Plant Soil
https://doi.org/10.1007/s11104-021-05226-y

REGULAR ARTICLE

Noccaea caerulescens seed endosphere: a habitat for an endophytic bacterial community preserved through generations and protected from soil influence

Alexis Durand · Pierre Leglize · Séverine Lopez · Thibault Sterckeman · Emile Benizri

Received: 20 July 2021 / Accepted: 10 November 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract
Background and aims While our understanding of seed microbiota has lagged far behind that of the rhizosphere and phyllosphere, many advances are now being made, particularly based on metagenomics studies. Today, our knowledge of seed microbiome assembly remains incomplete and the connections between seed and soil microbiomes are not yet fully understood, especially where hyperaccumulating plants are concerned. In this work, we assessed the structure and composition of the *Noccaea caerulescens* rhizosphere and endosphere-associated microbiota.
Methods A pot experiment was conducted for 6 months in a growth chamber, using two populations of the hyperaccumulator *Noccaea caerulescens* growing on their original soil (calamine or nonmetalliferous) and vice versa. The diversity of rhizosphere soil bacteria and bacterial endophytic communities present in the different habitats (initial seed, root, stem, leaves and new seed generation) was characterized by high-throughput 16S rRNA amplicon sequencing.
Results Bacterial communities from root endosphere, stem endosphere and leaf endosphere appeared to be soil-type dependent, contrary to the bacterial communities associated with seed endosphere habitats (initial seeds and new seed generation). Moreover, the seed endophytic bacterial communities of *Noccaea caerulescens* display a strong heritability across one plant generation. Indeed, a bacterial endophytic core-genome globally appeared to be constant between initial seeds and those obtained after the first generation.
Conclusion Our results suggest that *Noccaea caerulescens* may carry a selected bacterial community in its seeds across generations, despite soil environment changes.

Keywords Bacterial diversity · Hyperaccumulator · Endophyte · Trace element · High throughput sequencing · Seed

Responsible Editor: Stéphane Compant

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11104-021-05226-y.

A. Durand · P. Leglize · T. Sterckeman · E. Benizri (✉)
Université de Lorraine, INRAE, LSE, F-54000 Nancy, France
e-mail: emile.benizri@univ-lorraine.fr

S. Lopez
INRAE, Bordeaux Sciences Agro, ISVV, SAVE, F-33140 Villenave d'Ornon, France

Introduction

More than 700 plant taxa that hyperaccumulate one or more heavy metals and metalloids in their aerial parts have so far been identified (Krämer 2010; Reeves et al. 2018). These plants have been described from

Published online: 21 January 2022

metalliferous soils in disparate geographical regions (Reeves et al. 2021). *Noccaea caerulescens* (J. & C. Presl) F. K. Meyer is a pseudometallophyte species occurring on non-metalliferous soils, as well as on serpentine soil which is naturally enriched with nickel (Ni) or on calamine soil enriched following anthropogenic activities with cadmium (Cd), lead (Pb) and zinc (Zn) (Reeves et al. 2001). This species is not only tolerant to Cd, Pb, Ni and Zn, but is also able to hyperaccumulate these metals (apart from Pb). This hyperaccumulating plant species, which is a member of the *Brassicaceae* family, has been at the forefront of research concerning hyperaccumulation and recognized as a model species when studying hyperaccumulation (Assunção et al. 2003; Peer et al. 2003).

Plants have co-evolved with complex microbial communities. Recently, the holobiont and hologenome concepts have become widely used by the scientific community. In fact, host plant and microorganisms could be considered together as plant-microbiome superorganisms (Bosch and McFall-Ngai 2011; Durand et al. 2021b). The associated microbial part can increase the adaptability of the plant-microbiome superorganism to its environment and the microbiota can benefit plant fitness (Herrera Paredes and Lebeis 2016). Indeed, the plant microbiota can influence multiple plant traits such as biomass by producing phytohormones and enzymes, such as indoleacetic acid (IAA) (Asaf et al. 2017; Defez et al. 2017; Hamayun et al. 2017) and 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (Glick et al. 1997; Zhang et al. 2011), by inducing the production of particular metabolites, such as salicylic and jasmonic acids (Badri et al. 2013; Fesel and Zuccaro 2016), by increasing drought and /or metal tolerance (Rolli et al. 2015; Ma et al. 2017; Wang et al. 2020), by modifying the flowering time (Panke-Buisse et al. 2015; Dombrowski et al. 2017) and by improving disease resistance (Busby et al. 2016; Ritpitakphong et al. 2016).

Seed endophytic bacterial communities have been little studied to date, particularly in the case of hyperaccumulating plants. Endophytes refer to those microbes which inhabit the interior of plant tissues and form a range of different relationships with the host plant including symbiotic, mutualistic, and commensal links, without causing any harm to the host (Smith 1911; Perotti 1926). In addition, evidence has shown that such endophytes isolated from hyperaccumulators possess particular traits (i.e., Plant Growth Promoting traits) which promote plant growth by various mechanisms such as nitrogen fixation, solubilization of minerals such as phosphate, production of phytohormones and siderophores, utilization of 1-aminocyclopropane-1-carboxylic acid as sole nitrogen source (Santoyo et al. 2016; Yadav 2017). Moreover, Wang et al. (2020) showed that the endophytic bacteria *Sphingomonas* SaMR12 inoculation could alleviate metal stress of the Zn-hyperaccumulator *Sedum alfredii* by decreasing shoot hydrogen peroxide (H_2O_2) content and inducing a decrease in the membrane lipid peroxidation. Thus, several studies have focused on the benefits for the hyperaccumulating plants of endophytic microbes and their role in metals hyperaccumulation (Lodewyckx et al. 2002b; Weyens et al. 2009; Visioli et al. 2014; Ma et al. 2015b; Benizri et al. 2021; Durand et al. 2021a).

Concerning endophytic bacteria, colonization of the internal tissues of plants could confer an ecological advantage over bacteria that can only colonize plants epiphytically (i.e., rhizosphere bacteria), by providing a more protective environment against extreme environmental conditions such as high soil metal concentrations (Lodewyckx et al. 2002b). Nonetheless, if the interior of plant tissues constitutes a protective niche for endophytic bacteria, the survival of those endophytes will require specific adaptation. In their study, Ma et al. (2015a) reported that several bacterial endophytes, isolated from the hyperaccumulator *Sedum plumbizincicola,* exhibited metal resistance and were able to grow in a medium supplemented with 100 mg L^{-1} of Cd or Zn. Similarly, high Cd tolerance was found for the bacterial endophytes of a Cd-hyperaccumulator (*Solanum nigrum*) growing on mine tailings (Luo et al. 2011). Thus, high metal tolerance is generally associated with bacterial growth within metal-hyperaccumulating plants. However, there was some difference in the resistance-level depending on the plant compartment from which endophytes were isolated. In their study, Lodewyckx et al. (2002a) isolated endophytic inhabitants of the roots and shoots of the Zn-hyperaccumulator *Noccaea* (previously named *Thlaspi*) *caerulescens* subsp. *calaminaria*. Although similar endophytic species were isolated from the two compartments, those from roots showed lower resistance to Zn and Cd than the endophytic bacteria isolated from the shoots. Furthermore, they showed that some species were exclusively

abundant in shoots. As these authors pointed out, these differences might possibly reflect altered metal speciation in the different plant compartments.

Endophytes colonize plant tissues in several ways (Frank et al. 2017), roots being the primary site. This might explain similarities between endophytic and rhizosphere-colonizing bacteria, as many facultative endophytic bacteria can survive as rhizosphere bacteria. Concerning the root compartment, the main entrance for endophytic bacteria appears to be through wounds that occur naturally during plant growth, or through root hairs and at epidermal junctions (Kandel et al. 2017). In addition, some of the epiphytic microorganisms can become endophytes of plant aerial parts through leaf stomata, lenticel and hydathode penetration (Hardoim 2011). Endophytes can also be carried by various living vectors, such as phloem-feeding insects (Lòpez-Fernàndez et al. 2017), or herbivorous mammals (Frank et al. 2017). Once inside the plant, endophytes either reside in specific plant tissues like root cortex and xylem or colonize the plant systematically by transport through the vascular system or the apoplast. The endophytes acquired from both aboveground and underground plant environments, living within the plant tissues for only one part of their life cycle, are said to be transmitted horizontally, whereas those acquired by seeds, inhabiting plant tissues at all stage of the plant life cycle, are said to be transmitted vertically (Tannenbaum et al. 2020). Endophytes transmitted vertically in the seeds could have co-evolved with their host plant and shared a symbiotic relationship with it across generations (Wani et al. 2015; Berg et al. 2020).

Recently, microbiome research has revolutionized our understanding of plants as coevolved holobionts but also of indigenous microbiome-inoculant interactions and their potential for sustainable agriculture (Berg et al. 2021). Still, our knowledge of seed microbiome assembly remains incomplete and the connections between the seed and soil microbiomes are not yet fully understood, especially for hyperaccumulating plants (Nelson 2017; Nelson et al. 2018). Studies have indicated that bacteria may be recruited from the soils on which plants are grown (Klaedtke et al. 2016), but that neither environmental conditions nor host genotypes fully explain the assembly of the bacterial seed microbiota (Johnston-Monje et al. 2016). Nevertheless, it has been shown for several plants that the seed microbiome was plant dependent and was stable across sites and plant generations (Sánchez-López et al. 2018b; Walitang et al. 2018). This bring evidence that, at least in some cases, a part of the endophytic community is conserved across plant generations and originates from a seed stock. The study of 14 populations of seeds belonging to three genetic subunits of the hyperaccumulator *Nemoria caerulescens* and recovered across France from calamine and non-metalliferous sites, showed that the composition of the seed endophytic bacterial communities were very similar for all plant populations, whatever the edaphic characteristics and soil properties (Durand et al. 2021b). Moreover, 89% of the characterized seed OTUs were shared between all seed populations.

In this study, we used seeds belonging to two French populations of *N. caerulescens* and recovered from their native ecosystems, i.e., from a population growing on a non-metalliferous soil (Pic de Chenavari, France) and from another one on calamine soil (Largentière, France). The aim of this study was to characterize the endophytic bacterial communities in different plant parts (seed, root, shoot, leaf and the new seed generation) of the hyperaccumulator pseudometallophyte *N. caerulescens* by high-throughput 16S rRNA amplicon sequencing. The seeds of each population were sown on their original soil, but also on the original soil of the other population with the hypothesis that, an endophytic core microbiome in the new seed generation will be shared between the two populations and with the native seeds, whatever the soil type.

Materials and methods

Soil characteristics and experimental design

Seeds of the hyperaccumulating plant *N. caerulescens* (*Brassicaceae*) were collected from two plant populations from two sites in Ardèche in France (Gonneau et al. 2017). The first plant population develops on a non-metalliferous soil from the Chenavari site (abbreviated 'C', 44°35′58.80"N 04°41′04.20″E) and the second population on a calamine soil from the Largentière site (abbreviated 'L', 44°32′26.33"N 04°18′18.44″E). These two plant populations belong to the same genetic subunit (Gonneau et al. 2017). A pot experiment (1 L) was conducted for 6 months

in a growth chamber. The soils used for the experiment were collected from the top layer (10–40 cm) of the two sites around plant specimens. These soils were sieved to <5 mm and stored at 4 °C until analyses were carried out. Soil physicochemical properties (pH NF ISO 10390, organic carbon NF ISO 10694, total nitrogen NF ISO 13878, soil total trace element concentrations after hydrofluoric acid extraction NF ISO 22036 and NF EN ISO 1729-2, exchangeable cation concentrations using a hexammine cobalt trichloride solution NF ISO 23470 and exchangeable trace element concentrations using DTPA solution NF ISO 14870:2001) were determined by the Soil Analysis Laboratory of INRAE (Arras, France), (Tables 1, 2 and 3). The treatments were as follows: seeds of *N. caerulescens* originating from 'C' soil were sown in pots containing 'C' (518 g dry weight of soil) or 'L' (790 g dry weight of soil) soils and seeds of *N. caerulescens* originating from 'L' soil were also sown on both 'C' and 'L' ones. Thus, four treatments were considered: SCGC (initial soil from Chenavari site with initial seeds from Chenavari), SCGL (initial soil from Chenavari site with initial seeds from Largentière), SLGC (initial soil from Largentière site with initial seeds from Chenavari) and SLGL (initial soil from Largentière site with initial seeds from Largentière). For each treatment, only one plant per pot was grown, without any supply of nitrogen fertilizer. The experiment had a completely randomized block design with initially seven replicates of the four treatments, but some mesocosms failed to reach growth expectation and were consequently excluded from sampling, (Table 4). The growth conditions were: photoperiod 16 h, temperature 15 °C night and 20 °C day, relative humidity 70%, photosynthetic photon flux density: 350 mmol m^{-2} s^{-1}. The pots were watered daily with distilled water to 80% of soil water holding capacity.

Harvest, surface sterilization of plant organs and grinding of plant parts

Bacterial communities were analyzed in initial bulk soil (iS) and initial seed endosphere (iSE) samples from both sites (C and L). In addition, for the four treatments (SCGC, SCGL, SLGC, and SLGL), bacterial communities were analyzed in rhizosphere soil (RS), root endosphere (RE), stem endosphere (StE), leaf endosphere (LE) and seed endosphere (SE) samples by 16S rRNA gene amplicon sequencing.

Table 1 Altitude, soil pH, organic carbon (g kg^{-1}), total nitrogen (g kg^{-1}), C/N and total concentrations (mg kg^{-1}) of micro and macro elements for each site where seeds and soils were collected

Site	Altitude	pH	orga C	tot N	C/N	Al	Fe	Cr	Cu	Zn	Co	Pb	Cd
Chenavari	460	6.39	67.4	5.94	11.3	86,000	94,600	196	42.1	114	42.7	56.6	0.259
Largentière	264	6.8	18.8	0.64	24.8	31,900	7400	12	118	6670	5.53	39,600	37.9

Table 2 CEC (cmol+ kg^{-1}) and exchangeable concentrations with CEC (cmol+ kg^{-1}) of micro and macro elements for each site where seeds and soils were collected

Site	Ca	Mg	Na	K	Fe	Mn	Al
Chenavari	21.9	5.69	0.133	0.331	0.0069	0.0084	0.0369
Largentière	2.94	0.489	0.0076	0.0981	0.0144	0.006	0.0614

Table 3 Exchangeable elements with DTPA (mg kg^{-1}) of micro and macro elements for each site where seeds and soils were collected

Site	Cd	Cu	Ni	Pb	Zn	Cr	Fe	Mn
Chenavari	0.0732	1.02	1.02	4.07	2.57	< 0.2	186	4.66
Largentière	7.52	10.5	0.16	984	601	< 0.2	5.16	0.677

Table 4 Number of samples per modality. Compared to the initial design (7 replicates for the four treatments) some mesocosms did not reach growth expectation and were consequently not sampled. Three replicates were sampled for the initial soils and initial seeds. Negative numbers between brackets correspond to samples that were discarded during bioinformatic treatments, when data were subsampled at 3000 reads. Abbreviation for the four treatments are, SC: initial Soil from Chenavari site, SL: initial Soil from Largentière site, GC: initial « Graines » from Chenavari site (Seeds are translated into "Graines" in French), GL: initial « Graines » from Largentière site

Habitats	SCGC	SCGL	SLGC	SLGL	Total
Initial bulk soil (iS)		3		3	6
Initial seed endosphere (iSE)		4		4	8
Rhizosphere soil (RS)	7	7 (−1)	7	7	28
Root endosphere (RE)	7	7	4	4 (−1)	22
Stem endosphere (StE)	7	7	4	4	22
Leaf endosphere (LE)	7	7	4	4	22
Seed endosphere (SE)	4	4	3	3	14
					122 (−2)

Recovery of the bacterial DNA used to determine the composition of endophyte communities from those habitats is described below.

After the flowering stage (29 weeks after sowing) and the formation of the new seed generation, seeds were collected and the whole plants were carefully taken out of the pot and cut at the root/shoot boundary, then roots were shaken to discard the non-adhering soil. Moreover, rhizosphere soil, defined as the soil closely attached to each root system was also recovered. The root/shoot parts were washed clean of soil under tap water and gently blotted. Then the leaves were separated from the stems. Roots, stems and leaves were placed in separate 50 mL tubes. Seeds were recovered from the plants just before falling, the pods were opened, and undamaged seeds were selected for surface sterilization and transferred to 2 mL tubes.

Fresh tissues (50 mg of seeds, approximately 5 g for roots, stems and leaves), were immersed for 30 s in 1% NaClO solution supplemented with Triton 100X 0.1%, then washed for 30 s with EtOH (96%) and rinsed five times with sterile deionized water. Tissue sterilization was confirmed by plating 100 µL of the final rinsing water on 10% TSA medium (Tryptone Soy Agar) and by running a PCR on the last rinsing water (Sánchez-López et al. 2018a). PCR was design to amplify the bacterial 16S rRNA gene using the following primers: 27f (5′- AGA GTT TGA TCA TGG CTC A – 3′) and 1492r (5′- TAC GGT TAC CTT GTT ACG ACT T – 3′) (Eurofins Genomics, Paris, France) and by using the thermoscientific DreamTaq™ Green PCR Master Mix (2X) kit (Thermo Fisher Scientific, Carlsbad, California). For each PCR mix, 25 µL of Dream Taq Green master mix were used, each universal primer was adjusted to 0.5 µM, 5 µL of the last rinsing water were added and final volume was adjusted to 50 µL with nuclease-free water. DNA amplification was carried out in a thermocycler (Mastercycler gradient, Eppendorf, Hamburg, Germany) under the following conditions: 95 °C 2 min, 30 cycles 95 °C 30 s, 53 °C 30 s, and 72 °C 1 min, with an additional 10 min at 72 °C. Bacterial DNA from a previously isolated strain was used as positive control. Plant parts (seeds, roots, stems and leaves) and rhizosphere soil were stored at −80 °C. Plant parts that were successfully sterilized were then lyophilized, and ground in sterile conditions into a homogenous powder using a Mixer Mill for 30s at 30 Hz (model MM400; Retsch Inc., Newtown, Pennsylvania, USA). Grinding was carried out in 2 mL tubes with 0.5 mm zirconium beads previously rinsed with EtOH (96%). After removal of the beads, these homogenous powders and the previously recovered soil were used for the following steps.

DNA extraction and metabarcoding of endophytic bacterial communities

The DNA was obtained using precisely 250 mg dry weight of soils or plant tissues (50 mg for seeds). Total DNA was extracted with a modified hexadecyltrimethylammonium bromide (CTAB) chloroform alcohol protocol (Healey et al. 2014). Briefly,

the extraction protocol required 1 h at 65 °C with multiple agitations in the CTAB buffer (2 g CTAB, 4 mL EDTA 0.5 M, 10 mL TrisHCl 1 M and 86 mL NaCl 1.4 M in 100 mL), a heat shock (−80 °C to 65 °C) and enzymatic digestions with proteinase K, α-amylase from *Aspergillus oryzae* and RNAse A. The DNA precipitation was obtained firstly with isopropanol (at ambient temperature, 15 min) and next with ethanol 70% (at 4 °C, 30 min). A purification step was added using a QIAquick® PCR Purification Kit (Qiagen, Germany). The quantity and quality of the purified DNA were assessed using electrophoresis migration on a 1% agarose gel and with a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, made in USA). The PCR targeted the V5-V6 hypervariable regions of the 16S rRNA gene with chloroplast DNA excluding 799f (5′- AACMGGATTAGATACCCKG - 3′) and 1115r (5′- AGG GTT GCG CTC GTT G - 3′) primers, resulting in an amplicon of a small size (~316 bp), appropriate for Illumina sequencing (Kembel et al. 2014). The resulting amplicons were purified with AMpure magnetic beads (Agencourt) and pooled in equimolar concentrations before the sequencing performed with an Illumina MiSeq platform (ADNid, France). These libraries were mixed with Illumina-generated PhiX control libraries (5%) and sequenced with the MiSeq Reagent Kit V3–600 cycles (Illumina Inc., San Diego, USA).

Bioinformatic and statistical analysis of bacterial diversity

Reads were assigned to each of the 122 samples (Table 4) according to a unique barcode, and contigs were then assigned using the MOTHUR v.1.40.5 (last update 06/19/2018) (Schloss et al. 2009). Raw reads were joined then filtered by quality. Good quality sequences were then conserved if corresponding to the 16S rRNA gene, if long enough, if non-chimerical, and if at least 8 reads per sequence were found. OTUs were derived using Needleman distance and average neighbor clustering at a distance of 0.03. Taxonomic assignments were made with the SILVA ribosomal RNA databases v1.3.8 (Dec 16, 2019) (Quast et al. 2013).

The following analyses were performed based on a dataset in which the number of reads per sample were rarefied to 3000. OTU-based analysis of alpha diversity was performed with the following functions using Mothur calculators: ‘sobs’, “chao”, “ace” ‘shannoneven’, ‘shannon’, ‘invsimpson’, and ‘coverage’. These estimates included: observed OTU richness, Chao estimation of OTUs richness (Chao 1949), ACE (Abundance-based Coverage Estimator) estimation of OTU richness, Shannon diversity index, inverse Simpson diversity index, a measurement of evenness based on the Shannon index and coverage. The coverage calculator returned a Good’s coverage for an OTU definition (Good 1953). Coverage was calculated using the following equation: $C = [1 - (n/N)]*100$ (%), where ‘n’ is the number of OTUs and ‘N’ the number of sequences. For the following analyses, we used the R version 3.6.3 (latest update 02/29/2020) (R Core Team 2019). A 2-dimensional non-metric multi-dimensional scaling (NMDS) plot was calculated using the Bray Curtis method ($k = 3$) based on the standardized (Wisconsin double) and square root transformation of OTU abundance using the ‘metaMDS’ function in the ‘vegan’ package. We used the ‘anosim’ function in the ‘vegan’ package to perform an ANalysis Of SIMilarities (ANOSIM) to obtain P (i.e., significance levels) and R (i.e., the strength of the factors on the samples) values. A heatmap of Spearman’s correlations between the most abundant phyla in each modality was created with ‘heatmap.2’ from the ‘gplots’ package. The numbers of OTUs that were shared between modalities were visualized using Venn diagrams implemented in Mothur with the function ‘venn’. A network analysis was performed using the software QIIME (Quantitative Insights Into Microbial Ecology, version 1.8.0, (Caporaso et al. 2010) and was built with Cytoscape software version 3.8.2, (Shannon et al. 2003).

Results

Soil properties

Noccaea caerulescens seeds were collected from calamine and non-metalliferous sites. Soil pH was neutral to slightly acidic, which moderately promote the availability of trace elements (Table 1). The results showed that the two soils were clearly differentiated, with higher organic carbon concentration associated with lower C/N ratio for the non-metalliferous site than for the calamine one (Table 1). The higher C/N ratio observed for the calamine site could be related to

a decrease in the rate of degradation of organic matter due to higher trace elements concentrations (McEnroe and Helmisaari 2001). This trend was confirmed by the meta-analysis done by Zhou et al. (2016) suggesting that the larger C/N ratio in metal-polluted soil than in control soil might arise from changes in the rate of degradation of plant residues. HF-total (hydrofluoric extraction) and DTPA-exchangeable concentrations of Cd, Cu, Pb and Zn were much higher in the calamine soil than in the non-metalliferous one (Tables 1 and 3). Higher levels of Cd, Zn and Pb are characteristic of calamine edaphic conditions (Baker and Brooks 1989). Furthermore the Ca, Mg, K and Na exchangeable concentrations were lower for the calamine soil compared to non-metalliferous one (Table 2), this lower nutritional status perhaps enhancing the toxicity of divalent heavy metal cations (Preite et al. 2019). This distinction between edaphic types could be linked to the history of the two sites: previous industrial activities (mining/smelting) for Largentière and natural conditions for Chenavari.

Sequencing results and driving factors of the bacterial diversity

At the end of the experiment, a total of 122 samples were obtained (Table 4). Illumina sequencing of the V5-V6 hypervariable regions of the 16S rRNA gene targeted resulted in a total of 2,450,251 reads. After pre-processing, quality filtering, and subsampling, we reached 3000 effective sequences for each sample, except for two of them that were discarded (Table 4). One hundred and twenty samples and 360,000 effective sequences shared between a total of 2247 OTUs remained. To identify the major factors influencing bacterial diversity and community composition of the dataset, an analysis of similarities (ANOSIM) was performed to test the effect of initial soil, initial seed, habitats, as well as their interactions (initial soil: initial seed, initial soil: habitat, initial seed: habitat, initial soil: initial seed: habitat) (Table 5). The ANOSIM of 14 modalities considering the combination of initial soil (Chenavari or Largentière) and habitats (bulk soils, rhizosphere soils, roots, stems, leaves and both initial and new seeds) provided the best explanatory results (Table 5). In this study, 'habitats' were considered as the biotopes that hosted microbial life (e.g. rhizosphere, stem endosphere or seed endosphere) while

Table 5 ANalysis Of SIMilarities (ANOSIM) is similar to an ANOVA hypothesis test, but it uses a dissimilarity matrix (OTU's table) as input instead of raw data. It shows p value (i.e., significance levels of clustering) and R2 (i.e., the level of dissimilarities between clusters, where 1 is maximum dissimilarity). This ANOSIM was performed on three factors (initial soil, initial seed, and habitat) and their interactions to determine which groups were the most relevant to comparison. There is 14 modalities considered for the following analyses: C-iS (Chenavari initial bulk soil), L-iS (Largentière initial bulk soil), C-iSE (Chenavari initial seed endosphere), L-iSE (Largentière initial seed endosphere), C-RS (Chenavari rhizosphere soil), L-RS (Largentière rhizosphere soil), C-RE (Chenavari root endosphere), L-RE (Largentière root endosphere), C-StE (Chenavari stem endosphere), L-StE (Largentière stem endosphere), C-LE (Chenavari leaf endosphere), L-LE (Largentière leaf endosphere), C-SE (Chenavari seed endosphere), and L-SE (Largentière seed endosphere)

Factors and interactions between factors	R^2	p value
Initial soil	0.32	1.10^{-4}
Initial seed	0.07	0.004
Habitat	0.21	1.10^{-4}
Initial soil: Initial seed	0.25	1.10^{-4}
Initial soil: Habitat	0.55	1.10^{-4}
Initial seed: Habitat	0.16	1.10^{-4}
Initial soil: Initial seed: Habitat	0.54	1.10^{-4}

'modalities' corresponded to habitat x initial soil (e.g. C-RS, L-RS, C-StE or L-StE). This ANOSIM showed that the interaction between initial soil (Largentière and Chenavari) and habitats (bulk soils, rhizosphere soil, root endosphere, stem endosphere, leaf endosphere and seed endosphere) significantly affected the bacterial communities, as shown by the *R*-value of 0.55 ($P < 0.0001$). Meanwhile, variation in bacterial communities due to the initial seed, alone or in interaction with other factors, resulted in low *R*-values, falling between 0.07 and 0.25.

Diversity indices associated with bacterial communities

Coverage (Good's coverage) of the microbial communities from the various modalities ranged from 86.2 to 95.8% (Table 6). Analysis of the alpha diversity of the dataset revealed that there were no significant difference in species richness indices (Observed OTUs, Chaos1 index, and ACE index) and Shannon diversity index when comparing pairs of the same microbial habitats of the two soil types (i.e., C-RS vs L-RS, or C-StE vs L-StE) (Table 6). The only exceptions were for the comparison of observed OTUs in L-iS (272 ± 8), which was significantly lower than C-iS (471 ± 7) and for the comparison of Shannon index of L-iS (2.96 ± 0.04), which was also clearly lower than for C-iS (5.27 ± 0.02). Inverse Simpson index and Shannon evenness followed the same trend, yet additionally, these two diversity indices were higher in the stem endospheres and leaf endospheres of plants developed on Largentière soil (L-StE and L-LE) than in those developed on Chenavari soil (C-StE and C-LE). For both soil types, bacterial diversity indices and species richness were globally lower in the root endosphere (RE) than in other plant-related microbial habitats.

Description of the composition of the microbial communities

At the phylum level, there were 9 bacterial taxa with more than 1% relative abundance present within modalities (Fig. 1). There were 16 identified phyla that had a relative abundance inferior to 1% each (rare phyla on the Fig. 1), that when grouped together, represented 3.01% of the effective sequences. Moreover, 3.56% of sequences were unclassified at the phylum taxonomic level (unclassified phyla on the Fig. 1). In the entire dataset, the bacterial phyla (and the subphyla *Alphaproteobacteria* and *Gammaproteobacteria*) with a mean relative abundance greater than 1% were: *Actinobacteriota* (39.32%), *Chloroflexi* (15.82%), *Alphaproteobacteria* (13.03%), *Patescibacteria* (9.02%), *Bacteroidota* (6.20%), *Gammaproteobacteria* (5.79%), *Gemmatimonadota* (1.48%), *Deinococcota* (1.45%) and *Acidobacteriota* (1.32%). The bacterial community from L-iS was significantly enriched with *Patescibacteria* when compared to all other modalities, with a relative abundance of 63.77% of the sequences dispatched in 23 *Patescibacteria* OTUs (Fig. 1). This illustrated the low diversity in this modality, as previously revealed with alpha-diversity indices. The compositions of the bacterial communities of the initial bulk soils from the both sites (C-iS and L-iS) were very different, while endophyte communities associated to initial seeds recovered from both sites (C-iSE and L-iSE) were very similar (Fig. 1). Indeed, L-iS, compared to C-iS, was significantly enriched with *Patescibacteria* (63.77% vs 0.87%), while C-iS, was significantly enriched with rare phyla (9.53% vs 4.26%), *Actinobacteriota* (25.20% vs 1.90%), *Bacteroidota* (26.07% vs 6.00%), and *Gammaproteobacteria* (14.23% vs 3.60%). In contrast, relative abundances of the phyla that composed L-iSE and C-iSE endophyte bacterial communities showed no significant difference. When comparing pairs of below-ground habitats (RS and RE) from C and L soils, the bacterial diversity, at the phylum level, was very different (C-RS vs L-RS and C-RE vs L-RE).

In contrast, habitats from above-ground showed a very similar bacterial diversity at the phyla taxonomical level (C-StE vs L-StE and C-LE vs L-LE), especially in the seed endosphere (C-SE vs L-SE) (Fig. 1). In the rhizosphere, C-RS, compared to L-RS, was significantly enriched with rare phyla (2.57% vs 1.53%), *Actinobacteriota* (57.87% vs 24.77%) and *Alphaproteobacteria* (18.47% vs 11.67%), while L-RS was significantly enriched with unclassified phyla (6.10% vs 1.87%), *Chloroflexi* (30.10% vs 6.87%), *Deinococcota* (4.40% vs 0.03%), *Gemmatimonadota* (2.07% vs 0.80%), and *Patescibacteria* (8.87% vs 2.33%). The root endosphere C-RE, compared to L-RE, was significantly enriched with rare phyla (2.10% vs 1.13%), *Bacteroidota* (6.43% vs 3.97%), *Deinococcota* (2.07% vs 0.67%), *Alphaproteobacteria* (16.93% vs 10.93%),

Table 6 Observed OTUs and diversity indices. All diversity statistics were calculated using an OTU threshold of ≥97% sequence similarity on randomly sub-sampled data (3000 reads). Mean of the samples that composed the modalities and standard deviation (± SD) values are provided. Richness was calculated using the number of OTUs, Chao1 and ACE estimators. Diversity was estimated from the Shannon–Wiener (H), Inverse Simpson's (1/D), and Shannon Index Evenness (E) indices. Values designated with the same letters were not significantly different (Kruskal–Wallis test. $P < 0.05$). For each modality, (n) is the number of replicates. Modality names are given in Table 5

Habitats (n)	Observed OTUs	Chao1 index	ACE index	Shannon index (H)	Inverse Simpson index (1/D)	Shannon evenness	Coverage (%)
C-iS (3)	471 (±7)bcd	671 (±29)acd	663 (±20)abd	5.27 (±0.02)d	86.24 (±2.81)ef	0.86 (±0.001)c	94,2 (±0.4)de
C-RS (13)	488 (±13)ce	759 (±32)cd	911 (±81)bef	4.91 (±0.04)d	45.39 (±2.13)ade	0.79 (±0.004)c	92.7 (±0.3)cd
C-RE (14)	307 (±27)a	472 (±58)a	570 (±99)a	4.44 (±0.09)bc	36.73 (±3.77)ab	0.78 (±0.007)c	95.8 (±0.6)e
C-StE (14)	485 (±19)c	776 (±40)cd	941 (±61)def	4.84 (±0.15)cd	43.12 (±8.73)ad	0.78 (±0.019)c	92.7 (±0.3)cd
C-LE (14)	459 (±18)c	722 (±40)c	873 (±73)bef	4.79 (±0.12)cd	41.54 (±8.29)abc	0.78 (±0.014)c	93.2 (±0.3)bd
C-SE (8)	545 (±28)cd	901 (±62)de	1124 (±94)e	5.15 (±0.12)d	69.50 (±12.15)cdf	0.82 (±0.013)c	91.4 (±0.4)c
C-iSE (4)	613 (±16)de	698 (±11)bcd	769 (±11)aef	5.00 (±0.03)cd	63.00 (±3.19)bdf	0.66 (±0.001)b	87.2 (±0.5)a
L-iS (3)	272 (±8)a	441 (±23)ab	570 (±25)ab	2.96 (±0.04)a	4.55 (±0.15)a	0.53 (±0.004)a	95.7 (±0.2)e
L-RS (14)	481 (±55)c	745 (±82)ce	851 (±114)bcf	4.88 (±0.29)cd	58.85 (±17.03)bdf	0.79 (±0.033)c	92.9 (±0.8)cd
L-RE (7)	327 (±96)ab	507 (±150)ab	602 (±196)ac	3.93 (±0.54)b	28.87 (±12.82)ab	0.69 (±0.061)b	95.3 (±1.5)e
L-StE (8)	521 (±40)cd	782 (±74)cd	868 (±115)bcef	5.11 (±0.19)d	71.24 (±12.81)df	0.82 (±0.021)c	92.5 (±0.7)cd
L-LE (8)	503 (±42)cd	771 (±71)cd	901 (±114)bcef	5.09 (±0.20)d	71.79 (±13.67)df	0.82 (±0.021)c	92.7 (±0.7)cd
L-SE (6)	562 (±17)cd	937 (±52)d	1087 (±104)ef	5.25 (±0.10)d	80.93 (±11.26)f	0.83 (±0.012)c	91.3 (±0.4)bc
L-iSE (4)	626 (±11)d	619 (±6)bcd	796 (±9)aef	5.01 (±0.02)cd	62.63 (±2.79)bdf	0.66 (±0.001)b	86.2 (±0.3)a

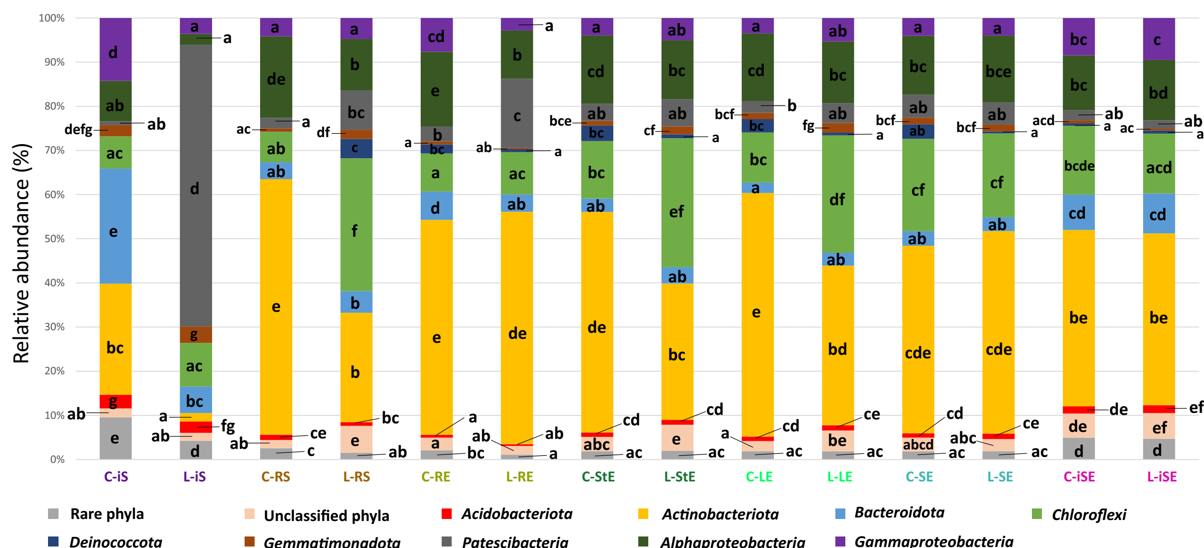

Fig. 1 Mean relative abundances (%) of bacterial phyla from the various modalities. Different letters correspond to significant differences (Kruskal–Wallis test. P < 0.05). All phyla corresponding to less than 1% relative abundance are grouped together as 'Rare Phyla' corresponding to the 16 following phyla: *Abditibacteriota, Armatimonadota, Bdellovibrionota, Desulfobacterota, Elusimicrobiota, Entotheonellaeota,* FCPU426, *Firmicutes, Latescibacterota, Myxococcota, Nitrospirota,* Unclassified *Proteobacteria*, RCP2–54, SAR324_clade(Marine_group_B), *Spirochaetota*, and *Verrucomicrobiota.* Modality names are given in Table 5

and *Gammaproteobacteria* (0.70% vs 0.30%), while L-RE was significantly enriched with *Patescibacteria* (15.63% vs 3.33%). The stem endosphere, C-StE, compared to L-StE, was significantly enriched with *Actinobacteria* (49.93% vs 30.90%) and *Deinococcota* (3.53% vs 0.87%), while L-StE was significantly enriched with unclassified phyla (5.90% vs 3.27%), and *Chloroflexi* (29.20% vs 13.00%). Similarly, the leaf endosphere C-LE, compared to L-LE, was significantly enriched with *Actinobacteria* (55.17% vs 36.20%) and *Deinococcota* (3.03% vs 0.63%), while L-LE was significantly enriched with unclassified phyla (4.70% vs 2.30%), and *Chloroflexi* (26.57% vs 11.33%). Finally, the relative abundance of the phyla that composed L-SE and C-SE endophytic bacterial communities showed no significant difference. Moreover, comparing bacterial endophytic communities of the two seed generations associated with the same soil type (C-iSE vs C-SE and L-iSE vs L-SE) revealed few differences. The initial seeds (iSE), compared to the new seeds (SE) were significantly enriched with rare phyla (4.83% vs 1.87%), unclassified phyla (5.65% vs 2.93%), *Bacteroidota* (8.53% vs 3.22%), and *Gammaproteobacteria* (8.98% vs 4.03%). Overall, the same tendencies were revealed at the taxonomical level of the class and can be found in the supplementary data (Supplementary Fig. 1).

Analyses of similarities and dissimilarities of the bacterial communities

A heatmap (Fig. 2), comparing the relative abundance of phyla in the bacterial communities of the 14 modalities corroborate previous findings. The high relative abundance of *Patescibacteria* in L-iS bacterial community discriminated it from the communities of all other modalities. The higher relative abundance of *Bacteroidota* and *Gammaproteobacteria* in endophytic bacterial communities from initial seeds (C-iSE and L-iSE) set them apart from communities in new seeds (C-SE and L-SE). The relative abundances of *Actinobacteria, Chloroflexi*, and *Alphaproteoacteria* were variable between all modalities and globally, *Actinobacteria* and *Alphaproteoacteria* relative abundances were higher in habitats related to the 'C' site, while *Chloroflexi* relative abundances were higher in habitats related to the 'L' one. Consequently, the hierarchical clustering of the modalities based on the phyla compositions of the endophytic bacterial communities showed a clear dichotomy

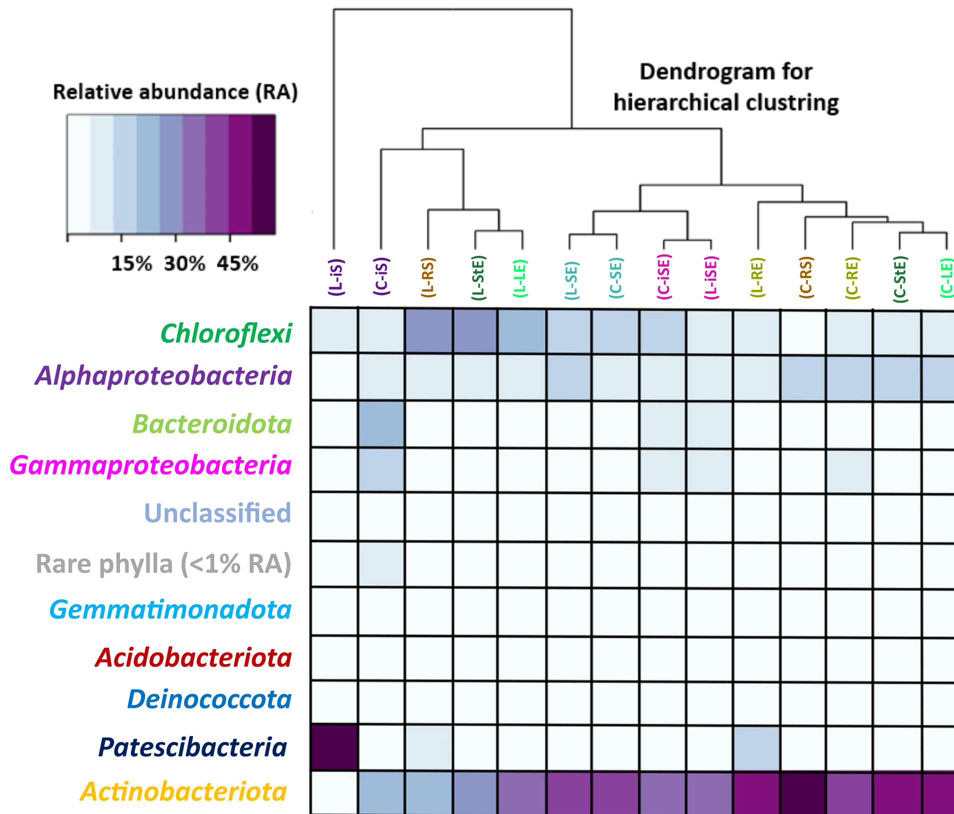

Fig. 2 Heat map and hierarchical cluster analysis of the relative abundance of bacterial phyla (and subphyla for *Proteobacteria*) from the various modalities. The dendrogram represents linkage clustering using Euclidean distance measurements. Modality names are given in Table 5

between L-RS, L-StE and L-LE on the one hand and C-RS, C-StE and C-LE on the other hand (Fig. 2). It also revealed that all seed endosphere bacterial communities (C-iSE, C-SE, L-iSE and L-SE) are clustered together. Nonetheless, the root endosphere bacterial communities from Largentière (L-RE) were more closely clustered with communities associated to Chenavari: C-RE, C-RS, C-StE and C-LE.

The graphical representation NMDS (Non-metric MultiDimensional Scaling) allowed dissimilarities to be observe between all the initial bulk soils, rhizosphere soils and endophytic bacterial communities studied (Fig. 3). The stress value was 0.11, which makes it possible to exploit these results (Clarke and Ainsworth 1993). This visualization revealed a clear clustering of the bacterial diversity in samples depending on soil types and habitats, with most of the samples attached to the Chenavari soil in the lower left half (61/70 circular dots) and most of the samples attached to the Largentière one in the upper right half (36/50 triangle dots) of the bidimensional representation (NMDS1xNMDS2). In fact, the clustering of samples appeared soil-type-dependent, apart from bacterial communities associated with seed endosphere modalities (C-iSE, C-SE, L-iSE and L-SE). Indeed, the clustering was especially strong for bacterial communities from below-ground habitats (iS, RS and RE), with 30/30 samples associated with those habitats attached to the Chenavari soil in the lower left half and 23/24 samples to the Largentière one in the upper right half. Albeit less clear, the soil-type dependent clustering was also observed for bacterial communities of above-ground habitats StE and LE, with 20/24 samples associated to those habitats attached to Chenavari soil in the lower left half and 12/16 samples to Largentière one in the upper right half. Concerning bacterial communities associated to the seed endosphere modalities (C-iSE, L-iSE, C-SE

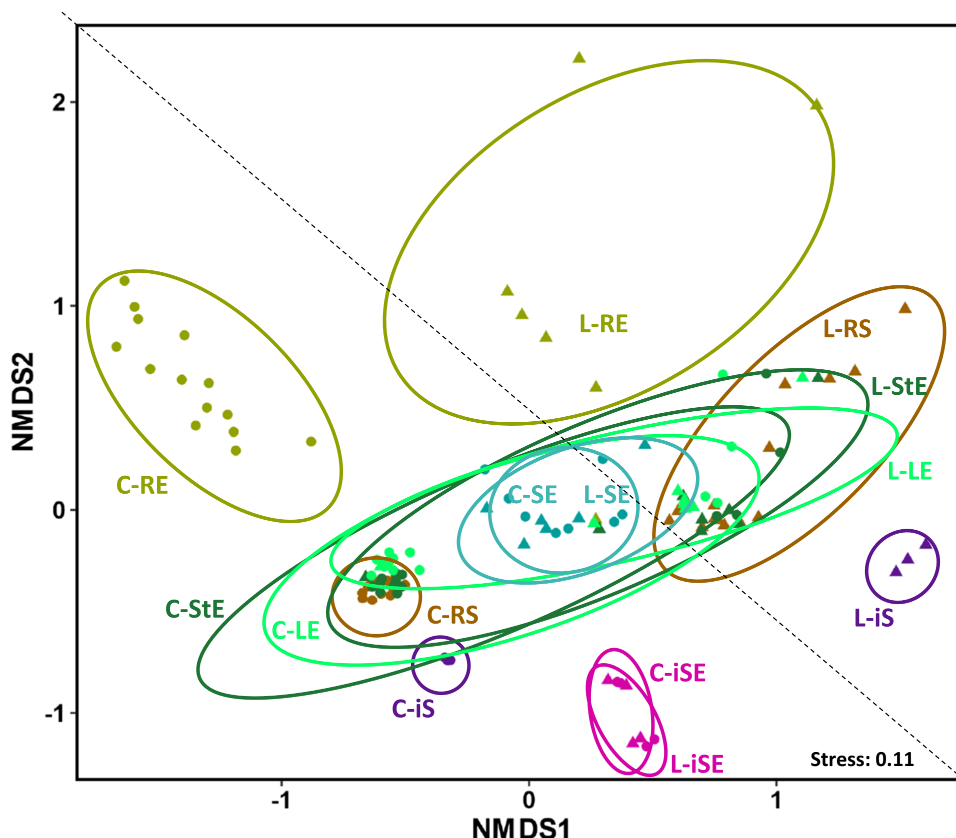

Fig. 3 Non-metric multidimensional scaling (NMDS) plot of bacterial communities. Each point coordinate depends on the diversity the bacterial community described in each sample, for a total of 120 samples (see Table 1). The shape of each point is based on the initial soil on which plants have grown and colors correspond to the type of habitat sampled. The confidence area of the ellipse = 0.95. Stress of the representation in reduced dimensions is shown in the bottom right corner. Modality names are given in Table 5

and L-SE), replicates were very close for each habitat and were positioned on the median part of this plan, especially for C-SE and L-SE (dotted line). Thus, the clustering of samples from seed endosphere habitats appeared less soil-type dependent than all other habitats. However, new seed bacterial endophytic communities (SE) formed a clustered together which was different from a second cluster formed by bacterial endophytic communities from initial seeds (iSE).

To assess the similarity of the bacterial community composition among the four seed endosphere modalities (L-iSE, L-SE, C-iSE and C-SE), a Venn diagram was used (Fig. 4). In these four modalities, a total of 2064 OTUs were identified, with a total of 1686 in C-SE, 1716 in L-iSE, 1343 in C-SE and 1227 in L-SE. New seed endophytic bacterial communities presented fewer OTUs than those of initial seeds. The diagram revealed that 802 OTUs were shared among the four groups, which represented 38.9% (802/2064) of the total OTUs. Endophytic bacterial communities from initial seeds shared a majority of the OTUs that composed them. Indeed, C-iSE and L-iSE bacterial communities shared OTUs that represent respectively 73.9% (993/1343) and 80.9% (993/1227) of the OTUs number. The Venn diagram also showed that in new seeds most of the OTUs that composed the communities were already present within the initial seeds. Indeed, C-SE communities shared 80.9% (1087/1343) of their OTUs with C-iSE communities, while L-SE communities shared 83.6% (1026/1227) of their OTUs with L-iSE communities.

Focusing on the four modalities corresponding to the initial (C-iSE and L-iSE) and new generation seeds (C-SE and L-SE), the obtained network

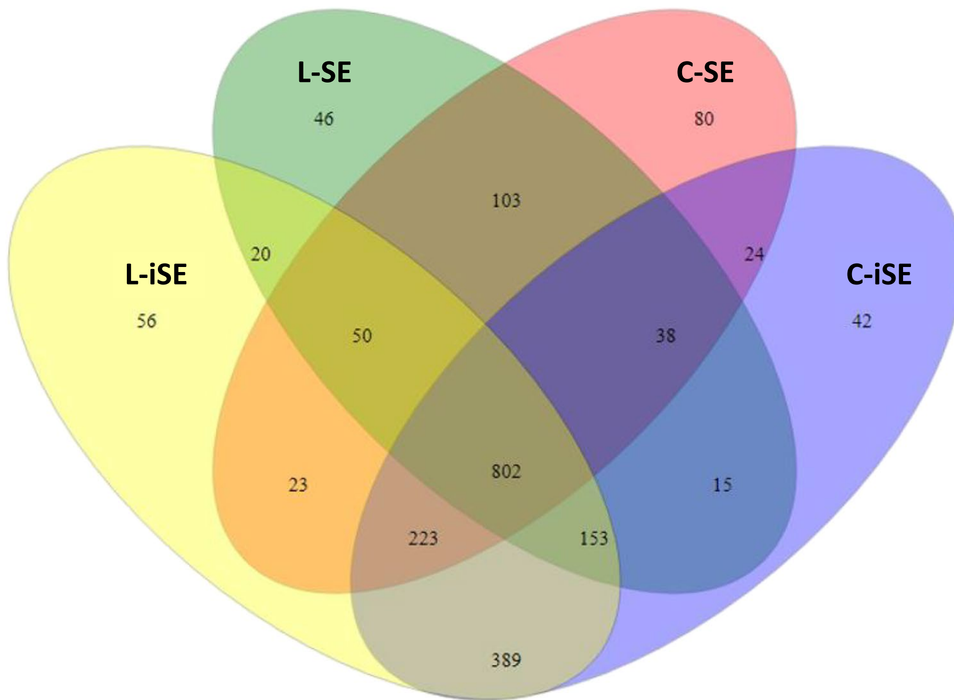

Fig. 4 Venn diagram showing shared and unique OTUs at 97% identity among the four seed endosphere modalities. Each modality OTU composition was the result of the sum of the replicates. Modality names are given in Table 5

of seed endosphere bacterial communities was composed by 28 OTUs, which had a relative abundance greater than 1% (Fig. 5). This network revealed that 4 OTUs were shared between the 4 modalities whatever the soil type (C or L) and the seed generation (initial or new). This core of seed endosphere bacteria contained four OTUs which were OTU0001 *Gaiellales*, OTU0002 *Xanthobacteraceae*, OTU0009 *Streptomyces*, and OTU0013 *Streptomycetaceae*. When we compared C-iSE and L-iSE bacterial communities, only 2 OTUs were specific to C-iSE (OTU0044 *Modestobacter* and OTU0698 *Micrococcaceae*) and none to L-iSE. Similarly, comparing C-SE and L-SE bacterial communities, only 3 OTUs were specific to C-SE (OTU007 *Deinococcaceae*, OTU0024 *Saccharimonadales* and OTU0027 *Roseiflexaceae*) and 2 were specific to L-SE (OTU0018 *Kribbella* and OTU0041 *Methylobacterium-Methylorubrum*). Moreover, 3 OTUs were shared only by L-iSE and C-iSE (OTU36 *Paenathrobacter*, OTU121 *Enterobactriaceae* and OTU258 *Enterobactriaceae*), while 9 OTUs were shared by L-SE and C-SE (OTU0001 *Gaiellales*, OTU0002 *Xanthobacteraceae*, OTU0004 *Bradyrhizobium*, OTU0005 *Chloroflexia*, OTU0006 *Lechevalieria*, OTU0009 *Streptomyces*, OTU0013 *Streptomycetaceae*, OTU0014 *Sphingomonas* and OTU0016 *Streptomyces*).

Discussion

In this study, we used seeds belonging to two *N. caerulescens* populations recovered from a population growing on a non-metalliferous soil ('C') and from a calamine one ('L'). Seeds of each population were sown on their soil of origin but also on the soil of origin of the other plant population. We investigated the structure of bacterial communities associated with different habitats during one successive generation, from initial seeds to the new seed generation. Because the microbiota can benefit plant development and fitness, studying the processes that drive microbiota assemblies from various habitats (in this study: bulk soil, rhizosphere soil, root endosphere, stem endosphere, leaf endosphere and seed endosphere), over space (in this study: Largentière

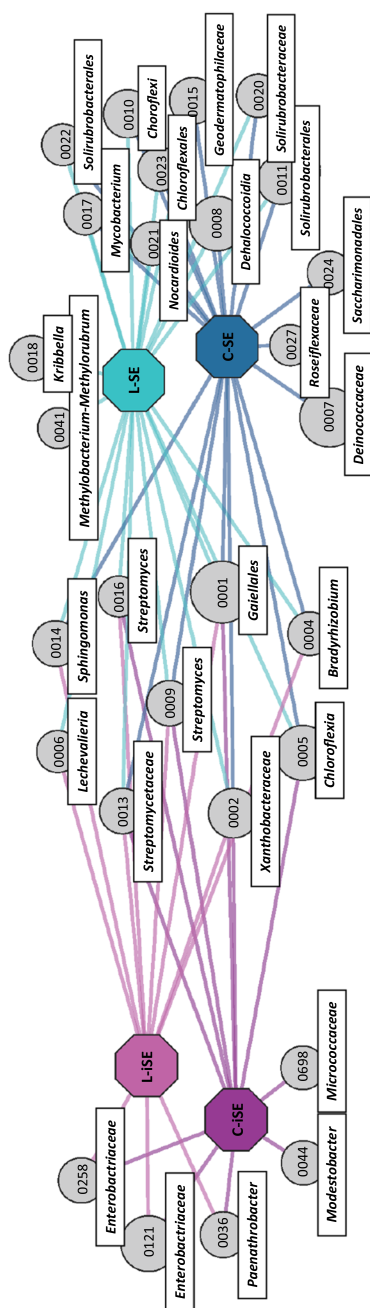

Fig. 5 Bacterial network analysis of the major OTUs (> 1%), categorized by the four modalities (C-iSE, L-iSE, C-SE and L-SE). Octagons represent modalities and circles represent the different OTUs, weighted according to their relative abundance (1.0 to 3.4%). The OTU number is indicated inside the circle and the taxonomic affiliation below. The line color indicates the presence of an OTU with a relative abundance >1% in a modality. Modality names are given in Table 5

and Chenavari soils) and time (in this study: new seed generation) may lead to constructive results (Herrera Paredes and Lebeis 2016). Indeed, deciphering whether selection by environment (e.g. soil physiochemical properties) and selection by host (e.g. plant genotype) influence the associated microbiota composition near to and within plant organs could be crucial questions for a better understanding of the interactions between host plants and microbial symbionts (Wagner et al. 2016; Dombrowski et al. 2017).

Diversity of endophytic bacterial communities associated to hyperaccumulating plants

Researchers have investigated the assembly of bacterial communities on and inside various plant organs, mainly focusing on comparing the rhizosphere soil, the root endosphere, the phyllosphere and leaf endosphere (Bulgarelli et al. 2013; Müller et al. 2016). They have all conclusively demonstrated that the bacterial plant microbiota is composed of only a few dominant phyla, mainly *Proteobacteria* and *Actinobacteriota*, and to a lesser extent *Bacteroidota* and *Firmicutes*. Our analyses partially corroborated these results, since the *Actinobacteriota* phylum was dominant in all endosphere habitats followed by *Proteobacteria* (corresponding here to *Alphaproteobacteria* plus *Gammaproteobacteria*) (Fig. 1). However, in our study, *Bacteroidota*, *Gemmatimonadota*, *Deinococcota* and *Acidobacteriota* had low relative abundances whatever the habitats considered (between 1.3 to 6.2%). *Firmicutes* were even rarer with less than 1% relative abundances in the various habitats. In contrast, the *Chloroflexi* phylum appeared in the various plant organs as a dominant phylum. Indeed, *Chloroflexi* was the third most represented phylum in the whole dataset (15.82%). This phylum is already known to be abundant in extreme and stressful conditions (Yamada et al. 2005) and was also found to be dominant in the *Odontarrhena chalcidica* rhizosphere when present in their natural habitats (ultramafic soil rich in nickel) in the Balkans (Lopez et al. 2017). We can hypothesize that trace elements found both in the soil from the site 'C' and 'L' induced stressful conditions not only in the rhizosphere but also in the plant tissues and finally favored this phylum. Pseudo-total concentrations of Cd, Cu, Pb and Zn were higher in the calamine soil ('L' site) than in the

non-metalliferous one ('C' site) and consequently, the relative abundance of *Chloroflexi* phylum was higher in L-RS compared to C-RS (30.10 vs 6.87%). These results indicated that these microorganisms could tolerate heavy metals well, can adapt to survive and reproduce both in these environments and in hyperaccumulating plant tissues. The abundance of *Chloroflexi* phylum may explain the low abundance of other phyla such as *Bacteroidota*, *Gemmatimonadota*, *Deinococcota*, *Acidobacteriota,* and *Firmicutes* compared to studies of non-hyperaccumulating plants (Bulgarelli et al. 2013; Müller et al. 2016).

Soil dependent response of the bacterial community diversity

The two initial soils have contrasting physiochemical properties with 'L' soil enriched with Pb, Zn, and Cd. It has been shown that Pb contamination implies a significant increase in *Patescibacteria* phylum (Niu et al. 2020). In our study, the concentrations of available-Pb (DTPA) reached 984 and 4.07 $mg.kg^{-1}$, respectively in L and C soils, which can explain the relatively high abundance of this phylum (63.77%) in L initial bulk soil (L-iS) in comparison with C-iS (0.9%). This dominance of a few OTUs from *Patescibacteria* phylum in L-iS explained the significantly lower diversity indices compared to C-iS. In addition, L-RS and L-RE microbial communities were also significantly enriched with *Patescibacteria* phylum compared to C-RS and C-RE, while no significant differences were found for above-ground habitats. Given that *N. caerulescens* does not hyperaccumulate Pb, this may be related to a Pb contamination of below-ground habitats, while above-ground habitats were not. Indeed, it is known that Cd and Zn are generally more easily translocated and accumulated than Pb in plants and particularly in *N. caerulescens* (Mirecki et al. 2015). Furthermore, Herrmann et al. (2019) showed that *Patescibacteria* survived well under oligotrophic conditions and was adapted to nutrient-poor ecosystems, which was the case for the Largentière soil.

When bacterial community pairs in the same microbial habitats but of the two different soil types were compared, the number of abundant phyla (> 1% relative abundance) with significantly different relative abundances, decreased following a gradient from below-ground to above-ground habitats (Fig. 1). In fact, this comparison revealed a contrast between the bacterial communities in the initial soils, as well as for below-ground habitats and to a lesser extent for above-ground habitats, but the soil type had almost no influence on phyla relative abundance of the seed endosphere bacterial communities. Indeed, while 5 of 9 abundant phyla showed a significantly different relative abundance in the root endosphere communities, only 3 of 9 were significantly different in the leaf endosphere or stem endosphere and none were significantly different in the seed endosphere. These dissimilarities were also very visible on the NMDS and followed the same trend. The ability of hyperaccumulating plants to recruit different endophytes in their shoot depending on the soil was previously elucidated for *Sedum alfredii* on Cd/Zn contaminated soils (Qiong et al. 2021). However, our results are the first, to our knowledge to demonstrate a gradient of influence of the soil on endophytic communities, from root to shoot endosphere of hyperaccumulators.

The diversity of seed endophytic bacterial communities across seed generations

In comparison to the phyllosphere (and leaf endosphere) and to the rhizosphere (and root endosphere), our knowledge of microbial communities associated to the other plant habitats, such as seeds, is quite limited. In the past, there have been numerous reports on indigenous endophytic bacteria in various plant tissues. Few of them included seeds (Mundt and Hinkle 1976) and these studies focused on vegetables, woody plants or weeds (Truyens et al. 2015). For example, Rezki et al. (2018) analyzed the structure and diversity of seed-associated microbial communities on radish seed samples collected from individual plants during three successive generations grown on the same experimental site. In their study, Lopez-Velasco et al. (2013) used pyrosequencing of 16 s rRNA gene amplicons to examine the bacterial communities of spinach seeds, while the bacterial community compositions in *Chenopodium album* and *Stellaria media* seeds were characterized by van Overbeek et al. (2011). However, little information is available concerning the structure and diversity of endophytic bacterial communities of hyperaccumulating plants. For example, Mastretta et al. (2009) isolated endophytic bacteria from *Nicotiana tabacum* seeds, which is considered to be a Cd and Zn accumulator. Their results

highlighted the fact that most of the isolated cultivable bacteria belonged to the pseudomonad family. In the same way, a cultivation-independent approach depicted the shoot endophytic bacterial populations of *Thlaspi goesingense* growing in a serpentine soil in eastern Austria (Idris et al. 2004). Endophytes were mainly α-*Proteobacteria* and high-G+C gram-positive bacteria. Most isolates (42%) showed close similarity to *Methylobacterium mesophilicum*. In addition, *Sphingomonas* sp. strains were highly abundant (37%). In their study, all remaining endophytes showed high homology to the genera *Rhodococcus*, *Curtobacterium*, and *Plantibacter*. Durand et al. (2021b) compared the bacterial structure and diversity of seed endophytic bacterial (SEB) communities of the hyperaccumulator pseudometallophyte *N. caerulescens* through the 16S rRNA profiling of seeds belonging to 14 European populations of *N. caerulescens* recovered from their native ecosystems across various edaphic types.

In our study, for each type of site ('C' and 'L'), if we compared the phyla present in initial and new seed generations, except for rare and unclassified phyla, overall, the same phyla were shared between these two habitats (SE and iSE). Indeed, the relative abundances of the different phyla in initial seeds from C and L (C-iSE and L-iSE), but also for the new seed generations from 'C' and 'L' (C-SE and L-SE) were similar and no statistical difference could be noted between the relative abundances of all phyla. Thus, the soil properties of the sites where the seeds were recovered and of the soils used in this study as substrates for plant growth did not result in seed endophytic bacterial communities changes in *N. caerulescens*. These results corroborate previous studies that have shown the low impact of soil property variations on the endophytic bacterial communities found in hyperaccumulator seeds (Durand et al. 2021b). Johnston-Monje et al. (2014) also showed the slight influences of a change of soil on the diversity of the endophyte community of very young plants (20 days after germination) of pre-domesticated, ancient, and modern maize. Nonetheless, in our experiment the initial seeds (iSE), compared to the new seeds (SE) were significantly enriched with rare phyla, unclassified phyla, *Bacteroidota*, and *Gammaproteobacteria*. The new generation of seeds, no matter the soil, hosted a seed endophytic bacterial community with a smaller species richness that the seeds recovered from their natural environment. Indeed, C-SE bacterial community hosted a total of 1343 OTUs, while C-iSE were composed of 1686 OTUs and similarly, the L-SE bacterial community was composed of 1227 OTUs, while 1716 OTUs constituted the L-iSE bacterial community. Moreover, the remaining OTUs, especially the most abundant, in SE bacterial communities were mostly (around 80%) initially present in iSE bacterial communities. We could conclude that the new seeds' endophytic communities derived from a reduction of the seeds' endophyte communities from the initial seeds communities, although this decrease in the number of OTUs did not result in a decrease in the Shannon index. This evolution may be related to the experiment conducted in controlled conditions that was not designed to be a simulation of the fluctuation of weather conditions from the natural sites. This highlights the fact that experiments conducted in controlled conditions cannot completely simulate real environmental conditions. Indeed, the influence of annual fluctuations in local weather conditions on bacterial communities associated with radish (*Raphanus sativus* var. Flamboyant5) seeds was previously revealed in a study where geographic distance was not involved (Rezki et al. 2018). In their study, fluctuations in microbial community profiles were observed relating to changes in community membership and composition across plant (and seed) generations. This observation was made across years of culture and were correlated to weather fluctuations. Note that these seeds were not surface-sterilized.

In our study, we observed that development of the host plant appeared to be the major driver of rhizosphere bacterial community assembly, regardless soil characteristics. Indeed, while the bacterial diversity differed between the two bulk soils considered (C-iS and L-iS), bacterial communities seemed more similar in rhizosphere soil (C-RS and L-RS). A second selection was related to the soil type which follows a gradient from below-ground habitats where its influence was strong, to above-ground habitats where the influence of this driver decreased and seemed to have almost no influence on seed endosphere microbial communities. The hierarchical cluster analysis of the bacterial community similarity (Fig. 2) showed that samples associated to seed endosphere habitats (L-iSE, C-iSE, L-SE and C-SE) were clearly separated from others, indicating that soil type ('C' and 'L') did not influence the bacterial diversity in

seeds despite differences in soil physicochemical parameters.

The resulting seed endosphere bacterial network (Fig. 5) consisted of 28 nodes corresponding to abundant OTUs with relative abundance greater than >1%. The network analysis corroborated previous findings of this study; the soil had a low effect on the composition of the seed endosphere bacterial communities. Indeed, there were few abundant OTUs (> 1%) specific to C-iSE compared to L-iSE and vice versa, and few abundant specific OTUs in C-SE bacterial community compared to L-SE and vice versa. In addition, this network revealed the most common OTUs that composed a core of seed endosphere bacteria present whatever the soil type and the seed generation (OTU0001 *Gaiellales*, OTU0002 *Xanthobacteraceae*, OTU0009 *Streptomyces*, and OTU0013 *Streptomycetaceae*). The OTU0001 *Gaiellales* (mean relative abundance in the seed endosphere 3.43%) belongs to a clad that was found in seeds of sugar beet and that was colonizing rhizosphere of *Rhizoctonia*-tolerant cultivars (Wolfgang et al. 2020). In another study, an enrichment in *Gaiellales* was observed in the potato periderm of the potato cultivar Kariera, which was resistant to common scab (Kopecky et al. 2019). In both cases, *Gaiellales* were considered as potential disease-suppressive agents. Previous studies have shown that *Streptomyces* bacteria were specifically recruited to the endosphere of the model plant *Arabidopsis thaliana* and that some could confer growth-promoting benefits on *A. thaliana* including production of indole-3-acetic acid (IAA), siderophores, and aminocyclopropane-1-carboxylate (ACC) deaminase while others might be exploited to protect crops against disease (Worsley et al. 2020). Moreover, *Streptomycetaceae* (e.g. *Streptomycetes*) have been identified as the common metabolically active bacteria in metal-contaminated soils (Gremion et al. 2003), which could also produce many metabolic products affecting host physiology and growth (e.g. phytohormones, bacterial siderophores, and ACC deaminase) (Nimnoi et al. 2010). In their study, Luo et al. (2019) showed that of the vertically transmitted endophytes, *Streptomycetaceae* was one of the most abundant taxa, and that a majority of members (particularly the *Streptomyces* genus) are known to secrete substantial amounts of antimicrobial compounds which provide favorable environments for their producers by inhibiting pathogens or other taxa competing for the habitat space. These authors also underlined that, members of the *Streptomycetaceae* family were able to affect both growth and Cd/Zn accumulation of *Sedum alfredii* stronger than other enriched bacterial taxa. In agreement with previous reports (Luo et al. 2019), the seed endosphere bacterial community we observed was also largely dominated by members of *Xanthobacteraceae* family. This taxa is known to contain a range of metal resistance and tolerance mechanisms and previously, it was revealed that *Xanthobacteraceae* bacteria are among the marker families in the root microbiome of *Juncus acutus* growing in soil enriched with Cd, Ni, Zn (Syranidou et al. 2018).

Among the nine OTUs only shared by L-SE and C-SE and not found in the initial seed generations (L-iSE and C-iSE), OTUs belonging to *Chloroflexi* phylum and *Mycobacterium* and *Nocardioides* genera were found with a relative abundance higher than 1%. In our study, the phyla to which these OTUs belong (*Chloroflexi* and *Actinobacteriota*) were present in the initial bulk (iS) and rhizosphere (RS) soils whatever the site (C and L), but considering OTUs with a relative abundance greater than 1%, we only found OTUs belonging to these phyla in the new generation seeds. This result supported the concept that soil rhizosphere bacteria are the prime source of endophytic bacteria, as mentioned by Seghers et al. (2004) and Peng et al. (2015). Indeed, previous studies have shown that endophytic bacteria are mainly derived from soil bacteria because the latter can pass into the plants through root fractures (McCully 2001), natural openings or wounds (Kluepfel 1993) and seeds (Lodewyckx et al. 2002a). Moreover, among these OTUs only present in new seed generation, endophytic *Actinobacteriota* are known for their potential to affect plant growth and nutrient uptake (Rajkumar et al. 2006). They are reported as plant growth promoters in various plants (Nimaichand et al. 2016), but were also reported to release metal binding secondary metabolites that mobilize Zn and Cd from the soil, enhancing their accumulation in the plants (Dimkpa et al. 2009; Kuffner et al. 2010).

Conclusion

The seed-associated microbial communities are ecologically interesting because they represent both an

endpoint and a starting point for community assembly of the plant microbiota (Shade et al. 2017). Moreover seed-associated microorganisms contribute to seed preservation (Chee-Sanford et al. 2006). Hyperaccumulators generally evolve in stress-full conditions and we can hypothesize that if a seed endophytic bacterial core-genome appeared to be constant between seed generations, that these bacterial communities play an important role in the fitness of these particular plants.

In contrast with other plants (Lopez-Velasco et al. 2013; Barret et al. 2015; Klaedtke et al. 2016; Rezki et al. 2018), this study underlines that some OTUs were consistently present in the consecutive generations of *N. caerulescens* seeds growing on metal-rich soils and constituted a core members of the seed endophyte bacteria. Nonetheless, there were variations between the composition of the initial and new seed endophytic communities. Certain OTUs were recruited from the soil and were only present in the new generation seeds. All these OTUs were known to produce many metabolic products affecting host physiology and growth and to favor metal accumulation in plants. Therefore, these bacteria could provide net benefits to the hosts and were thus maintained in the consecutive seed generations and/or recruited from soil. Thus, it is conceivable that bacteria trans-generationally preserved can contribute to the establishment and further growth of new generations of plants (Sánchez-López et al. 2018b; Walitang et al. 2018). Metatranscriptomic approaches can help to elucidate evidence about the possible roles of the members of the core seed endophyte bacterial community.

References

Asaf S, Khan MA, Khan AL, Waqas M, Shahzad R, Kim AY, Kang SM, Lee IJ (2017) Bacterial endophytes from arid land plants regulate endogenous hormone content and promote growth in crop plants: an example of *Sphingomonas* sp. and *Serratia marcescens*. J Plant Interactions 12:31–38. https://doi.org/10.1080/17429145.2016.1274060

Assunção AGL, Schat H, Aarts MGM (2003) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New Phytol 159:351–360. https://doi.org/10.1046/j.1469-8137.2003.00820.xAna

Badri DV, Zolla G, Bakker MG, Manter DK, Vivanco JM (2013) Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. New Phytol 198:264–273. https://doi.org/10.1111/nph.12124

Baker AJM, Brooks RR (1989) Terrestrial higher plants which hyperaccumulate metallic elements: a review of their distribution, ecology and phytochemistry. Biorecovery 1:81–126

Barret M, Briand M, Bonneau S, Préveaux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A (2015) Emergence shapes the structure of the seed microbiota. Appl Environ Microbiol 81:1257–1266. https://doi.org/10.1128/AEM.03722-14

Benizri E, Lopez S, Durand A, Kidd PS (2021) Diversity and role of endophytic and rhizosphere microbes associated with hyperaccumulator plants during metal accumulation. In: Agromining: Farming for metals. Springer, Cham, pp 239–279

Berg G, Kusstatscher P, Abdelfattah A, Cernava T, Smalla K (2021) Microbiome modulation: toward a better understanding of plant microbiome response to microbial inoculants. Front Microbiol 12:1–12. https://doi.org/10.3389/fmicb.2021.650610

Berg G, Rybakova D, Fischer D, Cernava T, Vergès MCC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH, Kazou M, Kinkel L, Lange L, Lima N, Loy A, Macklin JA, Maguin E, Mauchline T, McClure R et al (2020) Microbiome definition re-visited: old concepts and new challenges. Microbiome 8:1–22. https://doi.org/10.1186/s40168-020-00875-0

Bosch TCG, McFall-Ngai MJ (2011) Metaorganisms as the new frontier. Zoology 114:185–190. https://doi.org/10.1016/j.zool.2011.04.001

Bulgarelli D, Schlaeppi K, Spaepen S, Van Themaat EVL, Schulze-Lefert P (2013) Structure and functions of the bacterial microbiota of plants. Annual Rev Plant Biol 64:807–838. https://doi.org/10.1146/annurev-arplant-050312-120106

Busby PE, Ridout M, Newcombe G (2016) Fungal endophytes: modifiers of plant disease. Plant Mol Biol 90:645–655. https://doi.org/10.1007/s11103-015-0412-0

Caporaso J, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Pena A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303

Chao A (1949) On the estimation of the number of classes in a population. Annals Mathematical Statistics 20:572–579. https://doi.org/10.1214/aoms/1177729949

Chee-Sanford JC, Williams MM, Davis AS, Sims GK (2006) Do microorganisms influence seed-bank dynamics? Weed Sci 54:575–587. https://doi.org/10.1614/WS-05-055R.1

Clarke KR, Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. Marine Ecol Progress Series 92:205–219. https://doi.org/10.3354/meps092205

Defez R, Andreozzi A, Bianco C (2017) The overproduction of indole-3-acetic acid (IAA) in endophytes upregulates nitrogen fixation in both bacterial cultures and inoculated rice plants. Microbial Ecol 74:441–452. https://doi.org/10.1007/s00248-017-0948-4

Dimkpa CO, Merten D, Svatoš A, Büchel G, Kothe E (2009) Metal-induced oxidative stress impacting plant growth in contaminated soil is alleviated by microbial siderophores. Soil Biol Biochem 41:154–162. https://doi.org/10.1016/j.soilbio.2008.10.010

Dombrowski N, Schlaeppi K, Agler MT, Hacquard S, Kemen E, Garrido-Oter R, Wunder J, Coupland G, Schulze-Lefert P (2017) Root microbiota dynamics of perennial *Arabis alpina* are dependent on soil residence time but independent of flowering time. ISME J 11:43–55. https://doi.org/10.1038/ismej.2016.109

Durand A, Leglize P, Benizri E (2021a) Are endophytes essential partners for plants and what are the prospects for metal phytoremediation? Plant Soil 460:1–30. https://doi.org/10.1007/s11104-020-04820-w

Durand A, Sterckeman T, Gonnelli C, Coppi A, Bacci G, Leglize P, Benizri E (2021b) A core seed endophytic bacterial community in the hyperaccumulator *Noccaea caerulescens* across 14 sites in France. Plant Soil 459:203–216. https://doi.org/10.1007/s11104-020-04743-6

Fesel PH, Zuccaro A (2016) Dissecting endophytic lifestyle along the parasitism/mutualism continuum in *Arabidopsis*. Current Opinion Microbiol 32:103–112. https://doi.org/10.1016/j.mib.2016.05.008

Frank A, Saldierna Guzmán J, Shay J (2017) Transmission of bacterial endophytes. Microorganisms 5:70. https://doi.org/10.3390/microorganisms5040070

Glick B, Penrose D, Li J (1997) A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. J Theoretic Biol 190:63–68. https://doi.org/10.1006/jtbi.1997.0532

Gonneau C, Noret N, Godé C, Frérot H, Sirguey C, Sterckeman T, Pauwels M (2017) Demographic history of the trace metal hyperaccumulator *Noccaea caerulescens* (J. Presl and C. Presl) F. K. Mey. In Western Europe. Mol Ecol 26:904–922. https://doi.org/10.1111/mec.13942

Good IJ (1953) The population frequencies of species and the estimation of population parameters. Biometrika 40:237–264. https://doi.org/10.1093/biomet/40.3-4.237

Gremion F, Chatzinotas A, Harms H (2003) Comparative 16S rDNA and 16S rRNA sequence analysis indicates that *Actinobacteria* might be a dominant part of the metabolically active bacteria in heavy metal-contaminated bulk and rhizosphere soil. Environ Microbiol 5:896–907. https://doi.org/10.1046/j.1462-2920.2003.00484.x

Hamayun M, Hussain A, Khan SA, Kim HY, Khan AL, Waqas M, Irshad M, Iqbal A, Rehman G, Jan S, Lee IJ (2017) Gibberellins producing endophytic fungus *Porostereum spadiceum* AGH786 rescues growth of salt affected soybean. Front Microbiol 8:686. https://doi.org/10.3389/fmicb.2017.00686

Hardoim PR (2011) Bacterial endophytes of rice: Their diversity, characteristics and perspectives. University of Groningen

Healey A, Furtado A, Cooper T, Henry RJ (2014) Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant Methods 10:21. https://doi.org/10.1186/1746-4811-10-21

Herrera Paredes S, Lebeis SL (2016) Giving back to the community: microbial mechanisms of plant–soil interactions. Funct Ecol 30:1043–1052. https://doi.org/10.1111/1365-2435.12684

Herrmann M, Wegner CE, Taubert M, Geesink P, Lehmann K, Yan L, Lehmann R, Totsche KU, Küsel K (2019) Predominance of Cand. *Patescibacteria* in groundwater is caused by their preferential mobilization. Front Microbiol 10:1–15. https://doi.org/10.3389/fmicb.2019.01407

Idris R, Trifonova R, Puschenreiter M, Wenzel WW, Sessitsch A (2004) Bacterial communities associated with flowering plants of the Ni hyperaccumulator *Thlaspi goesingense*. Appl Environ Microbiol 70:2667–2677. https://doi.org/10.1128/AEM.70.5.2667-2677.2004

Johnston-Monje D, Lundberg DS, Lazarovits G, Reis VM, Raizada MN (2016) Bacterial populations in juvenile maize rhizospheres originate from both seed and soil. Plant Soil 405:337–355. https://doi.org/10.1007/s11104-016-2826-0

Johnston-Monje D, Mousa WK, Lazarovits G, Raizada MN (2014) Impact of swapping soils on the endophytic bacterial communities of pre-domesticated, ancient and modern maize. BMC Plant Biol 14:233. https://doi.org/10.1186/s12870-014-0233-3

Kandel S, Joubert P, Doty S (2017) Bacterial endophyte colonization and distribution within plants. Microorganisms 5:77. https://doi.org/10.3390/microorganisms5040077

Kembel SW, O'Connor TK, Arnold HK, Hubbell SP, Wright SJ, Green JL (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. Proc Natl Acad Sci U S A 111:13715–13720. https://doi.org/10.1073/pnas.1216057111

Klaedtke S, Jacques MA, Raggi L, Préveaux A, Bonneau S, Negri V, Chable V, Barret M (2016) Terroir is a key driver of seed-associated microbial assemblages. Environ Microbiol 18:1792–1804. https://doi.org/10.1111/1462-2920.12977

Kluepfel DA (1993) The behavior and tracking of bacteria in the rhizosphere. Annu Rev Phytopathol 31:441–472. https://doi.org/10.1146/ANNUREV.PY.31.090193.002301

Kopecky J, Samkova Z, Sarikhani E, Kyselková M, Omelka M, Kristufek V, Divis J, Grundmann GG, Moënne-Loccoz Y, Sagova-Mareckova M (2019) Bacterial, archaeal and micro-eukaryotic communities characterize a disease-suppressive or conducive soil and a cultivar resistant or susceptible to common scab. Sci Rep 9:1–14. https://doi.org/10.1038/s41598-019-51570-6

Krämer U (2010) Metal hyperaccumulation in plants. Annual Rev Plant Biol 61:517–534. https://doi.org/10.1146/annurev-arplant-042809-112156

Kuffner M, De Maria S, Puschenreiter M, Fallmann K, Wieshammer G, Gorfer M, Strauss J, Rivelli AR, Sessitsch A (2010) Culturable bacteria from Zn- and cd-accumulating *Salix caprea* with differential effects on plant growth and heavy metal availability. J Appl Microbiol 108:1471–1484. https://doi.org/10.1111/j.1365-2672.2010.04670.x

Lodewyckx C, Mergeay M, Vangronsveld J, Clijsters H, Van Der Lelie D (2002a) Isolation, characterization, and identification of bacteria associated with the zinc hyperaccumulator *Thlaspi caerulescens* subsp. calaminaria. Int

J Phytoremediation 4:101–115. https://doi.org/10.1080/15226510208500076

Lodewyckx C, Vangronsveld J, Porteous F, Moore ERB, Taghavi S, Mezgeay M, Van der Lelie D (2002b) Endophytic bacteria and their potential applications. Critical Rev Plant Sci 21:583–606. https://doi.org/10.1080/0735-260291044377

Lòpez-Fernàndez S, Mazzoni V, Pedrazzoli F, Pertot I, Campisano A (2017) A phloem-feeding insect transfers bacterial endophytic communities between grapevine plants. Front Microbiol 8:1–17. https://doi.org/10.3389/fmicb.2017.00834

Lopez-Velasco G, Carder PA, Welbaum GE, Ponder MA (2013) Diversity of the spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. FEMS Microbiol Lett 346:146–154. https://doi.org/10.1111/1574-6968.12216

Lopez S, Piutti S, Vallance J, Morel JL, Echevarria G, Benizri E (2017) Nickel drives bacterial community diversity in the rhizosphere of the hyperaccumulator *Alyssum murale*. Soil Biol Biochem 114:121–130. https://doi.org/10.1016/j.soilbio.2017.07.010

Luo J, Tao Q, Jupa R, Liu Y, Wu K, Song Y, Li J, Huang Y, Zou L, Liang Y, Li T (2019) Role of vertical transmission of shoot endophytes in root-associated microbiome assembly and heavy metal hyperaccumulation in *Sedum alfredii*. Environ Sci Technol 53:6954–6963. https://doi.org/10.1021/acs.est.9b01093

Luo SL, Chen L, Liang CJ, Xiao X, Ying XT, Wan Y, Rao C, Bin LC, Tang LY, Lai C, Ming ZG (2011) Analysis and characterization of cultivable heavy metal-resistant bacterial endophytes isolated from cd-hyperaccumulator *Solanum nigrum* L. and their potential use for phytoremediation. Chemosphere 85:1130–1138. https://doi.org/10.1016/j.chemosphere.2011.07.053

Ma Y, Oliveira RS, Nai F, Rajkumar M, Luo Y, Rocha I, Freitas H (2015a) The hyperaccumulator *Sedum plumbizincicola* harbors metal-resistant endophytic bacteria that improve its phytoextraction capacity in multi-metal contaminated soil. J Environ Manag 156:62–69. https://doi.org/10.1016/j.jenvman.2015.03.024

Ma Y, Rajkumar M, Moreno A, Zhang C, Freitas H (2017) Serpentine endophytic bacterium *Pseudomonas azotoformans* ASS1 accelerates phytoremediation of soil metals under drought stress. Chemosphere 185:75–85. https://doi.org/10.1016/j.chemosphere.2017.06.135

Ma Y, Rajkumar M, Rocha I, Oliveira RS, Freitas H (2015b) Serpentine bacteria influence metal translocation and bioconcentration of *Brassica juncea* and *Ricinus communis*. Front Plant Sci 5:757. https://doi.org/10.3389/fpls.2014.00757

Mastretta C, Taghavi S, Van Der Lelie D, Mengoni A, Galardi F, Gonnelli C, Barac T, Boulet J, Weyens N, Vangronsveld J (2009) Endophytic bacteria from seeds of *Nicotiana tabacum* can reduce cadmium phytotoxicity. Int J Phytoremediation 11:251–267. https://doi.org/10.1080/15226510802432678

McCully M (2001) Niches for bacterial endophytes in crop plants: a plant biologist's view. Aust J Plant Physiol 28:983–990

McEnroe NA, Helmisaari HS (2001) Decomposition of coniferous forest litter along a heavy metal pollution gradient, south-West Finland. Environ Pollution 113:11–18. https://doi.org/10.1016/S0269-7491(00)00163-9

Mirecki N, Agič R, Šunić L, Milenković L, Ilić ZS (2015) Transfer factor as indicator of heavy metals content in plants. Fresenius Environ Bull 24:4212–4219

Müller DB, Vogel C, Bai Y, Vorholt JA (2016) The plant microbiota: systems-level insights and perspectives. Annual Rev Genetics 50:211–234. https://doi.org/10.1146/annurev-genet-120215-034952

Mundt OJ, Hinkle NF (1976) Bacteria within ovules and seeds. Appl Environ Microbiol 32:694–698. https://doi.org/10.1128/aem.32.5.694-698.1976

Nelson EB (2017) The seed microbiome: origins, interactions, and impacts. Plant Soil 422:7–34. https://doi.org/10.1007/s11104-017-3289-7

Nelson EB, Simoneau P, Barret M, Mitter B, Compant S (2018) Editorial special issue: the soil, the seed, the microbes and the plant. Plant Soil 422:1–5. https://doi.org/10.1007/s11104-018-3576-y

Nimaichand S, Devi AM, Li W-J (2016) Direct plant growth-promoting ability of Actinobacteria in grain legumes. In: Plant growth promoting *Actinobacteria*: a new avenue for enhancing the productivity and soil fertility of grain legumes. Springer, Singapore, pp 1–16

Nimnoi P, Pongsilp N, Lumyong S (2010) Endophytic actinomycetes isolated from *Aquilaria crassna* Pierre ex Lec and screening of plant growth promoters production. World J Microbiol Biotechnol 26:193–203. https://doi.org/10.1007/s11274-009-0159-3

Niu X, Zhou J, Wang X, Su X, Du S, Zhu Y, Yang J, Huang D (2020) Indigenous bacteria have high potential for promoting *Salix integra* thunb. Remediation of lead-contaminated soil by adjusting soil properties. Front Microbiol 11:1–13. https://doi.org/10.3389/fmicb.2020.00924

Panke-Buisse K, Poole AC, Goodrich JK, Ley RE, Kao-Kniffin J (2015) Selection on soil microbiomes reveals reproducible impacts on plant function. ISME J 9:980–989. https://doi.org/10.1038/ismej.2014.196

Peer WA, Mamoudian M, Lahner B, Reeves RD, Murphy AS, Salt DE (2003) Identifying model metal hyperaccumulating plants: germplasm analysis of 20 *Brassicaceae* accessions from a wide geographical area. New Phytol 159:421–430. https://doi.org/10.1046/j.1469-8137.2003.00822.x

Peng A, Liu J, Ling W, Chen Z, Gao Y (2015) Diversity and distribution of 16S rRNA and phenol monooxygenase genes in the rhizosphere and endophytic bacteria isolated from PAH-contaminated sites. Sci Rep 5:1–12. https://doi.org/10.1038/srep12173

Perotti R (1926) On the limits of biological enquiry in soil science. Proc Int Soc Soil Sci 2:146–161

Preite V, Sailer C, Syllwasschy L, Bray S, Ahmadi H, Krämer U, Yant L (2019) Convergent evolution in *Arabidopsis halleri* and *Arabidopsis arenosa* on calamine metalliferous soils. Philos Trans Royal Soc B: Biol Sci 374:20180243. https://doi.org/10.1098/rstb.2018.0243

Qiong W, Fengshan P, Xiaomeng X, Rafiq MT, Xiao'e Y, Bao C, Ying F (2021) Cadmium level and soil type played a

selective role in the endophytic bacterial community of hyperaccumulator *Sedum alfredii* Hance. Chemosphere 263:127986. https://doi.org/10.1016/j.chemosphere.2020.127986
Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:590–596. https://doi.org/10.1093/nar/gks1219
R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing
Rajkumar M, Nagendran R, Lee KJ, Lee WH, Kim SZ (2006) Influence of plant growth promoting bacteria and Cr6+ on the growth of Indian mustard. Chemosphere 62:741–748. https://doi.org/10.1016/j.chemosphere.2005.04.117
Reeves RD, Baker AJM, Jaffré T, Erskine PD, Echevarria G, van der Ent A (2018) A global database for plants that hyperaccumulate metal and metalloid trace elements. New Phytol 218:407–411. https://doi.org/10.1111/nph.14907
Reeves RD, Schwartz C, Morel JL, Edmondson J (2001) Distribution and metal-accumulating behavior of *Thlaspi caerulescens* and associated metallophytes in France. Int J Phytoremediation 3:145–172. https://doi.org/10.1080/15226510108500054
Reeves RD, van der Ent A, Baker AJM (2021) Global distribution and ecology of hyperaccumulator plants. In: Agromining: farming for metals. Springer, Cham, pp 133–154
Rezki S, Campion C, Simoneau P, Jacques MA, Shade A, Barret M (2018) Assembly of seed-associated microbial communities within and across successive plant generations. Plant Soil 422:67–79. https://doi.org/10.1007/s11104-017-3451-2
Ritpitakphong U, Falquet L, Vimoltust A, Berger A, Metraux JP, L'Haridon F (2016) The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. New Phytol 210:1033–1043. https://doi.org/10.1111/nph.13808
Rolli E, Marasco R, Vigani G, Ettoumi B, Mapelli F, Deangelis ML, Gandolfi C, Casati E, Previtali F, Gerbino R, Pierotti Cei F, Borin S, Sorlini C, Zocchi G, Daffonchio D (2015) Improved plant resistance to drought is promoted by the root-associated microbiome as a water stress-dependent trait. Environ Microbiol 17:316–331. https://doi.org/10.1111/1462-2920.12439
Sánchez-López AS, Pintelon I, Stevens V, Imperato V, Timmermans J-P, González-Chávez MDCA, Carrillo-González R, Van Hamme J, Vangronsveld J, Thijs S (2018a) Seed endophyte microbiome of *Crotalaria pumila* unpeeled: identification of plant-beneficial methylobacteria. Int J Mol Sci 19:291. https://doi.org/10.3390/IJMS19010291
Sánchez-López AS, Thijs S, Beckers B, Gonzalez-Chavez MC, Weyens N, Carrillo-Gonzalez R, Vangronsveld J (2018b) Community structure and diversity of endophytic bacteria in seeds of three consecutive generations of *Crotalaria pumila* growing on metal mine residues. Plant Soil 422:51–66. https://doi.org/10.1007/s11104-017-3176-2
Santoyo G, Moreno-Hagelsieb G, del Carmen O-MM, Glick BR (2016) Plant growth-promoting bacterial endophytes. Microbiol Res 183:92–99. https://doi.org/10.1016/j.micres.2015.11.008
Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541. https://doi.org/10.1128/AEM.01541-09
Seghers D, Wittebolle L, Top EM, Verstraete W, Siciliano SD (2004) Impact of agricultural practices on the *Zea mays* L. endophytic community. Appl Environ Microbiol 70:1475–1482. https://doi.org/10.1128/AEM.70.3.1475-1482.2004
Shade A, Jacques M-A, Barret M (2017) Ecological patterns of seed microbiome diversity, transmission, and assembly. Current Opinion Microbiol 37:15–22. https://doi.org/10.1016/j.mib.2017.03.010
Shannon P, Markiel A, Ozier O, Baliga N, Wang J, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303.metabolite
Smith EF (1911) Bacteria in relation to plant diseases. Carnegie Institute, Washington
Syranidou E, Thijs S, Avramidou M, Weyens N, Venieri D, Pintelon I, Vangronsveld J, Kalogerakis N (2018) Responses of the endophytic bacterial communities of *Juncus acutus* to pollution with metals, emerging organic pollutants and to bioaugmentation with indigenous strains. Front Plant Sci 871:1–14. https://doi.org/10.3389/fpls.2018.01526
Tannenbaum I, Kaur J, Mann R, Sawbridge T, Rodoni B, Spangenberg G (2020) Profiling the *Lolium perenne* microbiome: from seed to seed. Phytobiomes J 4:281–289. https://doi.org/10.1094/PBIOMES-03-20-0026-R
Truyens S, Weyens N, Cuypers A, Vangronsveld J (2015) Bacterial seed endophytes: genera, vertical transmission and interaction with plants. Environ Microbiol Reports 7:40–50. https://doi.org/10.1111/1758-2229.12181
van Overbeek LS, Franke AC, Nijhuis EHM, Groeneveld RMW, da Rocha UN, Lotz LAP (2011) Bacterial communities associated with *Chenopodium album* and *Stellaria media* seeds from arable soils. Microbial Ecol 62:257–264. https://doi.org/10.1007/s00248-011-9845-4
Visioli G, D'Egidio S, Vamerali T, Mattarozzi M, Sanangelantoni AM (2014) Culturable endophytic bacteria enhance Ni translocation in the hyperaccumulator *Noccaea caerulescens*. Chemosphere 117:538–544. https://doi.org/10.1016/j.chemosphere.2014.09.014
Wagner MR, Lundberg DS, Del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T (2016) Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat Commun 7:1–15. https://doi.org/10.1038/ncomms12151
Walitang DI, Sunyoung CK, Yeongyeong J, Sa T (2018) Conservation and transmission of seed bacterial endophytes across generations following crossbreeding and repeated inbreeding of rice at different geographic locations. Microbiol Open 8:e006. https://doi.org/10.1002/mbo3.662

Wang Q, Ge C, Xu S, Wu Y, Sahito ZA, Ma L, Pan F, Zhou Q, Huang L, Feng Y, Yang X (2020) The endophytic bacterium *Sphingomonas* SaMR12 alleviates cd stress in oilseed rape through regulation of the GSH-AsA cycle and antioxidative enzymes. BMC Plant Biol 20:1–14. https://doi.org/10.1186/s12870-020-2273-1

Wani ZA, Ashraf N, Mohiuddin T, Riyaz-Ul-Hassan S (2015) Plant-endophyte symbiosis, an ecological perspective. Appl Microbiol Biotechnol 99:2955–2965. https://doi.org/10.1007/s00253-015-6487-3

Weyens N, Van Der Lelie D, Taghavi S, Vangronsveld J (2009) Phytoremediation: plant– endophyte partnerships take the challenge. Current Opinion Biotechnol 20:248–254. https://doi.org/10.1016/j.copbio.2009.02.012

Wolfgang A, Zachow C, Müller H, Grand A, Temme N, Tilcher R, Berg G (2020) Understanding the impact of cultivar, seed origin, and substrate on bacterial diversity of the sugar beet rhizosphere and suppression of soil-borne pathogens. Front Plant Sci 11:1–15. https://doi.org/10.3389/fpls.2020.560869

Worsley SF, Newitt J, Rassbach J, Batey SFD, Holmes NA, Murrell JC, Wilkinson B, Hutchings MI (2020) Streptomyces endophytes promote host health and enhance growth across plant species. Appl Environ Microbiol 86:1–17. https://doi.org/10.1128/AEM.01053-20

Yadav A (2017) Exploring the potential of endophytes in agriculture: a minireview. Advancesn Plants Agric Res 6:00221. https://doi.org/10.15406/apar.2017.06.00221

Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H (2005) Diversity, localization, and physiological properties of filamentous microbes belonging to *Chloroflexi* subphylum I in mesophilic and thermophilic methanogenic sludge granules. 71:7493–7503. https://doi.org/10.1128/AEM.71.11.7493

Zhang YF, Yan HL, Jin CZ, Ya WQ, Qian M, Fang SX (2011) Characterization of ACC deaminase-producing endophytic bacteria isolated from copper-tolerant plants and their potential in promoting the growth and copper accumulation of *Brassica napus*. Chemosphere 83:57–62. https://doi.org/10.1016/j.chemosphere.2011.01.041

Zhou T, Li L, Zhang X, Zheng J, Zheng J, Joseph S, Pan G (2016) Changes in organic carbon and nitrogen in soil with metal pollution by cd, cu, Pb and Zn: a meta-analysis. Eur J Soil Sci 67:237–246. https://doi.org/10.1111/ejss.12327

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.