Current Microbiology (2020) 77:1097–1103
https://doi.org/10.1007/s00284-020-01908-5

Complete Genome Sequence of *Paenibacillus* sp. JZ16, a Plant Growth Promoting Root Endophytic Bacterium of the Desert Halophyte *Zygophyllum Simplex*

Abdul Aziz Eida[1] · Salim Bougouffa[2,3] · Intikhab Alam[2] · Heribert Hirt[1,4] · Maged M. Saad[1]

Received: 7 December 2019 / Accepted: 29 January 2020 / Published online: 3 February 2020

Abstract

Paenibacillus sp. JZ16 is a gram-positive, rod-shaped, motile root endophytic bacterium of the pioneer desert halophytic plant *Zygophyllum simplex*. JZ16 was previously shown to promote salinity stress tolerance in *Arabidopsis thaliana* and possesses a highly motile phenotype on nutrient agar. JZ16 genome sequencing using PacBio generated 82,236 reads with a mean insert read length of 11,432 bp and an estimated genome coverage of 127X, resulting in a chromosome of 7,421,843 bp with a GC content of 49.25% encoding 6710 proteins, 8 rRNA operons, 117 ncRNAs and 73 tRNAs. Whole-genome sequencing analysis revealed a potentially new species for JZ16. Functional analysis revealed the presence of a number of enzymes involved in the breakdown of plant-based polymers. JZ16 could be of potential use in agricultural applications for promoting biotic and abiotic stress tolerance and for biotechnological processes (e.g., as biocatalysts for biofuel production). The culture-dependent collection of bacterial endophytes from desert plants combined with genome sequence mining provides new opportunities for industrial applications.

Introduction

Feeding a dramatically growing human population in a sustainable manner is a global challenge in agriculture. Microbes, such as bacteria, are emerging as a sustainable, green technology to replace or reduce the application of agrochemicals, such as chemical fertilizers and pesticides, which can have harmful long-term effects on the environment and human health [1, 2]. They are also showing promising results for increasing yields of important crops, especially in light of biotic and abiotic stresses and global warming [3–7]. Since its reclassification as a separate genus more than 25 years ago [8], the bacterial genus *Paenibacillus* has expanded to encompass several species that have shown plant growth promoting (PGP) and biocontrol activity [9]. For example, *P. polymyxa* isolates have been shown to mediate protection from the pathogen *Erwinia carotovora* (biotic stress) and drought (abiotic stress) in the model plant *Arabidopsis thaliana* [10]. It was also shown to promote growth of maize (*Zea mays*) and to possess antagonistic activity against the fungal plant pathogens *Fusarium oxysporum*, *F. graminearum*, and *Botrytis cinerea* [11, 12]. Inoculation of *P. yonginensis* was shown to promote aluminum, drought and salt stress tolerance in *A. thaliana* and salt stress tolerance in the medicinal plant *Panax ginseng* [13, 14].

Sequence Accession Number: The data for the bacterial genome assembly of Paenibacillus sp. JZ16 and sequencing were deposited in NCBI/DDBJ/EMBL database under the accession± number CP017659, BioSample SAMN05828174 and BioProject PRJNA345401. The annotations obtained by in-house INDIGO pipeline are available through the KAUST library repository (https://doi.org/10.25781/KAUST-0XG5M).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00284-020-01908-5) contains supplementary material, which is available to authorized users.

✉ Maged M. Saad
Maged.saad@kaust.edu.sa

[1] DARWIN21, Biological and Environmental Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

[2] Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

[3] BioScience Core Lab, Abdullah University of Science and Technology, Thuwal 23955-6900, King, Saudi Arabia

[4] Max F. Perutz Laboratories, University of Vienna, Dr. Bohrgasse 9, 1030 Vienna, Austria

Previously, *Paenibacillus* sp. JZ16 was isolated from the root endosphere of the annual desert halophyte *Zygophyllum simplex* [15]. Using qualitative agar-based plate assays, JZ16 was demonstrated to promote salinity stress tolerance in *A. thaliana* and, therefore, the genome of JZ16 was sequenced and presented here.

Materials and Methods

Growth Conditions and Genomic DNA Extraction

Pure cultures of *Paenibacillus* sp. JZ16 were regularly stored in 20% glycerol at -80 °C. The strain was regularly grown in Luria–Bertani LB broth (Lennox L Broth Base, Invitrogen) or on LB agar plates at 28 °C. Fresh, pure bacterial cultures were used for total genomic DNA extraction using Sigma's GenElute bacterial genomic DNA kit (Sigma-Aldrich) following the manufacturer's protocol. DNA quality and quantity were assessed by 0.7% agarose gel electrophoresis (35 V, 12 h), NanoDrop 2000 (Thermo Fisher Scientific) and Qubit dsDNA BR assay kit (Thermo Fisher Scientific).

Genome Sequencing and Assembly

DNA was size selected to 10 kb using the BluePippin™ Size-Selection System (Sage Science), following the "High-Pass™ DNA Size Selection of ~ 20 kb SMRTbell™ Templates" manual. The SMRTbell™ template library was prepared according to the instructions from Pacific Biosciences's "Procedure & Checklist—20 kb Template Preparation using BluePippin™ Size-Selection System" guide. The SMRT cells were run at the KAUST Bioscience Core Labs on the PacBio *RSII* (Pacific Biosciences) sequencing platform using P6-C4 chemistry. Raw data from PacBio's platform were assembled into a draft assembly using the Hierarchical Genome Assembly Process v3 (HGAP3) [16] from PacBio's SMRT Analysis pipeline v2.3.0.140936 patch 5. The assembly workflow can be broken down to three main steps: a preassembly step that mapped single pass reads to seed reads to generate consensus reads that were then quality trimmed. De novo assembly was done using the overlap layout consensus approach. The final step is consensus polishing using Quiver to reduce indels and base substitution using quality scores embedded in the raw data. To determine whether assembled contigs are circular, dot plots were generated using Gepard [17] for detecting overlaps at the peripheries. Overlaps were collapsed and genome was closed using Minimus2 [18]. Finally, additional polishing rounds were performed using Quiver by applying quality scores from raw data to correct for indels and base substitutions where the output from one round is inputted to the next.

Genome Annotation

Genome annotation was carried out using the Automatic Annotation of Microbial Genomes (AAMG) which is an integrated module in the in-house INDIGO-Desert v1.1 pipeline [19]. Briefly, gene prediction was done using prodigal v2.6.1 [20]. Functional annotation was done using a multitude of tools and databases. InterProScan [21] was used to assign domain information, Gene Ontology (GO) terms and KEGG pathways. Predicted genes were compared using BLAST against UniProt (https://www.uniprot.org/) for generic annotations and cross-references to COGs (Cluster of Orthologous Genes (COGs). For annotation of gene function, genes were compared to KEGG database (Functional Kyoto Encyclopedia of Genes and Genomes) [22]. RPS-BLAST [23] was used to identify conserved domains and COG (Clusters of Orthologous Groups). Predicted genes were also BLAST-ed against NCBI-nr, UniProt and KEGG. Ribosomal RNAs (rRNAs), transfer RNAs (tRNAs) and other non-coding RNAs (ncRNAs) were predicted using RNAmmer 1.2 [24], tRNAscan-SE 2.0 [25], and Infernal software [26], respectively. Function and pathway analysis were also performed using BlastKOALA web tool of KEGG database [22]. Identification of gene clusters responsible for the biosynthesis of secondary metabolites was performed using antiSMASH v.4.2.0 [27]. Chromosome map was generated using DNAPlotter release 18.0.2 [28].

Phylogenetic Analysis

The 16S rRNA gene sequences of isolate JZ16 were predicted using RNAmmer 1.2 and the most common and identical sequences of the eight copies were compared to known sequences listed in NCBI's GenBank using BLASTn [29]. The sequences with the highest similarity in terms of sequence identity and query coverage, along with other type strains from similar and distant genera were used for the phylogenetic tree construction. Multiple alignment of the nucleotide sequences was performed using MUSCLE [30]. The phylogenetic tree was constructed by the Neighbor-Joining method [31], based on the Kimura 2-parameter model [32], with bootstrap analysis (1,000 replications) using the software MEGA version 7 [33]. For distinguishing between *Paenibacillus* strains at species level, Digital DNA-DNA Hybridization (dDDH) and Average Nucleotide Identity (ANI) calculations were performed. Pairwise BLAST-based Average Nucleotide Identity values (ANIb) were obtained using JSpecies [34]. The genome sequence data were uploaded to the Type (Strain) Genome Server (TYGS), a free bioinformatics platform

available under https://tygs.dsmz.de, for a whole-genome-based taxonomic analysis [35]. All pairwise comparisons among the set of genomes were conducted using GBDP and accurate intergenomic distances inferred under the algorithm 'trimming' and distance formula d5 [36]. 100 distance replicates were calculated each. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 2.1 [36].

Electron Microscopy

For electron microscopy, pure colonies were picked and grown in LB broth at 28 °C overnight, sub-cultured the next day to OD_{600} of 0.1. Cells from the exponential phase were harvested by centrifugation at 3000×*g*, washed twice and resuspended in 0.1 M phosphate buffer (PBS). Bacterial cells were then fixed with 2.5% glutaraldehyde in cacodylate buffer (0.1 M, pH 7.4) overnight and rinsed with 0.1 M cacodylate buffer. Fixed samples were post-fixed with 2% osmium tetroxide, 1.5% potassium ferrocyanide in 0.1 M cacodylate buffer for 1 h and then washed with water. Samples were then dehydrated in ethanol series and embedded in Epon epoxy resin. Contrasting sections were stained with uranyl acetate and lead citrate and imaging was performed using a Titan 80–300 S/TEM (Titan Cryo Twin; FEI Company) operating at 300 kV. The sample preparation and imaging were performed at the KAUST Imaging and Characterization Core Labs.

Motility In Vitro Assay

For the motility assay, a single colony was streaked at the top of 12 × 12 cm square Petri dishes containing LB agar and the dishes were incubated at 28 °C for three days. Photographs were taken every 24 h after incubation for 3 days. The experiment was performed using three biological replicates.

Sequence Accession Number

The data for the bacterial genome assembly of *Paenibacillus* sp. JZ16 and sequencing were deposited in NCBI/DDBJ/EMBL database under the accession number CP017659, BioSample SAMN05828174 and BioProject PRJNA345401. The annotations obtained by in-house INDIGO pipeline are available through the KAUST library repository (https://doi.org/10.25781/KAUST-0XG5M).

Results and Discussion

Paenibacillus sp. JZ16 is a gram-positive, rod-shaped (~0.4 × 3.5 µm) bacterium with flagella (ESM_1 Fig. S1a). It was isolated from the root endosphere of the annual desert halophyte *Z. simplex* and exhibited salinity stress tolerance promotion (SSTP) on *A. thaliana* [15]. Phenotypically, on nutrient agar plates, JZ16 displays extensive motility behavior of motile microcolonies (spreading over the agar surface at an approximate speed of ~ 1.75 mm/hour) (ESM_1 Fig. S1b), which is typical for some members of the *Paenibacillus* genus [8, 37, 38]. Whole-genome sequencing by PacBio sequencing platform using one SMRT cell, taking one 360 min movie, generated 82,236 reads with a mean insert read length of 11,432 bp and an estimated genome coverage of 127X. Genome assembly and analysis of JZ16 generated one single circular chromosome of approximately 7.42 Mbp with a GC content of 49.25% (Table 1; Fig. 1a). By INDIGO annotation 6924 open reading frames (ORF) were identified and, thus, the chromosome of JZ16 contains a gene density of approximately 933 genes/Mbp. Of the 6924 genes, 97% (6,710) were protein coding sequences (CDS). The chromosome contained eight 16S-23S-5S rRNA operons, 117 ncRNAs, and 73 tRNAs.

Phylogenetic analysis using 16S rRNA sequences of JZ16 and closest relatives (based on top BLAST hits) showed clustering of JZ16 with the clade of *Paenibacillus lautus* (Fig. 1b). Using the whole-genome sequence, BLAST results revealed alignment to *P. lautus* strain E7593-69 (CP032412) at 90.82% identity with query cover of 78%. Whole-genome sequence phylogenetic analysis of JZ16 with other *Paenibacillus* type strains revealed clustering of JZ16 with *P. lautus* NBRC 15,380 (ESM_1 Fig. S2.). *P. lautus* NBRC 15,380 showed the highest similarity with JZ16 with an ANIb value of 91.48% and dDDH of 68.10% (formula d0) and 65% (formula d6) (Table 2; ESM_2 Table S1.). However, these values do not exceed the species cut-off of 95% and 70% for ANI and dDDH, respectively, and, thus, JZ16 is a potential new species.

Table 1 Summary of JZ16 genome features

Feature	Chromosome
Genome size (bp)	7,421,843
GC content (%)	49.25
ORF	6,924
Gene density (genes/Mbp)	932.92
CDS	6,710
Genes assigned to:	
UniProt	6,541 (97.48%)
COG	4,277 (63.74%)
KEGG	5,053 (75.30%)
rRNAs	24
16S-23S-5S operons	8
ncRNAs	117
tRNAs	73

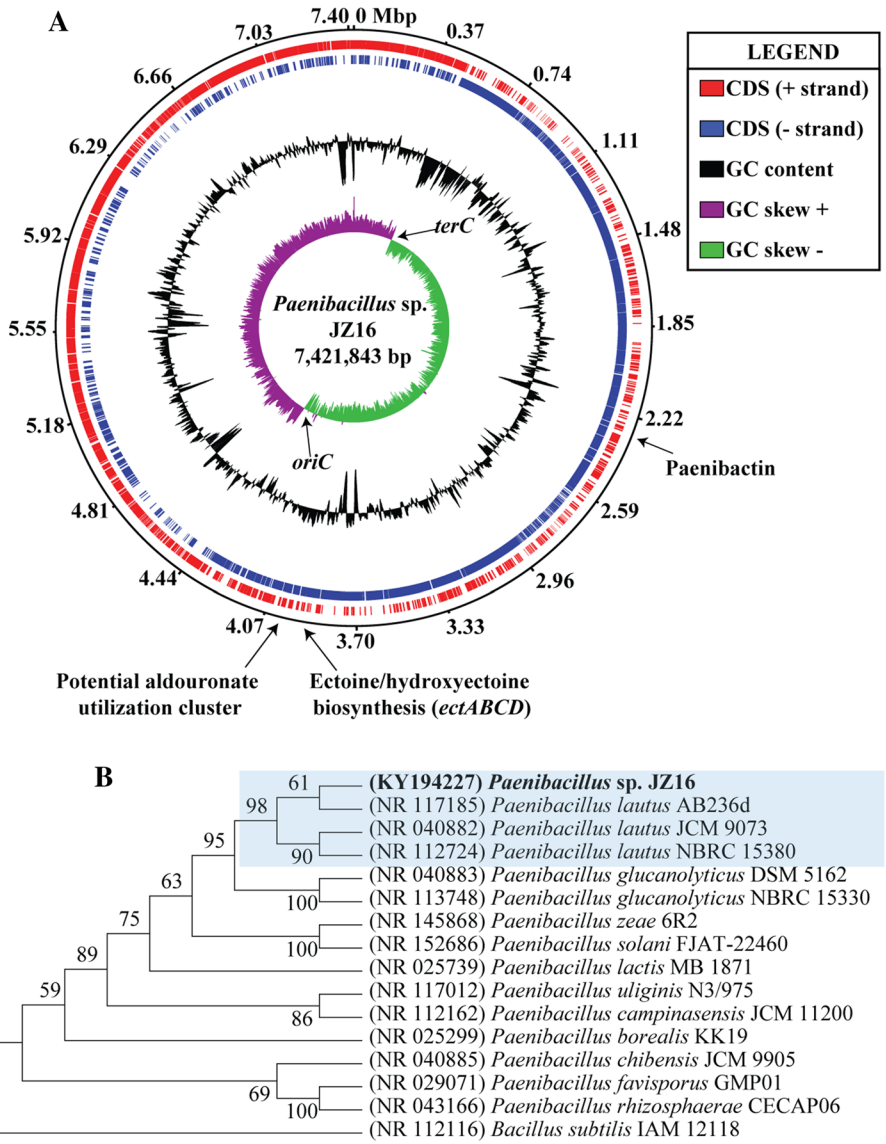

Fig. 1 Phylogenetic analysis of JZ16 and circular representation of chromosome. **a** Phylogenetic tree generated using 16S rRNA sequences, alignment of sequences using MUSCLE, evolutionary relationships inferred using the Neighbor-Joining method and the evolutionary distances computed using the Kimura 2-parameter method. GenBank accession numbers of isolates are presented before taxonomic names in parenthesis. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. **b** The tracks from inside to outside: GC skew [(G−C/(G+C)] positive (purple) and negative (green), % GC content (black), coordinates in mega base pair (Mbp), reverse-strand CDSs (blue), forward-strand CDSs (red). Arrows on the GC skews indicate the origin of replication (*oriC*) and replication termination (*terC*) regions where the shifts occur (Color figure online)

AntiSMASH analysis identified nine secondary metabolite regions, but only two clusters were similar to known clusters: Paenibactin (100% similarity, BGC0000401: paenibactin biosynthetic gene cluster from *Paenibacillus elgii* B69) and Ectoine (75% similarity, BGC0000853: ectoine biosynthetic gene cluster from *Streptomyces anulatus*) biosynthesis (Fig. 1a). Paenibactin is a catecholate-type siderophore, while ectoine is a compatible solute prominent in halophilic and halotolerant bacteria [39, 40]. The genes for encoding for ectoine biosynthesis (*ectABC*) and its conversion to hydroxyectoine (*ectD*) were present in JZ16, and their presence could contribute to the osmotic and salt stress tolerance of JZ16 [15].

Functional analysis using BlastKOALA identified 3061 CDSs (45.6% of total CDSs) with assigned functions in JZ16. Among the genes identified in the genome is a chitinase (three copies), an enzyme that breaks down chitin found in the cell wall of fungi and, thus, JZ16 could be a potential candidate as a biocontrol agent and pathogen antagonist [41]. Two-component systems are important signal transduction systems allowing bacteria to respond and adapt to changes in the external environment [42]. We found many copies of the two-component system (TCS) (YesN/YesM), comprising of a sensor histidine kinase (*yesM*, 52 copies) and response regulator (*yesN*, 47 copies). Even more copies of the aldouronate ABC transporter (*lplABC*) were found. This system comprises of a substrate-binding protein (*lplA*, 71 copies) and two protein with permease components (*lplB*, 66 copies and *lplC*, 66 copies) (ESM_2 Table S2.).

Chow et al. [43] demonstrated the structure, function and regulation of the aldouronate utilization gene cluster in *Paenibacillus* sp. strain JDR-2. Aldouronates are the by-products of the hemicellulose degradation process. The xylan backbone is depolymerized, via the enzymatic

Table 2 Pairwise comparisons of *Paenibacillus* sp. JZ16 versus type strain genomes

Subject strain	dDDH*			Biosample accession	Assembly accession
	(d0, in %)	(d4, in %)	(d6, in %)		
Paenibacillus lautus NBRC 15,380	68.1	47.5	65	SAMD00034170	GCA_004000945
Paenibacillus glucanolyticus DSM 5162	41	28.2	36.9	SAMN04432816	GCF_001632305
Paenibacillus solani FJAT-22460	34	27.2	31.4		
Paenibacillus macerans NCTC 6355	13.1	25.2	13.5	SAMEA4076714	GCA_900454495
Paenibacillus ihbetae JCM 31131 T	23.9	23.9	22.8	SAMN05335590	GCF_002741055
Paenibacillus wulumuqiensis Y24	12.8	23.2	13.2	SAMN03447299	GCA_000971965
Paenibacillus faecis DSM 23,593	13.2	22.8	13.5	SAMN12429299	GCA_008084145
Paenibacillus pinistramenti ASL46	13	22.3	13.3	SAMN11522047	GCA_005869875
Paenibacillus etheri SH7	13	21.8	13.3	SAMN03372544	GCA_001012825
Paenibacillus zeisoli 3–5-3	13	21.6	13.4	SAMN10579155	GCA_003994465
Paenibacillus peoriae KCTC 3763	12.9	21.6	13.3	SAMN02470185	GCA_000236805
Paenibacillus glacialis DSM 22,343	13	21.3	13.4	SAMN04574067	GCA_001637205
Paenibacillus macquariensis DSM 2	12.9	21.3	13.3	SAMN04574061	GCA_001637165
Paenibacillus uliginis N3/975	15.9	20.6	15.9	SAMN05661091	GCA_900177425
Paenibacillus chibensis NBRC 15,958	13.4	20.3	13.7	SAMD00034158	GCA_004001045

*The following table contains the pairwise dDDH values between JZ16 genome and the selected type strain genomes. The dDDH values are provided for the three different GBDP formulas: d0 (a.k.a. GGDC formula 1): length of all HSPs divided by total genome length; d4 (a.k.a. GGDC formula 2): sum of all identities found in HSPs divided by overall HSP length; d6 (a.k.a. GGDC formula 3): sum of all identities found in HSPs divided by total genome length

action of the endoxylanases (e.g. endo-1,4-beta-xylanase, *xynA*), and the aldotetrauronates (e.g. methylglucuronoxylotriose) and xylooligosaccharides are then converted into fermentable xylose [43]. The aldouronate utilization gene cluster was comprised of genes encoding the TCS YesN/YesM (*yesNM*), aldouronate ABC transporter (*lplABC*), and the enzymes α-glucuronidase (*aguA*), endoxylanase (*xynA*) and β-xylosidase/α-arabinofuranosidase (*xynB*). Genome analysis of JZ16 showed presence of five copies of *xynA* and two copies of each of *xynB* and *aguA* (ESM_2 Table S2.). The presence of a single cluster containing all the required genes was not found in JZ16. However, one region (4035047–4047414, Fig. 1a) contained *aguA*, *xynA*, *lplC*, *lplB*, *lplA*, *yesM*, *yesN* and a glucuronate isomerase (*uxaC*) that is often found in the *xytB* locus of xylan utilization systems in *Xanthomonas* [44].

A recent study using an integrated proteomics and lipidomics approach revealed the possible involvement of lipid modifications and key enzymes in metabolic pathways between different modes of motility (swimming vs. swarming) in *P. polymyxa* [45]. JZ16 contains many of the genes that *P. polymyxa* was suggested to be responsible for the motility phenotype. For example, JZ16 possesses all genes, except *plsY*, involved in the metabolic pathways of sn-glycerol-3-phosphate, a key precursor for the biosynthesis of phospholipids (ESM_2 Table S2.). PlsY converts sn-glycerol-3-phosphate to lysophosphatidic acid (LysoPA) and LysoPA is converted to PA by either PlsC or a diacylglycerol kinase (*dagK*). Poudel et al. [45] indicated that the swarming bacteria accumulated PA, while swimming bacteria favored LysoPA accumulation. Furthermore, JZ16 also contains a number of enzymes that either assist in the catabolic processes involved during motility or in the degradation and metabolism of plant polymers. These include alpha-L-fucosidase (*FUCA, AXY8*), α- and β-galactosidase (*bgaB*), β-glucosidase (*bglX*), endoglucanase, pectin esterase, glucuronoarabinoxylan endo-1,4-beta-xylanase (*xynC*), arabinoxylan arabinofuranohydrolase (*xynD*), hydrolases, pullulanase (*pulA*), fructan beta-fructosidase (*fruA*), xylose isomerase (*xylA*), and alpha-D-xyloside xylohydrolase (*xylS*) (ESM_2 Table S2.). Therefore, *Paenibacillus* sp. JZ16 could have potential agricultural and biotechnological applications, serving as biocontrol agents and/or biocatalysts in the conversion of plant-based biomass to biofuels and chemicals, especially in the light of climate change.

Acknowledgements The work presented is part of the DARWIN21 project (https://www.darwin21.org/), with the objective to improve sustainable agriculture on arid lands by exploiting microbes isolated from pioneer desert plants that are able to survive in extreme environmental conditions. The authors would thank all members of Hirt lab, CDA management team and the Bioscience Core Labs in KAUST for the technical assistance and for their help in many aspects of this work.

Funding The work was funded by KAUST baseline research project BAS/1/1062–01-01 of H.H., 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.

References

1. Ahmad M, Pataczek L, Hilger TH, Zahir ZA, Hussain A, Rasche F, Schafleitner R, Solberg SØ (2018) Perspectives of microbial inoculation for sustainable development and environmental management. Front Microbiol 9:436. https://doi.org/10.3389/fmicb.2018.02992
2. Alori ET, Babalola OO (2018) Microbial inoculants for improving crop quality and human health in Africa. Front Microbiol. https://doi.org/10.3389/fmicb.2018.02213
3. Daur I, Saad MM, Eida AA, Ahmad S, Shah ZH, Ihsan MZ, Muhammad Y, Sohrab SS, Hirt H (2018) Boosting alfalfa (*Medicago sativa* L.) production with rhizobacteria from various plants in Saudi Arabia. Front Microbiol. https://doi.org/10.3389/fmicb.2018.00477
4. de Zélicourt A, Synek L, Saad MM, Alzubaidy H, Jalal R, Xie Y, Andrés-Barrao C, Rolli E, Guerard F, Mariappan KG, Daur I, Colcombet J, Benhamed M, Depaepe T, Van Der Straeten D, Hirt H (2018) Ethylene induced plant stress tolerance by *Enterobacter* sp. SA187 is mediated by 2-keto-4-methylthiobutyric acid production. PLoS Genet 14(3):e1007273. https://doi.org/10.1371/journal.pgen.1007273
5. Pengnoo A, Kusongwiriyawong C, Nilratana L, Kanjanamaneesathian M (2000) Greenhouse and field trials of the bacterial antagonists in pellet formulations to suppress sheath blight of rice caused by *Rhizoctonia solani*. Biocontrol 45(2):245–256. https://doi.org/10.1023/A:1009948404423
6. Rozier C, Hamzaoui J, Lemoine D, Czarnes S, Legendre L (2017) Field-based assessment of the mechanism of maize yield enhancement by *Azospirillum lipoferum* CRT1. Sci Rep 7(1):7416. https://doi.org/10.1038/s41598-017-07929-8
7. Sarwar A, Latif Z, Zhang S, Hao J, Bechthold A (2019) A potential biocontrol agent *Streptomyces violaceusniger* AC12AB for managing potato common scab. Front Microbiol. https://doi.org/10.3389/fmicb.2019.00202
8. Ash C, Priest FG, Collins MD (1993) Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Antonie Van Leeuwenhoek 64(3):253–260. https://doi.org/10.1007/bf00873085
9. Grady EN, MacDonald J, Liu L, Richman A, Yuan Z-C (2016) Current knowledge and perspectives of *Paenibacillus*: a review. Microb Cell Fact 15(1):203. https://doi.org/10.1186/s12934-016-0603-7
10. Timmusk S, Wagner EGH (1999) The plant-growth-promoting rhizobacterium *Paenibacillus polymyxa* induces changes in *Arabidopsis thaliana* gene expression: a possible connection between biotic and abiotic stress responses. Mol Plant Microb Interact 12(11):951–959. https://doi.org/10.1094/mpmi.1999.12.11.951
11. Mohd Din ARJ, Rosli MA, Mohamad Azam Z, Othman NZ, Sarmidi MR (2019) *Paenibacillus polymyxa* role involved in phosphate solubilization and growth promotion of *Zea mays* under abiotic stress condition. Proc Natl Acad Sci India Sect B. https://doi.org/10.1007/s40011-019-01081-1
12. Timmusk S, Copolovici D, Copolovici L, Teder T, Nevo E, Behers L (2019) *Paenibacillus polymyxa* biofilm polysaccharides antagonise *Fusarium graminearum*. Sci Rep 9(1):662. https://doi.org/10.1038/s41598-018-37718-w
13. Sukweenadhi J, Kim Y-J, Choi E-S, Koh S-C, Lee S-W, Kim Y-J, Yang DC (2015) *Paenibacillus yonginensis* DCY84T induces changes in *Arabidopsis thaliana* gene expression against aluminum, drought, and salt stress. Microbiol Res 172:7–15. https://doi.org/10.1016/j.micres.2015.01.007
14. Sukweenadhi J, Balusamy SR, Kim Y-J, Lee CH, Kim Y-J, Koh SC, Yang DC (2018) A growth-promoting bacteria, *Paenibacillus yonginensis* DCY84T enhanced salt stress tolerance by activating defense-related systems in *Panax* ginseng. Front Plant Sci. https://doi.org/10.3389/fpls.2018.00813
15. Eida AA, Ziegler M, Lafi FF, Michell CT, Voolstra CR, Hirt H, Saad MM (2018) Desert plant bacteria reveal host influence and beneficial plant growth properties. PLoS ONE 13(12):e0208223. https://doi.org/10.1371/journal.pone.0208223
16. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563. https://doi.org/10.1038/nmeth.2474
17. Krumsiek J, Arnold R, Rattei T (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics 23(8):1026–1028. https://doi.org/10.1093/bioinformatics/btm039
18. Sommer DD, Delcher AL, Salzberg SL, Pop M (2007) Minimus: a fast, lightweight genome assembler. BMC Bioinform 8(1):64. https://doi.org/10.1186/1471-2105-8-64
19. Alam I, Antunes A, Kamau AA, Kalkatawi M, Stingl U, Bajic VB (2013) INDIGO—INtegrated Data Warehouse of MIcrobial GenOmes with examples from the Red Sea extremophiles. PLoS ONE 8(12):e82210. https://doi.org/10.1371/journal.pone.0082210
20. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 11(1):119. https://doi.org/10.1186/1471-2105-11-119
21. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics 30(9):1236–1240. https://doi.org/10.1093/bioinformatics/btu031
22. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44(D1):D457–D462. https://doi.org/10.1093/nar/gkv1070
23. Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res 30(1):281–283. https://doi.org/10.1093/nar/30.1.281
24. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35(9):3100–3108. https://doi.org/10.1093/nar/gkm160
25. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25(5):955–964. https://doi.org/10.1093/nar/25.5.955
26. Nawrocki EP, Eddy SR (2013) Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics 29(22):2933–2935. https://doi.org/10.1093/bioinformatics/btt509
27. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH (2015) AntiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43(W1):W237–W243. https://doi.org/10.1093/nar/gkv437
28. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2008) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25(1):119–120. https://doi.org/10.1093/bioinformatics/btn578
29. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402. https://doi.org/10.1093/nar/25.17.3389

30. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797. https://doi.org/10.1093/nar/gkh340
31. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4(4):406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454
32. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16(2):111–120. https://doi.org/10.1007/BF01731581
33. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874. https://doi.org/10.1093/molbev/msw054
34. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J (2015) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32(6):929–931. https://doi.org/10.1093/bioinformatics/btv681
35. Meier-Kolthoff JP, Göker M (2019) TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10(1):2182. https://doi.org/10.1038/s41467-019-10210-3
36. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14(1):60. https://doi.org/10.1186/1471-2105-14-60
37. Kobayashi K, Kanesaki Y, Yoshikawa H (2016) Genetic analysis of collective motility of *Paenibacillus* sp. NAIST15-1. PLoS Genet 12(10):e1006387. https://doi.org/10.1371/journal.pgen.1006387
38. Fünfhaus A, Göbel J, Ebeling J, Knispel H, Garcia-Gonzalez E, Genersch E (2018) Swarming motility and biofilm formation of *Paenibacillus larvae*, the etiological agent of American Foulbrood of honey bees (*Apis mellifera*). Sci Rep 8(1):8840. https://doi.org/10.1038/s41598-018-27193-8
39. Wen Y, Wu X, Teng Y, Qian C, Zhan Z, Zhao Y, Li O (2011) Identification and analysis of the gene cluster involved in biosynthesis of paenibactin, a catecholate siderophore produced by *Paenibacillus elgii* B69. Environ Microbiol 13(10):2726–2737. https://doi.org/10.1111/j.1462-2920.2011.02542.x
40. Czech L, Hermann L, Stöveken N, Richter AA, Höppner A, Smits SHJ, Heider J, Bremer E (2018) Role of the extremolytes ectoine and hydroxyectoine as stress protectants and nutrients: genetics, phylogenomics, biochemistry, and structural analysis. Genes 9(4):177. https://doi.org/10.3390/genes9040177
41. Veliz EA, Martínez-Hidalgo P, Hirsch AM (2017) Chitinase-producing bacteria and their role in biocontrol. AIMS Microbiol 3(3):689–705. https://doi.org/10.3934/microbiol.2017.3.689
42. Parkinson JS, Kofoid EC (1992) Communication modules in bacterial signaling proteins. Annu Rev Genet 26(1):71–112. https://doi.org/10.1146/annurev.ge.26.120192.000443
43. Chow V, Nong G, Preston JF (2007) Structure, function, and regulation of the aldouronate utilization gene cluster from *Paenibacillus* sp. strain JDR-2. J Bacteriol 189(24):8863. https://doi.org/10.1128/JB.01141-07
44. Déjean G, Blanvillain-Baufumé S, Boulanger A, Darrasse A, de Bernonville TD, Girard A-L, Carrére S, Jamet S, Zischek C, Lautier M, Solé M, Büttner D, Jacques M-A, Lauber E, Arlat M (2013) The xylan utilization system of the plant pathogen *Xanthomonas campestris* pv *campestris* controls epiphytic life and reveals common features with oligotrophic bacteria and animal gut symbionts. New Phytol 198(3):899–915. https://doi.org/10.1111/nph.12187
45. Poudel S, Giannone RJ, Farmer AT, Campagna SR, Bible AN, Morrell-Falvey JL, Elkins JG, Hettich RL (2019) Integrated proteomics and lipidomics reveal that the swarming motility of Paenibacillus polymyxa is characterized by phospholipid modification, surfactant deployment, and flagellar specialization relative to swimming motility. Front Microbiol. https://doi.org/10.3389/fmicb.2019.02594