

Dark matter of an orchid: metagenome of the microbiome associated with the rhizosphere of *Dactylorhiza traunsteineri*

Gabriel A. Vignolle^{1,3}, Leopold Zehetner¹, Christian Zimmerman¹, Domenico F. Savio¹, Ovidiu Paun², Robert L. Mach¹, Astrid R. Mach-Aigner¹, Julien Charest^{1§}

¹ Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Strasse 1a, 1060 Wien, Austria

² Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Wien, Austria

³ Center for Health & Bioresources, Competence Unit Molecular Diagnostics, AIT Austrian Institute of Technology GmbH, Giefinggasse 4, 1210 Vienna, Austria

Email addresses:

GAV: gabriel.vignolle@ait.ac.at

LZ: leopold.zehetner@gmx.at

CZ: christian.zimmerman@tuwien.ac.at

DFS: domenico.savio@gmx.at

OP: ovidiu.paun@univie.ac.at

RLM: robert.mach@tuwien.ac.at

ARMA: astrid.mach-aigner@tuwien.ac.at

JC: julien.charest@tuwien.ac.at

ABSTRACT

Plant microbiota forms complex co-associations with its host, promoting health in natural environments. Root-associated bacteria (RAB) colonize root compartments and can modulate plant functions by producing bioactive compounds as part of their secondary metabolism. We present an in-depth analysis of the rhizosphere-associated microbiome of the endangered marsh orchid *Dactylorhiza traunsteineri*. Using deep sequencing of 16S rRNA genes, we identified *Proteobacteria*, *Actinobacteria*, *Myxococcota*, *Bacteroidota*, and *Acidobacteria* as predominant phyla associated with the rhizosphere of *D. traunsteineri*. Using deep shotgun metagenomics and de novo assembly, we extracted high-quality metagenome-assembled genomes (MAGs), revealing significant metabolic and biosynthetic potentials of the *D. traunsteineri*'s RABs. Our study offers a comprehensive investigation into the microbial community of the *D. traunsteineri* rhizosphere, highlighting a potential novel source of critical bioactive substances. Our study offers novel insights and a robust platform for future investigations into *D. traunsteineri*'s rhizosphere, crucial for understanding plant-microbe interactions and aiding conservation efforts for endangered orchids.

KEYWORDS

Metagenome, *Dactylorhiza traunsteineri*, Rhizobiome, Holobiont, Secondary metabolites, Biosynthetic gene clusters, Orchid

INTRODUCTION

Every large multicellular organism studied so far, including plants, interacts with a plethora of microorganisms in its environment. The plant microbiota, which includes bacteria, fungi, protists, nematodes, and viruses, can form complex co-associations with its host. Plant microbial communities have been shown to promote plant growth, nutrient uptake, as well as pathogen resistance, thus promoting plant health in its natural environment (Richardson and Simpson, 2011, Pieterse et al., 2014, Trivedi et al.,

2016, Backer et al., 2018, Gouda et al., 2018, Trivedi et al., 2020). However, our understanding of plant-microbial alliances' functional mechanisms and broad ecological implications is still in its early stages (Martin et al., 2017).

Among the bacterial communities interacting with plants, a subset known as the root-associated bacteria (RAB) colonizes root compartments, such as the rhizosphere (i.e. the narrow zone surrounding and influenced by plant roots growth and metabolic activity) and intraradical regions and has the ability to modulate plant functions (Mendes et al., 2013, Qu et al., 2020, Dobbelaere et al., 2003, Walker et al., 2003). It has been suggested that the holobiont, defined as the host and its associated microbiome, forms a discrete symbiotic ecological unit that has co-evolved to maintain host function and fitness (Hassani et al., 2018, Uroz et al., 2019, Vandenkoornhuyse et al., 2015, Trivedi et al., 2020, Liu et al., 2019). Plants selectively recruit RAB depending on different factors such as host genotype, metabolic profiles, root exudates, and soil physicochemical properties (Bulgarelli et al., 2013, Hassani et al., 2018, Liu et al., 2019, Uroz et al., 2019). Although the plant and its exudates mostly influence this microenvironment, environmental influences also significantly shape its composition (Vives-Peris et al., 2020, Najera et al., 2020). Furthermore, diverse endophytic bacteria possess the ability to produce bioactive compounds of usually low molecular mass, with a wide range of reported functions, as part of their secondary metabolism. These compounds, which are not crucial for the fundamental cellular processes of bacteria in controlled laboratory conditions, often confer selective advantages within natural ecosystems. Such secondary metabolites (SMs) include antibiotics, pigments, toxins, enzyme inhibitors, immunomodulators, effectors of ecological competition or symbiosis, as well as compounds with hormonal activity or particular effects on lipids or carbohydrate metabolism (reviewed in (Fouillaud and Dufosse, 2022)). Many RABs have been reported to enhance plant growth through processes such as nutrient acquisition, hormonal regulation, protection against pathogens, and mitigation of abiotic stresses (reviewed and discussed in (Narayanan and Glick, 2022, Glick, 1995, Olanrewaju et al., 2017, Glick, 2012, Hassani et al., 2018, Mendes et al., 2013, Qu et al., 2020, Trivedi et al., 2020, Uroz et al., 2019, Vandenkoornhuyse et al., 2015)). These RAB, conferring growth benefits to plant hosts, are collectively

referred to as plant growth-promoting rhizobacteria (PGPR), which typically belong to the *Proteobacteria*, *Actinobacteria*, *Firmicutes*, and *Bacteroidetes* phyla (Santoyo et al., 2016). Within these phyla, genera such as *Rhizobium*, *Bacillus*, *Pseudomonas*, and *Burkholderia* are prominently recognized as PGPR and have demonstrated functional contributions in mediating seed germination, plant growth and biocontrol of plant disease, notably by conveying secondary metabolites inclusive of antibiotics, volatile compounds, phytohormones and siderophores (Glick, 2012, Kang et al., 2012, Walia et al., 2013, Bloemberg and Lugtenberg, 2001). Notably, the publication of whole-genome sequences of rhizosphere-inhabiting bacteria facilitates the identification of genes involved in the regulation and production of SMs by comparative and functional genomics. However, despite the established knowledge that symbiotic associations between roots and PGPR provide multiple growth benefits to host plants, little is known regarding the role of PGPR and other RAB in shaping plant ecology through their secondary metabolism (Siefert et al., 2019, Kaur and Sharma, 2021).

Orchids, which account for approximately 10% of all seed plants, exhibit intricate associations with diverse microorganisms in their natural habitat (Givnish et al., 2016, Givnish et al., 2015). One example is mycoheterotrophy, an alternative nutritional strategy observed in certain plant species, particularly orchids, which derive sugars and essential nutrients from symbiotic associations with mycorrhizal fungi (Timilsena et al., 2023). The marsh orchid *Dactylorhiza traunsteineri* is an allotetraploid ($2n = 80$), mainly found in the eastern hemisphere (Paun et al., 2010). *Dactylorhiza traunsteineri* is part of the *D. majalis* complex, which has evolved polytopically from unidirectional hybridization between the diploid species *Dactylorhiza fuchsii* and *D. incarnata* (Pillon et al., 2007, Brandrud et al., 2020). It is slender, usually low-grown with narrow leaves and a lax few-flowered inflorescence, has narrow soil moisture and pH tolerance, and grows in calcareous fens, typically on naturally open sites associated with seepage zones (Kull and Hutchings, 2006, Soó, 1980, Baumann and Künkele, 1988, Delforge, 2001). The species occurs in low-nutrient soils, characterized in particular by very low levels of available nitrate and other macro and micronutrients (Wolfe et al., 2023). A recent study documented, at two different localities, a high diversity of fungi associated with the roots of *D.*

traunsteineri in its natural environment, in comparison to its sibling *D. majalis* (Emelianova et al., 2024). Although hybrids and allopolyploids are known to be more resistant to environmental changes due to their larger genomic potential, *D. traunsteineri* is classified as endangered by several organizations in Europe and has proven challenging to grow and maintain under artificial environmental conditions, which is a requirement for ex-situ conservation efforts (Paun et al., 2007, Blinova and Uotila, 2012, Bartók et al., 2018, Paun et al., 2011).

This study presents a comprehensive analysis of the rhizosphere-associated bacterial microbiome of *Dactylorhiza traunsteineri*. Our community analysis, based on 16S ribosomal DNA (rDNA)-targeted sequencing, identified *Proteobacteria*, *Actinobacteria*, *Myxococcota*, *Bacteroidota*, and *Acidobacteria* as the predominant phyla within the *D. traunsteineri* rhizosphere. Leveraging deep shotgun metagenomics and de novo metagenome assembly, we extracted 47 MAGs, unveiling significant metabolic and biosynthetic capabilities of the orchid's RAB. Notably, this study represents the first comprehensive investigation into the microbial community of the *D. traunsteineri* rhizosphere, highlighting a potential novel source of critical bioactive substances. Our findings illuminate previously uncharacterized PGPR SM pathways and the discovery of "dark matter" biosynthetic gene clusters, underscoring the rhizosphere's untapped potential for biotechnological and ecological applications. This research advances our understanding of *D. traunsteineri*'s microbial interactions and sets the stage for future studies to exploit these microbial consortia for conservation and bioprospecting efforts.

MATERIALS & METHODS

Sampling and sample processing

Dactylorhiza traunsteineri individuals (from the *turfosa* subspecies) were isolated from the Kitzbühel alpine town region, in the western Austrian province of Tyrol on the 25th of May 2018. The average temperatures recorded in this region during summer are between 13°C and 26°C, and between 0°C and 6°C during winter, with an annual total rainfall estimated to 1,443 mm, the majority occurring primarily during summer in June (average 188 mm). The individuals were potted and maintained at the

Botanical Garden of the University of Vienna under conditions mimicking those of the sampling region. The sampling of the rhizosphere soil was performed on 03.02.2020 under sterile laboratory conditions. The plantlets were carefully uprooted, the soil discarded by shaking the roots, and the adherent soil directly in contact with the root system collected. Additionally, the roots of the plantlets were washed for 5 min with sterile DNA-free laboratory-grade Sartorius avium® filtered water. Samples were collected in sterile 50 ml Cellstar tubes. Both samples were shock-frozen in liquid nitrogen and kept at -80°C until the metagenomic DNA extraction was performed.

Metagenomic DNA extraction

Isolation of the metagenomic DNA was performed from the rhizosphere soil and the plantlet wash lyophilized samples in a sterile environment. The root wash samples were lyophilized using the FreeZone Freeze Dryer system [LabConco; Cat. 700201000]. 500 mg of lyophilized mass from each sample was transferred to a 2 mL screw cap reaction tube with sterile glass beads. Samples were first disrupted using the FastPrep-24 Classic homogenizer [MPBio; Cat. 116004500] at setting level 6 for 30 seconds. 1 mL of 2xCTAB buffer (pH 8.0) [100 mM Tris (pH 8.0); 20 mM EDTA (pH 8.0); 1.4 M NaCl; 2% (w/v) CTAB] with 4 mL β-mercaptoethanol, was preheated at 55°C and added to the samples, before disrupting the biomass a second time using the FastPrep-24 Classic homogenizer, twice for 30 seconds at setting level 5. The samples were then incubated for 20 minutes at 65°C and centrifuged for 1 minute at 10,000 rpm before transferring the supernatant and foam to fresh, sterilized 2 mL reaction tubes. Next, 400 µL Phenol and 400 µL SEVAG [4% (v/v) isoamyl alcohol; 96% (v/v) chloroform] were added to the samples before vigorous mixing and incubation for 1 hour at room temperature. The aqueous phases were transferred to fresh, sterilized 2 mL reaction tubes, and 800 µL chloroform was added to the samples. The samples were mixed vigorously and centrifuged for 1 minute at 10 000 rpm, before transferring the aqueous phase to fresh sterilized 2 mL reaction tubes. 2 µL RNase A [10 mg/ml [ThermoFisher; Cat: EN0531]] were then added, and the samples mixed by inversion and incubated for 15 minutes at 60°C.

Next, 2 volumes of ice-cold 96% ethanol were added to the samples and incubated overnight at -20°C for RNase A inactivation and DNA precipitation. Following overnight incubation, the samples were centrifuged for 20 minutes at 10,000 rpm at 4°C. The supernatants were discarded, and the DNA pellets were washed once in 500 µL 96% ethanol. The samples were then centrifuged for 10 minutes at 10,000 rpm at 4°C, and the DNA pellets washed twice with 500 µL 70% ethanol. After centrifuging for an additional 10 minutes at 10,000 rpm at 4°C, the DNA pellets were dried and resuspended in 100 µL ultra-pure water and incubated overnight at 4°C, before quality assessment. Samples were stored at 4°C.

16S rDNA amplicon sequencing and analysis

For bacterial community composition analysis, the V3-V4 fragment of the 16S rDNA gene was amplified from the genomic DNA extracted from the soil and root wash samples. 16S rRNA fragment concentrations in the DNA extracts were quantified using domain-specific quantitative PCR (Fontaine et al., 2023). DNA extracts were normalized based on the 16S rRNA gene concentrations to ensure 16S rRNA gene template molarity for amplification and a two-step barcoding procedure. A first amplification reaction was performed using the 16S rDNA 341F and 805R primer set, containing adapters for the introduction of Illumina adapters (341F + adapter: ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTACGGGNGGCWGCAG; 805R + adapter: AGACGTGTGCTCTTCCGATCTGACTACHVGGGTATCTAATCC), using Q5 high-fidelity DNA polymerase [New England Biolabs; Cat: M0491] and the following cycling conditions: 98°C for 1 minute, followed by 20 cycles of 98°C for 10 seconds, 62°C for 30 seconds, 72°C for 30 seconds, and a final extension at 72°C for 2 minutes. The 16S rRNA gene copy concentrations in DNA extracts were determined through quantitative PCR and normalized to equal 16S rRNA fragment copy number to enhance comparability and reduce PCR bias, and purified using the Agencourt AMPure XP [Beckman; Cat: B23319] purification system. Sequencing indexes were introduced in a second amplification reaction with sample-specific barcodes (FWD: AATGATACGGCGACCACCGAGATCTACAC-[index]-ACACTCTTTCCCTACACGACG; REV: CAAGCAGAAGACGGCATAACGAGAT-[index]-

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT) using Q5 high-fidelity DNA polymerase and the following cycling conditions: 98°C for 1 minute, followed by 15 cycles of 98°C for 10 seconds, 66°C for 30 seconds, 72°C for 30 seconds, and a final extension at 72°C for 2 minutes. Amplicon libraries were purified using the Agencourt AMPure XP purification system and quantified using PicoGreen kit [ThermoFisher; Cat: P7589]. The libraries were pooled and diluted to 4 nM before being sequenced for 300 cycles pair-end on the Illumina MiSeq system using the MiSeq® Reagent Kit v2 [Illumina; Cat: MS-102-2002].

Following sequencing, the quality of the 16S amplicon sequences was verified using FastQC (v.0.11.9) (Andrews, 2010) and sequencing adapter sequences, as well as low-quality reads, were removed from the raw read data using Trimmomatic v.0.40 (Bolger et al., 2014) with the following parameters: *ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36*. 16S amplicon sequencing data analysis has been performed using the R package dada2 (v.1.26.0) (Callahan et al., 2016), which includes quality filtering, dereplication, dataset-specific error model determination, ASV interference, chimera removal, as well as taxonomic assignment, as part of its standard workflow.

Shotgun Sequencing

Shotgun metagenomics sequencing was conducted on the genomic DNA extracted from the soil and root samples. Evaluation of the initial DNA concentrations in the prepared samples was carried out using the Qubit dsDNA HS Assay Kit [ThermoFisher; Cat: Q32851]. The libraries were prepared using 1000 ng of genomic DNA, which was fragmented using the Bioruptor sonication device [Diagenode; Cat: B01020001]. Fragments were cleaned up using the GeneJet PCR Purification Kit [ThermoFisher; Cat: #K0701] and size-selected for a 300 bp using Agencourt AMPure XP beads. Sequencing libraries were prepared following the standard NEBNext Ultra II DNA Library Prep with Sample Purification Beads [New England Biolabs; Cat: E7103] procedure. The final library concentrations were determined using the Qubit dsDNA HS Assay Kit and the average size of the library was determined using the Agilent 5200

Fragment Analyzer System [Agilent; Cat: M5310AA]. The libraries were pooled and diluted to 4 nM before being sequenced for 600 cycles pair-end on the Illumina MiSeq system using the MiSeq® Reagent Kit v3 [Illumina; Cat: MS-102-3303].

Metagenome assembly, genome binning and quality assessment

The quality of the metagenomic reads was verified using FastQC (v.0.11.9) (Andrews, 2010) and sequencing adapter sequences, as well as low-quality reads, were removed from the sequenced data using Trimmomatic v.0.40 (Bolger et al., 2014) with the following parameters: *ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36*. The quality trimmed reads from both samples were combined into one large dataset, containing the reads from both samples. The combined dataset was assembled using MEGAHIT (v.1.2.9) (Li et al., 2015) (parameters *--k-min 27 --k-max 247 --k-step 10*). Next, the paired-end metagenomic trimmed reads were mapped back to the generated contigs using bowtie2 (v.2.3.5.1) (Langmead and Salzberg, 2012) to generate coverage information, and sorted using SAMtools (v.1.10). To generate MAGs, contigs were binned using Autometa (v2.0) (Miller et al., 2019). Unclustered contigs were succinctly binned using MetaBAT2 (v.1.7) (Kang et al., 2015). For further improvement, the corresponding reads from each MAGs' contigs were extracted and reassembled using SPAdes (v.3.14.1) (Nurk et al., 2017, Bankevich et al., 2012) using the meta option. Finally, the quality of each MAG was evaluated using CheckM (v.1.1.3) (Parks et al., 2015) and QUAST (v.5.0.2) (Mikheenko et al., 2018), to determine their completeness, contamination and heterogeneity.

MAG analysis and taxonomical classification

For each MAG, ribosomal sequences were extracted based on the KEGG (Kanehisa, 2019, Kanehisa et al., 2023, Kanehisa and Goto, 2000) and PANNZER2 (Toronen et al., 2018) annotations, and quantified by type to identify the most abundant representatives in each MAG. For each binned MAG,

taxonomic and lowest common ancestor predictions were generated using Autometa (Miller et al., 2019), based on the non-redundant NCBI database (O'Leary et al., 2016). Furthermore, a library containing genomic sequences of 199 reference organisms (Supp. Table 1), obtained from the NCBI database, was generated to calculate the average nucleotide identify (ANI) of each MAG (Supp. Figure 1) using FastANI (v.1.33) (Jain et al., 2018).

Gene prediction and annotation

For each generated MAG, genes and proteins were predicted using Prodigal (v.2.6.3) (Hyatt et al., 2010), and the sequences annotated using the KEGG (Kanehisa, 2019, Kanehisa et al., 2023, Kanehisa and Goto, 2000) and PANNZER2 (Toronen et al., 2018) databases. The KEGG annotation was further used to investigate metabolic pathways, whereas PANNZER2 was used to annotate gene ontology and functional domains. Additionally, biosynthetic gene clusters (BGCs) for each MAG were predicted using antiSMASH (v.6.1.1) (Blin et al., 2017, Blin et al., 2021), DeepBGC (v.0.1.31) (Hannigan et al., 2019), and GECCO (v.0.9.10) (Carroll et al., 2021), annotated using the KEGG and PANNZER2 databases and evaluated for completeness.

RESULTS

16S rRNA gene amplicon sequencing/taxonomic identification

To gain insights into the rhizosphere of *D. traunsteineri*, a potted orchid (Individuum 1 from Kitzbühel, sampled on 25.05.2018), maintained in the botanical garden of the University of Vienna (Universität Wien), has been transferred to our laboratory environment (Figure 1). To taxonomically characterize the predominant bacterial and archaeal community inhabiting the rhizosphere of *D. traunsteineri*, we conducted PCR amplification and sequencing of the V3-V4 variable region of the 16S rRNA gene from the soil and root wash samples in technical duplicates. Following paired-end sequencing, a total of 240,935 reads were generated, each with a read length of 300 bp (Table 1). After rigorous processing steps, including trimming, filtering of low-quality reads, merging of forward and reverse

reads, and removal of chimeric sequences, 106,648 sequences were clustered into 2350 amplicon sequence variants (ASV) with taxonomic assignments spanning 31 phyla and over 84 classes (Figure 2). In total, we identified 2349 bacterial and 1 archaeal ASVs. The prevalence of the five major bacterial phyla ranged from 38.43% to 41.33% for *Proteobacteria*, 11.09% to 13.55% for *Actinobacteriota*, 9.40% to 11.90% for *Myxococcota*, 9.72% to 11.37% for *Bacteriodota*, and from 8.72% to 9.26% for *Acidobacteriota* (Table 2). Comparisons of ASV abundances across the four samples using permutational multivariate analysis of variance (PERMANOVA) with the R package *vegan* revealed no significant differences (similarity p-value < 0.09), indicating a consistent sampling and sequencing approach. Encouraged by the stability of the bacterial composition in the soil surrounding the roots of *D. traunsteineri* and the bacteria adhering to the roots themselves (root wash samples), we opted to analyze and process the shotgun metagenomic sequences as a unified sample.

Metagenome assembled genomes (MAGs) generation

To investigate the functional properties of the *D. traunsteineri* rhizobiome, we proceeded to shotgun metagenomic sequencing using short-read technology. We generated 33,043,213 paired-end metagenomic reads from the soil and root wash samples, totalling over 9 Gbp, with each sample ranging from over 1.8 to over 4.9 Gbp (Table 3). After in silico pruning of adapter sequences and low-quality reads, 31,541,795 reads were retained, representing over 95% of the initial dataset. *De novo* metagenomic assembly on the combined datasets using MEGAHIT (Li et al., 2015, Li et al., 2016) resulted in a total of 859,083 contigs, totalling over 7.5 Gbp (Table 4).

The obtained contigs were binned using Autometa (Miller et al., 2019), resulting in 37 predicted MAGs, as well as a group of unclustered contigs, which was further binned using MetaBAT2 (Kang et al., 2015), thus resulting in 10 additional MAGs (Figure 3). Overall, from the 248,298 contigs that were binned as Bacteria, 44,138 contigs could be grouped into 47 MAGs, from which 36 MAGs could successfully be reassembled using SPAdes (Bankevich et al., 2012, Nurk et al., 2017). These MAGs represent 359 Mbp of assembled reads, while 14.86% of all reads mapped back to the sequences

contained in the combined bacterial MAGs (Table 5). MAG size ranges from 361 kbp to 45.81 Gbp, with N50 ranging from 0.9 kbp to 221 kbp (Table 6). The average predicted completeness of the MAGs stands at 33.64%, with a mean coverage spanning from 2.11 to 25.00x. Following a quality assessment, 13 initially generated MAGs were excluded due to either a contamination score exceeding 50% or completeness falling below 10%. The taxonomic distribution of the top-represented groups revealed that 28 MAGs were classified as *Proteobacteria*, eight as *Acidobacteria*, and eight as *Bacteroidetes* (Table 7). Notably, our assembly efforts yielded one high-quality draft genome (MAG_Dt_26), satisfying the stringent Minimum Information about a Metagenome-Assembled Genome (MIMAG) criteria, boasting less than 5% contamination and over 90% completeness, as well as three medium-quality draft genomes (MAG_Dt_01, MAG_Dt_03, MAG_Dt_23), with estimated completeness above 50% with less than 10% estimated contamination (Bowers et al., 2017).

Gene prediction and secondary metabolite potential

Gene prediction was performed using PRODIGAL (Hyatt et al., 2010) for each MAG, and the resulting predicted proteins were annotated using the KEGG (Kanehisa and Goto, 2000, Kanehisa et al., 2023) and PANNZER2 (Toronen et al., 2018) databases for metabolic pathway enrichment (Table 8). Furthermore, each MAG was mined for SM BSGs using the nf-core funscan pipeline (Ewels et al., 2020). A total of 1,741 BGCs were predicted from the *D. traunsteineri* rhizosphere bacterial MAGs (Figure 4; Table 9).

In MAG_Dt_26 (*Proteobacteria*; completeness of 92.03%; contamination of 2.53%), a total of 181 SM BGCs were predicted. These included six nonribosomal peptide (NRP) clusters, 24 polyketide clusters, 22 ribosomally synthesized and post-translationally modified peptide (RiPP) clusters, two RiPP-like clusters, five saccharide clusters, as well as four terpene clusters (Supp. Table 2). Additionally, 101 gene clusters were categorized as “Other”, potentially representing multiple product types and rarer SM classes, while 17 clusters were classified as “Unknown” (Supp. Table 2). From MAG_Dt_01 (*Actinobacteria*; completeness of 80.24%; contamination of 4.58%), a total of nine SM BGCs were

predicted, representing one NRP cluster, one polyketide cluster, two RiPP clusters, two terpene clusters, as well as three “Other” product class clusters (Supp. Table 2). From MAG_Dt_03 (*Acidobacteria*; completeness of 65.94%; contamination of 3.99%), four SM BGCs were predicted, representing one NRP cluster, one terpene cluster, and two “Other” product class clusters (Supp. Table 2). From MAG_Dt_23 (*Proteobacteria*; completeness of 61.01%; contamination of 2.68%), four SM BGCs were predicted, representing two RiPP-like and two terpene clusters (Supp. Table 2).

DISCUSSION

Through comprehensive 16S rDNA profiling, a total of 2,349 bacterial and one archaeal ASVs were identified within the rhizosphere of *D. traunsteineri*. The predominant bacterial phyla detected included *Proteobacteria*, *Actinobacteria*, *Myxococcota*, *Bacteroidota*, and *Acidobacteria*, which collectively accounted for 95% of the sequences identified. Importantly, there were no significant differences in the bacterial community composition between the soil and the root-adhering bacteria associated with *D. traunsteineri*. This consistency underscores the robustness and reproducibility of our sample preparation and sequencing methodologies. Moreover, in our review of existing literature on orchid RAB ecology, we validated that our approach successfully recovered representatives from all phyla previously reported in orchid rhizospheres (Kaur and Sharma, 2021). Notably, the scientific documentation of orchid RAB through culture-independent methods is relatively scarce. Previous studies have predominantly utilized conventional Sanger sequencing methods, while others employing high-throughput sequencing technologies have primarily focused on 16S rDNA profiling of bacterial communities directly extracted from root samples. Nevertheless, the field of orchid RAB research is undergoing a transformative shift with the advent of innovative methodologies. In this context, we introduce a pioneering approach that leverages shotgun metagenomics, followed by the reconstruction of MAGs. This cutting-edge technique promises to offer deeper insights into the complex ecosystem of the orchid rhizobiome, providing unparalleled understanding of its microbial diversity and functional potential. Utilizing shotgun sequencing and *de novo* metagenomic assembly, we successfully obtained 47 bacterial MAGs from the

rhizobiome of *D. traunsteineri*. Our detailed analysis focused on four high-quality MAGs: MAG_Dt_01, MAG_Dt_03, MAG_Dt_23, and MAG_Dt_26. Specifically, MAG_Dt_01 was taxonomically assigned to *Actinobacteria bacterium* (taxid 1883427), MAG_Dt_03 to *Acidobacteria bacterium* (taxid 1978231), and MAG_Dt_23 and MAG_Dt_26 to *Dongia* sp. URHE0060 (taxid 1380391) and *Steroidobacter* (taxid 469322), respectively, both belonging to the phylum *Proteobacteria*.

The phyla *Proteobacteria* and *Actinobacteria*, generally enriched in the rhizosphere, alongside *Acidobacteria*, typically enriched in bulk soil, play pivotal roles in soil carbon and nitrogen cycles as well as the decomposition of organic matter (Eichorst et al., 2018, Song et al., 2021, Ling et al., 2022). The *Actinobacteria* phylum represents one of the largest and most diverse bacterial group, inhabiting soil, water and plant tissue (Diab et al., 2024, Boubekri et al., 2022), and are prolific producers of SMs including well-known antibiotics such as streptomycin, tetracycline, erythromycin, and vancomycin (Berdy, 2005). Beyond antibiotics, they produce a wide array of SMs with potential applications as biopesticides (Barka et al., 2016). For example, polyketides, such as macrolides, exhibit antifungal activity, while non-ribosomal peptides (NRPs), like cyclopeptides, demonstrate insecticidal properties (Gaynor and Mankin, 2003, Quiroz-Carreno et al., 2022). The *Acidobacteria* phylum, although highly prevalent within soil ecosystems, is characterized by its challenging cultivation (Kalam et al., 2020). This phylum fulfills crucial ecological functions through its significant involvement in carbon, nitrogen, and sulphur biogeochemical processes (Ward et al., 2009, Eichorst et al., 2018, Hausmann et al., 2018). Furthermore, *Acidobacteria* possess genes conducive to survival and competitive colonization within the rhizosphere, thereby facilitating the formation of symbiotic relationships with plants (Kalam et al., 2020, Liu et al., 2024). Members of the *Proteobacteria* phylum, such as the *Dongia* and *Steroidobacter* genera, have been frequently isolated from rhizosphere soil and display the potential to degrade complex organic compounds, including various polysaccharides (Ikenaga et al., 2021, Huang et al., 2019, Liu et al., 2024, Baik et al., 2013, Kim et al., 2016, Liu et al., 2010, Palla et al., 2022). The high-quality MAGs obtained in this study present a promising opportunity to elucidate in greater detail the roles of these phyla in plant

microbiome interactions, particularly regarding their secondary metabolite potential and their impact on plant health and development.

Given that our sampling and sequencing efforts were confined to the rhizobiome of a solitary specimen of *D. traunsteineri*, caution must be exercised in generalizing the presence and prevalence of the identified MAGs across all populations of *D. traunsteineri*. However, it is important to note that the sampled individual originated from an extant population and has been maintained under conditions that closely simulate its natural habitat at the Botanical Garden of the University of Vienna. These conditions include soil composition, humidity, and temperature, which are critical for mimicking the plant's natural environment. The controlled yet naturalistic setting in which the specimen was studied provides a valuable context for our findings. Despite the inherent limitations of a single-sample study, our research offers a foundational framework for future interaction studies between RABs and *D. traunsteineri*. This study lays the groundwork for more extensive investigations into the microbiome of this orchid species across different populations and environmental conditions. Furthermore, our findings open new avenues for exploring the complex processes involved in the adaptation, ecology, and evolution of this allopolyploid orchid. Understanding the interactions between *D. traunsteineri* and its associated microbiome can shed light on the mutualistic relationships that facilitate plant health, nutrient acquisition, and stress resilience. Such insights are crucial for conservation strategies, especially for orchids that are often sensitive to environmental changes and habitat disturbances. Additionally, the use of advanced metagenomic techniques in our study underscores the potential for discovering novel microbial taxa and functional genes that contribute to the unique ecology of orchid rhizospheres. This approach not only enhances our understanding of plant-microbe interactions but also provides a deeper comprehension of the evolutionary dynamics within the orchid family. Our study, therefore, serves as a stepping stone for future research aimed at unraveling the intricate web of interactions that sustain and influence the biodiversity and ecological success of *D. traunsteineri*.

CONCLUSION

This study provided a comprehensive overview of the microbiome inhabiting the rhizosphere of the endangered orchid *Dactylorhiza traunsteineri*. By conducting an in-depth bacterial community composition analysis through 16S rDNA-targeted sequencing, we identified *Proteobacteria*, *Actinobacteria*, *Myxococcota*, *Bacteroidota*, and *Acidobacteria* as the most abundant phyla within the *D. traunsteineri* rhizosphere. This foundational data highlights the complexity and diversity of the microbial communities associated with this orchid species. Our investigation extended beyond community profiling to include deep shotgun metagenomics sequencing and subsequent *de novo* metagenome assembly. This approach enabled us to extract high-quality metagenome-assembled genomes (MAGs), providing a rich resource for further functional analysis. Each MAG was meticulously analyzed for metabolic pathway enrichment and secondary metabolite (SM) biosynthetic gene clusters (BGCs), revealing a previously unexplored niche of the rhizobiome. These analyses uncovered potential metabolic capabilities and biosynthetic potentials that are integral to the rhizosphere's ecological dynamics.

The insights gained from this study not only enhance our understanding of the microbial diversity within the *D. traunsteineri* rhizosphere but also lay the groundwork for future research aimed at unravelling the intricate interconnections between the rhizobiome and its host. The high-quality MAGs we have made available serve as a valuable resource for further exploration of microbial functions and their contributions to the health, adaptation, and survival of *D. traunsteineri*. Moreover, our findings underscore the importance of employing advanced metagenomic techniques to study plant-associated microbiomes, particularly for endangered species. By elucidating the functional roles of key microbial players, we can develop more effective conservation strategies and foster a deeper understanding of the symbiotic relationships that underpin plant resilience and ecological success. In conclusion, our study represents a significant step forward in orchid microbiome research, providing novel insights and a robust platform for future investigations into the dynamic interactions within the rhizosphere of *D. traunsteineri*. These efforts are crucial for advancing our knowledge of plant-microbe interactions and supporting the conservation of this and other endangered orchid species.

412 Software

Software	Version	Reference
antiSMASH	v.6.1.1	(Blin et al., 2017, Blin et al., 2021)
Autometa	v.2.0	(Miller et al., 2019)
bowtie2	v.2.3.5.1	(Langmead and Salzberg, 2012)
CheckM	v.1.1.3	(Parks et al., 2015)
dada2 (R package)	v.1.16.0	(Callahan et al., 2016)
DeepBGC	v. 0.1.31	(Hannigan et al., 2019)
FastANI	v.1.33	(Jain et al., 2018)
FastQC	v.0.11.9	(Andrews, 2010)
funcscan	v.1.1.5	(Ewels et al., 2020)
GECCO	v.0.9.10	(Carroll et al., 2021)
ggtree (R package)	v.3.0.4	(Yu, 2020, Yu et al., 2018, Yu et al., 2016)
MEGAHIT	v.1.2.9	(Li et al., 2016, Li et al., 2015)
MetaBAT2	v.17	(Kang et al., 2015)
OrthoFinder	v.2.5.2	(Emms and Kelly, 2019)
Prodigal	v.2.6.3	(Hyatt et al., 2010)
QUAST	v.5.0.2	(Mikheenko et al., 2018)
SAMTools	v.1.10	(Li et al., 2009)
SPAdes	v.3.14.1	(Bankevich et al., 2012, Nurk et al., 2017)
Trimmomatic	v.0.40	(Bolger et al., 2014)

413

414 Abbreviations

415 ANI – average nucleotide identity

416 ASV – amplicon sequence variant

417 BGC – biosynthetic gene cluster

418 MAG – metagenome assembled genome

419 MIMAG - Minimum Information about a Metagenome-Assembled Genome

420 OTU - operational taxonomic unit

421 PERMANOVA - permutational multivariate analysis of variance

422 PGPR - plant growth promoting Rhizobacteria

423 RAB – root associated bacteria

424 SM – secondary metabolite

425

426 **Authors' contribution**

427 GAV, CZ, DFS and OV conceptualized and designed the study. GAV and LZ generated the datasets. GAV
428 and JC performed the data analysis. GAV and JC generated the figures and drafted the manuscript. ARMA
429 and RLM revised and contributed to the manuscript. All authors read and approved the final manuscript.

430

431 **Acknowledgements**

432 We would like to thank David Pressler and the Botanical Garden of the University of Vienna for
433 maintaining the plant in cultivation, and the Tyrol County administration for issuing necessary permits.

434

435 **Funding**

436 This study was supported by a grant from the Austrian Science Fund (FWF): P29556 given to RM, and
437 the PhD program TU Wien bioactive.

438

439 **Availability of data and materials**

440 The datasets supporting the conclusions of this article are publicly available in the National Center for
441 Biotechnology Information (NCBI) repository under BioProject PRJNA1124010.

442

443 **Ethics approval and consent to participate**

444 Not applicable.

445

446 **Consent for publication**

447 Not applicable.

448

Competing interests

The authors declare that they have no competing interests.

Corresponding author

Correspondence to Julien Charest (julien.charest@tuwien.ac.at)

REFERENCES

- ANDREWS, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data* [Online]. Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [Accessed].
- BACKER, R., ROKEM, J. S., ILANGUMARAN, G., LAMONT, J., PRASLICKOVA, D., RICCI, E., SUBRAMANIAN, S. & SMITH, D. L. 2018. Plant Growth-Promoting Rhizobacteria: Context, Mechanisms of Action, and Roadmap to Commercialization of Biostimulants for Sustainable Agriculture. *Front Plant Sci*, 9, 1473.
- BAIK, K. S., HWANG, Y. M., CHOI, J. S., KWON, J. & SEONG, C. N. 2013. *Dongia rigui* sp. nov., isolated from freshwater of a large wetland in Korea. *Antonie Van Leeuwenhoek*, 104, 1143-50.
- BANKEVICH, A., NURK, S., ANTIPOV, D., GUREVICH, A. A., DVORKIN, M., KULIKOV, A. S., LESIN, V. M., NIKOLENKO, S. I., PHAM, S., PRJIBELSKI, A. D., PYSHKIN, A. V., SIROTKIN, A. V., VYAHHI, N., TESLER, G., ALEKSEYEV, M. A. & PEVZNER, P. A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*, 19, 455-77.
- BARKA, E. A., VATSA, P., SANCHEZ, L., GAVEAU-VAILLANT, N., JACQUARD, C., MEIER-KOLTHOFF, J. P., KLENK, H. P., CLEMENT, C., OUHDOUCH, Y. & VAN WEZEL, G. P. 2016. Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol Mol Biol Rev*, 80, 1-43.
- BARTÓK, A., BRĂDEANU, A., DULUGEAC, R., BOBOCEA, M.-M. & ŞESAN, T. E. 2018. *Dactylorhiza traunsteineri* (Saut. ex Reichenb.) Soó Románia flórájában. *Kitaibelia*, 23, 15-24.

474 BAUMANN, H. & KÜNKELE, S. 1988. *Die Orchideen Europas*, Stuttgart, Kosmos Naturführe.

475 BERDY, J. 2005. Bioactive microbial metabolites. *J Antibiot (Tokyo)*, 58, 1-26.

476 BLIN, K., SHAW, S., KLOOSTERMAN, A. M., CHARLOP-POWERS, Z., VAN WEZEL, G. P.,
477 MEDEMA, M. H. & WEBER, T. 2021. antiSMASH 6.0: improving cluster detection and
478 comparison capabilities. *Nucleic Acids Res*, 49, W29-W35.

479 BLIN, K., WOLF, T., CHEVRETTE, M. G., LU, X., SCHWALEN, C. J., KAUTSAR, S. A., SUAREZ
480 DURAN, H. G., DE LOS SANTOS, E. L. C., KIM, H. U., NAVE, M., DICKSCHAT, J. S.,
481 MITCHELL, D. A., SHELEST, E., BREITLING, R., TAKANO, E., LEE, S. Y., WEBER, T. &
482 MEDEMA, M. H. 2017. antiSMASH 4.0-improvements in chemistry prediction and gene cluster
483 boundary identification. *Nucleic Acids Res*, 45, W36-W41.

484 BLINOVA, I. V. & UOTILA, P. 2012. *Dactylorhiza traunsteineri* (Orchidaceae) in Murmansk Region
485 (Russia). *Memoranda Societatis pro Fauna et Flora Fennica*, 88, 67-79.

486 BLOEMBERG, G. V. & LUGTENBERG, B. J. 2001. Molecular basis of plant growth promotion and
487 biocontrol by rhizobacteria. *Curr Opin Plant Biol*, 4, 343-50.

488 BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina
489 sequence data. *Bioinformatics*, 30, 2114-20.

490 BOUBEKRI, K., SOUMARE, A., MARDAD, I., LYAMLOULI, K., OUHDOUCH, Y., HAFIDI, M. &
491 KOUISNI, L. 2022. Multifunctional role of Actinobacteria in agricultural production
492 sustainability: A review. *Microbiol Res*, 261, 127059.

493 BOWERS, R. M., KYRPIDES, N. C., STEPANAUSKAS, R., HARMON-SMITH, M., DOUD, D.,
494 REDDY, T. B. K., SCHULZ, F., JARETT, J., RIVERS, A. R., ELOE-FADROSH, E. A.,
495 TRINGE, S. G., IVANOVA, N. N., COPELAND, A., CLUM, A., BECRAFT, E. D.,
496 MALMSTROM, R. R., BIRREN, B., PODAR, M., BORK, P., WEINSTOCK, G. M., GARRITY,
497 G. M., DODSWORTH, J. A., YOOSEPH, S., SUTTON, G., GLOCKNER, F. O., GILBERT, J.
498 A., NELSON, W. C., HALLAM, S. J., JUNGBLUTH, S. P., ETTEMA, T. J. G., TIGHE, S.,
499 KONSTANTINIDIS, K. T., LIU, W. T., BAKER, B. J., RATTEI, T., EISEN, J. A., HEDLUND,

B., MCMAHON, K. D., FIERER, N., KNIGHT, R., FINN, R., COCHRANE, G., KARSCH-
MIZRACHI, I., TYSON, G. W., RINKE, C., GENOME STANDARDS, C., LAPIDUS, A.,
MEYER, F., YILMAZ, P., PARKS, D. H., EREN, A. M., SCHRIML, L., BANFIELD, J. F.,
HUGENHOLTZ, P. & WOYKE, T. 2017. Minimum information about a single amplified genome
(MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat*
Biotechnol, 35, 725-731.

BRANDRUD, M. K., BAAR, J., LORENZO, M. T., ATHANASIADIS, A., BATEMAN, R. M., CHASE,
M. W., HEDREN, M. & PAUN, O. 2020. Phylogenomic Relationships of Diploids and the
Origins of Allotetraploids in *Dactylorhiza* (Orchidaceae). *Syst Biol*, 69, 91-109.

BULGARELLI, D., SCHLAEPPI, K., SPAEPEN, S., VER LOREN VAN THEMAAT, E. & SCHULZE-
LEFERT, P. 2013. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant*
Biol, 64, 807-38.

CALLAHAN, B. J., MCMURDIE, P. J., ROSEN, M. J., HAN, A. W., JOHNSON, A. J. & HOLMES, S.
P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*,
13, 581-3.

CARROLL, L. M., LARRALDE, M., FLECK, J. S., PONNUDURAI, R., MILANESE, A., CAPPIO, E.
& ZELLER, G. 2021. Accurate de novo identification of biosynthetic gene clusters with GECCO.
bioRxiv, 2021.05.03.442509.

DELFORGE, P. 2001. *Guide des Orchidées d'Europe, d'Afrique du Nord et du Proche-Orient*, Lausanne
& Paris, Delachaux et Niestlé.

DIAB, M. K., MEAD, H. M., AHMAD KHEDR, M. M., ABU-ELSAOUD, A. M. & EL-SHATOURY, S.
A. 2024. Actinomycetes are a natural resource for sustainable pest control and safeguarding
agriculture. *Arch Microbiol*, 206, 268.

DOBBELAERE, S., VANDERLEYDEN, J. & OKON, Y. 2003. Plant Growth-Promoting Effects of
Diazotrophs in the Rhizosphere. *Critical Reviews in Plant Sciences*, 22, 107-149.

- EICHORST, S. A., TROJAN, D., ROUX, S., HERBOLD, C., RATTEI, T. & WOEBKEN, D. 2018. Genomic insights into the Acidobacteria reveal strategies for their success in terrestrial environments. *Environ Microbiol*, 20, 1041-1063.
- EMELIANOVA, K., HAWRANEK, A.-S., ERIKSSON, M. C., WOLFE, T. M. & PAUN, O. 2024. Ecological divergence is marked by species specific plastic gene expression and distinct rhizosphere among sibling allopolyploid marsh orchids (Dactylorhiza). *bioRxiv*, 2024.04.24.590886.
- EMMS, D. M. & KELLY, S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*, 20, 238.
- EWELS, P. A., PELTZER, A., FILLINGER, S., PATEL, H., ALNEBERG, J., WILM, A., GARCIA, M. U., DI TOMMASO, P. & NAHNSEN, S. 2020. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol*, 38, 276-278.
- FONTAINE, L., PIN, L., SAVIO, D., FRIBERG, N., KIRSCHNER, A. K. T., FARNLEITNER, A. H. & EILER, A. 2023. Bacterial bioindicators enable biological status classification along the continental Danube river. *Commun Biol*, 6, 862.
- FOUILLAUD, M. & DUFOSSÉ, L. 2022. Microbial Secondary Metabolism and Biotechnology. *Microorganisms*, 10.
- GAYNOR, M. & MANKIN, A. S. 2003. Macrolide antibiotics: binding site, mechanism of action, resistance. *Curr Top Med Chem*, 3, 949-61.
- GIVNISH, T. J., SPALINK, D., AMES, M., LYON, S. P., HUNTER, S. J., ZULUAGA, A., DOUCETTE, A., CARO, G. G., MCDANIEL, J., CLEMENTS, M. A., ARROYO, M. T. K., ENDARA, L., KRIEBEL, R., WILLIAMS, N. H. & CAMERON, K. M. 2016. Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. *Journal of Biogeography*, 43, 1905-1916.
- GIVNISH, T. J., SPALINK, D., AMES, M., LYON, S. P., HUNTER, S. J., ZULUAGA, A., ILES, W. J., CLEMENTS, M. A., ARROYO, M. T., LEEBENS-MACK, J., ENDARA, L., KRIEBEL, R.,

NEUBIG, K. M., WHITTEN, W. M., WILLIAMS, N. H. & CAMERON, K. M. 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc Biol Sci*, 282.

GLICK, B. R. 1995. The Enhancement of Plant-Growth by Free-Living Bacteria. *Canadian Journal of Microbiology*, 41, 109-117.

GLICK, B. R. 2012. Plant growth-promoting bacteria: mechanisms and applications. *Scientifica (Cairo)*, 2012, 963401.

GOUDA, S., KERRY, R. G., DAS, G., PARAMITHIOTIS, S., SHIN, H. S. & PATRA, J. K. 2018. Revitalization of plant growth promoting rhizobacteria for sustainable development in agriculture. *Microbiol Res*, 206, 131-140.

HANNIGAN, G. D., PRIHODA, D., PALICKA, A., SOUKUP, J., KLEMPER, O., RAMPULA, L., DURCAK, J., WURST, M., KOTOWSKI, J., CHANG, D., WANG, R., PIZZIZZI, G., TEMESI, G., HAZUDA, D. J., WOELK, C. H. & BITTON, D. A. 2019. A deep learning genome-mining strategy for biosynthetic gene cluster prediction. *Nucleic Acids Res*, 47, e110.

HASSANI, M. A., DURAN, P. & HACQUARD, S. 2018. Microbial interactions within the plant holobiont. *Microbiome*, 6, 58.

HAUSMANN, B., PELIKAN, C., HERBOLD, C. W., KOSTLBACHER, S., ALBERTSEN, M., EICHORST, S. A., GLAVINA DEL RIO, T., HUEMER, M., NIELSEN, P. H., RATTEI, T., STINGL, U., TRINGE, S. G., TROJAN, D., WENTRUP, C., WOEBKEN, D., PESTER, M. & LOY, A. 2018. Peatland Acidobacteria with a dissimilatory sulfur metabolism. *ISME J*, 12, 1729-1742.

HUANG, J. W., HU, S. L., CHENG, X. K., CHEN, D., KONG, X. K. & JIANG, J. D. 2019. Steroidobacter soli sp. nov., isolated from farmland soil. *Int J Syst Evol Microbiol*, 69, 3443-3447.

HYATT, D., CHEN, G. L., LOCASCIO, P. F., LAND, M. L., LARIMER, F. W. & HAUSER, L. J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119.

576 IKENAGA, M., KATAOKA, M., YIN, X., MUROUCHI, A. & SAKAI, M. 2021. Characterization and
577 Distribution of Agar-degrading *Steroidobacter agaridevorans* sp. nov., Isolated from Rhizosphere
578 Soils. *Microbes Environ*, 36.

579 JAIN, C., RODRIGUEZ, R. L., PHILLIPPY, A. M., KONSTANTINIDIS, K. T. & ALURU, S. 2018.
580 High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat*
581 *Commun*, 9, 5114.

582 KALAM, S., BASU, A., AHMAD, I., SAYYED, R. Z., EL-ENSHASY, H. A., DAILIN, D. J. &
583 SURIANI, N. L. 2020. Recent Understanding of Soil Acidobacteria and Their Ecological
584 Significance: A Critical Review. *Front Microbiol*, 11, 580024.

585 KANEHISA, M. 2019. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*,
586 28, 1947-1951.

587 KANEHISA, M., FURUMICHI, M., SATO, Y., KAWASHIMA, M. & ISHIGURO-WATANABE, M.
588 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*, 51,
589 D587-D592.

590 KANEHISA, M. & GOTO, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids*
591 *Res*, 28, 27-30.

592 KANG, D. D., FROULA, J., EGAN, R. & WANG, Z. 2015. MetaBAT, an efficient tool for accurately
593 reconstructing single genomes from complex microbial communities. *PeerJ*, 3, e1165.

594 KANG, S. M., KHAN, A. L., HUSSAIN, J., ALI, L., KAMRAN, M., WAQAS, M. & LEE, I. J. 2012.
595 Rhizonin A from *Burkholderia* sp. KCTC11096 and its growth promoting role in lettuce seed
596 germination. *Molecules*, 17, 7980-8.

597 KAUR, J. & SHARMA, J. 2021. Orchid Root Associated Bacteria: Linchpins or Accessories? *Front Plant*
598 *Sci*, 12, 661966.

599 KIM, D. U., LEE, H., KIM, H., KIM, S. G. & KA, J. O. 2016. *Dongia soli* sp. nov., isolated from soil
600 from Dokdo, Korea. *Antonie Van Leeuwenhoek*, 109, 1397-402.

KULL, T. & HUTCHINGS, M. J. 2006. A comparative analysis of decline in the distribution ranges of orchid species in Estonia and the United Kingdom. *Biological Conservation*, 129, 31-39.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9, 357-9.

LI, D., LIU, C. M., LUO, R., SADAKANE, K. & LAM, T. W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31, 1674-6.

LI, D., LUO, R., LIU, C. M., LEUNG, C. M., TING, H. F., SADAKANE, K., YAMASHITA, H. & LAM, T. W. 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102, 3-11.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078-2079.

LING, N., WANG, T. & KUZUYAKOV, Y. 2022. Rhizosphere bacteriome structure and functions. *Nat Commun*, 13, 836.

LIU, F., HEWEZI, T., LEBEIS, S. L., PANTALONE, V., GREWAL, P. S. & STATON, M. E. 2019. Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC Microbiol*, 19, 201.

LIU, Y., JIN, J. H., LIU, Y. H., ZHOU, Y. G. & LIU, Z. P. 2010. *Dongia mobilis* gen. nov., sp. nov., a new member of the family Rhodospirillaceae isolated from a sequencing batch reactor for treatment of malachite green effluent. *Int J Syst Evol Microbiol*, 60, 2780-2785.

LIU, Y., XU, Z., CHEN, L., XUN, W., SHU, X., CHEN, Y., SUN, X., WANG, Z., REN, Y., SHEN, Q. & ZHANG, R. 2024. Root colonization by beneficial rhizobacteria. *FEMS Microbiol Rev*, 48.

MARTIN, F. M., UROZ, S. & BARKER, D. G. 2017. Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science*, 356.

MENDES, R., GARBEVA, P. & RAAIJMAKERS, J. M. 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol Rev*, 37, 634-63.

MIKHEENKO, A., PRJIBELSKI, A., SAVELIEV, V., ANTIPOV, D. & GUREVICH, A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, 34, i142-i150.

MILLER, I. J., REES, E. R., ROSS, J., MILLER, I., BAXA, J., LOPERA, J., KERBY, R. L., REY, F. E. & KWAN, J. C. 2019. Autometa: automated extraction of microbial genomes from individual shotgun metagenomes. *Nucleic Acids Res*, 47, e57.

NAJERA, F., DIPPOLD, M. A., BOY, J., SEGUEL, O., KOESTER, M., STOCK, S., MERINO, C., KUZUYAKOV, Y. & MATUS, F. 2020. Effects of drying/rewetting on soil aggregate dynamics and implications for organic matter turnover. *Biology and Fertility of Soils*, 56, 893-905.

NARAYANAN, Z. & GLICK, B. R. 2022. Secondary Metabolites Produced by Plant Growth-Promoting Bacterial Endophytes. *Microorganisms*, 10.

NURK, S., MELESHKO, D., KOROBAYNIKOV, A. & PEVZNER, P. A. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*, 27, 824-834.

O'LEARY, N. A., WRIGHT, M. W., BRISTER, J. R., CIUFO, S., HADDAD, D., MCVEIGH, R., RAJPUT, B., ROBBERTSE, B., SMITH-WHITE, B., AKO-ADJEI, D., ASTASHYN, A., BADRETDIN, A., BAO, Y., BLINKOVA, O., BROVER, V., CHETVERNIN, V., CHOI, J., COX, E., ERMOLAEVA, O., FARRELL, C. M., GOLDFARB, T., GUPTA, T., HAFT, D., HATCHER, E., HLAVIDA, W., JOARDAR, V. S., KODALI, V. K., LI, W., MAGLOTT, D., MASTERSON, P., MCGARVEY, K. M., MURPHY, M. R., O'NEILL, K., PUJAR, S., RANGWALA, S. H., RAUSCH, D., RIDDICK, L. D., SCHOCH, C., SHKEDA, A., STORZ, S. S., SUN, H., THIBAUD-NISSEN, F., TOLSTOY, I., TULLY, R. E., VATSAN, A. R., WALLIN, C., WEBB, D., WU, W., LANDRUM, M. J., KIMCHI, A., TATUSOVA, T., DICUCCIO, M., KITTS, P., MURPHY, T. D. & PRUITT, K. D. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*, 44, D733-45.

- OLANREWAJU, O. S., GLICK, B. R. & BABALOLA, O. O. 2017. Mechanisms of action of plant growth promoting bacteria. *World J Microbiol Biotechnol*, 33, 197.
- PALLA, M., TURRINI, A., CRISTANI, C., BONORA, L., PELLEGRINI, D., PRIMICERIO, J., GRASSI, A., HILAJ, F., GIOVANNETTI, M. & AGNOLUCCI, M. 2022. Impact of sheep wool residues as soil amendments on olive beneficial symbionts and bacterial diversity. *Bioresour Bioprocess*, 9, 45.
- PARKS, D. H., IMELFORT, M., SKENNERTON, C. T., HUGENHOLTZ, P. & TYSON, G. W. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*, 25, 1043-55.
- PAUN, O., BATEMAN, R. M., FAY, M. F., HEDREN, M., CIVEYREL, L. & CHASE, M. W. 2010. Stable epigenetic effects impact adaptation in allopolyploid orchids (Dactylorhiza: Orchidaceae). *Mol Biol Evol*, 27, 2465-73.
- PAUN, O., BATEMAN, R. M., FAY, M. F., LUNA, J. A., MOAT, J., HEDREN, M. & CHASE, M. W. 2011. Altered gene expression and ecological divergence in sibling allopolyploids of Dactylorhiza (Orchidaceae). *BMC Evol Biol*, 11, 113.
- PAUN, O., FAY, M. F., SOLTIS, D. E. & CHASE, M. W. 2007. Genetic and epigenetic alterations after hybridization and genome doubling. *Taxon*, 56, 649-56.
- PIETERSE, C. M., ZAMIOUDIS, C., BERENDSEN, R. L., WELLER, D. M., VAN WEES, S. C. & BAKKER, P. A. 2014. Induced systemic resistance by beneficial microbes. *Annu Rev Phytopathol*, 52, 347-75.
- PILLON, Y., FAY, M. F., HEDREN, M., BATEMAN, R. M., DEVEY, D. S., SHIPUNOV, A. B., VAN DER BANK, M. & CHASE, M. W. 2007. Evolution and temporal diversification of western European polyploid species complexes in Dactylorhiza (Orchidaceae). *Taxon*, 56, 1185-1208.
- QU, Q., ZHANG, Z., PEIJNENBURG, W., LIU, W., LU, T., HU, B., CHEN, J., CHEN, J., LIN, Z. & QIAN, H. 2020. Rhizosphere Microbiome Assembly and Its Impact on Plant Growth. *J Agric Food Chem*, 68, 5024-5038.

QUIROZ-CARRENO, S., MUNOZ-NUNEZ, E., SILVA, F. L., DEVOTTO-MORENO, L., SEIGLER, D.
S., PASTENE-NAVARRETE, E., CESPEDES-ACUNA, C. L. & ALARCON-ENOS, J. 2022.
Cyclopeptide Alkaloids from *Discaria chacaya* (Rhamnaceae) as Result of Symbiosis with
Frankia (Actinomycetales). *Chem Biodivers*, 19, e202200630.

RICHARDSON, A. E. & SIMPSON, R. J. 2011. Soil microorganisms mediating phosphorus availability
update on microbial phosphorus. *Plant Physiol*, 156, 989-96.

SANTOYO, G., MORENO-HAGELSIEB, G., OROZCO-MOSQUEDA MDEL, C. & GLICK, B. R.
2016. Plant growth-promoting bacterial endophytes. *Microbiol Res*, 183, 92-9.

SIEFERT, A., ZILLIG, K. W., FRIESEN, M. L. & STRAUSS, S. Y. 2019. Mutualists Stabilize the
Coexistence of Congeneric Legumes. *Am Nat*, 193, 200-212.

SONG, D. D., REN, L., LI, X., MA, D. L. & ZANG, S. Y. 2021. Soil Bacterial Diversity and
Composition of Different Forest Types in Greater Xing'an Mountains, China. *Applied Ecology
and Environmental Research*, 19, 1983-1997.

SOÓ, R. D. 1980. *Dactylorhiza*. In: TUTIN, T. G., HEYWOOD, V. H., BURGESS, N. A., MOORE, D. M.,
VALENTINE, D. H., WALTERS, S. M. & WEBB, D. A. (eds.) *Flora Europea*. Cambridge, UK:
Cambridge University Press

TIMILSENA, P. R., BARRETT, C. F., PINEYRO-NELSON, A., WAFULA, E. K., AYYAMPALAYAM,
S., MCNEAL, J. R., YUKAWA, T., GIVNISH, T. J., GRAHAM, S. W., PIRES, J. C., DAVIS, J.
I., ANE, C., STEVENSON, D. W., LEEBENS-MACK, J., MARTINEZ-SALAS, E., ALVAREZ-
BUYLLA, E. R. & DEPAMPHILIS, C. W. 2023. Phylotranscriptomic Analyses of
Mycoheterotrophic Monocots Show a Continuum of Convergent Evolutionary Changes in
Expressed Nuclear Genes From Three Independent Nonphotosynthetic Lineages. *Genome Biol
Evol*, 15.

TORONEN, P., MEDLAR, A. & HOLM, L. 2018. PANNZER2: a rapid functional annotation web server.
Nucleic Acids Res, 46, W84-W88.

704 TRIVEDI, P., LEACH, J. E., TRINGE, S. G., SA, T. & SINGH, B. K. 2020. Plant-microbiome
705 interactions: from community assembly to plant health. *Nat Rev Microbiol*, 18, 607-621.

706 TRIVEDI, P., TRIVEDI, C., GRINYER, J., ANDERSON, I. C. & SINGH, B. K. 2016. Harnessing Host-
707 Vector Microbiome for Sustainable Plant Disease Management of Phloem-Limited Bacteria.
708 *Front Plant Sci*, 7, 1423.

709 UROZ, S., COURTY, P. E. & OGER, P. 2019. Plant Symbionts Are Engineers of the Plant-Associated
710 Microbiome. *Trends Plant Sci*, 24, 905-916.

711 VANDENKOORNHUYSE, P., QUAISER, A., DUHAMEL, M., LE VAN, A. & DUFRESNE, A. 2015.
712 The importance of the microbiome of the plant holobiont. *New Phytol*, 206, 1196-206.

713 VIVES-PERIS, V., DE OLLAS, C., GOMEZ-CADENAS, A. & PEREZ-CLEMENTE, R. M. 2020. Root
714 exudates: from plant to rhizosphere and beyond. *Plant Cell Rep*, 39, 3-17.

715 WALIA, A., MEHTA, P., CHAUHAN, A. & SHIRKOT, C. K. 2013. Effect of *Bacillus subtilis* Strain
716 CKT1 as Inoculum on Growth of Tomato Seedlings Under Net House Conditions. *Proceedings of*
717 *the National Academy of Sciences, India Section B: Biological Sciences*, 84, 145-155.

718 WALKER, T. S., BAIS, H. P., GROTEWOLD, E. & VIVANCO, J. M. 2003. Root exudation and
719 rhizosphere biology. *Plant Physiol*, 132, 44-51.

720 WARD, N. L., CHALLACOMBE, J. F., JANSSEN, P. H., HENRISSAT, B., COUTINHO, P. M., WU, M.,
721 XIE, G., HAFT, D. H., SAIT, M., BADGER, J., BARABOTE, R. D., BRADLEY, B., BRETTIN,
722 T. S., BRINKAC, L. M., BRUCE, D., CREASY, T., DAUGHERTY, S. C., DAVIDSEN, T. M.,
723 DEBOY, R. T., DETTER, J. C., DODSON, R. J., DURKIN, A. S., GANAPATHY, A., GWINN-
724 GIGLIO, M., HAN, C. S., KHOURI, H., KISS, H., KOTHARI, S. P., MADUPU, R., NELSON,
725 K. E., NELSON, W. C., PAULSEN, I., PENN, K., REN, Q., ROSOVITZ, M. J., SELENGUT, J.
726 D., SHRIVASTAVA, S., SULLIVAN, S. A., TAPIA, R., THOMPSON, L. S., WATKINS, K. L.,
727 YANG, Q., YU, C., ZAFAR, N., ZHOU, L. & KUSKE, C. R. 2009. Three genomes from the
728 phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl*
729 *Environ Microbiol*, 75, 2046-56.

WOLFE, T. M., BALAO, F., TRUCCHI, E., BACHMANN, G., GU, W., BAAR, J., HEDREN, M.,
WECKWERTH, W., LEITCH, A. R. & PAUN, O. 2023. Recurrent allopolyploidizations diversify
ecophysiological traits in marsh orchids (*Dactylorhiza majalis* s.l.). *Mol Ecol*, 32, 4777-4790.

YU, G. 2020. Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinformatics*, 69,
e96.

YU, G., LAM, T. T., ZHU, H. & GUAN, Y. 2018. Two Methods for Mapping and Visualizing Associated
Data on Phylogeny Using Ggtree. *Mol Biol Evol*, 35, 3041-3043.

YU, G., SMITH, D. K., ZHU, H., GUAN, Y., LAM, T. T. Y. & MCINERNY, G. 2016. GGTREE:an R
package for visualization and annotation of phylogenetic trees with their covariates and other
associated data. *Methods in Ecology and Evolution*, 8, 28-36.

Table 1 – Sequencing statistics (16S profiling)

Samples	SRA accession no.	No. of raw sequence reads	No. of reads that passed quality control	ASVs after chimera removal
Soil-R1	SRR29434315	36,584	25,981	14,743
Soil-R2	SRR29434314	95,543	68,635	45,317
Wash-R1	SRR29434311	29,562	20,461	11,303
Wash-R2	SRR29434310	79,246	55,090	35,285

Table 2 – ASVs abundance (by Phylum) (%)

	Soil-R1	Soil-R2	Wash-R1	Wash-R2
Proteobacteria	41.33	40.11	39.98	38.43
Actinobacteriota	12.41	11.09	12.20	13.55
Myxococcota	10.15	9.40	11.90	10.84
Bacteroidota	10.40	11.37	10.56	9.72
Acidobacteriota	8.72	8.85	9.02	9.26
Verrucomicrobiota	4.92	5.04	4.29	4.97
Planctomycetota	3.59	4.36	3.64	4.33
Firmicutes	3.34	2.88	3.54	3.40
Chloroflexi	1.82	2.24	1.87	2.03
Cyanobacteria	0.65	0.66	0.70	0.84
Patescibacteria	0.54	1.02	0.40	0.66

Gemmatimonadota	0.24	0.57	0.41	0.48
Bdellovibrionota	0.41	0.70	0.27	0.28
Armatimonadota	0.22	0.48	0.41	0.29
Dependentiae	0.35	0.34	0.17	0.31
Desulfobacterota	0.31	0.15	0.09	0.14
Latescibacterota	0.11	0.13	0.12	0.09
FCPU4260	0.07	0.12	0.12	0.07
Elusimicrobiota	0.12	0.09	0.08	0.06
MBNT15	0	0.10	0.13	0.07
Fibrobacterota	0.01	0.13	0.04	0.01
Spirochaetota	0.03	0.01	0.12	0.04
NB1-j	0.05	0.02	0.04	0.02
WPS-2	0.09	0.01	0	0
Sumerlaeota	0.03	0.03	0	0.02
Nanoarchaeota	0.05	0	0	0
Abditibacteriota	0	0.01	0.02	0.02
WS2	0	0.01	0	0.04
Hydrogenedentes	0	0.02	0	0.01
Deinococcota	0	0.01	0	0
Halanaerobiaeota	0	0.01	0	0

Table 3 – Sequencing statistics (Shotgun sequencing)

Samples	SRA accession no.	No. of raw sequence reads	No. of reads that passed quality control
Soil	SRR29434313	6,252,114	6,050,295
Wash	SRR29434312	10,152,392	9,850,373

Table 4 – Metagenomic assembly statistics

No. of contigs	859,083
Largest contig	368,968 bp
Total length	757,260,241 bp
N ₅₀	829
L ₅₀	252795

Table 5 – Metagenomic binning statistics

	Bacteria	Archaea	Eukaryota	Virus	Plasmid
No. of reads	4,909,848	1,308	23,299	6,397	4,894
Percentage (%)	14.86%	0.004%	0.07%	0.01%	0.01%
No. of contigs	248,298	87	2,289	148	41
Largest contig (bp)	368,968	5,444	23,737	28,310	45,246
Total length (bp)	366,412,546	105,809	2,577,161	376,352	88,614
N ₅₀	1,372	1,155	1,046	3,945	45,246
L ₅₀	64,708	32	860	23	1

Bins (MAGs)	47	0	0	7	41
-------------	----	---	---	---	----

749

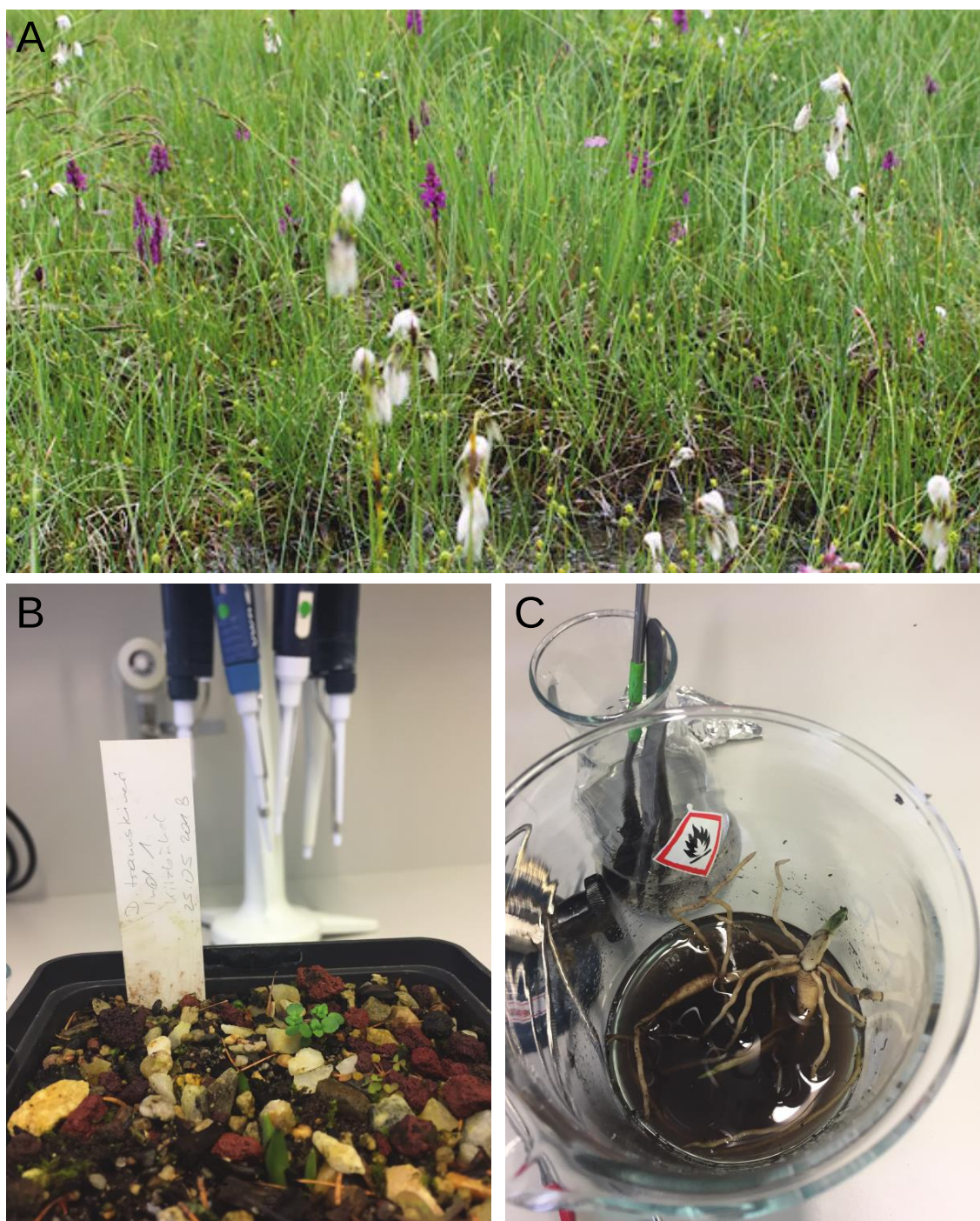


Figure 1 – *Dactylorhiza traunsteineri* sampling. *D. traunsteineri* individuals from the Kitzbühel region (Tyrol, Austria) (A).

D. traunsteineri individuum 1 from Kitzbühel in the laboratory environment (B). Washing of the roots with laboratory-grade

water (C).

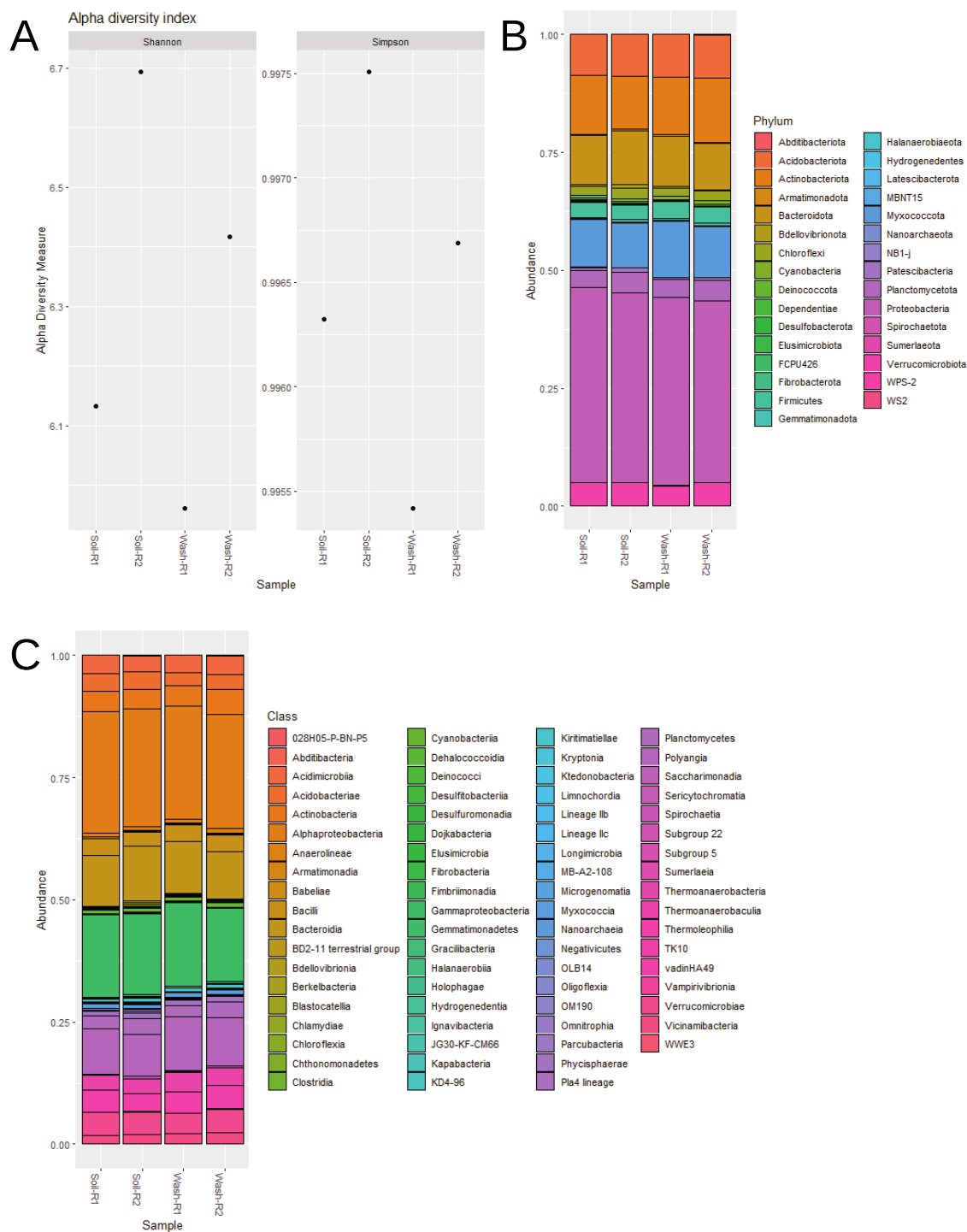


Figure 2 – Bacterial Community Composition Analysis based on 16S rDNA sequencing. Alpha-diversity (Shannon and

Simpson indexes) (A). Relative abundance of identified ASVs at Phylum (B) and Class (C) levels.

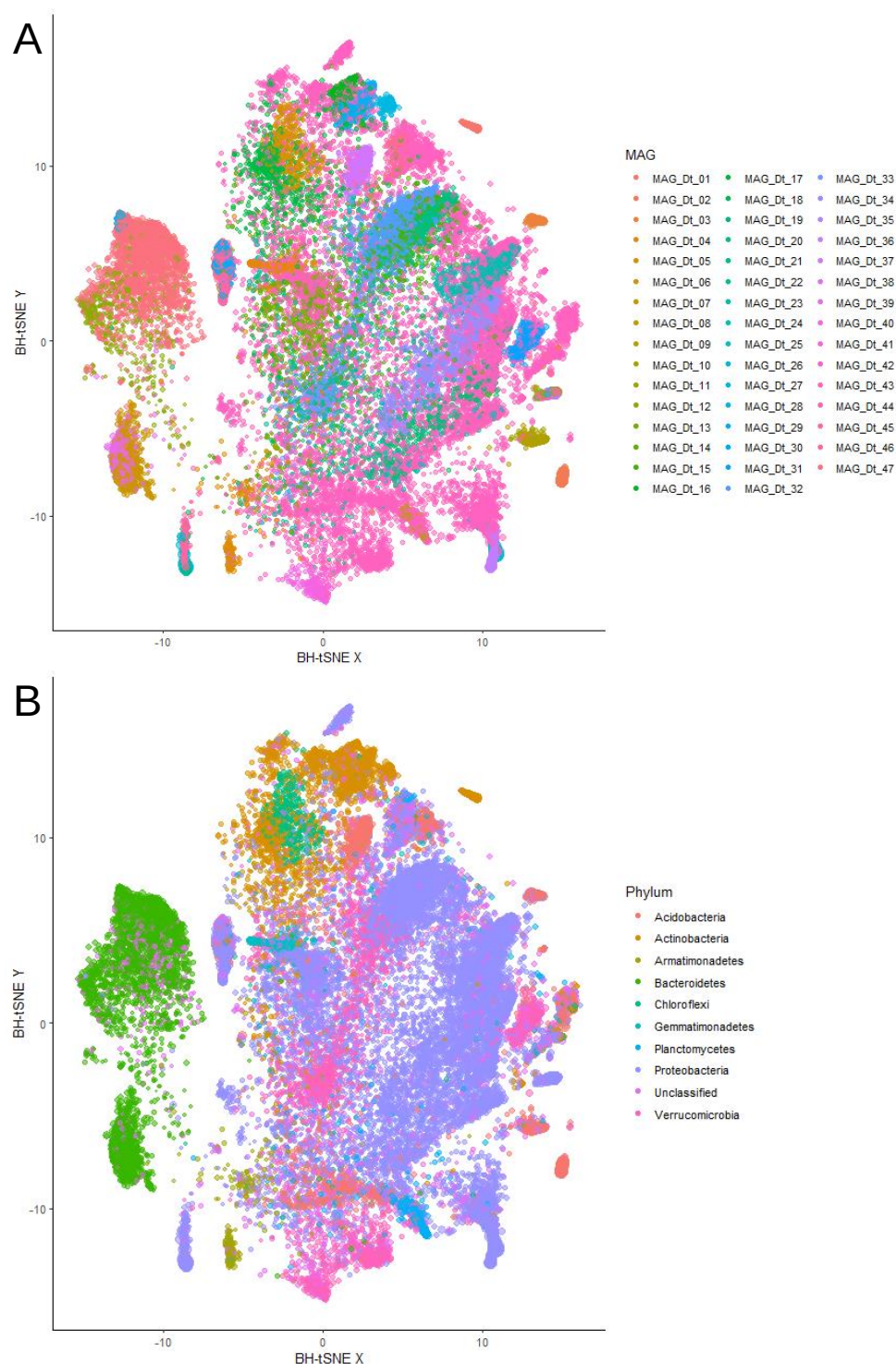


Figure 3 – Barnes-Hut t-SNE representation of MAG contigs. BH-t-SNE representation of MAG contigs coloured by

Figure 4 – Overview of number of predicted SM BGCs per MAG and their product classes.

Dark matter of an orchid: metagenome of the microbiome associated with the rhizosphere of *Dactylorhiza traunsteineri*

Supplementary Information

Gabriel A. Vignolle^{1,3}, Leopold Zehetner¹, Christian Zimmerman¹, Domenico F. Savio¹, Ovidiu Paun², Robert L. Mach¹, Astrid R. Mach-Aigner¹, Julien Charest^{1§}

¹ Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Strasse 1a, 1060 Wien, Austria

² Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Wien, Austria

³ Center for Health & Bioresources, Competence Unit Molecular Diagnostics, AIT Austrian Institute of Technology GmbH, Giefinggasse 4, 1210 Vienna, Austria

Email addresses:

GAV: gabriel.vignolle@ait.ac.at

LZ: leopold.zehetner@gmx.at

CZ: christian.zimmerman@tuwien.ac.at

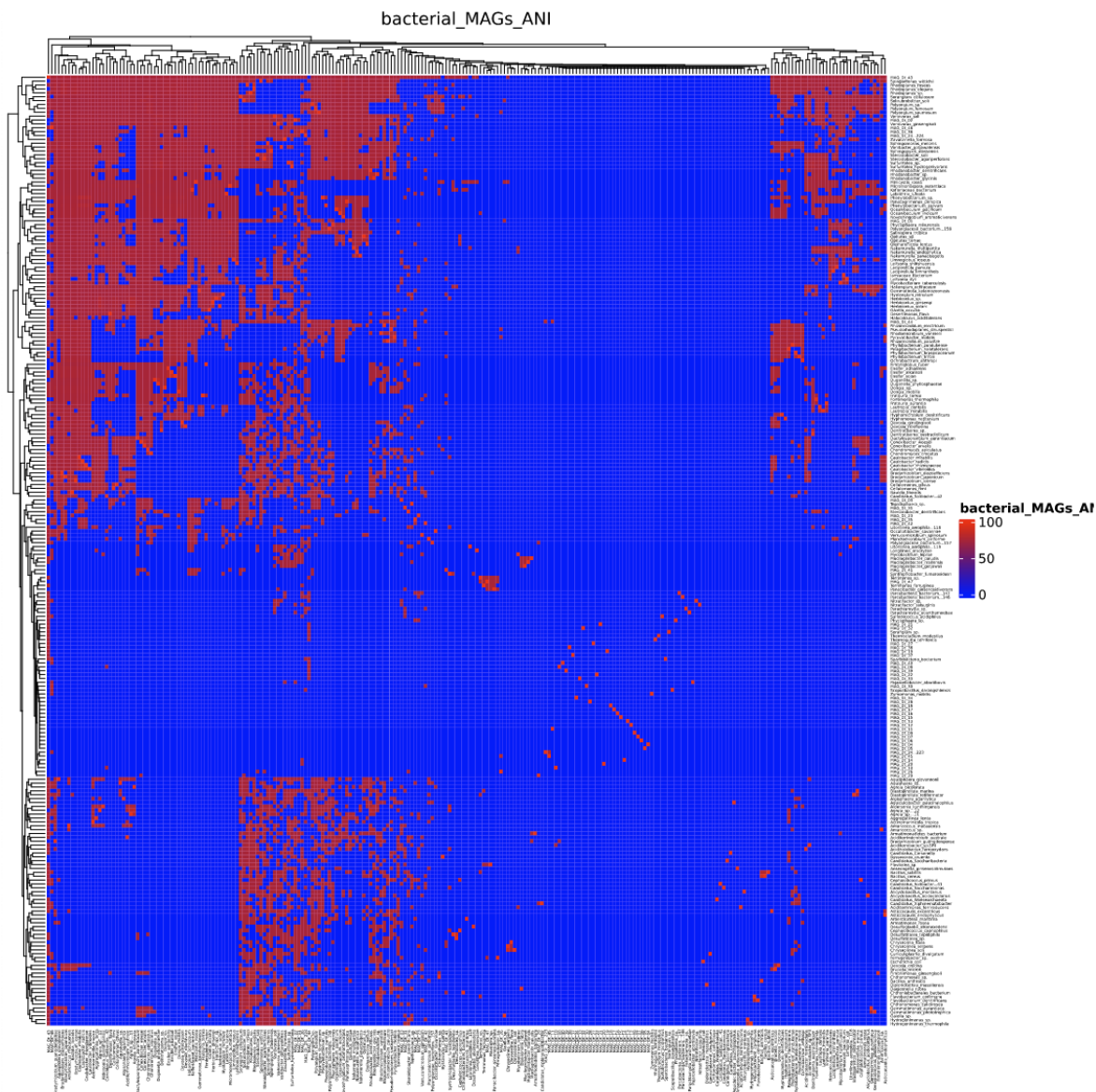
DFS: domenico.savio@gmx.at

OP: ovidiu.paun@univie.ac.at

773 RLM: robert.mach@tuwien.ac.at

774 ARMA: astrid.mach-aigner@tuwien.ac.at

775 JC: julien.charest@tuwien.ac.at



Supplementary Figure 1 – ANI of MAGs and reference organisms.

776

777

Supplementary Table 1 – Genomes used for comparative analysis

Genome	GenBank Assembly	Reference
<i>Acidiferrimicrobium australe</i> strain USS-CCA1	GCA_009697115.1	(Gonzalez et al., 2020)
<i>Acidiferrobacter</i> sp. SPIII/3	GCA_003184265.1	N/A
<i>Acidihalobacter ferrooxydans</i> strain V8	GCA_001975725.1	(Khaleque et al., 2019)
<i>Aciditerrimonas ferrireducens</i> JCM 15389	GCA_001311945.1	(Yamazaki et al., 2010)
<i>Actinomarinicola tropica</i> sp. nov.	GCA_009650215.1	(He et al., 2020b)
<i>Aggregatilinea lenta</i> sp. nov.	GCA_003569045.1	(Nakahara et al., 2019)
<i>Agreia bicolorata</i> strain AC-1804	GCA_000938265.1	(Evtushenko et al., 2001)
<i>Agreia</i> sp. COWG	GCA_904528075.1	N/A
<i>Agreia</i> sp. Leaf244	GCA_001422695.1	(Bai et al., 2015)
<i>Aldersonia kunmingensis</i> strain DSM 45001	GCA_001646865.1	(Sangal et al., 2016)
<i>Algisphaera agarilytica</i> strain DSM 103725	GCA_014207595.1	(Kyrpides et al., 2014)
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> DSM 446	GCA_000024285.1	(Mavromatis et al., 2010)
<i>Alicyclobacillus montanus</i> strain USB-GBX-503	GCA_900142255.1	(Lopez et al., 2018)
<i>Amaricoccus macauensis</i> strain DSM 101730	GCA_014201905.1	(Kyrpides et al., 2014)
<i>Amaricoccus solimangrovi</i> strain HB172011	GCA_006385685.1	(Mo et al., 2020)
<i>Anseongella ginsenosidimutans</i> strain Gsoil 524	GCA_008033235.1	(Siddiqi et al., 2016)
<i>Aquidulcibacter paucihalophilus</i> strain TH1-2	GCA_002105465.1	(Cai et al., 2017)
<i>Aquisphaera insulae</i> strain JC669	GCA_011064615.1	(Kumar et al., 2021)
<i>Aquisphaera giovannonii</i> strain OJF2	GCA_008087625.1	(Bondoso et al., 2011)
<i>Ardenticatena maritima</i> strain 110S	GCA_001306175.1	(Kawaichi et al., 2013)
<i>Armatimonadetes bacterium</i> Uphyl-Ar1	CP021423.1	(Woodhouse et al., 2017)
<i>Armatimonas rosea</i> strain DSM 23562	GCA_014202505.1	(Tamaki et al., 2011)
<i>Asticcacaulis endophyticus</i> strain KCTC 32296	GCA_014652575.1	(Zhu et al., 2014)
<i>Asticcacaulis excentricus</i> CB 48	GCA_000175215.2	(Liu et al., 2005)
<i>Bacillus anthracis</i> strain Ames	GCA_000007845.1	(Read et al., 2003)
<i>Bacillus cereus</i> ATCC 14579	GCA_000007825.1	(Ivanova et al., 2003)
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> strain 168	GCA_000009045.1	(Kunst et al., 1997)
<i>Bauldia litoralis</i> ATCC 35022	GCA_900104485.1	(Yee et al., 2010)
<i>Blastopirellula marina</i> DSM 3645	GCA_000153105.1	(Schlesner et al., 2004)
<i>Blastopirellula retiformator</i>	GCA_007859755.1	(Kallscheuer et al., 2020)
<i>Bradyrhizobium diazoefficiens</i> USDA 110	GCA_000011365.1	(Kaneko et al., 2002)
<i>Bradyrhizobium guangdongense</i> strain CCBAU 51649	GCA_004114975.1	N/A
<i>Bradyrhizobium icense</i> strain LMTR 13	GCA_001693385.1	(Ormeno-Orrillo et al., 2018)
<i>Bradyrhizobium japonicum</i> USDA 6	GCA_000284375.1	(Kaneko et al., 2011)
<i>Brucella microti</i> CCM 4915	GCA_000022745.1	(Audic et al., 2009)
<i>Byssovorax cruenta</i> JCM 12614	GCA_001312805.1	N/A
<i>Candidatus Carsonella ruddii</i> isolate JRPAMB4	GCA_013463375.1	N/A
<i>Candidatus Protochlamydia phocaeensis</i> strain C2	GCA_001545115.1	N/A
<i>Candidatus Saccharibacteria bacterium</i> RAAC3_TM7_1	GCA_000503915.1	(Kantor et al., 2013)
<i>Candidatus Saccharimonas aalborgensis</i> isolate TM71	GCA_000392435.1	(Albertsen et al., 2013)
<i>Candidatus Solibacter</i> sp.	GCA_011375985.1	N/A
<i>Candidatus Solibacter usitatus</i> Ellin6076	GCA_000014905.1	N/A
<i>Candidatus Woesearchaeota archaeon</i>	GCA_016432305.1	(He et al., 2020a)

<i>Candidatus Xiphinematobacter</i> sp. Idaho Grape	GCA_001318295.1	(Brown et al., 2015)
<i>Caulobacter mirabilis</i> strain FWC 38	GCA_002749615.1	N/A
<i>Caulobacter radialis</i> strain 695	GCA_003094615.1	N/A
<i>Caulobacter rhizosphaerae</i> strain KCTC 52515T	GCA_010977555.1	N/A
<i>Caulobacter crescentus</i> CB15	GCA_000006905.1	(Nierman et al., 2001)
<i>Cellulomonas fimi</i> ATCC 484	GCA_000212695.1	N/A
<i>Cellulomonas gilvus</i> ATCC 13127	GCA_000218545.1	(Christopherson et al., 2013)
<i>Cephalotococcus capnophilus</i> strain CV41	GCA_001580045.1	N/A
<i>Cephalotococcus primus</i> strain CAG34	GCA_001580015.1	N/A
<i>Chitinophaga rupis</i> strain DSM 21039	GCA_900109685.1	N/A
<i>Chondromyces apiculatus</i> DSM 436	GCA_000601485.1	N/A
<i>Chondromyces crocatus</i> strain Cm c5	GCA_001189295.1	(Zaburanyi et al., 2016)
<i>Chryseolinea flava</i> strain SDU1-6	GCA_003286915.1	(Wang et al., 2018)
<i>Chryseolinea serpens</i> strain DSM 24574	GCA_900129725.1	N/A
<i>Chryseolinea soli</i> strain KIS68-18	GCA_003589925.1	(Lee et al., 2019b)
<i>Chthoniobacteriales bacterium</i>	GCA_003244145.1	(Ji et al., 2017)
<i>Chthonomonas calidirosea</i> T49	GCA_000427095.1	(Lee et al., 2014)
<i>Chthonomonas</i> sp. UBA2362	GCA_002344135.1	(Parks et al., 2017)
<i>Conexibacter arvalis</i> strain DSM 23288	GCA_014199525.1	N/A
<i>Conexibacter woesei</i> DSM 14684	GCA_000025265.1	(Pukall et al., 2010)
<i>Cuniculiplasma divulgatum</i> strain PM4	GCA_900090055.1	(Golysina et al., 2016)
<i>Dactylosporangium aurantiacum</i> strain NRRL B-8018	GCA_000716715.1	(Doroghazi et al., 2014)
<i>Daejeonella rubra</i> strain DSM 24536	GCA_900103545.1	N/A
<i>Denitratisoma oestradiolicum</i> strain DSM16959	GCA_902813185.1	N/A
<i>Denitratisoma</i> sp. DHT3	GCA_007833355.1	(Wang et al., 2020)
<i>Desertimonas flava</i> strain SYSU D60003	GCA_003426815.1	(Asem et al., 2018)
<i>Desulfatitalea</i> sp. UBA4791	GCA_002402905.1	(Parks et al., 2017)
<i>Desulfatitalea tepidiphila</i> strain S28bF	GCA_001293685.1	(Higashioka et al., 2013)
<i>Desulfoglaeba alkanexedens</i> ALDC	GCA_005377625.1	(Kropp et al., 2000)
<i>Devosia crocina</i> strain IPL20	GCA_900116545.1	N/A
<i>Devosia ginsengisoli</i> strain Gsoil 520	GCA_007859655.1	(Quan et al., 2020)
<i>Devosia riboflavina</i> strain IFO13584	GCA_000743575.1	(Hassan et al., 2014)
<i>Diplorickettsiella massiliensis</i> 20B	GCA_000257395.1	(Mathew et al., 2012)
<i>Dongia mobilis</i> strain CGMCC 1.7660	GCA_004363235.1	(Whitman et al., 2015)
<i>Dongia</i> sp. URHE0060	GCA_000620685.1	N/A
<i>Duganella phyllosphaerae</i> strain T54	GCA_001758785.1	(Haack et al., 2016)
<i>Duganella dendranthematis</i> strain AF9R3	GCA_012849375.1	N/A
<i>Ensifer adhaerens</i> strain Casida A	GCA_000697965.2	(Williams et al., 2017)
<i>Ensifer alkalisoli</i> strain YIC4027	GCA_008932245.1	(Dang et al., 2019)
<i>Ensifer sojae</i> CCBau 05684	GCA_002288525.1	N/A
<i>Escherichia coli</i> str. K-12 substr. MG1655	GCA_000005845.2	(Blattner et al., 1997)
<i>Fimbrioglobus ruber</i> strain SP5	GCA_002197845.1	(Ravin et al., 2018)
<i>Fimbriimonas ginsengisoli</i> Gsoil 348	GCA_000724625.1	(Hu et al., 2014)
<i>Flavitalea</i> sp. OAE831	GCA_014873345.1	N/A
<i>Flavobacterium commune</i> strain PK15	GCA_001857965.1	N/A

<i>Flavobacterium denitrificans</i> DSM 15936	GCA_000425445.1	N/A
<i>Fontimonas thermophila</i> strain DSM 23609	GCA_900113085.1	N/A
<i>Frateuria aurantia</i> DSM 6220	GCA_000242255.3	N/A
<i>Frateuria terrea</i> strain DSM 26515	GCA_900109025.1	N/A
<i>Gaiella occulta</i> strain F2-233	GCA_003351045.1	(Severino et al., 2019)
<i>Gaiella</i> sp. SCGC AG-212-M14	GCA_001644475.1	(Cabello-Yeves et al., 2018)
<i>Gemmatimonas aurantiaca</i> T-27	GCA_000010305.1	(Zhang et al., 2003)
<i>Gemmatimonas phototrophica</i> strain AP64	GCA_000695095.2	(Zeng et al., 2016)
<i>Gemmatirosa kalamazoonesis</i> strain KBS708	GCA_000522985.1	(Debruyne et al., 2014)
<i>Haliangium ochraceum</i> DSM 14365	GCA_000024805.1	(Ivanova et al., 2010)
<i>Halocalculus aciditolerans</i> strain JCM 19596	GCA_014647475.1	(Wu and Ma, 2019)
<i>Hanamia caeni</i> strain BO-59	GCA_003721595.1	(Choi et al., 2022)
<i>Herbiconiux ginsengi</i> CGMCC 4.3491	GCA_900107435.1	N/A
<i>Herbiconiux solani</i> NBRC 106740	GCA_001571005.1	(Behrendt et al., 2011)
<i>Herbiconiux</i> sp. SALV-R1	GCA_013113715.1	N/A
<i>Hyalangium minutum</i> DSM 14724	GCA_000737315.1	N/A
<i>Hydrogenimonas</i> sp.	GCA_003945285.1	(Mino et al., 2018)
<i>Hydrogenimonas thermophila</i> EP1-55-1	GCA_900115615.1	N/A
<i>Hyphomicrobium denitrificans</i> ATCC 51888	GCA_000143145.1	N/A
<i>Hyphomonas neptunium</i> ATCC 15444	GCA_000013025.1	(Badger et al., 2006)
<i>Kofleriaceae bacterium</i> H4 MYX2	GCA_015075455.1	(Liu et al., 2020)
<i>Labilithrix luteola</i> strain DSM 27648	GCA_001263205.1	N/A
<i>Lacipirellula limnantheis</i> strain I41	GCA_007746075.1	(Kallscheuer et al., 2021)
<i>Lacipirellula parvula</i> PX69	GCA_009177095.1	(Dedysh et al., 2020)
<i>Lautropia dentalis</i> KCOM 2505	GCA_003892345.1	(Lim et al., 2019)
<i>Lautropia mirabilis</i> strain NCTC12852	GCA_900637555.1	N/A
<i>Leifsonia shinshuensis</i> strain INR9	GCA_014217625.1	(Benaud et al., 2021)
<i>Leifsonia xyli</i> subsp. cynodontis DSM 46306	GCA_000470775.1	(Monteiro-Vitorello et al., 2013)
<i>Limnoglobus roseus</i> strain PX52	GCA_008254045.1	(Kulichevskaya et al., 2020)
<i>Litorilinea aerophila</i> ATCC BAA-2444	GCA_006569185.1	(Maurais et al., 2022)
<i>Litorilinea aerophila</i> PRI-431	GCA_002148365.1	N/A
<i>Longilinea arvoryzae</i> KOME-1	GCA_001050235.2	(Matsuura et al., 2015)
<i>Micromonospora aurantiaca</i> ATCC 27029	GCA_000145235.1	N/A
<i>Minicystis rosea</i> strain DSM 24000	GCA_001931535.1	(Pal et al., 2021)
<i>Mucilaginibacter gotjawali</i>	GCA_002355435.1	N/A
<i>Mucilaginibacter mallensis</i> strain MP1X4	GCA_900105165.1	N/A
<i>Mucilaginibacter paludis</i> DSM 18603	GCA_000166195.3	N/A
<i>Mycobacterium tuberculosis</i> H37Rv	GCA_000195955.2	(Cole et al., 1998)
<i>Mycobacterium leprae</i> TN	GCA_000195855.1	(Cole et al., 2001)
<i>Nakamurella endophytica</i> CGMCC 4.7308	GCA_014646215.1	N/A
<i>Nakamurella multipartita</i> DSM 44233	GCA_000024365.1	(Tice et al., 2010)
<i>Nakamurella panacisegetis</i> strain P4-7	GCA_900104535.1	N/A
<i>Nitratifactor salsuginis</i> DSM 16511	GCA_000186245.1	(Anderson et al., 2011)
<i>Nitratifactor</i> sp.	GCA_015488535.1	(Reysenbach et al., 2020)
<i>Novosphingobium aromaticivorans</i> DSM 12444	GCA_000013325.1	N/A

<i>Occallatibacter savannae</i> AB23	GCA_003131205.1	N/A
<i>Oceanibaculum indicum</i> USBA 36	GCA_003633955.1	N/A
<i>Oceanibaculum pacificum</i> MCCC 1A02656	GCA_001618175.1	N/A
<i>Ochrobactrum anthropi</i> PBO	GCA_015326295.1	(Gao and Sun, 2021)
<i>Oleiharenicola lentus</i> TWA-58	GCA_004118375.1	(Chen et al., 2020)
<i>Opitutis</i> sp. GAS368	GCA_900104925.1	N/A
<i>Opitutis terrae</i> PB90-1	GCA_000019965.1	(van Passel et al., 2011)
<i>Pajaroellobacter abortibovis</i> strain BTF92-0548A	GCA_001931505.1	(Welly et al., 2017)
<i>Panacagrimonas perspica</i> DSM 26377	GCA_004364935.1	N/A
<i>Panacibacter ginsenosidivorans</i> strain Gsoil1550	GCA_007971225.1	N/A
<i>Parachlamydia acanthamoebae</i> UV-7	GCA_000253035.1	(Collingro et al., 2011)
<i>Parcubacteria bacterium</i> RAAC4_OD1_1	GCA_000505065.1	(Kantor et al., 2013)
<i>Parcubacteria bacterium</i> SCGC AAA011-A08	GCA_000380845.1	(Rinke et al., 2013)
<i>Pelagibacterium halotolerans</i> strain ANSP101	GCA_013004005.1	N/A
<i>Peristeroidobacter agariperforans</i> strain KA5-B	GCA_004138335.1	N/A
<i>Peristeroidobacter soli</i> strain JW-3	GCA_004138485.1	(Huang et al., 2019)
<i>Phenylobacterium parvum</i> strain HYN0004	GCA_003150835.1	(Baek et al., 2019)
<i>Phenylobacterium</i> sp. Root1290	GCA_001425305.1	(Bai et al., 2015)
<i>Phycisphaera mikurensis</i> strain DSM 103959	GCA_014207395.1	N/A
<i>Phycisphaera</i> sp. TMED24	GCA_002168105.2	(Tully et al., 2017)
<i>Phyllobacterium brassicacearum</i> strain 29-15	GCA_004362695.1	N/A
<i>Phyllobacterium trifolii</i> strain CECT 7015	GCA_014192095.1	N/A
<i>Phyllobacterium zundukense</i> strain Tri-48	GCA_002764115.1	(Safronova et al., 2018)
<i>Planctomicrobium piriforme</i> strain DSM 26348	GCA_900113665.1	N/A
<i>Polyangiaceae bacterium</i> isolate DT_87	GCA_013214025.1	(Boeuf et al., 2019)
<i>Polyangiaceae bacterium</i> UTPRO1	GCA_002050235.1	(Lawson et al., 2017)
<i>Polyangium fumosum</i> strain DSM 14668	GCA_005144585.1	N/A
<i>Polyangium aurulentum</i> strain SDU3-1	GCA_005144635.2	N/A
<i>Polyangium spumosum</i> strain DSM 14734	GCA_009649845.1	N/A
<i>Pseudorhodoplanes sinuspersici</i> strain RIPI110	GCA_002119765.1	N/A
<i>Pyruvatibacter mobilis</i> strain CGMCC 1.15125	GCA_014640905.1	N/A
<i>Rhabdotherrmincola sediminis</i> SYSU G02662	GCA_014805525.1	(Liu et al., 2019)
<i>Rhizomicrobium electricum</i> strain DSM 21034	GCA_011762045.1	N/A
<i>Rhizomicrobium palustre</i> strain DSM 19867	GCA_011761565.1	N/A
<i>Rhodanobacter denitrificans</i> strain 2APBS1	GCA_000230695.3	N/A
<i>Rhodanobacter glycinis</i> strain T01E-68	GCA_008000795.1	(Lee et al., 2019a)
<i>Rhodanobacter</i> sp. Root480	GCA_001426645.1	(Bai et al., 2015)
<i>Rhodomicrobium vanniellii</i> ATCC 17100	GCA_000166055.1	(Connors et al., 2021)
<i>Rhodoplanes elegans</i> strain DSM 11907	GCA_003258805.1	(LaSarre et al., 2018)
<i>Rhodoplanes roseus</i> strain DSM 5909	GCA_003258865.1	(LaSarre et al., 2018)
<i>Rhodoplanes</i> sp. Z2-YC6860	GCA_001579845.1	N/A
<i>Salinispora tropica</i> CNB-440	GCA_000016425.1	(Eustaquio et al., 2009)
<i>Scopulibacillus darangshiensis</i> strain DSM 19377	GCA_004341035.1	N/A
<i>Solirubrobacter soli</i> DSM 22325	GCA_000423665.1	N/A
<i>Sorangium cellulosum</i> 'So ce 56'	GCA_000067165.1	(Schneiker et al., 2007)

<i>Sorangium</i> sp. isolate HyVt-558	GCA_011374705.1	(Zhou et al., 2020)
<i>Spartobacteria bacterium</i> AMD-G4	GCA_002290555.1	(Cabello-Yeves et al., 2017)
<i>Sphingomonas melonis</i> strain ZJ26	GCA_002504265.1	N/A
<i>Sphingopyxis alaskensis</i> RB2256	GCA_000013985.1	(Lauro et al., 2009)
<i>Sphingomonas wittichii</i> RW1	GCA_000016765.1	(Miller et al., 2010)
<i>Steroidobacter denitrificans</i> strain DSM 18526	GCA_001579945.1	(Yang et al., 2016)
<i>Sulfodiicoccus acidiphilus</i> strain JCM 31740	GCA_014648235.1	N/A
<i>Sulfuritalea hydrogenivorans</i> sk43H	GCA_000828635.1	(Sakai and Kurosawa, 2019)
<i>Sulfuritalea</i> sp. isolate Bjer_18-Q3-R1-45_MAXAC.014	GCA_016708585.1	(Singleton et al., 2021)
<i>Syntrophobacter fumaroxidans</i> MPOB	GCA_000014965.1	(Plugge et al., 2012)
<i>Tepidisphaera</i> sp. isolate RI_211	GCA_016124375.1	N/A
<i>Terricaulis silvestris</i> strain 0127_4	GCA_009792355.1	(Vieira et al., 2020)
<i>Terrimonas ferruginea</i> DSM 30193	GCA_000425585.1	N/A
<i>Terrimonas</i> sp. strain NS-102	GCA_003075435.1	N/A
<i>Thermocladium modestius</i> strain JCM 10088	GCA_014646535.1	N/A
<i>Thermogutta terrifontis</i> strain R1	GCA_002277955.1	(Elcheninov et al., 2017)
<i>Variibacter gotjawalensis</i> GJW-30	GCA_002355335.1	(Lee et al., 2016)
<i>Variovorax ginsengisoli</i> strain S09.D	GCA_006438845.1	(Altshuler et al., 2019)
<i>Variovorax soli</i> NBRC 106424	GCA_001591385.1	N/A
<i>Verrucomicrobium spinosum</i> DSM 4136	GCA_000172155.1	(Xing et al., 2019)
<i>Zavarzinella formosa</i> DSM 19928	GCA_000255705.1	(Guo et al., 2012)
<i>Zymomonas mobilis</i> subsp. <i>mobilis</i> ZM4	GCA_000007105.1	(Seo et al., 2005)

REFERENCES

- ALBERTSEN, M., HUGENHOLTZ, P., SKARSHEWSKI, A., NIELSEN, K. L., TYSON, G. W. & NIELSEN, P. H. 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol*, 31, 533-8.
- ALTSHULER, I., HAMEL, J., TURNEY, S., MAGNUSON, E., LEVESQUE, R., GREER, C. W. & WHYTE, L. G. 2019. Species interactions and distinct microbial communities in high Arctic permafrost affected cryosols are associated with the CH(4) and CO(2) gas fluxes. *Environ Microbiol*, 21, 3711-3727.
- ANDERSON, I., SIKORSKI, J., ZEYTUN, A., NOLAN, M., LAPIDUS, A., LUCAS, S., HAMMON, N., DESHPANDE, S., CHENG, J. F., TAPIA, R., HAN, C., GOODWIN, L., PITLUCK, S., LIOLIOS, K., PAGANI, I., IVANOVA, N., HUNTEMANN, M., MAVROMATIS, K., OVCHINIKOVA, G., PATI, A., CHEN, A., PALANIAPPAN, K., LAND, M., HAUSER, L.,

BRAMBILLA, E. M., NGATCHOU-DJAO, O. D., ROHDE, M., TINDALL, B. J., GOKER, M.,
 DETTER, J. C., WOYKE, T., BRISTOW, J., EISEN, J. A., MARKOWITZ, V., HUGENHOLTZ,
 P., KLENK, H. P. & KYRPIDES, N. C. 2011. Complete genome sequence of *Nitratifactor*
salsuginis type strain (E9I37-1). *Stand Genomic Sci*, 4, 322-30.

ASEM, M. D., SHI, L., JIAO, J. Y., WANG, D., HAN, M. X., DONG, L., LIU, F., SALAM, N. & LI, W.
 J. 2018. *Desertimonas flava* gen. nov., sp. nov. isolated from a desert soil, and proposal of
Ilumatobacteraceae fam. nov. *Int J Syst Evol Microbiol*, 68, 3593-3599.

AUDIC, S., LESCOT, M., CLAVERIE, J. M. & SCHOLZ, H. C. 2009. *Brucella microti*: the genome
 sequence of an emerging pathogen. *BMC Genomics*, 10, 352.

BADGER, J. H., HOOVER, T. R., BRUN, Y. V., WEINER, R. M., LAUB, M. T., ALEXANDRE, G.,
 MRAZEK, J., REN, Q., PAULSEN, I. T., NELSON, K. E., KHOURI, H. M., RADUNE, D.,
 SOSA, J., DODSON, R. J., SULLIVAN, S. A., ROSOVITZ, M. J., MADUPU, R., BRINKAC, L.
 M., DURKIN, A. S., DAUGHERTY, S. C., KOTHARI, S. P., GIGLIO, M. G., ZHOU, L., HAFT,
 D. H., SELENGUT, J. D., DAVIDSEN, T. M., YANG, Q., ZAFAR, N. & WARD, N. L. 2006.
 Comparative genomic evidence for a close relationship between the dimorphic prosthecate
 bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. *J Bacteriol*, 188, 6841-50.

BAEK, C., SHIN, S. K. & YI, H. 2019. *Phenylobacterium parvum* sp. nov., isolated from lake water. *Int J*
Syst Evol Microbiol, 69, 1169-1172.

BAI, Y., MULLER, D. B., SRINIVAS, G., GARRIDO-OTER, R., POTTHOFF, E., ROTT, M.,
 DOMBROWSKI, N., MUNCH, P. C., SPAEPEN, S., REMUS-EMSERMANN, M., HUTTEL,
 B., MCHARDY, A. C., VORHOLT, J. A. & SCHULZE-LEFERT, P. 2015. Functional overlap of
 the *Arabidopsis* leaf and root microbiota. *Nature*, 528, 364-9.

BEHRENDT, U., SCHUMANN, P., HAMADA, M., SUZUKI, K. I., SPROER, C. & ULRICH, A. 2011.
 Reclassification of *Leifsonia ginsengi* (Qiu et al. 2007) as *Herbiconiux ginsengi* gen. nov., comb.
 nov. and description of *Herbiconiux solani* sp. nov., an actinobacterium associated with the
 phyllosphere of *Solanum tuberosum* L. *Int J Syst Evol Microbiol*, 61, 1039-1047.

BENAUD, N., EDWARDS, R. J., AMOS, T. G., D'AGOSTINO, P. M., GUTIERREZ-CHAVEZ, C.,
MONTGOMERY, K., NICETIC, I. & FERRARI, B. C. 2021. Antarctic desert soil bacteria
exhibit high novel natural product potential, evaluated through long-read genome sequencing and
comparative genomics. *Environ Microbiol*, 23, 3646-3664.

BLATTNER, F. R., PLUNKETT, G., 3RD, BLOCH, C. A., PERNA, N. T., BURLAND, V., RILEY, M.,
COLLADO-VIDES, J., GLASNER, J. D., RODE, C. K., MAYHEW, G. F., GREGOR, J.,
DAVIS, N. W., KIRKPATRICK, H. A., GOEDEN, M. A., ROSE, D. J., MAU, B. & SHAO, Y.
1997. The complete genome sequence of Escherichia coli K-12. *Science*, 277, 1453-62.

BOEUF, D., EDWARDS, B. R., EPPLEY, J. M., HU, S. K., POFF, K. E., ROMANO, A. E., CARON, D.
A., KARL, D. M. & DELONG, E. F. 2019. Biological composition and microbial dynamics of
sinking particulate organic matter at abyssal depths in the oligotrophic open ocean. *Proc Natl
Acad Sci U S A*, 116, 11824-11832.

BONDOSO, J., ALBUQUERQUE, L., NOBRE, M. F., LOBO-DA-CUNHA, A., DA COSTA, M. S. &
LAGE, O. M. 2011. Aquisphaera giovannonii gen. nov., sp. nov., a planctomycete isolated from a
freshwater aquarium. *Int J Syst Evol Microbiol*, 61, 2844-2850.

BROWN, A. M., HOWE, D. K., WASALA, S. K., PEETZ, A. B., ZASADA, I. A. & DENVER, D. R.
2015. Comparative Genomics of a Plant-Parasitic Nematode Endosymbiont Suggest a Role in
Nutritional Symbiosis. *Genome Biol Evol*, 7, 2727-46.

CABELLO-YEVES, P. J., GHAI, R., MEHRSHAD, M., PICAZO, A., CAMACHO, A. & RODRIGUEZ-
VALERA, F. 2017. Reconstruction of Diverse Verrucomicrobial Genomes from Metagenome
Datasets of Freshwater Reservoirs. *Front Microbiol*, 8, 2131.

CABELLO-YEVES, P. J., ZEMSKAYA, T. I., ROSSELLI, R., COUTINHO, F. H., ZAKHARENKO, A.
S., BLINOV, V. V. & RODRIGUEZ-VALERA, F. 2018. Genomes of Novel Microbial Lineages
Assembled from the Sub-Ice Waters of Lake Baikal. *Appl Environ Microbiol*, 84.

841 CAI, H., SHI, Y., WANG, Y., CUI, H. & JIANG, H. 2017. *Aquidulcibacter paucihalophilus* gen. nov., sp.
842 nov., a novel member of family Caulobacteraceae isolated from cyanobacterial aggregates in a
843 eutrophic lake. *Antonie Van Leeuwenhoek*, 110, 1169-1177.

844 CHEN, W. M., CHEN, T. Y., YANG, C. C. & SHEU, S. Y. 2020. *Oleiharenicola lentus* sp. nov., isolated
845 from irrigation water. *Int J Syst Evol Microbiol*, 70, 3440-3448.

846 CHOI, G. M., LIU, Q., LIU, Q., JUN, M. O., CHOI, W. J., YONG KIM, S., WEE, J. H. & IM, W. T.
847 2022. *Hanamia caeni* gen. nov., sp. nov., a Member of the Family Chitinophagaceae Isolated from
848 Activated Sludge in Korea. *Curr Microbiol*, 79, 134.

849 CHRISTOPHERSON, M. R., SUEN, G., BRAMHACHARYA, S., JEWELL, K. A., AYLWARD, F. O.,
850 MEAD, D. & BRUMM, P. J. 2013. The genome sequences of *Cellulomonas fimi* and "*Cellvibrio*
851 *gilvus*" reveal the cellulolytic strategies of two facultative anaerobes, transfer of "*Cellvibrio*
852 *gilvus*" to the genus *Cellulomonas*, and proposal of *Cellulomonas gilvus* sp. nov. *PLoS One*, 8,
853 e53954.

854 COLE, S. T., BROSCHE, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON,
855 S. V., EIGLMEIER, K., GAS, S., BARRY, C. E., 3RD, TEKAIA, F., BADCOCK, K., BASHAM,
856 D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R., DEVLIN, K.,
857 FELTWELL, T., GENTLES, S., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K.,
858 KROGH, A., MCLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., OSBORNE, J., QUAIL,
859 M. A., RAJANDREAM, M. A., ROGERS, J., RUTTER, S., SEEGER, K., SKELTON, J.,
860 SQUARES, R., SQUARES, S., SULSTON, J. E., TAYLOR, K., WHITEHEAD, S. &
861 BARRELL, B. G. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the
862 complete genome sequence. *Nature*, 393, 537-44.

863 COLE, S. T., EIGLMEIER, K., PARKHILL, J., JAMES, K. D., THOMSON, N. R., WHEELER, P. R.,
864 HONORE, N., GARNIER, T., CHURCHER, C., HARRIS, D., MUNGALL, K., BASHAM, D.,
865 BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R. M., DEVLIN, K., DUTHOY,
866 S., FELTWELL, T., FRASER, A., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K.,

867 LACROIX, C., MACLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., QUAIL, M. A.,
868 RAJANDREAM, M. A., RUTHERFORD, K. M., RUTTER, S., SEEGER, K., SIMON, S.,
869 SIMMONDS, M., SKELTON, J., SQUARES, R., SQUARES, S., STEVENS, K., TAYLOR, K.,
870 WHITEHEAD, S., WOODWARD, J. R. & BARRELL, B. G. 2001. Massive gene decay in the
871 leprosy bacillus. *Nature*, 409, 1007-11.

872 COLLINGRO, A., TISCHLER, P., WEINMAIER, T., PENZ, T., HEINZ, E., BRUNHAM, R. C., READ,
873 T. D., BAVOIL, P. M., SACHSE, K., KAHANE, S., FRIEDMAN, M. G., RATTEI, T., MYERS,
874 G. S. & HORN, M. 2011. Unity in variety--the pan-genome of the Chlamydiae. *Mol Biol Evol*,
875 28, 3253-70.

876 CONNERS, E. M., DAVENPORT, E. J. & BOSE, A. 2021. Revised Draft Genome Sequences of
877 Rhodomicrobium vannielii ATCC 17100 and Rhodomicrobium udaipurensis JA643. *Microbiol*
878 *Resour Announc*, 10.

879 DANG, X., XIE, Z., LIU, W., SUN, Y., LIU, X., ZHU, Y. & STAEHELIN, C. 2019. The genome of
880 Ensifer alkanisoli YIC4027 provides insights for host specificity and environmental adaptations.
881 *BMC Genomics*, 20, 643.

882 DEBRUYN, J. M., RADOSEVICH, M., WOMMACK, K. E., POLSON, S. W., HAUSER, L. J., FAWAZ,
883 M. N., KORLACH, J. & TSAI, Y. C. 2014. Genome Sequence and Methylome of Soil Bacterium
884 Gemmatirosa kalamazoonensis KBS708T, a Member of the Rarely Cultivated
885 Gemmatimonadetes Phylum. *Genome Announc*, 2.

886 DEDYSH, S. N., KULICHEVSKAYA, I. S., BELETSKY, A. V., IVANOVA, A. A., RIJPSTRA, W. I. C.,
887 DAMSTE, J. S. S., MARDANOV, A. V. & RAVIN, N. V. 2020. Lacipirellula parvula gen. nov.,
888 sp. nov., representing a lineage of planctomycetes widespread in low-oxygen habitats, description
889 of the family Lacipirellulaceae fam. nov. and proposal of the orders Pirellulales ord. nov.,
890 Gemmatales ord. nov. and Isosphaerales ord. nov. *Syst Appl Microbiol*, 43, 126050.

DOROGHAZI, J. R., ALBRIGHT, J. C., GOERING, A. W., JU, K. S., HAINES, R. R., TCHALUKOV, K. A., LABEDA, D. P., KELLEHER, N. L. & METCALF, W. W. 2014. A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol*, 10, 963-8.

ELCHENINOV, A. G., MENZEL, P., GUDBERGSDOTTIR, S. R., SLESAREV, A. I., KADNIKOV, V. V., KROGH, A., BONCH-OSMOLOVSKAYA, E. A., PENG, X. & KUBLANOV, I. V. 2017. Sugar Metabolism of the First Thermophilic Planctomycete *Thermogutta terrifontis*: Comparative Genomic and Transcriptomic Approaches. *Front Microbiol*, 8, 2140.

EUSTAQUIO, A. S., MCGLINCHEY, R. P., LIU, Y., HAZZARD, C., BEER, L. L., FLOROVA, G., ALHAMADSHEH, M. M., LECHNER, A., KALE, A. J., KOBAYASHI, Y., REYNOLDS, K. A. & MOORE, B. S. 2009. Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine. *Proc Natl Acad Sci U S A*, 106, 12295-300.

EVTUSHENKO, L. I., DOROFEEVA, L. V., DOBROVOLSKAYA, T. G., STRESHINSKAYA, G. M., SUBBOTIN, S. A. & TIEDJE, J. M. 2001. *Agreia bicolorata* gen. nov., sp. nov., to accommodate actinobacteria isolated from narrow reed grass infected by the nematode *Heteroanguina graminophila*. *Int J Syst Evol Microbiol*, 51, 2073-2079.

GAO, R. & SUN, C. 2021. A marine bacterial community capable of degrading poly(ethylene terephthalate) and polyethylene. *J Hazard Mater*, 416, 125928.

GOLYSHINA, O. V., LUNSDORF, H., KUBLANOV, I. V., GOLDENSTEIN, N. I., HINRICHS, K. U. & GOLYSHIN, P. N. 2016. The novel extremely acidophilic, cell-wall-deficient archaeon *Cuniculiplasma divulgatum* gen. nov., sp. nov. represents a new family, Cuniculiplasmataceae fam. nov., of the order Thermoplasmatales. *Int J Syst Evol Microbiol*, 66, 332-340.

GONZALEZ, D., HUBER, K. J., TINDALL, B., HEDRICH, S., ROJAS-VILLALOBOS, C., QUATRINI, R., DINAMARCA, M. A., IBACACHE-QUIROGA, C., SCHWARZ, A., CANALES, C. & NANCUCHEO, I. 2020. *Acidiferrimicrobium australe* gen. nov., sp. nov., an acidophilic and obligately heterotrophic, member of the Actinobacteria that catalyses

917 dissimilatory oxido-reduction of iron isolated from metal-rich acidic water in Chile. *Int J Syst*
918 *Evol Microbiol*, 70, 3348-3354.

919 GUO, M., HAN, X., JIN, T., ZHOU, L., YANG, J., LI, Z., CHEN, J., GENG, B., ZOU, Y., WAN, D., LI,
920 D., DAI, W., WANG, H., CHEN, Y., NI, P., FANG, C. & YANG, R. 2012. Genome sequences of
921 three species in the family Planctomycetaceae. *J Bacteriol*, 194, 3740-1.

922 HAACK, F. S., POEHLEIN, A., KROGER, C., VOIGT, C. A., PIEPENBRING, M., BODE, H. B.,
923 DANIEL, R., SCHAFER, W. & STREIT, W. R. 2016. Molecular Keys to the Janthinobacterium
924 and Duganella spp. Interaction with the Plant Pathogen Fusarium graminearum. *Front Microbiol*,
925 7, 1668.

926 HASSAN, Y. I., LEPP, D., HE, J. & ZHOU, T. 2014. Draft Genome Sequences of Devosia sp. Strain 17-
927 2-E-8 and Devosia riboflavina Strain IFO13584. *Genome Announc*, 2.

928 HE, C., KEREN, R., WHITTAKER, M., FARAG, I. F., DOUDNA, J., CATE, J. H. D. & BANFIELD, J.
929 2020a. Huge and variable diversity of episymbiotic CPR bacteria and DPANN archaea in
930 groundwater ecosystems. *bioRxiv*, 2020.05.14.094862.

931 HE, Y. Q., CHEN, R. W., LI, C., SHI, S. B., CUI, L. Q., LONG, L. J. & TIAN, X. P. 2020b.
932 Actinomarinicola tropica gen. nov. sp. nov., a new marine actinobacterium of the family
933 Iamiaceae, isolated from South China Sea sediment environments. *Int J Syst Evol Microbiol*, 70,
934 3852-3858.

935 HIGASHIOKA, Y., KOJIMA, H., WATANABE, M. & FUKUI, M. 2013. Desulfatitalea tepidiphila gen.
936 nov., sp. nov., a sulfate-reducing bacterium isolated from tidal flat sediment. *Int J Syst Evol*
937 *Microbiol*, 63, 761-765.

938 HU, Z. Y., WANG, Y. Z., IM, W. T., WANG, S. Y., ZHAO, G. P., ZHENG, H. J. & QUAN, Z. X. 2014.
939 The first complete genome sequence of the class Fimbriimonadia in the phylum
940 Armatimonadetes. *PLoS One*, 9, e100794.

941 HUANG, J. W., HU, S. L., CHENG, X. K., CHEN, D., KONG, X. K. & JIANG, J. D. 2019.
942 Steroidobacter soli sp. nov., isolated from farmland soil. *Int J Syst Evol Microbiol*, 69, 3443-3447.

943 IVANOVA, N., DAUM, C., LANG, E., ABT, B., KOPITZ, M., SAUNDERS, E., LAPIDUS, A., LUCAS,
944 S., GLAVINA DEL RIO, T., NOLAN, M., TICE, H., COPELAND, A., CHENG, J. F., CHEN, F.,
945 BRUCE, D., GOODWIN, L., PITLUCK, S., MAVROMATIS, K., PATI, A., MIKHAILOVA, N.,
946 CHEN, A., PALANIAPPAN, K., LAND, M., HAUSER, L., CHANG, Y. J., JEFFRIES, C. D.,
947 DETTER, J. C., BRETTIN, T., ROHDE, M., GOKER, M., BRISTOW, J., MARKOWITZ, V.,
948 EISEN, J. A., HUGENHOLTZ, P., KYRPIDES, N. C. & KLENK, H. P. 2010. Complete genome
949 sequence of *Haliangium ochraceum* type strain (SMP-2). *Stand Genomic Sci*, 2, 96-106.

950 IVANOVA, N., SOROKIN, A., ANDERSON, I., GALLERON, N., CANDELON, B., KAPATRAL, V.,
951 BHATTACHARYYA, A., REZNIK, G., MIKHAILOVA, N., LAPIDUS, A., CHU, L., MAZUR,
952 M., GOLTSMAN, E., LARSEN, N., D'SOUZA, M., WALUNAS, T., GRECHKIN, Y., PUSCH,
953 G., HASELKORN, R., FONSTEIN, M., EHRLICH, S. D., OVERBEEK, R. & KYRPIDES, N.
954 2003. Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*.
955 *Nature*, 423, 87-91.

956 JI, M., GREENING, C., VANWONTERGHEM, I., CARERE, C. R., BAY, S. K., STEEN, J. A.,
957 MONTGOMERY, K., LINES, T., BEARDALL, J., VAN DORST, J., SNAPE, I., STOTT, M. B.,
958 HUGENHOLTZ, P. & FERRARI, B. C. 2017. Atmospheric trace gases support primary
959 production in Antarctic desert surface soil. *Nature*, 552, 400-403.

960 KALLSCHEUER, N., RAST, P., JOGLER, M., WIEGAND, S., KOHN, T., BOEDEKER, C., JESKE, O.,
961 HEUER, A., QUAST, C., GLOCKNER, F. O., ROHDE, M. & JOGLER, C. 2021. Analysis of
962 bacterial communities in a municipal duck pond during a phytoplankton bloom and isolation of
963 *Anatolimnocola aggregata* gen. nov., sp. nov., *Lacipirellula limnantheis* sp. nov. and *Urbifossiella*
964 *limnaea* gen. nov., sp. nov. belonging to the phylum Planctomycetes. *Environ Microbiol*, 23,
965 1379-1396.

966 KALLSCHEUER, N., WIEGAND, S., HEUER, A., RENSINK, S., BOERSMA, A. S., JOGLER, M.,
967 BOEDEKER, C., PEETERS, S. H., RAST, P., JETTEN, M. S. M., ROHDE, M. & JOGLER, C.

2020. *Blastopirellula retiformator* sp. nov. isolated from the shallow-sea hydrothermal vent system close to Panarea Island. *Antonie Van Leeuwenhoek*, 113, 1811-1822.

KANEKO, T., MAITA, H., HIRAKAWA, H., UCHIIKE, N., MINAMISAWA, K., WATANABE, A. & SATO, S. 2011. Complete Genome Sequence of the Soybean Symbiont *Bradyrhizobium japonicum* Strain USDA6T. *Genes (Basel)*, 2, 763-87.

KANEKO, T., NAKAMURA, Y., SATO, S., MINAMISAWA, K., UCHIUMI, T., SASAMOTO, S., WATANABE, A., IDESAWA, K., IRIGUCHI, M., KAWASHIMA, K., KOHARA, M., MATSUMOTO, M., SHIMPO, S., TSURUOKA, H., WADA, T., YAMADA, M. & TABATA, S. 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res*, 9, 189-97.

KANTOR, R. S., WRIGHTON, K. C., HANDLEY, K. M., SHARON, I., HUG, L. A., CASTELLE, C. J., THOMAS, B. C. & BANFIELD, J. F. 2013. Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *mBio*, 4, e00708-13.

KAWAICHI, S., ITO, N., KAMIKAWA, R., SUGAWARA, T., YOSHIDA, T. & SAKO, Y. 2013. *Ardenticatena maritima* gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum 'Chloroflexi' isolated from an iron-rich coastal hydrothermal field, and description of *Ardenticatena* classis nov. *Int J Syst Evol Microbiol*, 63, 2992-3002.

KHALEQUE, H. N., GONZALEZ, C., KAKSONEN, A. H., BOXALL, N. J., HOLMES, D. S. & WATKIN, E. L. J. 2019. Genome-based classification of two halotolerant extreme acidophiles, *Acidihalobacter prosperus* V6 (=DSM 14174 =JCM 32253) and '*Acidihalobacter ferrooxidans*' V8 (=DSM 14175 =JCM 32254) as two new species, *Acidihalobacter aeolianus* sp. nov. and *Acidihalobacter ferrooxydans* sp. nov., respectively. *Int J Syst Evol Microbiol*, 69, 1557-1565.

KROPP, K. G., DAVIDOVA, I. A. & SUFLITA, J. M. 2000. Anaerobic oxidation of n-dodecane by an addition reaction in a sulfate-reducing bacterial enrichment culture. *Appl Environ Microbiol*, 66, 5393-8.

993 KULICHEVSKAYA, I. S., NAUMOFF, D. G., MIROSHNIKOV, K. K., IVANOVA, A. A., PHILIPPOV,
994 D. A., HAKOBYAN, A., RIJPSTRA, W. I. C., DAMSTE, J. S. S., LIESACK, W. & DEDYSH, S.
995 N. 2020. *Limnoglobus roseus* gen. nov., sp. nov., a novel freshwater planctomycete with a giant
996 genome from the family Gemmataceae. *Int J Syst Evol Microbiol*, 70, 1240-1249.

997 KUMAR, G., LHINGJAKIM, K. L., UPPADA, J., AHAMAD, S., KUMAR, D., KASHIF, G. M.,
998 SASIKALA, C. & RAMANA, C. V. 2021. *Aquisphaera insulae* sp. nov., a new member in the
999 family Isosphaeraceae, isolated from the floating island of Loktak lake and emended description
1000 of the genus *Aquisphaera*. *Antonie Van Leeuwenhoek*, 114, 1465-1477.

1001 KUNST, F., OGASAWARA, N., MOSZER, I., ALBERTINI, A. M., ALLONI, G., AZEVEDO, V.,
1002 BERTERO, M. G., BESSIERES, P., BOLOTIN, A., BORCHERT, S., BORRISS, R.,
1003 BOURSIER, L., BRANS, A., BRAUN, M., BRIGNELL, S. C., BRON, S., BROUILLET, S.,
1004 BRUSCHI, C. V., CALDWELL, B., CAPUANO, V., CARTER, N. M., CHOI, S. K., CORDANI,
1005 J. J., CONNERTON, I. F., CUMMINGS, N. J., DANIEL, R. A., DENZIOT, F., DEVINE, K. M.,
1006 DUSTERHOFT, A., EHRLICH, S. D., EMMERSON, P. T., ENTIAN, K. D., ERRINGTON, J.,
1007 FABRET, C., FERRARI, E., FOULGER, D., FRITZ, C., FUJITA, M., FUJITA, Y., FUMA, S.,
1008 GALIZZI, A., GALLERON, N., GHIM, S. Y., GLASER, P., GOFFEAU, A., GOLIGHTLY, E. J.,
1009 GRANDI, G., GUISEPPI, G., GUY, B. J., HAGA, K., HAIECH, J., HARWOOD, C. R.,
1010 HENAUT, A., HILBERT, H., HOLSAPPEL, S., HOSONO, S., HULLO, M. F., ITAYA, M.,
1011 JONES, L., JORIS, B., KARAMATA, D., KASAHARA, Y., KLAERR-BLANCHARD, M.,
1012 KLEIN, C., KOBAYASHI, Y., KOETTER, P., KONINGSTEIN, G., KROGH, S., KUMANO, M.,
1013 KURITA, K., LAPIDUS, A., LARDINOIS, S., LAUBER, J., LAZAREVIC, V., LEE, S. M.,
1014 LEVINE, A., LIU, H., MASUDA, S., MAUEL, C., MEDIGUE, C., MEDINA, N., MELLADO,
1015 R. P., MIZUNO, M., MOESTL, D., NAKAI, S., NOBACK, M., NOONE, D., O'REILLY, M.,
1016 OGAWA, K., OGIWARA, A., OUDEGA, B., PARK, S. H., PARRO, V., POHL, T. M.,
1017 PORTELLE, D., PORWOLLIK, S., PRESCOTT, A. M., PRESECAN, E., PUJIC, P.,

1018 PURNELLE, B., et al. 1997. The complete genome sequence of the gram-positive bacterium
1019 *Bacillus subtilis*. *Nature*, 390, 249-56.

1020 KYRPIDES, N. C., WOYKE, T., EISEN, J. A., GARRITY, G., LILBURN, T. G., BECK, B. J.,
1021 WHITMAN, W. B., HUGENHOLTZ, P. & KLENK, H. P. 2014. Genomic Encyclopedia of Type
1022 Strains, Phase I: The one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci*, 9,
1023 1278-84.

1024 LASARRE, B., KYSELA, D. T., STEIN, B. D., DUCRET, A., BRUN, Y. V. & MCKINLAY, J. B. 2018.
1025 Restricted Localization of Photosynthetic Intracytoplasmic Membranes (ICMs) in Multiple
1026 Genera of Purple Nonsulfur Bacteria. *mBio*, 9.

1027 LAURO, F. M., MCDOUGALD, D., THOMAS, T., WILLIAMS, T. J., EGAN, S., RICE, S., DEMAERE,
1028 M. Z., TING, L., ERTAN, H., JOHNSON, J., FERRIERA, S., LAPIDUS, A., ANDERSON, I.,
1029 KYRPIDES, N., MUNK, A. C., DETTER, C., HAN, C. S., BROWN, M. V., ROBB, F. T.,
1030 KJELLEBERG, S. & CAVICCHIOLI, R. 2009. The genomic basis of trophic strategy in marine
1031 bacteria. *Proc Natl Acad Sci U S A*, 106, 15527-33.

1032 LAWSON, C. E., WU, S., BHATTACHARJEE, A. S., HAMILTON, J. J., MCMAHON, K. D., GOEL, R.
1033 & NOGUERA, D. R. 2017. Metabolic network analysis reveals microbial community interactions
1034 in anammox granules. *Nat Commun*, 8, 15416.

1035 LEE, J. S., LEE, K. C., KIM, K. K. & LEE, B. 2016. Complete genome sequence of the *Variibacter*
1036 *gotjawalensis* GJW-30(T) from soil of lava forest, Gotjawal. *J Biotechnol*, 218, 64-5.

1037 LEE, K. C., MORGAN, X. C., DUNFIELD, P. F., TAMAS, I., MCDONALD, I. R. & STOTT, M. B.
1038 2014. Genomic analysis of *Chthonomonas calidirosea*, the first sequenced isolate of the phylum
1039 *Armatimonadetes*. *ISME J*, 8, 1522-33.

1040 LEE, S. A., KANTH, B. K., KIM, H. S., KIM, T. W., SANG, M. K., SONG, J. & WEON, H. Y. 2019a.
1041 Complete genome sequence of the plant growth-promoting endophytic bacterium *Rhodanobacter*
1042 *glycinis* T01E-68 isolated from tomato (*Solanum lycopersicum* L.) plant roots. *Korean J.*
1043 *Microbiol.*, 55, 422-424.

1044 LEE, S. A., KIM, Y., SANG, M. K., SONG, J., KWON, S. W. & WEON, H. Y. 2019b. *Chryseolinea soli*
1045 sp. nov., isolated from soil. *J Microbiol*, 57, 122-126.

1046 LIM, Y. K., PARK, S. N., LEE, W. P., SHIN, J. H., JO, E., SHIN, Y., PAEK, J., CHANG, Y. H., KIM, H.
1047 & KOOK, J. K. 2019. *Lautropia dentalis* sp. nov., Isolated from Human Dental Plaque of a
1048 Gingivitis Lesion. *Curr Microbiol*, 76, 1369-1373.

1049 LIU, L., WANG, Y., CHE, Y., CHEN, Y., XIA, Y., LUO, R., CHENG, S. H., ZHENG, C. & ZHANG, T.
1050 2020. High-quality bacterial genomes of a partial-nitrification/anammox system by an iterative
1051 hybrid assembly method. *Microbiome*, 8, 155.

1052 LIU, Z. P., WANG, B. J., LIU, S. J. & LIU, Y. H. 2005. *Asticcacaulis taihuensis* sp. nov., a novel stalked
1053 bacterium isolated from Taihu Lake, China. *Int J Syst Evol Microbiol*, 55, 1239-1242.

1054 LIU, Z. T., JIAO, J. Y., LIU, L., LI, M. M., MING, Y. Z., SONG, J. L., LV, A. P., XIAN, W. D., FANG, B.
1055 Z. & LI, W. J. 2019. *Rhabdothermincola sediminis* gen. nov., sp. nov., a new actinobacterium
1056 isolated from hot spring sediment, and emended description of the family Iamiaceae. *Int J Syst*
1057 *Evol Microbiol*, 71.

1058 LOPEZ, G., DIAZ-CARDENAS, C., DAVID ALZATE, J., GONZALEZ, L. N., SHAPIRO, N., WOYKE,
1059 T., KYRPIDES, N. C., RESTREPO, S. & BAENA, S. 2018. Description of *Alicyclobacillus*
1060 *montanus* sp. nov., a mixotrophic bacterium isolated from acidic hot springs. *Int J Syst Evol*
1061 *Microbiol*, 68, 1608-1615.

1062 MATHEW, M. J., SUBRAMANIAN, G., NGUYEN, T. T., ROBERT, C., MEDIANNIKOV, O.,
1063 FOURNIER, P. E. & RAOULT, D. 2012. Genome sequence of *Diplorickettsia massiliensis*, an
1064 emerging *Ixodes ricinus*-associated human pathogen. *J Bacteriol*, 194, 3287.

1065 MATSUURA, N., TOURLOUSSE, D. M., OHASHI, A., HUGENHOLTZ, P. & SEKIGUCHI, Y. 2015.
1066 Draft Genome Sequences of *Anaerolinea thermolimosa* IMO-1, *Bellilinea caldifistulae* GOMI-1,
1067 *Leptolinea tardivitalis* YMTK-2, *Levilinea saccharolytica* KIBI-1, *Longilinea arvoryzae* KOME-
1068 1, Previously Described as Members of the Class Anaerolineae (Chloroflexi). *Genome Announc*,
1069 3.

1070 MAURAS, E. G., IANNAZZI, L. C. & MACLEA, K. S. 2022. Genome Sequence of Litorilinea
1071 aerophila, an Icelandic Intertidal Hot Springs Bacterium. *Microbiol Resour Announc*, 11,
1072 e0120621.

1073 MAVROMATIS, K., SIKORSKI, J., LAPIDUS, A., GLAVINA DEL RIO, T., COPELAND, A., TICE, H.,
1074 CHENG, J. F., LUCAS, S., CHEN, F., NOLAN, M., BRUCE, D., GOODWIN, L., PITLUCK, S.,
1075 IVANOVA, N., OVCHINNIKOVA, G., PATI, A., CHEN, A., PALANIAPPAN, K., LAND, M.,
1076 HAUSER, L., CHANG, Y. J., JEFFRIES, C. D., CHAIN, P., MEINCKE, L., SIMS, D.,
1077 CHERTKOV, O., HAN, C., BRETTIN, T., DETTER, J. C., WAHRENBURG, C., ROHDE, M.,
1078 PUKALL, R., GOKER, M., BRISTOW, J., EISEN, J. A., MARKOWITZ, V., HUGENHOLTZ,
1079 P., KLENK, H. P. & KYRPIDES, N. C. 2010. Complete genome sequence of Alicyclobacillus
1080 acidocaldarius type strain (104-IA). *Stand Genomic Sci*, 2, 9-18.

1081 MILLER, T. R., DELCHER, A. L., SALZBERG, S. L., SAUNDERS, E., DETTER, J. C. & HALDEN, R.
1082 U. 2010. Genome sequence of the dioxin-mineralizing bacterium Sphingomonas wittichii RW1. *J*
1083 *Bacteriol*, 192, 6101-2.

1084 MINO, S., YONEYAMA, N., NAKAGAWA, S., TAKAI, K. & SAWABE, T. 2018. Enrichment and
1085 Genomic Characterization of a N(2)O-Reducing Chemolithoautotroph From a Deep-Sea
1086 Hydrothermal Vent. *Front Bioeng Biotechnol*, 6, 184.

1087 MO, K. L., HUANG, H. Q., WU, Q. J. & HU, Y. H. 2020. Amaricoccus solimangrovi sp. nov., isolated
1088 from mangrove soil. *Int J Syst Evol Microbiol*, 70, 5389-5393.

1089 MONTEIRO-VITORELLO, C. B., ZERILLO, M. M., VAN SLUYS, M. A., CAMARGO, L. E. &
1090 KITAJIMA, J. P. 2013. Complete Genome Sequence of Leifsonia xyli subsp. cynodontis Strain
1091 DSM46306, a Gram-Positive Bacterial Pathogen of Grasses. *Genome Announc*, 1.

1092 NAKAHARA, N., NOBU, M. K., TAKAKI, Y., MIYAZAKI, M., TASUMI, E., SAKAI, S., OGAWARA,
1093 M., YOSHIDA, N., TAMAKI, H., YAMANAKA, Y., KATAYAMA, A., YAMAGUCHI, T.,
1094 TAKAI, K. & IMACHI, H. 2019. Aggregatilinea lenta gen. nov., sp. nov., a slow-growing,
1095 facultatively anaerobic bacterium isolated from subseafloor sediment, and proposal of the new

1096 order Aggregatilineales ord. nov. within the class Anaerolineae of the phylum Chloroflexi. *Int J*
1097 *Syst Evol Microbiol*, 69, 1185-1194.

1098 NIERMAN, W. C., FELDBLYUM, T. V., LAUB, M. T., PAULSEN, I. T., NELSON, K. E., EISEN, J. A.,
1099 HEIDELBERG, J. F., ALLEY, M. R., OHTA, N., MADDOCK, J. R., POTOCKA, I., NELSON,
1100 W. C., NEWTON, A., STEPHENS, C., PHADKE, N. D., ELY, B., DEBOY, R. T., DODSON, R.
1101 J., DURKIN, A. S., GWINN, M. L., HAFT, D. H., KOLONAY, J. F., SMIT, J., CRAVEN, M. B.,
1102 KHOURI, H., SHETTY, J., BERRY, K., UTTERBACK, T., TRAN, K., WOLF, A.,
1103 VAMATHEVAN, J., ERMOLAEVA, M., WHITE, O., SALZBERG, S. L., VENTER, J. C.,
1104 SHAPIRO, L. & FRASER, C. M. 2001. Complete genome sequence of *Caulobacter crescentus*.
1105 *Proc Natl Acad Sci U S A*, 98, 4136-41.

1106 ORMENO-ORRILLO, E., ROGEL, M. A., ZUNIGA-DAVILA, D. & MARTINEZ-ROMERO, E. 2018.
1107 Complete Genome Sequence of the Symbiotic Strain *Bradyrhizobium icense* LMTR 13(T),
1108 Isolated from Lima Bean (*Phaseolus lunatus*) in Peru. *Genome Announc*, 6.

1109 PAL, S., SHARMA, G. & SUBRAMANIAN, S. 2021. Complete genome sequence and identification of
1110 polyunsaturated fatty acid biosynthesis genes of the myxobacterium *Minicystis rosea* DSM
1111 24000(T). *BMC Genomics*, 22, 655.

1112 PARKS, D. H., RINKE, C., CHUVOCHINA, M., CHAUMEIL, P. A., WOODCROFT, B. J., EVANS, P.
1113 N., HUGENHOLTZ, P. & TYSON, G. W. 2017. Recovery of nearly 8,000 metagenome-
1114 assembled genomes substantially expands the tree of life. *Nat Microbiol*, 2, 1533-1542.

1115 PLUGGE, C. M., HENSTRA, A. M., WORM, P., SWARTS, D. C., PAULITSCH-FUCHS, A. H.,
1116 SCHOLTEN, J. C., LYKIDIS, A., LAPIDUS, A. L., GOLTSMAN, E., KIM, E., MCDONALD,
1117 E., ROHLIN, L., CRABLE, B. R., GUNSALUS, R. P., STAMS, A. J. & MCINERNEY, M. J.
1118 2012. Complete genome sequence of *Syntrophobacter fumaroxidans* strain (MPOB(T)). *Stand*
1119 *Genomic Sci*, 7, 91-106.

1120 PUKALL, R., LAPIDUS, A., GLAVINA DEL RIO, T., COPELAND, A., TICE, H., CHENG, J. F.,
1121 LUCAS, S., CHEN, F., NOLAN, M., BRUCE, D., GOODWIN, L., PITLUCK, S.,

1122 MAVROMATIS, K., IVANOVA, N., OVCHINNIKOVA, G., PATI, A., CHEN, A.,
1123 PALANIAPPAN, K., LAND, M., HAUSER, L., CHANG, Y. J., JEFFRIES, C. D., CHAIN, P.,
1124 MEINCKE, L., SIMS, D., BRETTIN, T., DETTER, J. C., ROHDE, M., GOKER, M., BRISTOW,
1125 J., EISEN, J. A., MARKOWITZ, V., KYRPIDES, N. C., KLENK, H. P. & HUGENHOLTZ, P.
1126 2010. Complete genome sequence of *Conexibacter woesei* type strain (ID131577). *Stand*
1127 *Genomic Sci*, 2, 212-9.

1128 QUAN, X. T., SIDDIQI, M. Z., LIU, Q. Z., LEE, S. M. & IM, W. T. 2020. *Devosia ginsengisoli* sp. nov.,
1129 isolated from ginseng cultivation soil. *Int J Syst Evol Microbiol*, 70, 1489-1495.

1130 RAVIN, N. V., RAKITIN, A. L., IVANOVA, A. A., BELETSKY, A. V., KULICHEVSKAYA, I. S.,
1131 MARDANOV, A. V. & DEDYSH, S. N. 2018. Genome Analysis of *Fimbrioglobus ruber* SP5(T),
1132 a Planctomycete with Confirmed Chitinolytic Capability. *Appl Environ Microbiol*, 84.

1133 READ, T. D., PETERSON, S. N., TOURASSE, N., BAILLIE, L. W., PAULSEN, I. T., NELSON, K. E.,
1134 TETTELIN, H., FOUTS, D. E., EISEN, J. A., GILL, S. R., HOLTZAPPLE, E. K., OKSTAD, O.
1135 A., HELGASON, E., RILSTONE, J., WU, M., KOLONAY, J. F., BEANAN, M. J., DODSON, R.
1136 J., BRINKAC, L. M., GWINN, M., DEBOY, R. T., MADPU, R., DAUGHERTY, S. C.,
1137 DURKIN, A. S., HAFT, D. H., NELSON, W. C., PETERSON, J. D., POP, M., KHOURI, H. M.,
1138 RADUNE, D., BENTON, J. L., MAHAMOUD, Y., JIANG, L., HANCE, I. R., WEIDMAN, J. F.,
1139 BERRY, K. J., PLAUT, R. D., WOLF, A. M., WATKINS, K. L., NIERMAN, W. C., HAZEN, A.,
1140 CLINE, R., REDMOND, C., THWAITE, J. E., WHITE, O., SALZBERG, S. L., THOMASON,
1141 B., FRIEDLANDER, A. M., KOEHLER, T. M., HANNA, P. C., KOLSTO, A. B. & FRASER, C.
1142 M. 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related
1143 bacteria. *Nature*, 423, 81-6.

1144 REYSENBACH, A. L., ST JOHN, E., MENEGHIN, J., FLORES, G. E., PODAR, M., DOMBROWSKI,
1145 N., SPANG, A., L'HARIDON, S., HUMPHRIS, S. E., DE RONDE, C. E. J., CARATORI
1146 TONTINI, F., TIVEY, M., STUCKER, V. K., STEWART, L. C., DIEHL, A. & BACH, W. 2020.

1147 Complex subsurface hydrothermal fluid mixing at a submarine arc volcano supports distinct and
1148 highly diverse microbial communities. *Proc Natl Acad Sci U S A*, 117, 32627-32638.

1149 RINKE, C., SCHWIENTEK, P., SCZYRBA, A., IVANOVA, N. N., ANDERSON, I. J., CHENG, J. F.,
1150 DARLING, A., Malfatti, S., SWAN, B. K., GIES, E. A., DODSWORTH, J. A., HEDLUND,
1151 B. P., TSIAMIS, G., SIEVERT, S. M., LIU, W. T., EISEN, J. A., HALLAM, S. J., KYRPIDES, N.
1152 C., STEPANAUSKAS, R., RUBIN, E. M., HUGENHOLTZ, P. & WOYKE, T. 2013. Insights into
1153 the phylogeny and coding potential of microbial dark matter. *Nature*, 499, 431-7.

1154 SAFRONOVA, V. I., SAZANOVA, A. L., KUZNETSOVA, I. G., BELIMOV, A. A., ANDRONOV, E. E.,
1155 CHIRAK, E. R., POPOVA, J. P., VERKHOZINA, A. V., WILLEMS, A. & TIKHONOVICH, I.
1156 A. 2018. *Phyllobacterium zundukense* sp. nov., a novel species of rhizobia isolated from root
1157 nodules of the legume species *Oxytropis triphylla* (Pall.) Pers. *Int J Syst Evol Microbiol*, 68,
1158 1644-1651.

1159 SAKAI, H. D. & KUROSAWA, N. 2019. Complete genome sequence of the *Sulfodiicoccus acidiphilus*
1160 strain HS-1(T), the first crenarchaeon that lacks polB3, isolated from an acidic hot spring in
1161 Ohwaku-dani, Hakone, Japan. *BMC Res Notes*, 12, 444.

1162 SANGAL, V., GOODFELLOW, M., JONES, A. L., SCHWALBE, E. C., BLOM, J., HOSKISSON, P. A.
1163 & SUTCLIFFE, I. C. 2016. Next-generation systematics: An innovative approach to resolve the
1164 structure of complex prokaryotic taxa. *Sci Rep*, 6, 38392.

1165 SCHLESNER, H., RENSMANN, C., TINDALL, B. J., GADE, D., RABUS, R., PFEIFFER, S. &
1166 HIRSCH, P. 2004. Taxonomic heterogeneity within the Planctomycetales as derived by DNA–
1167 DNA hybridization, description of *Rhodopirellula baltica* gen. nov., sp. nov., transfer of *Pirellula*
1168 marina to the genus *Blastopirellula* gen. nov. as *Blastopirellula marina* comb. nov. and emended
1169 description of the genus *Pirellula*. *International Journal of Systematic and Evolutionary*
1170 *Microbiology*, 54, 1567-1580.

1171 SCHNEIKER, S., PERLOVA, O., KAISER, O., GERTH, K., ALICI, A., ALTMAYER, M. O., BARTELS,
1172 D., BEKEL, T., BEYER, S., BODE, E., BODE, H. B., BOLTEN, C. J., CHOUDHURI, J. V.,

1173 DOSS, S., ELNAKADY, Y. A., FRANK, B., GAIGALAT, L., GOESMANN, A., GROEGER, C.,
1174 GROSS, F., JELSBK, L., JELSBK, L., KALINOWSKI, J., KEGLER, C., KNAUBER, T.,
1175 KONIETZNY, S., KOPP, M., KRAUSE, L., KRUG, D., LINKE, B., MAHMUD, T.,
1176 MARTINEZ-ARIAS, R., MCHARDY, A. C., MERAI, M., MEYER, F., MORMANN, S.,
1177 MUNOZ-DORADO, J., PEREZ, J., PRADELLA, S., RACHID, S., RADDATZ, G., ROSENAU,
1178 F., RUCKERT, C., SASSE, F., SCHARFE, M., SCHUSTER, S. C., SUEN, G., TREUNER-
1179 LANGE, A., VELICER, G. J., VORHOLTER, F. J., WEISSMAN, K. J., WELCH, R. D.,
1180 WENZEL, S. C., WHITWORTH, D. E., WILHELM, S., WITTMANN, C., BLOCKER, H.,
1181 PUHLER, A. & MULLER, R. 2007. Complete genome sequence of the myxobacterium
1182 *Sorangium cellulosum*. *Nat Biotechnol*, 25, 1281-9.

1183 SEO, J. S., CHONG, H., PARK, H. S., YOON, K. O., JUNG, C., KIM, J. J., HONG, J. H., KIM, H.,
1184 KIM, J. H., KIL, J. I., PARK, C. J., OH, H. M., LEE, J. S., JIN, S. J., UM, H. W., LEE, H. J., OH,
1185 S. J., KIM, J. Y., KANG, H. L., LEE, S. Y., LEE, K. J. & KANG, H. S. 2005. The genome
1186 sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat Biotechnol*, 23, 63-8.

1187 SEVERINO, R., FROUFE, H. J. C., BARROSO, C., ALBUQUERQUE, L., LOBO-DA-CUNHA, A., DA
1188 COSTA, M. S. & EGAS, C. 2019. High-quality draft genome sequence of *Gaiella occulta* isolated
1189 from a 150 meter deep mineral water borehole and comparison with the genome sequences of
1190 other deep-branching lineages of the phylum Actinobacteria. *Microbiologyopen*, 8, e00840.

1191 SIDDIQI, M. Z., LIU, Q., KANG, M. S., KIM, M. S. & IM, W. T. 2016. *Anseongella ginsenosidimutans*
1192 gen. nov., sp. nov., isolated from soil cultivating ginseng. *Int J Syst Evol Microbiol*, 66, 1125x-
1193 1130.

1194 SINGLETON, C. M., PETRIGLIERI, F., KRISTENSEN, J. M., KIRKEGAARD, R. H., MICHAELSEN,
1195 T. Y., ANDERSEN, M. H., KONDROTAITE, Z., KARST, S. M., DUEHOLM, M. S., NIELSEN,
1196 P. H. & ALBERTSEN, M. 2021. Connecting structure to function with the recovery of over 1000
1197 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing.
1198 *Nat Commun*, 12, 2009.

1199 TAMAKI, H., TANAKA, Y., MATSUZAWA, H., MURAMATSU, M., MENG, X. Y., HANADA, S.,
1200 MORI, K. & KAMAGATA, Y. 2011. *Armatimonas rosea* gen. nov., sp. nov., of a novel bacterial
1201 phylum, Armatimonadetes phyl. nov., formally called the candidate phylum OP10. *Int J Syst Evol*
1202 *Microbiol*, 61, 1442-1447.

1203 TICE, H., MAYILRAJ, S., SIMS, D., LAPIDUS, A., NOLAN, M., LUCAS, S., GLAVINA DEL RIO, T.,
1204 COPELAND, A., CHENG, J. F., MEINCKE, L., BRUCE, D., GOODWIN, L., PITLUCK, S.,
1205 IVANOVA, N., MAVROMATIS, K., OVCHINNIKOVA, G., PATI, A., CHEN, A.,
1206 PALANIAPPAN, K., LAND, M., HAUSER, L., CHANG, Y. J., JEFFRIES, C. D., DETTER, J.
1207 C., BRETTIN, T., ROHDE, M., GOKER, M., BRISTOW, J., EISEN, J. A., MARKOWITZ, V.,
1208 HUGENHOLTZ, P., KYRPIDES, N. C., KLENK, H. P. & CHEN, F. 2010. Complete genome
1209 sequence of *Nakamurella multipartita* type strain (Y-104). *Stand Genomic Sci*, 2, 168-75.

1210 TULLY, B. J., SACHDEVA, R., GRAHAM, E. D. & HEIDELBERG, J. F. 2017. 290 metagenome-
1211 assembled genomes from the Mediterranean Sea: a resource for marine microbiology. *PeerJ*, 5,
1212 e3558.

1213 VAN PASSEL, M. W., KANT, R., PALVA, A., COPELAND, A., LUCAS, S., LAPIDUS, A., GLAVINA
1214 DEL RIO, T., PITLUCK, S., GOLTSMAN, E., CLUM, A., SUN, H., SCHMUTZ, J., LARIMER,
1215 F. W., LAND, M. L., HAUSER, L., KYRPIDES, N., MIKHAILOVA, N., RICHARDSON, P. P.,
1216 JANSSEN, P. H., DE VOS, W. M. & SMIDT, H. 2011. Genome sequence of the
1217 verrucomicrobium *Opitutus terrae* PB90-1, an abundant inhabitant of rice paddy soil ecosystems.
1218 *J Bacteriol*, 193, 2367-8.

1219 VIEIRA, S., PASCUAL, J., BOEDEKER, C., GEPPERT, A., RIEDEL, T., ROHDE, M. & OVERMANN,
1220 J. 2020. *Terricaulis silvestris* gen. nov., sp. nov., a novel prosthecate, budding member of the
1221 family Caulobacteraceae isolated from forest soil. *Int J Syst Evol Microbiol*, 70, 4966-4977.

1222 WANG, J. J., CHEN, Q. & LI, Y. Z. 2018. *Chryseolinea flava* sp. nov., a new species of *Chryseolinea*
1223 isolated from soil. *Int J Syst Evol Microbiol*, 68, 3518-3522.

1224 WANG, P. H., CHEN, Y. L., WEI, S. T., WU, K., LEE, T. H., WU, T. Y. & CHIANG, Y. R. 2020.
1225 Retroconversion of estrogens into androgens by bacteria via a cobalamin-mediated methylation.
1226 *Proc Natl Acad Sci U S A*, 117, 1395-1403.

1227 WELLY, B. T., MILLER, M. R., STOTT, J. L., BLANCHARD, M. T., ISLAS-TREJO, A. D.,
1228 O'ROURKE, S. M., YOUNG, A. E., MEDRANO, J. F. & VAN EENENNAAM, A. L. 2017.
1229 Genome Report: Identification and Validation of Antigenic Proteins from Pajaroellobacter
1230 abortibovis Using De Novo Genome Sequence Assembly and Reverse Vaccinology. *G3*
1231 (*Bethesda*), 7, 321-331.

1232 WHITMAN, W. B., WOYKE, T., KLENK, H. P., ZHOU, Y., LILBURN, T. G., BECK, B. J., DE VOS, P.,
1233 VANDAMME, P., EISEN, J. A., GARRITY, G., HUGENHOLTZ, P. & KYRPIDES, N. C. 2015.
1234 Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III: the genomes of soil and
1235 plant-associated and newly described type strains. *Stand Genomic Sci*, 10, 26.

1236 WILLIAMS, L. E., BALTRUS, D. A., O'DONNELL, S. D., SKELLY, T. J. & MARTIN, M. O. 2017.
1237 Complete Genome Sequence of the Predatory Bacterium *Ensifer adhaerens* Casida A. *Genome*
1238 *Announc*, 5.

1239 WOODHOUSE, J. N., MAKOWER, A. K., GROSSART, H. P. & DITTMANN, E. 2017. Draft Genome
1240 Sequences of Two Uncultured *Armatimonadetes* Associated with a *Microcystis* sp.
1241 (*Cyanobacteria*) Isolate. *Genome Announc*, 5.

1242 WU, L. & MA, J. 2019. The Global Catalogue of Microorganisms (GCM) 10K type strain sequencing
1243 project: providing services to taxonomists for standard genome sequencing and annotation. *Int J*
1244 *Syst Evol Microbiol*, 69, 895-898.

1245 XING, J., LI, X., SUN, Y., ZHAO, J., MIAO, S., XIONG, Q., ZHANG, Y. & ZHANG, G. 2019.
1246 Comparative genomic and functional analysis of *Akkermansia muciniphila* and closely related
1247 species. *Genes Genomics*, 41, 1253-1264.

1248 YAMAZAKI, Y., AKASHI, R., BANNO, Y., ENDO, T., EZURA, H., FUKAMI-KOBAYASHI, K.,
1249 INABA, K., ISA, T., KAMEI, K., KASAI, F., KOBAYASHI, M., KURATA, N., KUSABA, M.,

1250 MATUZAWA, T., MITANI, S., NAKAMURA, T., NAKAMURA, Y., NAKATSUJI, N.,
1251 NARUSE, K., NIKI, H., NITASAKA, E., OBATA, Y., OKAMOTO, H., OKUMA, M., SATO, K.,
1252 SERIKAWA, T., SHIROISHI, T., SUGAWARA, H., URUSHIBARA, H., YAMAMOTO, M.,
1253 YAOITA, Y., YOSHIKI, A. & KOHARA, Y. 2010. NBRP databases: databases of biological
1254 resources in Japan. *Nucleic Acids Res*, 38, D26-32.

1255 YANG, F. C., CHEN, Y. L., TANG, S. L., YU, C. P., WANG, P. H., ISMAIL, W., WANG, C. H., DING, J.
1256 Y., YANG, C. Y., YANG, C. Y. & CHIANG, Y. R. 2016. Integrated multi-omics analyses reveal
1257 the biochemical mechanisms and phylogenetic relevance of anaerobic androgen biodegradation in
1258 the environment. *ISME J*, 10, 1967-83.

1259 YEE, B., OERTLI, G. E., FUERST, J. A. & STALEY, J. T. 2010. Reclassification of the polyphyletic
1260 genus Prosthecomicrobium to form two novel genera, Vasilyevaea gen. nov. and Bauldia gen.
1261 nov. with four new combinations: Vasilyevaea enhydra comb. nov., Vasilyevaea mishustinii comb.
1262 nov., Bauldia consociata comb. nov. and Bauldia litoralis comb. nov. *Int J Syst Evol Microbiol*,
1263 60, 2960-2966.

1264 ZABURANNYI, N., BUNK, B., MAIER, J., OVERMANN, J. & MULLER, R. 2016. Genome Analysis
1265 of the Fruiting Body-Forming Myxobacterium Chondromyces crocatus Reveals High Potential
1266 for Natural Product Biosynthesis. *Appl Environ Microbiol*, 82, 1945-1957.

1267 ZENG, Y., BAUMBACH, J., BARBOSA, E. G., AZEVEDO, V., ZHANG, C. & KOBLIZEK, M. 2016.
1268 Metagenomic evidence for the presence of phototrophic Gemmatimonadetes bacteria in diverse
1269 environments. *Environ Microbiol Rep*, 8, 139-49.

1270 ZHANG, H., SEKIGUCHI, Y., HANADA, S., HUGENHOLTZ, P., KIM, H., KAMAGATA, Y. &
1271 NAKAMURA, K. 2003. Gemmatimonas aurantiaca gen. nov., sp. nov., a gram-negative, aerobic,
1272 polyphosphate-accumulating micro-organism, the first cultured representative of the new
1273 bacterial phylum Gemmatimonadetes phyl. nov. *Int J Syst Evol Microbiol*, 53, 1155-1163.

1274 ZHOU, Z., LIU, Y., XU, W., PAN, J., LUO, Z. H. & LI, M. 2020. Genome- and Community-Level
1275 Interaction Insights into Carbon Utilization and Element Cycling Functions of
1276 Hydrothermarchaeota in Hydrothermal Sediment. *mSystems*, 5.

1277 ZHU, L., LONG, M., SI, M., WEI, L., LI, C., ZHAO, L., SHEN, X., WANG, Y. & ZHANG, L. 2014.
1278 *Asticcacaulis endophyticus* sp. nov., a prosthecate bacterium isolated from the root of *Geum*
1279 *aleppicum*. *Int J Syst Evol Microbiol*, 64, 3964-3969.

1280

1281