A footprint of plant eco-geographic adaptation on the composition of the barley rhizosphere bacterial microbiota.

Rodrigo Alegria Terrazas[1], Katharin Balbirnie-Cumming[1], Jenny Morris[2], Pete E Hedley[2], Joanne Russell[2], Eric Paterson[3], Elizabeth M Baggs[4], Eyal Fridman[5] and Davide Bulgarelli[1*]

[1]University of Dundee, Plant Sciences, School of Life Sciences, Dundee, United Kingdom;
[2]Cell and Molecular Sciences, The James Hutton Institute, Dundee, United Kingdom;
[3]Ecological Sciences, The James Hutton Institute, Aberdeen, United Kingdom;
[4]Global Academy of Agriculture and Food Security, University of Edinburgh, Royal (Dick) School of Veterinary Studies, Midlothian, United Kingdom.
[5]Institute of Plant Sciences, Agricultural Research Organization (ARO), The Volcani Center, Bet Dagan, Israel.

***Correspondence:**
Davide Bulgarelli
d.bulgarelli@dundee.ac.uk

Abstract

The microbiota thriving in the rhizosphere, the thin layer of soil surrounding plant roots, plays a critical role in plant's adaptation to the environment. Domestication and breeding selection have progressively differentiated the microbiota of modern crops from the ones of their wild ancestors. However, the impact of eco-geographical constraints faced by domesticated plants and crop wild relatives on recruitment and maintenance of the rhizosphere microbiota remains to be fully elucidated. Here we performed a comparative 16S rRNA gene survey of the rhizosphere of 4 domesticated and 20 wild barley (*Hordeum vulgare*) genotypes grown in an agricultural soil under controlled environmental conditions. We demonstrated the enrichment of individual bacteria mirrored the distinct eco-geographical constraints faced by their host plants. Unexpectedly, Elite varieties exerted a stronger genotype effect on the rhizosphere microbiota when compared with wild barley genotypes adapted to desert environments with a preferential enrichment for members of Actinobacteria. Finally, in wild barley genotypes, we discovered a limited, but significant, correlation between microbiota diversity and host genomic diversity. Our results revealed a footprint of the host's adaptation to the environment on the assembly of the bacteria thriving at the root-soil interface. In the tested conditions, this recruitment cue layered atop of the distinct evolutionary trajectories of wild and domesticated plants and, at least in part, is encoded by the barley genome. This knowledge will be critical to design experimental approaches aimed at elucidating the recruitment cues of the barley microbiota across a range of soil types.

Introduction

By 2050 the world's population is expected to reach 9.5 billion and, to ensure global food security, crop production has to increase by 60% in the same timeframe[1]. This target represents an unprecedented challenge for agriculture as it has to be achieved while progressively decoupling yields from non-renewable inputs in the environment[2] and amidst climatic modifications which are expected to intensify yield-limiting events, such as water scarcity and drought[3].

A promising strategy proposes to achieve this task by capitalising on the microbiota inhabiting the rhizosphere, the thin layer of soil surrounding plant roots[4]. The rhizosphere microbiota plays a crucial role in plant's adaptation to the environment by facilitating, for example, plant mineral uptake[5] and enhancing plant's tolerance to both abiotic and biotic stresses[6].

Plant domestication and breeding selection, which have progressively differentiated modern cultivated crops from their wild relatives[7], have impacted on the composition and functions of the rhizosphere microbiota[8]. These processes were accompanied by an erosion of the host genetic diversity[9] and there are growing concerns that, in turn, these limited the metabolic diversity of the microbiota of cultivated plants[10]. Thus, to fully unlock the potential of rhizosphere microbes for sustainable crop production, it is necessary to study the microbiota thriving at the root-soil interface in the light of the evolutionary trajectories of its host plants[11].

Barley (*Hordeum vulgare* L.), a global crop[12] and a genetically tractable organism[13], represents an ideal model to study host-microbiota interactions within a plant domestication framework, due to the fact that wild relatives (*H. vulgare* ssp. *spontaneum*) of domesticated varieties (*H. vulgare* ssp. *vulgare*) are accessible for experimentation[14]. We previously demonstrated that domesticated and wild barley genotypes host contrasting bacterial communities[15] whose metabolic potential modulates the turn-over of the organic matter in the rhizosphere[16]. However, the impact of eco-geographical constraints faced by domesticated plants and crop wild relatives on recruitment and maintenance of the rhizosphere microbiota remains to be fully elucidated. Tackling this knowledge gap is a key pre-requisite to capitalise on plant-microbiota interactions to achieve the objectives of climate-smart agriculture, in particular sustainably enhancing crop production[17].

Here we investigated whether exposure to different environmental conditions during evolution left a footprint on the barley's capacity of shaping the rhizosphere bacterial

microbiota. We characterised twenty wild barley genotypes from the 'B1K' collection sampled in the Southern levant geographic region, one of the centres of domestication of barley[18,19]. This material represents the three-major barley 'Ecotypes' adapted to different habitats in the region[20]: the Golan Heights and northern Galilee, ('North Ecotype'); the coastal Mediterranean strip, ('Coast Ecotype'); and the arid regions along the river Jordan and southern Negev ('Desert Ecotype'). We further subdivided these 'Ecotypes' into 5 groups of sampling locations according to the average rainfall of the areas, as a proxy for plant's adaptation to limiting growth conditions: 'Coast 1', 'Coast 2', 'Desert 1' and 'Desert 2' and 'North', respectively. (Table 1; Figure 1). These wild barley genotypes were grown in a previously characterised soil, representative of barley agricultural areas of Scotland, under controlled environmental conditions, alongside four cultivated 'Elite' varieties encompassing the main usage and genetic diversity of the cultivated germplasm (Table 1). We used an Illumina MiSeq 16S rRNA gene amplicon survey to characterise the microbiota inhabiting the rhizosphere and unplanted soil samples. By using ecological indexes, multivariate statistical analyses and barley genome information we elucidated the impact of eco-geographical constraints and host genetics on the composition of the microbial communities thriving at the barley root-soil interface.

Results

Evolutionary trajectories and eco-geographic adaptation impact on plant growth

Aboveground dry weight from the barley genotypes was measured at early stem elongation as a proxy for plant growth: this allowed us to identify a 'biomass gradient' across the tested material. The 'Elite' varieties, outperforming wild barley plants, and wild barley genotypes adapted to the more extreme desert environments (i.e., 'Desert 2') defined the uppermost and lowermost ranks of this gradient, respectively ($P < 0.05$, Kruskal–Wallis non-parametric analysis of variance followed by Dunn's post hoc test; Figure 1). Conversely, when we inspected the ratio between above- and belowground biomass we noticed an opposite trend: almost invariably wild barley genotypes allocated more resources than 'Elite' varieties to root growth compared to stem growth ($P < 0.05$, Kruskal–Wallis non-parametric analysis of variance followed by Dunn's post hoc test; Figure 1). As we sampled plants at a comparable developmental stage (Methods; Figure S1), these observations indicate different growth responses of in wild and domesticated genotypes in the tested conditions.

Taxonomic diversification of the barley microbiota across barley genotypes

To study the impact of these differential responses on the composition of the barley microbiota we generated 6,646,864 16S rRNA gene sequencing reads from 76

rhizosphere and unplanted soil specimens. These high-quality sequencing reads yielded 11,212 Operational Taxonomic Units (OTUs) at 97% identity (Supplementary Dataset 1: worksheet 2). A closer inspection of the taxonomic affiliation of the retrieved OTUs revealed that members of five bacterial phyla, namely *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Proteobacteria* and *Verrucomicrobia*, accounted for more than 97.8% of the observed reads (Figure 2, Supplementary Dataset 1: worksheet 3). Among these dominant phyla, *Bacteroidetes* and *Proteobacteria* were significantly enriched in rhizosphere compared to bulk soil profiles (ANCOM, cut-off 0.6, alpha 0.05, taxa-based corrected, Supplementary Dataset 1: worksheets 4-5).

Next, we investigated the lower ranks of the taxonomic assignments (i.e., OTU level) and computed the Observed OTUs, Chao1 and Shannon indexes for each sample type. This analysis further supported the notion of the rhizosphere as a 'reduced complexity community', as both the Observed OTUs and Shannon indexes, but not the projected Chao1, identified significantly richer and more even communities in the bulk soil samples compared to plant-associated specimens ($P < 0.05$, Mann–Whitney U test; Figure S3). Interestingly, when we compared the Chao1 index within rhizosphere samples, we observed that members of the 'Desert 1' group assembled a richer and more even community compared with the other genotypes ($P < 0.05$, Kruskal–Wallis non-parametric analysis of variance followed by Dunn's post hoc test; Figure S3).

To gain further insights into the impact of the sample type on the barley microbiota we generated a canonical analysis of principal coordinates (CAP) using the weighted Unifrac distance, which is sensitive to OTU relative abundance and phylogenetic relatedness. This analysis revealed a marked effect of the microhabitat, i.e., either bulk soil or rhizosphere, on the composition of the microbiota as evidenced by the spatial separation on the axis accounting for the major variation (Figure 3). Interestingly, we observed a clustering of bacterial community composition within rhizosphere samples, which was more marked between 'Desert' and 'Elite' samples (Figure 3). These observations were corroborated by a permutational analysis of variance which attributed ~30% of the observed variation to the microhabitat and, within rhizosphere samples, ~17% of the variation to the individual eco-geographic groups (Permanova P <0.01, 5,000 permutations, Table 2). Strikingly similar results were obtained when we computed a Bray-Curtis dissimilarity matrix, which is sensitive to OTUs relative abundance only (Table 2; Figure S4).

Taken together, these data indicate that the composition of the barley microbiota is fine-tuned by plant recruitment cues which progressively differentiate between unplanted soil and rhizosphere samples and, within these latter, wild ecotypes from elite varieties.

A footprint of host eco-geographic adaptation shapes the wild barley rhizosphere microbiota

To gain insights into the bacteria underpinning the observed microbiota diversification we performed a series of pair-wise comparisons between 'Elite' genotypes and each group of the wild barley ecotypes. This approach revealed a marked specialisation of the members of the 'Desert' ecotype compared to 'Elite' varieties as evidenced by the number of OTUs differentially recruited between members of these groups (Wald test, P <0.05, FDR corrected; Figure 4; Supplementary Dataset 1: worksheets 7-11). Thus, the wild barley 'Ecotype' emerged as an element shaping the recruitment cues of the barley rhizosphere microbiota.

A closer inspection of the OTUs differentially recruited between 'Desert' wild barley and 'Elite' varieties revealed that the domesticated material exerted the greatest selective impact on the soil biota, as the majority of the differentially enriched OTUs were enriched in 'Elite' varieties (Wald test, P <0.05, FDR corrected; Supplementary Dataset 1: worksheets 7 and 8). Next, the taxonomic assignments of these 'Elite-enriched' OTUs versus the 'Desert' microbiota followed distinct patterns: while the comparison 'Elite'-'Desert 1' produced a subset of enriched OTUs assigned predominantly to *Actinobacteria*, *Bacteroidetes* and *Proteobacteria*, the comparison 'Elite'-'Desert 2' displayed a marked bias for members of the *Actinobacteria* (i.e., 44 out of 104 enriched OTUs, Figure 5). Consistently, the cumulative abundance of sequencing reads assigned of those Actinobacterial OTUs in 'Elite' samples nearly doubled the one recorded for 'Desert 2' samples (Figure S5). Within this phylum, we identified a broader taxonomic distribution, as those OTUs were assigned to the families *Intrasporangiaceae*, *Micrococcaceae*, *Micromonosporaceae*, *Nocardioidaceae*, *Pseudonocardiaceae*, *Streptomycetaceae*, as well as members of the order *Frankiales*. Interestingly, when we inspect intra-ecotype diversification we identified diagnostic OTUs capable of discriminating between 'Desert 1' and 'Desert 2' (Wald test, P <0.05, FDR corrected; Supplementary Dataset 1: worksheets 12 and 13), while no such a feature was identified discriminating between 'Coast 1' and 'Coast 2' at the statistical test imposed. Taken together, our data indicate that wild barley 'Ecotype' (i.e., the differential effect of 'North', 'Coast, and 'Desert' versus 'Elite') acts as a determinant for the rhizosphere barley microbiota whose composition is ultimately fine-

tuned by a sub-specialisation within the 'Ecotype' itself (i.e., the differential effect of 'Desert 1' and 'Desert 2').

These observations prompted us to investigate whether the differential microbiota recruitment between the tested plants was encoded, at least in part, by the barley genome. We therefore generated a dissimilarity matrix using Single Nucleotide Polymorphisms (SNPs) available for the tested genotypes and we inferred their genetic relatedness using a simple matching coefficient (Supplementary Dataset 1: worksheet 14). With few notable exceptions, this analysis revealed three distinct clusters of genetically related plants, represented by and reflecting the 'Elite' material, the 'Desert' and the 'Coast' wild barley genotypes (Figure S6). The genetic diversity between domesticated material exceeded their microbial diversity (compare relatedness of "Elite" samples in Figure 3 with the ones of Figure S6) as further evidenced by the fact that we failed to identify a significant correlation between these parameters (*P* value > 0.05). However, when we focused the analysis solely on the pool of wild barley genotypes, we obtained a significant correlation between genetic and microbial distances (Mantel test r = 0.230; *P* value < 0.05; Figure 6).

Taken together, this revealed a footprint of barley host's adaptation to the environment on the assembly of the bacteria thriving at the root-soil interface. This recruitment cue interjected the distinct evolutionary trajectories of wild and domesticated plants and, at least in part, is encoded by the barley genome.

Discussion

In this study we investigated how plant genotypes adapted to different eco-geographic niches may recruit a distinct microbiota once exposed to a common environment.

As we performed a 'common environment experiment' in a Scottish agricultural soil, we first determined how the chosen experimental conditions related to the ones witnessed by wild barleys in their natural habitats. Strikingly, the aboveground biomass gradient observed in our study, with 'Elite' material almost invariably outperforming wild genotypes and material sampled at the locations designated 'Desert 2' at the bottom of the ranking, "matched" the phenotypic characterisation of members of the 'B1K' collection grown in a 'common garden experiment' in a local Israeli soil[18]. Conversely, belowground resource allocation followed an opposite pattern as evidenced by an increased root:shoot dry weight ratio in wild genotypes compared to 'Elite' varieties. As responses to edaphic stress, such as drought tolerance, may modulate the magnitude of above-belowground

resource partitioning in plants[21] and root traits[22], our data might reflect the adaptation of the wild barley exposure to dry areas. Taken together, these results suggest adaptive that responses to eco-geographic constraints in barley have a genetic inheritance component which can be detected and studied in controlled conditions.

As genetically-inherited root traits have been implicated in shaping the rhizosphere microbiota in barley[23] and other crops[24], these observations motivated us to examine whether these below-ground differences were reflected by changes in microbiota recruitment. The distribution of reads assigned to given phyla appears distinct in plant-associated communities which are dominated in terms of abundance by members of the phyla *Acidobacteria, Actinobacteria, Bacteroidetes and Proteobacteria*, with these two latter phyla significantly enriched in rhizosphere samples compared to bulk soil controls. This taxonomic affiliation is consistent with previous investigations in barley in either the same[25] or in a different soil type[15] as well as in other crop plants[26]. In summary, these data indicate that the higher taxonomic ranks of the barley rhizosphere microbiota are conserved across soil types as well as wild and domesticated genotypes.

The characterisation of the microbiota at lower taxonomic ranks, i.e., the OTU-level, revealed a significant effect of the microhabitat (i.e., either bulk soil or rhizosphere) and, within plant-associated communities, a footprint of eco-geographic adaptation. For instance, alpha diversity indexes clearly pointed at selective processes modulating bacterial composition as the number of Observed OTUs and the Shannon index indicate simplified and reduced-complexity communities inhabiting the rhizosphere compared to unplanted soil. This can be considered a hallmark of the rhizosphere microbiota as it has been observed in multiple plant species and across soils[6]. Conversely, within rhizosphere samples, alpha-diversity analysis failed to identify a clear pattern, except for the Chao1 index revealing a potential for a richer community associated with plants sampled at the 'Desert 1' locations. This motivated us to further explore the between-sample diversity, which is beta-diversity. This analysis revealed a clear host-dependent diversification of the bacteria associated to barley plants manifested by ~17% of the variance of the rhizosphere microbiota explained by the eco-geographical location of the sampled material. This value exceeded the host genotype effect on the rhizosphere microbiota we previously observed in wild and domesticated barley plants[15], but is aligned with the magnitude of host effect observed in the rhizosphere microbiota of modern and ancestral genotypes of rice[27] and common bean[28]. As these studies were conducted in different soil

types, our data suggest that the magnitude of host control on the rhizosphere microbiota is ultimately fine-tuned by and in response to soil characteristics.

The identification of the bacteria underpinning the observed microbiota diversification led to three striking observations. First, the comparison between 'Elite' varieties and the material representing the 'Desert' ecotype was associated with the largest number of differentially recruited OTUs, while the other wild barley genotypes appeared to share a significant proportion of their microbiota with domesticated plants. A prediction of this observation is that the distinct evolutionary trajectories of wild and domesticated plants *per se* cannot explain the host-mediated diversification of the barley microbiota. As aridity and temperature played a prominent role in shaping the phenotypic characteristics of barley[19,29] it is tempting to speculate that the adaptation to these environmental parameters played a predominant role also in shaping microbiota recruitment.

Second, it is the domesticated material which exerted a stronger effect on microbiota recruitment, manifested by the increased number of host-enriched OTUs compared to wild barley genotypes. This suggests that the capacity of shaping the rhizosphere microbiota has not been "lost" during barley domestication and breeding selection. Our findings are consistent with data gathered for domesticated and ancestral common bean genotypes, which revealed that shifts from native soils to agricultural lands led to a stronger host-dependent effect on rhizosphere microbes[30]. Due to the intrinsic limitation of 16S rRNA gene profiles of predicting the functional potential of individual bacteria, it will be necessary to complement this investigation with whole-genome surveys[31,32] and metabolic analyses[16,33] to fully discern the impact of the host genotype on the functions provided by the rhizosphere microbiota to their hosts.

The third observation is the marked quantitative enrichment of OTUs assigned to the phylum *Actinobacteria* in 'Elite' varieties when compared to members of the 'Desert' ecotype, in particular plants of the 'Desert 2' locations. At first glance, the 'direction' of this bacterial enrichment is difficult to reconcile with the eco-geographic adaptation of wild barleys and, in particular, the fact that *Actinobacteria* are more tolerant to arid conditions[34] and, consequently, more abundant in desert vs. non-desert soils[35]. However, the enrichment of *Actinobacteria* in modern crops compared to ancestral relatives has recently emerged as a distinctive feature of the microbiota of multiple plant species[36]. Although the ecological significance of this trait of the domesticated microbiota remains to be fully elucidated, studies conducted in rice[37] and other grasses, including barley[38], indicate a

277 relationship between drought stress and *Actinobacteria* enrichments. These observations
278 suggest that the wild barley genome has evolved the capacity to recognise microbes
279 specifically adapted to the local conditions and, in turn, to repress the growth of others. For
280 instance, among the bacteria differentially enriched between 'Desert 1' and 'Desert 2' we
281 identified genera, such as *Arthrobacter sp.*, adapted to extreme environments and long-
282 term nutrient starvation[39], possibly reflecting the differential adaptation of 'Desert 1' and
283 'Desert 2' plants to soil with limited organic matter[40].

284 Interestingly, we were able to trace the host genotype effect on rhizosphere
285 microbes to the genome of wild barley. This suggests that, similar to other wild species[11],
286 microbiota recruitment co-evolved with other adaptive traits. Conversely, the genetic
287 diversity in 'Elite' material largely exceeded microbiota diversity. This is reminiscent of
288 studies conducted in maize which failed to identify a significant correlation between
289 polymorphisms in the host genome and alpha- and beta-diversity characteristics of the
290 rhizosphere microbiota[41,42]. Yet, and again similar to maize[43], our data indicate that the
291 recruitment of individual bacterial OTUs in the 'Elite' varieties, rather than community
292 composition as a whole, is the feature of the rhizosphere microbiota under host genetic
293 control.

294 Although these findings were gathered from the individual soil tested and further
295 validation across a range of soil types is required, a prediction from these observations is
296 that the host control of the rhizosphere microbiota is exerted by a limited number of loci in
297 the genome with a relatively large effect. This is congruent with our previous observation
298 that mono-mendelian mutations in a single root trait, root hairs, impact on ~18% of the
299 barley rhizosphere microbiota[25].

300 Likewise, this scenario is compatible with a limited number of genes controlling the
301 biosynthesis and rhizodeposition of defensive secondary metabolites which have been
302 implicated in shaping the plant microbiota[44]. Among these compounds, the indol-alkaloids
303 benzoxazinoids recently gained centre-stage as master regulators of the maize-associated
304 microbial communities[45-47]. Interestingly, *Hordeum vulgare* has evolved a distinct indol-
305 alkaloid compound, gramine[48], which is preferentially accumulated in the tissues of the
306 wild genotypes compared to 'Elite' varieites[49] and whose physiological properties are
307 comparable to the ones of benzoxazinoids[50]. Whether gramine or other species-specific
308 secondary metabolites contribute, at least in part, to shape the barley microbiota will be
309 the focus of future investigations.

Since modern varieties have been selected with limited or no knowledge of belowground interactions, how was the capacity of shaping the rhizosphere microbiota retained within the cultivated germplasm? The recent observation that genes controlling reproductive traits display pleiotropic effects on root system architecture[51] could provide a direct link between crop selection and microbiota recruitment in modern varieties. These traits, and in particular genes encoding flower developments, show a marked footprint of eco-geographic adaptation and have been selected during plant domestication and breeding[29]. By manipulating those genes, breeders may have manipulated also belowground traits, and in turn, the microbiota thriving at the root-soil interface. With an increased availability of genetic[52] and genomic[53] resources for wild and domesticated barleys, this hypothesis can now be experimentally tested and the adaptive significance of the barley rhizosphere microbiota ultimately deciphered. Specifically, interspecific populations within the wild[54] as well as between wild and cultivated[52] germplasm, could be deployed in genetic mapping experiments aimed at identifying barley genetic determinants of the rhizosphere microbiota.

Conclusions

Our results revealed a footprint of host's adaptation to the environment on the assembly of the bacteria thriving at the root-soil interface in barley. This recruitment cue layered atop of the distinct evolutionary trajectories of wild and domesticated plants and, at least in part, is encoded by the barley genome. Although our study was limited to the individual soil investigated, our sequencing survey will provide a reference dataset for the development of indexed bacterial collections of the barley microbiota. These can be used to infer causal relationships between microbiota composition and plant traits, as demonstrated for *Arabidopsis thaliana*[55] and rice[56]. Furthermore, this knowledge is critical for the establishment of reciprocal transplantation experiments aimed at elucidating the adaptive value of crop-microbiota interactions, similar to what has recently been proposed for the model plant *A. thaliana*[57]. However, for crop plants like barley, this will necessarily be conditioned by two elements: identifying the host genetic determinants of the rhizosphere microbiota and inferring microbial metabolic potential *in situ*. Ultimately, this will help devising strategies aimed at sustainably enhancing crop production for climate-smart agriculture.

342 # Methods

343 ## Soil

344 The soil was sampled from the agricultural research fields of the James Hutton
345 Institute, Invergowrie, Scotland, UK in the Quarryfield site (56°27'5"N 3°4'29"W; Sandy Silt
346 Loam, pH 6.2; Organic Matter 5%; Table S1). This field was left unplanted and unfertilised
347 in the three years preceding the investigations and previously used for barley-microbiota
348 interactions investigations[25].

349

350 ## Plant genotypes

351 Twenty wild barley genotypes (*Hordeum vulgare* ssp. *spontaneum*) and four 'Elite'
352 cultivars (*Hordeum vulgare* ssp. *vulgare*) were used and described in Table 1. Wild barley
353 genotypes were selected representing eco-geographical variation of the 'B1K'
354 collection[18,19]. The 'Elite' genotypes were selected as a representation of different types of
355 spring barley in plant genetic studies. The cultivar 'Morex' is an American six-row malting
356 variety whose genome was the first to be sequenced[58]. The cultivars 'Bowman' and 'Barke'
357 are two-row varieties, developed in US for feed and in Germany for malting, respectively,
358 whereas Steptoe is an American six-row type used for animal feed[52,59,60].

359 ## Plant growth conditions

360 Barley seeds were surface sterilized as previously reported[61] and germinated on
361 0.5% agar plates at room temperature. Seedlings displaying comparable rootlet
362 development after 5 days post-plating were sown individually in 12-cm diameter pots
363 containing approximately 500g of the 'Quarryfield' soil, plus unplanted pots filled with bulk
364 soil as controls. Plants were arranged in a randomised design with this number of
365 replicates: 'Coast1' number of replicates n=12; 'Coast2' n=12; 'Desert1' n=11; 'Desert2'
366 n=12; 'North' n=12; 'Elite' n=13 (Supplementary Dataset 1: worksheet 1). Plants were
367 grown for 5 weeks in a glasshouse at 18/14 °C (day/night) temperature regime with 16 h
368 day length and watered every two days with 50 ml of deionized water.

369 ## Bulk soil and rhizosphere DNA preparation

370 At early stem elongation, corresponding to Zadoks stages 30-32 [62], plants were
371 pulled from the soil and the stems and leaves were separated from the roots (Figure S1).
372 Above-ground plant parts were dried at 70 °C for 72 h and the dry weight recorded. The
373 roots were shaken manually to remove excess of loosely attached soil. For each barley
374 plant, the top 6 cm of the seminal root system and the attached soil layer was collected
375 and placed in sterile 50 ml falcon tube containing 15 ml phosphate-buffered saline solution

(PBS). Rhizosphere was operationally defined, for these experiments, as the soil attached to this part of the roots and extracted through this procedure. The samples were then vortexed for 30s and aseptically transferred to a second 50ml falcon containing 15ml PBS and vortexed again for 30s to ensure the dislodging and suspension of the rhizosphere soil. Then, the two falcon tubes with the rhizosphere suspension were mixed and centrifuged at 1,500 x *g* for 20min, the supernatant was removed, with the rhizosphere soil collected as the pellet, flash frozen with liquid nitrogen and stored at -80°C, until further use. After the rhizosphere extraction step, these parts of the roots were combined with the rest of the root system for each plant, thoroughly washed with water removing any attached soil particles and dried at 70°C for 72h for root biomass measurement. Bulk soil samples were collected from the 6cm below the surface of unplanted pots and subjected to the same procedure as above.

DNA was extracted from the rhizosphere samples using FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, USA) according to the manufacturer's recommendations. The concentration and quality of DNA was checked using a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, USA) spectrophotometer and stored at -20°C until further use. DNA concentration was used as a proxy for the proportion of the sampled microbiota and evaluated across sample type (Figure S2).

Preparation of 16 rRNA gene amplicon pools

The hypervariable V4 region of the small subunit rRNA gene was the target of amplification using the PCR primer pair 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3'). The PCR primers had incorporated an Illumina flow cell adapter at their 5' end and the reverse primers contained 12bp unique 'barcode' for simultaneous sequencing of several samples[63]. PCR, including No-Template Controls (NTCs) for each barcoded primer, was performed as previously reported with the exception of the BSA at 10mg/ml concentration per reaction[25]. Only samples whose NTCs yielded an undetectable PCR amplification were retained for further analysis.

Illumina 16S rRNA gene amplicon sequencing

The pooled amplicon library was submitted to the Genome Technology group, The James Hutton Institute (Invergowrie, UK) for quality control, processing and sequencing as previously described[25,64-66]. Briefly, samples were sequenced using an Illumina MiSeq platform with the 2 x 150bp chemistry.

Sequencing reads processing

Sequencing reads were processed and analysed using a custom bioinformatics pipeline. First, QIIME (Quantitative Insights into Microbial Ecology) software, version 1.9.0, was used to process the FASTQ files following default parameters for each step[67]. The forward and reverse read files from individual libraries were decompressed and merged using the command join_paired_ends.py, with a minimum overlap of 30bp between reads. Then, the reads were demultiplexed according to the barcode sequences. Quality filtering was performed using the command split_libraries_fastq.py, imposing a minimum acceptable PHRED score '-q' of 20. Next, these high quality reads were truncated at the 250th nucleotide using the function 'fastq_filter' implemented in USEARCH[68]. Only these high-quality PE, length-truncated reads were used for clustering in Operational Taxonomic Units (OTUs) a 97% sequence identity. OTUs were identified using the 'closed reference' approach against Silva database (version 132)[69]. OTU-picking against the Silva database was performed using the SortMeRNA algorithm[70], producing in an OTU table containing the abundance of OTUs per sample plus a phylogenetic tree. To control for potential contaminant OTUs amplified during library preparation, we retrieved a list of potential environmental contaminant OTUs previously identified in our laboratory[65] and we used this list to filter the results of the aforementioned OTU-enrichment analysis. Additionally, singleton OTUs, (OTUs accounting for only one sequencing read in the whole dataset) and OTUs assigned to chloroplast and mitochondria (taken as plant derived sequences) were removed using the command filter_otus_from_otu_tables.py. Taxonomy matrices, reporting the number of reads assigned to individual phyla, were generated using the command summarize_taxa.py. The OTU table, the phylogenetic tree and the taxonomy matrix, were further used in R for visualizations and statistical analysis.

Statistical analyses I: univariate datasets and 16S rRNA gene alpha and beta-diversity calculations

Analysis of the data was performed in R[71] using a custom script with the following packages: Phyloseq[72] for processing, Alpha and Beta-diversity metrics; ggplot2 [73] for data visualisations; Vegan[74] for statistical analysis of beta-diversity; Ape[75] for phylogenetic tree analysis. For any univariate dataset used (e.g., aboveground biomass; DNA concentration) the normality of the data's distribution was checked using Shapiro–Wilk test. Non-parametric analysis of variance were performed by Kruskal-Wallis Rank Sum Test, followed by Dunn's post hoc test with the functions kruskal.test and the posthoc.kruskal.dunn.test, respectively, from the package PMCMR.

For Alpha-diversity analysis, the OTU table was rarefied at 11,180 reads per sample and this allowed us to retain 8,744 OTUs for downstream analyses (Supplementary Dataset 1: worksheet 6). The Chao1, Observed OTUs and Shannon indices calculated using the function estimate richness in Phyloseq package. Beta-diversity was analysed using a normalized OTU table (i.e., not rarefied) for comparison. For the construction of the normalized OTU table, low abundance OTUs were further filtered removing those not present at least 5 times in 20% of the samples, to improve reproducibility. Then, to control for the uneven number of reads per specimen, individual OTU counts in each sample were divided over the total number of generated reads for that samples and converted in counts per million. Beta-diversity was analysed using two metrics: Bray-Curtis that considers OTUs relative abundance and Weighted Unifrac that additionally is sensitive to phylogenetic classification[76]. These dissimilarity matrices were visualized using Canonical Analysis of Principal coordinates (CAP)[77] using the ordinate function in the Phyloseq package and its significance was inspected using a permutational ANOVA over 5,000 permutations.

Beta-diversity dissimilarity matrices were assessed by Permutational Multivariate Analysis of Variance (Permanova) using Adonis function in Vegan package over 5,000 permutations to calculate effect size and statistical significance.

Statistical analyses II: analysis of Phyla and OTUs differentially enriched among samples

The analysis of the Phyla whose abundances differentiated among rhizosphere and bulk soil samples was performed with analysis of composition of microbiomes (ANCOM)[78] imposing 0.6 cut-off and 0.05 alpha value (taxa-based corrected) as previously described[79].

The analysis of the OTUs whose abundances differentiated among samples was performed a) between individual eco-geographic groups and bulk soil samples to assess the rhizosphere effect and b) between the rhizosphere samples to assess the eco-geographic effect. The eco-geographic effect was further corrected for a microhabitat effect (i.e., for each group, only OTUs enriched against both unplanted soil and at least another barley genotype were retained for further analysis). The analysis was performed using the DESeq2 method[80] with an adjusted *P* value < 0.05 (False Discovery Rate, FDR corrected). This method was selected since it outperforms other hypothesis-testing approaches when data are not normally distributed and a limited number of individual replicates per condition (i.e., approximately 10) are available[81]. DESeq2 was performed

using the eponymous named package in R with the OTU table filtered for low abundance OTUs as an input.

The number of OTUs differentially recruited in the pair-wise comparisons between 'Elite' and wild barley genotypes was visualised using the package UpSetR[82].

The phylogenetic tree was constructed using the representative sequences of the OTUs significantly differentiating 'Elite' genotypes and either 'Desert1' or 'Desert2' samples annotated with iTOL[83].

Statistical analyses III: Correlation plot genetic distance-microbial distance.

To assess the genetic variation on the barley germplasm we used the SNP platform 'BOPA1'[84] comprising 1,536 single nucleotide polymorphisms. We used GenAlex 6.5[85,86] to construct a genetic distance matrix using the simple matching coefficient. Genetic distance for the barley genotypes was visualised by hierarchical clustering using the function hclust in R. Microbial distance was calculated on the average distances for each ecogeographic group using the Weighted Unifrac metric. Correlation between the plant's genetic and microbial distances was performed using a mantel test with the mantel.rtest of the package ade4 in R. The correlation was visualised using the functions ggscatter of the R packages ggpbur.

Availability of Materials and Data

The sequences generated in the 16S rRNA gene sequencing survey are deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB35359. The version of the individual packages and scripts used to analyse the data and generate the figures of this study are available at https://github.com/BulgarelliD-Lab/Barley_B1K

References

1 Alexandratos, N. & Bruinsma, J. World agriculture towards 2030/2050: the 2012 revision. (2012).

2 Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20260-20264, doi:10.1073/pnas.1116437108 (2011).

3 Porter, J. R. *et al.* Climate change 2014: impacts, adaptation, and vulnerability. Part A: global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *Food security and food production systems*, 485-533 (2014).

4 Schlaeppi, K. & Bulgarelli, D. The plant microbiome at work. *Molecular plant-microbe interactions : MPMI* **28**, 212-217, doi:10.1094/MPMI-10-14-0334-FI (2015).

5 Alegria Terrazas, R. *et al.* Plant-Microbiota Interactions as a Driver of the Mineral Turnover in the Rhizosphere. *Adv Appl Microbiol* **95**, 1-67, doi:10.1016/bs.aambs.2016.03.001 (2016).

6 Bulgarelli, D., Schlaeppi, K., Spaepen, S., Van Themaat, E. V. L. & Schulze-Lefert, P. Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology* **64**, 807-838 (2013).

7 Purugganan, M. D. & Fuller, D. Q. The nature of selection during plant domestication. *Nature* **457**, 843-848 (2009).

8 Perez-Jaramillo, J. E., Mendes, R. & Raaijmakers, J. M. Impact of plant domestication on rhizosphere microbiome assembly and functions. *Plant Mol Biol* **90**, 635-644, doi:10.1007/s11103-015-0337-7 (2016).

9 Doebley, J. F., Gaut, B. S. & Smith, B. D. The molecular genetics of crop domestication. *Cell* **127**, 1309-1321 (2006).

10 Cordovez, V., Dini-Andreote, F., Carrión, V. J. & Raaijmakers, J. M. Ecology and Evolution of Plant Microbiomes. *Annual review of microbiology* **73** (2019).

11 Escudero-Martinez, C. & Bulgarelli, D. Tracing the evolutionary routes of plant–microbiota interactions. *Current opinion in microbiology* **49**, 34-40 (2019).

12 Newton, A. C. *et al.* Crops that feed the world 4. Barley: a resilient crop? Strengths and weaknesses in the context of food security. *Food Security* **3**, 141, doi:10.1007/s12571-011-0126-3 (2011).

13 Mascher, M. *et al.* A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**, 426-+, doi:10.1038/nature22043 (2017).

14 Milner, S. G. *et al.* Genebank genomics highlights the diversity of a global barley collection. *Nature genetics* **51**, 319-326 (2019).

15 Bulgarelli, D. *et al.* Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell host & microbe* **17**, 392-403, doi:10.1016/j.chom.2015.01.011 (2015).

16 Mwafulirwa, L. *et al.* Barley genotype influences stabilization of rhizodeposition-derived C and soil organic matter mineralization. *Soil Biology and Biochemistry* **95**, 60-69, doi:10.1016/j.soilbio.2015.12.011 (2016).

17 Lipper, L. *et al.* Climate-smart agriculture for food security. *Nature climate change* **4**, 1068-1072 (2014).

18 Hubner, S. *et al.* Phenotypic landscapes: phenological patterns in wild and cultivated barley. *Journal of evolutionary biology* **26**, 163-174, doi:10.1111/jeb.12043 (2013).

19 Hubner, S. *et al.* Strong correlation of wild barley (Hordeum spontaneum) population structure with temperature and precipitation variation. *Molecular ecology* **18**, 1523-1536, doi:10.1111/j.1365-294X.2009.04106.x (2009).

20 Hübner, S. *et al.* Phenotypic landscapes: phenological patterns in wild and cultivated barley. *Journal of Evolutionary Biology* **26**, 163-174 (2013).

21 Bloom, A. J., Chapin III, F. S. & Mooney, H. A. Resource limitation in plants-an economic analogy. *Annual review of Ecology and Systematics* **16**, 363-392 (1985).

22 Comas, L., Becker, S., Cruz, V. M. V., Byrne, P. F. & Dierig, D. A. Root traits contributing to plant productivity under drought. *Frontiers in plant science* **4**, 442 (2013).

23 Robertson-Albertyn, S. *et al.* Root hair mutations displace the barley rhizosphere microbiota. *Frontiers in plant science* **8**, 1094 (2017).

24 Pérez-Jaramillo, J. E. *et al.* Linking rhizosphere microbiome composition of wild and domesticated Phaseolus vulgaris to genotypic and root phenotypic traits. *The ISME journal* **11**, 2244-2257 (2017).

25 Robertson-Albertyn, S. *et al.* Root Hair Mutations Displace the Barley Rhizosphere Microbiota. *Frontiers in plant science* **8**, 1094, doi:10.3389/fpls.2017.01094 (2017).

26 Hacquard, S. *et al.* Microbiota and host nutrition across plant and animal kingdoms. *Cell host & microbe* **17**, 603-616 (2015).

27 Edwards, J. *et al.* Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E911-920, doi:10.1073/pnas.1414592112 (2015).

28 Perez-Jaramillo, J. E. *et al.* Linking rhizosphere microbiome composition of wild and domesticated Phaseolus vulgaris to genotypic and root phenotypic traits. *The ISME journal* **11**, 2244-2257, doi:10.1038/ismej.2017.85 (2017).

29 Russell, J. *et al.* Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nature Genetics* **48**, 1024 (2016).

30 Pérez-Jaramillo, J. E. *et al.* Deciphering rhizosphere microbiome assembly of wild and modern common bean (Phaseolus vulgaris) in native and agricultural soils from Colombia. *Microbiome* **7**, 1-16 (2019).

31 Garrido-Oter, R. *et al.* Modular Traits of the Rhizobiales Root Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia. *Cell host & microbe* **24**, 155-+, doi:10.1016/j.chom.2018.06.006 (2018).

32 Karasov, T. L. *et al.* Arabidopsis thaliana and Pseudomonas pathogens exhibit stable associations over evolutionary timescales. *Cell host & microbe* **24**, 168-179. e164 (2018).

33 Paterson, E., Gebbing, T., Abel, C., Sim, A. & Telfer, G. Rhizodeposition shapes rhizosphere microbial community structure in organic soil. *New Phytologist* **173**, 600-610 (2007).

34 Neilson, J. W. *et al.* Significant impacts of increasing aridity on the arid soil microbiome. *MSystems* **2**, e00195-00116 (2017).

35 Fierer, N. *et al.* Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 21390-21395, doi:10.1073/pnas.1215210110 (2012).

36 Pérez-Jaramillo, J. E., Carrión, V. J., de Hollander, M. & Raaijmakers, J. M. The wild side of plant microbiomes. *Microbiome* **6**, 143, doi:10.1186/s40168-018-0519-z (2018).

37 Santos-Medellín, C., Edwards, J., Liechty, Z., Nguyen, B. & Sundaresan, V. Drought stress results in a compartment-specific restructuring of the rice root-associated microbiomes. *MBio* **8**, e00764-00717 (2017).

38 Naylor, D., DeGraaf, S., Purdom, E. & Coleman-Derr, D. Drought and host selection influence bacterial community dynamics in the grass root microbiome. *The ISME journal* **11**, 2691-2704 (2017).

39 Romaniuk, K., Golec, P. & Dziewit, L. Insight into the diversity and possible role of plasmids in the adaptation of psychrotolerant and metalotolerant Arthrobacter spp. to extreme Antarctic environments. *Frontiers in microbiology* **9**, 3144 (2018).

40 Hübner, S. *et al.* Strong correlation of wild barley (Hordeum spontaneum) population structure with temperature and precipitation variation. *Molecular Ecology* **18**, 1523-1536 (2009).

41 Bouffaud, M. L. *et al.* Is diversification history of maize influencing selection of soil bacteria by roots? *Molecular ecology* **21**, 195-206, doi:10.1111/j.1365-294X.2011.05359.x (2012).

42 Peiffer, J. A. *et al.* Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences* **110**, 6548-6553, doi:10.1073/pnas.1302837110 (2013).

43 Walters, W. A. *et al.* Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1800918115 (2018).

44 Rolfe, S. A., Griffiths, J. & Ton, J. Crying out for help with root exudates: adaptive mechanisms by which stressed plants assemble health-promoting soil microbiomes. *Current opinion in microbiology* **49**, 73-82 (2019).

45 Cotton, T. A. *et al.* Metabolic regulation of the maize rhizobiome by benzoxazinoids. *The ISME journal* **13**, 1647-1658 (2019).

46 Hu, L. *et al.* Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature communications* **9**, 1-13 (2018).

47 Kudjordjie, E. N., Sapkota, R., Steffensen, S. K., Fomsgaard, I. S. & Nicolaisen, M. Maize synthesized benzoxazinoids affect the host associated microbiome. *Microbiome* **7**, 1-17 (2019).

48 Grün, S., Frey, M. & Gierl, A. Evolution of the indole alkaloid biosynthesis in the genus Hordeum: distribution of gramine and DIBOA and isolation of the benzoxazinoid biosynthesis genes from Hordeum lechleri. *Phytochemistry* **66**, 1264-1272 (2005).

49 Larsson, K. A., Zetterlund, I., Delp, G. & Jonsson, L. M. N-Methyltransferase involved in gramine biosynthesis in barley: Cloning and characterization. *Phytochemistry* **67**, 2002-2008 (2006).

635 50 Matsuo, H. *et al.* Gramine increase associated with rapid and transient systemic
636 resistance in barley seedlings induced by mechanical and biological stresses. *Plant*
637 *and Cell Physiology* **42**, 1103-1111 (2001).
638 51 Voss-Fels, K. P. *et al.* VERNALIZATION1 modulates root system architecture in
639 wheat and barley. *Molecular plant* **11**, 226-229 (2018).
640 52 Maurer, A. *et al.* Modelling the genetic architecture of flowering time control in
641 barley through nested association mapping. *BMC genomics* **16**, 290 (2015).
642 53 Bayer, M. M. *et al.* Development and Evaluation of a Barley 50k iSelect SNP Array.
643 *Frontiers in plant science* **8**, doi:ARTN 1792

644 10.3389/fpls.2017.01792 (2017).
645 54 Bdolach, E. *et al.* Thermal plasticity of the circadian clock is under nuclear and
646 cytoplasmic control in wild barley. *Plant, cell & environment* **42**, 3105-3120 (2019).
647 55 Bai, Y. *et al.* Functional overlap of the Arabidopsis leaf and root microbiota. *Nature*
648 **528**, 364-369 (2015).
649 56 Zhang, J. *et al.* NRT1. 1B is associated with root microbiota composition and
650 nitrogen use in field-grown rice. *Nature biotechnology* **37**, 676-684 (2019).
651 57 Thiergart, T. *et al.* Root microbiota assembly and adaptive differentiation among
652 European Arabidopsis populations. *Nature Ecology & Evolution* **4**, 122-131 (2020).
653 58 The International Barley Genome Sequencing, C. *et al.* A physical, genetic and
654 functional sequence assembly of the barley genome. *Nature* **491**, 711,
655 doi:10.1038/nature11543

656 https://www.nature.com/articles/nature11543#supplementary-information (2012).
657 59 Druka, A. *et al.* Genetic dissection of barley morphology and development. *Plant*
658 *physiology* **155**, 617-627, doi:10.1104/pp.110.166249 (2011).
659 60 Kleinhofs, A. *et al.* A molecular, isozyme and morphological map of the barley
660 (Hordeum vulgare) genome. *Theoretical and Applied Genetics* **86**, 705-712 (1993).
661 61 Bulgarelli, D. *et al.* The CC-NB-LRR-Type Rdg2a Resistance Gene Confers
662 Immunity to the Seed-Borne Barley Leaf Stripe Pathogen in the Absence of
663 Hypersensitive Cell Death. *PloS one* **5**, doi:ARTN e12599

664 10.1371/journal.pone.0012599 (2010).
665 62 Tottman, D., Makepeace, R. & Broad, H. An explanation of the decimal code for the
666 growth stages of cereals, with illustrations. *Annals of Applied Biology* **93**, 221-234
667 (1979).

668 63 Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the
669 Illumina HiSeq and MiSeq platforms. *The ISME journal* **6**, 1621-1624,
670 doi:10.1038/ismej.2012.8 (2012).
671 64 Caradonia, F. *et al.* Nitrogen Fertilizers Shape the Composition and Predicted
672 Functions of the Microbiota of Field-Grown Tomato Plants. *Phytobiomes Journal* **3**,
673 315-325 (2019).
674 65 Pietrangelo, L., Bucci, A., Maiuro, L., Bulgarelli, D. & Naclerio, G. Unraveling the
675 Composition of the Root-Associated Bacterial Microbiota of Phragmites australis
676 and Typha latifolia. *Frontiers in Microbiology* **9**, doi:ARTN 1650

677 10.3389/fmicb.2018.01650 (2018).
678 66 Rathore, R. *et al.* Crop Establishment Practices Are a Driver of the Plant Microbiota
679 in Winter Oilseed Rape (Brassica napus). *Frontiers in Microbiology* **8**, doi:ARTN
680 1489

681 10.3389/fmicb.2017.01489 (2017).
682 67 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community
683 sequencing data. *Nature methods* **7**, 335-336, doi:10.1038/nmeth.f.303 (2010).
684 68 Edgar, R. C. Search and clustering orders of magnitude faster than BLAST.
685 *Bioinformatics* **26**, 2460-2461, doi:10.1093/bioinformatics/btq461 (2010).
686 69 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data
687 processing and web-based tools. *Nucleic acids research* **41**, D590-596,
688 doi:10.1093/nar/gks1219 (2013).
689 70 Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of
690 ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211-3217 (2012).
691 71 Team, R. C. R: A language and environment for statistical computing. (2013).
692 72 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive
693 analysis and graphics of microbiome census data. *PloS one* **8**, e61217,
694 doi:10.1371/journal.pone.0061217 (2013).
695 73 Wickham, H. *ggplot2: elegant graphics for data analysis*. (Springer, 2016).
696 74 Oksanen, J. *et al.* vegan: community ecology package. R package version 2.0–7.
697 2013. *URL http://CRAN. R-project. org/package= vegan* (2016).
698 75 Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and
699 evolution in R language. *Bioinformatics* **20**, 289-290 (2004).

700 76 Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. & Knight, R. UniFrac: an
701 effective distance metric for microbial community comparison. *The ISME journal* **5**,
702 169-172, doi:10.1038/ismej.2010.133 (2011).
703 77 Anderson, M. J. & Willis, T. J. Canonical analysis of principal coordinates: a useful
704 method of constrained ordination for ecology. *Ecology* **84**, 511-525 (2003).
705 78 Mandal, S. *et al.* Analysis of composition of microbiomes: a novel method for
706 studying microbial composition. *Microbial ecology in health and disease* **26**, 27663
707 (2015).
708 79 Morton, J. T. *et al.* Establishing microbial composition measurement standards with
709 reference frames. *Nature communications* **10**, 1-11 (2019).
710 80 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
711 dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550,
712 doi:10.1186/s13059-014-0550-8 (2014).
713 81 Weiss, S. *et al.* Normalization and microbial differential abundance strategies
714 depend upon data characteristics. *Microbiome* **5**, 27, doi:10.1186/s40168-017-
715 0237-y (2017).
716 82 Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the
717 visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938-2940
718 (2017).
719 83 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic
720 tree display and annotation. *Bioinformatics* **23**, 127-128 (2006).
721 84 Close, T. J. *et al.* Development and implementation of high-throughput SNP
722 genotyping in barley. *Bmc Genomics* **10**, doi:Artn 582

723 10.1186/1471-2164-10-582 (2009).
724 85 PE, P. R. S. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for
725 teaching and research–an update. *Bioinformatics* **28**, 2537 (2012).
726 86 Peakall, R. & Smouse, P. E. GENALEX 6: genetic analysis in Excel. Population
727 genetic software for teaching and research. *Molecular ecology notes* **6**, 288-295
728 (2006).

729

730

Acknowledgements

We are grateful to Prof Andy Flavell (University of Dundee) for providing us with the 'B1K' seeds used in this study. We thank Malcolm Macaulay for the technical assistance during the sequencing library preparation. We thank Dr Timothy George (The James Hutton Institute) for the critical comments on the manuscript.

Funding

This work was supported by a Royal Society of Edinburgh/Scottish Government Personal Research Fellowship co-funded by Marie Curie Actions awarded to DB and a Scottish Food Security Alliance-Crops PhD studentship awarded by the University of Dundee, the University of Aberdeen, and the James Hutton Institute to RAT. RAT and DB are currently supported by the H2020 Innovation Action 'Circles' (European Commission, Grant agreement 818290) awarded to the University of Dundee.

Authors information

Contributions

The study was conceived by RAT and DB with critical inputs from EP and EB. RAT and KBC performed the experiments. JM and PH generated the 16S rRNA sequencing reads. JR provided access to the molecular marker information of the barley genome. EF provided access to the eco-geographical and phenotypic data of the B1K accessions. RAT and DB analysed the data. All authors critically reviewed and edited the manuscript and approved its publication.

Corresponding author

Correspondence to Davide Bulgarelli

Ethics declarations

Competing Interests

The authors declare no competing interests

Figure legends

Figure 1 Plant growth parameters of the wild and domesticated barley genotypes. (**a**) Distribution of the twenty wild barley genotypes used in this study in the Israeli geographic region. Individual dots depict the approximate sampling location for a given genotype, colour-coded according to the designated ecogeographic area. (**b**) Stem dry weight of the barley genotypes at the time of sampling. (**c**) Ratio between root and shoot dry weight of the indicated samples. In **b** and **c**, upper and lower edges of the box plots represent the upper and lower quartiles, respectively. The bold line within the box denotes the median, individual shapes depict measurements of individual biological replicates/genotypes for a given group. Different letters within the individual plots denote statistically significant differences between means by Kruskal-Wallis non-parametric analysis of variance followed by Dunn's post-hoc test ($P < 0.05$).

Figure 2 The dominant phyla of the bulk soil and rhizosphere microbiota are conserved across barley genotypes. Average relative abundance (% of sequencing reads) of the dominant phyla retrieved from the microbial profiles of indicated samples. Only phyla displaying a minimum average relative abundance of 1% included in the graphical representation. Stars depict phyla enriched in and discriminating between rhizosphere and between bulk soil samples (ANCOM, cut-off 0.6, alpha 0.05, taxa-corrected).

Figure 3 Wild and elite barley genotypes fine-tune the composition of the rhizosphere bacterial microbiota. Principal Coordinates Analysis of the Weighted Unifrac dissimilarity matrix of the microbial communities retrieved from the indicated sample types. Individual shape depicts individual biological replicates colour-coded according to the designated ecogeographic area.

Figure 4 Enrichments of individual bacteria discriminates between elite varieties and wild barley genotypes. Horizontal black bars denote the number of OTUs differentially enriched (Wald test, P value < 0.05, FDR corrected) in the indicated pair-wise comparisons between elite varieties and groups of wild barley genotypes. Vertical bars depict the number of differentially enriched OTUs unique for or shared among two or more comparisons highlighted by the interconnected dots underneath the vertical bars. Coloured vertical bars depict differentially enriched OTUs unique for the pair-wise comparisons between 'Elite' and 'Coast 2' (C2, dark blue), 'Coast 1' (C1, light blue), 'North' (N, green), 'Desert 1' (D1, yellow) and 'Desert 2' (D2, orange), respectively.

792 **Figure 5. Actinobacteria are preferentially enriched in and discriminate between elite**
793 **genotypes and wild barley genotypes adapted to desert environments.** Individual
794 external nodes of the tree represent one of the OTUs enriched in the rhizosphere of elite
795 genotypes compared to either (or both) rhizosphere samples desert areas (Wald test, *P*
796 value < 0.05, FDR corrected). The colours reflect OTUs' taxonomic affiliation at Phylum
797 level. A black bar in the outer rings depicts whether that given OTU was identified in the
798 comparisons between 'Elite' and either 'Desert 1' or 'Desert 2' genotypes, respectively.
799 Phylogenetic tree constructed using OTUs 16S rRNA gene representative sequences.

800 **Figure 6**: **Mantel test between genetic distance and microbial distance in the wild**
801 **barley rhizosphere.** Individual dots depict individual comparison of any given pair of wild
802 barley genotypes between average value of weighted unifrac distance (y-axis) and genetic
803 distance shown as simple matching coefficients (x-axis). The blue line depicts the
804 regression line, the grey shadow the 95% confidence interval, respectively.

805

Tables

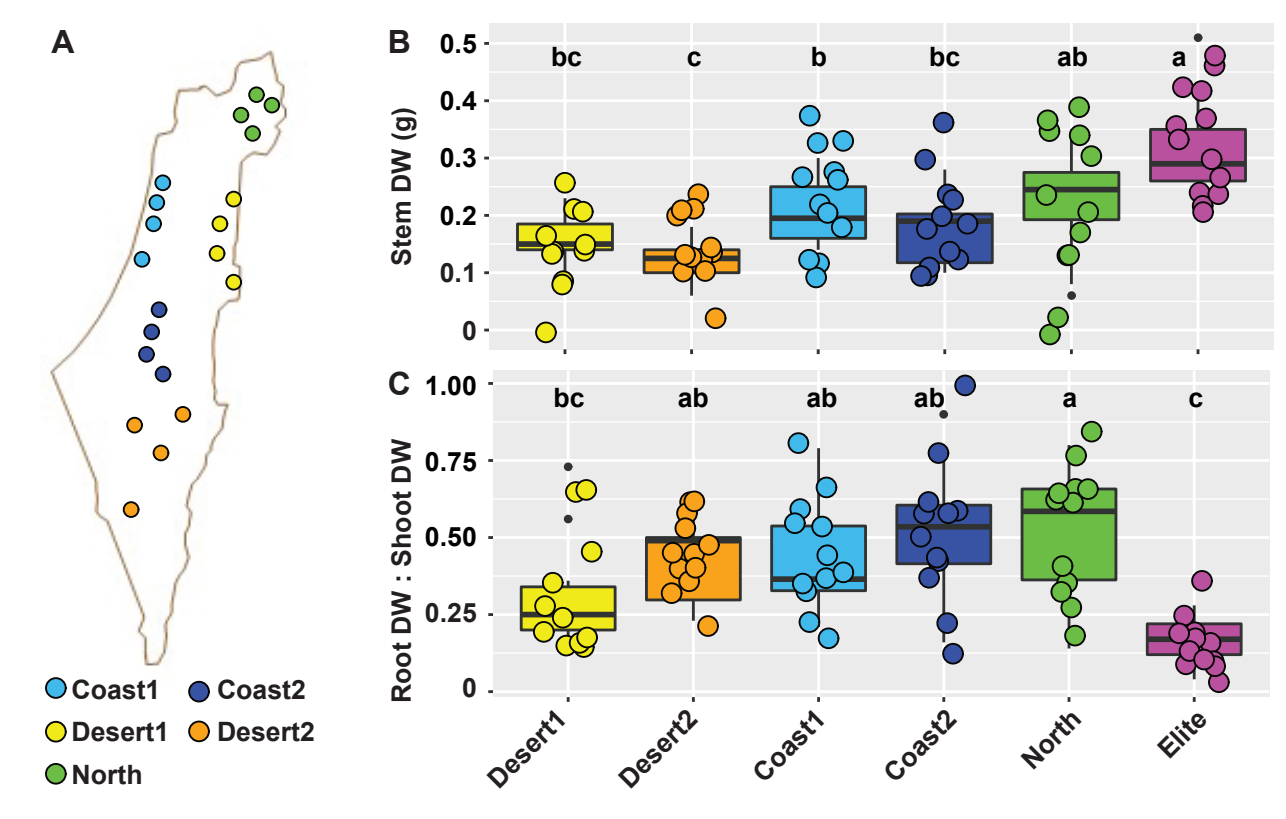
Table 1. Description of the genotypes used in this study. Eco-geographical group; sampling site or type of the Elite material, genotype ID; mean annual rainfall (MAR*), mid-day temperature in January (MDT1*), Elevation, and soil bulk density (Db*), organic matter content (OM*) of the 'B1K' sampling sites from[19,40]. (**) missing data

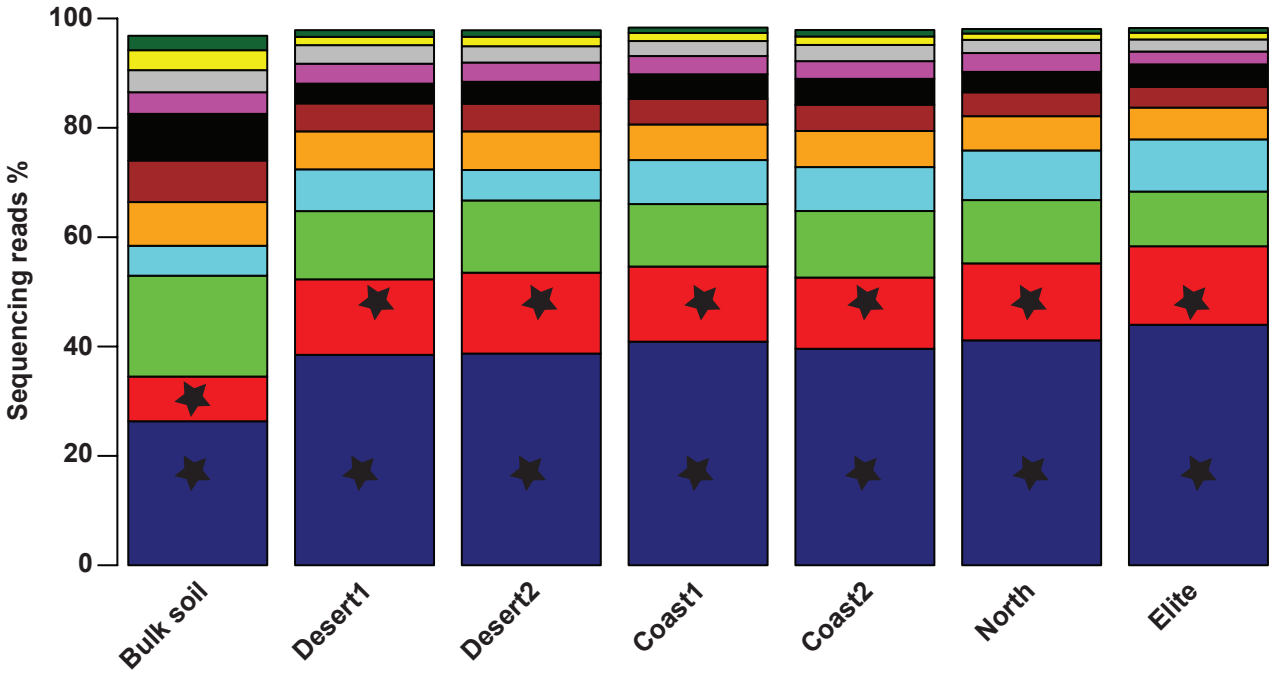
Eco-geographical group	Site/ Elite-type	Genotype ID	MAR* (mm)	MDT1* (℃)	Elevation* (m)	Db* (g/ml)	OM* (%)
Coast 1							
	Michmoret	B1K.03.09	569	12.3	19	1.32	0.979
	Dor	B1K.20.13	543	12.3	16	1.06	6.659
	Kerem Maharal	B1K.21.11	602	11.9	92	1.04	11.616
	Oren Canyon	B1K.30.07	623	11.9	98	1.02	9.430
Coast 2							
	Amatzya	B1K.17.10	366	10.5	355	1.21	2.564
	Shomerya	B1K.18.16	318	10.1	441	1.13	3.946
	Beit Govrin	B1K.35.11	386	10.8	303	0.97	5.251
	Sinsan Stream	B1K.48.06	471	10.4	358	1.05	6.242
Desert 1							
	Ein Prat	B1K.04.04	388	10.4	319	m.d. (**)	m.d. (**)
	Neomi	B1K.05.13	153	13.1	-245	1.28	4.460
	Talkid Stream	B1K.08.18	215	12.7	-253	1.05	2.077
	Kidron Stream	B1K.12.10	87	14	-380	1.38	1.609
Desert 2							
	Yeruham	B1K.02.18	112	9.9	535	1.41	0.175
	Shivta	B1K.11.11	88	10.7	358	1.43	1.138
	Mt. Harif	B1K.33.03	74	8.3	860	1.52	0.820
	Havarim Stream	B1K.34.20	93	10.1	485	1.32	1.337
North							
	Susita	B1K.14.04	444	10.5	51	0.93	7.551
	Hamat Gader	B1K.15.19	436	11.3	-69	0.96	4.122
	Avny hill	B1K.31.01	502	10.4	177	1.13	5.161
	Almagor	B1K.37.06	461	11.1	-37	1.11	6.096
Elite							
	Two-row/ malting	Barke					
	Two-row/ feeding	Bowman					
	Six-row/ malting	Morex					
	Six-row/ feeding	Steptoe					

Table 2. Proportion of rhizosphere microbiota variance explained by the indicated variables and corresponding statistical significance. Levels of the factor Microhabitat are either 'Bulk soil' or 'Rhizosphere'. Levels of the factor Eco-geography are the groups 'Coast 1'; 'Coast 2'; 'Desert 1'; 'Desert 2'; 'North'; and 'Elite', respectively.

Weighted Unifrac		
Factor	R2	Pr(>F)
Microhabitat	0.285	<0.001
Eco-geography*	0.168	<0.001
Bray-Curtis		
Factor	R2	Pr(>F)
Microhabitat	0.221	<0.001
Eco-geography*	0.129	<0.001

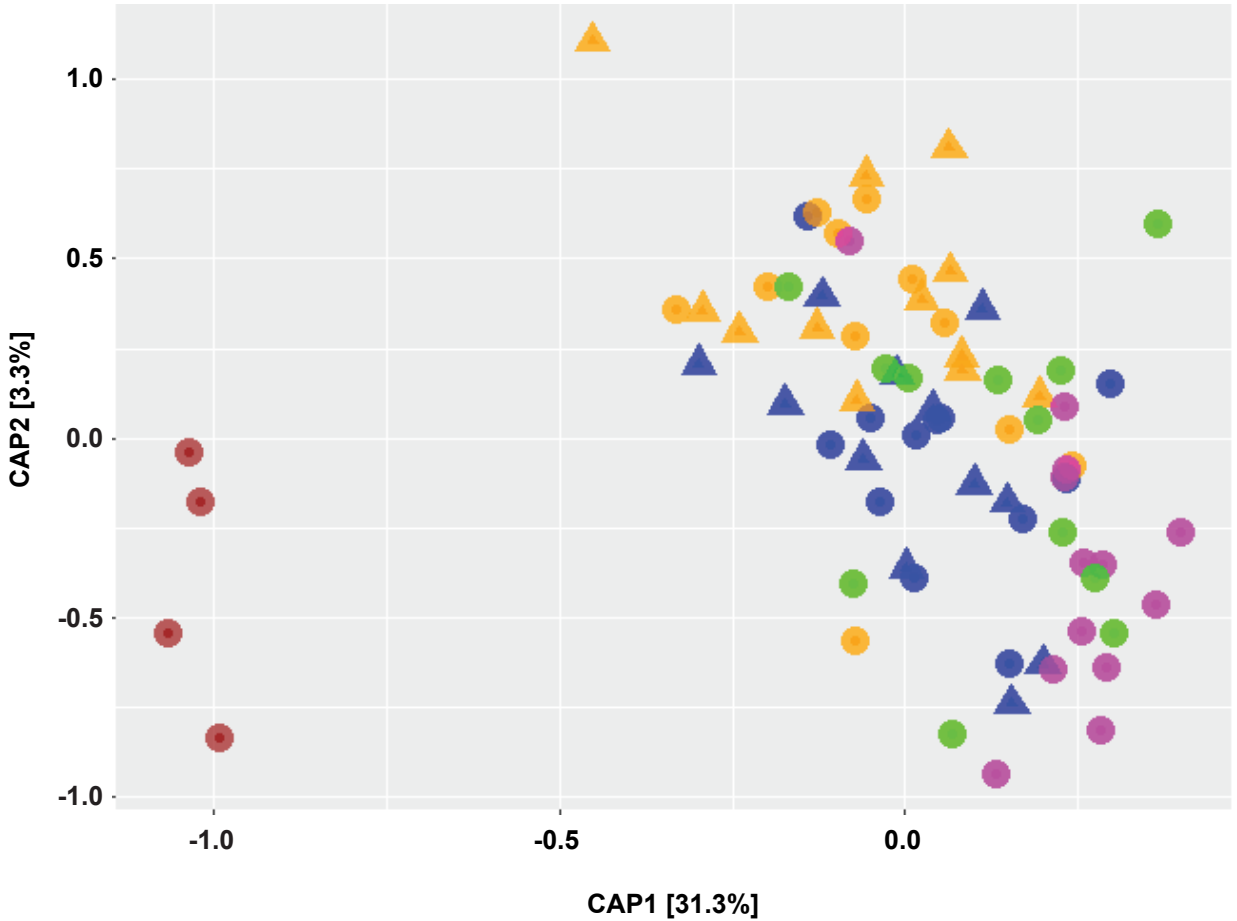
(*) Analysis performed in rhizosphere samples only

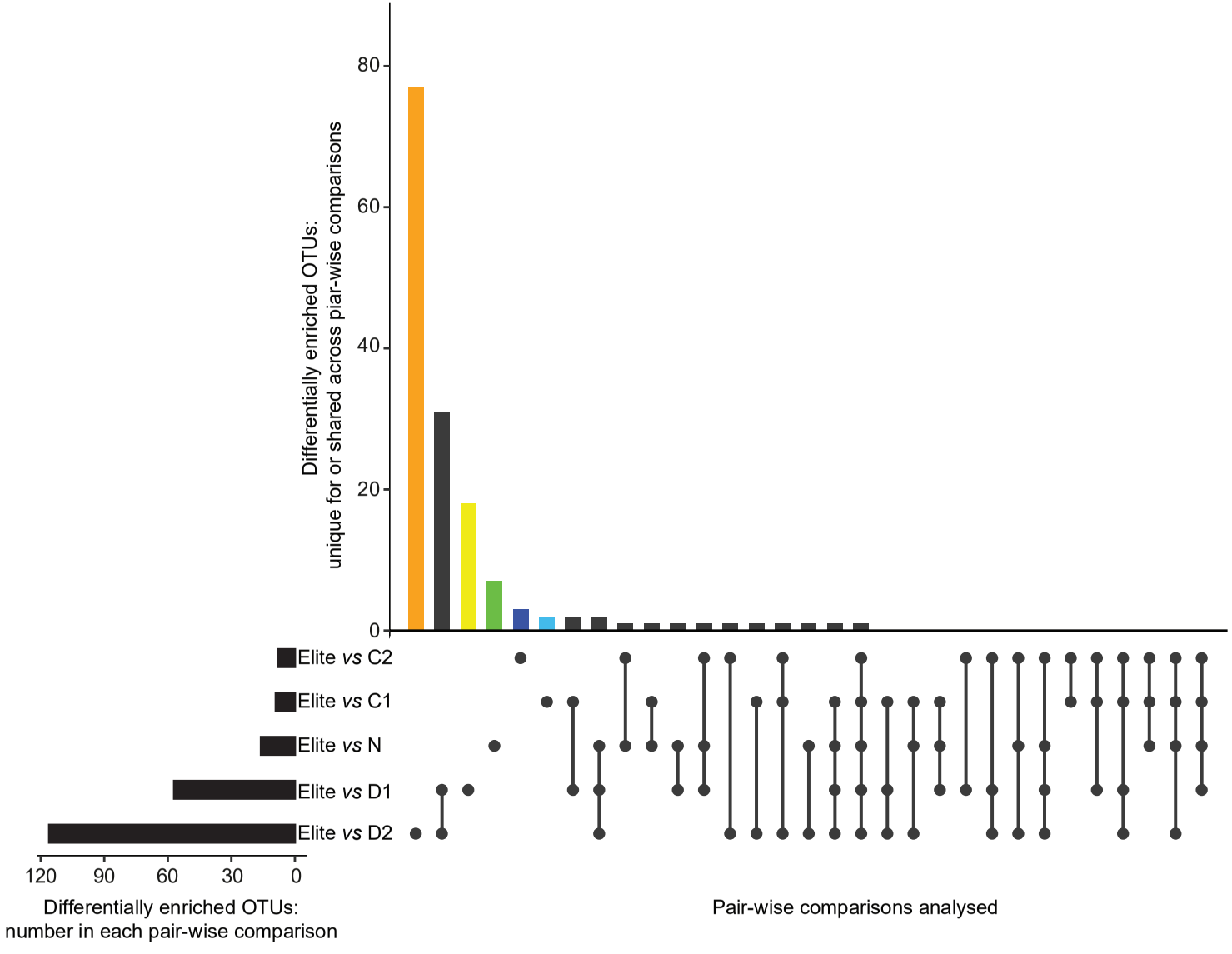

A
Coast1
Coast2
Desert1
Desert2
North
B
Stem DW (g)
0.5
0.4
0.3
0.2
0.1
0
bc
c
b
bc
ab
a
C
Root DW : Shoot DW
1.00
0.75
0.50
0.25
0
bc
ab
ab
ab
a
c
Desert1
Desert2
Coast1
Coast2
North
Elite

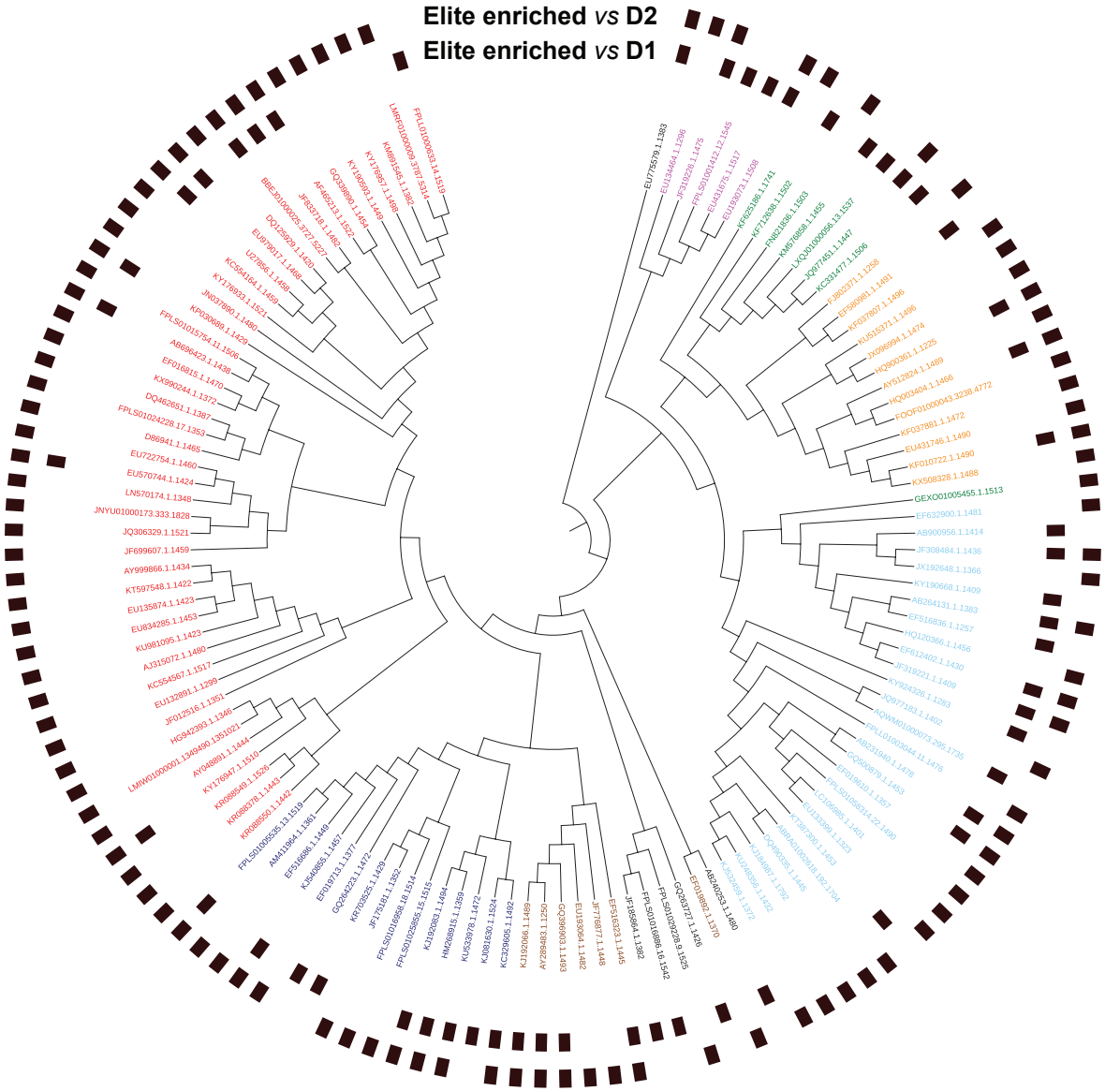


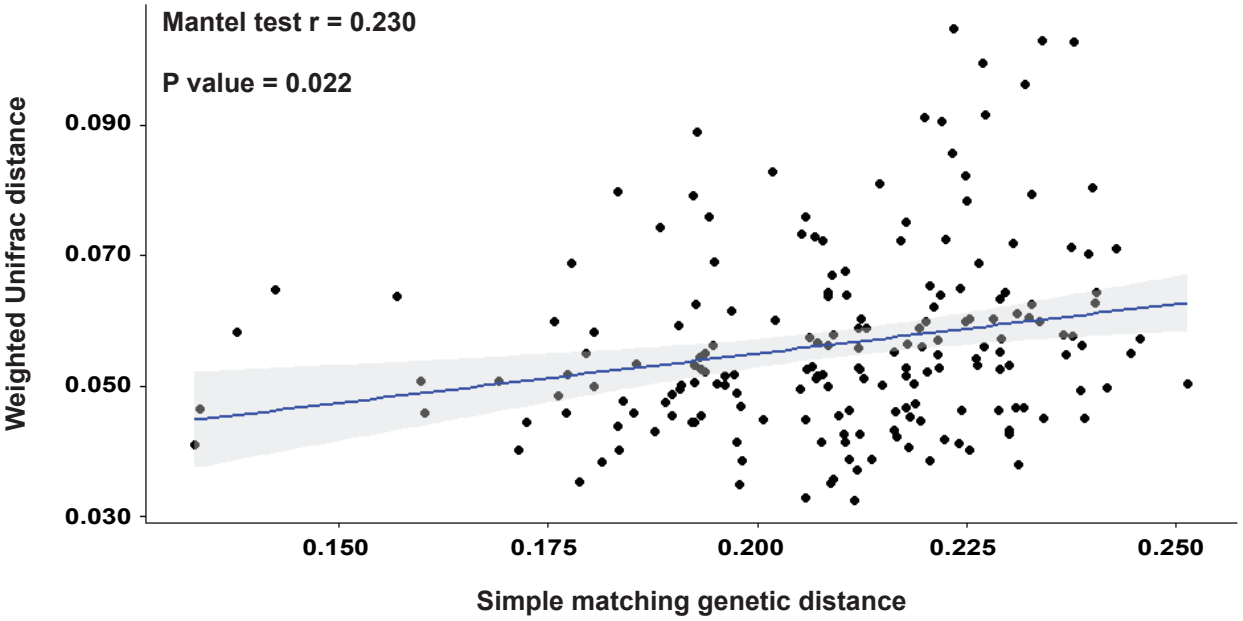
Taxonomy

- Acidobacteria
- Actinobacteria
- Bacteroidetes
- Chloroflexi
- Latescibacteria
- Gemmatimonadetes
- Planctomycetes
- Proteobacteria
- Rokubacteria
- Verrucomicrobia
- Thaumarchaeota


1.0
0.5
0.0
-0.5
-1.0
CAP2 [3.3%]
-1.0
-0.5
0.0
CAP1 [31.3%]
Sample type
Bulk soil
North
Coast1
Coast2
Desert1
Desert2
Elite


Differentially enriched OTUs:
unique for or shared across piar-wise comparisons
80
60
40
20
0
Elite vs C2
Elite vs C1
Elite vs N
Elite vs D1
Elite vs D2
120
90
60
30
0
Differentially enriched OTUs:
number in each pair-wise comparison
Pair-wise comparisons analysed


Elite enriched *vs* D2
Elite enriched *vs* D1
OTU taxonomy
Actinobacteria
Alphaproteobacteria
Bacteroidetes
Betaproteobacteria
Deltaproteobacteria
Gammaproteobacteria
Gemmatimonadetes
Other phyla


Mantel test r = 0.230
P value = 0.022
Weighted Unifrac distance
0.090
0.070
0.050
0.030
0.150
0.175
0.200
0.225
0.250
Simple matching genetic distance