Coupling of metataxonomics and culturing improves description of bacterial communities from a multi-contaminated waste sediment and identifies a novel *Rhizorhapis* sp. with metal resistance potential

José A. Siles[a,*], Andrew J. Hendrickson[a], Norman Terry[a]

[a]Department of Plant & Microbial Biology, University of California at Berkeley, Berkeley, CA 94720, USA

*Corresponding author. *E-mail address*: josesimartos@gmail.com

ORCID ID: José A. Siles (0000-0002-5850-0281)

Abstract

Long-term contaminated environments have been recognized as potential hotspots for bacterial discovery in taxonomic and functional terms for bioremediation purposes. Here, bacterial communities from waste sediment collected from a former industrial dumpsite and contaminated with petroleum hydrocarbon and heavy metals were investigated through the parallel application of culture-independent (16S rRNA gene amplicon sequencing) and -dependent (plate culturing followed by colony picking and isolates identification by 16S Sanger sequencing) approaches. The bacterial diversities retrieved by both approaches greatly differed. *Bacteroidetes* and *Proteobacteria* were dominant in the culture-independent community, while *Firmicutes* and *Actinobacteria* were the main culturable groups. Only 2.7% of OTUs (operational taxonomic units) in the culture-independent dataset were cultured. Most of the culturable OTUs were absent or in very low abundances in the culture-independent dataset, revealing that cultivation is a useful tool to study the rare bacterial biosphere. One culturable OTUs (comprising only the isolate SPR117) was identified as a potential new species in the genus *Rhizorhapis* (class *Alphaproteobacteria*) and was selected for further characterization. Phytopathogenicity tests showed that *Rhizorhapis* sp. strain SPR117 (ATCC TSD-228) is not pathogenic for lettuce, despite the only described species within this genus, *Rhizorhapis suberifaciens*, is causal agent of the lettuce corky root disease. The genome of the strain SPR117 was sequenced, assembled in 256 contigs with a length of 4,419,522 bp and a GC content of 59.9% and its further annotation revealed the presence of genes related to the resistance to arsenic, copper and iron, among other metals. Therefore, the coupling of metataxonomic and culturing is a useful tool to obtain not only an improved description of bacterial communities in contaminated environments, but also to isolate microorganisms with bioremediation potential.

Keywords: bacterial diversity, metataxonomics, culturing, rare biosphere, *Rhizorhapis* sp., metal resistance

1. Introduction

Pollutants such as heavy metals, total petroleum hydrocarbons (TPHs) and polycyclic aromatic hydrocarbons (PAHs) are continuously released into the environment as a consequence of industrial activities, their associated disposal operations and accidental spills (Brassington et al., 2007; Chen et al., 2015; Liu et al., 2018) . Most of them are not easily decomposable or not degradable at all and accumulate in terrestrial environments, exerting a negative impact on quality services of ecosystems, animal life and human health (Okereafor et al., 2020). Continuous and long-term exposure to contaminants usually leads to perturbations in bacterial communities (Feng et al., 2018). These alterations are usually driven by selective pressure, which favors those bacterial populations with the ability of tolerating the introduced contaminants and/or using them to their advantage (Lukhele et al., 2021; Thavamani et al., 2012; Valls and De Lorenzo, 2002). Mobile genetic elements have shown to have a critical role in this process (Dziewit et al., 2015). The study of bacterial communities in ecosystems contaminated with hydrocarbons and/or metals is thus of interest since they may be a valuable reservoir for bacterial discover in functional and taxonomic terms (Bao et al., 2017). In this context, with the introduction of high-throughput sequencing techniques, microbial metataxonomic and metagenomic studies in hydrocarbon- and metal-polluted sites have proliferated in detriment of culture-dependent surveys. This has shed light on the ecology and functionality of bacterial communities in these environments. For example, the “gamma-shift” is now a well-known phenomenon consisting in the enrichment of soil microbial communities in *Gammaproteobacteria* populations followed TPH entrance into the ecosystem (Dong et al., 2015; Militon et al., 2010; Siles and Margesin, 2018). Another example is the metagenomic work of Delmont et al. (2015), which showed that metals exert a positive selective pressure on genes related to their microbial resistance and reconstructed multiple bacterial genomes directly from metal-contaminated soils.

Nonetheless, in the last few years, as genomic sequencing data accumulate, there has been a new interest in culturing to obtain as-yet-uncultured bacteria since isolates are still required to fully discover physiology and ecology of bacteria and improve reference databases for sequence annotation (Lewis et al., 2020; Thrash, 2019, 2020). In fact, most of the genes recovered from bacterial genomes or metagenomes cannot currently be annotated, in part, because our ability to link DNA sequences to gene products is mainly based on the study of only a few well-characterized bacterial isolates (Carini, 2019). This

renewed motivation to culture bacteria from natural and polluted terrestrial ecosystems is also based on recent advances in culturing techniques and the enormous amount of molecular data probing the limitless microbial diversity in these environments, which we know very little about (Carini, 2019; Delgado-Baquerizo, 2019; Delgado-Baquerizo et al., 2018; Pham and Kim, 2012). Moreover, although the classical "1% culturability paradigm" (which means that around 99% of environmental microorganisms are unculturable) is no longer accurate, most bacterial taxa remain uncultured (Martiny, 2019; Steen et al., 2019). Innovations to cultivate new bacterial taxa include (i) membrane diffusion-based cultivation methods (e.g., i(isolation)Chip or soil substrate membrane system), (ii) microfluidics-based cultivation approaches (e.g., SlipChip), (iii) cell sorting-based techniques (e.g., Raman- activated cell sorting), (iv) fluorescence in situ hybridization of live cells and (v) reverse genomics (Lewis et al., 2020). These innovations have resulted in the isolation of new bacterial taxa that have aroused considerable interest because of the novelty of the isolated microbes or because the cultured microorganisms provided an improved understanding of certain natural processes (Lewis et al., 2020). Alternatively, traditional plate culturing with subsequent colony picking has still been shown to be a valid and successful method for the culture-dependent characterization of terrestrial bacterial communities and the isolation of "uncultured" bacteria (Nguyen et al., 2018; Thrash, 2020).

A new approach in the study of bacterial communities in polluted environments comprises the coupling of 16S rDNA amplicon sequencing and cultivation (Hamood Altowayti et al., 2020; Spini et al., 2018; Stefani et al., 2015). This provides insights into the taxa that culturing is biased towards and reveals whether the isolated strains belong to the abundant or rare bacterial biosphere (Pascual et al., 2016; Shade et al., 2012). Long-term industrial waste dumpsites have been shown to be hotspots for discovery of novel bacteria with bioremediation abilities (Muhonja et al., 2018; Singh et al., 2015). Therefore, the objective of the present work was to characterize the bacterial communities from waste sediment collected from a former industrial dumpsite (located in Bay Point, CA, USA) and contaminated with TPHs, PAHs and metals by applying culture-independent (16S rDNA amplicon sequencing) and -dependent (plate culturing with further colony picking and isolates identification by 16S rRNA gene amplification and Sanger sequencing) approaches. We hypothesized that bacterial culturable diversity would be greatly different and would represent only a small fraction of that retrieved through amplicon sequencing. We also hypothesized that

the studied sediment would be rich in potential novel species of bacteria due to its peculiar characteristics, which would be revealed by metataxonomics by yielding a high proportion of unclassified sequences at the different taxonomic levels. Since one of the isolates was a potential new species within the genus *Rhizorhapis* -with important agronomic and environmental implications-, we decided to gain more insight into the features of this strain through genome sequencing.

2. Materials and methods

2.1. Site and sampling description

The studied site is located in Bay Point, California, USA and is a former solid waste management unit (SWMU). Its dimensions are roughly 800 m (north to south) by 370 m (east to west), with a slightly trapezoidal shape. The SWMU (which is currently owned by The Pacific Gas and Electric (PG&E) Company) was constructed in the late 1940s to receive stormwater and wastewater discharges from commercial production plants and was active until 1980. Currently, the SWMU is composed of a layer of waste sediment (whose thickness ranges from 7.5 cm to a maximum of 150 cm) above natural peat and younger bay mud. The waste sediment is comprised of a varying mixture of residues from the former manufacturing plants, carbon black (fly ash residue from coal burning industry, particle size <10 µm) and organic matter (CH2M-Hill, 2010). The sediment is black, highly saline (due to the influence of marine waters) and presents contamination by TPHs, PAHs and metals (Khan et al., 2019; Xia et al., 2017). This site represents a serious environmental concern and will undergo a process of bioremediation in the near future. At the time of this study, the SWMU was partially covered by water to mitigate the odors that emanate from the sediment and to prevent the release of sediment dust. In the areas without water cap, SWMU was irregularly colonized by salgrass (*Distichlis spicata*), pickleweed (*Salicornia virginica*) and sea-purslane (*Sesuvium verrucosum*) (Xia et al., 2017). The climate in the area is typically Mediterranean, with a mean annual temperature of 15.4 ºC and annual precipitation of approximately 339 mm.

For sampling, an area of approximately 200 m^2, free of water cap, was delimited and three equidistant sampling spots were identified and selected to collect three replicate sediment samples (one from each spot). Each sample was a composite of five subsamples from the top 20 cm of the layer sediment: four subsamples orthogonally collected in a 2-m radius from a central subsample. In total, around 500 g of

sediment were collected per sample. Then, samples were transported to the lab in cooled boxes and subsequently homogenized, sieved (2 mm mesh) and stored at 4 °C for physicochemical characterization and bacterial plate culturing or at –80 °C for DNA extraction.

2.2. Physicochemical characterization of the waste sediment

Sediment samples were physicochemically characterized following standard methods at the UC Davis Analytical Lab (Davis, CA, USA). Levels of TPHs, PAHs and metals in sediment were measured according to the EPA (Environmental Protection Agency) methods 8015, 8270 and 6020, respectively, at Eurofins Calscience (Garden Grove, CA, USA).

2.3. Culture-independent bacterial diversity

Culture-independent bacterial diversity in the waste sediment was assessed through Illumina 16S rRNA gene amplicon sequencing (also known as 16S rDNA metabarcoding or metataxonomics). To do that, firstly, total DNA from each of the three sediment samples was extracted in triplicate from 250 mg of sediment fresh mass using DNeasy PowerSoil Kit (Qiagen) following the manufacturer's instructions. Triplicate DNA extracts from each sample were then pooled and DNAs were purified using DNA Clean & Concentrator Kit (Zymo Research, Irvine, USA) according to manufacturer's instructions. The quality of DNAs was spectrophotometrically checked by NanoDrop (Thermo Fisher Scientific Inc., Bremen, Germany) based on the absorbance ratios A260/A280 and A260/A230. Subsequently, DNAs were quantified using QuantiFluor™ dsDNA System (Promega, Madison, USA) and DNA concentration was standardized.

A fragment of the 16S rRNA gene, capturing the V1-V3 regions, was amplified using the primers 27F (5′- AGRGTTTGATCMTGGCTCA-3′) and 519R (5′- CCCCGYCAATTCMTTTRAGT -3′). PCR reactions were conducted using the HotStarTaq *Plus* Master Mix Kit (Qiagen, USA) and barcoded forward primers, under the following thermal conditions: (i) 94°C for 3 min, (ii) 30 cycles of 94°C for 30 s, 53°C for 40 s and 72°C for 60s; with (iii) a final elongation step at 72°C for 5 min. The success of the amplifications was checked in a 2% agarose gel and PCR products were subsequently purified using Agencourt AMPure XP magnetic beads kit (Beckman Coulter, Inc.), quantified with the QuantiFluor™ dsDNA System and pooled in equal proportions. The pooled product was then used to prepare the Illumina DNA library. Paired-end

(PE) sequencing (2 × 300) was performed on an Illumina MiSeq sequencing platform (Illumina, San Diego, USA) at MR DNA (www.mrdnalab.com, Shallowater, TX, USA).

Bioinformatic processing of the sequences was performed using the USEARCH pipeline and UPARSE-OTU algorithm (Edgar 2013). Briefly, PE sequences were firstly merged with the command -*fastq_mergepairs.* Then, reads were quality-filtered allowing a maximum e-value of 1.0, trimmed to 450-bp (base pair), dereplicated and sorted by abundance (removing singleton sequences), prior chimera detection and OTU (operational taxonomic unit) determination at 97% sequence identity. Finally, original trimmed and high-quality sequences were mapped to OTUs at the 97% identity threshold to obtain one OTU table. The taxonomic affiliation of each OTU was obtained using RDP (ribosomal database project) taxonomic classifier (Wang et al., 2007) against 16S rRNA training set 18 with a 80% confidence threshold. OTUs with average (*n*=3) relative abundances above 1% were considered as abundant "taxa", those with relative abundances below 0.1% were defined as rare and those with relative abundances between 0.1% and 1% were identified as "intermediate" OTUs (Fuhrman, 2009).

2.4. Culture-dependent bacterial diversity

Culturable heterotrophic bacteria harbored by the studied waste sediment were isolated by plate culturing (using two culture media differing in their nutrient contents) and subsequent colony picking and identification of isolates. To do this, one suspension per sample (*n*=3) was prepared by shaking 10 g of sediment with 90 ml of sterile saline solution (0.85% NaCl, w/v) for 30 min. Appropriate dilutions of soil suspensions were surface spread onto LB (Lysogeny Broth, Merck Millipore) and R2A (Reasoner's 2A, Research Products International) agar plates, with a pH of 8.2, and containing cycloheximide (200 µg mL^{-1}), to exclude fungal growth. While LB is a rich medium, R2A is a relatively nutrient-poor medium that favors the isolation of oligotrophic bacteria (Molina-Menor et al., 2021). The plates were wrapped in plastic bags to prevent evaporation and incubated during 50 days at 20 ºC. Afterwards, using the plates with the least number of colonies, 80 colonies growing on LB plates and 70 colonies growing on R2A plates (for a total of 150 colonies) were transferred to new Petri dishes and incubated for 2 weeks in order to identify unpurified and non-growing colonies. Finally, a total of 140 isolates -75 of them isolated from LB medium and 65 from R2A- were selected for taxonomic identification by 16S rRNA gene sequence analysis. To do

this, genomic DNA from each of the isolates was extracted using the DNeasy UltraClean Microbial Kit (Qiagen) following the manufacturer's instructions. The primers fD1 (5'- AGAGTTTGATCCTGGCTCAG - 3') and rP2 (5'-ACGGCTACCTTGTTACGACTT-3') were used for almost complete 16S rRNA gene amplification (Weisburg et al., 1991). PCR reactions were carried out in a final volume of 50 µL with 25 µL Taq 2X Master Mix (New England BioLabs), 1 µL each forward and reverse primers (10 µM) (Genewiz), 1 µL DNA solution and 23 µL H_2O. The thermal cycling program was as follows: (i) 95°C for 5 min; (ii) 35 cycles of 95 ºC for 30 s, 53°C for 30 s and 72°C for 60s; and (iii) a final elongation step of 72°C for 5 min. 16S rRNA gene PCR products were visualized on an agarose gel stained with SYBR™ safe (Thermo Fisher Scientific Inc.), purified using GENEJET PCR purification kit (Thermo Fisher Scientific Inc.) and Sanger sequenced (UC Berkeley DNA Sequencing Facility).

The sequences were manually edited using the software MEGA ver. X (Kumar et al., 2018) and the nearest phylogenetic neighbors were determined for each strain using the EzTazon-e Database (Kim et al., 2012). The information obtained from the taxonomic affiliation of each isolate was used to describe the culturable bacterial diversity in the waste sediment. Additionally, the 16S rRNA sequences of the isolates were clustered into OTUs and phylogenetically related to the OTUs obtained through 16S rRNA gene amplicon sequencing at a 97% identity threshold using the program VSEARCH (Rognes et al., 2016) and distance-based greedy clustering.

Almost complete 16S rRNA gene sequence of *Rhizorhapis* sp. strain SPR117 was deposited at DDBJ/EMBL/GenBank under the accession number OM362825. Phylogenetic analysis of SPR117 16S rRNA gene sequence and those of other related strains was performed using the software MEGA ver. X. Additional 16S rRNA gene sequences of related strains were retrieved from public databases. Sequences were firstly aligned with the Clustal W software (Thompson et al., 1994) and the phylogenetic tree was then reconstructed by the neighbour-joining algorithm (Saitou and Nei, 1987). The genetic distances were calculated by using Kimura's two-parameter model (Nishimaki and Sato, 2019) and the complete deletion option. The resultant tree topologies were evaluated by bootstrap analysis based on 1000 replicates.

2.5. Phytopathogenicity assay

The strain SPR117 was tested for pathogenicity to two different lettuce varieties: Waldmann's Dark Green (*Lactuca sativa* L. cv. Waldmann's dark green) and Romaine (*Lactuca sativa* L. var. *longifolia*). To do this, seeds of the two lettuce varieties were surface disinfected by immersion in 2% (v/v) hydrogen peroxide for 5 min. Seed germination was carried out at 25 ºC in trays containing sterilized vermiculite as substrate for one week. Afterwards, a 7-day old lettuce plant was planted in an 8-cm diameter pot containing sterilized potting mix substrate. In total, 24 pots were prepared for each lettuce variety. Three days after planting, and once a week afterwards, 12 of the 24 pots of each lettuce variety were inoculated with 5 mL of a cell suspension of SPR117 with an optical density of 0.380. The SPR117 inoculum was prepared by growing the strain during 7 days in R2A broth at 20 ºC under shaking conditions, centrifuging to remove culture medium and resuspending the pellet in sterile saline solution (0.85% NaCl, w/v). The 12 control pots received 5 mL of sterile saline solution. Lettuce plants were watered as needed with sterilized deionized water. The experiment was conducted in a greenhouse with controlled light and humidity conditions. Four weeks after the first bacterial inoculation, the plants were evaluated for lettuce corky root severity based on a 0–9 scale according to Brown and Michelmore (1988) and dry weight of shoot and root.

2.6. Genome sequence of Rhizorhapis *sp. strain SPR117 (ATCC TSD-228)*

For whole-genome sequencing, *Rhizorhapis sp.* strain SPR117 was grown during 7 days in R2A broth at 20 °C and its genomic DNA (gDNA) was extracted using the DNeasy UltraClean Microbial Kit (Qiagen, Germany) according to the manufacturer's instructions. GDNA was then quality-checked, quantified and used for the preparation of a NEBNext® Ultra™ DNA library (New England Biolabs) following the manufacturer's recommendations. The DNA library was loaded on an Illumina HiSeq instrument and sequenced using a 2x150 PE configuration in High Output Run mode at Genewiz (South Plainfield, NJ, USA).

Raw reads were quality-checked using FastQC v0.11.5 and then filtered using Trimmomatic v.0.38 (ILLUMINACLIP = TRUE; LEADING = 3; TRAILING = 3; SLIDINGWINDOW = 4:30; MINLEN = 90). Filtered PE sequences were then de novo assembled with SPAdes v. 3.13.0 (Bankevich et al. (2012); -no trim, 2 racon iterations, 2 pilon iterations, 5 minimum contig coverage and 300 bp minimum contig length-) integrated in PATRIC Bioinformatics Resource Center (Davis et al., 2020). Subsequently, contigs were

annotated using the RAST (Rapid Annotations using Subsystems Technology, Brettin et al. (2015)), Prokka v.1.14.6 (Seemann, 2014) and the NCBI Prokaryotic Genome Annotation Pipeline (Tatusova et al., 2016). CheckM algorithm was used for assessing the genome quality (Bankevich et al., 2012).

2.7. Statistical analyses

Non-parametric univariate PERMANOVA (Permutational Analysis of Variance) was used to check whether root and shoot dry weight of lettuce plants was significantly affected by the inoculation of *Rhizorhapis* sp. strain SPR117. For this analysis, the *adonis* function in the R package *vegan* was used (Oksanen et al., 2013). Data visualizations were performed using the R package *ggplot2* (Wickham, 2016) and CorelDRAW ver. 2020.

2.8. Strain and data accessibility

Rhizorhapis sp. strain SPR117 was deposited at the American Type Culture Collection (ATCC) under the strain designation ATCC TSD-228.

The draft-genome sequence of *Rhizorhapis* sp. strain SPR117 was deposited at DDBJ/EMBL/GenBank under the accession number JAKKIE000000000 (BioProject no. PRJNA800004 and BioSample no. SAMN25221871). The version described in this paper is the first version, JAKKIE000000000.

The raw bacterial sequences of the 16S rDNA amplicon sequencing analysis were deposited in the GenBank SRA database under BioProject accession number PRJNA800331.

3. Results and discussion

3.1. Physicochemical properties of the waste sediment.

The sediment was characterized by presenting an extremely high C/N ratio (199), alkaline pH (8.2), high salinity (electrical conductivity, 1.2 mS cm^{-1}) and high levels of contamination by TPHs (20.1 g kg^{-1} sediment), PAHs (370 mg kg^{-1} sediment; predominating: benzo(*ghi*)perylene, pyrene, indeno(1,2,3-*cd*)pyrene, benzo(*a*)pyrene, fluoranthene) and heavy metals (chromium, barium, zinc, copper, lead and

nickel) (Table S1). These data confirmed the peculiar physicochemical characteristics of the studied waste sediment.

3.2. Culture-independent and -dependent bacterial diversity in the waste sediment.

The metabarcoding assay yielded 98,948 PE high-quality 16S rDNA sequences, which were distributed among 625 OTUs. Of these OTUs, 621 were classified at kingdom level as *Bacteria* (using an 80% confidence threshold) and the remaining 4 as *Archaea*. Only bacterial OTUs were retained for further taxonomic analyses. Culture-independent bacterial diversity was distributed among 20 different phyla, being the predominant ones *Bacteroidetes* (the average percentage of classified sequences in this phylum was 52.4), *Proteobacteria* (14.2%), *Actinobacteria* (7.6%), *Planctomycetes* (4.5%), *Saccharibacteria* (3.3%), *Verrucomicrobia* (3.0%) and *Balneolaeota* (2.7%); these six phyla accounted for more than 87% of total sequences (Fig. 1a). *Flavobacteriia* class predominated among *Bacteroidetes*, *Gammaproteobacteria* among *Proteobacteria*, *Planctomycetacia* among *Plantomycetes* and *Acidimicrobiia* among *Actinobacteria* (Fig. 1b). These findings concur with those of previous works showing that *Bacteroidetes*, *Proteobacteria* and *Actinobacteria* usually predominate in long-term metal- and hydrocarbon soils (Khan et al., 2019; Pan et al., 2020; Siles and Margesin, 2018). At OTU level, 11 OTUs were regarded as abundant (i.e., with relative abundances >1%) and belonged to the phyla *Actinobacteria, Bacteroidetes*, *Chloroflexi*, *Proteobacteria*, *Saccharibacteria* and *Verrucomicrobia* (Table S2). Among these OTUs, the Otu1, with a relative abundance of 47%, clearly predominated and was taxonomically affiliated to the genus *Hoppeia* (although only with a confidence threshold of 56%, Table 2). Finding this microbial group as predominant makes sense since the only species described within this genus (namely, *Hoppeia youngheungensis*) was isolated from tidal flat sediment (Kwon et al., 2014).

76.1% of sequences could not be classified at genus level. This result can have two possible explanations: (i) the sequence length was too short for accurate classification or (ii) the reference database was not complete enough and some comparative elements were missing for the classification of all the bacterial sequences (Rachid et al., 2013). In the present study, since (i) a much lower proportion of unclassified sequences at different taxonomic levels was obtained in previous works using the same approaches as those utilized here for taxonomic affiliation of reads with a similar length (Margesin et al.,

2017) and (ii) most of the sequences were successfully classified at kingdom (*Bacteria*) and phylum levels, we explain the high proportion of sequences without an effective taxonomic classification as a consequence of the presence of a high proportion of undescribed bacterial diversity in the studied sediment (Delgado-Baquerizo, 2019). In light of these results, we presumed that bacterial culturing would result in the isolation of a high proportion of as-yet-uncultured taxa.

The culture-dependent characterization of the bacterial community in the waste sediment resulted in the isolation and successful identification of 140 strains. The culturable bacterial community was dominated by *Firmicutes* (56.4%), although isolates belonging to the phyla *Proteobacteria* (27.1%) and *Actinobacteria* (16.4%) were also detected (Fig. 1a). Despite we (i) used both rich and oligotrophic culture media for bacterial isolation, (ii) adjusted the pH of media to that of the sediment (i.e., 8.2) and (iii) used a rather long incubation time, the culturable diversity differed from that recovered through the 16S metabarcoding assay, as we initially hypothesized. In this way, only 2.3% of the OTUs in the metabarcoding library were cultured (Fig. 2 and Table 1). These dramatic differences between culture-dependent and -independent bacterial diversities have previously been reported (Lee et al., 2016; Siles et al., 2014) and are a consequence of our incapacity to reproduce in the laboratory the particular abiotic conditions (nutrient concentrations and forms, temperature, humidity, oxygen -or other gases- levels, light, etc.) that most bacteria need to growth (Hugenholtz, 2002; Molina-Menor et al., 2021). Furthermore, the growth of specific taxa depends on interactions between species and metabolic cooperation, which are largely unknown and really hard to mimic under experimental conditions (Molina-Menor et al., 2021). In this way, until recently, it has been believed that culturing only offers a biased and useless description of the bacterial communities of a given environment (Ritz, 2007). However, an increasing number of works are demonstrating that culturing and metataxonomics are complementary to fully describe the microbial diversity in a given environment (Shade et al., 2012; Stefani et al., 2015; Youseif et al., 2021). In our study, phylogenetic comparison of the culture-dependent and -independent communities at OTU level showed that a total of 21 OTUs were detected only by culturing (Fig. 2). Previous works dealing with this topic have found similar patterns. For example, Lee et al. (2016) found that 73 OTUs were detected only by cultivation after studying the bacterial diversity of tomato rhizosphere by coupling culturing and 16S metabarcoding. Shade et al. (2012) reported that up to 601 OTUs were uniquely retrieved by culture-dependent approaches during the

investigation of the bacterial communities in an agricultural soil. It has been argued that bacterial taxa recovered only by culturing are rare members of the microbial community, which are detected only by this approach because sequencing depth was insufficient (Hiergeist et al., 2015; Lee et al., 2016). We observed that most of the cultured OTUs in the metabarcoding dataset represented rare members of the culture-independent community (Table 1). Likewise, none of the abundant OTUs in the metabarcoding dataset was cultured (Table S2 and Table 1). Therefore, culturing seems to be especially useful to dig deeper into the rare biosphere of the waste sediment studied. Members of the rare biosphere are essential for the maintenance of microbial alpha and beta diversity and represent a seed bank of genetic and metabolic diversity that responds to different biotic or abiotic changes, contributing to the resilience of the community (Lasa et al., 2019; Lynch and Neufeld, 2015; Saw, 2021).

Approximately 79% of the strains shared ≥98.7% sequence identity with the closest known species in the EzTaxon Database, while ca. 21% of the isolates (sharing identities <98.7%) could constitute potential novel species (Table 1). These results suggest that plate culturing with subsequent random colony picking is still a suitable method for the isolation of as-yet-uncultured bacteria in terrestrial ecosystems, especially when long incubation periods are considered to facilitate cells awaking from dormant states (Thrash, 2020). However, it is well known that this approach requires substantial amounts of time and resources to be successful. New approaches in culturing tend thus to be more targeted (Carini, 2019). In this way, lists of most-wanted microbes are being created (Lewis et al., 2020). Reasons to include a specific microorganism in these lists are that it (i) belongs to a poorly characterized group, (ii) is key for a specific process, biogeochemical cycle or bioremediation and/or (iii) can potentially produce natural products (Carini, 2019; Lewis et al., 2020). Once one specific microorganism has been targeted for its isolation, particular cultivation strategies can be designed according to its lifestyle. For example, 79 different isolates of *Planctomycetes* were recently cultured for the first time using selective enrichment, antibiotic treatment and solid media streaking combined with colony picking (Wiegand et al., 2019). Complementary, it has been suggested that genomic data from uncultured microbes can be used to identify potential cultivation strategies through metabolic reconstruction (Carini, 2019).

Among all the potential novel strains discovered in the present work, the isolate SPR117 (which was isolated from the medium R2A and comprised the singleton Otu31, Table 1) captured our attention

since its closest known strain was, with a pairwise sequence similarity of only 96.88%, *Rhizorhapis suberifaciens* CA1^{T}. This finding suggests that the isolate SPR117 represents a new species within the genus *Rhizorhapis*. This assumption was supported by a neighbour-joining phylogenetic tree analysis, which showed that the strain SPR117 -forming a single tree branch- was closely related, but clearly distinct, from a cluster grouping different strains of *R. suberifaciens*. Unfortunately, these branches were not supported by high bootstrap values (Fig 3). *Rhizorhapis* cluster was closer related to *Sphingobium* cluster than the *Sphingomonas* one (Fig. 3). On other hand, the OTU containing the isolate SPR117 was not phylogenetically related to any of the OTUs in the amplicon sequencing dataset (Table 1). We thus assume that the isolate SPR117 is a rare member of the bacterial community in the waste sediment, which was not captured by the 16S amplicon sequencing assay. In light of the agronomic (Francis et al., 2014) and environmental importance (Brereton et al., 2020) of the genus *Rhizorhapis*, we decided to gain more insights into the characteristics of the strain SPR177.

3.3. Characterization of Rhizorhapis *sp. strain SPR117 (ATCC TSD-228)*

The genus *Rhizorhapis*, formerly known as *Rhizomonas*, belongs to the family *Sphingomonadaceae* (*Alphaproteobacteria* class) and currently comprises only the species type *Rhizorhapis suberifaciens,* which is causal agent of the lettuce (*Lactuca sativa* L.) corky root disease (Francis et al., 2014). Corky root has been one of the most important diseases of lettuce in production areas of USA, Europe and Australia (van Bruggen et al., 2015). In severe cases of the disease, the entire taproot becomes brown, severely cracked, non-functional and easily breakable (van Bruggen et al., 2014b). These symptoms at root level result in reduced head sizes of lettuce plants and decreasing crop yields. Corky root disease can be very severe —especially in intensive production systems and in warm climates—, with yield losses ranging from 30 to 70% (van Bruggen et al., 2014a). In this context, the evaluation of the phytopathogenic potential of the strain SPR117 was of interest.

We conducted a greenhouse experiment where the pathogenicity of SPR117 was tested in the lettuce varieties Romaine and Waldmann's Dark Green. After 30 days, the root and shoot dry weight of lettuce plants was not affected by the inoculation of SPR117 (Fig. 4). The potential corky root severity in the inoculated plants was also evaluated according to the guidelines provided by Brown and Michelmore

(1988), finding no symptoms related to this disease. Therefore, the strain SPR117 is not pathogenic to lettuce, at least to the varieties tested. Interestingly, some *Rhizorhapis* sp. strains have shown to provide specific control of corky root disease (van Bruggen et al., 2014b) and further studies will be needed in this respect for the strain SPR117.

Further insights into the strain SPR117 were obtained through the sequencing of its genome. Sequencing work yielded more than 33 M PE sequences, 18.79 M of which remained after quality-filtering and were subsequently used for assembling and annotating the genome. Main features of the assembly and annotation processes are summarized in Table 2. The genome of *Rhizorhapis* sp. SPR117 was assembled into 256 contigs (>300 bp) with a length of 4,419,522 bp long and an N_{50} score of 62,062 bp. It has a GC content of 59.9%, similar to that of *R. suberifaciens* (58.9%) (Francis et al., 2014), and a completeness of 98%. The annotation process predicted a total of 4,278 coding genes and coding sequences (CDSs), 3 rRNAs and 46 tRNAs (Table 2). 1129 CDSs were successfully classified into subsystems by RAST. A subsystem is a set of proteins that together implement a specific biological process or structural complex (Overbeek et al., 2005). The four most abundant subsystem categories were (i) membrane transport, (ii) amino acids and derivatives, (iii) carbohydrates, and (iv) protein metabolism (Fig. S1).

Members of the genus *Rhizorhapis* have been hypothesized to degrade petroleum hydrocarbons (Brereton et al., 2020); however, the genome of the strain SPR117 does not present genes compatible with this capability. This is seen in concordance with the results of a laboratory experiment where we observed lack of bacterial growth when the isolate SPR117 was cultivated in culture media containing alkanes (dodecane) or PAHs (anthracene, benzene and phenanthrene) as the only C source.

Some metals are required by bacteria in minute amounts for normal cell metabolism and as components of biological macromolecules, being their intake thus subjected to complex homeostatic mechanisms that ensure sufficient (but not excessive) acquisition (Chen et al., 2019; Valls and De Lorenzo, 2002). Instead, other metals do not have a relevant function. Both metal categories (i.e., with and without a biological function) can be toxic for bacteria when accumulated to a certain concentration. Consequently, emergence of metal resistance features in bacteria was an early event in evolution (Valls and De Lorenzo, 2002). Mechanisms of metal resistance include intra- or extracellular immobilization, biotransformation of

the toxic metal ion into a less noxious form, dissimilatory reduction and modulation of metal concentration homeostasis through cation efflux pumps (Romero et al., 2021). Since metals accumulate in the environment for extremely long periods and represent a strong selective pressure for bacterial life (Li et al., 2021), long-term metal polluted ecosystems, as the waste sediment studied here, is a potential hotspot for the isolation of as-yet-uncultured bacteria with metal resistance capabilities. In line with this argumentation, we found that the strain SPR117 has potential metal tolerance capabilities since its genome harbors a variety of genes encoding proteins involved in the resistance to cobalt, mercury, cadmium, zinc, and specially, arsenic, copper and iron (Table S3). This finding is seen in connection with the results of the work by Brereton et al. (2020), which showed that the relative abundance of OTUs taxonomically affiliated to the genus *Rhizorhapis* increased in the rhizosphere of *Festuca arundinacea* in comparison with unplanted control soil during processes of phytoremediation of metal-contaminated soils.

The genome of *Rhizorhapis* sp. strain SPR117 contains genes related to arsenic methylation (gene *arsH*) and reduction (*arsC*, *arsO*, *arsR* and *acr3*) (Table S3). The gene *arsC* encodes a glutaredoxin arsenate reductase that reduces arsenate to arsenite within the cytoplasmic membrane; arsenite is then exported out of the cell by an efflux bum (Abdullahi et al., 2021; Li et al., 2021). In fact, the gene *acr3* gene encodes an arsenite efflux permease (Li et al., 2021). The presence of the gene *arsH* in SPR117 may point out its capability to oxidize methyl arsenate compounds (Abbaszade et al., 2020). The four main pathways of microbial As detoxification in the environment are arsenite oxidation, arsenate reduction, arsenate respiration and arsenate methylation (Li et al., 2021).

Copper has an important role in cellular function. Due to its ability to undergo redox changes [Cu(I) ↔ Cu(II)], copper is an ideal cofactor for enzymes catalyzing electron transfers. However, this characteristic also makes copper dangerous for a cell as it is able to be involved in Fenton-like reactions creating reactive oxygen species. Cu(I) can attack and destroy iron-sulfur clusters, which releases iron that can in turn cause oxidative stress. Therefore, copper homeostasis has to be highly balanced to ensure proper cellular function and avoid cell damage (Rensing and McDevitt, 2013). Genes for copper resistance detected in the genome of the strain SPR117 included *copB*, *copC* and *copD* (Table S3). *copB* and *copD* encode outer and inner membrane proteins, respectively, while *copC* protein is believed to be a periplasmic copper chaperone (Lawton et al., 2016). Other genes for copper resistance were *cusA*, *cusB* and *CusC*, which encode three

proteins that are believed to form a complex that spans the periplasm and has the ability to remove copper and silver from the cell (Lawton et al., 2016).

The microbial iron homeostasis is highly regulated as this metal is required for cellular processes, but it becomes toxic at high concentrations. Several genes responsible for iron metabolism and transport were found in the genome of the strain SPR117 (Table S3). Iron-sulfur enzymes (as those synthetized by the genes *sufA*, *sufB*, *sufC*, *sufD*) are involved in important biological functional and processes such as electron transfer, regulating gene expression and substrate binding/activation, where the active site is made of iron and sulfur atoms (Abdullahi et al., 2021).

Although metals are mostly not susceptible to complete microbial degradation, metal pollution bioremediation uses microorganisms to immobilize or transform metals into non-bioavailable or less toxic forms (Ye et al., 2017). These approaches do not solve the problem completely, but they do help to protect polluted sites from noxious effects and isolate the contaminants as a contained, and sometimes, recyclable unit (Valls and De Lorenzo, 2002). Newly isolated bacteria, such as *Rhizorhapis* sp. strain SPR117, with metal tolerant features, are a potential source of bioremediation tools. For example, as shown by Teixeira et al. (2014) and Emenike et al. (2015), bioaugmentation of metal-contaminated soils has shown promising results. On the other hand, metal resistance genes are valuable resources for genetic engineering. For instance, the transformation of *Populus alba* × *P. tremula* var. *glandulosa* with a metal tolerance gene (*ScYCF1*, yeast cadmium factor 1) led to improved poplar growth and cadmium accumulation in aerial tissue (Shim et al., 2013). The overexpression of a phytochelatin synthase from *Schizosaccharomyces pombe* in *Escherichia coli* resulted in a higher cellular cadmium accumulation than that in the wild type (Kang et al., 2007). And the engineering of *Pseudomonas putida* for expression of the *arsM* arsenic(III) S-adenosine methyltransferase gene showed to be an effective approach to volatilize arsenic from polluted soils (Chen et al., 2014).

4. Conclusions

In the present work, we coupled 16S rDNA metabarcoding and plate culturing with further colony picking and identification of the isolates to describe the bacterial communities inhabiting multi-contaminated waste sediment (petroleum hydrocarbons and heavy metals) collected from a former dumpsite. The

bacterial diversities recovered by both approaches greatly differed and only a small fraction (2.7%) of the culture-independent bacterial community was cultured due to the current incapability to reproduce in the laboratory the specific abiotic and biotic conditions that most bacteria need to growth. Surprisingly, most of the culturable OTUs either were absent or found in very low abundances in the culture-independent community, which demonstrates that culturing complements the diversity retrieved through metataxonomics and that it is useful to investigate the rare biosphere. In fact, one member of the rare bacterial biosphere in the studied waste sediment was identified as a potential new species in the genus *Rhizorhapis*. The genome sequence analysis of *Rhizorhapis* sp. strain SPR117 (ATCC TSD-228) demonstrated that this strain may have bioremediation capabilities since genes related to metal resistance were annotated. This shows that rare biosphere is a reservoir of potential tools for bioremediation. Further studies will include a formal description of *Rhizorhapis* sp. strain SPR117 as a new species and will assess the real metal bioremediation capabilities of this bacterium.

Funding

This work was supported by the UC Berkeley grant number 51719.

CRediT authorship contribution statement

José A. Siles: Conceptualization, Methodology, Formal analysis, Investigation, Writing - Original Draft, Visualization. **Andrew J. Hendrickson:** Investigation, Writing - Review & Editing. **Norman Terry**: Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Ben LePage and Bob Grey from PG&E Company for providing us access to Shell Pond.

References

Abbaszade, G., Szabó, A., Vajna, B., Farkas, R., Szabó, C., Tóth, E., 2020. Whole genome sequence analysis of Cupriavidus campinensis S14E4C, a heavy metal resistant bacterium. Molecular Biology Reports 47, 3973-398510.1007/s11033-020-05490-8.

Abdullahi, S., Haris, H., Zarkasi, K.Z., Amir, H.G., 2021. Complete genome sequence of plant growth-promoting and heavy metal-tolerant Enterobacter tabaci 4M9 (CCB-MBL 5004). Journal of Basic Microbiology 61, 293-30410.1002/jobm.202000695.

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. Journal of Computational Biology 19, 455-47710.1089/cmb.2012.0021.

Bao, Y.J., Xu, Z., Li, Y., Yao, Z., Sun, J., Song, H., 2017. High-throughput metagenomic analysis of petroleum-contaminated soil microbiome reveals the versatility in xenobiotic aromatics metabolism. Journal of Environmental Sciences (China) 56, 25-3510.1016/j.jes.2016.08.022.

Brassington, K.J., Hough, R.L., Paton, G.I., Semple, K.T., Risdon, G.C., Crossley, J., Hay, I., Askari, K., Pollard, S.J.T., 2007. Weathered hydrocarbon Wastes: A Risk Management Primer. Critical Reviews in Environmental Science and Technology 37, 199-23210.1080/10643380600819625.

Brereton, N.J.B., Gonzalez, E., Desjardins, D., Labrecque, M., Pitre, F.E., 2020. Co-cropping with three phytoremediation crops influences rhizosphere microbiome community in contaminated soil. Science of The Total Environment 711, 135067https://doi.org/10.1016/j.scitotenv.2019.135067.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., 3rd, Stevens, R., Vonstein, V., Wattam, A.R., Xia, F., 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Scientific Reports 5, 8365-836510.1038/srep08365.

Brown, P., Michelmore, R., 1988. The genetics of corky root resistance in lettuce. Phytopathology 78, 1145-1150.

Carini, P., 2019. A “Cultural” Renaissance: Genomics Breathes New Life into an Old Craft. mSystems 4, e00092-0001910.1128/mSystems.00092-19.

CH2M-Hill, 2010. Summary of Findings Associated with Shell Pond Field Surveys, PG&E Shell Pond, Bay Point, California.

Chen, J., Li, J., Zhang, H., Shi, W., Liu, Y., 2019. Bacterial Heavy-Metal and Antibiotic Resistance Genes in a Copper Tailing Dam Area in Northern China. Frontiers in Microbiology 1010.3389/fmicb.2019.01916.

Chen, J., Sun, G.X., Wang, X.X., Lorenzo, V., Rosen, B.P., Zhu, Y.G., 2014. Volatilization of arsenic from polluted soil by Pseudomonas putida engineered for expression of the arsM Arsenic(III) S-adenosine methyltransferase gene. Environmental Science & Technology 48, 10337-1034410.1021/es502230b.

Chen, M., Xu, P., Zeng, G., Yang, C., Huang, D., Zhang, J., 2015. Bioremediation of soils contaminated with polycyclic aromatic hydrocarbons, petroleum, pesticides, chlorophenols and heavy metals by composting: Applications, microbes and future research needs. Biotechnology Advances 33, 745-75510.1016/j.biotechadv.2015.05.003.

Davis, J.J., Wattam, A.R., Aziz, R.K., Brettin, T., Butler, R., Butler, R.M., Chlenski, P., Conrad, N., Dickerman, A., Dietrich, E.M., Gabbard, J.L., Gerdes, S., Guard, A., Kenyon, R.W., Machi, D., Mao, C., Murphy-Olson, D., Nguyen, M., Nordberg, E.K., Olsen, G.J., Olson, R.D., Overbeek, J.C., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomas, C., VanOeffelen, M., Vonstein, V., Warren, A.S., Xia, F., Xie, D., Yoo, H., Stevens, R., 2020. The PATRIC Bioinformatics Resource

Center: expanding data and analysis capabilities. Nucleic Acids Research 48, D606-d61210.1093/nar/gkz943.

Delgado-Baquerizo, M., 2019. Obscure soil microbes and where to find them. The ISME journal 13, 2120-212410.1038/s41396-019-0405-0.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., Maestre, F.T., Singh, B.K., Fierer, N., 2018. A global atlas of the dominant bacteria found in soil. Science 359, 320-32510.1126/science.aap9516.

Delmont, T.O., Eren, A.M., Maccario, L., Prestat, E., Esen Ö, C., Pelletier, E., Le Paslier, D., Simonet, P., Vogel, T.M., 2015. Reconstructing rare soil microbial genomes using in situ enrichments and metagenomics. Frontiers in Microbiology 6, 35810.3389/fmicb.2015.00358.

Dong, Y., Lang, Z., Kong, X., Lu, D., Liu, Z., 2015. Kinetic and multidimensional profiling of accelerated degradation of oil sludge by biostimulation. Environmental Science: Processes & Impacts 17, 763-774.

Dziewit, L., Pyzik, A., Szuplewska, M., Matlakowska, R., Mielnicki, S., Wibberg, D., Schlüter, A., Pühler, A., Bartosik, D., 2015. Diversity and role of plasmids in adaptation of bacteria inhabiting the Lubin copper mine in Poland, an environment rich in heavy metals. Frontiers in Microbiology 610.3389/fmicb.2015.00152.

Emenike, C.U., Agamuthu, P., Fauziah, S.H., 2015. Blending Bacillus sp., Lysinibacillus sp. and Rhodococcus sp. for optimal reduction of heavy metals in leachate contaminated soil. Environmental Earth Sciences 75, 2610.1007/s12665-015-4805-9.

Feng, G., Xie, T., Wang, X., Bai, J., Tang, L., Zhao, H., Wei, W., Wang, M., Zhao, Y., 2018. Metagenomic analysis of microbial community and function involved in cd-contaminated soil. BMC Microbiology 18, 1110.1186/s12866-018-1152-5.

Francis, I.M., Jochimsen, K.N., De Vos, P., van Bruggen, A.H., 2014. Reclassification of rhizosphere bacteria including strains causing corky root of lettuce and proposal of Rhizorhapis suberifaciens gen. nov., comb. nov., Sphingobium mellinum sp. nov., Sphingobium xanthum sp. nov. and Rhizorhabdus argentea gen. nov., sp. nov. International Journal of Systematic and Evolutionary Microbiology 64, 1340-135010.1099/ijs.0.058909-0.

Fuhrman, J.A., 2009. Microbial community structure and its functional implications. Nature 459, 193-19910.1038/nature08058.

Hamood Altowayti, W.A., Almoalemi, H., Shahir, S., Othman, N., 2020. Comparison of culture-independent and dependent approaches for identification of native arsenic-resistant bacteria and their potential use for arsenic bioremediation. Ecotoxicology and Environmental Safety 205, 111267https://doi.org/10.1016/j.ecoenv.2020.111267.

Hiergeist, A., Gläsner, J., Reischl, U., Gessner, A., 2015. Analyses of Intestinal Microbiota: Culture versus Sequencing. ILAR Journal 56, 228-24010.1093/ilar/ilv017.

Hugenholtz, P., 2002. Exploring prokaryotic diversity in the genomic era. Genome Biology 3, reviews0003.000110.1186/gb-2002-3-2-reviews0003.

Kang, S.H., Singh, S., Kim, J.-Y., Lee, W., Mulchandani, A., Chen, W., 2007. Bacteria Metabolically Engineered for Enhanced Phytochelatin Production and Cadmium Accumulation. Applied and Environmental Microbiology 73, 6317-6320doi:10.1128/AEM.01237-07.

Khan, A.R., Reichmann, L.G., Ibal, J.C., Shin, J.H., Liu, Y., Collins, H., LePage, B., Terry, N., 2019. Variation in pickleweed root-associated microbial communities at different locations of a saline solid waste management unit contaminated with petroleum hydrocarbons. PloS One 14, e022290110.1371/journal.pone.0222901.

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H., Won, S., Chun, J., 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. International Journal of Systematic and Evolutionary Microbiology 62, 716-72110.1099/ijs.0.038075-0.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution 35, 1547-154910.1093/molbev/msy096.

Kwon, Y.M., Choi, D., Yang, S.H., Kwon, K.K., Kim, S.J., 2014. Hoppeia youngheungensis gen. nov., sp. nov., a member of the Flavobacteriaceae isolated from tidal flat sediment, and emended descriptions of the genus Sediminibacter and Sediminibacter furfurosus. International Journal of Systematic and Evolutionary Microbiology 64, 2053-205910.1099/ijs.0.061721-0.

Lasa, A.V., Fernández-González, A.J., Villadas, P.J., Toro, N., Fernández-López, M., 2019. Metabarcoding reveals that rhizospheric microbiota of Quercus pyrenaica is composed by a relatively small number of bacterial taxa highly abundant. Scientific Reports 9, 169510.1038/s41598-018-38123-z.

Lawton, T.J., Kenney, G.E., Hurley, J.D., Rosenzweig, A.C., 2016. The CopC Family: Structural and Bioinformatic Insights into a Diverse Group of Periplasmic Copper Binding Proteins. Biochemistry 55, 2278-229010.1021/acs.biochem.6b00175.

Lee, S.A., Park, J., Chu, B., Kim, J.M., Joa, J.H., Sang, M.K., Song, J., Weon, H.Y., 2016. Comparative analysis of bacterial diversity in the rhizosphere of tomato by culture-dependent and -independent approaches. Journal of microbiology 54, 823-83110.1007/s12275-016-6410-3.

Lewis, W.H., Tahon, G., Geesink, P., Sousa, D.Z., Ettema, T.J.G., 2020. Innovations to culturing the uncultured microbial majority. Nature Reviews Microbiology 19, 225–240, 10.1038/s41579-020-00458-8.

Li, X., Liu, X., Cao, N., Fang, S., Yu, C., 2021. Adaptation mechanisms of arsenic metabolism genes and their host microorganisms in soils with different arsenic contamination levels around abandoned gold tailings. Environmental Pollution 291, 117994, https://doi.org/10.1016/j.envpol.2021.117994.

Liu, L., Li, W., Song, W., Guo, M., 2018. Remediation techniques for heavy metal-contaminated soils: Principles and applicability. Science of The Total Environment 633, 206-219 https://doi.org/10.1016/j.scitotenv.2018.03.161.

Lukhele, T., Ogola, H.J.O., Selvarajan, R., Oruko, R.O., Nyoni, H., Mamba, B.B., Msagati, T.A.M., 2021. Metagenomic insights into taxonomic diversity and metabolic potential of bacterial communities associated with tannery waste-contaminated soils. International Journal of Environmental Science and Technology 10.1007/s13762-021-03298-y.

Lynch, M.D.J., Neufeld, J.D., 2015. Ecology and exploration of the rare biosphere. Nature Reviews Microbiology 13, 217-22910.1038/nrmicro3400.

Margesin, R., Siles, J.A., Cajthaml, T., Öhlinger, B., Kistler, E., 2017. Microbiology Meets Archaeology: Soil Microbial Communities Reveal Different Human Activities at Archaic Monte Iato (Sixth Century BC). Microbial Ecology 73, 925-93810.1007/s00248-016-0904-8.

Martiny, A.C., 2019. The '1% culturability paradigm' needs to be carefully defined. The ISME journal 14, 10–11, 10.1038/s41396-019-0507-8.

Militon, C., Boucher, D., Vachelard, C., Perchet, G., Barra, V., Troquet, J., Peyretaillade, E., Peyret, P., 2010. Bacterial community changes during bioremediation of aliphatic hydrocarbon-contaminated soil. FEMS Microbiology Ecology 74, 669-681.

Molina-Menor, E., Gimeno-Valero, H., Pascual, J., Peretó, J., Porcar, M., 2021. High Culturable Bacterial Diversity From a European Desert: The Tabernas Desert. Frontiers in Microbiology 1110.3389/fmicb.2020.583120.

Muhonja, C.N., Makonde, H., Magoma, G., Imbuga, M., 2018. Biodegradability of polyethylene by bacteria and fungi from Dandora dumpsite Nairobi-Kenya. PloS One 13, e019844610.1371/journal.pone.0198446.

Nguyen, T.M., Seo, C., Ji, M., Paik, M.-J., Myung, S.-W., Kim, J., 2018. Effective Soil Extraction Method for Cultivating Previously Uncultured Soil Bacteria. Applied and Environmental Microbiology 84, e01145-0111810.1128/aem.01145-18.

Nishimaki, T., Sato, K., 2019. An Extension of the Kimura Two-Parameter Model to the Natural Evolutionary Process. Journal of Molecular Evolution, 87, 60-6710.1007/s00239-018-9885-1.

Okereafor, U., Makhatha, M., Mekuto, L., Uche-Okereafor, N., Sebola, T., Mavumengwana, V., 2020. Toxic Metal Implications on Agricultural Soils, Plants, Animals, Aquatic life and Human Health. International Journal of Environmental Research and Public Health 17, 220410.3390/ijerph17072204.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.-Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E.D., Gerdes, S., Glass, E.M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A.C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G.D., Rodionov, D.A., Rückert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., Vonstein, V., 2005. The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes. Nucleic Acids Research 33, 5691-570210.1093/nar/gki866.

Pan, X., Zhang, S., Zhong, Q., Gong, G., Wang, G., Guo, X., Xu, X., 2020. Effects of soil chemical properties and fractions of Pb, Cd, and Zn on bacterial and fungal communities. Science of The Total Environment 715, 136904, https://doi.org/10.1016/j.scitotenv.2020.136904.

Pascual, J., Blanco, S., García-López, M., García-Salamanca, A., Bursakov, S.A., Genilloud, O., Bills, G.F., Ramos, J.L., van Dillewijn, P., 2016. Assessing Bacterial Diversity in the Rhizosphere of *Thymus zygis* Growing in the Sierra Nevada National Park (Spain) through Culture-Dependent and Independent Approaches. PloS one 11, e0146558, 10.1371/journal.pone.0146558.

Pham, V.H., Kim, J., 2012. Cultivation of unculturable soil bacteria. Trends in Biotechnology 30, 475-484, 10.1016/j.tibtech.2012.05.007.

Rachid, C.T., Santos, A.L., Piccolo, M.C., Balieiro, F.C., Coutinho, H.L., Peixoto, R.S., Tiedje, J.M., Rosado, A.S., 2013. Effect of sugarcane burning or green harvest methods on the Brazilian Cerrado soil bacterial community structure. PloS One 8, e59342, 10.1371/journal.pone.0059342.

Rensing, C., McDevitt, S.F., 2013. The copper metallome in prokaryotic cells. Metallomics and the Cell, 417-450.

Ritz, K., 2007. The plate debate: cultivable communities have no utility in contemporary environmental microbial ecology. FEMS Microbiology Rcology 60, 358-362, 10.1111/j.1574-6941.2007.00331.x.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584-e258410.7717/peerj.2584.

Romero, M.F., Gallego, D., Lechuga-Jiménez, A., Martínez, J.F., Barajas, H.R., Hayano-Kanashiro, C., Peimbert, M., Cruz-Ortega, R., Molina-Freaner, F.E., Alcaraz, L.D., 2021. Metagenomics of mine tailing rhizospheric communities and its selection for plant establishment towards bioremediation. Microbiological research 247, 126732https://doi.org/10.1016/j.micres.2021.126732.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4, 406-42510.1093/oxfordjournals.molbev.a040454.

Saw, J.H.W., 2021. Characterizing the Uncultivated Microbial Minority: towards Understanding the Roles of the Rare Biosphere in Microbial Communities. mSystems 6, e007732110.1128/mSystems.00773-21.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics (Oxford, England) 30, 2068-206910.1093/bioinformatics/btu153.

Shade, A., Hogan, C.S., Klimowicz, A.K., Linske, M., McManus, P.S., Handelsman, J., 2012. Culturing captures members of the soil rare biosphere. Environmental microbiology 14, 2247-225210.1111/j.1462-2920.2012.02817.x.

Shim, D., Kim, S., Choi, Y.-I., Song, W.-Y., Park, J., Youk, E.S., Jeong, S.-C., Martinoia, E., Noh, E.-W., Lee, Y., 2013. Transgenic poplar trees expressing yeast cadmium factor 1 exhibit the characteristics necessary for the phytoremediation of mine tailing soil. Chemosphere 90, 1478-1486https://doi.org/10.1016/j.chemosphere.2012.09.044.

Siles, J.A., Margesin, R., 2018. Insights into microbial communities mediating the bioremediation of hydrocarbon-contaminated soil from an Alpine former military site. Applied microbiology and biotechnology 102, 4409-442110.1007/s00253-018-8932-6.

Siles, J.A., Rachid, C.T.C.C., Sampedro, I., García-Romera, I., Tiedje, J.M., 2014. Microbial diversity of a Mediterranean soil and its changes after biotransformed dry olive residue amendment. PloS one 910.1371/journal.pone.0103035.

Singh, A.K., Garg, N., Lal, R., 2015. Pontibacter chinhatensis sp. nov., isolated from pond sediment containing discarded hexachlorocyclohexane isomer waste. Int J Syst Evol Microbiol 65, 2248-225410.1099/ijs.0.000241.

Spini, G., Spina, F., Poli, A., Blieux, A.-L., Regnier, T., Gramellini, C., Varese, G.C., Puglisi, E., 2018. Molecular and Microbiological Insights on the Enrichment Procedures for the Isolation of Petroleum Degrading Bacteria and Fungi. Frontiers in microbiology 910.3389/fmicb.2018.02543.

Steen, A.D., Crits-Christoph, A., Carini, P., DeAngelis, K.M., Fierer, N., Lloyd, K.G., Cameron Thrash, J., 2019. High proportions of bacteria and archaea across most biomes remain uncultured. The ISME journal10.1038/s41396-019-0484-y.

Stefani, F.O.P., Bell, T.H., Marchand, C., de la Providencia, I.E., El Yassimi, A., St-Arnaud, M., Hijri, M., 2015. Culture-Dependent and -Independent Methods Capture Different Microbial Community Fractions in Hydrocarbon-Contaminated Soils. PloS one 10, e012827210.1371/journal.pone.0128272.

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., Ostell, J., 2016. NCBI prokaryotic genome annotation pipeline. Nucleic acids research 44, 6614-662410.1093/nar/gkw569.

Teixeira, C., Almeida, C.M., Nunes da Silva, M., Bordalo, A.A., Mucha, A.P., 2014. Development of autochthonous microbial consortia for enhanced phytoremediation of salt-marsh sediments contaminated with cadmium. The Science of the total environment 493, 757-76510.1016/j.scitotenv.2014.06.040.

Thavamani, P., Malik, S., Beer, M., Megharaj, M., Naidu, R., 2012. Microbial activity and diversity in long-term mixed contaminated soils with respect to polyaromatic hydrocarbons and heavy metals. Journal of environmental management 99, 10-17, https://doi.org/10.1016/j.jenvman.2011.12.030.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680, 10.1093/nar/22.22.4673.

Thrash, J.C., 2019. Culturing the Uncultured: Risk versus Reward. mSystems 4, e00130-00119doi:10.1128/mSystems.00130-19.

Thrash, J.C., 2020. Towards culturing the microbe of your choice. Environmental Microbiology Reports n/ahttps://doi.org/10.1111/1758-2229.12898.

Valls, M., De Lorenzo, V., 2002. Exploiting the genetic and biochemical capacities of bacteria for the remediation of heavy metal pollution. FEMS microbiology Reviews 26, 327-338.

van Bruggen, A., Ochoa, O., Francis, I., Michelmore, R.W., 2014a. Differential interactions between strains of R hizorhapis, S phingobium, S phingopyxis or R hizorhabdus and accessions of L actuca spp. with respect to severity of corky root disease. Plant pathology 63, 1053-1061.

van Bruggen, A.H.C., Francis, I.M., Jochimsen, K.N., 2014b. Non-pathogenic rhizosphere bacteria belonging to the genera Rhizorhapis and Sphingobium provide specific control of lettuce corky root disease caused by species of the same bacterial genera. Plant Pathology 63, 1384-1394, 10.1111/ppa.12212.

van Bruggen, A.H.C., Francis, I.M., Krag, R., 2015. The vicious cycle of lettuce corky root disease: effects of farming system, nitrogen fertilizer and herbicide. Plant and Soil 388, 119-132, 10.1007/s11104-014-2312-5.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73, 5261-5267, 10.1128/aem.00062-07.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study. Journal of Bacteriology 173, 697-703, 10.1128/jb.173.2.697-703.1991.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wiegand, S., Jogler, M., Boedeker, C., Pinto, D., Vollmers, J., Rivas-Marín, E., Kohn, T., Peeters, S.H., Heuer, A., Rast, P., Oberbeckmann, S., Bunk, B., Jeske, O., Meyerdierks, A., Storesund, J.E., Kallscheuer, N., Lücker, S., Lage, O.M., Pohl, T., Merkel, B.J., Hornburger, P., Müller, R.-W., Brümmer, F., Labrenz, M., Spormann, A.M., Op den Camp, H.J.M., Overmann, J., Amann, R., Jetten, M.S.M., Mascher, T., Medema, M.H., Devos, D.P., Kaster, A.-K., Øvreås, L., Rohde, M., Galperin, M.Y., Jogler, C., 2019. Cultivation and functional characterization of 79 planctomycetes uncovers their unique biology. Nature Microbiology, 5, 126–140, 10.1038/s41564-019-0588-1.

Xia, M., Liu, Y., Taylor, A.A., Fu, D., Khan, A.R., Terry, N., 2017. Crude oil depletion by bacterial strains isolated from a petroleum hydrocarbon impacted solid waste management site in California. International Biodeterioration and Biodegradation 123, 70-77, 10.1016/j.ibiod.2017.06.003.

Ye, S., Zeng, G., Wu, H., Zhang, C., Dai, J., Liang, J., Yu, J., Ren, X., Yi, H., Cheng, M., Zhang, C., 2017. Biological technologies for the remediation of co-contaminated soil. Critical Reviews in Biotechnology: 37, 1062-1076, 10.1080/07388551.2017.1304357.

Youseif, S.H., Abd El-Megeed, F.H., Humm, E.A., Maymon, M., Mohamed, A.H., Saleh, S.A., Hirsch, A.M., 2021. Comparative Analysis of the Cultured and Total Bacterial Community in the Wheat Rhizosphere Microbiome Using Culture-Dependent and Culture-Independent Approaches. Microbiology Spectrum 9, e006782110.1128/Spectrum.00678-21.

Fig.1

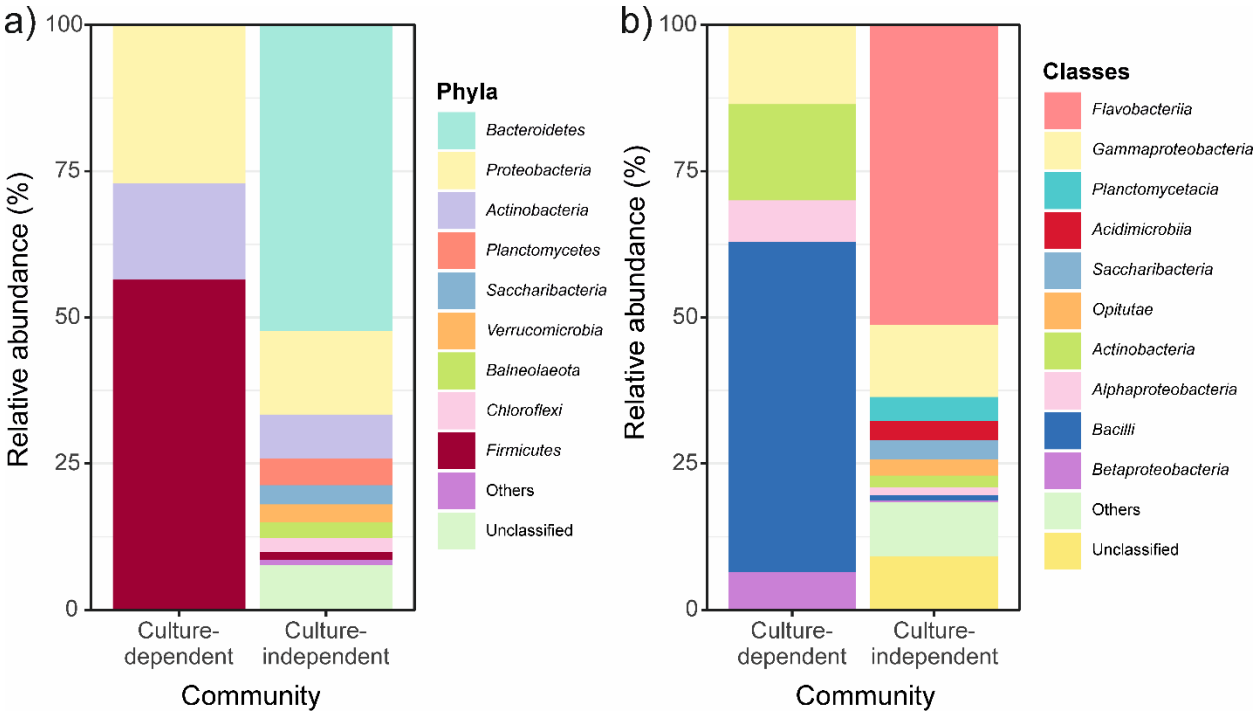

Fig. 1. Relative abundance of the phyla (a) and classes (b) conforming culture-dependent (culturing) and -independent (16S rRNA gene amplicon sequencing) bacterial communities in the studied waste sediment.

Fig. 2

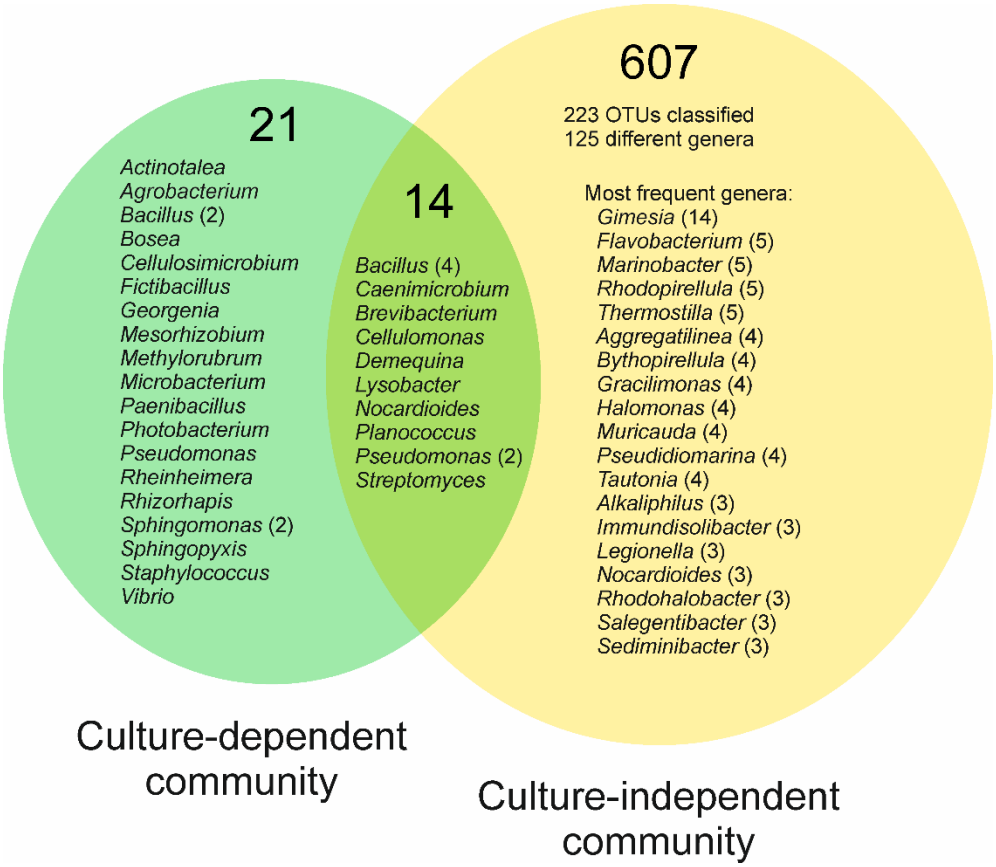

Fig. 2. Venn diagram showing the number and taxonomic affiliation (at genus level) of unique and shared OTUs in culture-dependent (culturing) and -independent (16S rRNA gene amplicon sequencing) bacterial communities from the studied waste sediment. Numbers in parenthesis denote the numbers of times that a specific genus was found.

Fig. 3

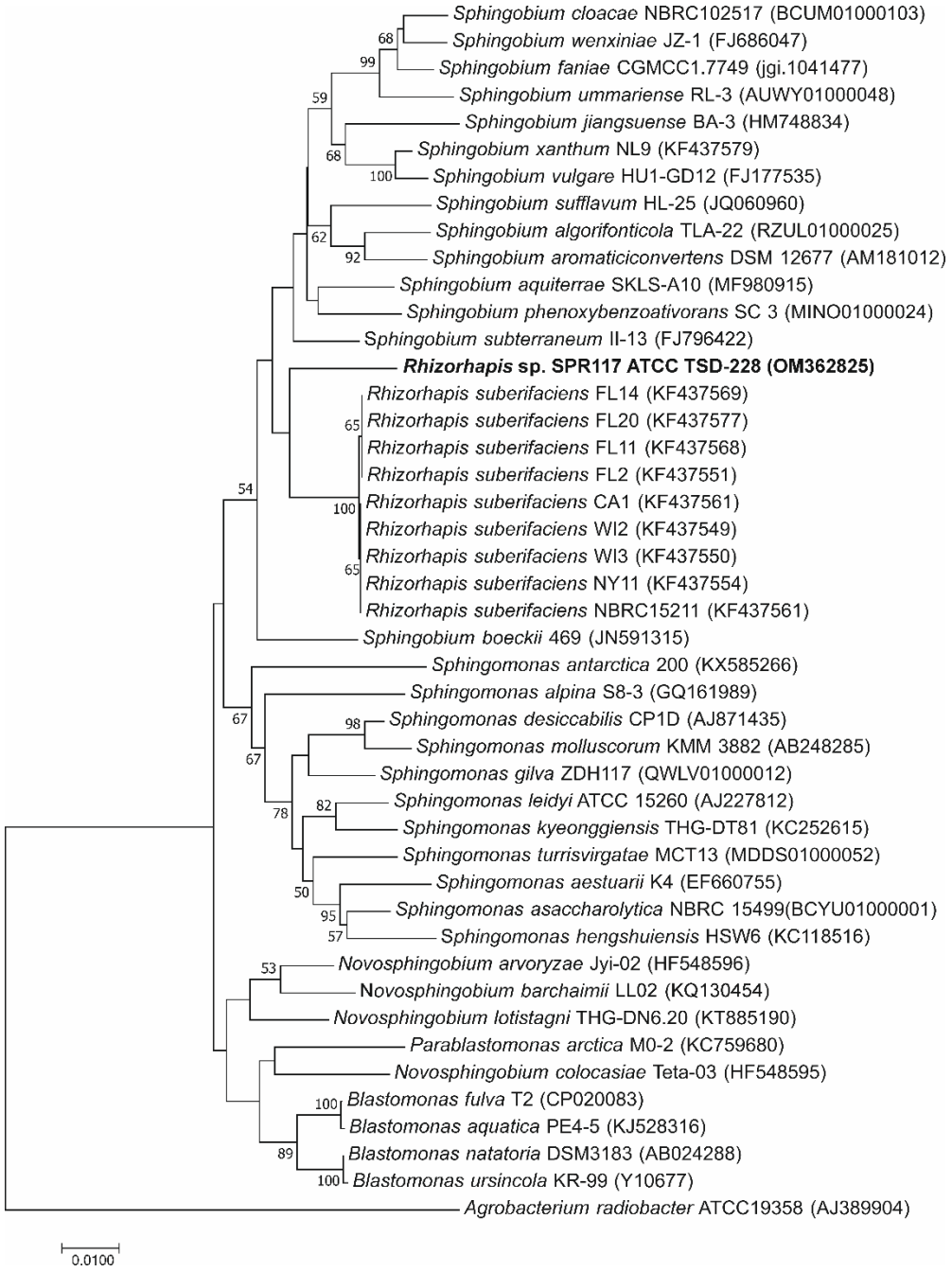

Fig. 3. Neighbour-joining tree, based on 16S rRNA gene sequence data, showing the phylogenetic position of *Rhizorhapis* sp. strain SPR117 and most related microorganisms according to EzBioCloud. The genetic distances were calculated by using Kimura's two-parameter model and the complete deletion option. Bootstrap values (calculated with 1000 resamplings) above 50% are shown in branch positions. *Agrobacterium radiobacter* ATCC19358 was used as outgroup organism. GenBank accession number of 16S rRNA gene sequences are given in parentheses. Bar, 0.01 expected changes per site.

Fig. 4

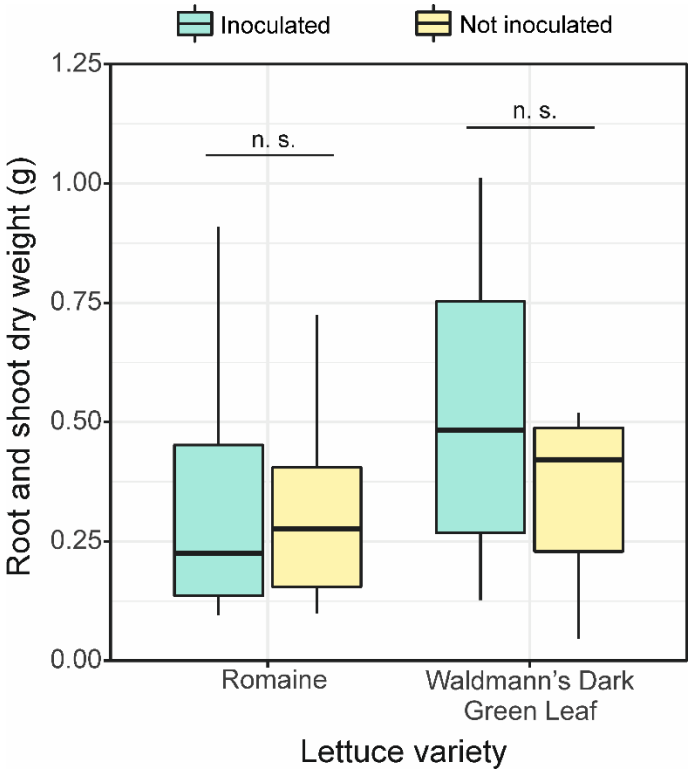

Fig. 4. Box plots comparing root and shoot dry weight of lettuce plants of the varieties Romaine and Waldmann's Dark Green Leaf after growing 28 days in sterilized potting substrate inoculated or not with *Rhizorhapis* sp. strain SPR117. The boxes represent the interquartile range (IQR) between the first and third quartiles (25th and 75th percentiles, respectively) and the vertical line inside the box defines the median. Whiskers represent the lowest and highest values within 1.5 times the IQR from the first and third quartiles, respectively. No significant differences were found between inoculated and non-inoculated plants at $p < 0.05$ according to PERMANOVA.

Tables

Table 1. Taxonomic information and relative abundance of the culturable OTUs and their phylogenetic relationship with the culture-independent OTUs.

Culturable OTU id.	**Number of isolates in the OTU**	**Relative abundance (%) of the culturable OTU**	**Representative strain in the OTU**	**Closest relative Eztaxon match (GenBank accession no.), %similarity of the representative strain**	**Phylum**	**Metabarcoding OTU related to the culturable OTU**	**Relative abundance (%) of the metabarcoding-based OTUs**
Otu1	39	27.86	SPR114	*Bacillus thioparans* BMP-1 (DQ371431), 100	*Firmicutes*	Otu440	0.05
Otu2	17	12.14	SPL145	*Bacillus oryzaecorticis* R1 (KF548480), 99.11	*Firmicutes*	Otu287	0.03
Otu3	13	9.29	SPR106	*Fictibacillus phosphorivorans* Ca7T (JX258924), 99.70	*Firmicutes*		
Otu4	7	5.00	SPR136	*Demequina activiva* BS-12M (KM591918), 99.35	*Actinobacteria*	Otu88	0.07
Otu5	6	4.29	SPR6	*Pseudomonas alcaligenes* NBRC 14159 (BATI01000076), 98.97	*Proteobacteria*	Otu36	0.23
Otu6	9	6.43	SPR140	*Caenimicrobium hargitense* CGII-59m2 (KM083134), 97.81	*Proteobacteria*	Otu307	0.01
Otu7	4	2.86	SPL106	*Lysobacter arseniciresistens* ZS79 (AVPT01000055), 97.27	*Proteobacteria*	Otu26	0.39
Otu8	4	2.86	SPL34	*Pseudomonas kunmingensis* HL22-2 (JQ246444), 99.14	*Proteobacteria*	Otu93	0.28
Otu9	5	3.57	SPR46	*Microbacterium oxydans* DSM 20578 (Y17227), 99.82	*Actinobacteria*		
Otu10	2	1.43	SPR17	*Bacillus zhangzhouensis* DW5-4 (JOTP01000061), 99.91	*Firmicutes*	Otu543	0.01
Otu11	1	0.71	SPL5	*Planococcus citreus* DSM 20549 (RCCP01000013), 100	*Firmicutes*	Otu580	0.02
Otu12	1	0.71	SPL154	*Streptomyces yanii* NBRC 14669 (AB006159), 100	*Actinobacteria*	Otu57	0.15
Otu13	1	0.71	SPL27	*Bacillus firmus* NBRC 15306 (BCUY01000205), 99.40	*Firmicutes*	Otu360	0.05
Otu14	2	1.43	SPR19	*Bacillus onubensis* 0911MAR22V3 (NSEB01000017), 100	*Firmicutes*		
Otu15	2	1.43	SPR125	*Mesorhizobium sediminum* YIM M12096 (KX151664), 99.72	*Proteobacteria*		
Otu16	2	1.43	SPR132	*Georgenia muralis* DSM 14418 (RKRA01000001), 99.33	*Actinobacteria*		
Otu17	1	0.71	SPL20	*Brevibacterium sandarakinum* DSM 22082 (LT629739), 97.03	*Actinobacteria*	Otu202	0.01
Otu18	1	0.71	SPR144	*Cellulomonas carbonis* T26 (HQ702749), 99.25	*Actinobacteria*	Otu410	0.01
Otu19	2	1.43	SPL19	*Staphylococcus saprophyticus* ATCC 15305 (AP008934), 100	*Firmicutes*		

Otu20	2	1.43	SPL22	*Pseudomonas oleovorans* subsp. *lubricantis* RS1 (DQ842018), 98.45	*Proteobacteria*		
Otu21	2	1.43	SPR21	*Agrobacterium salinitolerans* YIC 5082 (MRDH01000011), 100	*Proteobacteria*		
Otu22	1	0.71	SPR141	*Nocardioides massiliensis* GD13 (CCXJ01000100), 97.29	*Actinobacteria*	Otu172	0.02
Otu23	3	2.14	SPR42	*Actinotalea ferrariae* CF5-4 (HQ730135), 98.03	*Actinobacteria*		
Otu24	1	0.71	SPL24	*Bosea robiniae* DSM 26672 (jgi.1085745), 99.18	*Proteobacteria*		
Otu25	1	0.71	SPL31	*Photobacterium halotolerans* MACL01 (AY551089), 97.65	*Proteobacteria*		
Otu26	1	0.71	SPL32	*Rheinheimera aquimaris* SW-353 (EF076757), 99.23	*Proteobacteria*		
Otu27	1	0.71	SPL37	*Vibrio japonicus* JCM 31412 (LC143378), 99.09	*Proteobacteria*		
Otu28	1	0.71	SPR109	*Sphingomonas laterariae* LNB2 (jgi.1118286), 97.66	*Proteobacteria*		
Otu29	1	0.71	SPR115	*Sphingopyxis alaskensis* RB2256 (CP000356), 99.83	*Proteobacteria*		
Otu30	1	0.71	SPR116	*Sphingomonas rubra* CGMCC 1.9113 (jgi.1058074), 97.68	*Proteobacteria*		
Otu31	1	0.71	SPR117	*Rhizorhapis suberifaciens* CA1 (KF437561), 96.88	*Proteobacteria*		
Otu32	1	0.71	SPR13	*Methylorubrum salsuginis* MR (EF015478), 99.19	*Proteobacteria*		
Otu33	1	0.71	SPR130	*Paenibacillus typhae* CGMCC 1.11012 (jgi.1076256), 98.34	*Firmicutes*		
Otu34	1	0.71	SPR131	*Bacillus niacini* IFO 15566 (AB021194), 99.51	*Firmicutes*		
Otu35	2	1.43	SPR126	*Cellulosimicrobium cellulans* LMG 16121 (CAOI01000359), 100	*Actinobacteria*		

785

Table 2. Assembly and annotation metrics of *Rhizorhapis* sp. strain SPR117

Genome feature	*Rhizorhapis* sp. strain SPR117
Sequence length (bp)	4,419,522
Contigs	256
Completeness (%)	98.4
Fine consistency (%)	96.2
GC content (%)	59.86
Longest contig length (bp)	190,373
N_{50}	62,062
L_{50}	23
Genes (total)	4,452
CDS (total)	4,399
Genes (coding)	4,278
CDS (coding)	4,278
rRNAs	3 (1 5S, 1 16S, 1 23S)
tRNAs	46
ncRNAs	4
Pseudogenes	121