

Comprehensive genomic analysis of *Burkholderia arboris* PN-1 reveals its biocontrol potential against *Fusarium solani* induced root rot in *Panax notoginseng*

Yun Yang

yang1605@user.ynnu.edu.cn

Yunnan Provincial Observation and Research Station of Soil Degradation and Restoration for Cultivating Plateau Traditional Chinese Medicinal Plants, Yunnan Normal University

Haoji Wang

Yunnan Provincial Observation and Research Station of Soil Degradation and Restoration for Cultivating Plateau Traditional Chinese Medicinal Plants, Yunnan Normal University

Jielei Tu

Yunnan Provincial Observation and Research Station of Soil Degradation and Restoration for Cultivating Plateau Traditional Chinese Medicinal Plants, Yunnan Normal University

Yan Li

Yunnan Provincial Rural Energy Management Station

Huilin Guan

Yunnan Provincial Observation and Research Station of Soil Degradation and Restoration for Cultivating Plateau Traditional Chinese Medicinal Plants, Yunnan Normal University

Research Article

Keywords: *Panax notoginseng*, *Burkholderia arboris*, biocontrol agent, genome, secondary metabolites

Posted Date: December 22nd, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3777315/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Current Genetics on March 30th, 2024. See the published version at <https://doi.org/10.1007/s00294-024-01288-4>.

Abstract

Panax notoginseng (Burkill) F.H. Chen, a valuable traditional Chinese medicine, faces significant yield and quality challenges stemming from root rot primarily caused by *Fusarium solani*. *Burkholderia arboris* PN-1, isolated from the rhizosphere soil of *P. notoginseng*, demonstrated a remarkable ability to inhibit the growth of *F. solani*. This study integrates phenotypic, phylogenetic, and genomic analyses to enhance our understanding of the biocontrol mechanisms employed by *B. arboris* PN-1. Phenotype analysis reveals that *B. arboris* PN-1 effectively suppresses *P. notoginseng* root rot both *in vitro* and *in vivo*. The genome of *B. arboris* PN-1 comprises three circular chromosomes (contig 1: 3,651,544 bp, contig 2: 1,355,460 bp, and contig 3: 3,471,056 bp), with a 66.81% GC content, housing 7,550 protein-coding genes. Notably, no plasmids were detected. Phylogenetic analysis places PN-1 in close relation to *B. arboris* AU14372, *B. arboris* LMG24066, and *B. arboris* MEC_B345. Average nucleotide identity (ANI) values confirm the PN-1 classification as *B. arboris*. Comparative analysis with seven other *B. arboris* strains identified 4,628 core genes in *B. arboris* PN-1. The pan-genome of *B. arboris* appears open but may approach closure. Whole genome sequencing revealed 265 carbohydrate-active enzymes and identified 9 gene clusters encoding secondary metabolites. This comprehensive investigation enhances our understanding of *B. arboris* genomes, paving the way for their potential as effective biocontrol agents against fungal plant pathogens in the future.

Introduction

Panax notoginseng (Burk.) F.H.Chen, known as “Sanqi”, holds significance in traditional Chinese medicine. However, the threat of root rot disease poses a considerable risk to the sustainability of Sanqi cultivation (Ma et al. 2019). *Fusarium solani* is the primary cause of root rot, substantially reducing Sanqi yield and quality (Deng et al. 2022). Chemical fungicides, though traditionally used, pose challenges including prolonged field application leading to environmental pollution and potential fungal resistance. Considering the environmental impact of chemical solutions, there is a growing need for biological control agents (Xie et al. 2023). This approach, marked by its environmental friendliness, safety, and lack of resistance, is crucial for effective root disease management in Sanqi. Numerous published studies highlight the efficacy of microorganisms such as *Trichoderma* (Li et al. 2022) and *Bacillus* (Wang et al. 2023) as biocontrol agents against Sanqi root rot.

Burkholderia genomes harbor a diverse array of natural products holding promising therapeutic and biotechnological potential (Kunakom and Eustáquio 2019). In recent years, numerous *Burkholderia* species have been found to secrete secondary metabolites, including antibacterial, anticancer, herbicidal, and insecticidal compounds, with applications spanning bioremediation, biocontrol, and plant growth promotion (Alam et al. 2022). Certain *Burkholderia* species, distinguished by unique biosynthetic gene clusters, act as promising biocontrol agents, producing antimicrobial metabolites (Petrova et al. 2022). Among these, *Burkholderia arboris*, commonly found in the *Burkholderia cepacia* complex and frequently isolated from the rhizosphere, exhibits notable anti-fungal activity against various fungal plant pathogens (Zhu et al. 2022). Extensive research underscores *B. arboris* as an effective biological control

agent (Sandani et al. 2019; Singh et al. 2020; Zhang et al. 2022). In this study, we introduce *B. arboris* PN-1, a potent biocontrol strain isolated from the rhizosphere of Sanqi. However, its potential in preventing Sanqi root rot disease caused by *F. solani* remains unknown, and its specific biocontrol mechanisms are yet to be elucidated. To unravel the molecular mechanisms of plant-microbe interactions, we present a high-quality genome assembly and annotation of *B. arboris* PN-1.

The aim is to discern the antagonistic activity of *B. arboris* PN-1 against *F. solani* in vitro and in vivo, along with an assessment of its biocontrol efficiency within a greenhouse setting. Additionally, PN-1 underwent a comparative analysis with seven other *B. arboris* strains, each with published complete genome sequences. This comparative evaluation utilized phylogenetic tree, average nucleotide identity (ANI), and nucleic acid co-linearity analysis. Subsequent to the genomic comparisons, leveraging the complete genome sequence of *B. arboris* PN-1, specific functional analyses were conducted. These encompassed an evaluation of Carbohydrate-Active Enzymes (CAZymes) and an exploration of secondary metabolites. The analysis aimed at predicting PN-1's capabilities, including biofilm formation, production of antifungal CAZymes, and synthesis of antagonistic secondary metabolites against pathogens. This research provides novel insights into the biocontrol mechanisms of *B. arboris*. Moreover, it offers a valuable resource for new antibiotic discovery, contributing to innovative approaches to disease control.

Materials and methods

Culture of Strain PN-1 and Assessment of Antifungal Efficacy

B. arboris PN-1 (CGMCC NO.9733) was cultivated in solid CYMG (casitone yeast extract magnesium chloride glycerol) medium at 28 °C for 2 days. To evaluate its antagonistic activity against the root rot pathogens of *F. solani*, a dual-culture assay was conducted. Initially, *F. solani* pathogens were cultivated on PDA plates at 28 °C for 7 days. Using a 6 mm diameter punch, a bacterial block was extracted along the mycelium's edge and transferred to the center of a fresh PDA plate. Subsequently, *B. arboris* PN-1 was inoculated 2.5 cm from the central mycelium block in four directions: east, south, west, and north. Following 7 days of cultivation at 28 °C, the growth rate of *F. solani* was measured relative to the control (without strain PN-1 on a plate) (Wang et al. 2023). The inhibition rate was calculated using the formula: (control colony diameter – treatment colony diameter) / control colony diameter × 100%.

Biocontrol potential of strain PN-1

To assess the plant growth promotion and biocontrol effects of *B. arboris* PN-1, this study conducted greenhouse experiments. Two-year-old Sanqi seedlings were utilized for pot experiments, with the greenhouse setup comprising four treatments, each with five replicates: sterile distilled water (CK), *F. solani* inoculation (Fs), *B. arboris* (PN-1), pre-inoculation with *B. arboris* (PN-1) three days before subsequent *F. solani* inoculation (PN-1 + Fs). All inoculations employed the root irrigation method, with a

B. arboris suspension concentration of 1×10^{10} CFU·ml⁻¹ and a spore suspension of *F. solani* at 1×10^7 CFU·ml⁻¹, using 150 mL for each treatment. Treatment involved a plastic bowl (15 centimeters in diameter and 12.5 centimeters in depth) containing five Sanqi plants, considered one experimental unit. At 45 days post inoculation (dpi) with PN-1 and Fs, plant height, fresh weight, disease severity, and disease index were recorded. Disease severity was assessed using a 0–4 disease scale: 0 = Healthy growing plant; 1 = Disease spots are visible at the ends of leaves; 2 = Less than 50% of the leaves on the plant are yellow; 3 = Over 50% of the leaves on the plant are yellow; 4: Plant has died. Disease index (Di) was calculated using the formula (Yang et al. 2023): $Di = \sum [(Dn \times Dg) / (Tn \times Mg)] \times 100$. where Dn is the number of plants in the batch sharing the same disease grade, Dg is the corresponding disease grade, Tn is the total plant number, and Mg is the highest disease grade for a plant in the batch. Relative control effect (REC) was calculated as the formula (Dong et al. 2023): $RCE (\%) = (Di \text{ of water treatment} - Di \text{ of PN-1 + Fs treatment}) / Di \text{ of water treatment} \times 100$.

Genome Sequencing, Assembly, and Annotation

The genomic DNA of *B. arboris* PN-1 was extracted using the TIANamp Bacteria DNA kit (TIANGEN Biotechnology, China) according to the manufacturer's instructions. Subsequently whole-genome sequencing was performed using the Pacific Biosciences platform (PacBio, Menlo Park, CA) and the Illumina Miseq platform at BioMarker Technologies Co., Ltd (Beijing, China). Gene prediction was performed utilizing Prodigal software v2.6.3 (Hyatt et al. 2010). Repetitive sequences within the genome were identified using Repeat Masker v4.0.5 software (Tarailo-Graovac and Chen 2009). The prediction of transfer RNA (tRNA) and ribosomal RNA (rRNA) sequences was performed using tRNA scan-SE v2.0 (Chan and Lowe 2019) and Infernal v1.1.3 (Nawrocki and Eddy 2013). CRISPRs in the genome were predicted using the CRTv1.2 tool (Bland et al. 2007), while gene islands in the bacterial genomes were identified using Island Path-DIMOB v0.2 (Bertelli et al. 2018). To predict coding DNA sequences (CDSs), Prodigal v2.6.3 was used (Hyatt et al. 2010). A comprehensive blast search for CDSs was then conducted against widely used databases, including NR, Swiss-Prot, COGs, GO, and KEGGs (Wei et al. 2023).

Phylogenetic analysis and genome comparisons

Phylogenetic analysis was conducted to elucidate the evolutionary relationships between *B. arboris* PN-1 and seven other *B. arboris* strains. This was achieved by analyzing homologous genes, wherein single-copy homologous genes from each strain underwent multiple sequence alignment and quality control comparisons. Subsequently, the phylogenetic tree was constructed using RAXML software, employing the maximum likelihood method based on both single gene and multiple genes (Luo et al. 2018). To further determine the taxonomic status of PN-1, average nucleotide identity (ANI) values were calculated between PN-1 and the seven other *B. arboris* strains in the database, utilizing the EzBioCloud server (<http://www.ezbiocloud.net/tools/ani>) (Tang et al. 2021). Strains with ANI values $\geq 95\%$ served as the classification threshold to distinguish species (Zeng et al. 2021a). The ANI values were then utilized to construct a heatmap using TBtools v2.012 software (Dong et al. 2023). Based on the results from phylogenetic and ANI analyses, three representative strains with released complete genomes *B. arboris*

AU14372, *B. arboris* LMG24066, and *B. arboris* MEC_B345 were selected for comparative genomic analysis. The genome collinearity among these strains was assessed using Mauve software version 2.3.1 (Xu et al. 2022).

Core pan-genome analysis

A comprehensive comparison of protein sequences was undertaken for *B. arboris* PN-1 and seven additional strains, employing the Bacterial Pan Genome Analysis (BPGA) software (Chaudhari et al. 2016; Guo et al. 2022). Orthologous clusters of protein sequences were defined using USEARCH, with a default 50% sequence identity cut-off (Svetlicic et al. 2022). To visually represent the relationships among the proteins in the 8 strains of *B. arboris*, a flower plot was generated using R 4.3.2 (Dong et al. 2023).

Analysis of CAZymes and secondary metabolic genes of *B. arboris* PN-1

Advanced annotation of Carbohydrate-Active enZymes (CAZy) was conducted using the CAZy database (<http://www.cazy.org/>) (Li et al. 2021). For the prediction of secondary metabolite clustering, the antiSMASH website (<https://antismash.secondarymetabolites.org/>) was employed (Xu et al. 2020).

Statistical Analysis

Statistical analysis was performed using SPSS Statistics v26.0 (IBM Corp.). Group comparisons were made using one-way analysis of variance (ANOVA). Mean values were compared using Duncan's multiple range test, with a significance level set at $P \leq 0.05$.

Results

Biocontrol activity in vitro and in vivo

As illustrated in Fig. 1, *B. arboris* PN-1 demonstrated a remarkable inhibition of mycelial growth in culture media. The control pathogen colony exhibited a diameter of 8.06 cm, while pathogen colony inoculated with the PN-1 plate had a diameter of 2.39 cm, resulting in an inhibition rate of 70.32% (Fig. 1). These findings strongly suggest that PN-1 holds the potential to suppress root rot in Sanqi.

To validate this hypothesis, greenhouse experiments were conducted, revealing that PN-1 significantly controlled root rot in Sanqi. After inoculating Fs, PN-1 + Fs treatment exhibited a significant enhancement in plant biomass. Plant height, fresh weight, and dry weight improved by 41.42%, 52.99%, and 52.43%, respectively, compared to the Fs treatment alone. Moreover, the plant height, fresh weight, and dry weight of the PN-1 treatment surpass the other three conditions. Visible differences were observed in Sanqi plants treated with Fs, exhibiting weaker growth, yellowing leaves, and significantly fewer and shorter fibrous roots compared to the control, PN-1, and PN-1 + Fs treatments (Fig. 2A, B). The disease index highlighted a substantial reduction in Di for the PN-1 + Fs treatment, decreasing by 58.90% compared to Fs (86.25%), with a corresponding RCE of 68.29% (Table 1).

Table 1
Evaluation of *B. arboris* PN-1 greenhouse biocontrol efficacy against *F. solani*

Treatment	Plant Height (cm)	Root Fresh Weight (g)	Root dry weight (g)	Disease Index (Di %)	Relative control Effect (RCE %)
CK	23.99 ± 0.11b	7.27 ± 0.28c	1.47 ± 0.03b	-	-
PN-1	27.86 ± 0.20a	9.81 ± 0.21a	1.89 ± 0.04a	-	-
Fs	17.77 ± 0.28c	5.36 ± 0.35d	1.03 ± 0.07c	86.25 ± 2.88a	-
PN-1 + Fs	25.13 ± 0.28b	8.20 ± 0.15b	1.57 ± 0.03b	27.35 ± 1.99b	68.29

Overview of the Genome of *B. arboris* PN-1

The complete genome sequence of *B. arboris* PN-1 was determined with the aim of exploring its biocontrol mechanisms and applications in sustainable agriculture. The PN-1 genome is comprised of three circular chromosomes, measuring 3,651,544 bp, 1,355,460 bp, and 3,471,056 bp, with an average GC content of 66.81% (Fig. 3). No plasmids were identified. Notably, the PN-1 genome was predicted to contain 7,550 protein-coding genes, along with 18 rRNAs, and 70 tRNAs. Furthermore, four CRISPRs and genomic islands were identified within the genome. Annotation of genes against various databases revealed the following counts: NR (7,479), Swiss-Prot (4,348), COG (6,290), GO (5,709), and KEGG (2,833).

Utilizing the COG database, 6,290 genes from *B. arboris* PN-1 were categorized into 23 COG functional groups (Fig. 4A). The group with unknown function represented the largest category (1,124 genes, comprising 17.87% of all CDSs), followed by transcription. Conversely, only a limited number of genes were assigned to the RNA processing and modification category. In terms of GO annotation, a total of 5,709 genes were classified into 40 functional groups, with a predominant representation in biological process-related groups (Fig. 4B). Within the biological process category, the highest number of genes was associated with metabolic processes, constituting 57.78% of the total. According to KEGG annotation, 2,833 genes (37.52% of all CDS) were assigned to 49 KEGG pathways, with the majority allocated to metabolism pathways. Among these pathways, biosynthesis of amino acids (155 genes, 2.05% of all CDS) emerged as the most represented, followed by carbon metabolism and metabolism of purine pathways (Fig. 4C).

Whole-Genome Phylogenetic Analysis of *B. arboris*

To elucidate the genetic relationship of *B. arboris* PN-1 with other strains, a phylogenetic tree was constructed based on single copy homologous genes, utilizing seven publicly available complete genome sequences. The genetic distance revealed that PN-1 was most closely related to *B. arboris* AU14372, *B. arboris* LMG24066, and *B. arboris* MEC_B345 (Fig. 5A). The accurate systematic position of PN-1 as *B. arboris* was confirmed through phylogenetic analysis. To further validate these findings, ANI values were

calculated for different strains. PN-1 exhibited the highest ANI value with *B. arboris* AU14372 at 99.51%, followed by *B. arboris* LMG24066 and *B. arboris* MEC_B345 with values of 99.50% and 99.45%, respectively. ANI values between PN-1 and the other four *B. arboris* strains ranged from 92.30 to 96.23% (Fig. 5B). Based on these analyses, strain PN-1 was conclusively identified as *B. arboris*.

To assess the evolutionary distance among *B. arboris* strains, a comparison of their whole genome sequences was conducted using Mauve (Fig. 6). Progressive Mauve Align revealed a highly conserved synteny between *B. arboris* PN-1 and strains *B. arboris* AU14372, *B. arboris* LMG24066, and *B. arboris* MEC_B345. Notably, 33, 74, and 5 LCBs with minimum weights of 183, 191, and 581,630 were identified between *B. arboris* PN-1 and these three strains, respectively. Comparing the genome of *B. arboris* PN-1 with that of *B. Arboris* AU14372, partial heterotopy and inversion were observed. However, *B. arboris* LMG24066 exhibited large-scale sequence rearrangements or inversions compared to the *B. arboris* PN-1 genome. These findings align with the results from genome distance calculations and phylogenetic tree analysis, collectively supporting the genetic relationships among these *B. arboris* strains.

The Pan-Genome Features of *B. arboris*

To elucidate the core and variable gene pool among *B. Arboris* PN-1 and seven other sequenced *B. arboris* strains, a pan-genome analysis was carried out using the Bacterial Pan Genome Analysis (BPGA) software. The combined pan-genome size was determined to be 9,256 genes. The analysis revealed distinct categories, including the core genome, which comprises 4,628 genes (50%) shared across all strains, the accessory genome consisting of 2,144 genes present in some, but not all strains, and the unique genome encompassing 284 genes specific to *B. Arboris* PN-1 (Fig. 7A). Additionally, the count of unique genes in each strain was as follows: 284 (*B. arboris* PN-1), 206 (*B. arboris* AU1125 206), 169 (*B. arboris* AU14372), 501 (*B. arboris* AU37640), 1190 (*B. arboris* AU38796), 193 (*B. arboris* BCC0049), 156 (*B. arboris* LMG24066), and 272 (*B. arboris* MEC_B345 272). These findings indicate a high degree of homology among the eight strains, while also revealing evolutionary differences manifested in unique genes.

The core-genome and pan-genome sizes among *B. arboris* PN-1 and seven other *B. arboris* strains were calculated by extrapolation of the selected genome data. The pan-genome curves of *B. arboris* were effectively modeled by Heaps law mathematical functions: $y = 6857.16 x^{0.21}$, where y represents the pan-genome size and x refers to the number of sequenced genomes. The dilution curve of core-pan proteins showed that the pan-genome of *B. arboris* is still open but may close in the near future. According to these equations, the pan-genome size of *B. arboris* appears to reach infinity when the number of genomes increases to infinity (Fig. 7B). The concept of an infinite pan-genome for *B. arboris* implies that the bacteria will likely continue acquiring new genes as they evolve independently over evolutionary time.

Genome CAZymes analysis of *B. arboris* PN-1

CAZyme prediction for *B. arboris* PN-1 revealed a repertoire of 265 CAZyme genes, distributed across various classes: 61 Glycoside Hydrolases (GHs), 86 Glycosyl Transferases (GTs), 58 Carbohydrate

Esterases (CEs), 35 Auxiliary Activities (AAs), 5 Polysaccharide Lyases (PLs), and 20 Carbohydrate-Binding Modules (CBMs). These enzymes play crucial roles, such as catalyzing glycosidic bond cleavage, facilitating glycosyl group transfer, and aiding in carbohydrate substrate binding. (Fig. 8). A comprehensive classification assigned the 265 genes to 72 CAZyme families. Notably, the CE1 family emerged as the most abundant, comprising 29 members. The CBM50 family, with 10 members, contributes to the antifungal activity by binding to the chitinous component of the fungal cell wall. Within the GH family, specific enzymes such as β -glucosidases (GH3), cellulases (GH5), Glucosidases (GH4 and GH13_29), chitinases (GH23 and GH18), and glucanases (GH5_2 and GH16) exhibited potential antifungal activities. These findings underscore the significant role of CAZymes in *B. arboris* PN-1 in exerting antifungal activity.

Analysis of genes encoding antifungal-related compounds

Utilizing the antiSMASH database, we identified 9 clusters related to secondary metabolite synthesis in PN-1. Notably, the ornibactin, pyochelin, and pyrrolnitrin clusters exhibited a 100% similarity to reported secondary metabolite synthesis gene clusters (Table 2). The *B. Arboris* PN-1 genome encompassed diverse gene clusters: four clusters encoding NRPS, one cluster for T1PKS, one cluster for terpene biosynthesis, one cluster for arylpolyene biosynthesis, one cluster for phosphonate biosynthesis, and one cluster encoding other secondary metabolites. This genomic repertoire positions *B. arboris* PN-1 as a potential source for antibacterial agent biosynthesis, regulation, transport, and plant growth promotion. The abundance of genes with antibacterial and growth-promoting potential further enhances the biocontrol capabilities of this strain.

Table 2
Genes and gene cluster encoding for secondary metabolites in *B. arboris* PN-1

Type	Most similar known cluster	Gene cluster ID similarity (%)	Gene location Start-end	Size (bp)
T1PKS	lipopolysaccharide	BGC0000774 (5)	782,306–829,651	47,346
NRP-metallophore; NRPS	ornibactin C8	BGC0002569 (100%)	1,788,287–1,853,284	64,998
arylpolyene	APE Vf	BGC0000837 (10%)	3,171,213–3,212,424	41,212
phosphonate	dehydrofosmidomycin	BGC0002036 (15%)	1,706,517–1,739,238	32,722
terpene	N-acyloxyacyl glutamine	BGC0001684 (50%)	1,889,328–1,910,392	21,065
NRP-metallophore; NRPS	pyochelin	BGC0002678 (100%)	3,041,314–3,095,166	53,853
NRPS; betalactone T1PKS,	occidiofungin A	BGC0001711 (88%)	132,581 – 267,689	135,109
NRPS, T1PKS	Pf-5 pyoverdine	BGC0000413 (1%)	354,030–411,296	57,267
other	pyrrolnitrin	BGC0000924 (100%)	683,059–724,144	41,086

Discussion

Previous studies have identified root rot as a key factor limiting the quality and yield of Sanqi, prompting focused research on the safe and effective prevention and control of Sanqi root rot. Within the *B. cepacia* complex, *B. arboris* has garnered attention for its ability to synthesize novel antifungal compounds, such as Hafk, exhibiting potent antifungal activity against diverse plant fungal pathogens (Zhu et al. 2022). This study was designed to expand our understanding of the molecular biocontrol mechanisms employed by *B. arboris* PN-1. Through a series of in vivo and in vitro biocontrol experiments, coupled with growth promotion assays, we validated its remarkable growth-promoting and biocontrol effects. To unravel the genetic blueprint underlying these traits, the complete genome of *B. arboris* PN-1 was sequenced. The genome, comprised of three circular chromosomes (3,651,544 bp, 1,355,460 bp, and 3,471,056 bp), boasts an average GC content of 66.81%. Comparative genomic analyses were then conducted, juxtaposing the PN-1 genome with seven other *B. arboris* strains.

The surge in the number of sequenced microbial genomes offers a unique opportunity to reevaluate approaches for deciphering bacterial phylogeny (Zeng et al. 2021b). While sequencing the 16S rRNA gene

has been a powerful tool for studying microbial communities and their responses to changing environmental conditions, its limitations in resolving relationships between closely related groups are evident (Wemheuer et al. 2020). In this study, a robust phylogenetic analysis was conducted using 4,165 single-copy core genes, categorizing them into three distinct clusters (Fig. 5A). To delve deeper into the gene relatedness of PN-1, an analysis based on ANI was performed (Fig. 5B). Notably, with the exception of *B. arboris* AU3876, which falls below 96%, all other strains exhibited ANI values greater than 96%. ANI values exceeding the taxonomic threshold of 95% are indicative of the same species (Mališová et al. 2021). Therefore, the designation of strain PN-1 as *B. arboris* aligns with both the ANI results and the phylogenetic tree. Whole-genome pairwise comparisons play an important role in diversity analysis across different species lineages, particularly in the evolutionary analysis of species and subspecies (Xu et al. 2022). High genomic similarity was observed among *B. arboris* PN-1, *B. arboris* AU14372, *B. arboris* LMG24066, and *B. arboris* MEC_ B345, consistent with their taxonomic classification. This alignment suggests that *B. arboris* PN-1 may harbor comparable biocontrol potential to other strains within the species.

Chitin, a predominant component of the cell wall in many pathogenic fungi, is the target for enzymatic degradation in biological control processes. Chitinase, a catalytic enzyme, hydrolyzes chitin to produce N-acetylglucosamine, disrupting the structural integrity of fungal cell walls and inhibiting spore germination and hyphal growth (Xie et al. 2023). In addition to chitinase, enzymes such as pectinase, cellulase, and protease also play roles in the biological control process, synergizing with chitinase to augment antibacterial effects. The CAZymes analysis conducted on *B. arboris* PN-1 revealed the presence of enzymes involved in cell wall degradation, including chitinase and β -1,3-glucanase. Previous research has found that these enzymes can modify the structure and function of fungal cell walls, ultimately inhibiting the growth of plant pathogenic fungal hyphae, spore germination, and plant infection. Consequently, it is speculated that the degradation of the cell wall in pathogenic fungi by *B. arboris* PN-1 constitutes one of its antibacterial mechanisms.

The genomic sequencing of *B. arboris* PN-1 has revealed the presence of gene clusters responsible for synthesizing various secondary metabolites, including ornibactin (Deng et al. 2017), pychelelin (Ong et al. 2017), and pyrrolnitrin (Ankenbauer et al. 1988). Ornibactin and pychelelin, identified as key siderophores (Wang et al. 2021), function as bacteriostatic agents, inhibiting the growth of pathogenic microorganisms by depleting iron in the soil (Deng et al. 2017). Additionally, the antibiotic pyrrolnitrin, derived from tryptophan, exhibits broad-spectrum antimicrobial activity, making it a valuable biocontrol agent against plant diseases (Liu et al. 2018). Previous studies have recognized pyrrolnitrin-producing bacteria as promising biopesticides (Keum et al. 2010), and it has been identified as a major factor in the biocontrol of Rhizoctonia stem rot of poinsettia by *B. cepacia* (Hwang et al. 2002). The genomic analysis of *B. ambifaria* T16 has also revealed several operons involved in the biosynthesis of antimicrobial compounds, such as pyrrolnitrin, ornibactin, occidiofungin, and the membrane-associated AFC-BC11 (Alvarez et al. 2022). In summary, the antibacterial effects of *B. arboris* PN-1 are primarily attributed to the encoding and synthesis of antimicrobial compounds.

In conclusion, the genomic analysis of *B. Arboris* PN-1 revealed the presence of 265 CAZymes and 9 gene clusters responsible for the biosynthesis of secondary metabolites, encoding various anti-microbial substances with the potential to inhibit root rot disease in Sanqi. This genome serves as a valuable resource for *B. arboris* biocontrol strains, laying the foundation for future exploration and discovery of novel anti-fungal substances.

Declarations

Data availability

The chromosome sequence of *B. arboris* PN-1 were deposited in NCBI under accession numbers CP109821, CP109822 and CP109823. The BioProject and BioSample designations for this project were PRJNA891013 and SAMN31305798, respectively.

Acknowledgements

We are extremely gratefully to Wenshan Miao Xiang Notoginseng Technology, Co., Ltd., for providing facilities that enable us to conduct this study. The authors would like to express their gratitude to EditSprings (<https://www.editsprings.cn>) for the expert linguistic services provided.

Funding Information

This work was supported by the Major Science and Technology Special Project of Yunnan Province (grant number: 202202AE090025); Yunnan Academician Expert Workstation (grant number: 202305AF150027); Yunnan (Kunming) Academician Expert Workstation (grant number: YSZJGZZ-2021062); Postgraduate research Innovation Fund Project in Yunnan Normal University (grant number: YJSJJ22-A15).

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Alam K, Zhao Y, Lu X, Gong K, Zhong L, Hao J, Islam MM, Islam S, Li G, Zhang Y (2022) Isolation, complete genome sequencing and in silico genome mining of *Burkholderia* for secondary metabolites. BMC Microbiol 22: 323
2. Alvarez F, Simonetti E, Draghi WO, Vinacour M, Palumbo MC, Do Porto DF, Montecchia MS, Roberts IN, Ruiz JA (2022) Genome mining of *Burkholderia ambifaria* strain T16, a rhizobacterium able to produce antimicrobial compounds and degrade the mycotoxin fusaric acid. World J Microb Biot 38:114

3. Ankenbauer RG, Toyokuni T, Staley A, Rinehart KL, Cox CD (1988) Synthesis and biological activity of pyochelin, a siderophore of *Pseudomonas aeruginosa*. J Bacteriol 170:5344-5351
4. Bertelli C, Brinkman FSL, Valencia A (2018) Improved genomic island predictions with IslandPath-DIMOB. Bioinformatics 34:2161-2167
5. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, Hugenholtz P (2007) CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics 8:209
6. Chan PP, Lowe TM (2019) tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Gene Prediction 1962:1-14
7. Chaudhari NM, Gupta VK, Dutta C (2016) BPGA- an ultra-fast pan-genome analysis pipeline. Sci Rep 6:24373
8. Deng J, Guan R, Liang T, Su L, Ge F, Cui X, Liu D (2022) Dirigent gene family is involved in the molecular interaction between *Panax notoginseng* and root rot pathogen *Fusarium solani*. Ind Crop Prod 178:114544
9. Deng P, Foxfire A, Xu J, Baird SM, Jia J, Delgado KH, Shin R, Smith L, Lu S-E, Drake HL (2017) The Siderophore Product Ornibactin Is Required for the Bactericidal Activity of *Burkholderia contaminans* MS14. Appl Environ Microb 83: e00051-17
10. Dong H, Gao R, Dong Y, Yao Q, Zhu H (2023) Whole-genome sequencing of a biocontrol *Myxococcus xanthus* R31 isolate and comparative genomic analysis. Gene 863:147286.
11. Dong Q, Liu Q, Goodwin PH, Deng X, Xu W, Xia M, Zhang J, Sun R, Wu C, Wang Q (2023) Isolation and Genome-Based Characterization of Biocontrol Potential of *Bacillus siamensis* YB-1631 against Wheat Crown Rot Caused by *Fusarium pseudograminearum*. J Fungi (Basel) 9:547
12. Guo Y, Zeng C, Ma C, Cai H, Jiang X, Zhai S, Xu X, Lin M (2022) Comparative genomics analysis of the multidrug-resistant *Aeromonas hydrophila* MX16A providing insights into antibiotic resistance genes. Front Cell Infect Microbiol 12:1042350
13. Hwang J, Chilton WS, Benson DM (2002) Pyrrolnitrin production by *Burkholderia cepacia* and biocontrol of Rhizoctonia stem rot of poinsettia. Biol Control 25:56-63
14. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119
15. Keum YS, Lee HR, Kim J-H (2010) Effects of Pesticides on the Bacterial Production of Pyrrolnitrin. J. Agric Food Chem 58:5531-5537
16. Kunakom S, Eustáquio AS (2019) *Burkholderia* as a Source of Natural Products. J Nat Prod 82:2018-2037
17. Li C, Cheng P, Zheng L, Li Y, Chen Y, Wen S, Yu G (2021) Comparative genomics analysis of two banana *Fusarium* wilt biocontrol endophytes *Bacillus subtilis* R31 and TR21 provides insights into their differences on phytobeneficial trait. Genomics 113:900-909

18. Li Y, Liu Y, Zhang Z, Li J, Zhu S, Yang M, Luo L (2022) Application of plant survival-promoting and pathogen-suppressing *Trichoderma* species for crop biofertilization and biocontrol of root rot in *Panax notoginseng*. J Plant Pathol 104:1361-1369
19. Liu X, Yu X, Yang Y, Heeb S, Gao S, Chan KG, Camara M, Gao K (2018) Functional identification of the prnABCD operon and its regulation in *Serratia plymuthica*. Appl Microbiol Biotechnol 102:3711-3721
20. Luo Y, Cheng Y, Yi J, Zhang Z, Luo Q, Zhang D, Li Y (2018) Complete Genome Sequence of Industrial Biocontrol Strain *Paenibacillus polymyxa* HY96-2 and Further Analysis of Its Biocontrol Mechanism. Front Microbiol 9:1520
21. Ma YN, Chen CJ, Li QQ, Xu FR, Cheng YX, Dong X (2019) Monitoring Antifungal Agents of *Artemisia annua* against *Fusarium oxysporum* and *Fusarium solani*, Associated with *Panax notoginseng* Root-Rot Disease. Molecules 24:213
22. Mališová L, Španělová P, Sedláček I, Pajer P, Musílek M, Puchálková B, Jakubů V, Žemličková H, Šafránková R (2021) The first case of *Planococcus glaciei* found in blood, a report from the Czech Republic. Folia Microbiologica 67:121-127
23. Nawrocki EP, Eddy SR (2013) Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics 29:2933-2935
24. Ong KS, Cheow YL, Lee SM (2017) The role of reactive oxygen species in the antimicrobial activity of pyochelin. J Adv Res 8:393-398
25. Petrova YD, Zhao J, Webster G, Mullins AJ, Williams K, Alswat AS, Challis GL, Bailey AM, Mahenthiralingam E (2022) Cloning and expression of *Burkholderia* polyene biosynthetic gene clusters in *Paraburkholderia* hosts provides a strategy for biopesticide development. Microb Biotechnol 15:2547-2561
26. Sandani HBP, Ranathunge NP, Lakshman PLN, Weerakoon WMW (2019) Biocontrol potential of five *Burkholderia* and *Pseudomonas* strains against *Colletotrichum truncatum* infecting chilli pepper. Biocontrol Sci Techn 29:727-745
27. Singh TB, Sahai V, Goyal D, Prasad M, Yadav A, Shrivastav P, Ali A, Dantu PK (2020) Identification, Characterization and Evaluation of Multifaceted Traits of Plant Growth Promoting Rhizobacteria from Soil for Sustainable Approach to Agriculture. Curr Microbiol 77:3633-3642
28. Svetlicic E, Jaen-Luchoro D, Klobucar RS, Jers C, Kazazic S, Franjevic D, Klobucar G, Shelton BG, Mijakovic I (2022) Genomic characterization and assessment of pathogenic potential of *Legionella* spp. isolates from environmental monitoring. Front Microbiol 13:1091964
29. Tang C, Fan C, Guo DB, Ma XJ, Cai QT, Chen XX, Zhang M, Li JY, An QY, Zhao R (2021) Identification of *Boseongicola sediminum* sp. nov., a Novel Decabromodiphenyl Ether (BDE-209)-tolerant Strain Isolated from Coastal Sediment in Xiamen, China. Biomed Environ Sci 34:656-661
30. Tarailo-Graovac M, Chen N (2009) Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences. Current Protocols in Bioinformatics 25:4.10.11-14.10.14
31. Wang C, Zhao X, Wu K, Liang C, Liu J, Yang H, Wang C, Yang B, Yin F, Zhang W (2023) Isolation and characterization of *Bacillus velezensis* strain B19 for biocontrol of *Panax notoginseng* root rot. Biol

Control 185:105311

32. Wang X, Zhang M, Loh B, Leptihn S, Ahmed T, Li B (2021) A novel NRPS cluster, acquired by horizontal gene transfer from algae, regulates siderophore iron metabolism in *Burkholderia seminalis* R456. *Int J Biol Macromol* 182:838-848
33. Wei J, Zhao J, Suo M, Wu H, Zhao M, Yang H (2023) Biocontrol mechanisms of *Bacillus velezensis* against *Fusarium oxysporum* from *Panax ginseng*. *Biol Control* 182:105222
34. Wemheuer F, Taylor JA, Daniel R, Johnston E, Meinicke P, Thomas T, Wemheuer B (2020) Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. *Environ Microbiome* 15:11
35. Xie L, Liu L, Luo Y, Rao X, Di Y, Liu H, Qian Z, Shen Q, He L, Li F (2023) Complete genome sequence of biocontrol strain *Bacillus velezensis* YC89 and its biocontrol potential against sugarcane red rot. *Front Microbiol.* 14:1180474
36. Xu P, Xie S, Liu W, Jin P, Wei D, Yaseen DG, Wang Y, Miao W (2020) Comparative Genomics Analysis Provides New Strategies for Bacteriostatic Ability of *Bacillus velezensis* HAB-2. *Front Microbiol* 11:594079
37. Xu S, Zhang Z, Xie X, Shi Y, Chai A, Fan T, Li B, Li L (2022) Comparative genomics provides insights into the potential biocontrol mechanism of two *Lysobacter enzymogenes* strains with distinct antagonistic activities. *Front Microbiol* 13:966986
38. Yang Y, Wei F, Li Z, Han F, Guan H, Huang J (2023) Improper crop rotation may enrich soil-borne pathogens of *Panax notoginseng*. *J Phytopathol* 171:567-576
39. Zeng Q, Xie J, Li Y, Gao T, Zhang X, Wang Q (2021a) Comprehensive Genomic Analysis of the Endophytic *Bacillus altitudinis* Strain GLB197, a Potential Biocontrol Agent of Grape Downy Mildew. *Front Genet* 12:729603
40. Zeng Q, Xie J, Li Y, Gao T, Zhang X, Wang Q (2021b) Comprehensive Genomic Analysis of the Endophytic *Bacillus altitudinis* Strain GLB197, a Potential Biocontrol Agent of Grape Downy Mildew. *Front Genet* 12:729603
41. Zhang R, Ouyang J, Xu X, Li J, Rehman M, Deng G, Shu J, Zhao D, Chen S, Sayyed RZ (2022) Nematicidal Activity of *Burkholderia arboris* J211 Against *Meloidogyne incognita* on Tobacco. *Front Microbiol* 13:915546
42. Zhu H, Xu C, Chen Y, Liang Y (2022) His-Ala-Phe-Lys peptide from *Burkholderia arboris* possesses antifungal activity. *Front Microbiol* 13:1071530

Figures

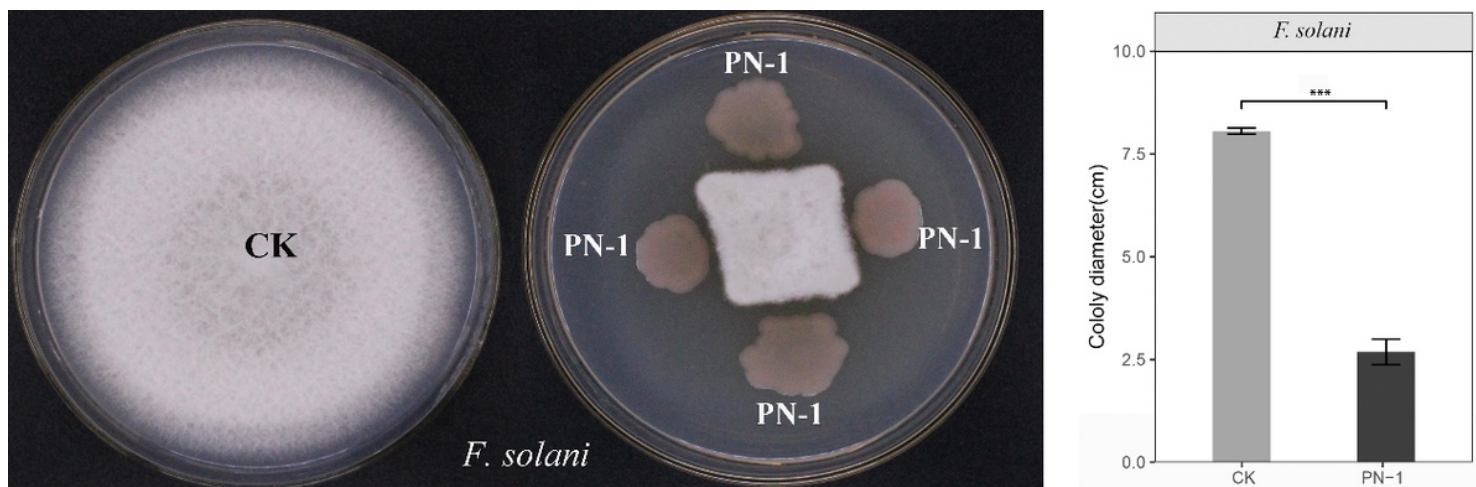
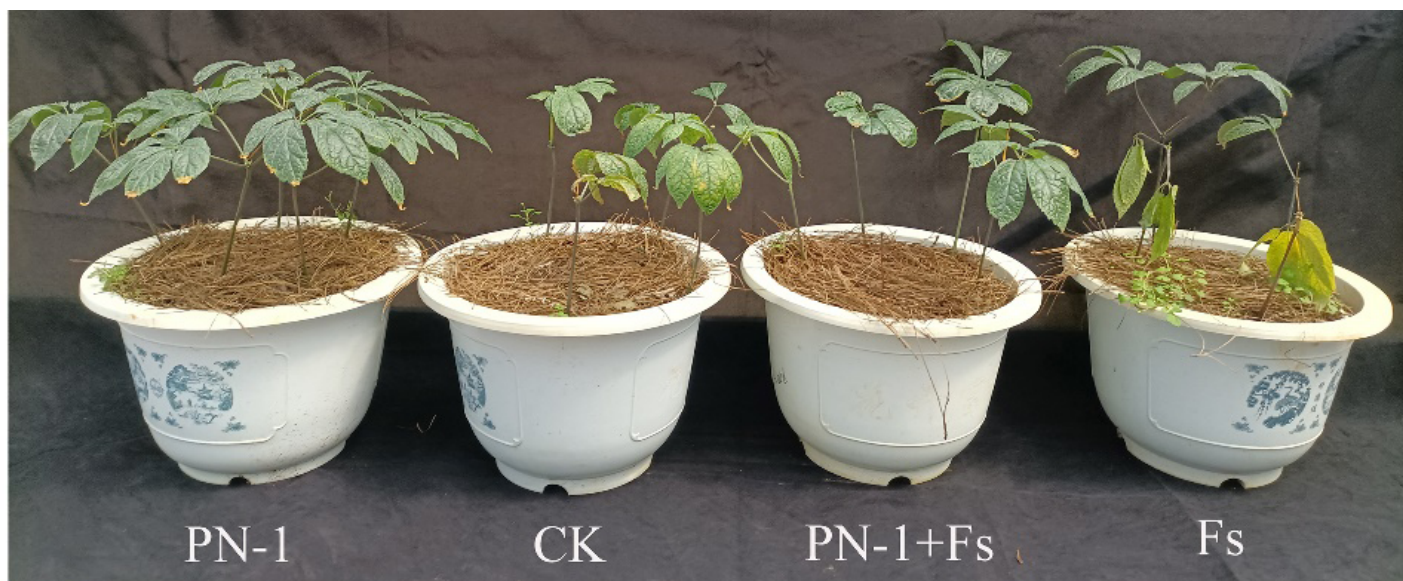


Figure 1

The inhibition of *F. solani* mycelial growth by PN-1. "***" indicates a significant difference at $P < 0.001$.

A



B

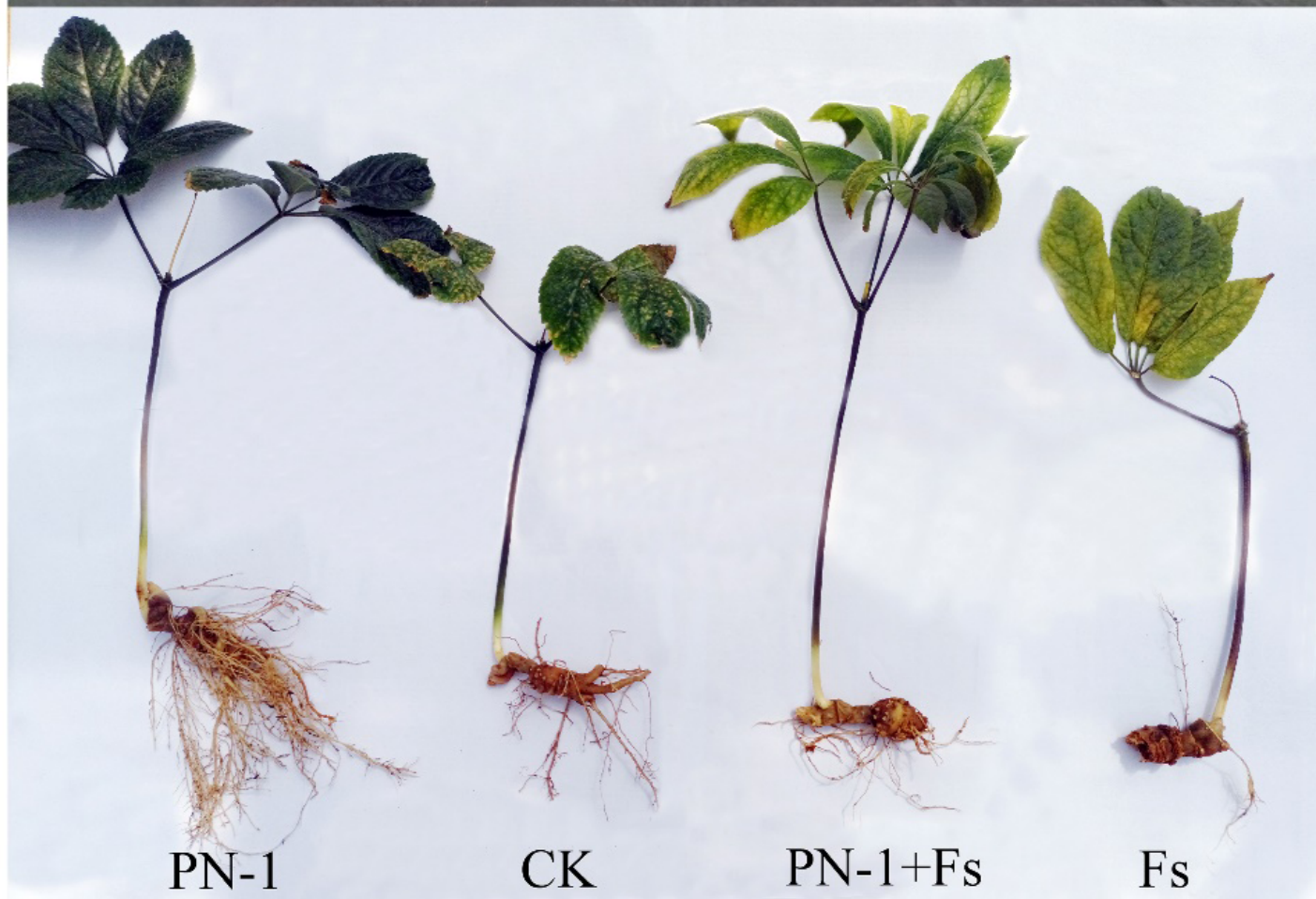


Figure 2

Greenhouse control efficacy of *B. arboris* PN-1 on Sanqi root rot.

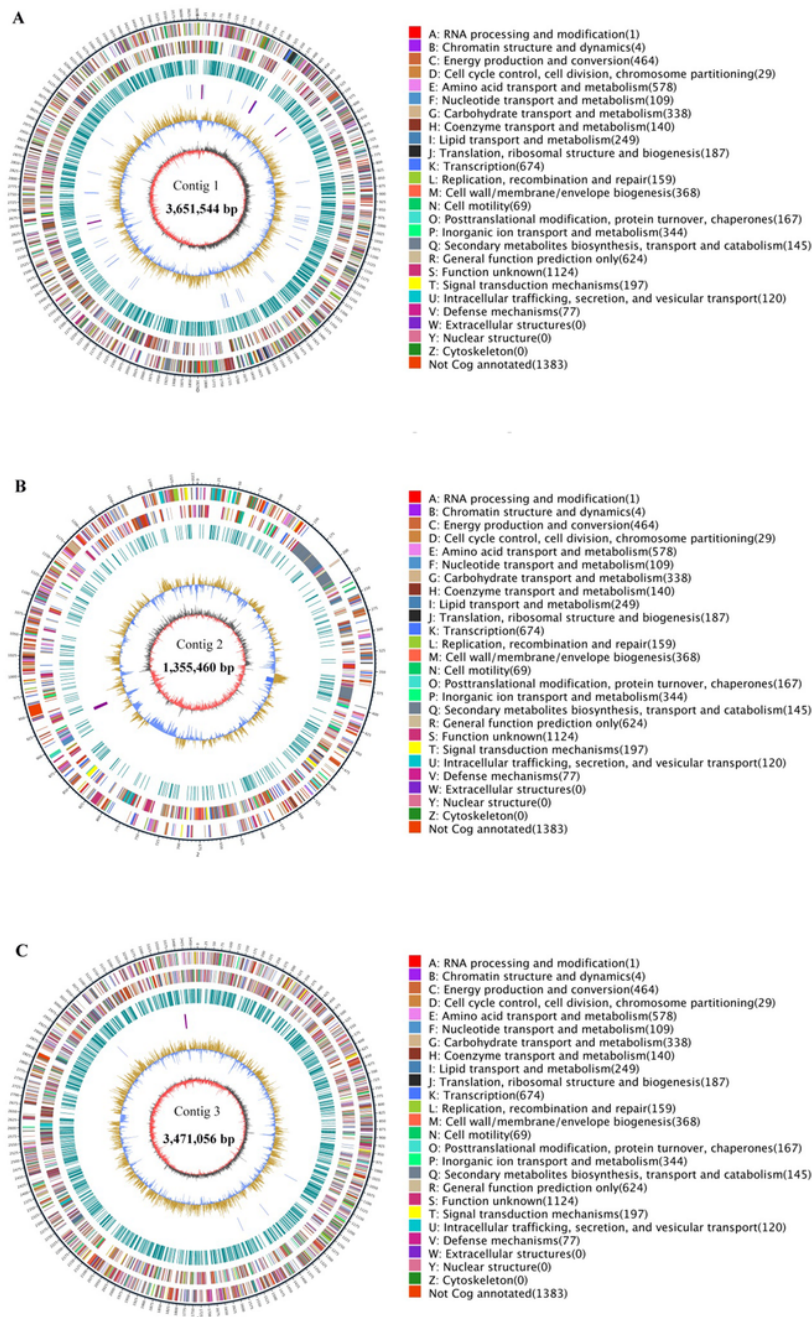


Figure 3

Genome map of *B. arboris* PN-1 showing 3 contigs (A: Contig 1, B: Contig 2, C: Contig 3). The circular map is organized into six concentric circles. From the outermost to the innermost: Circle1: represents the scale of the genome size; Circles 2 and 3: illustrate genes on the plus and minus strands of the genome, respectively; Circle 4: depicts repeat sequences; Circle 5: represents ribosomal RNA sequences (tRNA and rRNA); Circle 6 represents the GC content across the genome; Innermost Circle: Represents the GC skew.

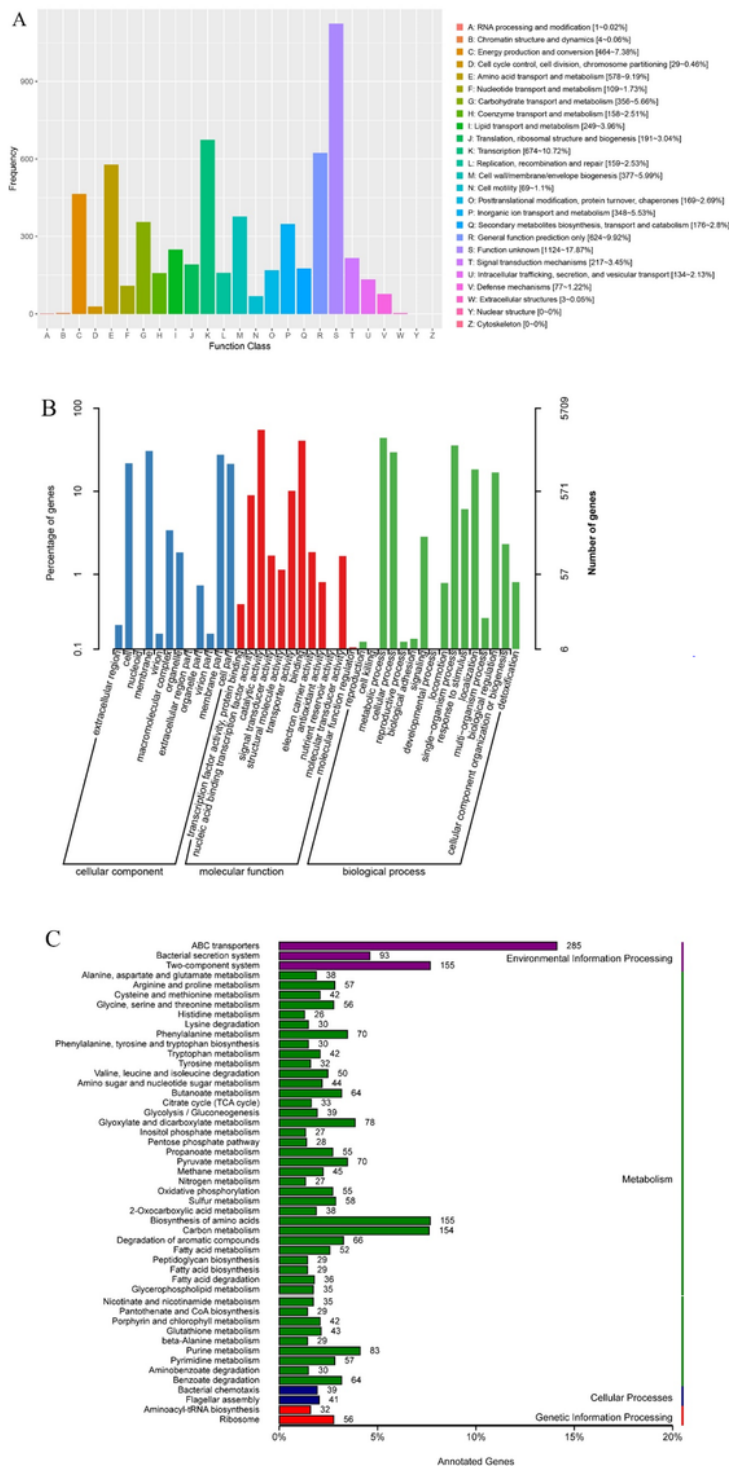


Figure 4

Gene annotation by COG, GO, and KEGG databases for *B. Arboris* PN-1. (A) COG function classification of genes in PN-1; (B) GO function classification of genes in PN-1; (C) KEGG pathway classification of genes in PN-1.

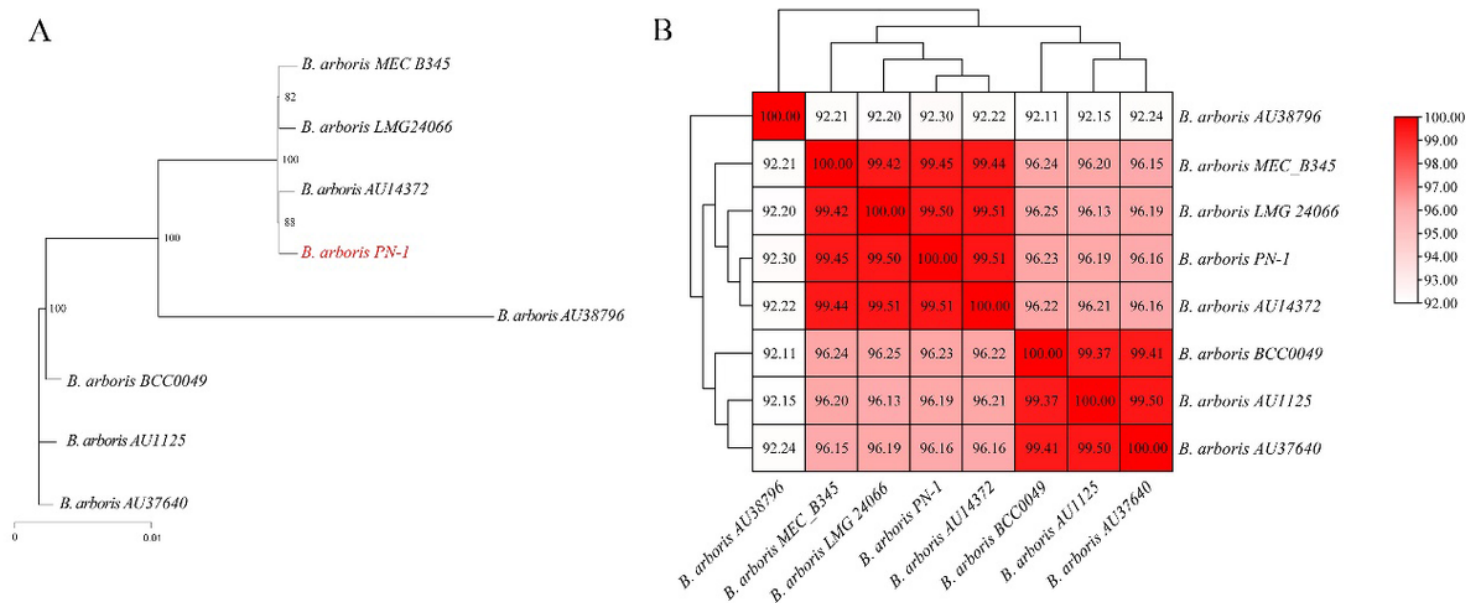


Figure 5

Phylogenetic analysis of *B. arboris*PN-1. (A) Phylogenetic tree for *B. arboris* PN-1 and the seven other *B. arboris* strains based on homologous genes. (B) Heat-map of ANI values among various strains of *B. arboris*.

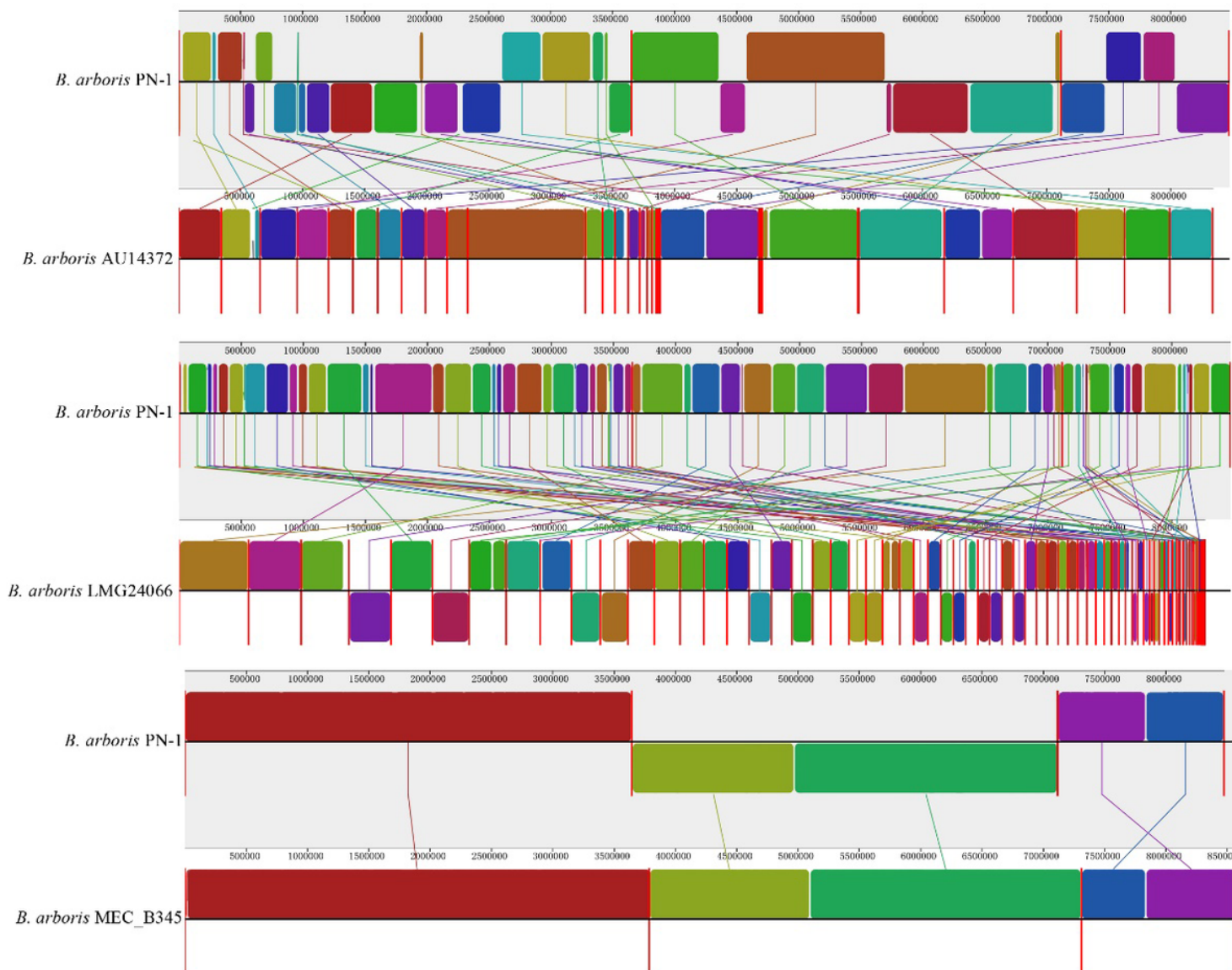


Figure 6

Mauve synteny alignment analysis of *B. arboris* Genomes. Mauve synteny alignment analysis illustrates the genomic sequences of *B. arboris* PN-1 and three other strains of *B. arboris*. Boxes of distinct colors represent the locally collinear blocks (LCBs). Connected lines of the same color indicate the position of LCBs on the corresponding genomes. LCBs above the reference line indicate sequence direction consistency with the reference genome. LCBs below the reference line indicate sequence direction opposition to the reference sequence. White areas signify genome-specific sequence elements, potentially containing unique genomic content or regions not adequately aligned between the compared genomes.

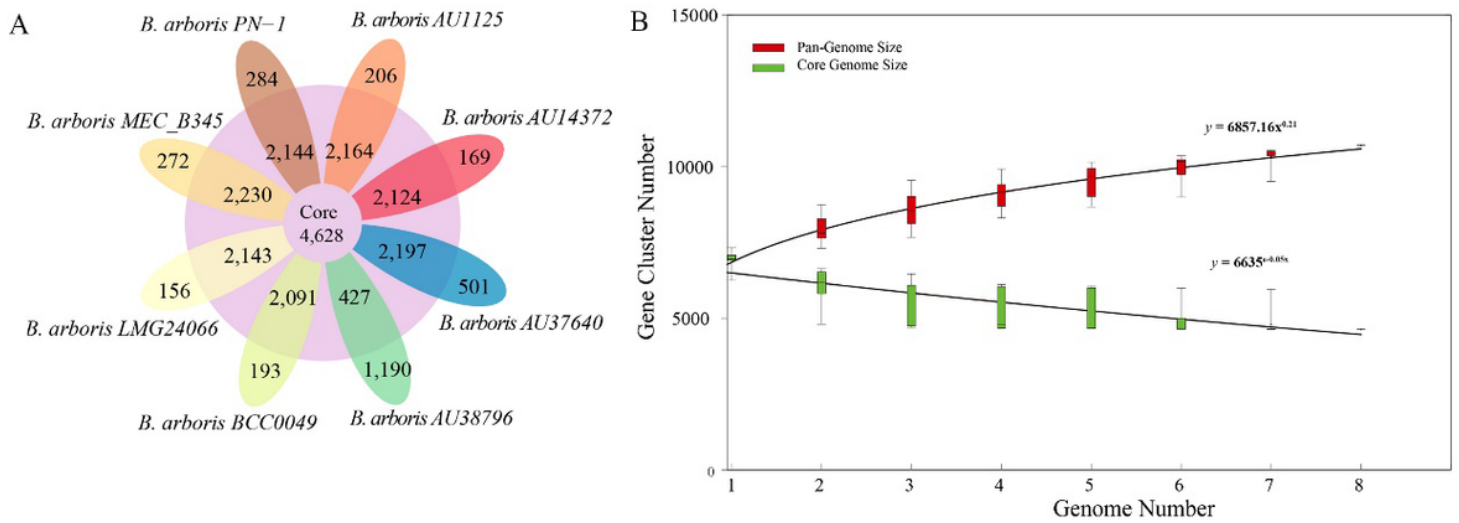


Figure 7

Pan-genomes of *B. arboris* strains. (A) Flower plot displaying the numbers of specific, unique, and core genes among the genomes of PN-1 and seven other strains of *B. arboris*. The inner circle illustrates the core genomes shared between all strains. The outer circles indicate specific genes for each strain. (B) Curves for *B. arboris* pan-genomes and core-genomes. Red boxes represent the pan-genome size of *B. arboris* for each genome comparison. Green boxes represent the core genome size of *B. arboris* for each genome comparison. Median values are linked to display the relationship between the number of genomes and gene families.

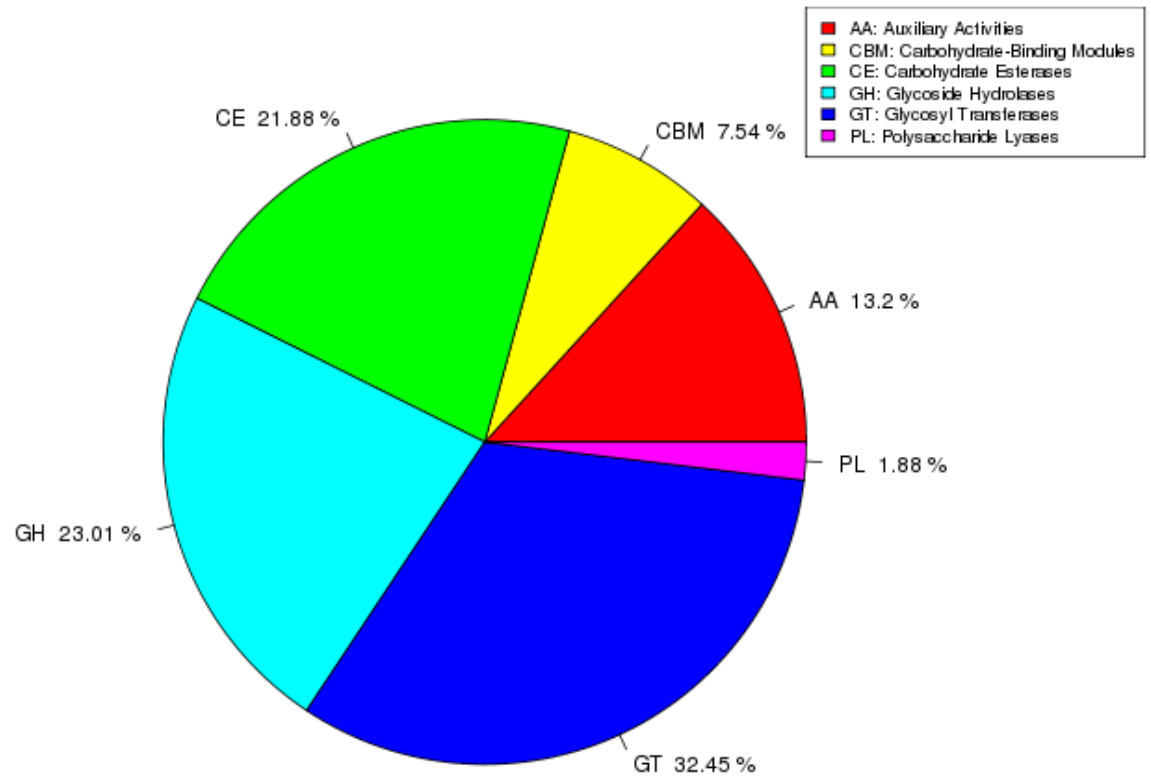


Figure 8

Gene count distributions of CAZy families in *B. arboris* PN-1 genome