Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2025**; 30(10): 46230
https://doi.org/10.31083/FBL46230

Original Research

Complete Genome Sequence Analysis of the Bacterial Wilt Disease Pathogen *Ralstonia pseudosolanacearum* Strain MLY158

Junxian Zou[1], Guan Lin[1], Juntao Gao[1], Denghui Li[1], Yu Cui[1], Bin Kong[2], Konghua Xie[2], Yong Liu[1,*], Xingjiang Chen[3,*]

[1]School of Biological & Chemical Engineering, Zhejiang University of Science & Technology, 310023 Hangzhou, Zhejiang, China
[2]Guizhou Provincial Tobacco Company Zunyi Company Yuqing County Branch, 564400 Zunyi, Guizhou, China
[3]Guizhou Academy of Tobacco Science, 550001 Guiyang, Guizhou, China
*Correspondence: 116047@zust.edu.cn (Yong Liu); chenxingjiang1@163.com (Xingjiang Chen)
Academic Editor: Ricardo Jorge Pinto Araujo
Submitted: 1 September 2025 Revised: 7 October 2025 Accepted: 11 October 2025 Published: 30 October 2025

Abstract

Background: The *Ralstonia solanacearum* species complex (RSSC) is a group of destructive plant-pathogenic bacteria that targets a wide range of economically important crops across the globe, including tomato, pepper, and tobacco. Extensive research on this plant pathogen is essential due to the severe losses it inflicts on agricultural production. **Methods**: We isolated strain MLY158 from diseased tobacco, identifying it as *Ralstonia pseudosolanacearum*. The strain was characterized genomically by biochemical profiling, genome sequencing, compositional and functional annotation, and comparative genomics. **Results**: MLY158 was capable of utilizing D-glucose (dGLU), sucrose (SAC), and D-trehalose dihydrate (dTRE). The genome had a total size of 5.88 MB and consisted of a circular chromosome and a circular megaplasmid. It contained 5485 coding genes and had a GC content of 67.50%. Comparative genomic results revealed that MLY158 is closely related to *R. pseudosolanacearum* strain GMI1000 (average nucleotide identity (ANI) value of 99.03%). MLY158 has 527 special genes and 13 homologous genes of species-specific gene families. The primary differences between MLY158 and genomes from other strains are located in the phage protein region and show characteristics of high genomic uniqueness. **Conclusions**: Complete genome sequence analysis of MLY158 has contributed important information regarding the genome of the bacterial wilt disease pathogen *R. pseudosolanacearum*. This work provides useful references for future research into molecular disease control strategies and disease-resistant breeding.

Keywords: *Ralstonia solanacearum*; genomics; plant disease; tobacco

1. Introduction

The United Nations (UN) Food and Agriculture Organization and the Organization for Economic Cooperation and Development have reported that rapid growth of the agricultural sector is now facing a major challenge from crop diseases, which are impeding further development [1]. *Ralstonia solanacearum* species complex (RSSC) is a phytopathogenic bacterium and a member of the *Proteobacteria* [2]. RSSC ranks among the most devastating phytopathogenic bacteria globally [3], with a wide distribution in tropical, subtropical and temperate regions [4]. RSSC infects over 450 plant species across 54 families, including economically important crops like tomato, peanut and tobacco [5,6]. RSSC infection of tobacco causes rapid wilting of the plant, resulting in a 10–30% loss in production worldwide [7].

RSSC is a soil-borne pathogen. Upon infecting a plant, it enters the vascular tissue and multiplies, ultimately leading to plant death [8]. RSSC can return to the soil through withered plants and survive under nutrient-poor conditions for several years in preparation for the next infestation [9]. The classification of RSSC is based on its geographic distribution, with four phylotypes identified to date: phylotype I (Asia), phylotype II (America), phylotype III (Africa), and phylotype IV (Indonesia) [10]. RSSC was also classified into five biochemical variants based on the utilization of carbohydrates (lactose, maltose, fibrous dextran, and mannitol, etc.) [11]. The RSSC classification system was revised in 2014, with division into three distinct species: *R. solanacearum*, *R. pseudosolanacearum*, and *R. syzygii*. These species are distinguished by their distinct phylotypes, with *R. solanacearum* predominantly comprising phylotype II, *R. pseudosolanacearum* primarily consisting of phylotypes I and III, and *R. syzygii* comprising phylotype IV [12–14]. This classification system is currently employed by the majority of researchers in the field.

Significant biochemical differences exist among different bacterial strains. The VITEK 2 Compact fully automated microbial identification system offers highly efficient and rapid biochemical identification and characterization of strains. This system enables rapid biochemical characterization by utilizing gram-negative identification (GN) cards designed specifically for Gram-negative bacteria, with the cards containing multiple, independent biochemical reactions [15]. It has been widely and successfully applied in comparative studies of clinical, foodborne, and environmental pathogens [16]. The continuous devel-

opment of microbial genome sequencing technology provides an important means to identify bacterial pathogenicity at the genetic level. Salanoubat *et al*. [17] isolated *R. pseudosolanacearum* strain GMI1000 from tomato plants and performed whole genome sequencing, thereby laying the foundation to study its pathogenesis. Subsequently, an increasing number of strains have undergone complete genome sequencing. The RSSC genome is approximately 5.9 MB in size and is primarily composed of a circular chromosome and a circular megaplasmid [18]. The pathogenicity of *R. pseudosolanacearum* is closely related to its virulence factors (VFs). *R. pseudosolanacearum* causes plant death by colonizing the vascular tissues and inducing the synthesis of large quantities of extracellular polysaccharides that impede water transport [19]. In addition, VFs including effector proteins, flagella, and cell surface appendages allow *R. pseudosolanacearum* to attach to plant roots during the early stages of infection [20]. The Type III secretion system (T3SS) has been studied extensively [21]. *R. pseudosolanacearum* utilizes T3SS to inject effector proteins into plant cells, thereby suppressing the host's defense mechanisms and in some cases triggering immune responses [22,23].

Although a growing number of strains have been sequenced, additional genomic data are required for a comprehensive analysis of this species. Furthermore, substantial variation in virulence is observed among strains from different geographical origins and host plants. Albuquerque *et al*. [24] reported two *R. pseudosolanacearum* strains that exhibit differing compatibility with the tomato cultivar Hawaii 7996, indicating the importance of variations in effector genes and pathogenic interactions among strains. Ding *et al*. [12] recently compared four strains derived from sunflowers. Despite originating from the same host, these strains exhibited significant differences in pathogenicity and genetic composition [12].

In the present study, we sequenced the complete genome of *R. pseudosolanacearum* strain MLY158. Compared to the reference genome, MLY158 carries special phage-associated genes and plasmid fragments. These findings not only suggest potential genomic plasticity and differences in virulence, but also provide valuable clues for deepening our understanding of gene transfer and evolution at the genetic level in *R. pseudosolanacearum*. Moreover, the comprehensive genome sequencing and annotation of strain MLY158 enriches the genomic database of *R. pseudosolanacearum* and provides a crucial reference for further research into molecular disease control strategies and disease-resistant breeding.

2. Materials and Methods

2.1 Isolation, Pathogenicity Testing and Biochemical Characterization of Strains

R. pseudosolanacearum strain MLY158 was isolated from a diseased tobacco stem leachate from Dejiang County, Tongren, Guizhou Province. After incubation in nutrient agar (NA) medium at 30 °C for 48 h, single colonies were selected for further inoculation into conical flasks containing nutritious broth (NB) liquid medium. This medium was shaken at 180 rpm/min at 30 °C overnight, and the cells were then collected by centrifugation at 4000 rpm/min for 5 min. The prepared strains were used for subsequent experiments. The OD_{600} was adjusted to 0.1 using sterile water, and the suspension injected into the leaf veins of tobacco seedlings using a sterile syringe. After inoculation, the seedlings were placed in a light incubator (Shanghai Boxun Industrial Co., Ltd. Medical Equipment Factory, Shanghai, China) at 30 °C with a 16-h light cycle and an 8-h dark cycle. The pathogen was then re-isolated and identified from the infected tobacco plants.

The prepared strains were suspended in sterile saline (0.45% to 0.50% NaCl solution, pH 4.5 to 7.0) and the McFarland turbidity was adjusted to 0.50–0.63. The bacterial suspension tubes and GN cards were placed in a card rack for online analysis of the biochemical reaction patterns. The 16S ribosomal RNA gene was amplified using the universal primers Eubac27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and Eubac1492R (5′-GGTTACCTTGTTACGACTT-3′). The resulting DNA sequences were measured and subjected to sequence comparisons on GenBank (https://www.ncbi.nlm.nih.gov/).

2.2 Genome Sequencing and Assembly

DNA from MLY158 was extracted using the MGIEasy Microbiome DNA Extraction Kit (1000027955, MGI Tech, Shenzhen, Guangdong, China) according to the manufacturer's instructions. The genome for this strain was sequenced using Pacbio and DNBSEQ platforms.

2.3 Prediction of Genomic Components

Gene prediction of the assembly results was performed using Glimmer v3.02 (The Center for Computational Biology at Johns Hopkins University, Baltimore, MD, USA) [25]. Non-coding RNA prediction was performed using RNAmmer v1.2 (The Center for Biological Sequence Analysis at the Technical University of Denmark, Kongens Lyngby, Denmark) [26], tRNA prediction by tRNAscan v1.3.1 (Lowe Lab at the University of California, Santa Cruz, CA, USA) [27], and sRNA prediction by Rfam v9.1 (https://rfam.org) [28]. Tandem repeats were predicted by Tandem Repeats Finder (TRF) v4.04 (Benson Lab at Boston University, Boston, MA, USA) [29]. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) were identified using CRISPRCasFinder v4.2.20 (https://crisprcas.i2bc.paris-saclay.fr/).

2.4 Genomic Function Annotation

Gene function annotation was performed based on Non-Redundant Protein Database (NR) v2023-4-20 (https://www.ncbi.nlm.nih.gov/refseq/

Table 1. Biochemical characterization of strain MLY158.

Well no.	Tests	Results	Well no.	Tests	Results	Well no.	Tests	Results
2	APPA	–	21	BXYL	–	42	SUCT	+
3	ADO	–	22	BAIap	–	43	NAGA	–
4	PyrA	–	23	ProA	+	44	AGAL	–
5	IARL	–	26	LIP	–	45	PHOS	–
7	dCEL	–	27	PLE	–	46	GlyA	–
9	BGAL	–	29	TyrA	–	47	ODC	–
10	H2S	–	31	URE	+	48	LDC	–
11	BNAG	–	32	dSOR	–	53	IHISa	–
12	AGLTp	–	33	SAC	+	56	CMT	–
13	dGLU	+	34	dTAG	–	57	BGUR	–
14	GGT	–	35	dTRE	+	58	O129R	–
15	OFF	–	36	CIT	–	59	GGAA	–
17	BGLU	–	37	MNT	–	61	IMLTa	–
18	dMAL	–	39	5KG	–	62	ELLM	–
19	dMAN	–	40	ILATk	–	64	ILATa	–
20	dMNE	–	41	AGLU	–			
			GN Card			Incubation time: 6.00 h		

"+" indicates that the bacterium is capable of producing a certain reaction or possessing a certain characteristic under certain conditions; "–" indicates that the bacterium is unable to produce a reaction or does not possess a characteristic under certain conditions. APPA, alanine-phenylalanine-proline arylamidase; ADO, adonitol; PyrA, L-pyrrolidonyl arylamidase; IARL, L-arabitol; dCEL, D-cellobiose; BGAL, β-galactosidase; H2S, H2S production; BNAG, β-N-acetylglucosaminidase; AGLTp, glutamyl arylamidase pNA; dGLU, D-glucose; GGT, γ-glutamyltransferase; OFF, Fermentation of glucose; BGLU, β-glucosidase; dMAL, D-maltose; dMAN, D-mannitol; dMNE, D-mannose; BXYL, β-xylosidase; BAIap, β-alanine arylamidase pNA; ProA, L-proline arylamidase; LIP, Lipase; PLE, Paleo sugar; TyrA, Tyrosine arylamidase; URE, Urease; dSOR, D-Sorbitol; SAC, Sucrose; dTAG, D-tagatose; dTRE, D-trehalose; CIT, Citrate (sodium); MNT, Malonate; 5KG, 5-Keto-D-gluconate; ILATk, L-Lactate alkalinisation; AGLU, α-glucosidase; SUCT, Succinate alkalinisation; NAGA, β-N-acetylgalactosaminidase; AGAL, α-galactosidase; PHOS, Phosphatase; GlyA, Glycine arylamidase; ODC, Ornithine decarboxylase; LDC, Lysine decarboxylase; IHISa, L-histidine assimilation; CMT, Coumarate; BGUR, β-glucuronidase; O129R, O/129 resistance; GGAA, Glutamate-glycine-arginine arylamidase; IMLTa, L-malate assimilation; ELLM, Ellman's reagent (reductive thiol groups); ILATa, L-lactate assimilation; GN, gram-negative identification.

about/nonredundantproteins/), Cluster of Orthologous Groups of proteins (COG) v202011-25 (https://www.ncbi.nlm.nih.gov/research/cog/) [30], Gene Ontology (GO) v2019-07-01 (https://www.geneontology.org/) [31], Kyoto Encyclopedia of Genes and Genomes (KEGG) v106.0 (https://www.genome.jp/kegg/) [32], and Swiss-Prot v2022_03 databases (https://www.expasy.org/resources/uniprotkb-swiss-prot) [33]. The type III effector protein of strain MLY158 was predicted by the T3SS v1.0 database [21]. Additional annotations were performed using IPR (https://www.ebi.ac.uk/interpro/), Antibiotic Resistance Genes Database (ARDB) v1.1 [34], virulence factor database (VFDB) v2023-4-28 (https://www.mgc.ac.cn/VFs/) [35], Carbohydrate-Active enZYmes Database (CAZy) v2022-9-15 (https://www.cazy.org/) [36], and The Comprehensive Antibiotic Resistance Database (CARD) v3.0.9 (https://card.mcmaster.ca/) [37]. Default parameters were used for all analyses.

2.5 Comparative Genome Analysis

Genomic similarity was analyzed based on the average nucleotide identity (ANI) using fastANI v1.32 software (The Center for Computational Biology at the University of California, San Diego, CA, USA) [38]. In addition, MUMmer v3.22 (The Center for Computational Biology at Johns Hopkins University, Baltimore, MD, USA) was used to align five *R. solanacearum* strains (RS-UW251, CFBP2957, RS-CIAT078, Po82, RS-Rs5), two *R. pseudosolanacearum* strains (GMI1000, CMR15), and three *R. syzygii* strains (PSI07, RS-T95, RS-SL2064). These were compared with the reference genome of MLY158 to determine the percentage of co-lined genes [39]. We also analyzed the core and pan genomes of these 11 strains, mainly using the software CD-HIT v4.6.6 (The Li Lab at the University of Texas Health Science Center at Houston, Houston, TX, USA) [40]. Phylogenetic trees were constructed using the PHYML (maximum likelihood method) algo-

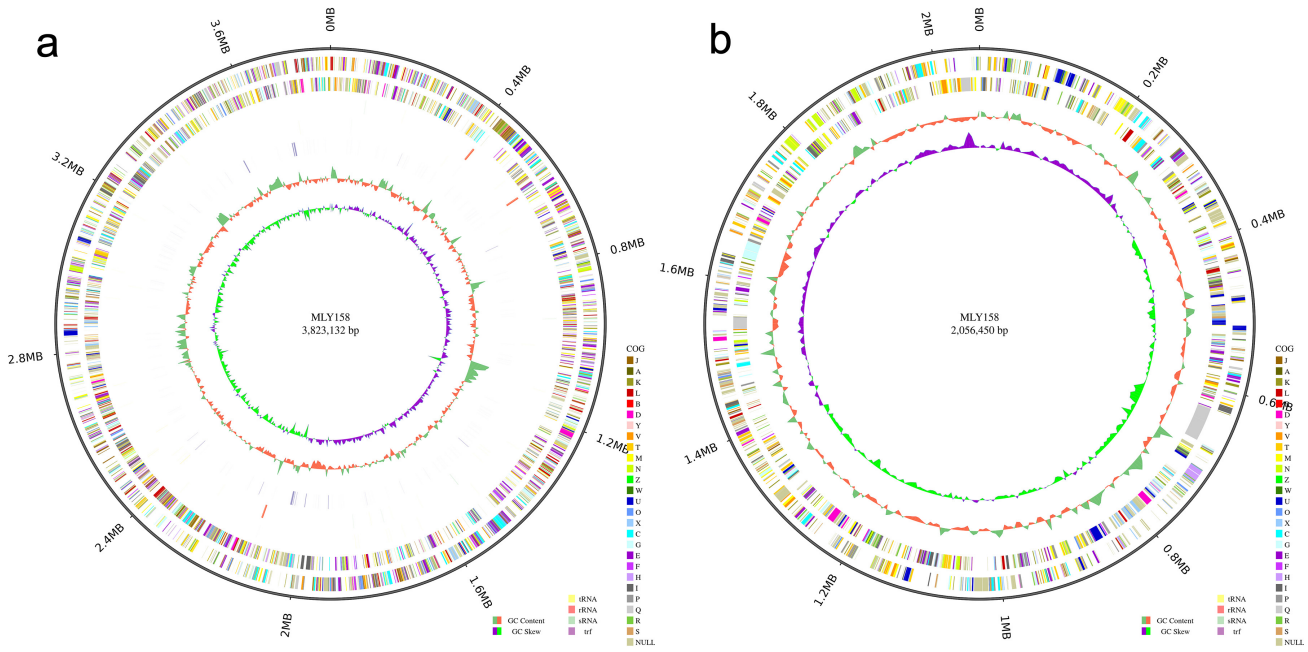

Fig. 1. Circle diagram. (a) The chromosome circle diagram of MLY158. (b) The giant plasmid circle diagram of MLY158.

rithm for the single nucleotide polymorphism (SNP) matrix of the population of the 11 strains. Default parameters were used for all analyses. Additionally, comparative genomic analysis of the MLY158 genome sequence was performed against 10 highly similar *R. pseudosolanacearum* strains. Plasmid-specific regions and mutation hotspots were identified by comparing the comparative genomic circles. We also performed comparative analysis of the 10 similar genome sequences against the large plasmid in the MLY158 genome and constructed comparative genomic circles, primarily using BRIG v0.9.5 software (The Australian Centre for Ecogenomics at the University of Queensland, Brisbane, QLD, Australia).

3. Results

3.1 Backcrossing Test and Biochemical Characterization of Strain MLY158

Following inoculation with MLY158, tobacco seedlings exhibited wilting symptoms, with blackened stems and extensive leaf yellowing and necrosis (**Supplementary Fig. 1a**). The strain was re-isolated from the blackened stems (**Supplementary Fig. 1b**), and 16S rRNA sequencing demonstrated 100% similarity to strain MLY158. Based on Koch's postulates, MLY158 was identified as *R. pseudosolanacearum*, the pathogen causing tobacco wilt.

Table 1 shows the biochemical characteristics of strain MLY158, as determined by GN card. Carbon sources for MLY158 include D-glucose (dGLU), sucrose (SAC) and D-trehalose (dTRE), but not D-cellobiose (dCEL), Fermentation of glucose (OFF), D-maltose monohydrate (dMAL), D-mannose (dMNE), Paleo sugar (PLE) and D-tagatose (dTAG). With regard to chemical sensitivity, MLY158 exhibited a positive reaction only for succinate alkalization (SUCT) among the tested enzymatic reactions. This indicates it has the ability to metabolize succinate and induce environmental alkalization. Additionally, the MLY158 strain demonstrated resistance to the antibiotic O129R.

Table 2. Genomic characterization of MLY158 isolated from diseased tobacco stalks.

Feature	Genome
Genome size	5.88 MB
GC content (%)	67.50
tRNA	60
5s_rRNA	4
16s_rRNA	4
23s_rRNA	4
sRNA	22
CRISPRs	2 (questional)
TRF	307
Minisatellite DNA	209
Microsatellite DNA	34

GC, G+C; CRISPRs, Clustered Regularly Interspaced Short Palindromic Repeats; TRF, Tandem Repeats Finder.

3.2 Characterization of the Genome

The genome of *R. pseudosolanacearum* strain MLY158 was sequenced using both the DNBSEQ and PacBio platforms to achieve a high-quality assembly. The 1270 MB (216×) Clean Data was generated with the DNBSEQ platform, while the PacBio platform was used to obtain 6263 MB (1065×) Subreads after removing

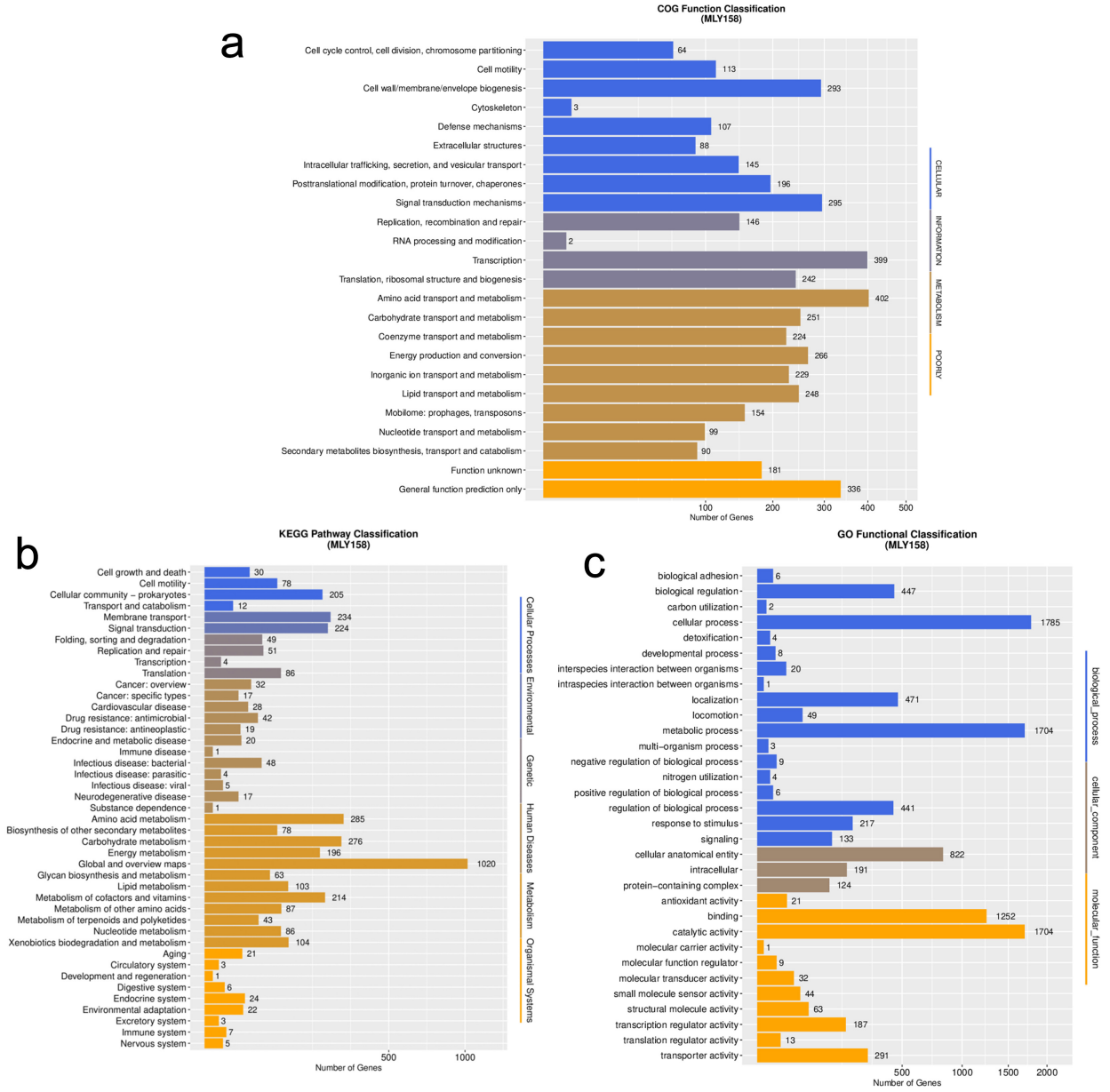

Fig. 2. Functional classification map of genes in MLY158. (a) The function of COG annotation. (b) The function of KEGG annotation. (c) The function of GO annotation. COG, Cluster of Orthologous Groups; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.

low-quality or ambiguous read lengths. The genome was 5.88 MB in size and consisted of a circular chromosome (Fig. 1a) and a circular giant plasmid (Fig. 1b). The genome depth was 45.31× and the GC content was 67.50%. A total of 5485 genes were predicted for MLY158, with the gene length distribution shown in **Supplementary Fig. 2**. The genome also encoded 60 tRNAs, 12 rRNAs, and 22 sRNAs (Table 2). Genomic analysis revealed the presence of 307 tandem repeats, 209 minisatellites, and 34 microsatellites.

In total, 5404 genes were annotated. The gene annotation rates are shown in **Supplementary Table 1**. Gene function annotation successfully categorized 3940 genes into 24 COGs (Fig. 2a). Among the assigned genes, most were involved in transcription (n = 399), general function (n = 336), amino acid transport and metabolism (n = 402), signal transduction mechanisms (n = 295), and cell wall/membrane/envelope biogenesis (n = 293). In addition, 3060 and 3097 annotated genes were categorized by GO and KEGG, respectively (Fig. 2b,c). Among these were *murI* (glutamate racemase), *tatC* (component of the Tat secretion system), and *dsbA* (catalyzes disulfide bond formation in periplasmic proteins), which are related to xylem sap fitness factors. Comparison with the VF database revealed that MLY158 contained 478 homologous genes, including *Cya*, *Hsp*, and flagella. We also annotated VF-related genes, including those involved in stress

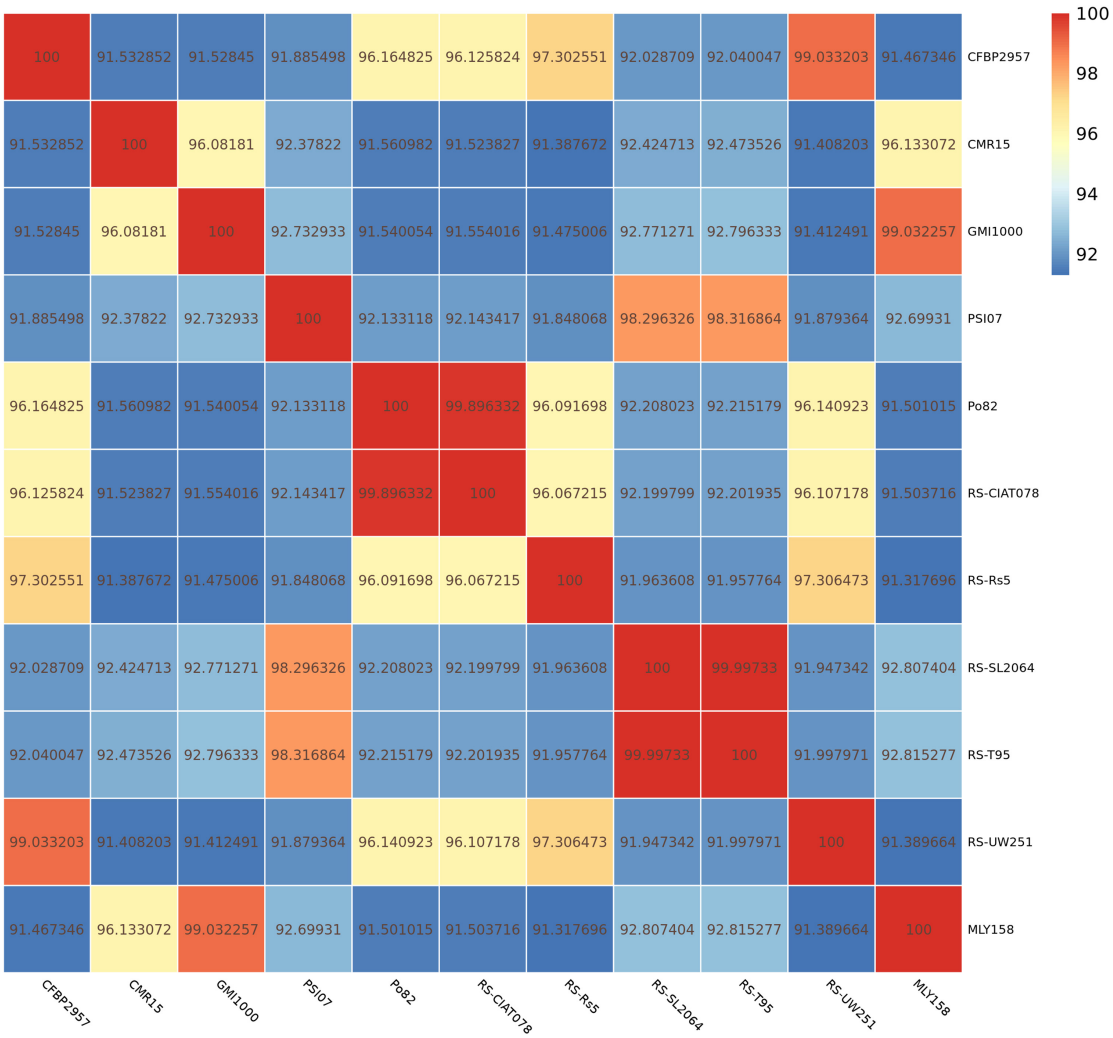

Fig. 3. ANI heat map of MLY158 and reference strains.

tolerance (*bcp*, *acrA*, *acrB*, and *dps*), motility (*pilA* and *filC*), cell wall degradation (*cbhA*), and exopolysaccharide (EPS) biosynthesis (*epsABCDEFP*). The antibiotic resistance genes identified include *varG* (from *Vibrio cholerae*) and *adeF*. These were mainly involved in antibiotic inactivation and efflux.

3.3 Comparative Genomic Analysis

Next, we determined the ANI (parameter: -c 1024) between the genome of strain MLY158 and the genome of the reference strains (CFBP2957, CMR15, GMI1000, PSI07, Po82, RS-CIAT078, RS-Rs5, RS-SL2064, RS-T95, RS-UW251). The results showed that MLY158 was most closely related to strain GMI1000 (ANI 99.03%), followed by strain CMR15 (ANI 96.13%) (Fig. 3). The ANI values between MLY158 and the remaining reference strains ranged from 91.32% to 92.82%. The ANI values between PSI07 and strains RS-T95 and RS-SL2064 were 98.32% and 98.30%, respectively. The ANI values between RS-Rs5 and RS-UW251 (97.31%), RS-CIAT078 (96.07%), Po82 (96.09%), and CFBP2957 (97.30%) strains were all >95%. Notably, the ANI value between RS-UW251 and CFBP2957 was 99.03%, and between Po82 and RS-CIAT078 it was 99.90%, indicating a high degree of similarity. To identify structural genomic variations that arose during evolution, the MLY158 genome was compared with the genomes of 10 reference strains using MUMmer software. Comparative analysis demonstrated nucleotide and amino acid sequence homology between strain MLY158 and the reference strains (**Supplementary Tables 2,3**, respectively). Further analysis showed that the highest match between MLY158 and the reference strains in terms of chromosomes was GMI1000 (91.19%), followed by strain CMR15 (82.26%). The highest match between MLY158 and reference strains in terms of giant plasmids was GMI1000 (94.74%), followed by strain CMR15 (81.48%).

Pan-genome analysis of strain MLY158 alongside the 10 reference strains identified 5505 homologous gene groups. Among these, 2998 core genes were shared across all 11 genomes (Fig. 4a). MLY158 contained 527 unique genes, the second-highest count among the differ-

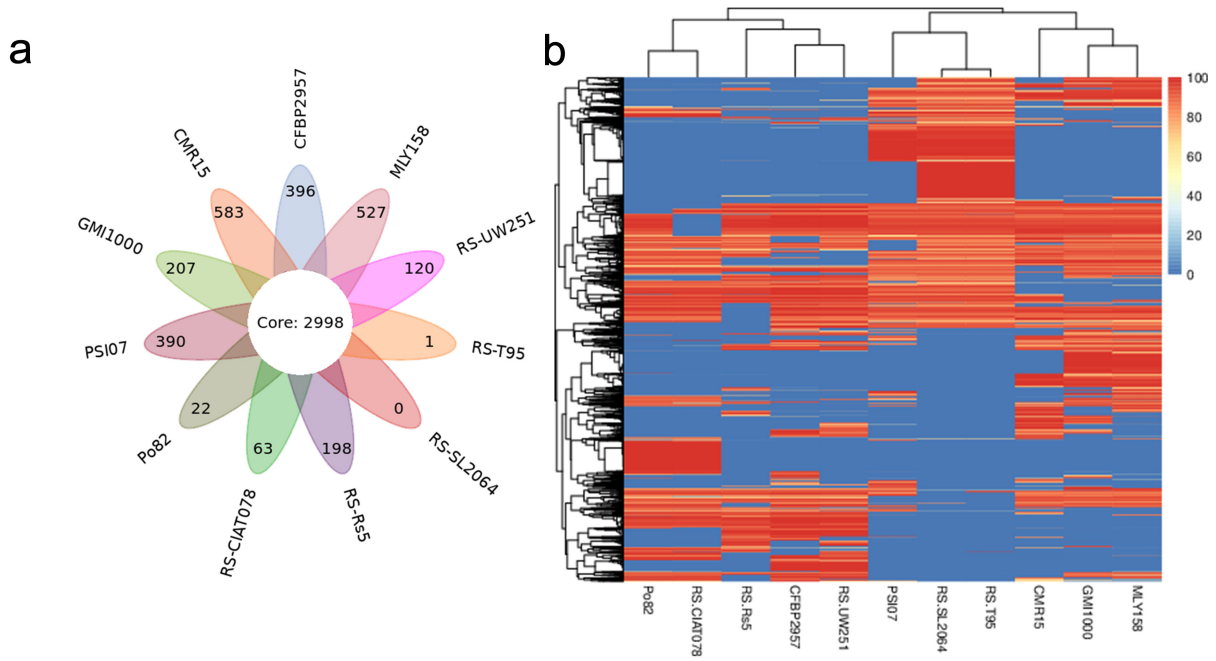

Fig. 4. Comparison of genes between MLY158 and the reference strains. (a) Venn diagram of Pan gene set for MLY158 and the reference strains. (b) Heatmap of dispensable genes for MLY158 and the reference strains.

Table 3. Number of homologous genes in the Dispensable gene family of MLY158 and the reference strains.

Sample ID	Gene number	Clustered gene	UnClustered gene	Family num	Unique family
CFBP2957	5065	4669	393	3097	3
CMR15	5123	4609	505	3056	9
GMI1000	5109	4928	181	3137	13
PSI07	4967	4611	356	2982	1
Po82	4649	4622	27	3089	0
RS-CIAT078	4548	4453	95	3031	2
RS-Rs5	4560	4405	155	2977	9
RS-SL2064	4719	4713	6	3174	0
RS-T95	4729	4727	2	3176	0
RS-UW251	4758	4660	98	3083	2
MLY158	5485	4955	530	3128	13

ent strains. MLY158-specific genes were annotated using COG (**Supplementary Fig. 3**), with 4 genes related to cellular, 17 genes related to information, 8 genes related to metabolism, and 8 genes related to poorly. According to gene family analysis (Table 3), the MLY158 and GMI1000 strains had the highest number of special gene families (n = 13), with MLY158 containing the most unclustered genes (n = 530). A heatmap of dispensable genes was also drawn based on their distribution in different samples (Fig. 4b). MLY158 and GMI1000 clustered together, forming a group with CMR15 that exhibited high similarity. RS-T95, RS-SL2064, and PSI07 clustered together. RS-UW251, CFBP2957, and RS-Rs5 also clustered together, whereas RS-CIAT078 and Po82 clustered separately.

A phylogenetic tree was constructed using the Neighbor-Joining method (bootstrap setting = 1000) based on the results of GeneFamily analysis. MLY158 was found to cluster with GMI1000, indicating a close relationship, and also formed a branch with CMR15 (Fig. 5). PSI07, RS-T95 and RS-SL2064 were closely related to each other and clustered together. RS-CIAT078 and Po82 were closely related, and RS-UW251 and CFBP2957 were also closely related, with each of these pairs forming a cluster. However, RS-Rs5 was more distant from the other strains and formed a unique branch.

We subsequently performed BLAST alignment of MLY158 against the NCBI database and selected 10 highly similar plasmid sequences (10 *R. pseudosolanacearum* plasmid sequences). The comparative genomic circle is shown in Fig. 6. The full-length MLY158 plasmid spans 2,056,450 bp, with a GC content of 67.08%. Four antimicrobial resistance genes were identified: the efflux pump-associated *adeFGH* gene *adeF*, the *mdtABC* gene *mdtB*, and the OXA-type β-lactamase gene *blaOXA*. Vir-

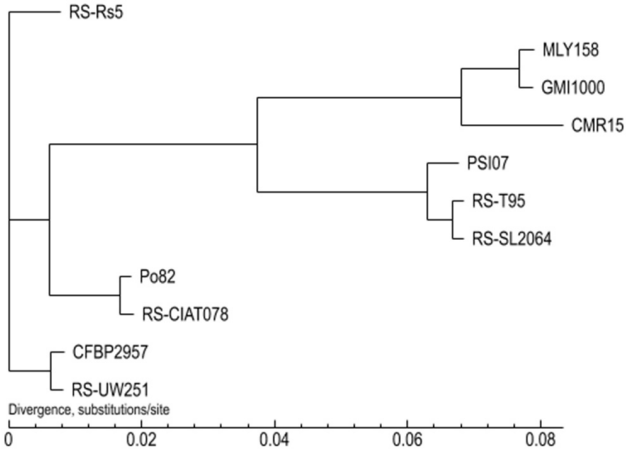

Fig. 5. Phylogenetic tree between MLY158 and reference strains.

ulence components included *flagella*, which is essential for cell motility and invasion, the *cyaE* gene which encodes a toxin with adenylate cyclase activity, and *Hsp* genes which encode anti-inflammatory proteins. Comparative genomic circle analysis revealed that differences between the MLY158 plasmid and other genomes occur primarily at phage protein locations (784,811–843,609 bp, and 1,420,968–1,421,687 bp).

4. Discussion

The whole genome sequence of *R. pseudosolanacearum* GMI1000 isolated from tomato plants was reported in 2002. This was the first whole genome sequence of the species and laid a solid foundation for elucidating its pathogenic mechanisms and evolutionary history [17]. With the continuous improvements in microbial genome sequencing technology, an increasing number of strains have now been sequenced. The complete genome of *R. pseudosolanacearum* strain MLY158, isolated from diseased tobacco, was sequenced here using a combination of PacBio and DNBSEQ technologies. The results provide key insights into the genomic diversity of this pathogen.

One of the complexities of RSSC infestation of plants is its ability to utilize different carbon sources as nutrients. The strain adapts to the xylem of plants in relation to bacterial metabolism. KEGG annotation identified the gene *scrA*, which is related to sucrose degradation metabolism, but which was not found in the CFBP2957 strain. The biochemical results for MLY158 showed that it could utilize dGLU, SAC and dTRE as carbon sources. Sucrose serves as an important carbon source in the xylem sap of plants, providing nutrients for *R. pseudosolanacearum* to colonize the xylem after infection [41]. Additionally, MLY158 carries the envelope-associated genes *murI* (glutamate racemase), *tatC* (component of the Tat secretion system) and *dsbA* (catalyzes disulfide bond formation in periplasmic proteins), which significantly enhance its adaptability in the xylem. Biochemical characterization of Gram-negative bacteria with the GN card significantly reduces the incubation time required for traditional biochemical testing (24–48 h). Biochemical test reports can be generated within 6–24 h, enabling rapid analysis of antibiotic susceptibility for multidrug-resistant Gram-negative bacteria (*P. aeruginosa*, *E. coli*, and *E. cloacae*) in clinical settings. This is critical for patients with severe infections [16]. Studies have shown that the GMI1000 and PSI07 strains require the auxiliary biosynthetic genes *thiC* (thiamine biosynthesis gene) and *purU* (putative formyltetrahydrofolate deformylase) to allow full growth in xylem sap [42]. MLY158 also possesses these two genes. *R. pseudosolanacearum* may consume SAC and dTRE during the infestation of plants by *R. pseudosolanacearum*. Moreover, *R. pseudosolanacearum* infestation has been shown to reduce soil ammonium nitrogen and total plant nitrogen, thereby severely damaging crops [43].

Genome sequencing can provide an important reference to understand plant-pathogen interactions at the genome level. The RSSC genome is approximately 5.9 MB in size and is composed primarily of a circular chromosome and a circular megaplasmid, with occasionally also a small plasmid (FJ1003 contains a 0.2 Mb small plasmid) [44]. Genome sequencing results revealed a smaller genome size for MLY158 compared to other strains isolated from tobacco such as gd-2 (5.93 Mb), CQPS-1 (5.89 Mb) and FJ1003 (5.90 Mb) [11]. However, the predicted genome of MLY158 (5485 genes) was larger than that of gd-2 (5074), CQPS-1 (5229) and FJ1003 (5010) [45]. In addition, strains isolated from different hosts each possess different genomic characteristics [44,46], highlighting the complex diversity of *R. pseudosolanacearum* [47]. Comparative genetics identified candidate VFs and antibiotic resistance genes, including stress tolerance (*bcp*, *acrA*, *acrB*, and *dps*), motility (*pilA* and *filC*), cell wall-degrading enzymes (*cbhA*), and EPS biosynthesis (*epsABCDEFP*)-related virulence genes. In addition, we identified *varG* (from *Vibrio cholerae*) and *adeF* resistance genes similar to gd-2 [11]. This study enriches the available gene pool, expands our understanding of virulence mechanisms, and informs future drug design [46]. Unlike GMI1000 and PSI07, MLY158 lacks the *SpeC* gene, which is essential for growth in rich media. KEGG annotation revealed that MLY158 carries the *phcA* gene, which positively regulates the production of bacterial EPS under conditions of high cell density [48]. The earliest reported genome of *R. pseudosolanacearum* was model strain GMI1000, which laid the foundation for understanding the effector gene pool and the regulation of pathogenicity mechanisms. However, GMI1000 originated from diseased tomato plants in Madagascar and shows genetic differences with strains that are prevalent in Asia. MLY158 originates from Guizhou Province in China. Genomic sequencing revealed this strain carries multiple unique plasmid fragments and phage-associated elements, suggesting

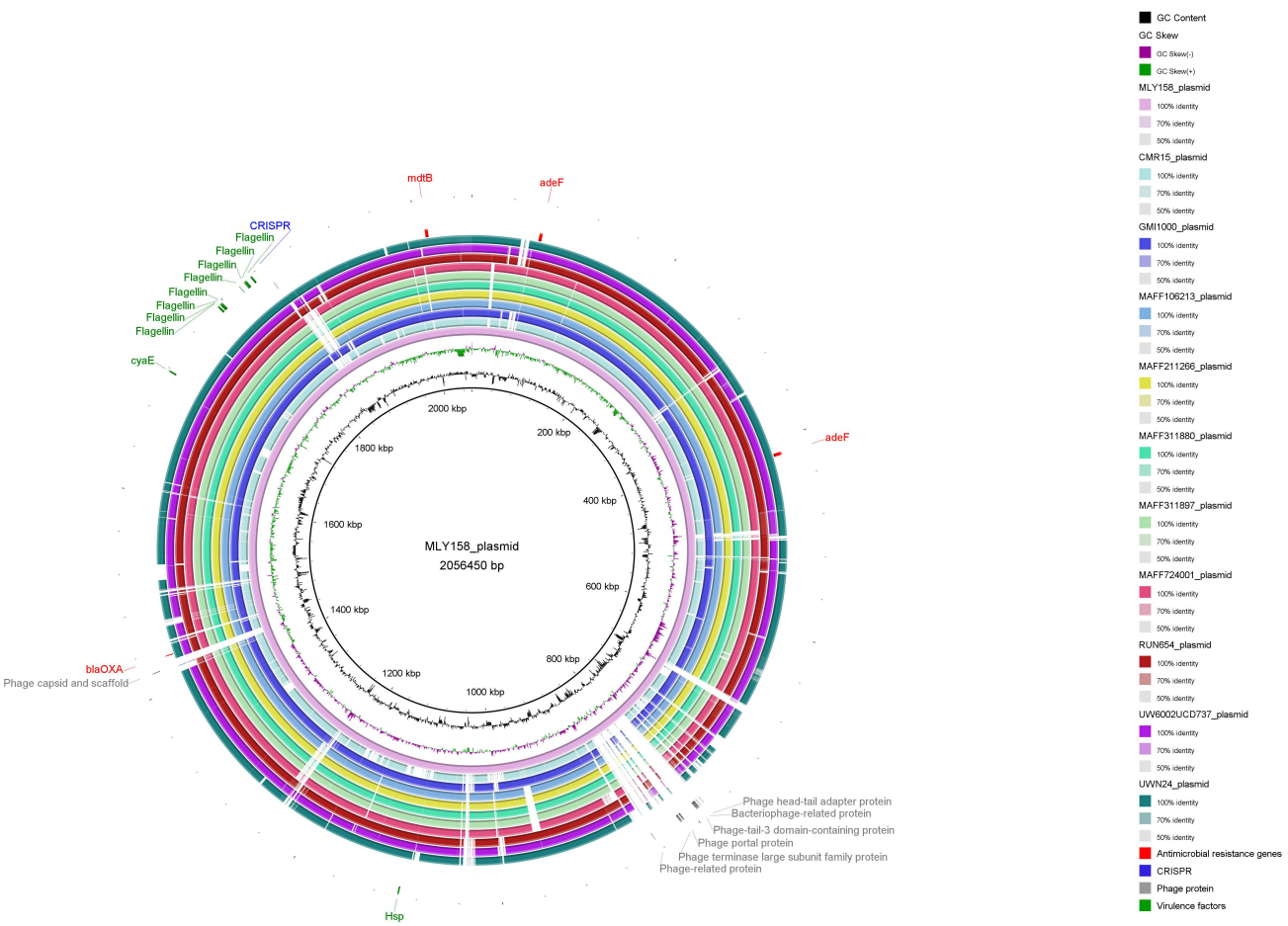

Fig. 6. Comparison of genome circle analysis of MLY158.

potentially novel characteristics in genomic plasticity and horizontal gene transfer. This discovery also provides new perspectives for investigating host-pathogen interactions and the ecological adaptation of this pathogen.

Comparison of strain MLY158 with 10 reference strains in this study revealed that MLY158 was more similar (ANI >96.13%) to two *R. pseudosolanacearum* strains (GMI1000 and CMR15) than to the other strains. According to taxonomic criteria, strains with ANI >95% are considered to belong to the same species [38]. Strain MLY158 therefore belongs to *R. pseudosolanacearum*. Despite the complex diversity of *R. pseudosolanacearum* strains, all strains require a core genome to maintain their basic biological characteristics [49]. The number of special genes varies considerably between strains, explaining the high level of complexity among *R. pseudosolanacearum* [46]. The MLY158 special genes identified through annotation are primarily associated with metabolism (defense mechanisms; intracellular trafficking, secretion, and vesicular transport; post-translational modification; protein turnover, chaperones; signal transduction mechanisms), information (replication, recombination and repair; transcription; translation, ribosomal structure and biogenesis), cellular (amino acid transport and metabolism; carbohydrate transport and metabolism; coenzyme transport and metabolism; mobilome: prophages, transposons), and poorly (general function prediction only). These genes may confer unique survival advantages to the MLY158 strain. In addition to special genes, dispensable genes also play an important role in the genetic diversity of species. Our results showed that MLY158 has high similarity with the GMI1000 and CMR15 strains. These strains clustered together into one group (Fig. 4b), and again in the phylogenetic tree (Fig. 5). In addition, the ANI of PSI07 and strains RS-T95 (98.32%) and RS-SL2064 (98.30%) were both >95%. The ANI of RS-Rs5 and strains RS-UW251 (97.31%), RS-CIAT078 (96.07%), Po82 (96.09%), and CFBP2957 (97.30%) were all >95%. Moreover, these strains clustered together in the dispensable gene heatmap (Fig. 4b), and occupied a similar position in the phylogenetic tree (Fig. 5). This indicates that ANI, dispensable gene heat maps, and phylogenetic trees all play important roles in deciphering the genetic evolution of species. Furthermore, comparison of plasmids from similar strains revealed that the primary genomic differences between the MLY158 plasmid and the others were located in the phage protein region, which may confer novel virulence traits to MLY158.

5. Conclusions

This study presented the complete genome sequence of *R. pseudosolanacearum* strain MLY158 isolated from infected tobacco plants. Strain MLY158 is able to utilize specific carbon sources, making it easier to infect plants. Functional annotation and comparative genomes provided evidence for the genetic diversity of *R. pseudosolanacearum*. In conclusion, these findings provide a foundation for understanding the underlying mechanisms of bacterial wilt in tobacco and for future studies on the *R. pseudosolanacearum*-tobacco interaction.

Availability of Data and Materials

The MLY158 genome sequence was deposited in the DDBJ/ENA/GenBank database (BioProject/BioSample: PRJNA1271622/SAMN48881384). The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

Conceptualization: JZ, GL, XC; Methodology: JZ, JG, DL, YC, YL; Formal analysis and investigation: BK, KX; Writing—original draft preparation: JZ, GL; Writing—review and editing: YL, XC; Funding acquisition: YL, XC. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The tobacco variety used for the experiment was *Yunyan 85*. The tobacco plants were supplied by the Guizhou Provincial Tobacco Company Zunyi Company Yuqing County Branch. According to China regulations and policies, this study does not involve human or animal subjects; therefore, ethical committee approval is not necessary.

Acknowledgment

We would also like to thank all the colleagues in our group who have supported us.

Funding

This research was supported by the R&S project of China Tobacco Hunan Industrial Co., Ltd. (grant no. KY2025JD0001).

Conflict of Interest

All authors declare no conflicts of interest. Despite they received sponsorship from Guizhou Provincial Tobacco Company Zunyi Company Yuqing County Branch and China Tobacco Hunan Industrial Co., Ltd, the judgments in data interpretation and writing were not influenced by this relationship.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/FBL46230.

References

[1] Mekouar MA. Food and Agriculture Organization of the United Nations (FAO). Yearbook of International Environmental Law. 2023; 34: yvae031. https://doi.org/10.1093/yiel/yvae031.

[2] Castillo JA. Analysis of the speciation process suggests a dual lifestyle in the plant pathogen *Ralstonia solanacearum* species complex. European Journal of Plant Pathology. 2023; 166: 251–257. https://doi.org/10.1007/s10658-023-02670-7.

[3] Tang Y, Zhou M, Yang C, Liu R, Du H, Ma M. Advances in isolated phages that affect *Ralstonia solanacearum* and their application in the biocontrol of bacterial wilt in plants. Letters in Applied Microbiology. 2024; 77: ovae037. https://doi.org/10.1093/lambio/ovae037.

[4] Li X, Huang X, Chen G, Zou L, Wei L, Hua J. Complete genome sequence of the sesame pathogen *Ralstonia solanacearum* strain SEPPX 05. Genes & Genomics. 2018; 40: 657–668. https://doi.org/10.1007/s13258-018-0667-3.

[5] Li N, Sun S, Kong L, Chen Z, Xin Y, Shao R, *et al*. Transcriptomic and physiological analyses reveal plant resistance against *Ralstonia solanacearum* involves salicylic acid-mediated defences in tomato leaves. Plant Pathology. 2025; 74: 123–136. https://doi.org/10.1111/ppa.14002.

[6] Zhang H, Liu Z, Geng R, Ren M, Cheng L, Liu D, *et al*. Genome-wide identification of the TIFY gene family in tobacco and expression analysis in response to *Ralstonia solanacearum* infection. Genomics. 2024; 116: 110823. https://doi.org/10.1016/j.ygeno.2024.110823.

[7] Alariqi M, Wei H, Cheng J, Sun Y, Zhu H, Wen T, *et al*. Large-scale comparative transcriptome analysis of Nicotiana tabacum response to *Ralstonia solanacearum* infection. Plant Biotechnology Reports. 2022; 16: 757–775. https://doi.org/10.1007/s11816-022-00765-x.

[8] Sarkar S, Yadav M, Dey U, Sharma M, Mukhopadhyay R, Kumar A. Exploring the multifaceted role of pehR in *Ralstonia solanacearum* pathogenesis: enzyme activity, motility, and biofilm formation. Microbiological Research. 2025; 290: 127925. https://doi.org/10.1016/j.micres.2024.127925.

[9] Chen K, Wang L, Chen H, Zhang C, Wang S, Chu P, *et al*. Complete genome sequence analysis of the peanut pathogen *Ralstonia solanacearum* strain Rs-P.362200. BMC Microbiology. 2021; 21: 118. https://doi.org/10.1186/s12866-021-02157-7.

[10] Vailleau F, Genin S. *Ralstonia solanacearum*: An Arsenal of Virulence Strategies and Prospects for Resistance. Annual Review of Phytopathology. 2023; 61: 25–47. https://doi.org/10.1146/annurev-phyto-021622-104551.

[11] Xiao Z, Li G, Yang A, Liu Z, Ren M, Cheng L, *et al*. Comprehensive genome sequence analysis of *Ralstonia solanacearum* gd-2, a phylotype I sequevar 15 strain collected from a tobacco bacterial phytopathogen. Frontiers in Microbiology. 2024; 15: 1335081. https://doi.org/10.3389/fmicb.2024.1335081.

[12] Ding S, Ma Z, Yu L, Lan G, Tang Y, Li Z, *et al*. Comparative genomics and host range analysis of four *Ralstonia pseudosolanacearum* strains isolated from sunflower reveals genomic and phenotypic differences. BMC Genomics. 2024; 25: 191. https://doi.org/10.1186/s12864-024-10087-7.

[13] Safni I, Cleenwerck I, De Vos P, Fegan M, Sly L, Kappler U. Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the descriptions of *Ralstonia solanacearum* and *Ralstonia syzygi*i and reclassify cur-

rent *R. syzygii* strains as *Ralstonia syzygii* subsp syzygii subsp nov., *R. solanacearum* phylotype IV strains as *Ralstonia syzygii* subsp indonesiensis subsp nov., banana blood disease bacterium strains as *Ralstonia syzygii* subsp celebesensis subsp nov and *R. solanacearum* phylotype I and III strains as *Ralstonia pseudosolanacearum* sp. nov. International Journal of Systematic and Evolutionary Microbiology. 2014; 64: 3087–3103. https://doi.org/10.1099/ijs.0.066712-0.

[14] Chen W, Zhang JW, Qin BX, Xie HT, Zhang Z, Qiao XZ, *et al*. Quantitative detection of the *Ralstonia solanacearum* species complex in soil by qPCR combined with a recombinant internal control strain. Microbiology Spectrum. 2023; 11: e0021023. https://doi.org/10.1128/spectrum.00210-23.

[15] Crowley E, Bird P, Fisher K, Goetz K, Boyle M, Benzinger MJ, Jr, *et al*. Evaluation of the VITEK 2 Gram-negative (GN) microbial identification test card: collaborative study. Journal of AOAC International. 2012; 95: 778–785. https://doi.org/10.5740/jaoacint.cs2011_17.

[16] Munoz-Dávila MJ, Yagüe G, Albert M, García-Lucas T. Comparative evaluation of Vitek 2 identification and susceptibility testing of Gram-negative rods directly and isolated from BacT/ALERT-positive blood culture bottles. European Journal of Clinical Microbiology & Infectious Diseases. 2012; 31: 663–669. https://doi.org/10.1007/s10096-011-1356-1.

[17] Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Arlat M, *et al*. Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature. 2002; 415: 497–502. https://doi.org/https://doi.org/10.1038/415497a.

[18] Gonçalves OS, Campos KF, de Assis JCS, Fernandes AS, Souza TS, do Carmo Rodrigues LG, *et al*. Transposable elements contribute to the genome plasticity of *Ralstonia solanacearum* species complex. Microbial Genomics. 2020; 6: e000374. https://doi.org/10.1099/mgen.0.000374.

[19] Ingel B, Caldwell D, Duong F, Parkinson DY, McCulloh KA, Iyer-Pascuzzi AS, *et al*. Revisiting the source of wilt symptoms: X-Ray microcomputed tomography provides direct evidence that *Ralstonia* biomass clogs xylem vessels. PhytoFrontiers. 2022; 2: 41–51. https://doi.org/10.1094/PHYTOFR-06-21-0041-R.

[20] Liu Q, Li C, Zhang X, Ding M, Liao X, Yan J, *et al*. PhcX Is a LqsR-family response regulator that contributes to *Ralstonia solanacearum* virulence and regulates multiple virulence factors. mBio. 2023; 14: e0202823. https://doi.org/10.1128/mbio.02028-23.

[21] Huang M, Tan X, Song B, Wang Y, Cheng D, Wang B, *et al*. Comparative genomic analysis of *Ralstonia solanacearum* reveals candidate avirulence effectors in HA4-1 triggering wild potato immunity. Frontiers in Plant Science. 2023; 14: 1075042. https://doi.org/10.3389/fpls.2023.1075042.

[22] Cheng D, Zhou D, Wang Y, Wang B, He Q, Song B, *et al*. *Ralstonia solanacearum* type III effector RipV2 encoding a novel E3 ubiquitin ligase (NEL) is required for full virulence by suppressing plant PAMP-triggered immunity. Biochemical and Biophysical Research Communications. 2021; 550: 120–126. https://doi.org/10.1016/j.bbrc.2021.02.082.

[23] Xian L, Yu G, Wei Y, Rufian JS, Li Y, Zhuang H, *et al*. A Bacterial Effector Protein Hijacks Plant Metabolism to Support Pathogen Nutrition. Cell Host & Microbe. 2020; 28: 548–557.e7. https://doi.org/10.1016/j.chom.2020.07.003.

[24] Albuquerque GMR, Souza EB, Silva AMF, Lopes CA, Boiteux LS, Fonseca MEDN. Genome Sequence of *Ralstonia pseudosolanacearum* Strains with Compatible and Incompatible Interactions with the Major Tomato Resistance Source Hawaii 7996. Genome Announcements. 2017; 5: e00982-17. https://doi.org/10.1128/genomeA.00982-17.

[25] Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics. 2007; 23: 673–679. https://doi.org/10.1093/bioinformatics/btm009.

[26] Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Research. 2007; 35: 3100–3108. https://doi.org/10.1093/nar/gkm160.

[27] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research. 1997; 25: 955–964. https://doi.org/10.1093/nar/25.5.955.

[28] Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, *et al*. Rfam: updates to the RNA families database. Nucleic Acids Research. 2009; 37: D136–D140. https://doi.org/10.1093/nar/gkn766.

[29] Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research. 1999; 27: 573–580. https://doi.org/10.1093/nar/27.2.573.

[30] Galperin MY, Wolf YI, Makarova KS, Vera Alvarez R, Landsman D, Koonin EV. COG database update: focus on microbial diversity, model organisms, and widespread pathogens. Nucleic Acids Research. 2021; 49: D274–D281. https://doi.org/10.1093/nar/gkaa1018.

[31] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics. 2000; 25: 25–29. https://doi.org/10.1038/75556.

[32] Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Research. 2016; 44: D457–D462. https://doi.org/10.1093/nar/gkv1070.

[33] Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, *et al*. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Research. 2016; 44: D286–D293. https://doi.org/10.1093/nar/gkv1248.

[34] Liu B, Pop M. ARDB–Antibiotic Resistance Genes Database. Nucleic Acids Research. 2009; 37: D443–D447. https://doi.org/10.1093/nar/gkn656.

[35] Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis–10 years on. Nucleic Acids Research. 2016; 44: D694–D697. https://doi.org/10.1093/nar/gkv1239.

[36] Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Research. 2009; 37: D233–D238. https://doi.org/10.1093/nar/gkn663.

[37] Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, *et al*. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Research. 2020; 48: D517–D525. https://doi.org/10.1093/nar/gkz935.

[38] Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature Communications. 2018; 9: 5114. https://doi.org/10.1038/s41467-018-07641-9.

[39] Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, *et al*. Versatile and open software for comparing large genomes. Genome Biology. 2004; 5: R12. https://doi.org/10.1186/gb-2004-5-2-r12.

[40] Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012; 28: 3150–3152. https://doi.org/10.1093/bioinformatics/bts565.

[41] Lowe-Power TM, Hendrich CG, von Roepenack-Lahaye E, Li B, Wu D, Mitra R, *et al*. Metabolomics of tomato xylem sap during bacterial wilt reveals *Ralstonia solanacearum* produces abundant putrescine, a metabolite that accelerates wilt disease. Environmental Microbiology. 2018; 20: 1330–1349. https://doi.org/10.1111/1462-2920.14020.

[42] Georgoulis SJ, Shalvarjian KE, Helmann TC, Hamilton CD, Carlson HK, Deutschbauer AM, *et al*. Genome-Wide Identification of Tomato Xylem Sap Fitness Factors for Three Plant-Pathogenic *Ralstonia* Species. mSystems. 2021; 6: e0122921. https://doi.org/10.1128/mSystems.01229-21.

[43] Wang Z, Zhang Y, Bo G, Zhang Y, Chen Y, Shen M, *et al*. *Ralstonia solanacearum* Infection Disturbed the Microbiome Structure Throughout the Whole Tobacco Crop Niche as Well as the Nitrogen Metabolism in Soil. Frontiers in Bioengineering and Biotechnology. 2022; 10: 903555. https://doi.org/10.3389/fbioe.2022.903555.

[44] Chen K, Zhuang Y, Wang L, Li H, Lei T, Li M, *et al*. Comprehensive genome sequence analysis of the devastating tobacco bacterial phytopathogen *Ralstonia solanacearum* strain FJ1003. Frontiers in Genetics. 2022; 13: 966092. https://doi.org/10.3389/fgene.2022.966092.

[45] Tan X, Dai X, Chen T, Wu Y, Yang D, Zheng Y, *et al*. Complete Genome Sequence Analysis of *Ralstonia solanacearum* Strain PeaFJ1 Provides Insights Into Its Strong Virulence in Peanut Plants. Frontiers in Microbiology. 2022; 13: 830900. https://doi.org/10.3389/fmicb.2022.830900.

[46] Geng R, Cheng L, Cao C, Liu Z, Liu D, Xiao Z, *et al*. Comprehensive Analysis Reveals the Genetic and Pathogenic Diversity of *Ralstonia solanacearum* Species Complex and Benefits Its Taxonomic Classification. Frontiers in Microbiology. 2022; 13: 854792. https://doi.org/10.3389/fmicb.2022.854792.

[47] Sharma P, Johnson MA, Mazloom R, Allen C, Heath LS, Lowe-Power TM, *et al*. Meta-analysis of the *Ralstonia solanacearum* species complex (RSSC) based on comparative evolutionary genomics and reverse ecology. Microbial Genomics. 2022; 8: 000791. https://doi.org/10.1099/mgen.0.000791.

[48] Kai K, Ohnishi H, Shimatani M, Ishikawa S, Mori Y, Kiba A, *et al*. Methyl 3-Hydroxymyristate, a Diffusible Signal Mediating phc Quorum Sensing in *Ralstonia solanacearum*. Chembiochem. 2015; 16: 2309–2318. https://doi.org/10.1002/cbic.201500456.

[49] Prior P, Ailloud F, Dalsing BL, Remenant B, Sanchez B, Allen C. Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. BMC Genomics. 2016; 17: 90. https://doi.org/10.1186/s12864-016-2413-z.