Challis, G. L. (2017) Antibiotics from Gram-negative bacteria: A comprehensive overview and selected biosynthetic highlights. *Natural Product Reports*, ***34***(7), 712–783

27. Zhong, Z., He, B., Li, J., & Li, Y.-X. (2020) Challenges and advances in genome mining of ribosomally synthesized and post-translationally modified peptides (RiPPs). *Synthetic and Systems Biotechnology*, ***5***(3), 155–172

28. Adamek, M., Spohn, M., Stegmann, E., & Ziemert, N. (2017) Mining Bacterial Genomes for Secondary Metabolite Gene Clusters. *Methods in Molecular Biology (Clifton, N.J.)*, *1520*, 23–4

29. Zhao, N., Pan, Y., Cheng, Z., & Liu, H. (2016) Lasso peptide, a highly stable structure and designable multifunctional backbone. *Amino Acids*, ***48***(6), 1347–1356.

30. Hegemann, J. D., Zimmermann, M., Zhu, S., Steuber, H., Harms, K., Xie, X., & Marahiel, M. A. (2014) Xanthomonins I-III: A new class of lasso peptides with a seven-residue macrolactam ring. *Angewandte Chemie (International Ed. in English)*, ***53***(8), 2230–2234.

31. Bouthillette, L. M., Darcey, C. A., Handy, T. E., Seaton, S. C., & Wolfe, A. L. (2017) Isolation of the antibiotic pseudopyronine B and SAR evaluation of C3/C6 alkyl analogs. *Bioorganic & Medicinal Chemistry Letters*, ***27***(12), 2762–2765.

32. Hudson, C. M., Lau, B. Y., & Williams, K. P. (2015) Islander: A database of precisely mapped genomic islands in tRNA and tmRNA genes. *Nucleic Acids Research*, ***43***(D1), D48–D53.

33. Bertelli, C., & Brinkman, F. S. L. (2018) Improved genomic island predictions with IslandPath-DIMOB. *Bioinformatics (Oxford, England)*, ***34***(13), 2161–2167.

34. Langille, M. G. I., Hsiao, W. W. L., & Brinkman, F. S. L. (2008) Evaluation of genomic island predictors using a comparative genomics approach. *BMC Bioinformatics*, ***9***, 1–10.

35. Julião, M. H. M., Silva, S. R., Ferro, J. A., & Varani, A. M. (2020) A genomic and transcriptomic overview of MATE, ABC, and MFS transporters in *Citrus sinensis* interaction with *Xanthomonas citri* subsp. *citri*. In *Plants* ***9***(6).

36. Li, T., Huang, S., Zhou, J., & Yang, B. (2013) Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. *Molecular Plant*, ***6***(3), 781–789

37. Soares, S. de C., Oliveira, L. de C., Jaiswal, A. K., & Azevedo, V. (2016) Genomic Islands: An overview of current software and future improvements. *Journal of Integrative Bioinformatics*, ***13***(1), 301

38. Harvey, K. L., Jarocki, V. M., Charles, I. G., & Djordjevic, S. P. (2019) The diverse functional roles of elongation factor Tu (EF-Tu) in microbial pathogenesis. *Frontiers in Microbiology*, ***10***, 2351

39. He, S. Y. (1998) Type III protein secretion systems in plant and animal pathogenic bacteria. *Annual Review of Phytopathology*, ***36***, 363–392.

40. An, S. Q., Potnis, N., Dow, M., Vorhölter, F. J., He, Y. Q., Becker, A., Teper, D., Li, Y., Wang, N., Bleris, L., & Tang, J. L. (2019) Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. *FEMS Microbiology Reviews*, ***44***(1), 1–32.

41. Alvarez-Martinez, C. E., Sgro, G. G., Araujo, G. G., Paiva, M. R. N., Matsuyama, B. Y., Guzzo, C. R., Andrade, M. O., & Farah, C. S. (2020) Secrete or perish: The role of secretion systems in *Xanthomonas* biology. *Computational and Structural Biotechnology Journal*, ***19***, 279–302.

42. Teper, D., Pandey, S. S., & Wang, N. (2021) The HrpG/HrpX Regulon of Xanthomonads - An insight to the complexity of regulation of virulence traits in phytopathogenic bacteria. *Microorganisms*, ***9***(1).

43. Tao, H., Fan, S.-S., Jiang, S., Xiang, X., Yan, X., Zhang, L.-H., & Cui, Z.-N. (2019) Small molecule inhibitors specifically targeting the type III secretion system of *Xanthomonas oryzae* on Rice. In *International Journal of Molecular Sciences* ***20***(4).

44. Lohou, D., Lonjon, F., Genin, S., & Vailleau, F. (2013) Type III chaperones & Co in bacterial plant pathogens: A set of specialized bodyguards mediating effector delivery. In *Frontiers in Plant Science* (Vol. 4, p. 435).

45. Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M. A., Toth, I., Salmond, G., & Foster, G. D. (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*, ***13***(6), 614–629.

46. Papaianni, M., Paris, D., Woo, S. L., Fulgione, A., Rigano, M. M., Parrilli, E., Tutino, M. L., Marra, R., Manganiello, G., Casillo, A., Limone, A., Zoina, A., Motta, A., Lorito, M., & Capparelli, R. (2020) plant dynamic metabolic response to bacteriophage treatment after *Xanthomonas campestris* pv. *campestris* infection. *Frontiers in Microbiology*, ***11***(April), 1–15.

47. Wang, Y. Q., Zhang, X. F., Li, N., & Liu, X. (2017) Comparison of cellular responses to *Xanthomonas perforans* infection between resistant and susceptible tomato accessions. *Journal of Plant Physiology*, ***209***, 105–114.

48. Hert, A. P., Marutani, M., Momol, M. T., Roberts, P. D., & Jones, J. B. (2009) Analysis of pathogenicity mutants of a bacteriocin producing *Xanthomonas perforans*. *Biological Control*, ***51***(3), 362–369

49. Gordon, J. L., Lefeuvre, P., Escalon, A., Barbe, V., Cruveiller, S., Gagnevin, L., & Pruvost, O. (2015) Comparative genomics of 43 strains *of Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. *BMC Genomics*, ***16***(1).

50. Fonseca, N. P., Patané, J. S. L., Varani, A. M., Felestrino, É. B., Caneschi, W. L., Sanchez, A. B., Cordeiro, I. F., Lemes, C. G. de C.,

Assis, R. de A. B., Garcia, C. C. M., Belasque, J., Martins, J., Facincani, A. P., Ferreira, R. M., Jaciani, F. J., de Almeida, N. F., Ferro, J. A., Moreira, L. M., & Setubal, J. C. (2019) Analyses of seven new genomes of *Xanthomonas citri* pv. *aurantifolii* strains, causative agents of citrus canker B and C, show a reduced repertoire of pathogenicity-related genes. *Frontiers in Microbiology*, ***10***(October), 1–15.

51. Jalan, N., Kumar, D., Andrade, M. O., Yu, F., Jones, J. B., Graham, J. H., White, F. F., Setubal, J. C., & Wang, N. (2013) Comparative genomic and transcriptome analyses of pathotypes of *Xanthomonas citri* subsp. *citri* provide insights into mechanisms of bacterial virulence and host range. *BMC Genomics*, ***14***(1).

52. Brunings, A. M., & Gabriel, D. W. (2003) *Xanthomonas citri*: Breaking the surface. *Molecular Plant Pathology*, ***4***(3), 141–157.

53. Lang, J. M., DuCharme, E., Ibarra Caballero, J., Luna, E., Hartman, T., Ortiz-Castro, M., Korus, K., Rascoe, J., Jackson-Ziems, T. A., Broders, K., & Leach, J. E. (2017) Detection and characterization of *Xanthomonas vasicola* pv. *vasculorum* (Cobb 1894) comb. nov. causing bacterial leaf streak of corn in the United States. *Phytopathology*, ***107***(11), 1312–1321.

54. Niño-Liu, D. O., Ronald, P. C., & Bogdanove, A. J. (2006) *Xanthomonas oryzae* pathovars: Model pathogens of a model crop. *Molecular Plant Pathology*, ***7***(5), 303–324.

55. Yang, B., & Bogdanove, A. (2013) Inoculation and virulence assay for bacterial blight and bacterial leaf streak of rice. *Methods in Molecular Biology*, ***956***, 249–255.

56. Henry, P. M., Gebben, S. J., Tech, J. J., Yip, J. L., & Leveau, J. H. J. (2016) Inhibition of *Xanthomonas fragariae*, causative agent of angular leaf spot of strawberry, through iron deprivation. *Frontiers in Microbiology*, ***7***(October), 1–11.

57. Ravanlou, A., & Babadoost, M. (2015) Development of bacterial spot, incited by *Xanthomonas cucurbitae*, in pumpkin fields. *HortScience*, ***50***(5), 714–720.