Original Article

Genome Analysis of *Bacillus paralicheniformis* AA1 isolated from a conventional milpa farming system in the southern region of Sonora, Mexico

Análisis del genoma de *Bacillus paralicheniformis* AA1 aislado de un sistema de milpa convencional en la región sur de Sonora, México

Andrés Chávez-Almanza[1]†, Eber Villa-Rodríguez[2]†, Jonathan Rojas-Padilla[1], Ernesto Cantú-Soto[1], Carlos Díaz-Quiroz[a] & Abel Verdugo-Fuentes[1]*✉

[1] Departamento de Biotecnología y Ciencias Alimentarias, Instituto Tecnológico de Sonora, 5 de febrero 818 Sur, CP 85000, Cd. Obregón, Sonora, México

[2] Department of Molecular Biology and Genetics, Aarhus University, Denmark

ABSTRACT

This study provides an in-depth genome analysis of *Bacillus paralicheniformis* AA1, a bacterial strain isolated from a traditional milpa farming system in Sonora, Mexico. The genomic analysis revealed a high level of completeness, demonstrated by the presence of a diverse and functionally significant repertoire of genes associated with fundamental biological processes, including nutrient assimilation, stress response, and cellular regulation. Notably, the genome also contains genes responsible for the biosynthesis of secondary metabolites, highlighting its potential for biotechnological applications. Taxonomic classification w as rigorously c onducted using integrated genome-wide approaches, which definitively confirmed the identification of iso late AA1 as belonging to the species *Bacillus paralicheniformis*. Comparative genomic analysis further established a high degree of genetic similarity between AA1 and other *B. paralicheniformis* strains with well-characterized biotechnological capabilities. This similarity strongly suggests that AA1 harbors genetic elements responsible for the synthesis of antimicrobial compounds, enzymes with industrial relevance, and metabolites that promote plant growth. The findings underscore the potential of *Bacillus paralicheniformis* AA1 as a valuable resource for biotechnology and sustainable agriculture. By enhancing our understanding of microbial diversity within traditional agroecosystems, this study contributes to the broader knowledge base required for the development of innovative agricultural practices. Future research should focus on functional validation of key genes to fully unlock AA1's potential as a bioresource for antimicrobial production, enzyme synthesis, and crop enhancement, paving the way for its application in environmentally sustainable farming systems.
Keywords: *Bacillus paralicheniformis*, whole genome, genomic analysis, sustainable agriculture.

RESUMEN

El presente estudio presenta un análisis genómico de *Bacillus paralicheniformis* AA1, una cepa bacteriana aislada de un sistema agrícola tradicional de milpa en Sonora, México. El análisis genómico reveló un alto nivel de integridad, evidenciado por la presencia de un repertorio diverso y funcionalmente significativo de genes asociados con procesos biológicos fundamentales, incluyendo la asimilación de nutrientes, la respuesta al estrés y la regulación celular. En particular, el genoma también contiene genes responsables de la biosíntesis de metabolitos secundarios, lo que resalta su potencial para aplicaciones biotecnológicas. La clasificación taxonómica se realizó rigurosamente mediante enfoques genómicos integrados, confirmando la identificación de la cepa AA1 como perteneciente a la especie *B. paralicheniformis*. El análisis genómico comparativo estableció además un alto grado de similitud genética entre *B. paralicheniformis* AA1 y otras cepas de *B. paralicheniformis* con capacidades biotecnológicas caracterizadas. Esta similitud sugiere que AA1 posee elementos genéticos responsables de la síntesis de compuestos antimicrobianos, enzimas de relevancia industrial y metabolitos que promueven el crecimiento vegetal. Los hallazgos resaltan el potencial de *B. paralicheniformis* AA1 como una cepa valiosa para la biotecnología y la agricultura sostenible. Este análisis contribuye al conocimiento necesario para el desarrollo de prácticas agrícolas innovadoras al mejorar nuestra comprensión de la diversidad microbiana en agroecosistemas tradicionales. Las investigaciones derivadas de este estudio deberán centrarse en la validación funcional de genes clave para comprender completamente el potencial de AA1 en la producción de antimicrobianos, la síntesis de enzimas y su aplicación en sistemas agrícolas sostenibles.
Palabras clave: *Bacillus paralicheniformis*, genoma completo, análisis genómico, agricultura sustentable.

INTRODUCTION

Bacteria exhibit a ubiquitous presence across diverse environments, spanning terrestrial and marine habitats, where they play diverse ecological roles (Delgado-Baquerizo *et al.*, 2018). Among these, a range of different species within the genus *Bacillus* are commercially available biopesticides because of their capacity to synthesize a diverse group of antimicrobials, nematicidal, and insecticidal metabolites (Steinke *et al.*, 2021; Su *et al.*, 2020).

In recent years, *Bacillus paralicheniformis* has gained attention for its remarkable characteristics, including its ability to produce antimicrobial bioactive substances and thrive in challenging environments (Ashajyothi *et al.*, 2024; Iqbal *et al.*, 2023). From an agrobiotechnology perspective, *B. paralicheniformis* is an interesting source for biofertilizer development due to its multifaceted capabilities (Swietczak *et al.*, 2023). These include the production of bioactive molecules that stimulate plant growth (*e.g.*, indoleacetic acid, siderophores, ACC deaminase), enhancement of plant nutrition through nutrient solubilization (*iron*, phosphorus, and potassium), and inhibition of phytopathogen attacks either via antimicrobial production or by triggering plant immune responses (Chavarria-Quicaño *et al.*, 2023; Jinal *et al.*, 2020). This particular *Bacillus* specie has demonstrated successful applications across a variety of economically significant crops, including tomato (Díaz-Manzano *et al.*, 2023), mango (Hemangini *et al.*, 2018), chili (Pawaskar, 2023), highlighting its biotechnological relevance.

Milpa farming systems, characterized by the cultivation of multiple crop species such as maize, beans, herbs, and grass, have been linked to elevated soil microbial diversity in comparison to monocultures (Fonteyne *et al.*, 2023). In these highly diverse microbiomes, microbe-microbe interactions, involving both competition and cooperation, are naturally intensified. Consequently, microbes thriving in such environments often possess an interesting set of genes enabling them to effectively compete and collaborate with other microbial inhabitants, while also adapting to the selective pressures exerted by the varied root exudates of different plant species (Compant *et al.*, 2019). This characteristic makes milpa production systems a promising reservoir of microbes with interesting biotechnological potential. In this work, we present the genome assembly of one of the most abundant bacterial isolates (strain AA1) recovered from a conventional milpa farming system in the northwestern region of Sonora, Mexico. The aim of this genome sequence announcement is to report the draft genome sequence of *Bacillus paralicheniformis* strain AA1, a rhizospheric bacterium isolated from a traditional milpa farming system in Sonora, Mexico. This genome provides insights into the genetic basis of its potential for biotechnological applications, including secondary metabolite biosynthesis, stress response, and plant growth promotion, and contributes to the understanding of microbial diversity in traditional agroecosystems.

MATERIALS AND METHODS
Sample collection
Soil samples were collected from a milpa production system at the Technological Institute of Sonora in the northwest region of Sonora, Mexico. The location is 27°29'46"N 109°58'18"W, and the elevation is 40 m above sea level. During May, the area typically receives temperatures of 27.5 °C and an average rainfall of 2.0 mm (based on a local weather station, CONAGUA). During the vegetative phase of the maize growth cycle in May 2023, three composite soil samples weighing about 1 kg each were randomly obtained from three separate locations. The depth at which the samples were collected was 20 cm, using sterile tools and sampling bags. The soil samples were subsequently examined for microbial analyses in the laboratory after being carried in a cooler maintained at 4 ± 2 °C.

Microbiological analysis
Bacillus isolation was carried out as follows; ten grams of each sample were dissolved in 90 mL of distilled water and heated to 80 °C for 15 min (Wen *et al.*, 2022). After diluting the solution with sterile water, 0.1 mL of the mixture was equally spread out on Luria Bertani (LB) (Sigma-Aldrich®) agar plates. The plates were then placed in an aerobic incubator at 37 °C for 24 h. The colonies were then prepared for microscopic morphological identification using Gram staining and subjected to purification. Among the isolates, strain AA1 was one of the most abundant and consistently observed colony types across replicate plates. This strain was purified by repeated streaking on fresh nutrient agar plates until a uniform, morphologically consistent colony phenotype was obtained, ensuring clonality of the isolate. The purified strain AA1 was stored at –80 °C in LB broth (Sigma-Aldrich®) supplemented with 20% glycerol for further molecular and genomic analyses (Lavanya *et al.*, 2021).

DNA extraction and sequencing
Strain AA1 was grown in Trypticase Soy Broth (TSB) at 37 °C for 24 h. An axenic culture of the strain was subsequently sent to The Sainsbury Laboratory (Norwich, United Kingdom), where the downstream genomic procedures were carried out. At The Sainsbury Laboratory, genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Cat. No. 69504), following the manufacturer's instructions. The purity and concentration of the extracted DNA were assessed spectrophotometrically. Library preparation and high-throughput sequencing were then performed on the Illumina NovaSeq platform (2x250 bp), according to standard protocols. All procedures related to DNA isolation, quality control, and sequencing were conducted by The Sainsbury Laboratory team.

Bioinformatic analyses
Genome assembly was conducted using SPAdes genome assembler (version 3.15.4) (Bankevich *et al.*, 2012). The resulting assembly was submitted to NCBI Prokaryotic Genome Annotation (PGAP) pipeline for i) prediction of protein-coding genes, ii) determination of RNAs, tRNAs, and pseudogenes, iii) gene annotation and iv) genome completeness and contamination assessment (Tatusova *et al.*, 2016). Additionally, KOs and COGs assignments were executed using anvi'o v8 (Eren *et al.*, 2020), these annotations provide insights into the metabolic potential and ecological roles of the strain by linking genes to known biological pathways (via KOs) and to evolutionarily conserved functional categories (via COGs). Phylogenomics analysis was performed using GtoTree pi-

peline (Lee, 2019), while a pangenomics-based tree and ANI identity matrix were generated using anvi'o v8 (Eren *et al*., 2020). The genome map was generated using Proksee (Grant *et al*., 2023).

Data availability

The whole genome sequence was submitted to DDBJ/ENA/GenBank and assigned the accession number JA-YMDM000000000. The sequenced strain's BioProject database accession number is PRJNA1061142. The project's Sequence Read Archive information can be accessed using the accession number SRR27458597.

RESULTS AND DISCUSSION

Milpa farming systems represent a valuable reservoir of microorganisms with potential biotechnological applications. To explore this microbial diversity, we collected rhizospheric soil samples from a traditional milpa production system located at the Technological Institute of Sonora in the northwest region of Sonora, Mexico. Bacterial isolation was performed using the serial dilution method on nutrient agar plates. After incubation, colonies with distinct morphological features were selected and subjected to Gram staining to facilitate preliminary identification. Among the isolates, one of the most abundant and consistently observed colonies was designated as strain AA1, which was further purified through successive streaking on fresh agar plates until a morphologically uniform culture was obtained. Strain AA1 was selected for whole genome sequencing based on its stable association with the rhizosphere environment, with the aim of exploring its genetic repertoire potentially linked to plant–microbe interactions and environmental adaptability.

AA1 sequencing results in a high-quality genome assembly

To gain insights into the genetic features facilitating the proliferation of this strain within the studied milpa system, genomic DNA (gDNA) was isolated, sequenced, and assembled the obtained sequencing reads using the SPAdes assembler. The sequencing achieved an average genome coverage of approximately 56× (mean coverage: 55.99; mean coverage excluding 0s: 55.99), considered adequate for high-quality genome assembly. The de novo genome assembly resulted in a total of 39 contigs, 10 contigs with sizes larger than 50 kb, and the largest contig of 1071 kb (Figure 1A and 1B) with a total genome length of 4,318,295 bp, a GC content of 46.06 %, and an N50 value of 701,491 bp. Next, we submitted our genome to the Prokaryotic Genome Annotation Pipeline (PGAP) to identify genome features of AA1 (*e.g*., CDSs, RNAs, tRNAs, and pseudogenes) and perform genome completeness and contamination assessment. This analysis indicated high genome completeness (99.41 %) and no contamination (0%) (Figure 1C). Additionally, PGAP predicted i) a total of 4400 genes, of which 4212 correspond to protein-coding sequences and 75 pseudogenes, ii) 26 ribosomal RNAs (rRNAs), of which 9, 7, and 12 correspond to 5S, 16S, and 23S rRNA,

respectively, and iii) 80 tRNAs and 5 ncRNAs (Figure 1C). According to NCBI Prokaryotic Genome Annotation Standards (Klimke *et al*., 2011), the minimum standards to consider a genome complete are i) at least one copy of each rRNA (5S, 16S, 23S), ii) at least one copy of tRNAs for each amino acid and iii) Protein-coding genes count divided by genome length close to 1. Overall, our genome assembly surpasses the assemblies of previous studies (Berais-Rubio *et al*., 2023; Berriel *et al*., 2021) and meets NCBI standards, indicating suitable assembly to explore the genomic information of this isolate.

The sequenced genomes of *Bacillus paralicheniformis* exhibit a consistent genome size ranging from approximately 4.16 to 4.55 Mb, with a stable GC content around 45.3% to 46.1% (Agersø *et al*., 2019; Albdaiwi *et al*., 2022; Chebotar *et al*., 2024; Olajide *et al*., 2021). These genomes typically encode between 3,981 and 4,478 protein-coding genes, alongside a complement of rRNAs and tRNAs, reflecting a well-conserved genetic architecture (Albdaiwi *et al*., 2022; Olajide *et al*., 2021).

Genome-wide analysis reveals AA1 is a *Bacillus paralicheniformis* strain

Taxonomic identification of bacterial isolates is a crucial step in bioprospection, as it enables us to gain insights into their biotechnological or pathogenic potential. In this work, we initially explore the taxonomic affiliation of strain AA1 by performing a BLAST using 16S rRNA, a highly conserved and widely used prokaryotic taxonomic marker (Kim and Chun, 2014). Thus, we observed that the 16S rRNA sequence of strain AA1 is highly conserved with some isolates of *B. licheniformis* (100% identity/coverage) and *B. paralicheniformis* (100% identity/coverage), suggesting that strain AA1 could belong to one of these taxonomic groups. While useful, 16S rRNA-based bacterial classification has the disadvantage that some bacteria may share high similarity with other members of the same family (Bars-Cortina *et al*., 2023), as observed here.

Hence, we conducted an integrated genome-wide approach to taxonomically classify strain AA1 (Figure 2). First, we performed a phylogenomic analysis using 119 single-copy gene (SCG) markers commonly used in firmicutes (Figure 2A), alongside 38 *Bacillus* representative accessions retrieved from the Genome Taxonomy Database (Parks *et al*., 2022). This analysis facilitated the identification of different *Bacillus* species (*B. paralicheniformis*, *B. licheniformis*, *B. haynesii*, *B. sweneyi*, *B. pumilus*) closely related to AA1 strain. Subsequently, a pangenome-based tree was constructed using high-quality complete genomes (including type strains) of closely related species, obtained from EZbiocloud (www.ezbiocloud.net). This approach clusters genomes based on the presence or absence of predicted coding regions. Our pangenome-based tree reveals three distinct groups, with the AA1 genome present in group 1, clustering alongside *B. paralicheniformis* genomes. This clustering strongly suggests AA1 affiliation with this taxonomic group. Lastly, we conducted an Average Nucleotide Identity (ANI) analysis using

A Assembly information

Number of contigs	Longest contig length (kb)	Shortest contig length (bp)	Number of contigs > 50 kb	L50	N50
39	1071	212	10	3	701

B Genome map
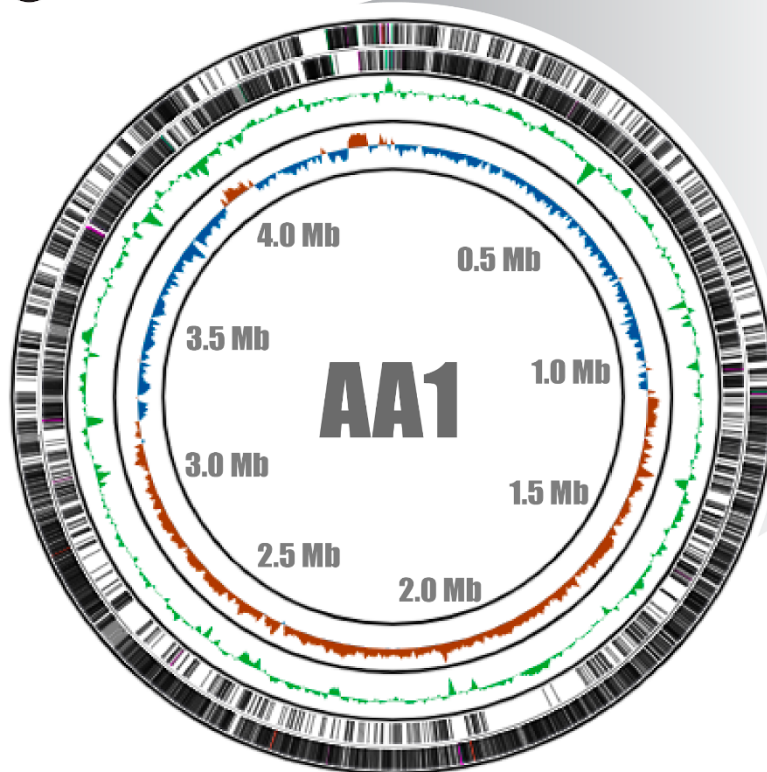
C Genome information

Genome size	4,3 Mb
GC content	46.06 %
Number of genes	4400
Protein-coding	4212
Pseudogenes	75
Ribosomal RNA 5S	9
Ribosomal RNA 16S	7
Ribosomal RNA 23S	12
ncRNAs	5
Transfer tRNAs	80
Completness	99.41 %
Contamination	0 %

Legend:
- CDS
- tRNAs
- rRNAs
- GC content
- GC view skew +
- GC view skew -

Figure 1. Genome assembly and statistics of AA1 genome. A) Quality information of AA1 genome assembly. B) AA1 circular genome map. From the innermost to the outermost rings; GC skew, GC content, CDS in minus (-) strand, CDS in plus (+) strand. C) AA1 genome information determined by PGAP pipeline.

Figura 1. Ensamblaje y estadísticas del genoma de AA1. A) Información de calidad del ensamblaje del genoma de AA1. B) Mapa circular del genoma de AA1. De los anillos más internos a los más externos: sesgo de GC, contenido de GC, CDS en la cadena negativa (-), CDS en la cadena positiva (+). C) Información genómica de AA1 determinada mediante el pipeline PGAP.

genomes within group 1. In bacterial taxonomy, species are typically distinguished based on an ANI threshold of 95% (Arahal, 2014). Our analysis revealed that AA1 exhibits an ANI value higher than 95% when compared to any of the *B. paralicheniformis* genomes included in the study. This result shows that strain AA1 can unequivocally be classified as *Bacillus paralicheniformis*. Noteworthy, this *Bacillus* species has been associated with multiple traits with biotechnological relevance (Hemangini *et al*., 2018; Jinal *et al*., 2020), suggesting that strain AA1 could harbor some of these traits.

Genome annotation and comparative genomics reveal functional traits in *Bacillus paralicheniformis* AA1 associated with biotechnological applications

In bacterial bioprospection, genome annotation is an essential step since it facilitates the identification of the functional elements found in the genome. This process provides valuable insights that could reveal potential biotechnological applications for bacterial isolates. In this sense, we perform the genome annotation of *Bacillus paralicheniformis* AA1 to gain insight into their potential biological functions. To this
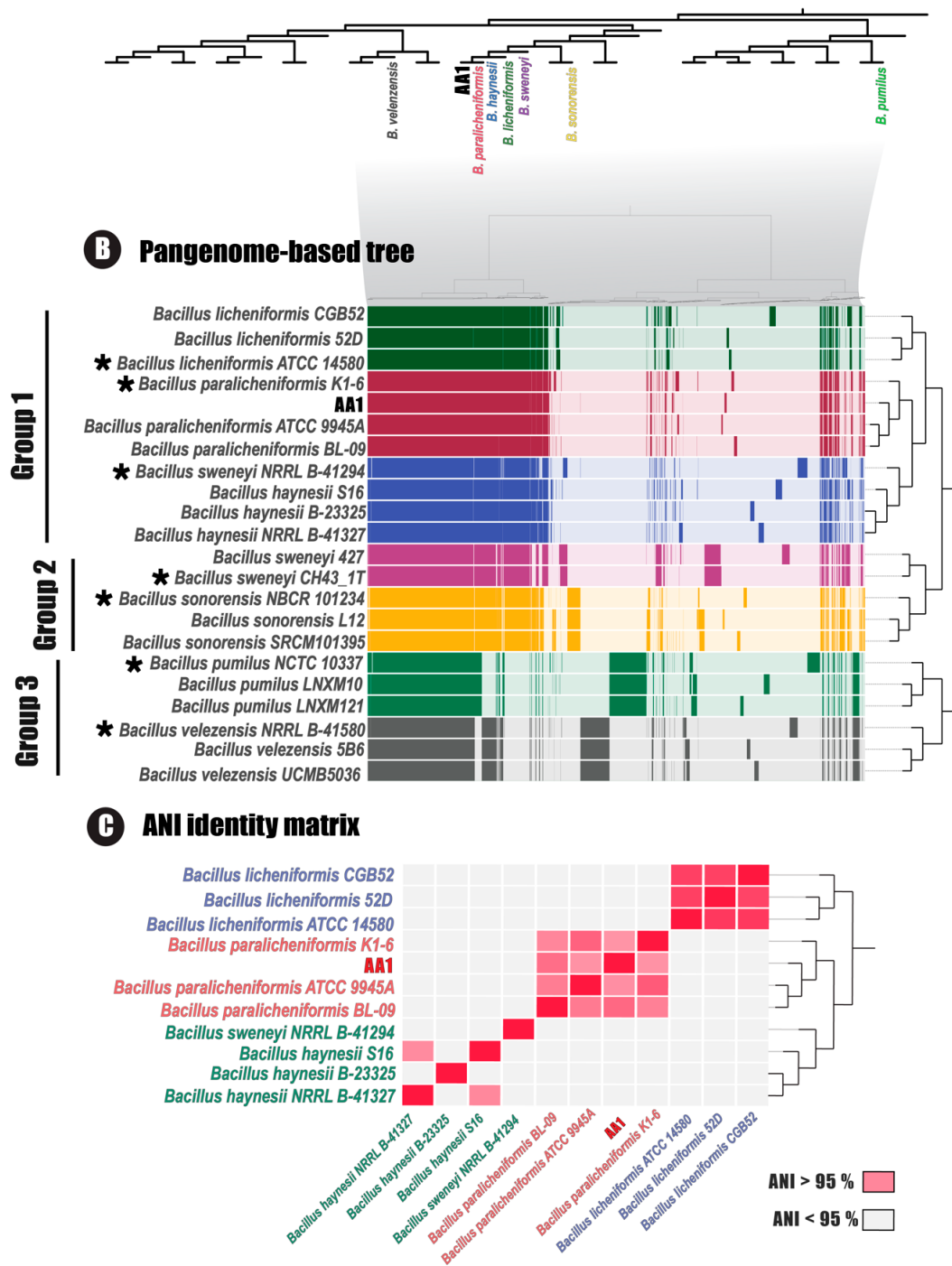
Figure 2. Genome-wide analysis determines AA1 taxonomy. A) Phylogenomic analysis using a single copy gene (SCG) set for Firmicutes (119 genes). B) Pangenome-based tree using closely related species determined in phylogenomic tree using SCG. Asterisks denote type strains retrieved from EZbiocloud (www.ezbiocloud.net). Each line represents a genome. Filled-colored lines (Green to gray for *B. licheniformis*, *B. paralicheniformis*, *B. haynesii*, *B. sweneyi*, *B. pumilus*, and *B. velezensis*, respectively) represent the presence of a gene, while lighter color represents the absence of a gene. C) ANI identity matrix using strains present AA1 clade in the pan genome-based tree.

Figura 2. El análisis a nivel genómico determina la taxonomía de AA1. A) Análisis filogenómico utilizando un conjunto de genes de copia única (SCG) para Firmicutes (119 genes). B) Árbol basado en el pan-genoma utilizando especies estrechamente relacionadas determinadas en el árbol filogenómico con SCG. Los asteriscos indican cepas tipo recuperadas de EZbiocloud (www.ezbiocloud.net). Cada línea representa un genoma. Las líneas coloreadas y rellenas (de verde a gris para *B. licheniformis*, *B. paralicheniformis*, *B. haynesii*, *B. sweneyi*, *B. pumilus* y *B. velezensis*, respectivamente) representan la presencia de un gen, mientras que los colores más claros indican su ausencia. C) Matriz de identidad ANI utilizando cepas presentes en el clado AA1 en el árbol basado en el pan-genoma.

end, we use an integrated approach using different annotation systems (PGAP, KEGG, COG). A total of 3827, 3427, and 2971 genes were annotated with PGAP, COG, and KEGG systems, respectively, displaying multiple overlaps in the annotated genes (Figure 3A). Overall, with our approach, we were able to retrieve biological information from around 91% of the coding sequences (Figure 3A). Within the AA1 genome, there is an overrepresentation of genes linked to fundamental biological processes essential for cellular function. These include genes involved in i) carbohydrate metabolism (345 genes), ii) transcription (303 genes), iii) amino acid metabolism and transport (301 genes), and iv) translation, ribosomal structure, and biogenesis (245 genes). Noteworthy, *Bacillus paralicheniformis* AA1 has 81 genes involved in secondary metabolite biosynthesis (Figure 3B), which suggests that this strain could produce several metabolites with biotechnological potential.
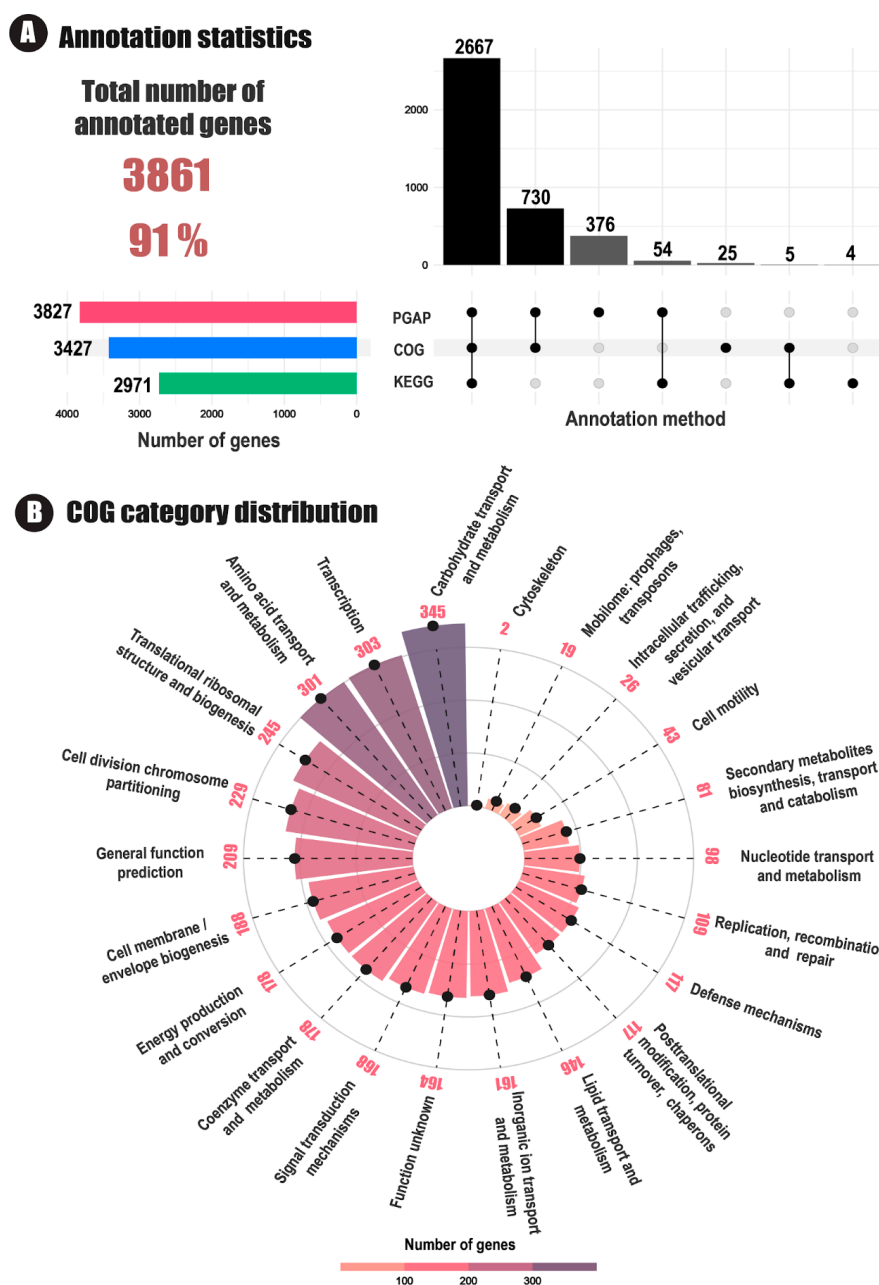

Figure 3. Genome annotation of *Bacillus paralicheniformis* AA1. A) Upset plot displaying overlaps of gene annotations by multiple annotation systems. B) Distribution of Clusters of Orthologous Genes (COG) categories within the AA1 genome, with respective gene counts indicated.
Figura 3. Anotación genómica de *Bacillus paralicheniformis* AA1. A) Gráfico de tipo Upset que muestra la superposición de anotaciones génicas por múltiples sistemas de anotación. B) Distribución de las categorías de Clusters of Orthologous Genes (COG) dentro del genoma de AA1, con el número respectivo de genes indicado.

Over the last years, several *Bacillus paralicheniformis* representatives with biotechnological potential have been studied. This can be exemplified in a recent study where a comprehensive genome analysis of *Bacillus paralicheniformis* BP9 revealed this strain has multiple interesting biotechnological traits, such as production of secondary metabolites (bacillibactin, fengycin, bacitracin, and lantibiotics), genes involved in niche competition, and genes involved in plant growth promotion (Asif *et al*., 2023a). Similarly, genome-wide analysis in other *Bacillus paralicheniformis* strains, TRQ65 and KMS 80, evidenced similar traits (Annapurna *et al*., 2018; Valenzuela-Ruiz *et al*., 2019).

Molecular markers such as the antibiotic resistance genes *aadK* and *aph* are conserved and co-located within the chromosome, while the presence of the fengycin biosynthetic operon distinguishes *B. paralicheniformis* from closely related *Bacillus* species (Agersø *et al*., 2019; Asif *et al*., 2023b). Importantly, the genomes harbor multiple biosynthetic gene clusters responsible for the production of secondary metabolites, including bacitracin, fengycin, butirosin, bacillibactin, pulcherriminic acid, schizokinen, and geobacillin II, underscoring their significant antimicrobial potential (Albdaiwi *et al*., 2022; Chebotar *et al*., 2024). Additionally, genes involved in abiotic stress tolerance, nitrogen fixation, phosphate mobilization, and auxin production highlight their capacity to promote plant growth and resilience, particularly under saline or stressful environmental conditions (Iqbal *et al*., 2023).

In this work, we perform a comparative genomic analysis using AA1 alongside these strains exhibiting biotechnological traits (Figure 4). Our analysis reveals that while each strain harbors a unique set of genes, a high overlap exists (ANI > 99 %), with more than 4300 genes shared among them. Furthermore, we observe that strain AA1 shares a closer genomic content, particularly in terms of gene presence/absence, with the BP9 strain. The high overlap of genes between strain AA1 and other *Bacillus paralicheniformis* strains known for their biotechnological applications indicates that AA1 harbors promising genetic elements. Collectively, these genomic features position *Bacillus paralicheniformi*s as a promising candidate for biotechnological applications, especially as a biocontrol agent and microbial inoculant in sustainable agriculture. Nonetheless, a future analysis focusing on identifying and validating AA1 biotechnological traits must be undertaken.
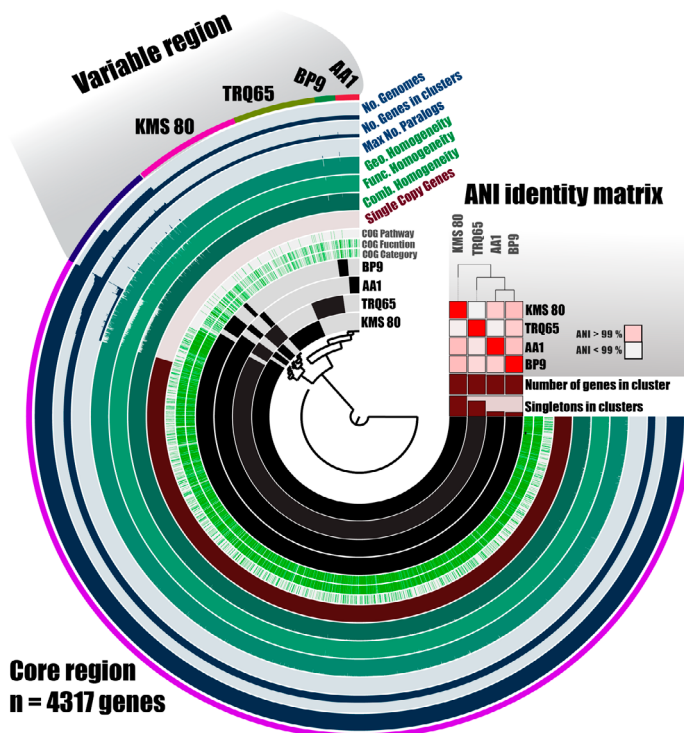

Figure 4. illustrates the comparative genomics of *Bacillus paralicheniformis* AA1 alongside other strains known for their biotechnologically relevant traits within the species. Each line in the graph represents a genome, with black-filled segments indicating the presence of a gene and lighter sections denoting gene absence. Additionally, unique genes present in the analyzed genomes are specifically highlighted.

Figura 4. Genómica comparativa de *Bacillus paralicheniformis* AA1 con otras cepas conocidas por sus rasgos relevantes en biotecnología dentro de la especie. Cada línea en el gráfico representa un genoma, con segmentos negros rellenos que indican la presencia de un gen y secciones más claras que denotan su ausencia. Además, se destacan específicamente los genes únicos presentes en los genomas analizados.

CONCLUSION

The genome analysis of *Bacillus paralicheniformis* AA1, isolated from a milpa farming system in Sonora, Mexico, reveals a wealth of genetic features that underscore its potential biotechnological applications. This study enhances our understanding of the microbial diversity associated with milpa ecosystems and highlights the capacity of such microorganisms to contribute to sustainable agricultural practices. The findings lay a foundation for exploring AA1's functional roles, particularly in antimicrobial compound synthesis, industrial enzyme production, and plant growth promotion. Future research focused on identifying and characterizing specific genes responsible for these traits will be pivotal in fully realizing the biotechnological promise of *Bacillus paralicheniformis* AA1, potentially offering innovative solutions for agriculture and biotechnology.

FUNDING INFORMATION

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

REFERENCES

Agersø, Y., Bjerre, K., Brockmann, E., Johansen, E., Nielsen, B., Siezen, R., Stuer-Lauridsen, B., Wels, M. and Zeidan, A.A. 2019. Putative antibiotic resistance genes present in extant *Bacillus licheniformis* and *Bacillus paralicheniformis* strains are probably intrinsic and part of the ancient resistome. *PLOS One*, *14*(1), e0210363. https://doi.org/10.1371/JOURNAL.PONE.0210363

Albdaiwi, R., Alhindi, T. and Hasan, S. 2022. Draft genome sequence of *Bacillus paralicheniformis* strain GSFE7, a halotolerant plant growth-promoting bacterial endophyte isolated from cultivated saline areas of the dead sea region. *Microbiology Resource Announcements*, *11*(9), e00425-22. https://doi.org/10.1128/MRA.00425-22

Annapurna, K., Govindasamy, V., Sharma, M., Ghosh, A. and Chikara, S. K. 2018. Whole genome shotgun sequence of *Bacillus paralicheniformis* strain KMS 80, a rhizobacterial endophyte isolated from rice (*Oryza sativa* L.). *3 Biotech*, *8*(5), 1-4. https://doi.org/10.1007/S13205-018-1242-y

Arahal, D.R. 2014. Whole-genome analyses: Average nucleotide identity. *Methods in Microbiology*, *41*, 103-122. https://doi.org/10.1016/BS.MIM.2014.07.002

Ashajyothi, M., Mahadevakumar, S., Venkatesh, Y.N., Sarma, P.V.S.R.N., Danteswari, C., Balamurugan, A., Prakash, G., Khandelwal, V., Tarasatyavathi, C., Podile, A.R., Kirankumar, M.S. and Chandranayaka, S. 2024. Comprehensive genomic analysis of *Bacillus subtilis* and *Bacillus paralicheniformis* associated with the pearl millet panicle reveals their antimicrobial potential against important plant pathogens.

BMC Plant Biology, *24*(1), 1-23. https://doi.org/10.1186/S12870-024-04881-4

Asif, M., Li-Qun, Z., Zeng, Q., Atiq, M., Ahmad, K., Tariq, A., Al-Ansari, N., Blom, J., Fenske, L., Alodaini, H.A. and Hatamleh, A.A. 2023a. Comprehensive genomic analysis of *Bacillus paralicheniformis* strain BP9, pan-genomic and genetic basis of biocontrol mechanism. *Computational and Structural Biotechnology Journal*, *21*, 4647-4662. https://doi.org/10.1016/J.CSBJ.2023.09.043

Asif, M., Li-Qun, Z., Zeng, Q., Atiq, M., Ahmad, K., Tariq, A., Al-Ansari, N., Blom, J., Fenske, L., Alodaini, H.A. and Hatamleh, A.A. 2023b. Comprehensive genomic analysis of *Bacillus paralicheniformis* strain BP9, pan-genomic and genetic basis of biocontrol mechanism. *Computational and Structural Biotechnology Journal*, *21*, 4647-4662. https://doi.org/10.1016/J.CSBJ.2023.09.043

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A. 2012. SPAdes: A new genome assembly Algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455. https://doi.org/10.1089/CMB.2012.0021

Bars-Cortina, D., Moratalla-Navarro, F., García-Serrano, A., Mach, N., Riobó-Mayo, L., Vea-Barbany, J., Rius-Sansalvador, B., Murcia, S., Obón-Santacana, M. and Moreno, V. 2023. Improving species level-taxonomic assignment from 16S rRNA sequencing technologies. *Current Protocols*, *3*(11), e930. https://doi.org/10.1002/CPZ1.930

Berais-Rubio, A., Morel-Revetria, M., Filippi, C.V., Reyno, R. and Monza, J. 2023. *Ensifer meliloti* elite strain U143 used as alfalfa inoculant in Uruguay: Characterization and draft genome sequence. *The Microbe*, *1*, 100008. https://doi.org/10.1016/J.MICROB.2023.100008

Berriel, V., Morel, M.A., Filippi, C.V. and Monza, J. 2021. Draft genome sequence of *Bradyrhizobium* sp. strain Oc8 isolated from *Crotalaria ochroleuca* nodule. *Current Research in Microbial Sciences*, *2*, 100074. https://doi.org/10.1016/J.CRMICR.2021.100074

Chavarria-Quicaño, E., Contreras-Jácquez, V., Carrillo-Fasio, A., De la Torre-González, F. and Asaff-Torres, A. 2023. Native *Bacillus paralicheniformis* isolate as a potential agent for phytopathogenic nematodes control. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/FMICB.2023.1213306

Chebotar, V.K., Gancheva, M.S., Keleinikova, O.V., Baganova, M.E., Zaplatkin, A.N., Pishchik, V.N. and Chizhevskaya, E.P. 2024. Whole-genome sequence of *Bacillus paralicheniformis* strain R243 isolated from rice (*Oryza sativa* L.) roots. *Microbiology Resource Announcements*, *13*(11), e00879-24. https://doi.org/10.1128/MRA.00879-24

Compant, S., Samad, A., Faist, H. and Sessitsch, A. 2019. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research*, *19*, 29-37. https://doi.org/10.1016/J.JARE.2019.03.004

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., Maestre, F.T., Singh, B.K. and Fierer, N. 2018. A global atlas of the dominant bacteria found in soil. *Science*, *359*(6373), 320-325. https://doi.org/10.1126/science.aap9516

Díaz-Manzano, F.E., Amora, D.X., Martínez-Gómez, Á., Moelbak, L. and Escobar, C. 2023. Biocontrol of *Meloidogyne* spp. in *Solanum lycopersicum* using a dual combination of *Bacillus* strains. *Frontiers in Plant Science*, 13, 1077062. https://doi.org/10.3389/FPLS.2022.1077062

Eren, A.M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M.S., Fink, I., Pan, J.N., Yousef, M., Fogarty, E.C., Trigodet, F., Watson, A.R., Esen, Ö.C., Moore, R.M., Clayssen, Q., Lee, M.D., Kivenson, V., Graham, E.D., Merrill, B.D., … Willis, A.D. 2020. Community-led, integrated, reproducible multi-omics with anvi'o. *Nature Microbiology 2020 6:1*, 6(1), 3-6. https://doi.org/10.1038/s41564-020-00834-3

Fonteyne, S., Castillo Caamal, J.B., Lopez-Ridaura, S., Van Loon, J., Espidio Balbuena, J., Osorio Alcalá, L., Martínez Hernández, F., Odjo, S. and Verhulst, N. 2023. Review of agronomic research on the milpa, the traditional polyculture system of Mesoamerica. *Frontiers in Agronomy*, 5, 1115490. https://doi.org/10.3389/FAGRO.2023.1115490

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.Y., Graham, M., Van Domselaar, G. and Stothard, P. 2023. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic acids research*, 51(W1), W484-W492. https://doi.org/10.1093/NAR/GKAD326

Hemangini, B.D., Kiran, C.S., Vrinda, T.S. and Bhatt Hemangini, C.D. 2018. Antimicrobial activity of *Bacillus paralicheniformis* SUBG0010 against plant pathogenic bacteria of mango. *Journal of Pharmacognosy and Phytochemistry*, 7(6), 449-455.

Iqbal, S., Qasim, M., Rahman, H., Khan, N., Paracha, R.Z., Bhatti, M.F., Javed, A., and Janjua, H.A. 2023. Genome mining, antimicrobial and plant growth-promoting potentials of halotolerant *Bacillus paralicheniformis* ES-1 isolated from salt mine. *Molecular Genetics and Genomics*, 298(1), 79-93. https://doi.org/10.1007/S00438-022-01964-5

Jinal, H.N., Sakthivel, K. and Amaresan, N. 2020. Characterisation of antagonistic *Bacillus paralicheniformis* (strain EAL) by LC–MS, antimicrobial peptide genes, and ISR determinants. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 113(8), 1167-1177. https://doi.org/10.1007/S10482-020-01423-4

Kim, M. and Chun, J. 2014. 16S rRNA gene-based identification of bacteria and archaea using the EzTaxon Server. *Methods in Microbiology*, 41, 61-74. https://doi.org/10.1016/BS.MIM.2014.08.001

Klimke, W., O'Donovan, C., White, O., Brister, J.R., Clark, K., Fedorov, B., Mizrachi, I., Pruitt, K.D. and Tatusova, T. 2011. Solving the problem: Genome annotation standards before the data deluge. *Standards in Genomic Sciences*, 5(1), 168. https://doi.org/10.4056/SIGS.2084864

Lavanya, J., Manjula, S., Chitra, A.G., Chaithanya, V. and Sivaramakrishnan, R. 2021. Efficacy of peptone glycerol broth in long-term storage of the bacterial and yeast cultures. *Journal of Clinical and Diagnostic Research*. https://doi.org/10.7860/JCDR/2021/47796.14524

Lee, M.D. 2019. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics*, 35(20), 4162-4164. https://doi.org/10.1093/BIOINFORMATICS/BTZ188

Olajide, A.M., Chen, S. and LaPointe, G. 2021. Markers to rapidly distinguish *Bacillus paralicheniformis* from the very close relative, *Bacillus licheniformis*. *Frontiers in Microbiology*, 11, 596828. https://doi.org/10.3389/FMICB.2020.596828

Parks, D.H., Chuvochina, M., Rinke, C., Mussig, A.J., Chaumeil, P.A. and Hugenholtz, P. 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research*, 50(D1), D785-D794. https://doi.org/10.1093/NAR/GKAB776

Pawaskar, M. 2023. *In-vivo biocontrol of Fusarium wilt in chilli plants using hypersaline Bacillus paralicheniformis strain MPSK23*. https://doi.org/10.21203/RS.3.RS-2927693/V1

Steinke, K., Mohite, O.S., Weber, T. and Kovács, Á.T. 2021. Phylogenetic distribution of secondary metabolites in the *Bacillus subtilis* species complex. *mSystems*, 6(2). https://doi.org/10.1128/MSYSTEMS.00057-21

Su, Y., Liu, C., Fang, H. and Zhang, D. 2020. *Bacillus subtilis*: a universal cell factory for industry, agriculture, biomaterials and medicine. *Microbial Cell Factories 2020 19:1*, 19(1), 1-12. https://doi.org/10.1186/S12934-020-01436-8

Swietczak, J., Kalwasińska, A., Felföldi, T. and Swiontek Brzezinska, M. 2023. *Bacillus paralicheniformis* 2R5 and its impact on canola growth and N-cycle genes in the rhizosphere. *FEMS Microbiology Ecology*, 99(9). https://doi.org/10.1093/FEMSEC/FIAD093

Tatusova, T., Dicuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M. and Ostell, J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research*, 44(14), 6614-6624. https://doi.org/10.1093/NAR/GKW569

Valenzuela-Ruiz, V., Robles-Montoya, R.I., Parra-Cota, F.I., Santoyo, G., del Carmen Orozco-Mosqueda, M., Rodríguez-Ramírez, R. and de los Santos-Villalobos, S. 2019. Draft genome sequence of *Bacillus paralicheniformis* TRQ65, a biological control agent and plant growth-promoting bacterium isolated from wheat (*Triticum turgidum* subsp. *durum*) rhizosphere in the Yaqui Valley, México. *3 Biotech*, 9(11), 1-7. https://doi.org/10.1007/S13205-019-1972-5

Wen, J., Smelt, J.P.P.M., Vischer, N.O.E., de Vos, A.L., Setlow, P. and Brul, S. 2022. Heat activation and inactivation of bacterial spores: Is there an overlap? *Applied and Environmental Microbiology*, 88(5). https://doi.org/https://doi.org/10.1128/aem.02324-21