Genomic and phylogenetic analysis of plant growth-promoting bacteria

Marcus Vinicius da Silva Rodrigues[1], João Paulo de Oliveira, Sandriele Aparecida Noriler[2], Angelo Barreiro Garcia[2], Ulisses de Pádua Pereira[2], Ulisses Nunes da Rocha[3], Admilton Gonçalves de Oliveira Junior[2], Glaucia Maria Bressan[1]

[1]Graduate Program in Bioinformatics
Universidade Tecnológica Federal do Paraná
Cornélio Procópio, PR, Brazil.

[2]Department of Microbiology
Universidade Estadual de Londrina
Londrina, PR, Brazil.

[3]Department of Microbial Ecology
Helmholtz Centre for Environmental Research
Leipizg, Germany

marcusr@alunos.utfpr.edu.br, jp.joaopaulo_oliveira@hotmail.com, norilersnd@gmail.com, angelo.b.garcia3@uel.br, upaduapereira@uel.br, Ulisses.rocha@ufz.de, admilton@uel.br, glauciabressan@utfpr.edu.br,

Abstract. *This study aims to analyze the genomic characteristics of a plant growth-promoting bacterium, focusing on genome assembly and annotation, as well as phylogenetic and metabolic analyses. The primary objective is to identify and validate genes related to potassium and phosphate solubilization, as well as other primary and secondary metabolic pathways. The research involves the selection of a strain, cultivation, DNA extraction, sequencing, genome assembly and annotation, followed by phylogenetic and comparative analyses. The analyzed Bacillus nitratireducensLABIM48 strain shows high genomic conservation, indicating a common origin and conserved evolution. Phylogenetic analyses confirmed that the strain belongs to the same mentioned species, with the potential to produce bioactive compounds of interest identified in silico. The ability to solubilize phosphate and potassium was validated through in vitro tests.*

Resumo. *Este estudo busca analisar as características genômicas de uma bactéria promotora de crescimento de plantas, com foco na montagem e anotação do genoma, além de análises filogenéticas e metabólicas. O objetivo principal é identificar e validar genes relacionados à solubilização de potássio e fosfato, bem como a outras vias metabólicas primárias e secundárias. A pesquisa envolve a seleção de uma linhagem, cultivo, extração de DNA, sequenciamento, montagem e anotação genômica, seguida por análises filogenéticas e comparativas. A linhagem de Bacillus nitratireducensLABIM48 analisada apresenta alta conservação genômica, indicando uma origem comum e evolução conservada. As análises filogenéticas confirmaram que a linhagem pertence à mesma espécie mencionada, com potencial para produzir compostos bioativos de interesse identificados in silico. A capacidade de solubilizar fosfato e potássio foi validada por testes in vitro.*

1. Introduction

Bioinformatics has played a fundamental role in agriculture and modern biotechnology, revolutionizing the way we analyze and understand complex biological data. By integrating biology, computer science, and statistics, bioinformatics facilitates the processing of large volumes of data, especially those generated by commonly used sequencing technologies like Illumina MiSeq [Sarvendra 2024]. This interdisciplinary field has enabled significant advances in genomics, allowing for the mapping and understanding of essential biological processes in agricultural organisms, such as plant growth-promoting rhizobacteria (PGPR) [Tongrui et al. 2024]. PGPR are microorganisms capable of interacting with plants and promoting their growth through various mechanisms, including nitrogen fixation, nutrient solubilization (such as phosphorus and potassium), and the production of bioactive compounds like phytohormones and natural antibiotics [Debasis et al. 2024]. These microorganisms have garnered significant interest in agricultural biotechnology due to their potential to replace or complement the use of synthetic fertilizers, offering a more sustainable and environmentally friendly alternative for crop management [Mahawar 2024].

Within the PGPR group, bacteria of the genus *Bacillus* have stood out for their beneficial properties for plant growth. In particular, *Bacillus nitratireducens* is a species with great potential in the agricultural context [Ramírez-Pool et al. 2024]. This bacterium can perform nutrient solubilization processes and produce metabolites that aid in plant growth. Moreover, genomic and phylogenetic analyses of these bacteria provide valuable insights into their capabilities and how they can be optimized for use as agricultural bioinputs [Kalyanasundaram et al. 2020].

The current agricultural model, heavily dependent on synthetic fertilizers, faces environmental and economic challenges. The excessive use of these inputs has caused soil degradation, contamination of groundwater, and greenhouse gas emissions, compromising the sustainability of agricultural production in the long term [Faizan et al. 2024]. Additionally, the production and use of these fertilizers involve high costs, especially in countries that rely on the import of chemical inputs. In this context, the identification and use of growth-promoting bacteria, such as *B. nitratireducens*, emerge as a promising alternative [Fernando et al. 2024]. By promoting nutrient absorption by plants and contributing to the biological balance of the soil, these bacteria can significantly reduce the need for synthetic fertilizers. Therefore, a detailed study of the genomic, phylogenetic, and metabolic characteristics of these bacteria is essential for their development as efficient bioinputs that can be applied on a large scale in sustainable agriculture [Abbas et al. 2024].

The study of the genomics and phylogeny of plant growth-promoting bacteria, such as *B. nitratireducens*, is of great importance for the advancement of sustainable agriculture. By better understanding the genetic and metabolic mechanisms of these bacteria, it becomes possible to develop more efficient and specific bioinputs for different types of soil and crops. These bioinputs can not only reduce the use of chemical fertilizers but also improve soil health and crop productivity, contributing to more sustainable agricultural practices [Zia 2024]. The use of bioinformatics tools to analyze these microorganisms allows for a more precise and effective approach in identifying strains with high application potential. In this sense, the present work represents an important

contribution to innovation in the agricultural sector by proposing science- and technology-based solutions that can reduce environmental impacts and production costs, promoting a more sustainable future for agriculture [Dixit et al. 2023]

The present study aims to analyze the genomic and phylogenetic characteristics of a potentially plant growth-promoting bacterium, with a special focus on the *B. nitratireducens*LABIM48 strain. The research seeks to understand the mechanisms by which this bacterium promotes plant growth, identifying genes involved in potassium and phosphate solubilization, as well as the production of primary and secondary metabolites. Additionally, the study aims to validate, through in vitro tests, the potential of this bacterium to promote plant growth. More specifically, the objectives include the assembly and annotation of the genomes of the LABIM48 strain, phylogenetic analyses, investigation of its metabolic profile, and identification of genes related to nutrient solubilization, as well as the in vitro validation of potassium and phosphate solubilization capabilities predicted by genomic analysis.

2. Materials and Methods

The microbial strain selected for this study was obtained from the collection of the Microbial Biotechnology Laboratory (LABIM) at the State University of Londrina (UEL), located in Londrina, Paraná. LABIM is widely recognized for its expertise in agricultural microbiology, which underscores the importance of the strains used. This strain was collected in the region of Pontal do Paraná - PR on January 20, 2020, at coordinates 25°36'06.2"S and 48°23'20.8"W. The careful selection of the strain and its collection site ensured that the data obtained were relevant and of high quality for the research. The bacterial strain was initially cultured on Tryptic Soy Agar (TSA) medium.

Subsequently, isolated colonies were inoculated into Tryptic Soy Broth (TSB) and incubated in a shaker at 150 rpm at 28°C for 18 hours. DNA extraction was performed using the DNeasy PowerSoil Pro kit (QIAGEN), following the recommended protocol. The purity and quality of the DNA were verified by gel electrophoresis and concentration measurement using Nanodrop. The extracted DNA was then sent for sequencing at the Helmholtz Centre for Environmental Research (UFZ) in Leipzig, Germany, where the MiSeq sequencing system was used with the MiSeq Reagent V2 Micro Kit (300 cycles).

After sequencing, the raw reads of the bacterial strain were deposited in the National Center for Biotechnology Information (NCBI), allowing public access for future research. The sequencing data underwent rigorous bioinformatics analysis for genome assembly. First, the TRIMMOMATIC v0.39 tool was used for trimming, removing low-quality sequences and adapters with a Phred quality filter of ≥ 30. Next, the assembly of contigs was performed with the SPADES v3.13 software, which uses de Bruijn graphs to assemble genomes. To confirm the strain's identity, the assembled contigs were compared to the GenBank database via BLASTn, and sequences with identity and coverage greater than 90% were selected for further analysis.

To validate the strain's identification, the corresponding sequences were analyzed using the Orthologous Average Nucleotide Identity (OAT) tool, where Genome-to-Genome Distance Calculator (GGDC) and Average Nucleotide Identity (ANI)

parameters were applied, with identity criteria above 95%, confirming the strain's identity. After confirmation, the assembly was refined using the MeDuSa v1.0 tool, which aligned the contigs with reference genomes, while the RagTag v2.1 tool was used to close gaps in the assembly, resulting in a more complete genome. The quality of the assembly was assessed with QUAST v5.2, which calculated metrics such as the number of contigs, the size of the largest contig, and the N50 value, providing an objective evaluation of the assembly quality relative to the reference genome.

After the genome's identification and assembly, strains related to *Bacillus nitratireducens* were downloaded from GenBank and subjected to phylogenetic analyses using the Type Strain Genome Server (TYGS). Using algorithms such as RAxML and FastME, a phylogenetic tree was constructed based on complete genomes, with bootstrap support to evaluate the robustness of evolutionary relationships. The 16S rRNA gene sequence was also analyzed using the maximum likelihood method, with bootstrap support. To represent the circular genome, the analyzed genomes were aligned with the refseq BM02 using the MeDuSa software. The largest scaffold of each strain was used for the representation of the circular genome.

To characterize the metabolic profile of this *Bacillus nitratireducens* strain, the GUTSMASH and ANTISMASH software were used to identify biosynthetic loci related to the production of primary and secondary metabolites. Additionally, an extensive literature review was conducted to identify genes associated with potassium and phosphate solubilization. These gene sequences were compared with databases such as UNIPROT to obtain information about their functions and structures and were verified in the strain's genome using tBLASTn

Finally, potassium and phosphate solubilization tests were performed on Petri dishes at 28°C for one month, in quintuplicate. Aleksandrov medium was used for potassium solubilization, and NBRIP medium was used for phosphate. During the incubation period, changes such as halo formation and color changes around the colonies were observed, indicating the strain's ability to solubilize these nutrients.

These tests, along with phylogenetic and genomic analyses, provided a comprehensive understanding of the LABIM48 strain, allowing the evaluation of its biotechnological potential, particularly concerning sustainable agriculture and its use as biofertilizers.

3. Results and Discussion

The genomic analysis of the *Bacillus nitratireducens* strain LABIM48, revealed a high level of genetic conservation, suggesting a common evolutionary origin. This conservation was observed both in the global genome alignment and in the comparison of specific regions (Figure 1). The genomic similarity reflects a conserved evolution within the species, indicating that strain retain important functional characteristics in common, potentially making them capable of promoting plant growth.

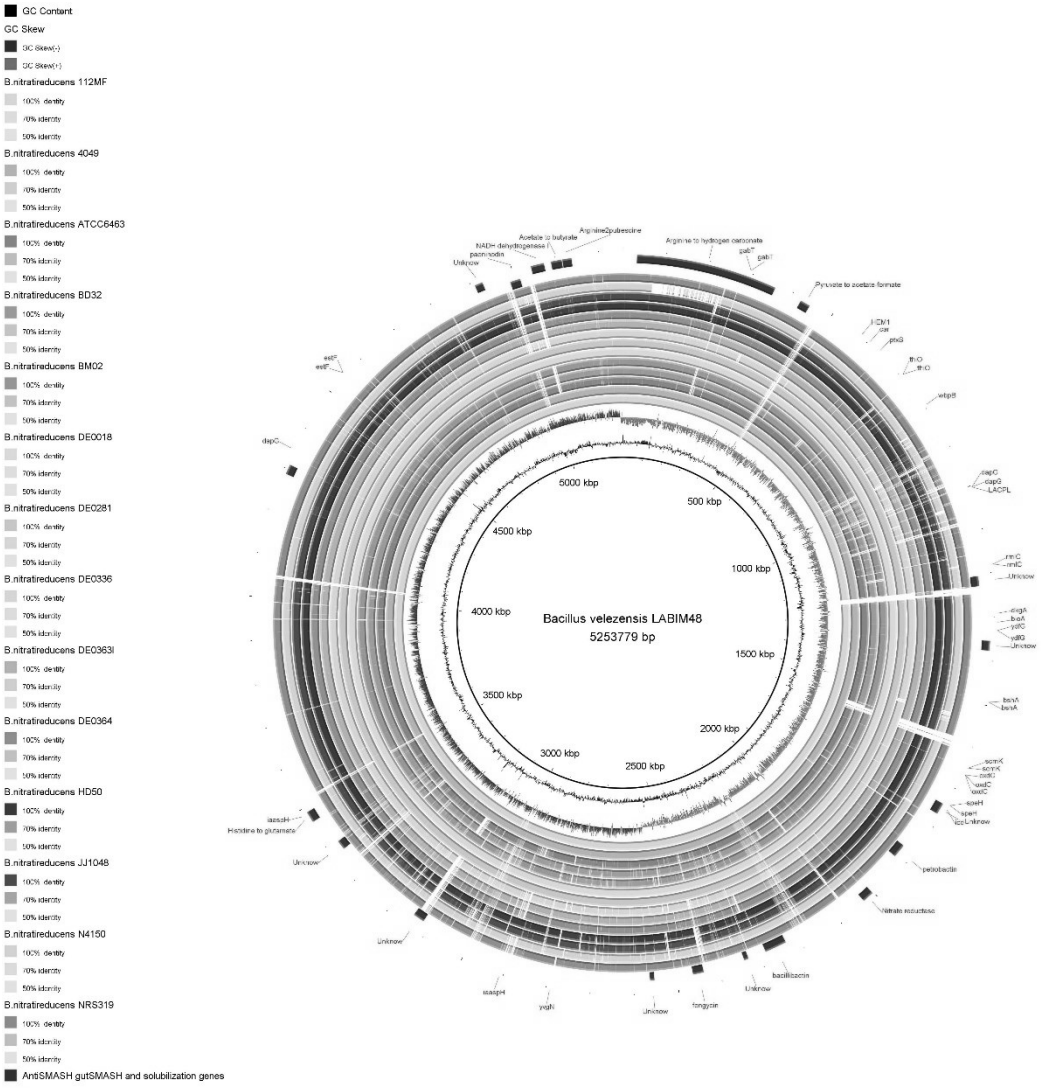

Figure 1. Circular genomic map of *LABIM48*, highlighting the GC content, GC skew, identity comparisons among strains of the same species, and genes involved in potassium and phosphate solubilization.

These findings are consistent with previous studies, such as Yin et al. [2019], which also found high genomic conservation in strains of *Bacillus velezensis*, suggesting that closely related strains maintain stable characteristics over time, facilitating their use as agricultural bioinputs. The high genetic conservation observed supports the idea that this strain can be studied for potential practical applications across different soils and crops.

The phylogenetic analysis, based on both complete genomes and the sequences of the ribosomal 16S rRNA gene (Figures 2 and 3), enabled the identification of evolutionary relationships among the strain from the Microbial Biotechnology Laboratory (LABIM) at UEL and other strains available in the GenBank datasets of the species *Bacillus nitratireducens* (14 strains), with another species from the genus *Bacillus* used for rooting the trees. The phylogenetic trees, constructed using maximum likelihood methods supported by bootstrap parameters, placed the studied strain within a well-defined clade, confirming that it belongs to the species *Bacillus nitratireducens*.

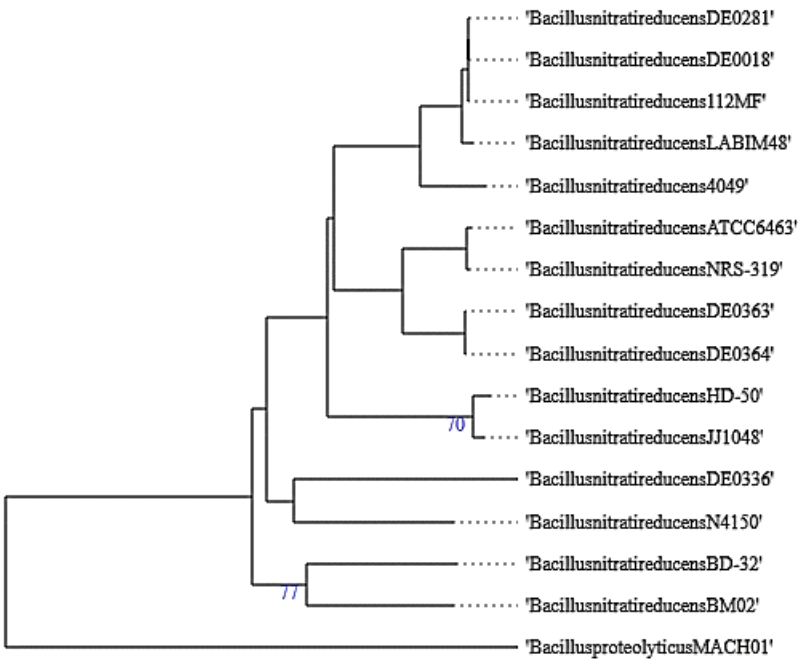

Figure 2. Phylogenetic tree of *Bacillus nitratireducens* based on complete genomes constructed using the maximum likelihood method.

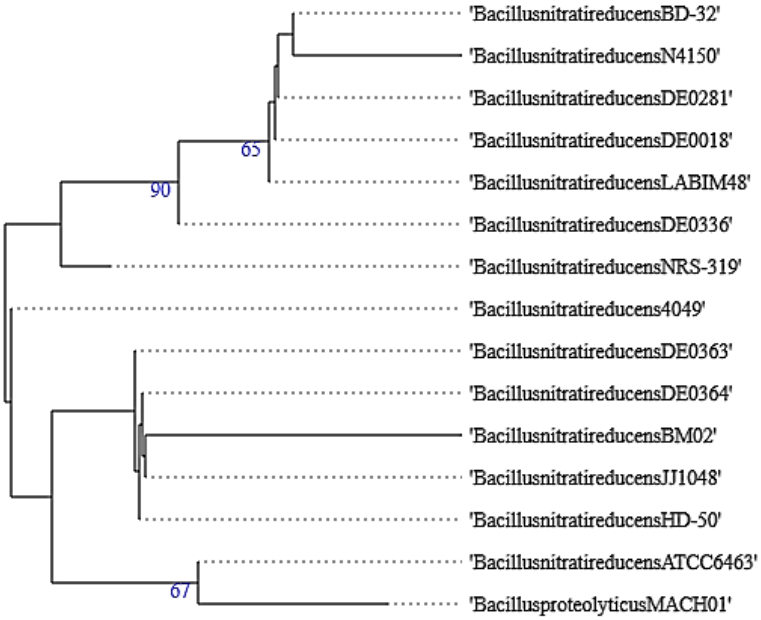

Figure 3. Phylogenetic tree of *Bacillus nitratireducens* based on the sequence of the 16S rRNA gene, constructed using maximum likelihood.

Phylogenetic analyses based on complete genomes (Figures 2 and 4) provide a detailed view of genetic variations among strains, allowing for the identification of subtle differences [Christensen and Olsen 2022]. This approach also facilitates the understanding of the conservation of metabolic pathways and the evolution of strains, as observed in plant growth-promoting bacteria, particularly those of the *Bacillus* spp. genus [Das et al. 2024]

Effective for taxonomic identification at the genus and species level, the ribosomal 16S rRNA gene, widely used due to its high conservation, improves phylogenetic resolution, allowing for the distinction of closely related strains as well as

understanding intraspecific differences among to other freely accessible bacteria of the same species [Nichols and Davenport 2024].

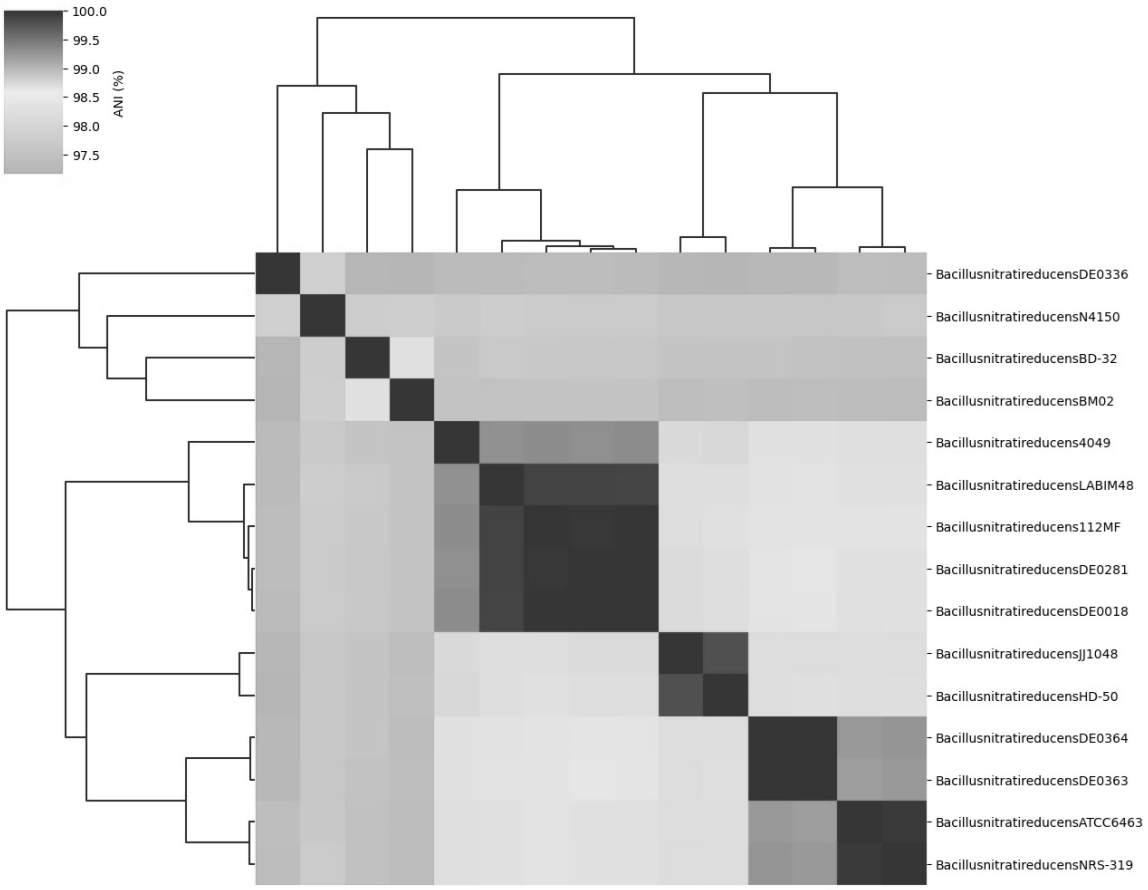

Figure 4 - Genomic Similarity Heatmap with the sixteen genomes analysed

The ANI comparison between *Bacillus nitratireducens* strains, where dark gray shades indicate high genomic similarity and light gray represents lower identity. Strains with more dark gray tones are genetically very close, while light gray areas show greater diversity. The dendrogram reflects these relationships, grouping strains with higher identity.

This analysis helps identify genetic proximity, useful for taxonomic studies and biotechnology applications, such as the development of agricultural biopesticides, by highlighting strains that are more similar and those with greater potential for functional diversity [Zia et al. 2024].

In the metabolic profile, *in silico* mining and gene analyses revealed (Table 1) the potential presence of metabolic pathways responsible for the production of organic acids related to potassium and phosphate solubilization, such as Gluconic acid, acetic acid, Fumaric acid, among others. Additionally, coding regions for primary (Table 2) and secondary metabolites (Table 3) were identified, providing insights into the potential of that strain to produce bioactive compounds of interest.

Table 1. Presence of genes related to potassium and phosphate solubilization.

Acid	Coverage	Identity	Expected action
Aspartic	100%	67.08%	Potassium solubilization
Pyruvic	100%	79.53%	Potassium solubilization
Acetic	100%	33.94%	Potassium and phosphate solubilization
Gluconic	99%	73.72%	Potassium and phosphate solubilization
Succinic	97%	64.79%	Potassium and phosphate solubilization
Fumaric	95%	58.13%	Potassium and phosphate solubilization
Malic	98%	61.46%	Potassium and phosphate solubilization
Isocitric	99%	61.83%	Phosphate solubilization
Ascorbic	98%	49.80%	Phosphate solubilization

Table 2. Identification of coding regions for primary metabolites

Cluster	**Metabolic product**	**Metabolic function**	**Associated enzyme(s)**
1	Arginine → Hydrogencarbonate	Nitrogen metabolism	Arginine degradation pathway enzymes
2	Pyruvate → Acetate-formate	Energy metabolism	Pyruvate Dehydrogenase Complex
3	TPP AA Metabolism	Amino acid synthesis	Transaldolase, Transketolase
4	Nitrate Reductase	Nitrogen cycle	Nitrate Reductase
5	Histidine → Glutamate (HutHGIU Operon)	Nitrogen Metabolism and Protein Synthesis	HutH, HutG, HutI, HutU
6	Fumarate → Succinate	Citric Acid Cycle	Fumarate Reductase
7	Putrescine → Spermidine	Polyamine synthesis	Ornithine Decarboxylase, Spermidine Synthase
8	NADH Dehydrogenase I	Electron Transport Chain	NADH Dehydrogenase
9	Acetate → Butyrate	Short-Chain Fatty Acid Production	Butyryl-CoA Dehydrogenase, Acetate Kinase
10	Arginine → Putrescine	Polyamine synthesis	Arginine Decarboxylase
11	Leucine (Reductive Branch)	Branched-Chain Amino Acid Synthesis	Leucine Reductase

Table 3 – Identification of coding regions for secondary metabolites.

Strain	Number of coding regions	Type of secondary metabolite	Main target metabolites
LABIM48	9	LAP, NI-siderophore, NRP-metallophore, NRPS, RiPP-like (2), Betalactone, Lanthipeptide-class-i, Terpene, Lassopeptide	Petrobactin, Bacillibactin, Fengycin, Paenimodin

Their arginine and histidine metabolisms, which produce compounds like bicarbonate and polyamines, are essential for survival in adverse conditions and nitrogen regulation. In anaerobic environments, the use of the pyruvate dehydrogenase complex and fumarate reductase is crucial for energy production [Rufus et al. 2024]. These metabolic pathways are exploited to enhance the production of enzymes, antibiotics, and biofertilizers, demonstrating the biotechnological potential of the genus [Jain et al. 2024].

Table 2 shows the coding regions and types of secondary metabolites produced by the LABIM48 strain. The strain also produces siderophores and antimicrobial compounds, as well as other types such as Lanthipeptide-class-i and Ranthipeptide. This metabolite synthesis capability could maximize its use in agricultural biotechnology and pathogen control [Yin et al. 2023].

The in vitro results confirmed that the *Bacillus nitratireducens*LABIM48 strain is effective in solubilizing both nutrientes (Figures 5 and 6). This capability was correlated with the production of organic acids, such as citric acid, acetic acid, and aspartic acid, which are known for their role in mineral solubilization in soil, aligning with the work of Meena et al. (2016), who also identified these acids as important agents for potassium and phosphorus solubilization.

Figure 5 - Potassium solubilization of *Bacillus nitratireducens* strains LABIM48

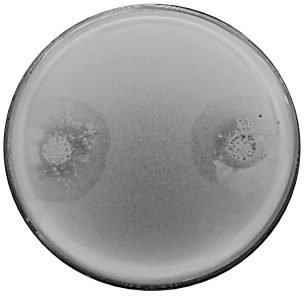

Figure 6 - Phosphate solubilization of *Bacillus nitratireducens* strains LABIM48

The *in vitro* tests provided experimental evidence that theat strain have significant biotechnological potential for improving soil fertility. Complementarily, studies such as Setiawati et al. (2022) confirmed that *Bacillus* strains are effective in nutrient solubilization under various soil conditions, further reinforcing the applicability of that strain in different agricultural contexts.

The implications of these results for sustainable agriculture are clear. The use of that *B. nitratireducens*LABIM48 strain as bioinputs can significantly reduce the need for synthetic fertilizers, promoting more sustainable soil management. The ability to effectively solubilize potassium and phosphate provides a natural source of nutrients for plants, which can minimize the environmental impacts associated with the use of chemical inputs. These results align with studies like Bahadur et al. (2019), which demonstrated that the application of potassium solubilizers can increase plant productivity while reducing the use of chemical fertilizers. Therefore, the use of that LABIM48 strain as agricultural bioinputs represents a viable and promising alternative for sustainable agricultural practices, contributing to soil preservation and improved agricultural productivity.

The genomic, phylogenetic, and metabolic analyses conducted in this study indicate a high genetic conservation among all the *Bacillus nitratireducens* strains analyzed, suggesting a common origin and stable biological characteristics. The ability of that strain to solubilize potassium and phosphate, confirmed by both *in silico* and *in vitro* experiments, highlights their great potential as agricultural bioinputs. This offers direct benefits for sustainability and productivity in the field, further emphasizing the crucial role of bioinformatics in accelerating the development of innovative solutions for agriculture.

4. Conclusion

This work provided a comprehensive analysis of the genomic, phylogenetic, and metabolic characteristics of the *Bacillus nitratireducens* strain LABIM48. The results highlight the close relationship between all the analyzed strains, as evidenced by their high genomic and phylogenetic similarity. Genomic analysis revealed a high level of conservation in the genome of this strain, indicating a recent common origin and conserved genetic evolution. Phylogeny based on the complete genome and the 16S rRNA gene confirmed that the strain belongs to the species *Bacillus nitratireducens*. Metabolic analysis identified regions encoding primary and secondary metabolites in this strain, demonstrating its potential capacity to produce a variety of bioactive compounds important for microbial survival and competition. Additionally, in vitro tests of phosphate and potassium solubilization corroborated the predictions made by the computational analysis of organic acids, demonstrating the strain's ability to solubilize these essential nutrients.

References

Abbas, N., Mazhar, S., Essa, R., Bukhari, B., Nadeem, A.A., Saleem, Y., Syed, Q., Abidi, S.H. and Riaz, S. (2024) 'Exploring Azotobacter species as soil biological enhancers for enhanced crop nutrition and stable yields: a review', doi: 10.32388/ovm1db.

Bahadur, I. et al. (2019) 'Potassium-solubilizing bacteria (KSB): a microbial tool for K-solubility, cycling, and availability to plants', in *Plant Growth Promoting Rhizobacteria for Agricultural Sustainability: From Theory to Practices*, pp. 257-265.

Christensen, H. and Olsen, J.E. (2022) 'Introduction to phylogenetic analysis of molecular sequence data', *Learning Materials in Biosciences*, doi: 10.1007/978-3-031-45293-2_6.

Das, V.M., Gautam, B., Yadav, P.K. and Singh, S. (2024) 'Identification of conserved pathways in *Bacillus* strains known for plant growth-promoting behavior using a multifaceted computational approach', *Agriculture*, doi: 10.3390/agriculture14060838.

Debasis, Mitra., Marika, P. and Beatriz, E.G.S. (2024) 'Interaction between plants and growth-promoting rhizobacteria (PGPR) for sustainable development', *Bacteria*, doi: 10.3390/bacteria3030009.

Dixit, R., Kapoor, N., Singh, N. and Kumar, A. (2023) 'Bioinformatics resources for sustainable agriculture', doi: 10.4018/979-8-3693-3026-5.ch015.

Faizan, M., Singh, A., Eren, A., Sultan, H., Sharma, M., Djalović, I. and Trivan, G. (2024) 'Small molecule, big impacts: nano-nutrients for sustainable agriculture and food security', *Journal of Plant Physiology*, doi: 10.1016/j.jplph.2024.154305.

Fernando, G.R., Abeysinghe, S., Rajapakse, R.G.S.C. and Gamage, W.S.M.K. (2024) 'Identification of plant growth-promoting rhizobacteria for the development of probiotic consortia', *Proceedings of the International Forestry and Environment Symposium*, doi: 10.31357/fesympo.v28.7039.

Jain, P.M., Cheruvari, A., Tamilselvan, H.M.J. and Kammara, R. (2024) '*Bacillus subtilis* comparative genomics', doi: 10.1099/acmi.0.000810.v1.

Kalyanasundaram, G.T., Syed, N. and Subburamu, K. (2020) 'Recent developments in plant growth-promoting rhizobacteria (PGPR) for sustainable agriculture', doi: 10.1016/B978-0-12-821406-0.00017-5.

Mahawar, S. (2024) 'Bioinoculants: an ecofriendly approach towards artificial fertilizers in sustainable agriculture', doi: 10.58532/v3bfbt1p1ch21.

Nichols, R.G. and Davenport, E. (2024) 'Clade-specific long-read sequencing increases the accuracy and specificity of the gyrB phylogenetic marker gene', doi: 10.1101/2024.07.15.603574.

Ramírez-Pool, J.A., Calderón-Pérez, B., Ruiz-Medrano, R., Ortiz-Castro, R. and Xoconostle-Cázares, B. (2024) '*Bacillus* strains as effective biocontrol agents against phytopathogenic bacteria and promoters of plant growth', *Microbial Ecology*, doi: 10.1007/s00248-024-02384-1

Rufus, A.C., Kalai, K. and Elangovan, E. (2024) 'Isolation and characterization of *Bacillus* species for the detection of siderophores', *Research Journal of Biotechnology*, doi: 10.25303/1908rjbt092099.

Sarvendra, V.S. (2024) 'Bioinformatics – Supporting modern life science research, applications, and challenges', *Brazilian Journal of Development*, doi: 10.34117/bjdv10n2-011.

Setiawati, T.C. et al. (2022) 'Use of *Bacillus* as a plant growth-promoting rhizobacteria to improve phosphate and potassium availability in acidic and saline soils', *KnE Life Sciences*, pp. 541–558.

Tongrui, Zhang., Qinhao, J., Li, G., Linlin, L., Fei, L. and Can, Z. (2024) 'Plant growth-promoting rhizobacteria (PGPR) improve the growth and quality of several crops', *Heliyon*, doi: 10.1016/j.heliyon.2024.e31553.

Yin, Q.J., Ying, T.T., Zhou, Z.Y., Hu, G., Yang, C., Hua, Y., Wang, H. and Wei, B. (2023) 'Species-specificity of the secondary biosynthetic potential in Bacillus', *Frontiers in Microbiology*, doi: 10.3389/fmicb.2023.1271418.

Zia, A. (2024) 'Plant growth promoting rhizobacteria: a novel approach towards sustainable agriculture', *Deleted Journal*, doi: 10.37446/corbio/ra/2.1.2024.16-28.