

Whole genome sequences of four Plant Growth Promoting Rhizobacteria strains from different Tunisian rhizospheres

Naima Sayahi^{1*‡}, Nourelhouha Abdelkefi^{2‡}, Chantal Ebel¹, Tahar Mechichi², Moez Hanin¹

1: Plant Physiology and Functional Genomics Research Unit, Institute of Biotechnology of Sfax, University of Sfax, 3038 Sfax, Tunisia

2: Laboratoire de Biochimie et de Génie Enzymatique des lipases, Ecole Nationale d'Ingénieurs de Sfax, Université de Sfax, Sfax, Tunisia

* Corresponding authors: naima.sayahi1993@gmail.com; abdelkefinourelhouda@gmail.com

‡: Both authors contributed equally to this publication

Abstract

This study reports the full genome sequencing of four plant growth promoting rhizobacteria strains that were isolated from rhizospheres of plants growing in different Tunisian arid regions. These genome sequences offer valuable insights into the diversity of the bacterial community interconnected with plant rhizosphere. Furthermore, they give a glimpse into the essential genes that are pivotal for their PGPR behaviors and their effectiveness in enhancing plant growth under harsh environmental conditions.

Introduction:

Plant Growth-Promoting Rhizobacteria (PGPR) are beneficial bacteria that inhabit the rhizosphere, which is the soil layer around plant roots (Santoyo et al., 2021). These bacteria have a positive impact on plant growth and development through various mechanisms (Olanrewaju et al., (2017)). PGPR can for instance solubilize essential nutrients such as phosphorus, making them more available to plants (Etesami et al., 2020). PGPR are also able to produce plant growth hormones, such as auxins, cytokinins, and gibberellins, which influence various aspects of plant growth and development, including root and shoot elongation (Tsukanova et al., 2017). In addition, PGPR can act as biocontrol agents against plant pathogens by producing antibiotics, competing for resources, and inducing systemic resistance in plants (Meena et al., 2020). These features were reported to be useful to enhance plant tolerance to various abiotic (i.e. drought and salinity) and biotic stresses (i.e. fungal and viral infections; Kumar et al 2019; Annapurna et al., 2013).

Therefore, the use of PGPR in agriculture has gained increased attention as a sustainable and eco-friendly approach to enhance crop productivity (Ahirwar et al., 2020; Aloo et al., 2019). Despite an extended literature describing the characterization of various PGPR, the specific mechanisms by which they enhance crop growth and yield remain unclear. Whole-genome sequencing (WGS) has become a pivotal tool for functional studies which allow the identification and annotation of genes within organisms (Chandra et al., 2021) and particularly the identification of biosynthetic gene clusters (BGCs), which are groups of genes responsible for the production of natural products (Li et al., 2020). For our work, WGS allows the comprehensive identification of genes and genetic pathways associated with beneficial traits of PGPR. This includes genes related to nitrogen fixation, phosphate solubilization, production of plant growth hormones, and biocontrol activities against plant pathogens (Gupta et al., 2014, Narayanasamy et al 2023). The entire genome annotation helps understanding the

molecular mechanisms behind the beneficial interactions between PGPR and plants (Surovyetal., 2019). WGS can also boost biotechnological applications by engineering genetically-modified PGPR with enhanced plant growth-promoting abilities to promote sustainable agriculture practices with reduced reliance on chemical fertilizers and pesticides (Adebayo et al., 2023). Furthermore, genome comparison of different strains or species helps in studying the evolutionary relationships among different PGPR strains which is valuable for tracing the origin of beneficial traits and understanding the evolution of PGPR-plant interactions (Hassen, et al., 2021).

Four PGPR strains were isolated from soil samples collected from rhizospheres of different plants growing in Tunisian arid regions. The PGPR were chosen on their PGP potential and their salt stress tolerance. These isolates were previously identified by analyzing the 16S rRNA gene sequences. In this study, we present the WGS of these bacterial strains which allowed a better identification of the bacterial species, genetic diversity, and will aid gene discovery for functional analysis and biotechnological valorization.

Materials and methods

Bacterial isolates

The strain C2 was isolated from the rhizosphere of *Opuntia ficus* grown in the region of Sfax (34°39'N, 10°43'E) as described in Sayahi et al. (2022). The strains GB7, All3 and Gb119 were isolated from Kebili (33° 42' N, 8° 58' E), Mednine (33° 20' N, 10° 29' E) and Gabes oasis (33° 53'N, 10° 5' E) respectively.

Analyzing bacterial genomic DNA

WGS was performed by Get Genome (<https://getgenome.net/>) through the sequencing service provider MicrobesNG (Birmingham, UK) who managed the bacterial culturing, DNA extraction, genome sequencing by Illumina, and assembly according to established protocols available at <https://microbesng.com/>. Trimmomatic version 0.39 (Bolger et al., 2014) was used to trim the sequencing reads, using a sliding window quality cutoff of Q15. Quality evaluation was carried out through a combination of in-house scripts, along with Samtools version 1.4 ([git://github.com/samtools/samtools.git](https://github.com/samtools/samtools.git)), BedTools version 2.18 (Quinlan and Hall, 2010), and bwa-mem (Li and Durbin, 2009) software. Following Gurevich et al. (2013), these sequence reads were assembled into contigs using Quast software version 5.0.2. Genomes were annotated using Prokka 1.14.3 (<https://github.com/tseemann/prokka>). Prediction of protein-

coding features and tRNA was performed utilizing Prodigal version 2.6 (Hyatt et al., 2010), while rRNA prediction was carried out using ARAGORN version 1.2 (Laslett and Canback, 2004).

Results:

Genome sequencing was used to analyze four PGPR strains isolated from diverse rhizospheres of plants grown in southern Tunisia. Quast was used to aggregate the sequence reads, which resulted in a variety of contigs, ranging from 40 to 187, with the largest contig measuring 1,616,315 base pairs. The four strains, C2, GB7, All3, and GB119, have accumulated contig lengths of 5,241,915 base pairs, 4,126,956 base pairs, 4,024,208 base pairs, and 5,824,939 base pairs respectively. The GC contents were 59.20% for C2, 43.66% for GB7, 43.76% for All3, and 45.61% for GB119. A total count of genes, which includes protein-coding genes, tRNA genes, and rRNA genes is shown in Table 1.

The whole genomes of the strain C2, GB119, All3, and GB7 present a percentage identity of 92.14%, 92.09%, 63.47%, and 89.98 % to *Serratia marcescens*, *Paenibacillus polymyxa*, *Bacillus subtilis*, *Bacillus halotolerans* respectively. It is worth noting that the C2 strain has previously been characterized as *Siccibacter* spp based on a partial 16S rRNA fragment (Sayahi et al. 2022) showing that WGS enabled a more precise and reliable identification.

Using the whole genomes of the four strains, we were able to identify numerous genes involved in the plant growth pathway (see table 2). The strain C2 had the highest number of genes involved in GABA production, organic P mineralization and glyphosate degradation, zinc solubilization, siderophore production, chitinase production, putrescine biosynthesis, chemotaxis, and the coenzyme PQQ synthesis. Furthermore, the strain G119, the strain GB7, and the strain All3 had the greatest number of genes for tryptophan biosynthesis (figure 1). A Venn diagram was constructed to perform a comparative analysis of the four bacterial strains based on the number of PGP related genes. After clustering, a total of 104, 30, 30 and 43 genes from C2, G119, All3 and GB7 respectively, was obtained (Figure 2). 86 and 18 genes were specific for C2 and GB7 respectively. Moreover, the number of overlapping genes is 5 between C2 and GB7, 7 between G119 and All3, 3 between G119, C2 and All3, 10 between G119, All3 and GB7 and 10 between C2, G119, All3 and GB7.

Data availability:

The entire genome shotgun has been deposited at DDBJ/ENA/GenBank with accession numbers JAXCHC000000000, JAXCHD000000000, JAXCHE000000000 and JAXCHF000000000. Raw sequence reads have been deposited in the NCBI Sequence Read Archive under BioProject number PRJNA1041092 and accession numbers SRS19575062, SRS19575063, SRS19575064, and SRS19575065 for the strain C2, GB119, All3, and GB7 respectively, as indicated in table 1.

Acknowledgements:

The sequencing of the bacterial strains was fully supported by GetGenome and the Sainsbury Laboratory, Norwich, UK, with contributions from the Gatsby Charitable Foundation, the Biotechnology, Biological Sciences Research Council (BBSRC) and The University of East Anglia. We are extremely grateful to Sophien Kamoun, James Canham and Joe Win from GetGenome for their valuable support.

References:

- Ahirwar, N. K., Singh, R., Chaurasia, S., Chandra, R., & Ramana, S. (2020). Effective role of beneficial microbes in achieving the sustainable agriculture and eco-friendly environment development goals: a review. *Front. Microbiol.*, 5, 111-123.
- Aloo, B. N., Makumba, B. A., & Mbega, E. R. (2019). The potential of Bacilli rhizobacteria for sustainable crop production and environmental sustainability. *Microbiological Research*, 219, 26-39.
- Annapurna, K., Kumar, A., Kumar, L. V., Govindasamy, V., Bose, P., & Ramadoss, D. (2013). PGPR-induced systemic resistance (ISR) in plant disease management. *Bacteria in agrobiology: disease management*, 405-425.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120
- Chandra, A., Chandra, P., & Tripathi, P. (2021). Whole genome sequence insight of two plant growth-promoting bacteria (*B. subtilis* BS87 and *B. megaterium* BM89) isolated and characterized from sugarcane rhizosphere depicting better crop yield potentiality. *Microbiological Research*, 247, 126733.

Etesami, H., & Adl, S. M. (2020). Plant growth-promoting rhizobacteria (PGPR) and their action mechanisms in availability of nutrients to plants. *Phyto-Microbiome in stress regulation*, 147-203.

Giannelli, G., Potestio, S., & Visioli, G. (2023). The Contribution of PGPR in Salt Stress Tolerance in Crops: Unravelling the Molecular Mechanisms of Cross-Talk between Plant and Bacteria. *Plants*, 12(11), 2197

Gupta A, Gopal M, Thomas GV, Manikandan V, Gajewski J, Thomas G, Seshagiri S, Schuster SC, Rajesh P, Gupta R. Whole genome sequencing and analysis of plant growth promoting bacteria isolated from the rhizosphere of plantation crops coconut, cocoa and arecanut. PLoS One. 2014 Aug 27;9(8):e104259. doi:

Gupta, G., Parihar, S. S., Ahirwar, N. K., Snehi, S. K., & Singh, V. (2015). Plant growth promoting rhizobacteria (PGPR): current and future prospects for development of sustainable agriculture. *J Microb Biochem Technol*, 7(2), 096-102.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075. <https://doi.org/10.1007/978-3-030-44368-9>

Hassen, A. I., Khambani, L. S., Swanevelder, Z. H., Mtsweni, N. P., Bopape, F. L., Van Vuuren, A., ... & Morey, L. (2021). Elucidating key plant growth-promoting (PGPR) traits in *Burkholderia* sp. Nafp2/4-1b (= SARCC-3049) using gnotobiotic assays and whole-genome-sequence analysis. *Letters in Applied Microbiology*, 73(5), 658-671.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, 11(1), 1-11. <https://doi.org/10.1186/1471-2105-11-119>

Ishitani, M., Rao, I., Wenzl, P., Beebe, S., & Tohme, J. (2004). Integration of genomics approach with traditional breeding towards improving abiotic stress adaptation: drought and aluminum toxicity as case studies. *Field Crops Research*, 90(1), 35-45.

Kumar, A., Patel, J. S., Meena, V. S., & Srivastava, R. (2019). Recent advances of PGPR based approaches for stress tolerance in plants for sustainable agriculture. *Biocatalysis and agricultural biotechnology*, 20, 101271.

Laslett, D., Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic acids research*, 32(1), 11-16. <https://doi.org/10.1093/nar/gkh152>

Li, H., Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, 25(14), 1754-1760. <https://doi.org/10.1093/bioinformatics/btp324>

Li, Z., Song, C., Yi, Y., & Kuipers, O. P. (2020). Characterization of plant growth-promoting rhizobacteria from perennial ryegrass and genome mining of novel antimicrobial gene clusters. *BMC genomics*, 21(1), 1-11

Meena, M., Swapnil, P., Divyanshu, K., Kumar, S., Harish, Tripathi, Y. N., ... & Upadhyay, R. S. (2020). PGPR-mediated induction of systemic resistance and physiochemical alterations in plants against the pathogens: Current perspectives. *Journal of Basic Microbiology*, 60(10), 828-861

Narayanasamy, S., Thankappan, S., Kumaravel, S., Ragupathi, S., & Uthandi, S. (2023). Complete genome sequence analysis of a plant growth-promoting phylloplane *Bacillus altitudinis* FD48 offers mechanistic insights into priming drought stress tolerance in rice. *Genomics*, 115(1), 110550.

Olanrewaju, O. S., Glick, B. R., & Babalola, O. O. (2017). Mechanisms of action of plant growth promoting bacteria. *World Journal of Microbiology and Biotechnology*, 33, 1-16.

Quinlan, A. R., Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841-842. <https://doi.org/10.1093/bioinformatics/btq033>

Santoyo, G., Urtis-Flores, C. A., Loeza-Lara, P. D., Orozco-Mosqueda, M. D. C., & Glick, B. R. (2021). Rhizosphere colonization determinants by plant growth-promoting rhizobacteria (PGPR). *Biology*, 10(6), 475.

Sayahi, N., Djemal, R., Ben Merdes, K., Saidii, M. N., Yengui, M., Gdoura, R., ... & Hanin, M. (2022). Characterization of *Siccibacter* sp. Strain C2 a Novel Rhizobacterium that Enhances Tolerance of Barley to Salt Stress. *Current Microbiology*, 79(8), 239.)

Surovy, M. Z., Gupta, D. R., Mahmud, N. U., Dame, Z. T., Roy, P. K., & Islam, M. T. (2019). Genomics and post-genomics approaches for elucidating molecular mechanisms of plant

growth-promoting Bacilli. *Bacilli and Agrobiotechnology: Phytostimulation and Biocontrol: Volume 2*, 161-200.

Tsukanova, K. A., Meyer, J. J. M., & Bibikova, T. N. (2017). Effect of plant growth-promoting Rhizobacteria on plant hormone homeostasis. *South African Journal of Botany*, 113, 91-102.

Table1:Genomes assembled from Illumina for the four PGPR strains

Bacterial strain	Rhizospheres area	Biosample	# contigs ≥1000 bp	Largest contig (bp)	Total length (bp)	GC (%)	Mean coverage	N50	CDS	tRNA	tmRNA	GenBank Accession (Assembly)	GenBank Accession (Raw reads)
Serratia sp. 201	Opuntia ficus	SAMN38274870	12	1222121616315	5241915	59,20	99,1792	1588841	4841	85	1	JAXCHC000000000	SRR26872287
Paenibacillus sp. 203	Mednine oasis	SAMN38274871	36	695128	5824939	45,61	63,4205	529095	5118	107	1	JAXCHD000000000	SRR26872286
Bacillus sp. 204(2023	Gabes oasis	SAMN38274872	21	787869	4024208	43,76	105,055	391810	3959	84	1	JAXCHE000000000	SRR26872285
Bacillus sp. 205(2023)	Kebili oasis	SAMN38274873	25	1458351	4126956	43,66	91,3469	318766	4089	84	1	JAXCHF000000000	SRR26872284

Table 2 : List of the genes assigned to plant growth promotion traits in the PGPR strains’ genome

Plant growth promotion traits	Genes with potential to confer PGP traits			
	C2	G119	All3	Gb7
IAA production	ipdC_1, ipdC_2, aspC	aspC	aspC	aspC
Tryptophan biosynthesis	tyrB_1, tyrB_2	trpA, trpB, trpC, trpF, trpD, pabA, trpE	trpA, trpB, trpC, trpF, trpD, pabA, trpE	trpA, trpB, trpF, trpC, trpD, trpE
Cytokinin biosynthesis	miaA, miaB	miaA, miaB	miaA, miaB	miaA, miaB
GABA production	puuC, pup, puuB_2, prr_1, prr_2, puuE, gbuA, puuC, puuA_1	-	-	puuB
GABA transport	gabP	gabP	gabP	gabP
GABA degradation	-	gabD	gabD	gabD
Inorganic P solubilisation	Gcd, ppaC, ppa, ppx	appA, pstB3, pstB3, pstS1	appA, pstB3, pstB3, pstS1	-
Organic P mineralization and glyphosate degradation	phnK, phnJ, phnI, phnH, phnG, phnP, phnN, phnM, phnL, phnW, phnX	-	-	appA, oppF_3, oppD_3
Zinc solubilization	pitA_2, pitA_1, znuA, znuB, znuA, cadA, cadC, cadB, zntA, zntB_1, zntB_2, zntR	znuB, znuC, znuC	znuB, znuC, znuC	znuA, znuC_2, znuB
Siderophore production	entC_2, entC_1, entF, entE, entB, entS_2, entS_1, exbD_1, exbD_2, bfr, bfd	-	-	-
Chitinase production	chiA1, chiB, chiA	-	-	-
Acetoin production	budB, ilvB_2, ilvN, ilvB_3, ilvI, ilvH, ilvM, ilvG	-	-	alsD, ilvB, budC, ilvH
Spermidine/Spermine biosynthesis	-	-	-	speH, speE, speG
Putrescine biosynthesis	speB, speA, speD, speC, speE_1, speE_2, speF, metK	speB, speE, speH, metK, speE	speB, speE, speH, metK, speE	speA_1, speA_2, speB, speE, speH, metK
Rhizopine degradation and transport	iolG_1, iolG_2, iolG_3, iolG_4, iolG_5	iolG	iolG	iolG_1
transporters	ugpB, ugpA, ugpC_3, ugpC_1	-	-	ugpC, pstA
Chemotaxis	cheW, cheY, cheA, cheZ, cheR, cheB, tsr_1, tap, tsr_2, tsr_3, tsr_4	cheA, cheW, cheC, cheD, cheY, cheY, cheR, cheB	cheA, cheW, cheC, cheD, cheY, cheY, cheR, cheB	cheA, cheW, cheY_2, cheY_1, cheR, cheB_2, cheB_3, cheB_4, cheB_1
Copper resistance	copA	-	-	copA
Plant polymerdegradation enzymes	puuE, iolB, uxaC	uxaC	uxaC	eglS
Coenzyme PQQ synthesis	pstC, pstA, pstB, phoB, pstS_1, phoR, pqqC, pqqB, pqqD, pqqB	-	-	-

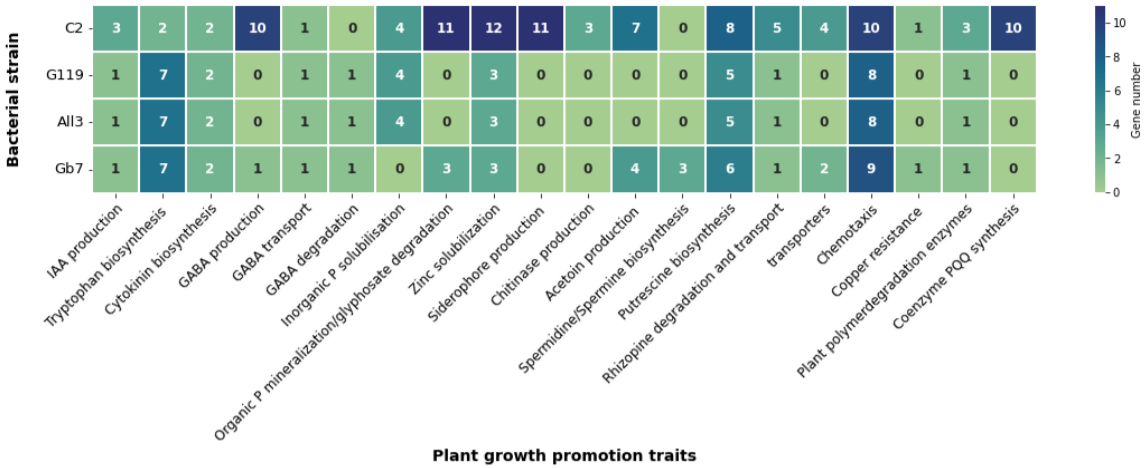


Figure 1:Annotated heatmap representation of the number of genes associated with plant growth promotion traits for individual strains

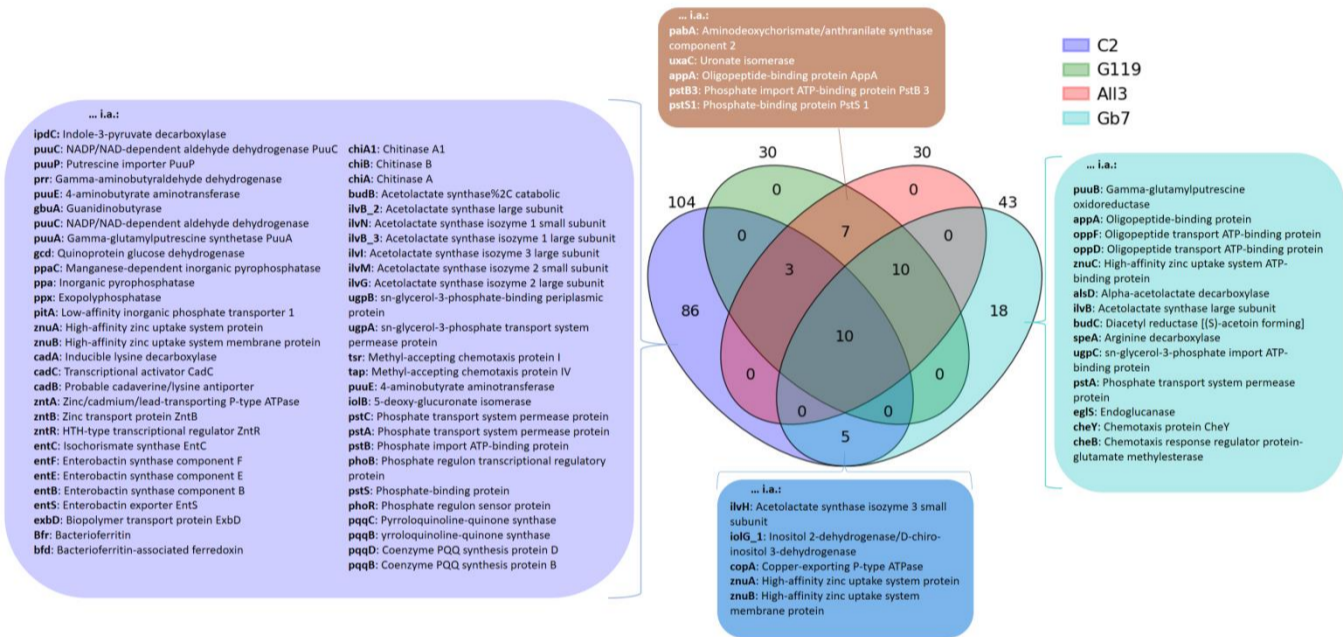


Figure 2: Venn diagram analysis illustrating shared and dissimilar genes/gene sets coding for plant growth promotion traits for individual strains. The numbers indicate the number of shared or strain-specific genes. The number on top of each ellipse indicates the total number of genes possessed by each strain