Microbiological Research 218 (2019) 32–40

Contents lists available at ScienceDirect

Microbiological Research

journal homepage: www.elsevier.com/locate/micres

Genome sequence and comparative genomics of *Rhizobium sp.* Td3, a novel plant growth promoting phosphate solubilizing *Cajanus cajan* symbiont

Bhagya Iyer, Shalini Rajkumar*

Institute of Science, Nirma University, Ahmedabad, 382 481, Gujarat, India

A R T I C L E I N F O

Keywords:
Rhizobia
P solubilization
Gluconic acid
Plant growth promotion

A B S T R A C T

Rhizobium sp. Td3 is a *Sesbania* plant growth promoting, *Cajanus cajan* nodulating rhizobia. Studying its whole genome was important as it is a potent phosphate solubilizer with constitutive gluconic acid production ability through operation of the periplasmic glucose oxidation pathway even under conditions of catabolite repression. This is in contrast to the other explored phosphate solubilizing Rhizobial isolates sequenced so far are known to lack components of the direct glucose oxidation pathway and [illegible] not produce gluconic acid on its own. Here, we present the genome sequence of *Rhizobium sp.* Td3. Genome comprises of a single chromosome of size 5,606,547 bp (5.6 Mb) with no symbiotic plasmid. *Rhizobium leguminosarum bv. viciae* USDA2370 was the closest whole genome known. 109 genes responsible for diverse plant growth promoting activities like P solubilization, synthesis of acetoin, nitric oxide, indole-3 acetic acid, exopolysaccharide, siderophore and trehalose have been identified. Flagellar proteins, genes e[illegible] [illegible]otic and metal resistance, enzymes required for combating oxidative stress as w[illegible] [illegible]chment and colonization in the plant rhizosphere are also present. Availability of genome sequence of [illegible] a [illegible] plant growth promoting agent will help in exploiting all the phyto-beneficial traits of Td3 for [illegible] [illegible]izer.

1. Introduction

Advances in DNA sequencing technology have dramatically changed the strategies for studying genetic [illegible] of various organisms. Complete genomic sequences not only provide the information obligatory to perform functional analysis of the [illegible]es but also furnishes new insights into gene function, gene evolution and genome evolution. Development of ultra-high throughput sequencing (UHTS) has been exceedingly influential [illegible] advancing research in all scientific areas, particularly in microbiology [illegible]re genomes are small (Bertelli and Greub, 2013). Fu[illegible] the a[illegible]ceme[illegible] and falling prices of sequencing techniques have [illegible]elerated [illegible] availability of genomic data of many bacte[illegible] species [illegible] comparison.

Biologi[illegible] [illegible] fixat[illegible] [illegible] nodule forming bacteria in symbiosis with legumes [illegible] significant role in agricultural systems as it provides N source [illegible] the legumes and improves legume growth and agricultural product[illegible]. In a strict sense, the term "rhizobia" refers to the members of genus *Rhizobium,* taxonomically, metabolically and genetically diverse N fixing legume nodulating soil bacteria (Boivin et al., 2009). Majority of these bacterial species belong to the genera *Rhizobium*, *Mesorhizobium*, *Ensifer* and *Bradyrhizobium* of *Rhizobiaceae* family in the class alpha proteobacteria. Recent research has shown the presence of many other new rhizobial species through lateral transfer of symbiotic genes. Members of *Phyllobacterium, Microvirga* (Ardley et al., 2012)*, Ochrobactrum* (Pineiro et al., 2007)*, Methylobacterium, Devosia* (Rivas et al., 2003) and *Shinella* (Lin et al., 2008) belonging to alpha proteobacteria and *members of Burkholderia,* and *Cupravidus* (Liu et al., 2011) belonging to beta proteobacteria have also been incorporated in rhizobia. The term "rhizobia" is now used for all the bacteria capable of N fixation and nodulation (Willems, 2006) which may either grow in soil as free living organisms or as N fixing symbionts within the root nodule. Use of rhizobia as a N biofertilizer is a well known agronomic practice to improve the growth of legumes and non legumes (Fraile et al., 2012) and is obviously a superior alternative to the application of harmful N fertilizers.

Whole genome information about plant growth promoting and phosphate solubilizing *Paraburkholderia tropica* (Kaur et al., 2016), *Pseudomonas spp.* (Duan et al., 2013; Kwak et al., 2016), etc. are already available. However, in order to more completely exemplify the genetics of plant growth promotion and phosphate solubilization in rhizobia, *Rhizobium sp.* Td3 was fully sequenced. Td3 could solubilize 423 μg/mL phosphate from tricalcium phosphate and 23 μg/mL from rock phosphate under 50 mM buffering conditions through production of 19.2 mM gluconic acid (manuscript under consideration). We hereby

---

* Corresponding author.
E-mail address: shalini.rjk@nirmauni.ac.in (S. Rajkumar).

https://doi.org/10.1016/j.micres.2018.09.007
Received 19 June 2018; Received in revised form 21 August 2018; Accepted 21 September 2018

0944-5013

Table 1
Genome characteristics of *Rhizobium sp.* Td3 and other rhizobia.

Organism	Genome size (Mb)	G + C content (%)	Protein coding genes	rRNA coding genes	tRNA coding genes
Td3	5.6	57.1	5448	6	44
Rhizobium sp. NGR234	3.9	63.03	3630	9	52
R. leguminosarum bv. trifolii WSM1325	4.7	61.09	4565	9	51
Sinorhizobium meliloti 1021	3.6	62.02	3359	9	51
R. etli CFN 42	4.4	61.27	4031	9	45
R. leguminosarum bv. viciae 3841	5.0	61.09	4694	9	52
Mesorhizobium sp. BNC1	4.4	61.07	4.64	6	49
M. loti MAFF303099	4.4	62.9	6743	6	50

report nucleotide sequence of phosphate solubilizing *Cajanus cajan* symbiont Td3; which has 5.6Mb genome and constitutively produces gluconic acid through operation of a periplasmic glucose oxidation pathway. Rhizobia sequenced so for are known to lack either periplasmic glucose dehydrogenase (Gcd) or pyrroloquinoline quinone (PQQ) biosynthesis genes making them dependent on PQQ contributed by the root exudates or other soil bacteria for periplasmic glucose oxidation and subsequent organic acid production (Bernardelli et al., 2001). Rhizobial isolate Td3 is novel as it bears genes responsible for PQQ biosynthesis along with multiple copies of periplasmic PQQ dependent glucose dehydrogenase, gluconate-2 dehydrogenase and 2-keto D-gluconate dehydrogenases. In addition, it also exhibits multiple plant beneficial traits like synthesis of acetoin, nitric acid, indole-3 acetic acid, exopolysaccharide, siderophore, trehalose, genes encoding antibiotic and metal resistance, attachment and colonization as well as enzymes required for combating oxidative stress.

2. Materials and methods

2.1. Strain and DNA library preparation

Rhizobium sp. Td3 was used as the DNA donor a[illegible]nic DN[illegible] was extracted using NEXTFlex DNA sequencing kit[illegible]nomi[illegible]NA was sheared to generate fragments of approximate[illegible]e 200-4[illegible]p using Covaris S220 system and was purified using High[illegible] b[illegible], [illegible] by end repair, A-tailing and ligation o[illegible]e Illu[illegible] multiplexing adaptors. Adaptor ligated DNA was pu[illegible]d using High[illegible] beads and subjected to 10 rounds of PCR (dena[illegible]t 98 °C for 2[illegible]n, cycling (98 °C for 30 s, 65 °C for 30 s and 72[illegible]C for [illegible]and final extension at 72 °C for 4 min) with Illumin[illegible]ompatible pri[illegible] PCR product was purified using HighPrep be[illegible], followed by quan[illegible]ication with Qubit fluorometer followed by[illegible]gment[illegible] distribution analysis on Agilent D1000 TapeStation.

2.2. Genome se[illegible]

The ge[illegible]e of [illegible]sp.Td3 was sequenced using Illumina NextSeq500 p[illegible]d sequencing with 150*2 paired end reads.

2.3. Sequence analy[illegible]ene prediction and annotation

Quality of Illumina paired end reads were checked using FastQC. Denovo assembly of Illumina NextSeq data was performed using SPAdes (Bankevich et al., 2012) assembler and scaffolding of the assembled contigs was carried out using SSPACE (Boetzer et al., 2011) program. For gene prediction and gene annotation, a multidatabase approach was followed using RAST server (Aziz et al., 2008), BLAST (GO) (Altschul et al., 1990), KASS (Moriya et al., 2007) and MISA. Assembled contigs were uploaded to RAST server for annotation. Predicted proteins were searched for similarity against Uniprot *Rhizobium* protein database for GO annotation using BLASTP (Altschul et al., 1990) program. Pathway analysis was done by using KAAS (Moriya et al. 2007) server and repeats were identified in each contig sequence using MISA perl script. Phylogenetic t[illegible]was[illegible]tructed with the closest hits using Phylogeny (Dereepe[illegible] al., 2008).

Accession numbers:

Rhizobium sp. Td3 has been s[illegible]mitte[illegible]NCMR (N[illegible]onal Centre for Microbial Resource), Pune u[illegible] the access[illegible]num[illegible]r MCC3648. This whole genome project ha[illegible]een de[illegible]ited t[illegible]BI BioProject under accession number PRJN[illegible]4840[illegible]

3. Results and [illegible]cussion

3.1. Genom[illegible]equen[illegible]sembly statistics and phylogeny

[illegible] presented he[illegible]e genome sequence of *Rhizobium sp.* Td3, a [illegible]osphate solubilizing *Sesbania rostrata* plant growth promoting and *C[illegible]nus cajan* no[illegible]ating rhizobia. A total of 7,595,422 reads were ass[illegible]led into 81[illegible]ontigs using SPAdes. The strain had 57.1% G + C con[illegible]with a[illegible]romosome size of 5,606,547 bp (5.6 Mb). Maximum contig [illegible]was 8,88,656 bp with N50 value 4,08,102. Contigs ob[illegible]ined were mapped and compared to the chromosomes of multiple [illegible]e rhizobial strains to deduce the closest known genome for further analysis. Table 1 indicates the genome characteristics of Td3 and its comparison to other rhizobial genomes. As 16S rRNA gene sequences are highly conserved among the same bacterial species, they are recurrently used to identify and categorize microorganisms. Td3 16S rRNA gene sequence showed highest similarity to that of *Rhizobium sp.* IRBG74, a member of *Rhizobium/Agrobacterium* group as shown in Fig. 1. However, as the number of rRNA genes in prokaryotic genomes can vary from one to as many as fifteen copies and the intragenic diversity ranges from 0.06% to 20.38% (Pei et al., 2010), they may not clearly reflect the exact phylogeny. Many studies have shown that "housekeeping” genes should be used to assist bacterial species classification (Peix et al., 2007; Duan et al., 2013). Phylogeny based on homology of housekeeping genes *dnaK* (chaperone protein), *recA* (DNA

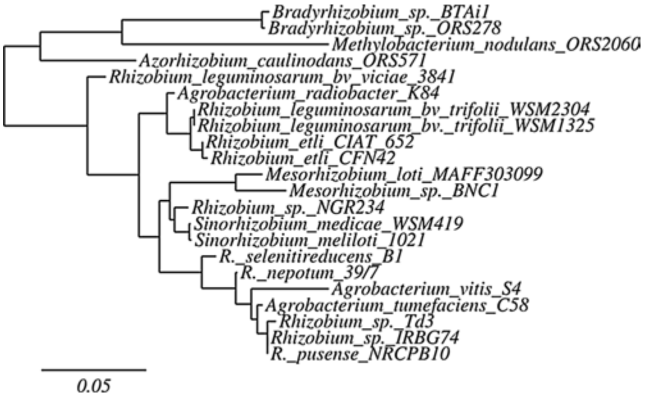

Fig. 1. Phylogenetic tree of Td3 16S rRNA gene homology. Td3 16s rRNA gene sequence was similarity searched using NCBI blast program. Top 19 hits were then considered for the generation of a phylogenetic tree using Phylogeny. Organisms belonging to the same genus have been represented with a similar colour code. Value 0.01 indicates number of nucleotide substitutions per site.

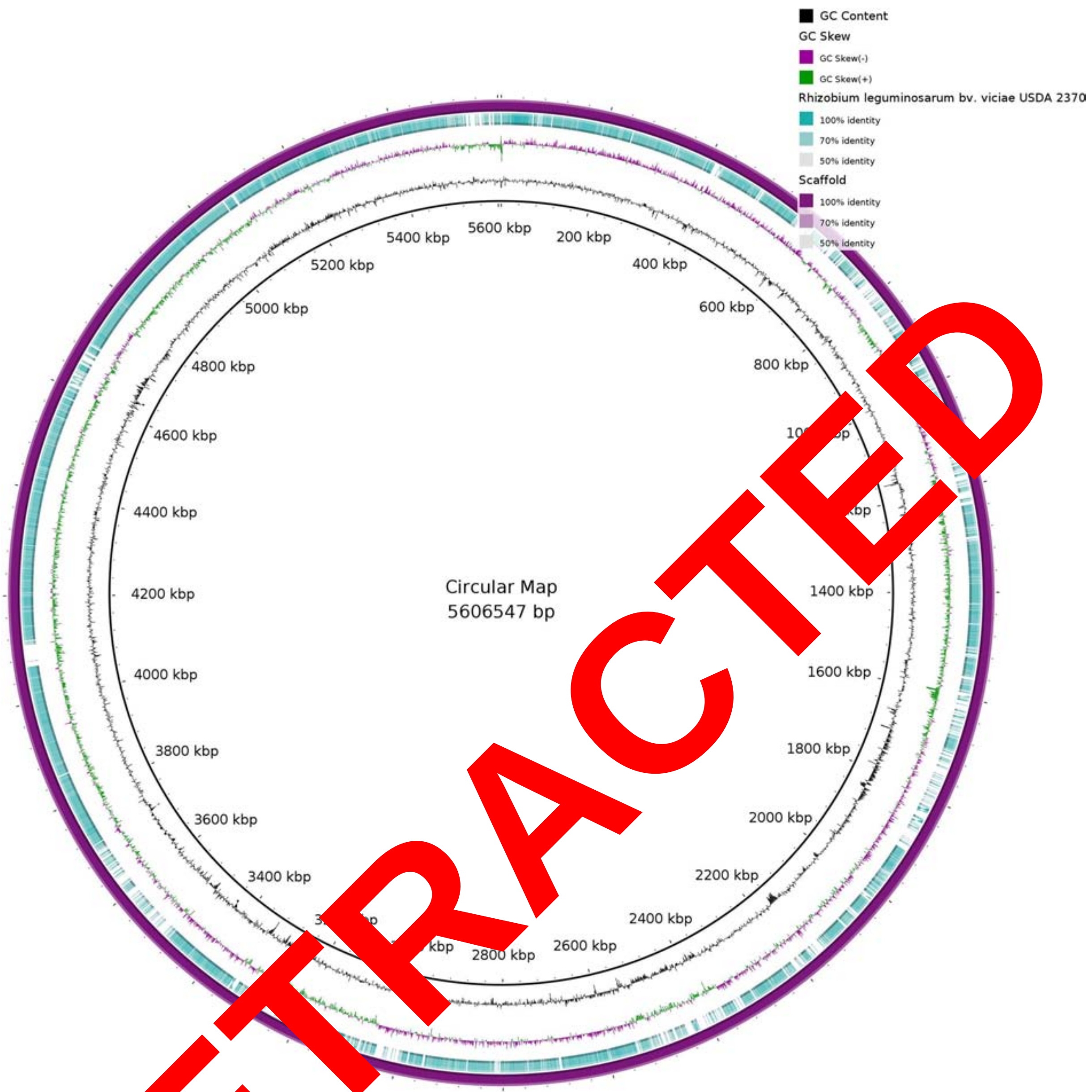

Fig. 2. Circular Genome Map. The assembled scaffolds were aligned against the reference genome *R. leguminosarum bv. viciae* USDA2370 genome. Purple lines indicate the GC skewness of [illegible] positive strand. Green lines indicate the GC skewness of the negative strand while the black circle indicates the GC content. (For interpretation of the references [illegible] colour in this figure legend, the reader is referred to the web version of this article).

recombinase ([illegible]) and *gl*[illegible] (glutamine synthetase) demonstrated that the isolate was closest to [illegible]F303099. Recent advances in genetic methods offers [illegible] increasing number of genomic sequences per species. ANI, average nucleotide identity calculated from pair-wise comparison of all sequences pooled between two strains has been proposed as a new means for bacterial species characterization and classification (Han et al., 2016). Td3 showed 78.26% and 78.18% ANI with *R. leguminosarum* bv. *trifolii* WSM1325 and *R. etli* CFN42 respectively. *Rhizobium leguminosarum bv. viciae* USDA2370 genome was the closest to Td3 followed by *R. leguminosarum* bv. *trifolii* WSM1325 and *R. leguminosarum* bv. *viciae* 3841. Fig. 2 depicts the circular genome map of Td3 with its closest genome.

3.2. Gene prediction and functional annotation

Proteins were annotated against all *Rhizobium* sequences from UniProt database and protein sequences with more than 30% identity were taken for further analysis. Td3 genome encoded 5448 protein coding genes, 44 tRNAs and 6 rRNAs. Of the 5448 protein coding genes, 4976 (91.3%) were annotated proteins while remaining 472 (8.7%) were unannotated proteins. As shown in Fig. 3, 36.69% of coding genes had molecular function whereas 40.9% and 19.43% genes were shown to code for cellular component or involved in a biological process respectively. Functional annotation of the genome sequence was done automatically using RAST server with all the functionally characterized genes distributed under 480 subsystems.

For comparative genomic analysis, one member of each rhizobial species with significant similarity to Td3 housekeeping genes (*dnaK*, *recA* and *glnA)* were selected. Table 2 shows the comparison of different subsystem categories of Td3 with *Sinorhizobium meliloti* 1021, *Mesorhizobium loti* MAFF303099, *Rhizobium leguminosarum bv. viciae* 3841, *Bradyrhizobium japonicum* USDA110 and *Azorhizobium caulinodans* ORS571. Table S1 compares the subsystem features and counts. Compared to the other genomes taken into consideration, Td3 genome had

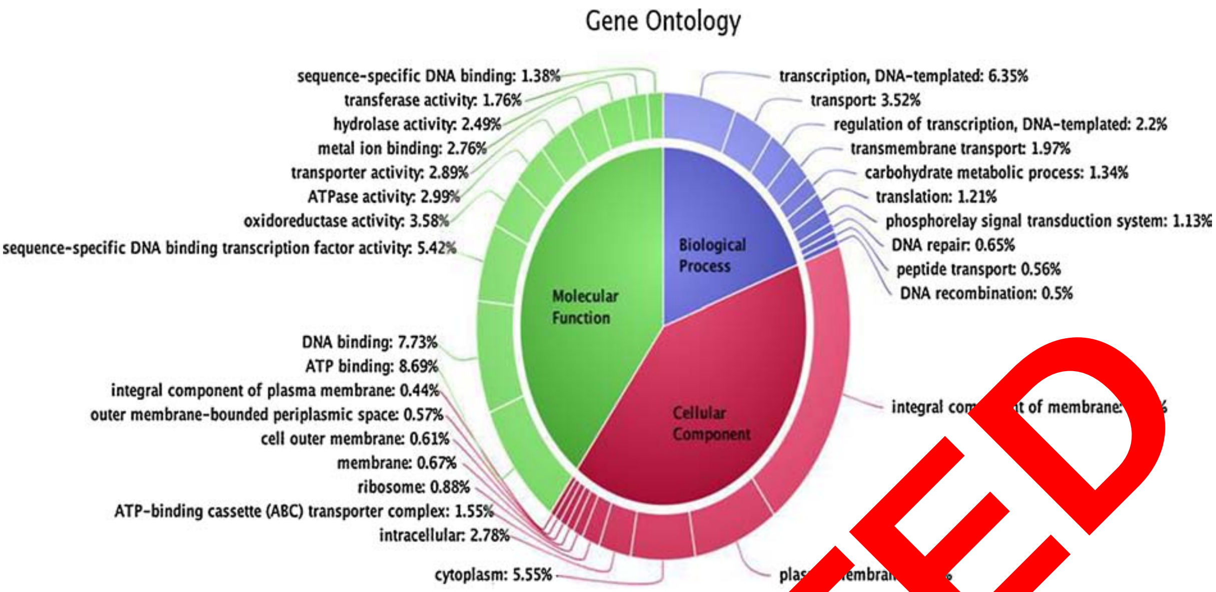

Fig. 3. Gene ontology characterization of Td3. All the coding genes have been classified into three [illegible]ies; those inv[illegible] biological process, those having a molecular function and those that are a part of cellular component. Top ten gene ontologies of e[illegible] of t[illegible]ree categorie[illegible]are indicated in the figure. Percentage values indicate the fraction of genes involved in a particular function from the total number of protein cod[illegible]enes.

less number of genes coding for synthesis of cofactors, vitamins, prosthetic groups and pigments; potassium metabolism; RNA metabolism; DNA metabolism; N metabolism; respiration and carbohydrate metabolism. Td3 genome encoded for hefty number of genes coding for membrane transport; iron acquisition and metabolism; cell divisio[illegible] cycle; motility and chemotaxis; fatty acid, lipids and isopre[illegible]ds synthesis; sulfur metabolism and stress response. Genome encoded [illegible]a to[illegible] of 1184 hypothetical proteins, 35 response regulator proteins and 2[illegible] transcriptio[illegible] regulators scattered among varied families. Family w[illegible] distribution [illegible] all the classes of transcriptional regulators encoded by [illegible] genom[illegible]d their general functions are shown in Table S2. Of the 2[illegible]iptional regulators, 86 regulators belonged to LysR fa[illegible]ily, followed by GntR, AraC and TetR families which were 34, 25 and [illegible]umber. 13 other transcriptional regulators were identified which

Table 2
Comparison of different subsystem categories of Td3 wi[illegible]other rh[illegible]bia.

	Num[illegible]volved in rhizobial subsystem					
Subsystem features	[illegible]	Rm1021	*M. loti*	*R. leg*	*B. jap*	*A. cau*
Cofactors, Vitamins, Prosthetic groups, Pig[illegible]	293	309	369	306	380	243
Cell wall and capsule	126	128	148	116	145	108
Virulence, Disease and Defense	88	92	86	94	102	57
Potassium metabolism	19	47	36	39	32	31
Photosynthesis	0	0	0	0	0	0
Miscellaneous	67	83	59	78	71	47
Phages, Prophages, Transpo[illegible]elem[illegible] Plasmids	7	0	14	18	0	17
Membrane transport	286	198	166	192	246	74
Iron acquisition an[illegible]olism	70	48	36	35	22	33
RNA metabolism	148	188	199	197	197	132
Nucleosides a[illegible]ucleoti[illegible]	147	155	160	149	133	100
Protein M[illegible]ism	257	254	244	251	269	225
Cell divisio[illegible]Cell[illegible]	36	29	28	32	27	24
Motility and C[illegible]s	122	61	55	80	106	58
Regulation and Ce[illegible]alling	81	91	83	120	95	52
Secondary metabolis[illegible]	4	4	6	5	17	7
DNA metabolism	124	154	127	164	145	74
Fatty acids, Lipids and Isoprenoids	147	103	127	124	234	102
Nitrogen metabolism	21	74	39	40	109	68
Dormancy and Sporulation	1	1	1	1	4	1
Respiration	139	239	206	198	280	210
Stress Response	177	184	155	163	166	128
Metabolism of Aromatic compounds	51	53	41	74	139	53
Amino acids and derivatives	557	511	540	513	609	403
Sulfur metabolism	81	41	43	61	98	64
Phosphorous metabolism	52	55	78	76	51	54
Carbohydrates	533	589	541	688	590	282
Total	3634	3691	3587	3814	4267	2647

Td3: Rhizobium sp.Td3Rm1021: *Sinorhizobium meliloti* 1021.
M. *loti: Mesorhizobium loti* MAFF303099R. leg: *Rhizobium leguminosarum bv. viciae* 3841.
B. *jap: Bradyrhizobium japonicum* USDA110A. *cau: Azorhizobium caulinodans* ORS571.

were not categorized into any of the known distinct classes.

3.3. Membrane transport, protein secretion systems and cytoskeletal structures

Td3 genome is rich in genes encoding ABC transporters dedicated for the transport of alkylphosphonates, branched chain amino acids, dipeptides and oligopeptides as well as Ton and Tol transport systems. Superior number of genes are dedicated for macromolecular synthesis (MMS) operon with nine variants of genes being involved. Genes coding for type II protein secretion system are highly conserved among gram negative bacteria and are an important determinant of bacterial virulence. Genes for type VII protein secretion system; imperative for interaction with the host (Tseng et al., 2009) and multisubunit cation antiporter shown to be a necessary component for proper invasion into plant root tissue (Yamaguchi et al., 2009) are also present in Td3. Cytoskeletal structures play imperative roles in cell division, cell polarity, cell shape regulation and plasmid partition (Graumann, 2004). Prokaryotic cells contain homologs for each of the three major groups of eukaryotic cytoskeletal proteins, i.e., actin, tubulin and intermediate filaments. In addition to these three, they have a fourth group, the MinD-ParA group, that appears to be unique to bacteria (Shih and Rothfield, 2006). Like *Neisseria,* Td3 has genes for actin, tubulin and intermediate filaments, but lacks genes belonging to MinD-ParA group.

3.4. Regulation and cell signalling

Compared to other bacterial species, Td3 has high number of cAMP signalling genes, but as compared to other rhizobia, they are few. It also a has putative hemin-binding lipoprotein and a clp protease. cAMP is a ubiquitous messenger integrating many cellular processes. Various families of adenylate cyclase and phosphodiesterase meticulously regulate the intracellular cAMP concentration and any variation in the cytosolic cAMP concentration has a weighty effect on the various cellular processes. Disruption of these cellular processes involves often the most critical event in pathogenesis as many pathogenic bacteria secrete toxins to modify the intracellular cAMP concentration (A[illegible] et al., 2004). Td3 genome also bears genes for toxin-antitoxin replicon stabilizing systems which consists of a stable toxin and [illegible] small, labile antitoxin. Under unfavourable conditions, the antitoxin [illegible] degraded, leading to activation of the toxin, resulting [illegible] growth arrest and possibly bacterial programmed cell death. They also act as stress response elements, helping the cells to survive unfavourable growth conditions (Pandey and Gerdes, 2005). Compared to other rhizobia, Td3 codes large number of genes for orphan regulatory proteins, LysR regulatory proteins, zinc regulated en[illegible] murein hydrolase regulation and cell death.

3.5. DNA metabolism

In addition [illegible] bacterial MutL-MutS and UvrABC DNA repair system common to other rhizobia, Td3 genome bears genes for RecA and a hypothetical protein based DNA repair machinery. However, RecX which is known to regulate RecA activity and often included in this cluster is lacking. It lacks DNA topoisomerase III. Td3 has multiple copies of ATP-dependent DNA ligase and Ku domain protein, absent in most bacteria. DNA methyltransferase subunit M (HsdM) and speficity subunit (HsdS) of type I restriction modification system are present, but restriction subunit (HsdR) is lacking. Unlike most bacterial species, Td3 bears hyp1, a protein involved in catabolism of external DNA.

3.6. Amino acid derivatives

Unlike other rhizobia taken for analysis, Td3 has glutamate and aspartate uptake machinery, urea carboxylase and allophanate hydrolase cluster, lysine biosynthesis DAP pathway and HMG CoA synthesis machinery. It also bears genes for aspartate ammonia lyase, L-asparaginase, glutaminase, glutamate racemase, leucine responsive regulatory protein and aspartate racemase, lacking in most bacteria. Among genes required for glutamate and aspartate uptake, Td3 has GltI (glutamate-aspartate transporter) and GltS (sodium/glutamate symport protein) (Reizer et al., 1994), lacks GltP (glutamate-aspartate carrier protein) (Raunser et al., 2006) and has a sigma-54 specific central transcriptional regulator of acidic amino acid uptake.

3.7. Phosphate transport and metabolism

Inorganic phosphate (Pi), the preferred source of phosphorus (P), is often limiting in the environment. Multiple [illegible] induced under Pi limiting conditions comprise the PHO operon involved in P assimilation and are coregulated by extracellular phosphate concentration. For phosphate transport, Td3 has PstS, a periplasmic Pi binding protein; PstACB, high-affinity inner membrane Pi uptake ABC transporter and PhoU, phosphate transport system regulatory protein. Repression of PHO regulon under conditions of Pi excess (above 4 microM) requires the PstSABC system and [illegible] present in m[illegible] bacteria. In addition to these genes, Td3 also has histidine kinase PhoR, response regulator PhoB for induction of PHO regulon under Pi-limiting conditions (Beneit, 201[illegible]).

Td3 bears genes [illegible] polyphosphate kinase, inorganic pyrophosphatase, exopolyphosphatase, secreted alkaline phosphatase, phosphate starvation inducible protein PhoH, low affinity inorganic phosphate transporter and sodium dependent phosphate transporter facilitating inorganic and organic phosphate utilization. In addition, Td3 also has the machinery for alkylphosphonate utilization. PhnG, PhnH, PhnI, PhnJ, PhnK, PhnL (phosphonate transport ATP binding protein), PhnM (metal dependent hydrolase involved in phosphonate metabolism) and PhnP (metal dependent hydrolase of the β-lactamase superfamily) are thought to comprise a multisubunit COP lyase. It also has PhnO, PhnF, PhnN (ATP binding protein), PhnA (alkylphosphonate utilization operon protein), PhnB and rcsF.

3.8. Sulfur metabolism

Td3 encodes genes for inorganic sulfur assimilation, utilization of glutathione as a sulfur source as well as machinery for alkane sulphonate utilization and assimilation. Td3 has CysAWTp/Sbp ABC-type heteromeric transporter which is a sulfate and thiosulfate binding protein. Similar to *E. coli,* the transporter can use both APS (adenosine phosphosulfate) and PAPS (phosphoadenosine phosphosulfate) as sulfite precursors by reducing them via sulfite reductase flavoprotein and hemoprotein (CysJI). For utilizing glutathione as a S source, Td3 has a glutathione ABC transporter and a gamma-glutamyltranspeptidase. Td3 has complete machinery for utilization and assimilation of alkane sulphonates but lacks organosulfonate utilization protein SsuF.

3.9. Genes conferring plant growth promotion traits in Td3

3.9.1. PQQ biosynthesis and periplasmic glucose oxidation for phosphate solubilization

Td3 has numerous genes responsible for multiple plant growth promotion traits and Table 3 lists the genes bestowing such properties. Td3 is a potent phosphate solubilizer and genome analysis revealed that it codes for exceptionally large number of genes involved in D-gluconate and ketogluconate metabolism. The ability of Td3 to solubilize mineral phosphates and zinc has been attributed to its ability to directly oxidize glucose in the periplasm to the corresponding acids by means of periplasmic dehydrogenases. Bioinformatic analysis revealed that, among all the rhizobial genomes analyzed, Td3 is the only rhizobia bearing all the genes involved in periplasmic glucose oxidation pathway along with PQQ biosynthesis genes. Five genes to be responsible for PQQ biosynthesis were identified in rhizobia namely

Table 3
Genes contributing to plant growth promoting traits in Td3.

Plant growth promotion trait	Gene	Gene function
Phosphate solubilization	*pqqB*	PQQ biosynthesis protein
	pqqC	PQQ biosynthesis protein
	pqqD	PQQ biosynthesis protein
	pqqE	PQQ biosynthesis protein
	phoD	Alkaline phosphatase D
	ldhA	D-lactate dehydrogenase
	icd	Isocitrate dehydrogenase
	gltA	Citrate synthase
	gcd	Quinoprotein glucose dehydrogenase
	gad	Gluconate 2- dehydrogenase
	2kgdh	2 ketogluconate dehydrogenase
	phoR	Phosphate regulon sensor protein
	pstS	Phosphate ABC transporter,
	pstC	periplasmic binding protein
	pstA	Phosphate transport system permease
	pstB	Phosphate transport system permease
	phoU	Phosphate transport ATP binding
	phoB	protein
	ppa	Phosphate transport system regulatory
		protein
		Phosphate regulon transcriptional
		regulatory protein
		Inorganic pyrophosphatase
Acetoin synthesis	*budA*	α-acetolactate decarboxylase
	budB	α-acetolactate synthase
	budC	Diacetyl reductase
Nitric oxide synthesis	*nirK*	Nitrite reductase
	nod	Nitric oxide dioxygenase
	nruU	Required for expression of nitric oxide
		and nitrite reductase
Auxin synthesis	*trpA*	Tryptophan synthase alpha chain
	trpB	Tryptophan synthase beta chain
	trpD	Anthranilate
	trpF	phosphoribosyltransferase
	ysnE	Phosphoribosylanthranilate isomerase
		IAA acetyltransferase
Nitrogen fixation	*fixL*	Two component [illegible] sensory
	fixJ	histidine kinase
	nifU	Two component N fixation regulatory
		protein
		N fixation Fe-S cluster scaffold protein
Exopolysaccharide synthesis	*exoD*	Exopolysaccharide synthesis protein
	exoR	[illegible]sacchatide production
	exoZ	negative regulator
	exoL	Exopolysaccharide production protein
	gtf	Succinoglycan biosynthesis protein
		Glycosyl transferase for EPS
		production
Siderophore production	[illegible]	Ferric enterobactin transport system
		permease
	[illegible]	[illegible] dependent siderophore receptor
	fecR	Ferric siderophore transport
	fhuB	periplasmic binding protein
	fhuD	Iron siderophore sensor protein
		Ferric hydroxamate ABC transporter,
		permease
		Ferric hydroxamate ABC transporter,
		periplasmic binding protein
		Ferrichydroxamate ABC transporter,
		ATP binding protein
Flagellar proteins	*flaA*	Flagellin protein
	fliP	Flagellar biosynthesis protein
	fliL	Flagellar biosynthesis protein
	flgH	Flagellar L- ring protein
	flgL	Flagellar P-ring protein
	flgA	Flagellar basal body P-ring formation
	flgG	protein
	fliE	Flagellar basal body rod protein
	flgC	Flagellar hook-basal body complex
	flgB	protein
	fliL	Flagellar basal body rod protein
	flgF	Flagellar basal body rod protein
	motA	Flagellar specific ATP synthase
	fliM	Flagellar basal body rod protein
	fliN	Flagellar motor rotation protein
	fliG	Flagellar motor switch protein
	flhB	Flagellar motor switch protein
	motB	Flagellar motor switch protein
	flgE	Flagellar biosynthesis protein
	flgK	Flagellar motor rotation protein
	flgL	Flagellar hook protein
	flaF	Flagellar hook associated protein
	flgD	Flagellar hook associated protein
	fliQ	[illegible]
	flhA	Flagellar basal body rod modification
	fliR	protein
	flgJ	Flagellar biosynthesis protein
		Flagellar biosynthesis protein
		Flagellar biosynthesis protein
		Flagellar [illegible]
Trehalose synthesis	*treZ*	Malto-oligosyltrehalose
	[illegible]	trehalose hydrolase
	[illegible]	Predicted cobalt transporter
		[illegible]E-UreJ family nickel and cobalt
		transporter
Nickel resistance	*nikA*	Nickel uptake ABC transporter,
	nikB	periplasmic binding protein
		Nickel uptake ABC transporter
		permease
Cobalt resistance	*corA*	Magnesium and cobalt transport
	corC	protein
	cbtA	Magnesium and cobalt efflux protein
	czcD	Predicted cobalt transporter
		Cobalt zinc cadmium resistance
Copp[illegible] protein	*cutE*	Copper homeostasis protein
	–	Copper ABC transporter, ATP binding
	–	component
	–	Copper ABC transporter, permease
		component
		Copper ABC transporter, periplasmic
		binding component
Arsenic resistance	–	Arsenate reductase
	arsH	Arsenic resistance protein
Enzymes important during oxidative stress	–	Peroxidase
	cat	Catalase
	sod	Super oxide dismutase
	gsh	Glutathionetransferase
Antibiotic resistance	*bla*	Beta lactamase
	pbp	Penicillin binding protein
Attachment and colonization in the plant rhizosphere	*fimA*	Major pilin protein
	virB5	Minor pilin of type IV secretion
	virB2	complex
	virB3	Major pilus subunit of type IV
	virB4	secretion complex
	flp	Type IV secretion protein channel
	tadD	protein
	tadC	Type IV secretion complex ATPases
	tadB	Pilus assembly protein, pilin
	tadA/	Flp pilus assembly protein
	virB11/	Type II/IV secretion system protein
	cpaF	Flp pilus assembly protein
	tadZ/	Type II/IV secretion system ATP
	cpaE	hydrolase
	cpaD	Type II/IV secretion system ATPase
	cpaC	Flp pilus assembly protein
	rcpA	Type ii/IV secretion system secretin
	tadV	Flp pilus assembly protein
		Type IV prepilin peptidase

pqqA, pqqB, pqqC, pqqD and *pqqE. S. meliloti* 1021 bears all the PQQ biosynthesis genes while *M. loti* MAFF303099 and *R. leguminosarum* WSM1325 lacks all of them. Td3 bears all PQQ coding genes except *pqqA*. Even though it lacks *pqqA,* it can still produce functional PQQ as reflected by its *in vitro* gluconic acid production ability. This is in accordance to an earlier report which shows that *pqqA* is not required for

Table 4
Comparison of proteins/ enzymes involved in D-gluconate and ketogluconate metabolism in *Rhizobium sp.* Td3 and other rhizobia.

Protein/ Enzyme	Number of genes coding for protein in rhizobial species					
	Td3	Rm1021	*M. loti*	*R. leg*	*B. jap*	*A. cau*
Glucose 1-dehydrogenase	1	0	1	0	1	0
Glucose dehydrogenase, PQQ dependent	3	1	1	2	0	0
Gluconate 2- dehydrogenase	6	0	0	0	0	2
2- keto D-gluconate dehydrogenase	9	0	1	0	0	1
2,5 diketo-D-gluconate reductase	1	0	0	1	0	0
Positive regulator of L-idonate catabolism	0	0	0	1	0	0
Gluconate transporter	1	0	0	0	0	1
L-idonate 5-dehydrogenase	3	1	0	1	2	1
5-keto-D-gluconate 5-reductase	1	1	1	2	1	1
Gluconokinase	1	2	1	2	1	0
Gluconate dehydratase	4	2	6	4	1	0
Phosphogluconate dehydratase	0	1	1	1	0	1
2-dehydro-3-deoxyphosphogluconate aldolase	0	2	1	1	1	1
6-phosphogluconate dehydrogenase	1	1	2	2	1	0
2-dehydro-3-deoxygluconate kinase	1	1	1	1	1	1
Total	32	12	16	18	9	9

Td3: *Rhizobium sp.* Td3; Rm1021: *Sinorhizobium meliloti* 1021; M. loti: *Mesorhizobium loti* MAFF303099; R.leg: *Rhizobium leguminosarum bv. trifolii* WSM1325; B.jap: *Bradyrhizobium japonicum* USDA110; A.cau: *Azorhizobium caulinodans* ORS571.

PQQ biosynthesis in *Methylobacterium extorquens* AM1 (Toyama and Lidstrom, 1998).

It oxidizes glucose to gluconates and ketogluconates. Genes of the entire periplasmic glucose oxidation pathway of [illegible] were [illegible] other rhizobia. Table 4 summarizes the 32 genes involved in D-gluconate and ketogluconate metabolism in Td3 and its comparison with other rhizobia. Td3 has three periplasmic glucose dehydrogenases each encoded by a single gene and designated [illegible] *gcd1, gcd2* and *gcd3* respectively. As shown in Fig. 4, Gcd1 is closest [illegible] Gcd of *Yangia sp.* and *Xanthomonas sp.* Gcd2 is closest to the other rhizobial Gcd proteins whereas Gcd3 is closest [illegible] Gcd of *Klebsiella pneumoniae* and *Gluconobacter sp.* It also bears two gluconate 2-dehydrogenase (*gad1* and *gad2*) and three 2-keto D-gluconate dehydrogenase (*2kgdh1, 2kgdh2* and *2kgdh3*); each encoded by three genes one for each component namely gamma, flavoprotein and cytochrome C subunits. Td3 also has a high affinity gluconate transporter, gluconokinase and gluconate dehydrogenase which are not commonly present in rhizobia.

3.9.2. Stress response

Oxidative stress is the main environmental problem which nearly all organisms should deal with. In case of pathogens, protection from reactive oxygen species is also an aspect of their pathogenicity. Many other stresses like starvation, pH, osmotic, temperature, etc. may also lead to oxidative stress (Seshasayee et al., 2006). Responses to H_2O_2, superoxide, singlet oxygen, peroxynitrite, etc. will depend on a particular organism's protective, responsive and enzymatic functions. Genome codes for multitude of genes operative in generating stress response and for betaine uptake and biosynthesis. Key enzymes regulating oxidative stress include peroxidase, catalase, superoxide dismutase and glutathione transferase, all of which are produced by Td3. Reactive oxygen species (ROS) generated as side products of metabolic pathways are degraded by multiple mechanisms. When the balance between generation and degradation is lost, oxidative stress occur. Peroxidases using some reductants and peroxides as cosubstrates, producing water and reduced products; Catalases decompose hydrogen peroxide, yielding water and molecular oxygen, superoxide dismutases speed up the dismutation of superoxide anion to molecular oxygen and hydrogen peroxide. Glutathione transferase detoxifies xenobiotics and limits oxidative damage and other plant stress responses (Gong et al., 2005).

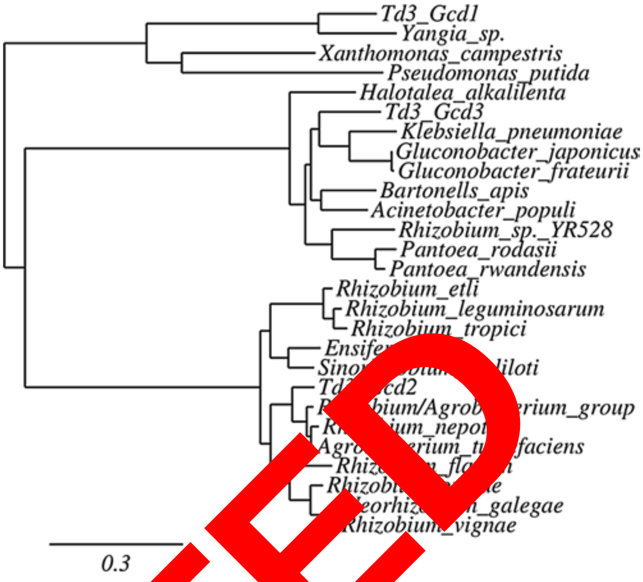

Fig. 4. Phylogenetic relationship of [illegible] Gcd proteins with other rhizobacteria. Phylogenetic tree [illegible] been constructed using Phylogeny. for the three Td3 quinoprotein glucose dehydrogenase with other rhizobacteria. Organisms belonging to the same genera have been represented with a similar colour code. Values [illegible] the scale bar indicates the number of substitutions per site.

Presence of one or more functional cassette depends on the environmental niche which a bacterium occupies and the probability of being subjected to oxidative damage. Genome also codes for several components of bacterial type II and type IV secretion system. CidA is a pro-lytic protein that increases murein hydrolysis and *cidABC* operon regulates programmed cell lysis (Rice et al., 2003). Betalactamase activity aids in multiplication in the rhizosphere and root infection. Osmoregulation genes are highly important for soil bacteria as variations in the osmotic environment within rhizosphere effects root colonization, nodule development and atmospheric nitrogen fixation efficiency. Glycine betaine (N,N,N-trimethylglycine) has been shown to be a very efficient osmolyte found in wide range of bacterial and plant species, where it is accumulated at high cytoplasmic concentrations in response to osmotic stress. *S. meliloti* (in contrast to *E. coli, B. subtilis*, and other bacteria), can use glycine betaine not only as osmoprotectants but as C, N and energy source (Boncompagni et al., 2000).

3.9.3. Metal resistance

Genome also bears genes for transport of Ni, Co, Cu, As, Zn, Cd thereby inducing metal resistance. It also has a zinc uptake regulation protein, a putative metal chaperone involved in Zn homeostasis and Zn dependent enzymes like GTP cyclohydrolase I, dihydroorotase, porphobilinogen synthase, carbonic anhydrase, phosphoribosyl-AMP cyclohydrolase, N-acetylmuramoyl-L-alanine amidase, cysteinyl-tRNA synthetase, etc. As zinc is a component of many proteins, in particular, DNA polymerases, proteases, ribosomal proteins, etc., bacteria must have effective systems for zinc transport.

3.9.4. Other plant growth promotion traits

Genome also codes for an exceptionally large number of genes vital for production of enterobactin and aerobactin siderophores and flagella formation. Td3 genome encodes significantly large number of genes involved in thiamin biosynthesis, multidrug resistance efflux pump, arsenic resistance and flagellar motility (flagellar motor switch proteins FliG, FliM, FliN) compared to the other genomes taken into consideration.

3.10. Other features

Unlike other rhizobial genomes taken into account, Td3 genome has genes vital for alginate and sialic acid metabolism, phage packaging machinery (phage tail fibre protein, phage capsid protein and phage lysis module), tRNA nucleotidyl transferases, dipeptidases, omega-peptidases, nudix proteins (nucleoside triphosphate hydrolases), bacterial chemotaxis serine chemoreceptor protein, chemotaxis protein (CheV), CheA protein activity positive regulator (CheW), CheA histidine kinase, chemotaxis regulator CheY, protein-glutamate methylesterase (CheB), chemotaxis protein methyltransferase (CheR) and chemotaxis protein). It also bears periplasmic binding protein for maltose/maltodextrin, ribose and dipeptide transport. Td3 also has propionyl CoA to succinyl CoA conversion module and large number of genes for glycerol and glycerol-3-phosphate uptake and utilization, can catabolize inositol, D-galactarate, D-glucarate and D-glycerate and can metabolize mannose.

4. Conclusion

This study showed that in addition to phosphate solubilization, Td3 genome encodes for genes involved in multiple phytobeneficial traits like synthesis of auxins, nitric oxide, siderophores, acetoin, exopolysaccharide, resistance to metal ions like Ni, Co, Cu, As, Zn, Cd, [illegible] can also produce certain enzymes conferring protection against oxidative stress, antibiosis and rhizospheric colonization. As rhizosphere [illegible] complex environmental conditions, these genes bestowing plant growth promotion may also be vital in bacterial survival and [illegible]ion. The availability of genetic information of such a versatile plant growth promoter Td3 may advance our understanding [illegible] phosph[illegible]bilization in rhizobia with multiple plant growth [illegible]om[illegible]its and may further provide insight into the practical applications [illegible]e strain as a biofertilizer in the field of agriculture.

Conflict of interest

The authors declare th[illegible] they h[illegible] no conflict of interest.

Ethical approval

This art[illegible] does no[illegible]ontain any studies with human participants or animals pe[illegible]ed b[illegible] authors.

Acknowledgements

The authors would like to acknowledge Department of Science and Technology (DST), Government of India for providing INSPIRE fellowship to Bhagya Iyer (IF130895) and Nirma Education and Research Foundation (NERF), Nirma university for providing basic infrastructure for carrying out this research. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi: https://doi.org/10.1016/j.micres.2018.09.007

References

Ahuja, N., Kumar, P., Bhatnagar, R., 2004. The adenylate cyclase toxins. Crit. Rev. Microbiol. 30 (3), 187–196.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search too. J. Mol. Biol. 215, 403–410.

Ardley, J.K., Parker, M.A., Meyer, S.E., Trengove, R.D., Hara, G.W., Reeve, W.G., Yates, R.J., Dilworth, M.J., Willems, A., Howieson, J.G., 2012. *Microvirga lupini sp. nov., Microvirga lotononidis sp. nov.* and*Microvirga zambiensis sp. nov*. are alphaproteobacterial root-nodule bacteria that specifically nodulate and fix nitrogen with geographically and taxonomically separate legume hosts. Int. J. Syst. Evol. Microbiol. 62, 2579–2588.

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., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vasseiva, O., Vonstein, V., Wilke, A., [illegible]O., 2008. The RAST Server: Rapid Annotations using Subsystems [illegible]nology. [illegible] Genomics 9, 75.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, [illegible]., Dvorkin, M., [illegible]likov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, [illegible], Pyshkin, A.V., [illegible]tkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., [illegible]. SPAdes: A [illegible] genome assembly algorithm and its applications to [illegible]e cell seq[illegible]ng. J. C[illegible]ut. Biol. 19 (5), 455–477.

Beneit, F.S., 2015. The Pho regu[illegible]a huge r[illegible]latory n[illegible] in bacteria. Front. Microbiol. 6, 402.

Bernardelli, C.E., Luna, M.F., Ga[illegible]M.L., [illegible]rdi, J.L., [illegible]01. Periplasmic PQQ-dependent glucose oxidation in [illegible] livin[illegible]ymbiotic [illegible]eria. Curr. Microbiol. 42 (5), 310–315.

Bertelli, C., Greub, [illegible]013. Rapid bac[illegible]me sequencing: methods and applications in cli[illegible]iology. Clin. M[illegible]ol. Infect. 3 (19), 803–813.

Boetzer, M., H[illegible]kel, C.V., [illegible]en, H.J., Butler, D., Pirovano, W., 2011. SSPACE: Scaffolding pre-assemb[illegible]ntigs using SSPACE. Bioinformatics 27 (4), 578–579.

Boivin, [illegible]raud, E., Perre[illegible]atut, J., 2009. Establishing nitrogen-fixing symbiosis [illegible] legumes: how many rh[illegible]bium recipes? Trends Microbiol. 17, 458–466.

Bo[illegible]mpagni, E., Dupont, L., Mignot, T., Osteras, M., Lambert, A., Poggi, M.C., Rudulier, [illegible]., 2000. Charac[illegible]zation of a Sinorhizobium meliloti ATP-Binding cassette histidine [illegible]ansporter also [illegible]lved in betaine and proline uptake. J. Bacteriol. 182 (13), [illegible]7–3725.

Dere[illegible]A., Gui[illegible], V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Gu[illegible]fort, V., Lescot, M., Claverie, J.M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 1 (36), [illegible]465–W469.

[illegible]iang, W., Cheng, Z., Heikkila, J.J., Glick, B.R., 2013. The complete genome sequence of the plant growth promoting bacterium *Pseudomonas* sp. UW4. PLoS One 8 (3). https://doi.org/10.1371/journal.pone.0058640. e58640.

Fraile, P.G., Carro, L., Robledo, M., Bahena, M.H., Felix, J.D., Fernandez, M.T., Mateos, P.F., Rivas, R., Igual, J.M., Molina, E.M., Peix, A., Velazquez, E., 2012. *Rhizobium* promotes nonlegumes growth and quality in several production steps: Towards a biofertilization of edible raw vegetables healthy for humans. PLoS One 7 (5). https://doi.org/10.1371/journal.pone.0038122.

Gong, H., Jiao, Y., Hu, W.W., 2005. Expression of glutathione-s- transferase and its role in plant growth and development in vivo and shoot morphogenesis in vitro. Plant Mol. Biol. 57 (1), 53–66.

Graumann, P.L., 2004. Cytoskeletal elements in bacteria. Curr. Opin. Microbiol. 7 (6), 565–571.

Han, N., Qiang, Y., Zhang, W., 2016. ANItools Web: A Web Tool for Fast Genome Comparison Within Multiple Bacterial Strais. Database 2016: baw084. PMCID: PMC4911789. .

Kaur, C., Selvakumar, G., Ganeshamurtha, A.N., 2016. Draft genome sequence of phosphate solubilizing bacterium *Paraburkholderia tropica* strain P-31 isolated from pomegranate *(Punica granatum)* rhizosphere. Genome Announc. 4 (4). https://doi.org/10.1128/genomeA.00844-16. e00844-16.

Kwak, Y., Park, G.S., Shin, J.H., 2016. High quality draft genome sequence of the type strain of *Pseudomonas lutea* OK2, a phosphate solubilizing rhizospheric bacterium. Stand. Genomic Sci. 11 (51). https://doi.org/10.1186/s40793-016-0173-7.

Lin, D.X., Wang, E.T., Tang, H., 2008. *Shinellakummerowiae*sp. nov., a symbiotic bacterium isolated from root nodules of the herbal legume*Kummerowiastipulacea*. Int. J. Syst. Evol. Microbiol. 58, 1409–1413.

Liu, X.Y., Wu, W., Wang, E.T., Zhang, B., Macdermott, J., Chen, W.X., 2011. Phylogenetic relationships and diversity of β-rhizobia associated with *Mimosa* species grown in Sishuangbanna. China. Int J Syst Evol Microbiol 61, 334–342.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A., Kanehisa, M., 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35, 182–185.

Pandey, D.P., Gerdes, K., 2005. Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. 33 (3), 966–976.

Pei, A.Y., Oberdorf, W.E., Nossa, C.W., 2010. Diversity of 1 6S rRNA genes within individual prokaryotic genomes. Appl. Environ. Microbiol. 76 (1–2), 3886–3897.

Peix, A., Valverde, A., Rivas, R., Igual, J.M., Ramirez-Bahena, M.H., Mateos, P.F., Santa-Regina, I., Rodriguez-Barrueco, C., Martinez-Molina, E., Velazquez, E., 2007. Reclassification of *Pseudomonas aurantiaca* as a synonym of *Pseudomonas chlororaphis* and proposal of three subspecies, *P. chlororaphis subsp. chlororaphis subsp. nov., P. chlororaphis subsp. aureofaciens subsp. nov., comb. nov.* and*P. chlororaphis subsp. aurantiaca subsp. nov., comb. nov.* Int. J. Syst. Evol. Microbiol. 57 (6), 1286–1290.

Pineiro, J.L., Rivas, R., Trujillo, M.E., Vizcaino, N., Carrasco, J.A., Chamber, M., Palomares, A., Mateos, P.F., Martinez-Molina, E., Velazquez, E., 2007.

Ochrobactrumcytisisp. nov., isolated from nodules of Cytisusscopariusin spain. Int. J. Syst. Evol. Microbiol. 57, 784–788.

Raunser, S., Appel, M., Ganae, C., Kaufer, G.U., Fendler, K., Kuhlbrandt, W., 2006. Structure and function of prokaryotic glutamate transporters from *Escherichia coli* and *Pyrococcus horikoshii*. Biochemistry 45 (42), 12796–12805.

Reizer, J., Reizer, A., Saier, M.H., 1994. A functional superfamily of sodium/solute symporters. Biochim. Biophys. Acta 1197, 133–166.

Rice, K.C., Firek, B.K., Nelson, J.B., Yang, S.J., Patton, T.G., Bayles, K.W., 2003. The Staphylococcus aureus cidAB operon: evaluation of its role in the regulation of murein hydrolase activity and penicillin tolerance. J. Bacteriol. 185, 2635–2643.

Rivas, R., Willems, A., Subba Rao, N.S., 2003. Description of *Devosianeptuniae*sp. nov. that nodulates and fixes nitrogen in symbiosis with*Neptunianatans*, an aquatic legume from India. Syst. Appl. Microbiol. 26, 47–53.

Seshasayee, A.S., Bertone, P., Fraser, G.M., Luscombe, N.M., 2006. Transcriptional regulatory networks in bacteria: from input signals to output responses. Curr. Opin. Microbiol. 9, 511–519.

Shih, Y.L., Rothfield, L., 2006. The bacterial cytoskeleton. Microbiol. Mol. Biol. Rev. 70 (3), 729–754.

Toyama, H., Lidstrom, M.E., 1998. *pqqA* is not required for biosynthesis of pyrroloquinoline quinone in *Methylobacterium extroquens* AM1. Microbiol 144, 183–191.

Tseng, T.T., Tyler, B.M., Setubal, J.C., 2009. Protein secretion systems in bacterial-host associations, and their descriptionn in the gene ontology. BMC Microbiol. 9 (Suppl. 1), S2 PMC2654662.

Willems, A., 2006. The taxonomy of rhizobia: an overview. Plant Soil 287, 3–14.

Yamaguchi, T., Tsutsumi, F., Putnoky, P., Fukuhara, M., Nakamura, T., 2009. pH-dependent regulation of the multi-subunit cation/proton antiporter Pha1 system from Sinorhizobium meliloti. Microbiology 155, 2750–2756.