Functional Insights into Novel Extremophilic Bacteria Isolated from the NASA Phoenix Mission Spacecraft Assembly Cleanrooms

Júnia Schultz
King Abdullah University of Science and Technology
Tahira Jamil
King Abdullah University of Science and Technology
Pratyay Sengupta
Indian Institute of Technology Madras
Shobhan Karthick Muthamilselvi Sivabalan
Indian Institute of Technology Madras
Anamika Rawat
King Abdullah University of Science and Technology
Niketan Patel
King Abdullah University of Science and Technology
Srinivasan Krishnamurthy
Institute of Microbial Technology
Intikhab Alam
King Abdullah University of Science and Technology
Nitin Kumar Singh
Jet Propulsion Laboratory
Karthik Raman
Indian Institute of Technology Madras
Alexandre Soares Rosado
King Abdullah University of Science and Technology
Kasthuri Venkateswaran
kjvenkat1955@gmail.com

Jet Propulsion Laboratory

Article

Keywords: Extreme environment, Cleanroom, Phylogenomics, Novel species, Phoenix mission, BGCs.

Posted Date: September 2nd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4830964/v1

Additional Declarations: No competing interests reported.

Functional Insights into Novel Extremophilic Bacteria Isolated from the NASA Phoenix Mission Spacecraft Assembly Cleanrooms

Júnia Schultz[1], Tahira Jamil[1], Pratyay Sengupta[2,3,4], Shobhan Karthick Muthamilselvi Sivabalan[4], Anamika Rawat[1], Niketan Patel[1], Srinivasan Krishnamurthy[5], Intikhab Alam[1], Nitin K. Singh[6], Karthik Raman[3,4,7], Alexandre Soares Rosado[1,8*], Kasthuri Venkateswaran[6*]

[1]Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Makkah, 23955, Saudi Arabia
[2]Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai, 600036, India
[3]Center for Integrative Biology and Systems Medicine (IBSE), Indian Institute of Technology Madras, Chennai, 600036, India
[4]Robert Bosch Centre for Data Science and Artificial Intelligence (RBCDSAI), Indian Institute of Technology Madras, Chennai, 600036, India
[5]Microbial Type Culture Collection and Gene Bank (MTCC), Institute of Microbial Technology, Chandigarh, 160036, India
[6]NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[7]Department of Data Science and AI, Wadhwani School of Data Science and AI, Indian Institute of Technology Madras, Chennai, Tamil Nadu, 600036, India
[8]Bioscience Program, Biological and Environmental Science and Engineering (BESE), Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah, 23955, Saudi Arabia

* These authors share senior authorship.

Correspondence:
kjvenkat1955@gmail.com; Orcid: 0000-0002-6742-0873
alexandre.rosado@kaust.edu.sa; Orcid: 0000-0001-5135-1394

Keywords: Extreme environment, Cleanroom, Phylogenomics, Novel species, Phoenix mission, BGCs.

Abstract

Human-designed oligotrophic environments, such as cleanrooms, harbor unique microbial communities shaped by selective pressures like temperature, humidity, nutrient availability, cleaning reagents, and radiation. During the Phoenix spacecraft mission, genomes of 215 bacterial isolates were sequenced and based on overall genome related indices, 53 strains belong to 26 novel species were recognized. Metagenome mapping indicated less than 0.1% of the reads associated with novel species, suggesting their rarity. Genes responsible for biofilm formation, such as BolA (COG0271) and CvpA (COG1286), were predominantly found in proteobacterial members but were absent in other non-spore-forming and spore-forming species. YqgA (COG1811) was detected in most spore-forming members but was absent in *Paenibacillus* and non-spore-forming species. Cell fate regulators, COG1774 (YaaT), COG3679 (YlbF, YheA/YmcA), and COG4550 (YmcA, YheA/YmcA), controlling sporulation, competence, and biofilm development processes, were observed in all spore-formers but were missing in non-spore-forming species. COG analyses further revealed resistance-conferring proteins in all spore-forming novel species (n=13) and eight actinobacteria, responsible for enhanced membrane transport and signaling under radiation (COG3253), transcription regulation under radiation stress (COG1108), and DNA repair and stress responses (COG2318). Additional functional analysis revealed that *Agrococcus phoenicis, Microbacterium canaveralium,* and *Microbacterium jpeli* contained biosynthetic gene clusters (BGCs) for ε-Poly-L-lysine, beneficial in food preservation and biomedical applications. Two novel *Sphingomonas* species exhibited zeaxanthin, an antioxidant beneficial for eye health. *Paenibacillus canaveralius* harbored genes for bacillibactin, crucial for iron acquisition. *Georgenia phoenicis* had BGCs for alkylresorcinols, compounds with antimicrobial and anticancer properties used in food preservation and pharmaceuticals. Despite strict decontamination and controlled environmental conditions, cleanrooms foster novel bacterial species which can form biofilms, resist various stressors, and produce biotechnologically valuable compounds.

Introduction

Cleanrooms and other human-designed oligotrophic environments present distinct ecosystems that may expedite microbial speciation due to unique selective pressures[1]. These pressures may arise from specialized construction materials, controlled temperature and humidity, and exposure to cleaning agents, diverging from more nutrient-rich natural settings[2]. Such environments select microbes that can survive nutrient-poor conditions, potentially giving rise to new species[3]. Globally, human activities transport microbes to different oligotrophic environments, like cleanrooms, facilitating distinctive evolutionary trajectories[2]. Despite the resource-limited conditions, microbial communities in these controlled environments are complex and competitive. This complexity fosters the isolation of rare microbes, shaped by specific microbial interactions and environmental conditions, unfolding over years to decades.

Ensuring the biological cleanliness of the National Aeronautics and Space Administration's (NASA) mission-associated cleanrooms, where spacecraft are assembled and tested, is imperative to meet planetary protection requirements[4]. These facilities undergo constant monitoring to detect and assess the presence of any microorganisms that could potentially survive a transfer to an extraterrestrial environment via robotic exploration devices[5-7]. Despite meticulous control measures, including regulation of airflow, humidity, temperature, and air particulate concentrations, along with rigorous cleaning using chemical detergents, UV radiation, and hydrogen peroxide, certain microorganisms can persist in this challenging and nutrient-limited environment[8-11].

The "cleanroom effect" may provide a platform for microorganisms to adapt to selective pressures (i.e., extremely oligotrophic, low-humidity, and desiccation conditions), bolstering their growth, survival, lifestyle, and resilience under extreme conditions, and the production of specialized metabolites[11,12]. It is crucial to characterize these resistant microbes, which defy conventional biological control measures and potentially identify novel microbial species. This effort is pivotal for monitoring the risk of forward microbial contamination and safeguarding extraterrestrial environments against unintentional colonization of exploring planets[13].

During the Phoenix mission, 215 strains were isolated from the Kennedy Space Center - Payload Hazardous Servicing Facility (KSC-PHSF) cleanroom floors under various extreme conditions[9] and whole genome sequencing (WGS) of all 215 isolates was performed. The central objectives of this study were to characterize a cohort of 53 strains, representing 26 previously unidentified bacterial species discovered among Phoenix mission isolates. These strains were subjected to extensive examination, which included characterizing their physiological attributes, and conducting thorough genome analysis, followed by in-depth phylogenomic assessments. Evaluations were performed to determine the incidence, prevalence, and persistence of these novel species even after nine years by analyzing metagenomic reads sourced from several NASA cleanrooms, including KSC-PHSF. In parallel, an investigation into the genomic functions of these extremophiles was undertaken, with a particular emphasis on the discovery of potential genes responsible for radiation resistance and secondary metabolites, indicative of their adaptive capacity and biotechnological applications (Supplementary Figure S1).

Results

Based on WGS, the bacterial strains (n=215) isolated from the KSC-PHSF were classified into three phyla: *Actinomycetota, Bacillota,* and *Pseudomonadota*. Furthermore, around 25% of the bacterial strains (53 out of 215 isolates), were novel species and most of them belong to the members of the class *Bacilli* (47.7%), *Alphaproteobacteria* (24.5%), *Gammaproteobacteria* (13.9%), and *Actinomycetia* (13.9%). The percent occurrence of the novel species at family level is given in Supplementary Figure S2.

Among 53 strains, 26 belong to yet-to-be-described bacterial species, encompassing 18 genera. Within these 53 novel bacterial strains, 33 strains were isolated before the arrival of the Phoenix mission spacecraft to the KSC-PHSF cleanroom (21 novel species), 7 were cultured during the assembly and testing of the spacecraft (3 novel species), and 13 were isolated from the cleanroom floors after moving the spacecraft for the launch (2 novel species). Among the 53 novel extremo-tolerant strains, 22 were isolated under the alkaline condition (>pH 10; alkalophile), eight after heat-shock (80˚C; 15 min; heat-tolerant), seven grown at 4˚C (psychrophile), six at 25˚C (mesophile), five under anaerobic atmosphere, and five after exposing to UVC condition (254_{nm}; 1,000 J/m^2).

Genome features and relatedness indices

The isolation source, conditions, and WGS assembly statistics of the 53 novel strains is presented in Supplementary Table S1. The draft genomes of the novel species generated using the Nanopore platform were constructed with high-quality sequences, with assembly quality ranging from the complete genome (*n*=20) to 8 scaffolds, and many of the strains exhibited >99% completeness. The similarities among the closely related species of the novel species based on marker genes (16S rRNA and gyrB), average nucleotide index (ANI), average amino acid index (AAI), and digital DNA:DNA hybridization (dDDH) are given in Table 1. Moreover, ANI indices (<95%) and dDDH values (<70%) fell below the threshold levels of bacterial species identity, confirming that the examined Phoenix mission strains (n=53) were novel species. The ANI index ranged from 79 to 94%, with most of the 53 novel strains having less than 90% of ANI similarity with the closest relatives. Since no set threshold values for AAI and bacterial genus discrimination exist, it could not be definitively determined whether any of these novel species belong to new genera.

Phylogenomic analysis

The phylogenomic analysis based on 16S rRNA gene, gyrB, and WGS was performed, and these novel organisms were placed in their respective phylogenetic trees to determine their precise taxonomic placement.

Members of *Actinomycetota* phylum showed varied ANI index similarities when compared to established species. The ANI index of *Agrococcus phoenicis* 1P02AA revealed a low similarity (79-86%) with already recognized *Agrococcus* species, with *Agrococcus carbonis* being the closest species at 86% ANI. However, based on the single-copy core genes, *Agrococcus baldri* was the closest species, with 85.63% ANI. Our three strains from *Arthrobacter phoenicis* exhibited 100% similarity among themselves and were closely related to *Arthrobacter oryzae*, with 83% ANI. The species *Curtobacterium phoenicis* was closely related to *Curtobacterium luteum*, exhibiting an ANI similarity of 89%. The strains belonging to *Georgenia phoenicis* (1P01AC and 1P07AB) were 100% similar to each other and presented an ANI of 89% to the closest relative, *Georgenia satyanarayanai*. The species belonging to *Microbacterium* genus (*M. canaveralium, M. jepli*, *M. phoenicis* and *M. pratiae)* presented ANI values ranging from 84 to 93% compared with their closest relatives.

Four novel species were identified as belonging to *Pseudomonadota* phylum. *Noviherbaspirillum phoenicis* were closely related to *Noviherbaspirillum soli,* exhibiting an ANI of 94%. The novel species *Brevundimonas phoenicis* comprising 18 strains, clustered together with 100% ANI and showed 93% ANI similarity with *Brevundimonas diminuta*. Similarly, the four strains of *Pseudomonas phoenicis* were grouped with 100% ANI and showed ~86% similarity with the closest species *Pseudomonas cremoricolorata*. The novel species *Sphingomonas canaveralia* was placed near *Sphingomonas jatrophae* with an 79% ANI, and *Sphingomonas phoenicis* was adjacent to *Sphingomonas metalli* with an of 83% ANI.

Additionally, among the strains belonging to the *Bacillota* phylum, the *Alkalihalobacillus* and *Shouchella*

genera were placed in the same phylogenomic tree due to their similarity. *Alkalihalobacillus phoenicis* 1P02AB was closest to *Alkalihalobacillus alcalophilus* with an ANI of 92%, while the strains of *Shouchella phoenicis* were similar among themselves and closest to *Shouchella hunanensis* with an ANI of 81%. The novel species *Bacillus jepli,* and *B. kalamii* were not closely related, with ANI value of 80%, and similar patterns were observed when compared with other strains of *Bacillus* genus (ANI ranging from 76 to 83%). *Lysinibacillus canaveralius* clustered with *Lysinibacillus odysseyi,* presenting an ANI of 84%, while *Lysinibacillus phoenicis* was closely related to *Lysinibacillus fusiformis* with an ANI of 85%. Two species of *Neobacillus*, *N. canaveralius* and *N. phoenicis*, were distant from each other, with an ANI of 79%. Phylogenetic analysis revealed that *N. canaveralius* is closer to *N. niacini*, with an ANI of 87%, while *N. phoenicis* is closer to *N. bataviensis*, with an ANI of 79%. The single species *Oceanobacillus phoenicis* presented a high ANI percentage with its closest relative, *Oceanobacillus kimchi* with ANI of 90%. The novel representatives of *Paenibacillus, P. jepli* and *P. canaveralius*, were closely related to *P. daejeonensis* (ANI 81%) and *P. chitinolyticus* (ANI 90%), respectively. Two strains of *Peribacillus phoenicis* clustered together with 100% ANI, and showed 94% similarity to *Peribacillus frigoritolerans*, its closest relative. The species *Robertmurraya phoenicis* was similar to *Robertmurraya massiliosenegalensis,* with an 91% ANI.

To further validate the placement of the novel species within the bacterial tree of life, a phylogenetic tree was generated by comparing them with 4,441 complete, non-anomalous representative genomes of bacteria (Supplementary Figure S3). The tree of life showed that these novel genomes are almost distributed across the entire spectrum, indicating that spacecraft assembly cleanrooms can harbor a wide range of bacterial diversity. Additionally, 17 phylogenetic trees were constructed at the genus level, with Figure 1 representing non-spore-formers and Figure 2 representing spore-formers.

Morphological characterization

To analyze the bacterial isolates in further detail, Gram staining was performed on each isolate. Of all the isolates, 69% were Gram-negative, while the rest (31%) were characterized as Gram-positive bacteria. For in-depth morphological characterization scanning electron microscopy (SEM) analysis was carried out for all the isolates characterized as novel species. Many of the bacterial cells exhibited round or rod-shaped morphologies, presenting either as single cells or in aggregation of multiple cells. The details of the microscopic characterization of each isolate are presented in Table 2, based on SEM images (Figure 3) and Gram staining images (Supplementary Figure S4). The novel species etymologies are given in Table 2.

Persistence of novel species

Quality-filtered shotgun metagenomic reads were mapped onto 26 isolated novel species to assess their abundance based on the fraction of mapped reads and coverage breadth. Non-spore-formers had significantly more reads than spore-formers (Figure 4A). Due to the limited proportion of mapped reads to novel species (<1%), a read assembly was conducted to assess coverage breadth against isolated genomes. The average coverage breadth ranged from 0.0007% to 64.4% in JPL-SAF during 2016, from 0.00045% to 3.93% in JPL-SAF during 2018, and from 0.0004% to 6.8% in KSC-PHSF during 2018 (Supplementary Table S2). Using a 1% cutoff, the distribution of coverage breadth for novel species showing >1% coverage (n=23 species) in at least one sample is plotted in Figure 4B. *B. phoenicis* demonstrated the highest mapping percentage, an anomaly, comprising 64.4% of total reads in a sample from location 9 in JPL-SAF. Additionally, *M. jepli* and *G. phoenicis* were present in more samples (n=108) with >1% coverage, followed by *P. phoenicis* (n=105) and *A. phoenicis* (n=104). Furthermore, three novel species *(A. phoenicis, O. phoenicis,* and *P. jpeli*) were <1% in

their abundance in any of the samples and are not shown in Figure 4B. This indicates that none of these 26 novel species dominate the cleanrooms and might be rare.

Functional characterization

Putative functions of the 26 novel bacterial species were annotated using Prokka and COG-classifier. A total of 212,520 CDS with 3,807 distinct COG annotations were identified (Supplementary Table S3). Among the annotated subsystems, the top categories based on average gene counts included amino acids transport and metabolisms (259 genes), followed by transcription (232 genes), translation, ribosomal structure and biogenesis (229 genes), and carbohydrate transport and metabolism (225 genes). Further analysis of these organisms from the Phoenix spacecraft mission revealed that, on average, they possessed 74 genes predicted for defense mechanisms, primarily related to resistance to antibiotics and toxic compounds, and invasion and intracellular resistance.

Key genes potentially related to radiation resistance were observed across different bacterial isolates (Figure 5A). The COG3253 proteins that were responsible for enhanced membrane transport and signaling under radiation were present in in all spore-formers (n=13) and eight novel actinobacterial species during this study. COG0608 genes, highlighting their role in DNA repair, were absent in all eight actinobacterial species but present in 18 other novel species. COG1108 genes, related to transcription regulation under radiation stress, were present in all novel species except alpha- and beta-proteobacteria (n=24). COG1971 proteins involved in DNA repair after radiation exposure were found in all 13 spore-formers and five out of 13 non-spore-forming novel species. COG2318 proteins, associated with DNA repair and stress responses, were identified in spore-formers and *A. phoenicis*. COG4365 genes, responsible for increased radiation resistance, were present in all spore-formers but absent in other novel species. The involvement of COG4119 proteins in nucleotide excision repair pathways was reported in *Bacillus subtilis*, and in this investigation, this protein was present only in *N. canaveralius* whereas 12 other novel spore-formers lacked it.

The KMAP approach was used to recover the dataset of proteins of interest (POIs) from the novel species (Supplemental Figure S4A). While exploring the metabolic potential of these novel species, various noteworthy observations were made, including the annotation of several hundred proteins in different application categories. Notably, higher numbers of proteins related to bioprocess engineering, medicine and pharmaceuticals, and analytics were observed, particularly those involved in synthesis, drug development, agriculture, the food industry, and molecular biology. POIs relevant to withstanding extremophilic conditions (such as high temperature and alkalinity) were also identified.

Biofilm formation

The biofilm-associated COG proteins observed across various bacterial isolates are depicted in Figure 5B. The DNA-binding global transcriptional regulator BolA, which affects cell shape, cell division, and biofilm formation (COG0271), was identified exclusively in proteobacterial members (5 species; 25 strains). Like BolA, the colicin V production accessory protein CvpA, a regulator of purF expression and biofilm formation (COG1286), was also present in proteobacterial members but absent in non-spore-forming species. Conversely, all novel spore-formers (12 species; 14 strains) except *Pa. jepli* contained COG1286. The membrane protein YqgA (COG1811), associated with biofilm formation, was found in most spore-forming members but was absent in both *Paenibacillus* species and all non-spore-forming members during this study. Membrane-bound acyltransferase YfiQ (COG3936), involved in biofilm formation and previously found in Yersinia pestis, was present only in *L. phoenicis* and not in any other 25 novel species identified in this research. A group

of functionally related cell fate molecular regulators that controlled sporulation, competence, and biofilm development processes and events through modulation of gene and protein expression, such as COG1774 (YaaT), COG3679 (YlbF, YheA/YmcA), and COG4550 (YmcA, YheA/YmcA), was detected in all spore-formers but was absent in non-spore-forming species during this study.

Antimicrobial resistance

Several AMR gene families were identified across the genomes, indicating resistance to ten distinct drug classes, with a predominance for fluoroquinolones, tetracyclines, disinfecting agents/antiseptics, phosphonic acids, and glycopeptides (Figure 5C). The 53 genomes exhibited potential resistance to vancomycin and tetracycline antibiotics (Supplementary Table S4). The species *B. canaveralius, B. jepli, B. phoenicis, L. phoenicis, L. canaveralius, P. canaveralius, P. jepli, R. phoenicis, N. canaveralius* and both strains of *Pe. phoenicis* (IP06PA-2 and 1P06PB) presented the higher amount of resistance genes. In terms of antibiotic resistance, five mechanisms were identified: the most common was antibiotic efflux, followed by antibiotic target alteration, antibiotic inactivation, and less commonly, antibiotic target protection and antibiotic target replacement. Overall, the genomic mining predicted the presence of 21 AMR genes, however, phenotypic investigation is necessary to validate the mechanism.

Biosynthetic Gene Clusters

A biosynthetic gene cluster (BGC) analysis revealed 11 cluster types across 26 novel species, with T3PKS and terpene clusters being the most abundant (Supplementary Figure S4B). *P. jepli* 1P07SE and *S. phoenicis* 1P01AA exhibited the highest number of BGCs, with 12 and 10 BGCs each, respectively. BGCs from isolates showing >80% similarity with known gene clusters, including alkylresorcinol, carotenoid, ε-Poly-L-lysine, and paeninodin, were observed in 17 isolates (Supplementary Table S5. The ε-Poly-L-lysine, known for its wide-spectrum inhibitory activity, heat stability, and biodegradability as a food preservative, was identified in three species (*A. phoenicis*, *M. canaveralium*, *M. jepli*) with 100% similarity. A gene cluster neighborhood comparison of ε-Poly-L-lysine with known producers revealed functional ε-Poly-L-lysine synthetase genes. Protein sequence comparison showed 48% identity with the fungal producer *Epichloe festucae* and around 67% identity with the bacterial producer *Corynebacterium variabile*, with the highest 70.8% identity in *M. canaveralium* (Figure 6). Domain analysis indicated conserved non-ribosomal peptide synthetases adenylation (A) and thiolation (T) domains, six transmembrane (TM) domains, and three C-terminal tandem domains, crucial for substrate binding and lysine polymerization. This suggests potential for producing ε-Poly-L-lysine, effective against foodborne pathogens like *E. coli* O157:H7, *Listeria monocytogenes*, *Staphylococcus aureus*, and *Serratia marcescens*.

Discussion

Several factors contributed to the higher percentage of novel cultivable species (~25%; 53 out of 215 strains) retrieved from cleanrooms compared to 6 to 12% in natural environments[14,15]. Studies demonstrated that extreme and controlled environments might select unique microbial communities capable of thriving under harsh conditions (low-nutrient, desiccation, etc.), which potentially drive microbial speciation and adaptation[16-20]. In cleanrooms, traditionally spore-formers are often reported. However, non-spore-formers such as *Arthrobacter, Brevundimonas, Georgenia, Microbacterium,* and *Pseudomonas* species, which can survive in oligotrophic, arid, and radiation conditions[21-25], should also be considered when setting bioburden requirements for future NASA missions. Additionally, the isolation of spore-formers like *Peribacillus* and *Shouchella* species, which require different cultural conditions compared to *Bacillus* species, underscores the importance of WGS in characterizing yet-to-be-recognized cultivable microbial species[26,27]. Research on microbial isolates from the Atacama

Desert further supported the notion that oligotrophic conditions and unique environmental pressures led to the discovery of more novel microbial taxa[28,29]. The comprehensive genome analysis of the novel species revealed the presence of already established/peer-reviewed genetic adaptations that enable bacteria to survive extreme conditions, including genes responsible for resistance to radiation, desiccation, and other environmental stressors.

Experimental studies on *HemQ* (COG3253), also known as coproheme decarboxylase/chlorite dismutase, have demonstrated its significant role in coenzyme transport and metabolism, as well as inorganic ion transport and metabolism. In gram-positive bacteria (*Bacillota* and *Actinomycetota*) *HemQ* plays an essential role and has been associated with respiration, detoxification of reactive oxygen (ROS) and nitrogen species, gas sensing, and transport[30], a crucial property for stress survivability. This linkage (COG3253) was observed in all spore-formers (n=13) and eight novel actinobacterial species in this study. Knockout experiments of COG0608 genes resulted in increased radiation sensitivity, demonstrating their role in DNA repair[31]. These genes were absent in all eight actinobacterial species found in this study but present in 18 other novel species. Except in alpha- and beta-proteobacteria, all other novel species (n=24) exhibited COG1108 genes, which are related to transcription regulation under radiation stress, potentially confirming their protective role[32]. All 13 spore-formers and five out of 13 non-spore-forming novel species exhibited the presence of COG1971 proteins involved in DNA repair after radiation exposure, as reported to be upregulated in *D. radiodurans*[33]. Spore-formers and *A. phoenicis* had COG2318 proteins, which were experimentally proved to respond to radiation using a transcriptomic study in *D. radiodurans*, indicating potential roles in DNA repair and stress responses. All spore-formers, but not other novel species, exhibited the presence of COG4365 genes that were shown to be responsible for increased radiation resistance, confirming their potential role in DNA repair[34]. Radiation exposure studies in *B. subtilis* confirmed the involvement of COG4119 proteins in nucleotide excision repair pathways[35]. However, the absence of COG4119 proteins in 12 out of 13 spore-forming novel species in this study requires further investigation. Despite rigorous decontamination procedures, microbes possessing these traits likely contribute to their persistence in cleanroom environments.

The metagenome analysis, which aimed to correlate the persistence of novel microbes within the assembly facility after more than a decade of their isolation, revealed that these novel bacterial species were rare microbial species due to their low incidence in shotgun metagenomes and the overall breadth of coverage for their genomes. Although individually rare, members of these novel bacterial communities collectively might have played crucial roles in ecosystem functioning and stability, including nutrient cycling, decomposition, and symbiotic interactions, potentially leading to the discovery of novel bioactive compounds, enzymes, and metabolic pathways[36-38].

Insights into the survival strategies of these extremophilic bacteria, thriving under the unique conditions of cleanrooms, were gathered through comprehensive genomic analyses. Genes responsible for the synthesis of compounds such as unknown NAGGN, extensively found in the novel strains, aided the bacteria in facing osmotic stress. The synthesis of NAGGN was induced to enhance bacterial colonization in various ecological niches[39]. This functional property, along with other traits like the presence of genes encoding proteins involved in stress response and adaptation, such as heat shock proteins, cold shock proteins, and chaperones, facilitated survival under harsh cleanroom conditions. This is of particular interest for future NASA missions, where understanding microbial resilience is crucial[40,41].

Biofilms are associated with antibiotic resistance, likely due to their organization, which protects bacteria in the inner layers from antimicrobial agents and promotes horizontal gene transfer of resistance genes[42-44]. BolA (COG0271) noticed in proteobacterial members of this study was shown

to be highly expressed in bacteria during the stationary phase and under stress conditions, suggesting its role in biofilm formation[45]. Overexpression of BolA in *E. coli* which promoted biofilm formation, while its absence produced thinner biofilms was reported[46]. Stress conditions such as nutrient depletion or oxidative stress resulted in significantly lower biofilm production in BolA mutants compared to the wild-type strain. *Brevundimonas* species during this study also possessed BolA that was reported to be forming biofilms with higher concentrations of antibiotic-resistant bacteria under disinfection pressure from chlorination and chloramination, increasing antibiotic resistance in tap water[47]. The membrane protein YqgA (COG1811) that was found to affect biofilm formation in *E. coli*[48] also retrieved in majority of the spore-formers during this study. In *Y. pestis*, biofilm formation increased significantly in cobB and yfiQ (COG3936) mutants, suggesting that they were the key players in biofilm formation. The cell fate regulators YmcA, YlbF, and YaaT (COG1744, COG3679, COG4550) were required for sporulation, competence, and biofilm formation[49]. Multiple transcriptional regulators were involved in complex cell differentiation in actinobacteria, cyanobacteria, and sporulating bacillota[50]. Genetic screens for mutants blocked in biofilm formation revealed that ylbF and ymcA genes played crucial roles, with YlbF and YmcA forming a complex with YaaT. Mutants lacking YaaT also showed impaired biofilm formation, competence, and sporulation[49,51,52].

A. phoenicis, M. canaveralium, and *M. jpeli* genomes had BGCs related to potential production of ε-Poly-L-lysine which is a versatile biopolymer with significant potential across various industries due to its strong antimicrobial activity and biodegradability. Its applications range from food preservation to biomedical and industrial uses, making it a valuable compound in enhancing product safety and longevity[53]. Both *Sphingomonas* species (n=2) possess BGCs related to zeaxanthin, a carotenoid produced by other sphingomonads, which is significant for its strong antioxidant properties, protecting cells from oxidative stress[54]. It plays a crucial role in photoprotection by absorbing blue light and preventing damage from UV radiation. In biotechnology, zeaxanthin is valued for its potential health benefits, including eye health, reducing the risk of age-related macular degeneration, and other chronic diseases.

P. canaveralius showed BGCs related to the production of bacillibactin, which is a siderophore produced by certain *Bacillus* species[55]. Siderophores are small, high-affinity iron-chelating compounds that microorganisms synthesize and secrete to sequester iron from the environment, which is vital for their growth and metabolism, especially under iron-limiting conditions. *P. jepli* contains BGCs related to producing bacillopaline, which is often used in agriculture as biocontrol agent and biofertilizer. Bacillopaline's antimicrobial properties can protect plants from pathogenic microorganisms, thus promoting healthier plant growth. By inhibiting plant pathogens, bacillopaline-producing bacterial strains can reduce the reliance on chemical pesticides, offering a more sustainable and environmentally friendly approach to agriculture[56].

All four strains of *Ps. phoenicis* exhibited BGCs related to carotenoids, which are reported to serve as powerful antioxidants and photoprotective agents, protecting cells from oxidative damage and UV radiation. They also enhance bacterial survival by aiding quorum sensing and biofilm formation, with significant applications in pharmaceuticals, cosmetics, and as food additives[57]. Similarly, both genomes of *G. phoenicis* contain BGCs related to alkylresorcinols, which are bioactive compounds known for their antimicrobial, antifungal, and anticancer properties[58]. They play a role in bacterial defense mechanisms and biofilm formation. Additionally, alkylresorcinols are used in pharmaceuticals for their therapeutic potential and in the food industry as natural preservatives due to their inhibitory effects on spoilage organisms. BGCs related to the potential production of paeninodin were found in both strains of *Pe. phoenicis*. Paeninodin is a cyclic lipopeptide produced by

Paenibacillus species and exhibits significant antimicrobial properties, particularly against Gram-positive bacteria. This compound is noted for its potential in agricultural biocontrol, offering an environmentally friendly alternative to chemical pesticides. Furthermore, surfactant properties of paeninodin make it valuable in industrial applications, such as in the formulation of biosurfactants for bioremediation processes[59].

Using KMAP analysis, several biotechnological applications were predicted in the novel strains. Notably, genes encoding enzymes like polymerases and cellulases, which are relevant for survival in high temperature and alkalinity conditions, were observed. These extremozymes have significant industrial applications due to their stability and efficiency under extreme conditions, making them valuable for processes such as PCR and bioremediation[60,61]. Further exploration of these POIs from extremophilic organisms could enhance current industrial processes by comparing them with the best enzymes available, potentially leading to more efficient and robust biotechnological solutions[62].

Culturing methods may introduce biases, favoring certain microbial types over others[63,64]. However, WGS of novel cultivated species can contribute to metagenome sequence approaches. While comprehensive, technology development is needed for metagenomic analysis to include rare and low-abundant species or those with highly divergent genomes[65]. Future research should focus on further characterizing the functional properties of these novel species, exploring their applications in various industries, and developing improved contamination control strategies.

Material and methods

Samples were taken from the KSC-PHSF at three distinct times: first before the Phoenix spacecraft's arrival on April 25, 2007 (1P), next during the spacecraft's assembly and testing before its launch on June 27, 2007 (2P), and finally after the spacecraft had been moved to the launch pad on August 1, 2007 (3P). Sample collection and isolation of bacterial strains (n=215 strains) cultured under different extreme conditions were already published[9].

DNA extraction and whole-genome sequencing

For WGS, genomic DNA was extracted using the ZymoBIOMICS DNA MagBead kit. The DNA of 215 strains was assessed for the quality, normalized to 50 ng for library preparation, and barcoded with an Oxford Nanopore Technology transposase barcoding kit (SQK-RBK114.96, Oxford Nanopore, Oxford, UK). Finally, each pool of libraries was loaded onto a PromethION flowcell (FLO-PRO114M, R10.4.1) for long-read sequencing.

Genome assembly and relatedness indices

We conducted quality checks of the raw reads using FastQC v.0.12.0[66]. We utilized Unicycler v.0.5.0[67], Flye v.2.9.1[68], and Canu v.2.2[69] on the filtered reads for *de novo* assembly of the genome. To identify the optimal representative assembly from each genome group, genomes within each group were de-replicated using dRep v. 3.4.5[70]. Subsequently, each assembly was assessed for completeness and contamination by CheckM v.1.2.2[71].

To facilitate nucleotide-level comparisons of the genomes within their respective genera, the NCBI command line tool datasets v.15.23.0 was employed to obtain all validly described representative genomes of these 18 genera (https://github.com/ncbi/datasets). We then computed the pairwise ANI computations using FastANI v.1.34 with the novel strains as a query with representative genomes[72]. Furthermore, for estimating dDDH, the Genome-to-Genome Distance Calculator v.3.0 online tool was used with recommended Formula 2 utilizing the BLAST+ alignment tool[73]. In

addition, AAI values were computed using aai.rb function from the Enveomics collection toolbox, and the sequence identity for conserved protein *gyr*B was calculated using Blast v.2.13.0, respectively.

WGS-based phylogeny

For the *Actinomycetota* group (n=11 strains), a set of 138 single-copy genes (SCGs) and *Bacillota* group, 119 SCGs (n=15 strains) were utilized to construct phylogenetic trees at the genus level employing GToTree v.1.8.2[74]. For *Pseudomonadota* group (n=27 strains), a class level phylogenetic tree was generated using 117 SCGs for *Alphaproteobacteria,* 172 SCGs belonging to *Gammaproteobacteria,* and 203 SCGs of *Betaproteobacteria*. An appropriate outgroup was selected for each tree construction.

Subsequently, we employed IQTREE v.2.2.0.3[75] with ModelFinder-Plus[76] to construct the phylogenetic tree from the protein alignment generated by GToTree with 1,000 ultrafast bootstrap replicates. Additionally, we aimed to place the novel strains in the bacterial tree of life and hence retrieved 4,441 complete, non-anomalous representative genomes of bacteria from the NCBI Reference Sequence (RefSeq) database. We constructed a phylogenetic tree using the 16 SCG-set as previously described by Hug et al.[77]. All trees were then annotated and visualized using the interactive Tree Of Life (iTOL) v.6.7[78].

Microscopic characterization

Each bacterial strain was cultured on TSA medium incubated at 26°C for up to 48 hours before proceeding for Gram staining[79]. For SEM imaging analysis, the bacterial samples were loaded on silicon wafers and fixed in 4% glutaraldehyde in 0.1 M phosphate buffer for 2 hours at room temperature, followed by 3 washes of 5 minutes with 0.1 M phosphate buffer. The samples were then dehydrated in ascending isopropanol (IPA) and water series (25%, 30%, 50%, 70%, 80%, 90%, 95%, and 100%) each for 10 minutes, followed by the final 3 times rinsing in 100% IPA and then were critically point dried in EM CPD300 (Leica Company, Wetzlar, Germany). Finally, the silicon wafers carrying the bacterial samples were mounted on SEM stubs (Ted Pella Inc.) using carbon tape and coated with 2 nm of iridium using a sputter coater (Q300T T Plus; Electron Microscopy Sciences Company, Hartfield, PA, USA). The SEM images were collected on Quattro ESEM (ThermoFisher Company, Waltham, MA, USA).

Estimating the abundance of novel species in the cleanroom metagenomes

In order to investigate the presence of newly identified species within controlled cleanroom environments of NASA, we analyzed 164 metagenome samples obtained from Mars 2020 mission assembly cleanrooms: 140 samples from the Spacecraft Assembly Facility (SAF) at the Jet Propulsion Laboratory (JPL), California, and 24 samples from Payload Hazardous Servicing Facility (PHSF) at the Kennedy Space Center (KSC), Florida. Detailed information about the samples can be found in Supplementary Table S2. The samples treated with propidium monoazide (PMA) were considered for this study to capture only viable and intact cells. Initially, the samples were subjected to quality filtering using fastp v.0.22.0 with a phred-score cut-off of 15 and polyG tails trimming with a minimum length of 10[80] to eliminate low-quality reads. Then, we utilized Bowtie2 v.1.2.2 within MetaCompass v.2.0 to align the filtered reads to newly identified genomes and determine their abundance in the NASA cleanrooms based on mapped reads. Following this, we utilized MEGAHIT v.1.0.6 within MetaCompass to assemble the mapped reads and generate consensus sequences[81]. We quantified the percentage of reads aligned to these novel species and assessed the breadth of coverage of the consensus sequences in each sample.

Genome characterization and screening of secondary-metabolite biosynthetic potential

We identified open reading frames (ORFs) in the 53 novel strains using the command-line tool Prokka v.1.14.5, which employs Prodigal for gene annotation based on multiple reference databases[82]. For functional profiling, we utilized the Python-based tool cogclassifier v.1.0.5 (https://pypi.org/project/cogclassifier/) to retrieve Clusters of Orthologous Groups (COGs) from the annotated genomes. To detect antibiotic resistance genes and markers, we used the Resistance Gene Identifier (RGI) v.6.0.3, leveraging the Comprehensive Antibiotic Resistance Database (CARD) v.3.2.6[83]. Only "Perfect" and "Strict" matches were considered to ensure high confidence in the identified antibiotic-resistance genes. All genomes were also annotated using the KAUST Metagenomic Analyses Platform (KMAP)[84], which captures Proteins of Industrial Interest (POIs) based on a comprehensive dictionary of genes relevant to industries such as bioprocess engineering, medicine, pharmaceuticals, cosmetics, and detergents.

Secondary metabolite biosynthetic gene clusters (BGCs) were identified in each novel genome using antiSMASH v.7.0.0[85] with a "Relaxed" detection setting, and the identified BGCs were curated for functional annotation using MIBiG v.3.1[86]. We focused on one particular BGC, ε-Poly-L-lysine, present in three of the isolates with 100% similarity score. The gene neighborhood across this cluster was visualized using Clinker on the CAGECAT web server (https://cagecat.bioinformatics.nl/tools/clinker), comparing it with the known producers *Epichloe festucae* and *Corynebacterium variabile*. Additionally, we aligned the protein sequence of ε-Poly-L-lysine synthetase using the Clustal Omega web server (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) and visualized the conserved regions in different domains using the NCBI Multiple Sequence Alignment Viewer v.1.25.0.

Data availability

The 16S rRNA gene and the draft genome sequences of all the 53 novel strains characterized in this study were deposited in NCBI under BioProject PRJNA1048065. The 16S and WGS accession numbers are given in Table 1, and the genome versions described in this paper are the first versions. The codes used in this study are available at https://github.com/RamanLab/phoenix-novel-species/wiki.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' contribution

KV and NS managed the Phoenix strain collection. ASR and JS performed the Phoenix mission spacecraft strain genome sequencing. JS, ASR, and KV conceived and designed the study, and generated the draft of the manuscript with contributions from all authors. AR and NP conducted microscopy analysis and visualization. TJ and IA performed the genome assembly. PS, SKMS, and TJ conducted WGS-based phylogenetic placement, comparative genomics, genome annotation, and functional characterization with inputs from KR, NKS, and KV. PS performed biosynthetic gene cluster and biofilm-based gene analysis. PS and SKMS performed metagenomic mapping analysis. SK performed biochemical characterization of the novel species. ASR funded the project. JS and KV wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by Prof. Alexandre Soares Rosado's KAUST Baseline Grant (BAS/1/1096-01-01). Part of the research described in this publication was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with National Aeronautics and Space Administration. This research was funded by a 2012 Space Biology NNH12ZTT001N grant no. 19-12829-26 under Task Order NNN13D111T award to KV. PS is supported through the Prime Minister's Research Fellowship from the Ministry of Education, Government of India. The funders had no role in study design, data collection and interpretation, the writing of the manuscript, or the decision to submit the work for publication.

Acknowledgment

We thank Michael Williams, a student intern from CalPoly-Pomona, for reviving and purifying the strains, and Zymo Research Corp. for extracting DNA. We would like to thank the staff from the Bioscience Core Lab, Supercomputing Core Lab, and Imaging and Characterization Core Lab from the King Abdullah University of Science and Technology for their excellent support in conducting the analyses. We thank Abhay Bhat for his support in metagenomics data analysis. P.S. is a recipient of the Prime Minister's Research Fellowship (PMRF) from the Ministry of Education, Government of India. S.K.M.S. acknowledges the Half-Time Teaching Assistantship (HTTA) from the Ministry of Education, Government of India.

References:

1. Lax, S. et al. Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science* **345**, 1048-1052 (2014).
2. Mora, M. et al. Microorganisms in Confined Habitats: Microbial Monitoring and Control of Intensive Care Units, Operating Rooms, Cleanrooms and the International Space Station. *Frontiers in Microbiology* **7**(2016).
3. Miliotis, G. et al. Novel spore-forming species exhibiting intrinsic resistance to third- and fourth-generation cephalosporins and description of *Tigheibacillus jepli* gen. nov., sp. nov. *mBio* **15**, e00181-24 (2024).
4. NASA. Planetary Protection Provisions for Robotic Extraterrestrial Missions. NPR 8020.12D, April 2011. (National Aeronautics and Space Administration, Washington, D.C., 2011).
5. La Duc, M.T., Nicholson, W., Kern, R. & Venkateswaran, K. Microbial characterization of the Mars Odyssey spacecraft and its encapsulation facility. *Environ. Microbiol.* **5**, 977-985 (2003).
6. La Duc, M.T., Kern, R.G. & Venkateswaran, K. Microbial monitoring of spacecraft and associated environments. *Microbial Ecology* **47**, 150-158 (2004).
7. Chander, A.M. et al. Description and Genome Characterization of Three Novel Fungal Strains Isolated from Mars 2020 Mission-Associated Spacecraft Assembly Facility: Recommendations for Two New Genera and One Species. *Journal of Fungi* **9**, 31 (2023).
8. La Duc, M.T. et al. Isolation and Characterization of Bacteria Capable of Tolerating the Extreme Conditions of Clean Room Environments. *Applied and Environmental Microbiology* **73**, 2600-2611 (2007).
9. Ghosh, S., Osman, S., Vaishampayan, P. & Venkateswaran, K. Recurrent isolation of extremotolerant bacteria from the clean room where Phoenix spacecraft components were assembled. *Astrobiology* **10**, 325-35 (2010).
10. Blachowicz, A. et al. The Isolation and Characterization of Rare Mycobiome Associated With Spacecraft Assembly Cleanrooms. *Front Microbiol* **13**, 777133 (2022).
11. Hendrickson, R. et al. Clean room microbiome complexity impacts planetary protection bioburden. *Microbiome* **9**, 238 (2021).
12. Link, L., Sawyer, J., Venkateswaran, K. & Nicholson, W. Extreme spore UV resistance of *Bacillus pumilus* isolates obtained from an ultraclean spacecraft assembly facility. *Microb. Ecol.* **47**, 159-163 (2004).
13. Rummel, J.D. Planetary exploration in the time of astrobiology: protecting against biological contamination. *Proc. Nat. Acad. Sci. USA* **98**, 2128-2131 (2001).
14. Zhao, X. et al. Diversity of culturable heterotrophic bacteria from the Mariana Trench and their ability to degrade macromolecules. *Marine Life Science & Technology* **2**, 181-193 (2020).
15. Sefrji, F.O. et al. Insights Into the Cultivable Bacterial Fraction of Sediments From the Red Sea Mangroves and Physiological, Chemotaxonomic, and Genomic Characterization of *Mangrovibacillus cuniculi* gen. nov., sp. nov., a Novel Member of the Bacillaceae Family. *Front Microbiol* **13**, 777986 (2022).
16. Mahnert, A. et al. Cleanroom Maintenance Significantly Reduces Abundance but Not Diversity of Indoor Microbiomes. *PLOS ONE* **10**, e0134848 (2015).
17. Li, S.J. et al. Microbial communities evolve faster in extreme environments. *Sci Rep* **4**, 6205 (2014).
18. Moissl, C. et al. Molecular bacterial community analysis of clean rooms where spacecraft are assembled. *FEMS Microbiology Ecology* **61**, 509-521 (2007).
19. Moissl-Eichinger, C. et al. Quo vadis? Microbial profiling revealed strong effects of cleanroom maintenance and routes of contamination in indoor environments. *Scientific reports* **5**(2015).

20. Vaishampayan, P., Rabbow, E., Horneck, G. & Venkateswaran, K. Survival of *Bacillus pumilus* spores for a prolonged period of time in real space conditions. *Astrobiology* **12**, 487-497 (2012).
21. Han, X.Y. & Andrade, R.A. *Brevundimonas diminuta* infections and its resistance to fluoroquinolones. *J Antimicrob Chemother* **55**, 853-9 (2005).
22. Corretto, E. et al. Comparative Genomics of *Microbacterium* Species to Reveal Diversity, Potential for Secondary Metabolites and Heavy Metal Resistance. *Frontiers in Microbiology* **11**(2020).
23. Strong, L.C., Rosendahl, C., Johnson, G., Sadowsky, M.J. & Wackett, L.P. *Arthrobacter aurescens* TC1 metabolizes diverse s-triazine ring compounds. *Appl Environ Microbiol* **68**, 5973-80 (2002).
24. da Silva, A.C., Rachid, C.T.C.d.C., de Jesus, H.E., Rosado, A.S. & Peixoto, R.S. Predicting the biotechnological potential of bacteria isolated from Antarctic soils, including the rhizosphere of vascular plants. *Polar Biology* **40**, 1393-1407 (2017).
25. Senchenkov, V.Y. et al. Whole-Genome Sequencing and Biotechnological Potential Assessment of Two Bacterial Strains Isolated from Poultry Farms in Belgorod, Russia. *Microorganisms* **11**(2023).
26. Patel, S. & Gupta, R.S. A phylogenomic and comparative genomic framework for resolving the polyphyly of the genus *Bacillus*: Proposal for six new genera of *Bacillus* species, *Peribacillus* gen. nov., *Cytobacillus* gen. nov., *Mesobacillus* gen. nov., *Neobacillus* gen. nov., *Metabacillus* gen. nov. and *Alkalihalobacillus* gen. nov. *Int J Syst Evol Microbiol* **70**, 406-438 (2020).
27. Kim, K.H., Han, D.M., Lee, J.K. & Jeon, C.O. *Alkalicoccobacillus porphyridii* sp. nov., isolated from a marine red alga, reclassification of *Shouchella plakortidis* and *Shouchella gibsonii* as *Alkalicoccobacillus plakortidis* comb. nov. and *Alkalicoccobacillus gibsonii* comb. nov., and emended description of the genus *Alkalicoccobacillus* Joshi et al. 2022. *International Journal of Systematic and Evolutionary Microbiology* **73**(2023).
28. Paulino-Lima, I.G. et al. Isolation of UVC-tolerant bacteria from the hyperarid Atacama Desert, Chile. *Microb Ecol* **65**, 325-35 (2013).
29. Gaete, A., Mandakovic, D. & González, M. Isolation and Identification of Soil Bacteria from Extreme Environments of Chile and Their Plant Beneficial Characteristics. *Microorganisms* **8**, 1213 (2020).
30. Aftab, H. & Donegan, R.K. Regulation of heme biosynthesis via the coproporphyrin dependent pathway in bacteria. *Frontiers in Microbiology* **15**(2024).
31. Blasius, M., Sommer, S. & Hübscher, U. *Deinococcus radiodurans*: what belongs to the survival kit? *Crit Rev Biochem Mol Biol* **43**, 221-38 (2008).
32. Frenkiel-Krispin, D. et al. Structure of the DNA-SspC complex: implications for DNA packaging, protection, and repair in bacterial spores. *J Bacteriol* **186**, 3525-30 (2004).
33. Liu, Y. et al. Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. *Proc Natl Acad Sci U S A* **100**, 4191-6 (2003).
34. Eisen, J.A. & Hanawalt, P.C. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat Res* **435**, 171-213 (1999).
35. Cortesao, M. et al. *Bacillus subtilis* Spore Resistance to Simulated Mars Surface Conditions. *Front Microbiol* **10**, 333 (2019).
36. Simpson, A.C. et al. Phylogenomics, phenotypic, and functional traits of five novel (Earth-derived) bacterial species isolated from the International Space Station and their prevalence in metagenomes. *Scientific Reports* **13**, 19207 (2023).
37. Lam, K.N., Cheng, J., Engel, K., Neufeld, J.D. & Charles, T.C. Current and future resources for functional metagenomics. *Front Microbiol* **6**, 1196 (2015).

38. Bikel, S. et al. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Computational and Structural Biotechnology Journal* **13**, 390-401 (2015).
39. Sagot, B. et al. Osmotically induced synthesis of the dipeptide N-acetylglutaminylglutamine amide is mediated by a new pathway conserved among bacteria. *Proc Natl Acad Sci U S A* **107**, 12652-7 (2010).
40. Coker, J.A. Recent advances in understanding extremophiles. *F1000Res* **8**(2019).
41. Brock, T.D. The value of basic research: discovery of Thermus aquaticus and other extreme thermophiles. *Genetics* **146**, 1207-10 (1997).
42. Uruén, C., Chopo-Escuin, G., Tommassen, J., Mainar-Jaime, R.C. & Arenas, J. Biofilms as Promoters of Bacterial Antibiotic Resistance and Tolerance. *Antibiotics (Basel)* **10**(2020).
43. Costerton, J.W., Stewart, P.S. & Greenberg, E.P. Bacterial biofilms: a common cause of persistent infections. *Science* **284**, 1318-22 (1999).
44. Donlan, R.M. & Costerton, J.W. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15**, 167-93 (2002).
45. da Silva, A.A., Galego, L. & Arraiano, C.M. New Perspectives on BolA: A Still Mysterious Protein Connecting Morphogenesis, Biofilm Production, Virulence, Iron Metabolism, and Stress Survival. *Microorganisms* **11**, 632 (2023).
46. Vieira, H.L., Freire, P. & Arraiano, C.M. Effect of *Escherichia coli* morphogene bolA on biofilms. *Appl Environ Microbiol* **70**, 5682-4 (2004).
47. Zhang, J. et al. Impact of disinfectant on bacterial antibiotic resistance transfer between biofilm and tap water in a simulated distribution network. *Environ Pollut* **246**, 131-140 (2019).
48. Kvist, M., Hancock, V. & Klemm, P. Inactivation of efflux pumps abolishes bacterial biofilm formation. *Appl Environ Microbiol* **74**, 7376-82 (2008).
49. Carabetta, V.J. et al. A complex of YlbF, YmcA and YaaT regulates sporulation, competence and biofilm formation by accelerating the phosphorylation of Spo0A. *Molecular Microbiology* **88**, 283-300 (2013).
50. Tang, S.-K. et al. Cellular differentiation into hyphae and spores in halophilic archaea. *Nature Communications* **14**, 1827 (2023).
51. Branda, S.S. et al. Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. *J Bacteriol* **186**, 3970-9 (2004).
52. Kearns, D.B., Chu, F., Branda, S.S., Kolter, R. & Losick, R. A master regulator for biofilm formation by *Bacillus subtilis*. *Mol. Microbiol.* **55**, 739-749 (2005).
53. Chen, Y. et al. The structure, properties, synthesis method and antimicrobial mechanism of ε-polylysine with the preservative effects for aquatic products. *Trends in Food Science & Technology* **139**, 104131 (2023).
54. Breithaupt, D.E. Modern application of xanthophylls in animal feeding – a review. *Trends in Food Science & Technology* **18**, 501-506 (2007).
55. Saha, R., Saha, N., Donofrio, R.S. & Bestervelt, L.L. Microbial siderophores: a mini review. *J Basic Microbiol* **53**, 303-17 (2013).
56. Kloepper, J.W., Ryu, C.M. & Zhang, S. Induced Systemic Resistance and Promotion of Plant Growth by *Bacillus* spp. *Phytopathology* **94**, 1259-66 (2004).
57. Stafsnes, M.H. et al. Isolation and characterization of marine pigmented bacteria from Norwegian coastal waters and screening for carotenoids with UVA-blue light absorbing properties. *J Microbiol* **48**, 16-23 (2010).
58. Zabolotneva, A.A., Shatova, O.P., Sadova, A.A., Shestopalov, A.V. & Roumiantsev, S.A. An Overview of Alkylresorcinols Biological Properties and Effects. *J Nutr Metab* **2022**, 4667607 (2022).

59. Ongena, M. & Jacques, P. *Bacillus* lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol* **16**, 115-25 (2008).
60. Hough, D.W. & Danson, M.J. Extremozymes. *Curr Opin Chem Biol* **3**, 39-46 (1999).
61. Elleuche, S., Schröder, C., Sahm, K. & Antranikian, G. Extremozymes--biocatalysts with unique properties from extremophilic microorganisms. *Curr Opin Biotechnol* **29**, 116-23 (2014).
62. Burkhardt, C. et al. Mining thermophiles for biotechnologically relevant enzymes: evaluating the potential of European and Caucasian hot springs. *Extremophiles* **28**, 5 (2023).
63. Schultz, J., Modolon, F., Peixoto, R.S. & Rosado, A.S. Shedding light on the composition of extreme microbial dark matter: alternative approaches for culturing extremophiles. *Frontiers in Microbiology* **14**(2023).
64. Schultz, J., dos Santos, A., Patel, N. & Rosado, A.S. Life on the Edge: Bioprospecting Extremophiles for Astrobiology. *Journal of the Indian Institute of Science* **103**, 721-737 (2023).
65. McLaren, M.R., Willis, A.D. & Callahan, B.J. Consistent and correctable bias in metagenomic sequencing experiments. *eLife* **8**, e46923 (2019).
66. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. (2010).
67. Wick, R.R., Judd, L.M., Gorrie, C.L. & Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology* **13**, e1005595 (2017).
68. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* **37**, 540-546 (2019).
69. Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* **27**, 722-736 (2017).
70. Olm, M.R., Brown, C.T., Brooks, B. & Banfield, J.F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *Isme j* **11**, 2864-2868 (2017).
71. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P. & Tyson, G.W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, 1043-55 (2015).
72. Jain, C., Rodriguez, R.L., Phillippy, A.M., Konstantinidis, K.T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* **9**, 5114 (2018).
73. Meier-Kolthoff, J.P., Carbasse, J.S., Peinado-Olarte, R.L. & Göker, M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* **50**, D801-d807 (2022).
74. Lee, M.D. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* **35**, 4162-4164 (2019).
75. Minh, B.Q. et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530-1534 (2020).
76. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermiin, L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587-589 (2017).
77. Hug, L.A. et al. A new view of the tree of life. *Nature Microbiology* **1**, 16048 (2016).
78. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**, W293-w296 (2021).
79. Claus, D. A standardized Gram staining procedure. *World Journal of Microbiology and Biotechnology* **8**, 451-452 (1992).

80. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).
81. Luan, T. et al. MetaCompass: Reference-guided Assembly of Metagenomes. *arXiv preprint arXiv:2403.01578* (2024).
82. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-9 (2014).
83. Alcock, B.P. et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research* **51**, D690-D699 (2022).
84. Alam, I. et al. KAUST Metagenomic Analysis Platform (KMAP), enabling access to massive analytics of re-annotated metagenomic data. *Sci Rep* **11**, 11511 (2021).
85. Blin, K. et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Research* (2023).
86. Terlouw, B.R. et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res* **51**, D603-d610 (2023).
87. Pal, S. et al. Unraveling radiation resistance strategies in two bacterial strains from the high background radiation area of Chavara-Neendakara: A comprehensive whole genome analysis. *PLOS ONE* **19**, e0304810 (2024).

Figure Legends

Figure 1. Phylogenetic tree of non-spore-forming bacteria (*n*=38) spanning into nine genera from Phoenix spacecraft mission. Novel species are highlighted in red, and their corresponding NCBI accessions are provided. Bootstrap values (expressed as percentages) are indicated near the branches.

Figure 2. Phylogenetic tree of 15 novel strains of spore-forming bacteria from *Bacillaceae* family isolated from Phoenix spacecraft mission. Novel species are highlighted in red, and their corresponding NCBI accessions are provided. Bootstrap values (expressed as percentages) are indicated near the branches.

Figure 3. Scanning electron microscopy of the novel species isolated from the Phoenix spacecraft assembly cleanroom.

Figure 4. Metagenomic read mapping to novel isolates from NASA cleanrooms, highlighting temporal and spatial dynamics. A. Spatial distribution of mapped reads across 26 novel species, showing distinct signatures between spore-forming and non-spore-forming bacteria in different NASA cleanroom locations. B. Box plots illustrating the breadth of coverage (>1%) of consensus genomes constructed from mapped reads aligned to 23 novel species (out of 26). Reads were collected from cleanrooms at SAF JPL and PHSF KSC in 2016 (red) and PHSF in 2018 (blue).

Figure 5. Functional insights into novel species from NASA cleanrooms. A. Presence of radiation resistance COGs (from Pal et al.[87]) in the 26 novel species, revealing their genetic potential for radiation resilience. B. Identification of Proteins of Interest (PoI) using KMAP with applications in various industries, leveraging the unique functional capabilities of novel species from NASA cleanrooms.

Figure 6. Comparative analysis of ε-Poly-L-lysine synthetase in novel species. A. ε-Poly-L-lysine gene cluster comparison in *Epichloe festucae* (fungal producer), *Corynebacterium variabile* (bacterial producer), and three novel species from our study (*Agrococcus phoenicis*, *Microbacterium canaveralium*, *Microbacterium jepli*) and *Leifsonia virtsii* (isolated from ISS) show conserved gene cluster architecture. B. Protein sequence alignment of ε-Poly-L-lysine synthetase enzymes from these organisms exhibits conserved domains, including NRPS adenylation (A), thiolation (T), transmembrane (TM), and C-terminal tandem domains (C1, C2, C3).

Figure 1

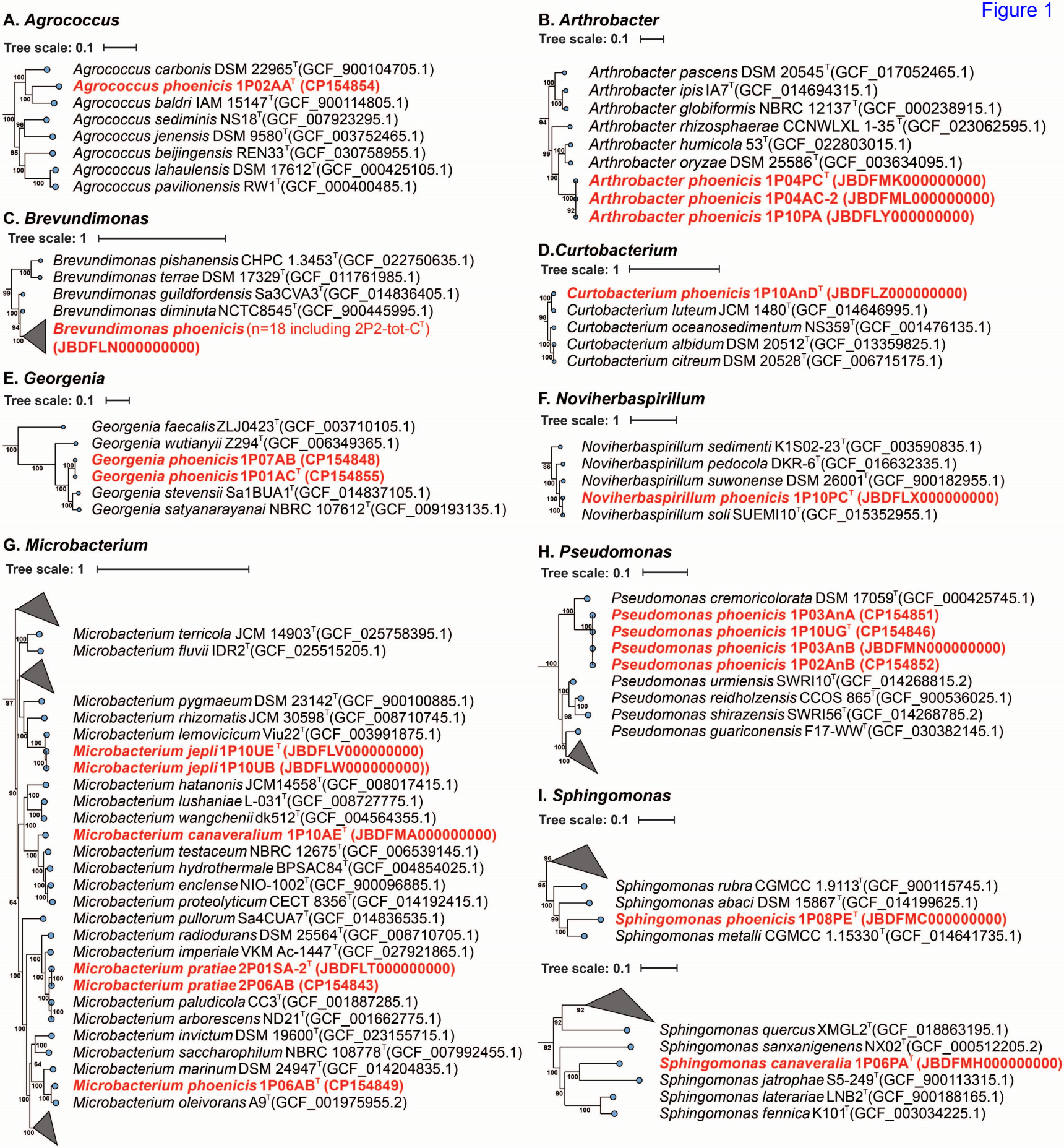
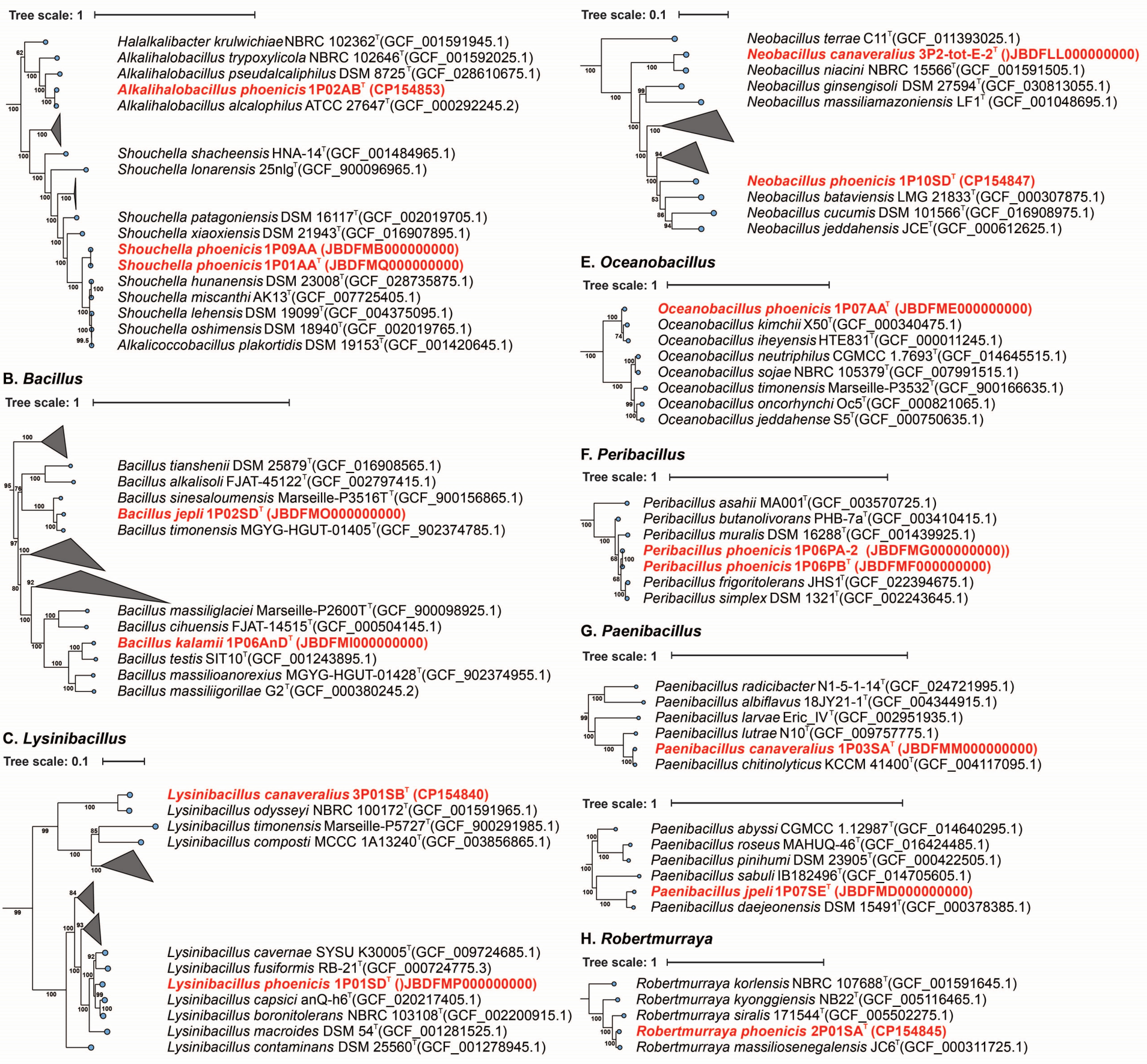
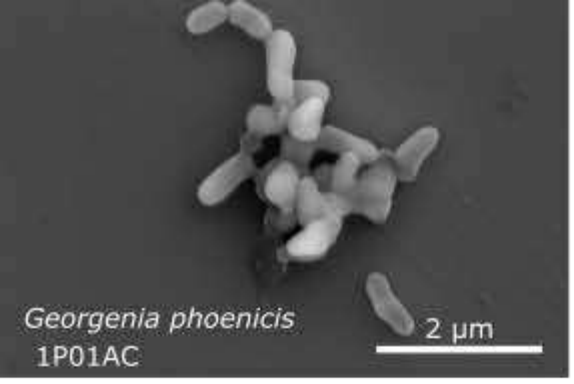
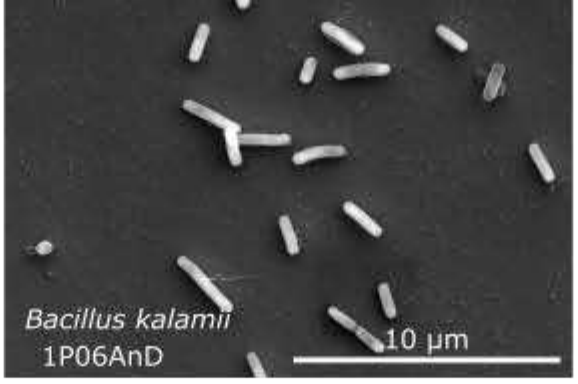
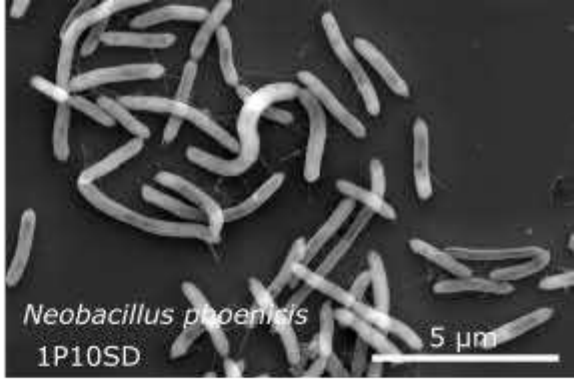

A. Agrococcus
Tree scale: 0.1
Agrococcus carbonis DSM 22965T(GCF_900104705.1)
Agrococcus phoenicis 1P02AAT (CP154854)
Agrococcus baldri IAM 15147T(GCF_900114805.1)
Agrococcus sediminis NS18T(GCF_007923295.1)
Agrococcus jenensis DSM 9580T(GCF_003752465.1)
Agrococcus beijingensis REN33T(GCF_030758955.1)
Agrococcus lahaulensis DSM 17612T(GCF_000425105.1)
Agrococcus pavilionensis RW1T(GCF_000400485.1)
B. Arthrobacter
Tree scale: 0.1
Arthrobacter pascens DSM 20545T(GCF_017052465.1)
Arthrobacter ipis IA7T(GCF_014694315.1)
Arthrobacter globiformis NBRC 12137T(GCF_000238915.1)
Arthrobacter rhizosphaerae CCNWLXL 1-35T(GCF_023062595.1)
Arthrobacter humicola 53T(GCF_022803015.1)
Arthrobacter oryzae DSM 25586T(GCF_003634095.1)
Arthrobacter phoenicis 1P04PCT (JBDFMK000000000)
Arthrobacter phoenicis 1P04AC-2 (JBDFML000000000)
Arthrobacter phoenicis 1P10PA (JBDFLY000000000)
C. Brevundimonas
Tree scale: 1
Brevundimonas pishanensis CHPC 1.3453T(GCF_022750635.1)
Brevundimonas terrae DSM 17329T(GCF_011761985.1)
Brevundimonas guildfordensis Sa3CVA3T(GCF_014836405.1)
Brevundimonas diminuta NCTC8545T(GCF_900445995.1)
Brevundimonas phoenicis (n=18 including 2P2-tot-CT)
(JBDFLN000000000)
D.Curtobacterium
Tree scale: 1
Curtobacterium phoenicis 1P10AnDT (JBDFLZ000000000)
Curtobacterium luteum JCM 1480T(GCF_014646995.1)
Curtobacterium oceanosedimentum NS359T(GCF_001476135.1)
Curtobacterium albidum DSM 20512T(GCF_013359825.1)
Curtobacterium citreum DSM 20528T(GCF_006715175.1)
E. Georgenia
Tree scale: 0.1
Georgenia faecalis ZLJ0423T(GCF_003710105.1)
Georgenia wutianyii Z294T(GCF_006349365.1)
Georgenia phoenicis 1P07AB (CP154848)
Georgenia phoenicis 1P01ACT (CP154855)
Georgenia stevensii Sa1BUA1T(GCF_014837105.1)
Georgenia satyanarayanai NBRC 107612T(GCF_009193135.1)
F. Noviherbaspirillum
Tree scale: 1
Noviherbaspirillum sedimenti K1S02-23T(GCF_003590835.1)
Noviherbaspirillum pedocola DKR-6T(GCF_016632335.1)
Noviherbaspirillum suwonense DSM 26001T(GCF_900182955.1)
Noviherbaspirillum phoenicis 1P10PCT (JBDFLX000000000)
Noviherbaspirillum soli SUEMI10T(GCF_015352955.1)
G. Microbacterium
Tree scale: 1
Microbacterium terricola JCM 14903T(GCF_025758395.1)
Microbacterium fluvii IDR2T(GCF_025515205.1)
Microbacterium pygmaeum DSM 23142T(GCF_900100885.1)
Microbacterium rhizomatis JCM 30598T(GCF_008710745.1)
Microbacterium lemovicicum Viu22T(GCF_003991875.1)
Microbacterium jepli 1P10UET (JBDFLV000000000)
Microbacterium jepli 1P10UB (JBDFLW000000000))
Microbacterium hatanonis JCM14558T(GCF_008017415.1)
Microbacterium lushaniae L-031T(GCF_008727775.1)
Microbacterium wangchenii dk512T(GCF_004564355.1)
Microbacterium canaveralium 1P10AET (JBDFMA000000000)
Microbacterium testaceum NBRC 12675T(GCF_006539145.1)
Microbacterium hydrothermale BPSAC84T(GCF_004854025.1)
Microbacterium enclense NIO-1002T(GCF_900096885.1)
Microbacterium proteolyticum CECT 8356T(GCF_014192415.1)
Microbacterium pullorum Sa4CUA7T(GCF_014836535.1)
Microbacterium radiodurans DSM 25564T(GCF_008710705.1)
Microbacterium imperiale VKM Ac-1447T(GCF_027921865.1)
Microbacterium pratiae 2P01SA-2T (JBDFLT000000000)
Microbacterium pratiae 2P06AB (CP154843)
Microbacterium paludicola CC3T(GCF_001887285.1)
Microbacterium arborescens ND21T(GCF_001662775.1)
Microbacterium invictum DSM 19600T(GCF_023155715.1)
Microbacterium saccharophilum NBRC 108778T(GCF_007992455.1)
Microbacterium marinum DSM 24947T(GCF_014204835.1)
Microbacterium phoenicis 1P06ABT (CP154849)
Microbacterium oleivorans A9T(GCF_001975955.2)
H. Pseudomonas
Tree scale: 0.1
Pseudomonas cremoricolorata DSM 17059T(GCF_000425745.1)
Pseudomonas phoenicis 1P03AnA (CP154851)
Pseudomonas phoenicis 1P10UGT (CP154846)
Pseudomonas phoenicis 1P03AnB (JBDFMN000000000)
Pseudomonas phoenicis 1P02AnB (CP154852)
Pseudomonas urmiensis SWRI10T(GCF_014268815.2)
Pseudomonas reidholzensis CCOS 865T(GCF_900536025.1)
Pseudomonas shirazensis SWRI56T(GCF_014268785.2)
Pseudomonas guariconensis F17-WWT(GCF_030382145.1)
I. Sphingomonas
Tree scale: 0.1
Sphingomonas rubra CGMCC 1.9113T(GCF_900115745.1)
Sphingomonas abaci DSM 15867T(GCF_014199625.1)
Sphingomonas phoenicis 1P08PET (JBDFMC000000000)
Sphingomonas metalli CGMCC 1.15330T(GCF_014641735.1)
Tree scale: 0.1
Sphingomonas quercus XMGL2T(GCF_018863195.1)
Sphingomonas sanxanigenens NX02T(GCF_000512205.2)
Sphingomonas canaveralia 1P06PAT (JBDFMH000000000)
Sphingomonas jatrophae S5-249T(GCF_900113315.1)
Sphingomonas laterariae LNB2T(GCF_900188165.1)
Sphingomonas fennica K101T(GCF_003034225.1)

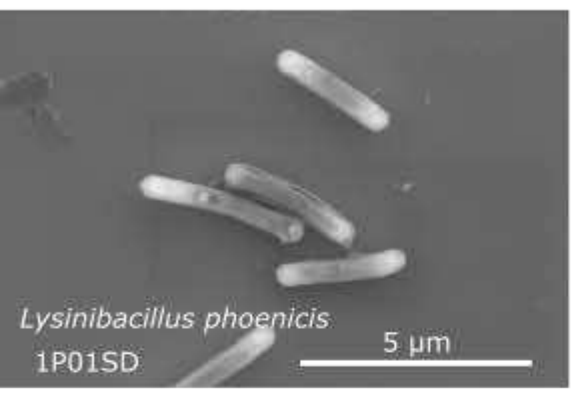
Figure 2

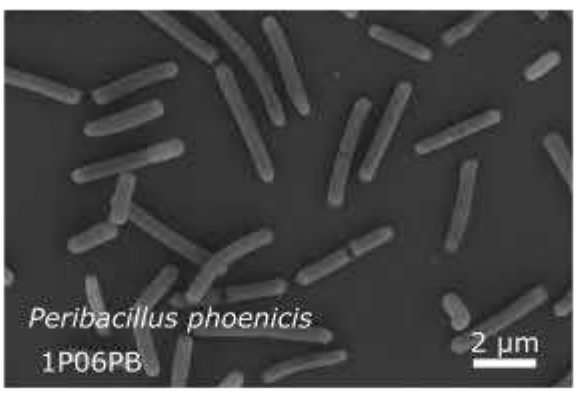

A. Alkalihalobacillus/ Shouchella
Tree scale: 1
Halalkalibacter krulwichiae NBRC 102362T(GCF_001591945.1)
Alkalihalobacillus trypoxylicola NBRC 102646T(GCF_001592025.1)
Alkalihalobacillus pseudalcaliphilus DSM 8725T(GCF_028610675.1)
Alkalihalobacillus phoenicis 1P02ABT (CP154853)
Alkalihalobacillus alcalophilus ATCC 27647T(GCF_000292245.2)
Shouchella shacheensis HNA-14T(GCF_001484965.1)
Shouchella lonarensis 25nlgT(GCF_900096965.1)
Shouchella patagoniensis DSM 16117T(GCF_002019705.1)
Shouchella xiaoxiensis DSM 21943T(GCF_016907895.1)
Shouchella phoenicis 1P09AA (JBDFMB000000000)
Shouchella phoenicis 1P01AAT (JBDFMQ000000000)
Shouchella hunanensis DSM 23008T(GCF_028735875.1)
Shouchella miscanthi AK13T(GCF_007725405.1)
Shouchella lehensis DSM 19099T(GCF_004375095.1)
Shouchella oshimensis DSM 18940T(GCF_002019765.1)
Alkalicoccobacillus plakortidis DSM 19153T(GCF_001420645.1)
B. Bacillus
Tree scale: 1
Bacillus tianshenii DSM 25879T(GCF_016908565.1)
Bacillus alkalisoli FJAT-45122T(GCF_002797415.1)
Bacillus sinesaloumensis Marseille-P3516TT(GCF_900156865.1)
Bacillus jepli 1P02SDT (JBDFMO000000000)
Bacillus timonensis MGYG-HGUT-01405T(GCF_902374785.1)
Bacillus massiliglaciei Marseille-P2600TT(GCF_900098925.1)
Bacillus cihuensis FJAT-14515T(GCF_000504145.1)
Bacillus kalamii 1P06AnDT (JBDFMI000000000)
Bacillus testis SIT10T(GCF_001243895.1)
Bacillus massilioanorexius MGYG-HGUT-01428T(GCF_902374955.1)
Bacillus massiliigorillae G2T(GCF_000380245.2)
C. Lysinibacillus
Tree scale: 0.1
Lysinibacillus canaveralius 3P01SBT (CP154840)
Lysinibacillus odysseyi NBRC 100172T(GCF_001591965.1)
Lysinibacillus timonensis Marseille-P5727T(GCF_900291985.1)
Lysinibacillus composti MCCC 1A13240T(GCF_003856865.1)
Lysinibacillus cavernae SYSU K30005T(GCF_009724685.1)
Lysinibacillus fusiformis RB-21T(GCF_000724775.3)
Lysinibacillus phoenicis 1P01SDT ()JBDFMP000000000)
Lysinibacillus capsici anQ-h6T(GCF_020217405.1)
Lysinibacillus boronitolerans NBRC 103108T(GCF_002200915.1)
Lysinibacillus macroides DSM 54T(GCF_001281525.1)
Lysinibacillus contaminans DSM 25560T(GCF_001278945.1)
D. Neobacillus
Tree scale: 0.1
Neobacillus terrae C11T(GCF_011393025.1)
Neobacillus canaveralius 3P2-tot-E-2T ()JBDFLL000000000)
Neobacillus niacini NBRC 15566T(GCF_001591505.1)
Neobacillus ginsengisoli DSM 27594T(GCF_030813055.1)
Neobacillus massiliamazoniensis LF1T (GCF_001048695.1)
Neobacillus phoenicis 1P10SDT (CP154847)
Neobacillus bataviensis LMG 21833T(GCF_000307875.1)
Neobacillus cucumis DSM 101566T(GCF_016908975.1)
Neobacillus jeddahensis JCET(GCF_000612625.1)
E. Oceanobacillus
Tree scale: 1
Oceanobacillus phoenicis 1P07AAT (JBDFME000000000)
Oceanobacillus kimchii X50T(GCF_000340475.1)
Oceanobacillus iheyensis HTE831T(GCF_000011245.1)
Oceanobacillus neutriphilus CGMCC 1.7693T(GCF_014645515.1)
Oceanobacillus sojae NBRC 105379T(GCF_007991515.1)
Oceanobacillus timonensis Marseille-P3532T(GCF_900166635.1)
Oceanobacillus oncorhynchi Oc5T(GCF_000821065.1)
Oceanobacillus jeddahense S5T(GCF_000750635.1)
F. Peribacillus
Tree scale: 1
Peribacillus asahii MA001T(GCF_003570725.1)
Peribacillus butanolivorans PHB-7aT(GCF_003410415.1)
Peribacillus muralis DSM 16288T(GCF_001439925.1)
Peribacillus phoenicis 1P06PA-2 (JBDFMG000000000))
Peribacillus phoenicis 1P06PBT (JBDFMF000000000)
Peribacillus frigoritolerans JHS1T(GCF_022394675.1)
Peribacillus simplex DSM 1321T(GCF_002243645.1)
G. Paenibacillus
Tree scale: 1
Paenibacillus radicibacter N1-5-1-14T(GCF_024721995.1)
Paenibacillus albiflavus 18JY21-1T(GCF_004344915.1)
Paenibacillus larvae Eric_IVT(GCF_002951935.1)
Paenibacillus lutrae N10T(GCF_009757775.1)
Paenibacillus canaveralius 1P03SAT (JBDFMM000000000)
Paenibacillus chitinolyticus KCCM 41400T(GCF_004117095.1)
Tree scale: 1
Paenibacillus abyssi CGMCC 1.12987T(GCF_014640295.1)
Paenibacillus roseus MAHUQ-46T(GCF_016424485.1)
Paenibacillus pinihumi DSM 23905T(GCF_000422505.1)
Paenibacillus sabuli IB182496T(GCF_014705605.1)
Paenibacillus jpeli 1P07SET (JBDFMD000000000)
Paenibacillus daejeonensis DSM 15491T(GCF_000378385.1)
H. Robertmurraya
Tree scale: 1
Robertmurraya korlensis NBRC 107688T(GCF_001591645.1)
Robertmurraya kyonggiensis NB22T(GCF_005116465.1)
Robertmurraya siralis 171544T(GCF_005502275.1)
Robertmurraya phoenicis 2P01SAT (CP154845)
Robertmurraya massiliosenegalensis JC6T(GCF_000311725.1)

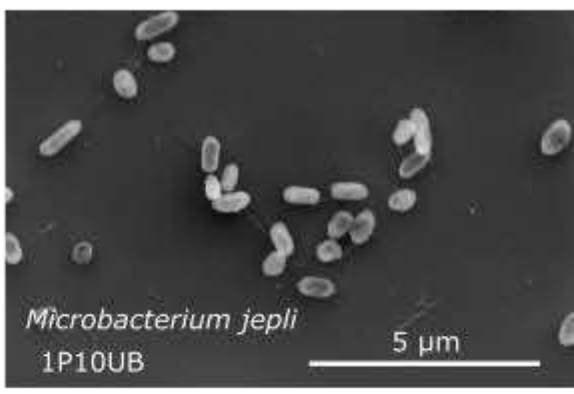

Georgenia phoenicis
1P01AC
2 µm

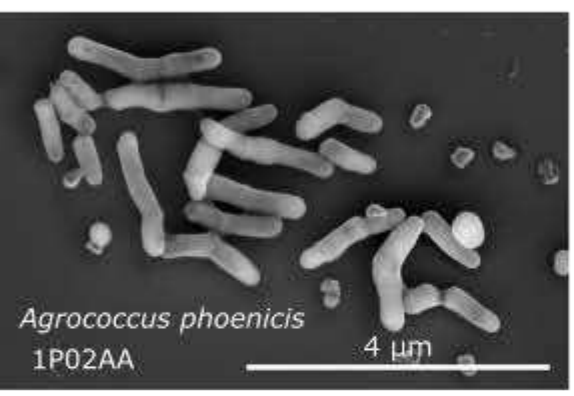

Bacillus kalamii
1P06AnD
10 µm

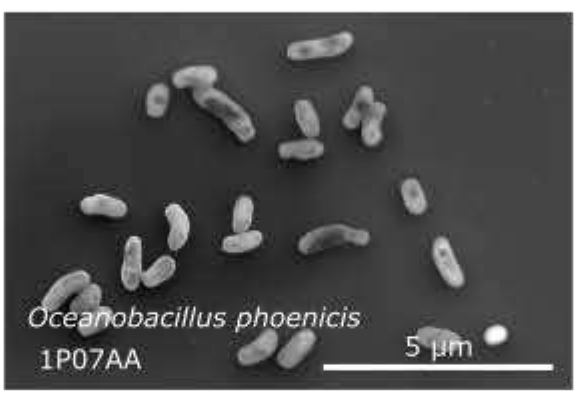

Neobacillus phoenicis
1P10SD
5 µm

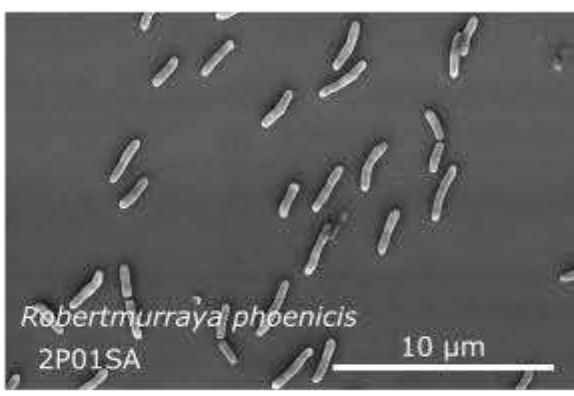

Lysinibacillus phoenicis
1P01SD
5 µm

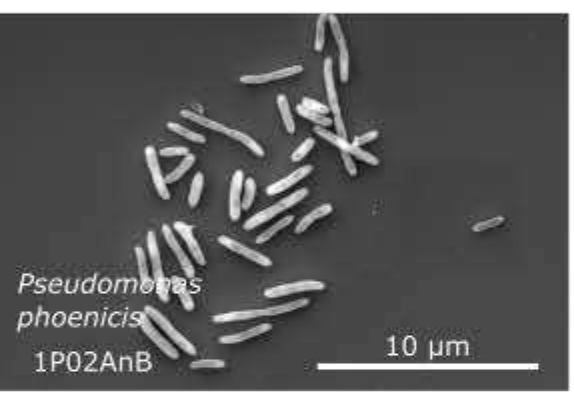

Peribacillus phoenicis
1P06PB
2 µm

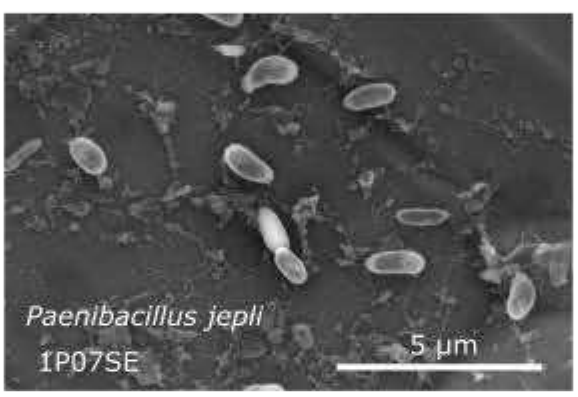

Microbacterium jepli
1P10UB
5 µm

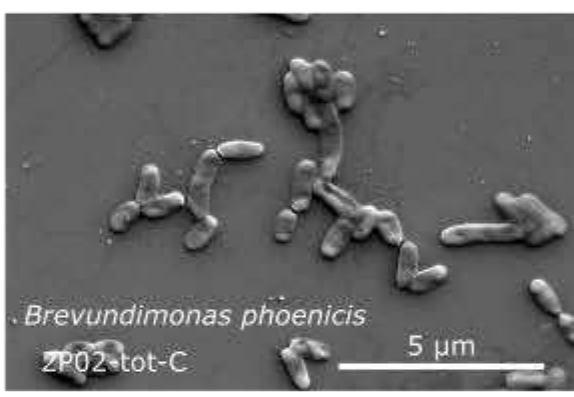

Agrococcus phoenicis
1P02AA
4 µm

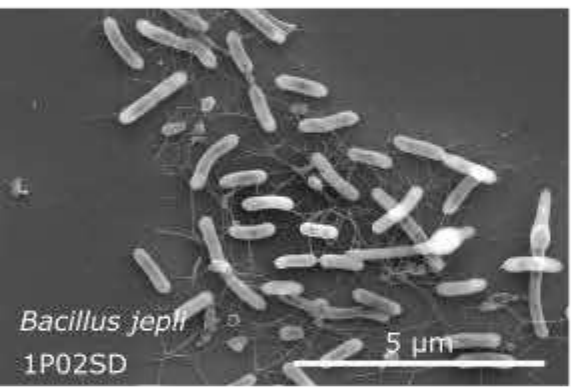

Oceanobacillus phoenicis
1P07AA
5 µm

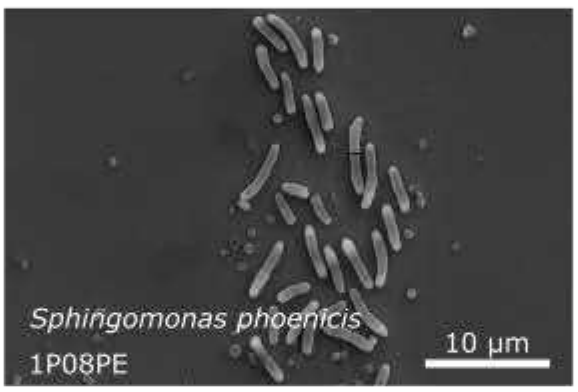

Robertmurraya phoenicis
2P01SA
10 µm

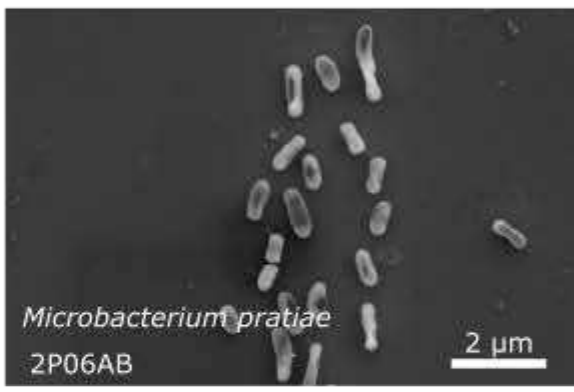

Pseudomonas phoenicis
1P02AnB
10 µm

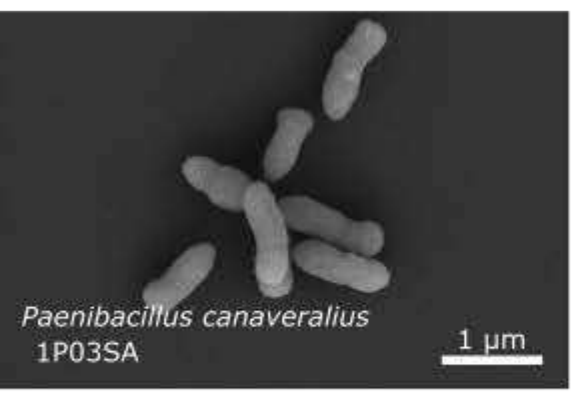

Paenibacillus jepli
1P07SE
5 µm

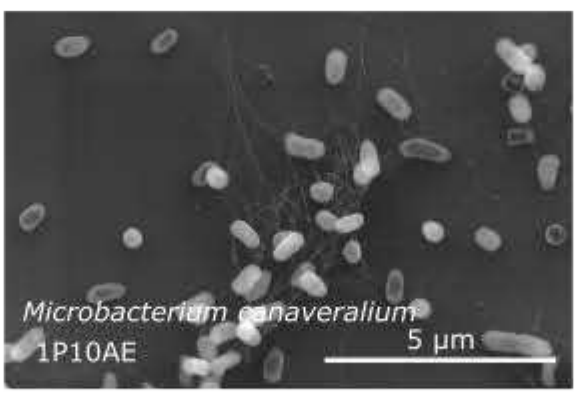

Brevundimonas phoenicis
2P02-tot-C
5 µm

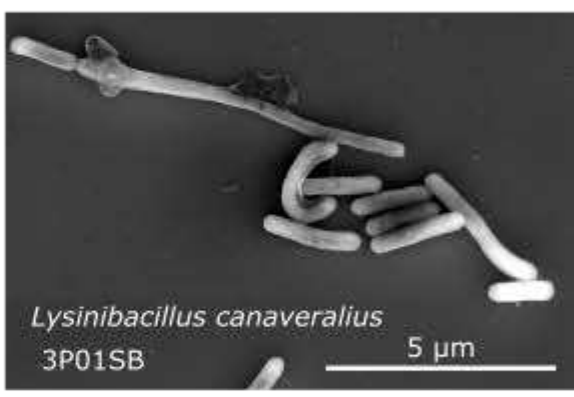

Bacillus jepli
1P02SD
5 µm

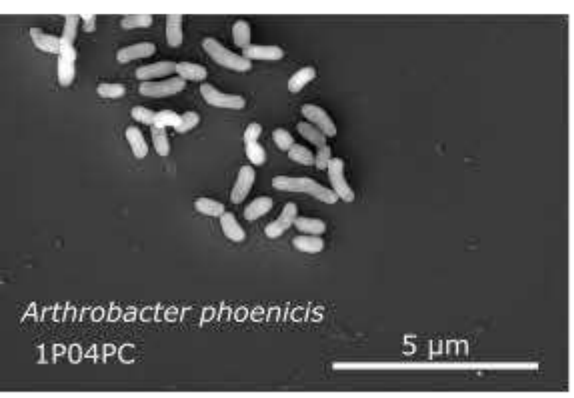

Sphingomonas phoenicis
1P08PE
10 µm

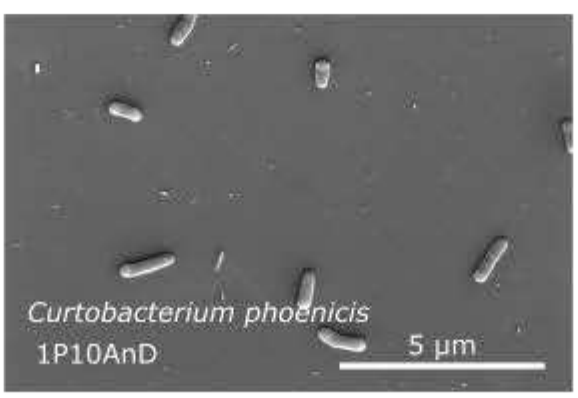

Microbacterium pratiae
2P06AB
2 µm

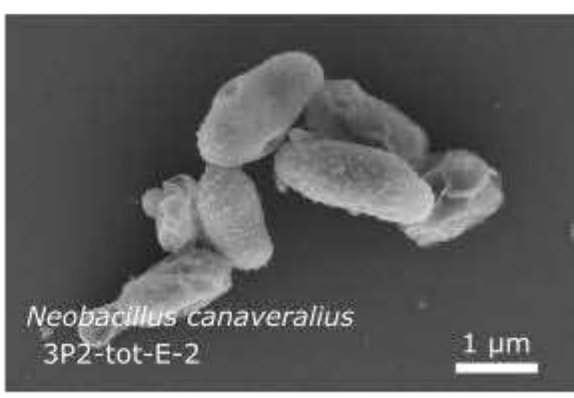

Paenibacillus canaveralius
1P03SA
1 µm


Microbacterium canaveralium
1P10AE
5 µm


Lysinibacillus canaveralius
3P01SB
5 µm


Arthrobacter phoenicis
1P04PC
5 µm


Curtobacterium phoenicis
1P10AnD
5 µm


Neobacillus canaveralius
3P2-tot-E-2
1 µm


Microbacterium phoenicis
1P06AB
2 µm

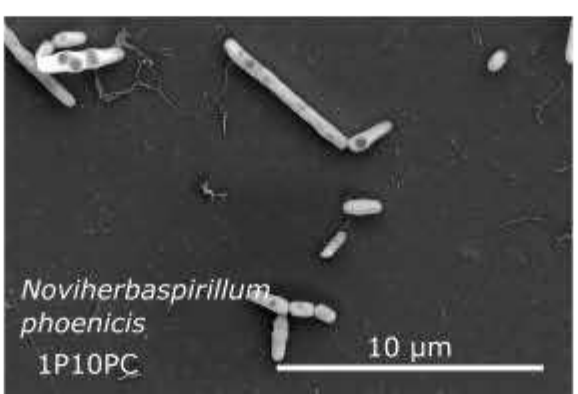
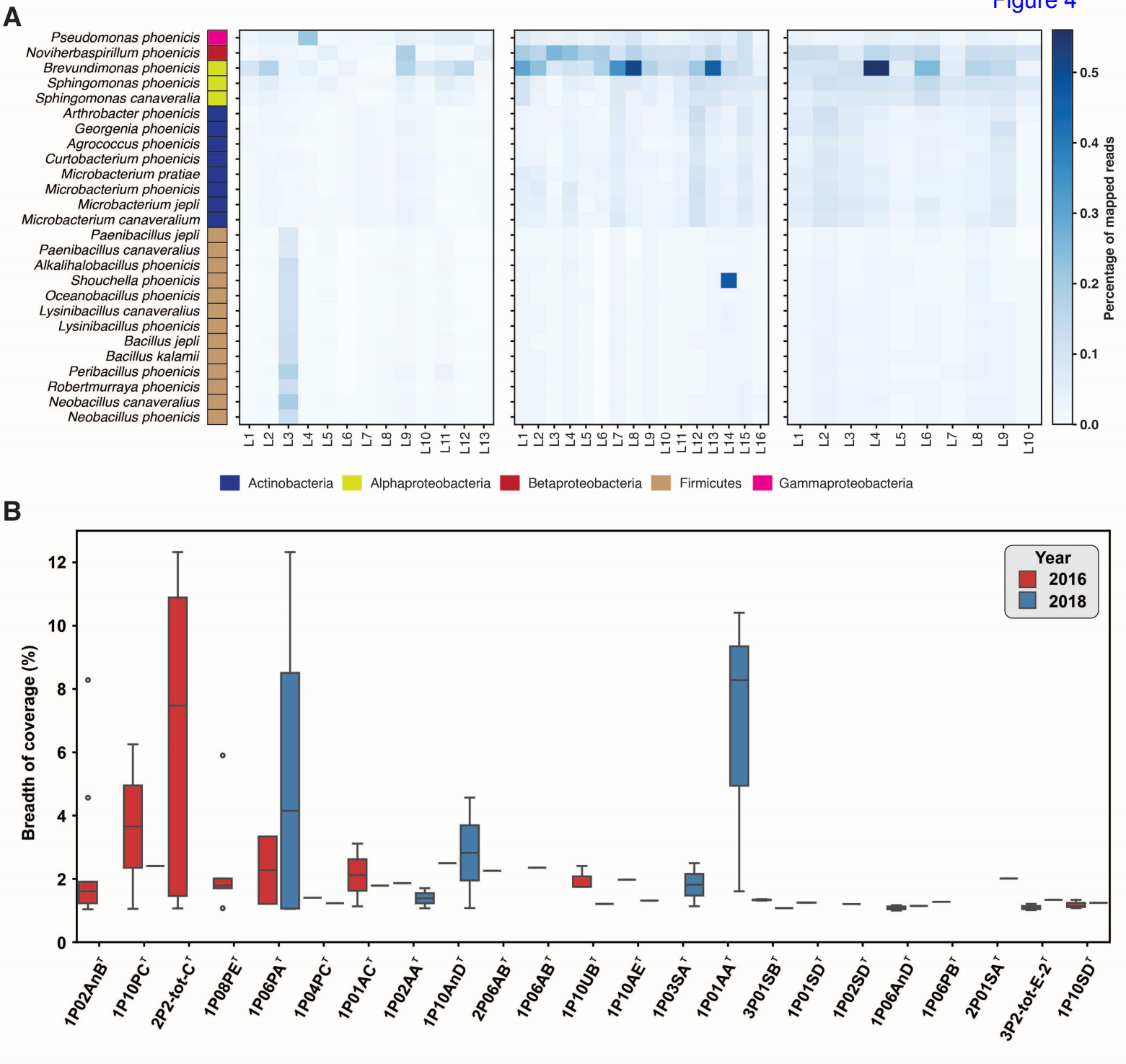

Noviherbaspirillum phoenicis
1P10PC
10 µm

Figure 3

Figure 4


A
Pseudomonas phoenicis
Noviherbaspirillum phoenicis
Brevundimonas phoenicis
Sphingomonas phoenicis
Sphingomonas canaveralia
Arthrobacter phoenicis
Georgenia phoenicis
Agrococcus phoenicis
Curtobacterium phoenicis
Microbacterium pratiae
Microbacterium phoenicis
Microbacterium jepli
Microbacterium canaveralium
Paenibacillus jepli
Paenibacillus canaveralius
Alkalihalobacillus phoenicis
Shouchella phoenicis
Oceanobacillus phoenicis
Lysinibacillus canaveralius
Lysinibacillus phoenicis
Bacillus jepli
Bacillus kalamii
Peribacillus phoenicis
Robertmurraya phoenicis
Neobacillus canaveralius
Neobacillus phoenicis
L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13
L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 L14 L15 L16
L1 L2 L3 L4 L5 L6 L7 L8 L9 L10
Percentage of mapped reads
0.0 0.1 0.2 0.3 0.4 0.5
Actinobacteria
Alphaproteobacteria
Betaproteobacteria
Firmicutes
Gammaproteobacteria
B
Breadth of coverage (%)
0 2 4 6 8 10 12
Year
2016
2018
1P02AnB
1P10PC
2P2-tot-C
1P08PE
1P06PA
1P04PC
1P01AC
1P02AA
1P10AnD
2P06AB
1P06AB
1P10UB
1P10AE
1P03SA
1P01AA
3P01SB
1P01SD
1P02SD
1P06AnD
1P06PB
2P01SA
3P2-tot-E-2
1P10SD

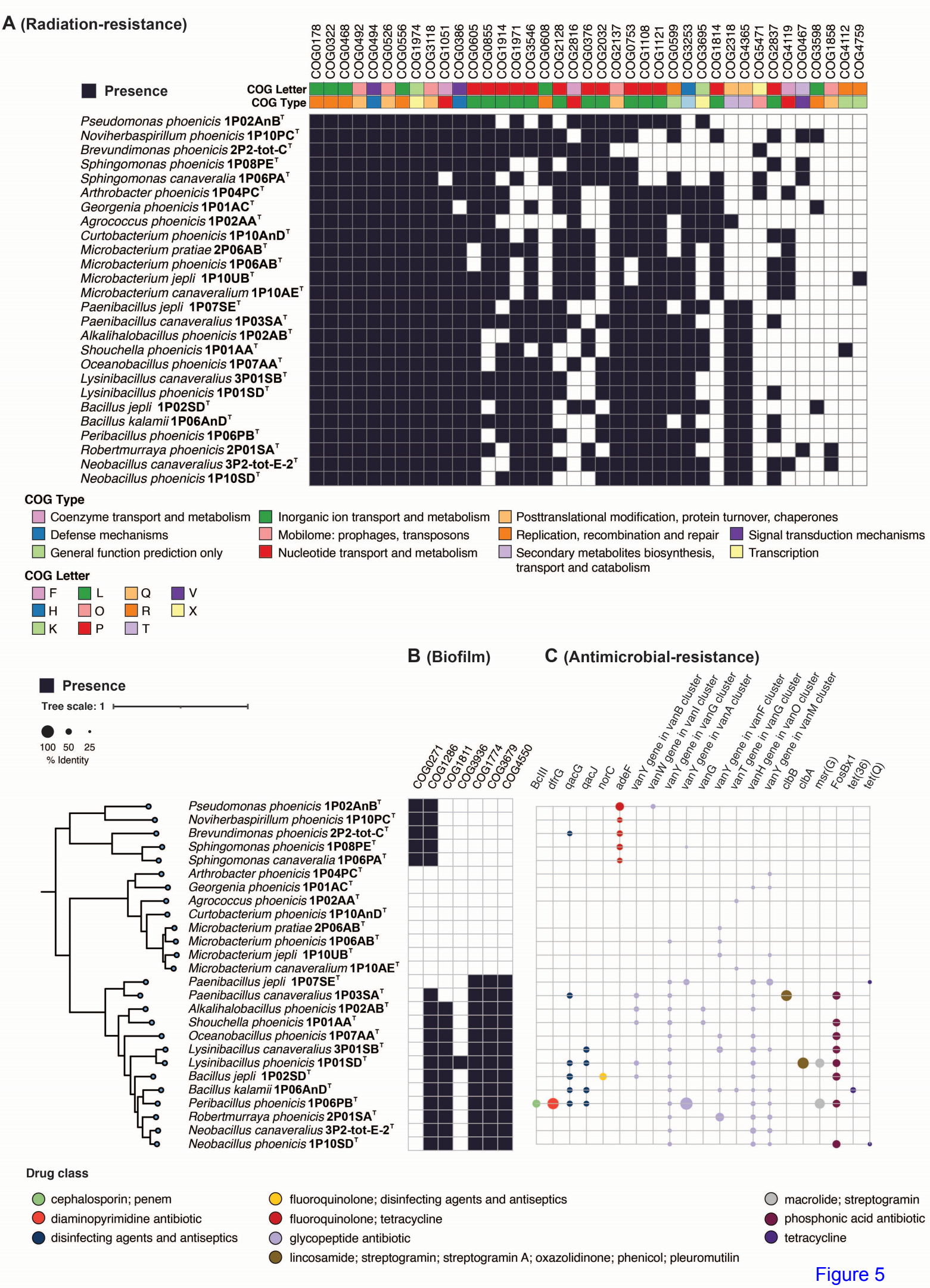

A (Radiation-resistance)
COG0178 COG0322 COG0468 COG0492 COG0494 COG0526 COG0556 COG1974 COG3118 COG1051 COG0386 COG0605 COG0855 COG1914 COG1971 COG3546 COG0608 COG2128 COG2816 COG0376 COG2032 COG2137 COG0753 COG1108 COG1121 COG0599 COG3253 COG3695 COG1814 COG2318 COG4365 COG5471 COG2837 COG4119 COG0467 COG3598 COG1858 COG4112 COG4759
Presence
COG Letter
COG Type
Pseudomonas phoenicis 1P02AnB^T
Noviherbaspirillum phoenicis 1P10PC^T
Brevundimonas phoenicis 2P2-tot-C^T
Sphingomonas phoenicis 1P08PE^T
Sphingomonas canaveralia 1P06PA^T
Arthrobacter phoenicis 1P04PC^T
Georgenia phoenicis 1P01AC^T
Agrococcus phoenicis 1P02AA^T
Curtobacterium phoenicis 1P10AnD^T
Microbacterium pratiae 2P06AB^T
Microbacterium phoenicis 1P06AB^T
Microbacterium jepli 1P10UB^T
Microbacterium canaveralium 1P10AE^T
Paenibacillus jepli 1P07SE^T
Paenibacillus canaveralius 1P03SA^T
Alkalihalobacillus phoenicis 1P02AB^T
Shouchella phoenicis 1P01AA^T
Oceanobacillus phoenicis 1P07AA^T
Lysinibacillus canaveralius 3P01SB^T
Lysinibacillus phoenicis 1P01SD^T
Bacillus jepli 1P02SD^T
Bacillus kalamii 1P06AnD^T
Peribacillus phoenicis 1P06PB^T
Robertmurraya phoenicis 2P01SA^T
Neobacillus canaveralius 3P2-tot-E-2^T
Neobacillus phoenicis 1P10SD^T
COG Type
Coenzyme transport and metabolism
Inorganic ion transport and metabolism
Posttranslational modification, protein turnover, chaperones
Defense mechanisms
Mobilome: prophages, transposons
Replication, recombination and repair
Signal transduction mechanisms
General function prediction only
Nucleotide transport and metabolism
Secondary metabolites biosynthesis, transport and catabolism
Transcription
COG Letter
F L Q V
H O R X
K P T
B (Biofilm)
C (Antimicrobial-resistance)
Presence
Tree scale: 1
100 50 25
% Identity
COG0271 COG1286 COG1811 COG3936 COG1774 COG3679 COG4550
BcIII dfrG qacG qacJ norC adeF vanY gene in vanB cluster vanW gene in vanI cluster vanY gene in vanG cluster vanY gene in vanA cluster vanG vanY gene in vanF cluster vanT gene in vanG cluster vanH gene in vanO cluster vanY gene in vanM cluster clbB clbA msr(G) FosBx1 tet(36) tet(Q)
Drug class
cephalosporin; penem
diaminopyrimidine antibiotic
disinfecting agents and antiseptics
fluoroquinolone; disinfecting agents and antiseptics
fluoroquinolone; tetracycline
glycopeptide antibiotic
lincosamide; streptogramin; streptogramin A; oxazolidinone; phenicol; pleuromutilin
macrolide; streptogramin
phosphonic acid antibiotic
tetracycline

Figure 5

Figure 6

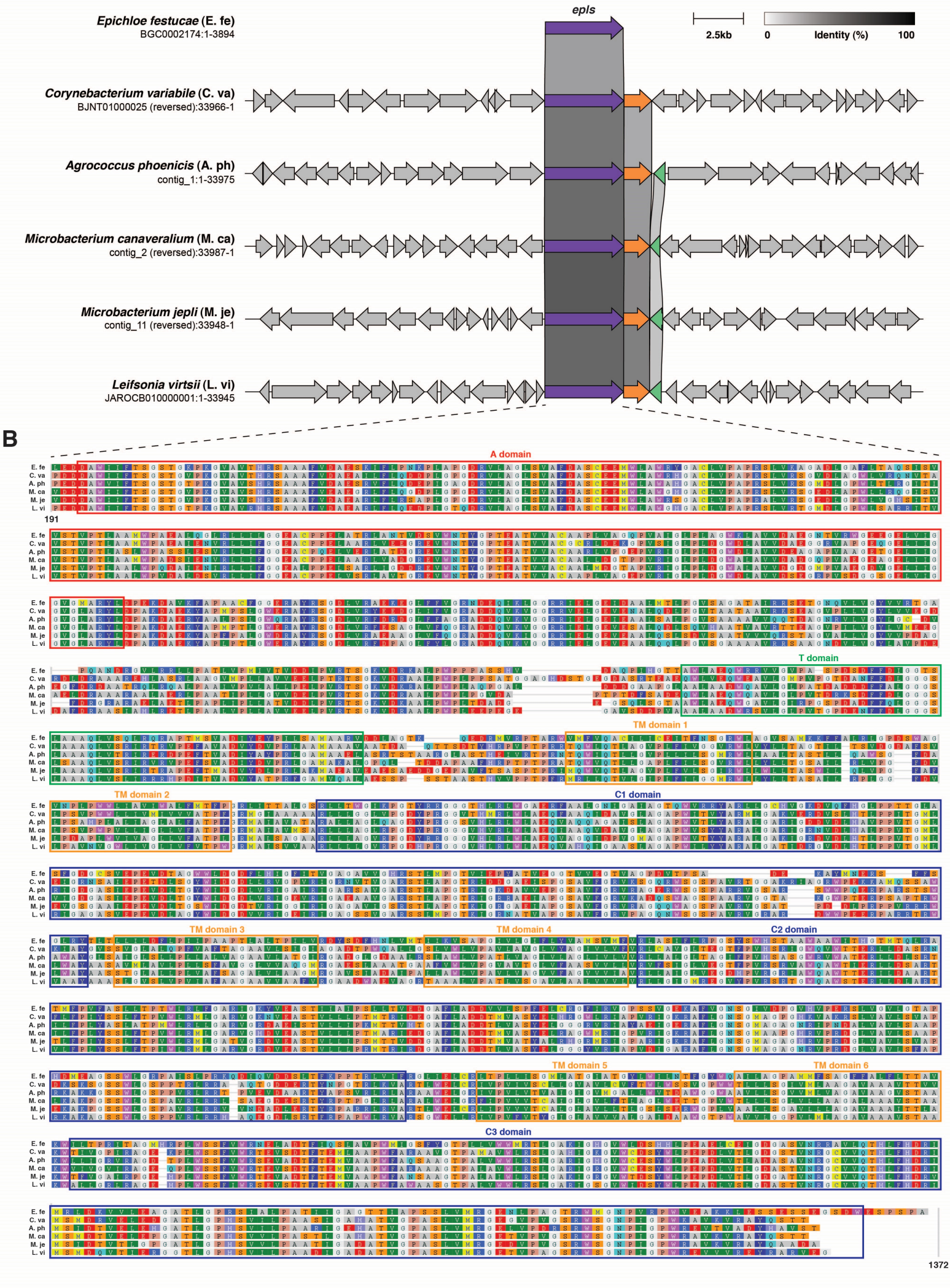

A
epls
2.5kb
0
Identity (%)
100
Epichloe festucae (E. fe)
BGC0002174:1-3894
Corynebacterium variabile (C. va)
BJNT01000025 (reversed):33966-1
Agrococcus phoenicis (A. ph)
contig_1:1-33975
Microbacterium canaveralium (M. ca)
contig_2 (reversed):33987-1
Microbacterium jepli (M. je)
contig_11 (reversed):33948-1
Leifsonia virtsii (L. vi)
JAROCB010000001:1-33945
B
A domain
T domain
TM domain 1
TM domain 2
C1 domain
TM domain 3
TM domain 4
C2 domain
TM domain 5
TM domain 6
C3 domain
E. fe
C. va
A. ph
M. ca
M. je
L. vi
191
1372

Table 1. Whole genome, marker genes sequence similarities, and dDDH values between novel bacterial species and nearest neighbor from the Phoenix spacecraft mission.

Phoenix novel Species	Phoenix strain #	GenBank accession # of 16S rRNA gene of Phoenix isolates	16S rRNA gene characteristics of the closest species			GenBank accession # of WGS of Phoenix strains	WGS characteristics of the closest species					
			Name	Accession #	Percent similarities		Name	Accession #	ANI (%)	AAI (%)	dDDH (%)	gyrB (%)
Shouchella phoenicis	1P01AA	EU977642	*Alkalihalobacillus miscanthi*	NR_180786.1	99.09	JBDFMQ000000000	*Shouchella hunanensis*	GCA_028735875.1	80.5	79.9	20.4	82.8
Georgenia phoenicis	1P01AC	EU977812	*Georgenia satyanarayanai*	NR_117051.1	99.30	CP154855	*"Oceanitalea stevensii"*	GCF_014837105.1	90.0	88.2	35.8	N/A
Lysinibacillus phoenicis	1P01SD	PP475405	*Lysinibacillus fusiformis*	NR_042072.1	98.62	JBDFMP000000000	*Lysinibacillus capsici*	GCF_020217405.1	87.0	90.0	32.3	91.0
Agrococcus phoenicis	1P02AA	PP475406	*Agrococcus terreus*	NR_116650.1	98.55	CP154854	*Agrococcus carbonis*	GCF_900104705.1	86.0	80.0	26.7	90.5
Alkalihalobacillus phoenicis	1P02AB	EU977645	*Alkalihalobacillus alcalophilus*	NR_036889.1	99.53	CP154853	*Alkalihalobacillus alcalophilus*	GCF_004802515.1	92.0	93.4	45.1	96.3
Pseudomonas phoenicis	1P02AnB	PP475407	*Pseudomonas juntendi*	NR_180497.1	98.96	CP154852	*Pseudomonas cremoricolorata*	GCF_000425745.1	85.6	85.1	27.1	87.2
Bacillus jepli	1P02SD	EU977769	*Bacillus timonensis*	NR_133024.1	99.39	JBDFMO000000000	*"Bacillus timonensis"*	GCF_902374785.1	83.0	85.3	26.7	87.3
Paenibacillus canaveralius	1P03SA	EU977770	*Paenibacillus chitinolyticus*	NR_112053.1	99.09	JBDFMM000000000	*Paenibacillus chitinolyticus*	GCF_004117095.1	90.1	91.9	38.7	94.8
Arthrobacter phoenicis	1P04PC	EU977744	*Pseudarthrobacter phenanthrenivorans*	NR_074770.2	98.23	JBDFMK000000000	*Arthrobacter oryzae*	GCF_003634095.1	83.0	80.2	24.6	N/A
Microbacterium phoenicis	1P06AB	EU977652	*Microbacterium oleivorans*	NR_042262.1	99.80	CP154849	*Microbacterium paludicola*	GCF_001975955.2	86.0	88.3	29.4	91.8
Bacillus kalamii	1P06AnD	PP475408	*Bacillus testis*	NR_144719.1	99.21	JBDFMI000000000	*"Bacillus testis"*	GCF_001243895.1	80.0	76.9	21.0	82.0
Sphingomonas canaveralia	1P06PA	EU977746	*Sphingomonas prati*	NR_152092.1	96.84	JBDFMH000000000	*Sphingomonas jatrophae*	GCF_900113315.1	79.7	66.8	20.2	78.9
Peribacillus phoenicis	1P06PB	EU977747	*Peribacillus frigoritolerans*	NR_115064.1	99.87	JBDFMF000000000	*Peribacillus frigoritolerans*	GCF_022394675.1	93.5	94.3	54.4	97.8
Oceanobacillus phoenicis	1P07AA	EU977643	*Oceanobacillus iheyensis*	NR_075027.2	98.98	JBDFME000000000	*Oceanobacillus kimchii*	GCF_000340475.1	90.4	92.3	40.3	93.2
Paenibacillus jepli	1P07SE	PP475409	*Paenibacillus hispanicus*	NR_152687.1	98.64	JBDFMD000000000	*Paenibacillus daejeonensis*	GCF_000378385.1	81.0	82.0	22.4	84.7
Sphingomonas phoenicis	1P08PE	EU977752	*Sphingomonas roseiflava*	NR_117716.1	97.29	JBDFMC000000000	*Sphingomonas metalli*	GCF_014641735.1	83.0	77.4	23.7	N/A
Microbacterium canaveralium	1P10AE	EU977655	*Microbacterium timonense*	NR_179660.1	98.36	JBDFMA000000000	*Microbacterium hydrothermale*	GCF_004854025.1	84.0	80.8	25.4	79.7
Curtobacterium phoenicis	1P10AnD	EU977722	*Curtobacterium pusillum*	NR_042315.1	99.47	JBDFLZ000000000	*Curtobacterium luteum*	GCF_014646995.1	89.0	90.1	33.4	80.0
Noviherbaspirillum phoenicis	1P10PC	EU977754	*Noviherbaspirillum aurantiacum*	NR_118040.1	99.67	JBDFLX000000000	*Noviherbaspirillum soli*	GCF_015352955.1	93.6	94.2	52.4	97.9
Neobacillus phoenicis	1P10SD	EU977785	*Neobacillus bataviensis*	NR_036766.1	98.60	CP154847	*"Bacillus salipaludis"*	GCF_004358205.1	79.0	73.2	20.8	78.3
Microbacterium jepli	1P10UB	EU977807	*Microbacterium timonense*	NR_179660.1	97.63	JBDFLW000000000	*Microbacterium lemovicicum*	GCF_003991875.1	93.0	94.2	47.5	95.4
Robertmurraya phoenicis	2P01SA	PP475410	*Robertmurraya massiliosenegalensis*	NR_125590.1	99.12	CP154845	*Robertmurraya massiliosenegalensis*	GCF_000311725.1	90.8	92.2	41.8	92.0
Microbacterium pratiae	2P06AB	EU977682	*Microbacterium arborescens*	NR_029265.1	99.93	CP154843	*Microbacterium oleivorans*	GCF_001887285.1	91.0	93.2	39.2	N/A
Brevundimonas phoenicis	2P2-tot-C	PP475411	*Brevundimonas diminuta*	NR_040805.1	99.17	JBDFLN000000000	*Brevundimonas diminuta*	GCF_900445995.1	93.3	92.0	47.0	97.0
Lysinibacillus canaveralius	3P01SB	EU977788	*Lysinibacillus odysseyi*	NR_025258.1	99.47	CP154840	*Lysinibacillus odysseyi*	GCF_001591965.1	84.0	87.6	28.1	88.2
Neobacillus canaveralius	3P2-tot-E-2	PP475412	*Neobacillus niacini*	NR_024695.1	99.34	JBDFLL000000000	*Neobacillus niacini*	GCF_001591505.1	87.0	87.8	33.4	90.6

Table 2. Species epithet and etymology of the novel species described during this study.

Species name	Strain number	Cell characteristics	Spore formation	Gram stain characteristics	Species Etymology
Shouchella phoenicis	1P01AA	Round-shaped; Cells are aerobic, motile, and rod-shaped. Colony on TSA medium is beige, circular, entire margin, smooth, non-transparent, and raised.	Yes	Gram-positive	*Shouchella phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Georgenia phoenicis	1P01AC	Round-shaped; Cells are aerobic, motile, short rods (0.5 µm ± 0.2 µm in width and 0.5 µm in length). Colonies on R2A medium are light yellow, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	Georgenia phoenicis (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Lysinibacillus phoenicis	1P01SD	Rod-shaped; Cells are aerobic, motile, 2.5 µm ± 0.1 µm long and 0.5 µm in width. Colonies on R2A medium are white, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Lysinibacillus phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Agrococcus phoenicis	1P02AA	Round-shaped; Cells are aerobic, non-motile, short rods (0.3 µm in width and 0.9 ± 0.4 µm in length). Colonies on R2A medium are light yellow, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Agrococcus phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Alkalihalobacillus phoenicis	1P02AB	Rod-shaped cells, aerobic, and endospore forming and motile. Colony on TSA medium is beige to transparent color, circular, entire margin, smooth, and raised.	Yes	Gram-positive	*Alkalihalobacillus phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Pseudomonas phoenicis	1P02AnB	Round-shaped; Cells are aerobic, motile, with 1.8 ± 0.5 µm in length and 0.5 ± 0.1 µm in width. Colonies on R2A medium are light yellow, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Pseudomonas phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Bacillus jepli	1P02SD	Round-shaped; Cells are aerobic, motile, with 1.3 ± 0.5 µm in length and 0.4 ± 0.1 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Bacillus jepli* (jep'li. N.L. gen. n. *jepli*, arbitrary name derived from the abbreviation JPL, meaning of or pertaining to the NASA's Jet Propulsion Laboratory, where the type strain of the species was isolated).
Paenibacillus canaveralius	1P03SA	Rod-shaped; Cells are aerobic, motile, short rods (0.4 µm ± 0.1 µm in width and 0.9 ± 0.2 µm in length). Colonies on R2A medium are white, irregular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-positive	*Paenibacillus canaveralius* (ca.na.ve.ra'li.us. N.L. masc. adj. *canaveralius* pertaining to (Cape) Canaveral, isolated from walls and floors of the Kennedy Space Center at Cape Canaveral).
Arthrobacter phoenicis	1P04PC	Rod-shaped; Cells are aerobic, non-motile, short rods (0.5 µm in width and 1.0 ± 0.2 µm in length). Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Arthrobacter phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Microbacterium phoenicis	1P06AB	Round-shaped; Cells are aerobic, non-motile, short rods (0.2 µm ± 0.1 µm in width and 0.5 ± 0.2 µm in length). Colonies on R2A medium are dark beige, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-positive	*Microbacterium phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Bacillus kalamii	1P06AnD	Rod-shaped; Cells are aerobic, motile, with 2.1 ± 0.7 µm in length and 0.6 ± 0.1 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Bacillus kalamii* (ka.lam'i.i. N.L. gen. n. *kalamii* referring to Abdul Kalam, a well-known scientist who advanced space research in India).
Sphingomonas canaveralia	1P06PA	Cells are aerobic, non-motile, short rods or ovoid. Colonies on R2A medium are bright yellow, small, circular, and entire margin, smooth, and raised after incubation at 25˚C for 2 to 7 days.	No	Gram-negative	*Sphingomonas canaveralia* (ca.na.ve.ra'li.a. N.L. fem. adj. *canaveralia* pertaining to (Cape) Canaveral, isolated from walls and floors of the Kennedy Space Center at Cape Canaveral).
Peribacillus phoenicis	1P06PB	Rod-shaped; Cells are aerobic, motile, with 3.0 ± 1.1 µm in length and 0.6 ± 0.1 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Peribacillus phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Oceanobacillus phoenicis	1P07AA	Rod-shaped; Cells are aerobic, motile, short rods (0.4 µm ± 0.1 µm in width and 0.9 ± 0.3 µm in length). Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-positive	*Oceanobacillus phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).

Species name	Strain number	Cell characteristics	Spore formation	Gram stain characteristics	Species Etymology
Paenibacillus jepli	1P07SE	Rod-shaped; Cells are aerobic, motile, 1.4 ± 0.3 µm in length and 0.7 µm ± 0.1 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Paenibacillus jepli* (jep'li. N.L. gen. n. *jepli*, arbitrary name derived from the abbreviation JPL, meaning of or pertaining to the NASA's Jet Propulsion Laboratory, where the type strain of the species was isolated.
Sphingomonas phoenicis	1P08PE	Rod-shaped; Cells are aerobic, non-motile, 1.6 ± 0.5 µm in length and 0.5 µm ± 0.1 µm in width. Colonies on R2A medium are orangeish, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Sphingomonas phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Microbacterium canaveralium	1P10AE	Round-shaped; Cells are aerobic, non-motile, short rods (0.6 µm ± 0.1 µm in width and 1.1 ± 0.4 µm in length). Colonies on R2A medium are yellow, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-positive	*Microbacterium canaveralium* (ca.na.ve.ra'li.um. N.L. neut. adj. *canaveralium* pertaining to (Cape) Canaveral, isolated from walls and floors of the Kennedy Space Center at Cape Canaveral).
Curtobacterium phoenicis	1P10AnD	Round-shaped; Cells are aerobic, motile, short rods (0.4 µm in width and 1.0 ± 0.2 µm in length). Colonies on R2A medium are yellow, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Curtobacterium phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Noviherbaspirillum phoenicis	1P10PC	Round-shaped; Cells are aerobic, motile, 1.4 ± 0.6 µm in length and 0.6 µm ± 0.1 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Noviherbaspirillum phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Neobacillus phoenicis	1P10SD	Rod-shaped; Cells are aerobic, motile, 1.8 ± 1.0 µm in length and 1.1 µm ± 1.1 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-positive	*Neobacillus phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Microbacterium jepli	1P10UB	Round-shaped; Cells are aerobic, non-motile, short rods (0.4 µm ± 0.1 µm in width and 0.7 ± 0.2 µm in length). Colonies on R2A medium are white, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Microbacterium jepli* (jep'li. N.L. gen. n. *jepli*, arbitrary name derived from the abbreviation JPL, meaning of or pertaining to the NASA's Jet Propulsion Laboratory, where the type strain of the species was isolated.
Robertmurraya phoenicis	2P01SA	Rod-shaped; Cells are aerobic, motile, 2.1 µm ± 0.5 µm long and 0.5 µm ± 0.1 µm in width. Colonies on R2A medium are light white, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-positive	*Robertmurraya phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Microbacterium pratiae	2P06AB	Round-shaped; Cells are aerobic, non-motile, short rods (0.4 µm ± 0.1 µm in width and 1.0 ± 0.4 µm in length). Colonies on R2A medium are dark beige, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-positive	*Microbacterium pratiae* (pra'ti'ae. N.L. gen. fem. n. *pratiae* referring to Dr. Lisa Pratt, a biogeochemist and astrobiologist who previously served as the Planetary Protection Officer for NASA).
Brevundimonas phoenicis	2P2-tot-C	Round-shaped; Cells are aerobic, motile, short rods (0.6 µm ± 0.1 µm in width and 1.2 ± 0.4 µm in length). Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	No	Gram-negative	*Brevundimonas phoenicis* (phoe'ni.cis. L. gen. n. *phoenicis* of phoenix, isolated from the surface of the clean room where the Mars Phoenix spacecraft was assembled).
Lysinibacillus canaveralius	3P01SB	Rod-shaped; Cells are aerobic, motile, 1.7 µm ± 0.7 µm long and 0.4 µm ± 0.1 µm in width. Colonies on R2A medium are white, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Lysinibacillus canaveralius* (ca.na.ve.ra'li.us. N.L. masc. adj. *canaveralius* pertaining to (Cape) Canaveral, isolated from walls and floors of the Kennedy Space Center at Cape Canaveral).
Neobacillus canaveralius	3P2-tot-E-2	Rod-shaped; Cells are aerobic, motile, 2.3 ± 1.0 µm in length and 1.0 µm ± 0.5 µm in width. Colonies on R2A medium are light beige, circular and opaque after incubation at 25˚C for 48 h.	Yes	Gram-negative	*Neobacillus canaveralius* (ca.na.ve.ra'li.us. N.L. masc. adj. *canaveralius* pertaining to (Cape) Canaveral, isolated from walls and floors of the Kennedy Space Center at Cape Canaveral).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarymaterialPhoenixManuscript.pdf
- zT.SupplementaryTablesPhoenixManuscript.xlsx