ARTICLES
https://doi.org/10.1038/s41559-019-1063-3
nature ecology & evolution

Root microbiota assembly and adaptive differentiation among European *Arabidopsis* populations

Thorsten Thiergart[1,7], Paloma Durán[1,7], Thomas Ellis[2], Nathan Vannier[1], Ruben Garrido-Oter[1,3], Eric Kemen[4], Fabrice Roux[5], Carlos Alonso-Blanco[6], Jon Ågren[2*], Paul Schulze-Lefert[1,3*] and Stéphane Hacquard[1,3*]

Factors that drive continental-scale variation in root microbiota and plant adaptation are poorly understood. We monitored root-associated microbial communities in *Arabidopsis thaliana* and co-occurring grasses at 17 European sites across 3 years. We observed strong geographic structuring of the soil biome, but not of the root microbiota. A few phylogenetically diverse and geographically widespread bacteria consistently colonized plant roots. Among-site and across-year similarity in microbial community composition was stronger for the bacterial root microbiota than for filamentous eukaryotes. In a reciprocal transplant between two *A. thaliana* populations in Sweden and Italy, we uncoupled soil from location effects and tested their contributions to root microbiota variation and plant adaptation. Community differentiation in plant roots was explained primarily by location for filamentous eukaryotes and by soil origin for bacteria, whereas host genotype effects were marginal. Strong local adaptation between the two *A. thaliana* populations was observed, with differences in soil properties and microbes of little importance for the observed magnitude of adaptive differentiation. Our results suggest that, across large spatial scales, climate is more important than soil conditions for plant adaptation and variation in root-associated filamentous eukaryotic communities, whereas soil properties are primary drivers of bacterial community differentiation in roots.

Plants interact with multi-kingdom microbial communities (for example, bacteria, fungi and oomycetes) that can impact host fitness—either directly, or indirectly through microbe–microbe interactions[1–5]. The immune systems of plants, rhizodeposits and microbial interactions are known determinants of root-associated microbial assemblages and make them distinct from the surrounding soil biota[3,6–8]. Large-scale spatial variation in the composition of the soil biota has been associated with differences in edaphic and climatic conditions[9]. Particularly, local edaphic factors such as pH primarily predict the geographic distribution of soil bacteria[10–12], whereas climatic variables better predict fungal distribution in soil[13]. However, systematic field studies exploring the extent to which variations in soil and climatic conditions impact root microbiota composition and adaptive differentiation in plants are lacking.

Reciprocal transplants have shown local adaptation in a large number of plant species, including the model plant *Arabidopsis thaliana*[14,15]. However, the relative importance of different abiotic and biotic factors for the evolution and maintenance of local adaptation is poorly known[16–18]. Soil edaphic factors and soil microbes are known to influence flowering phenology and modulate host fitness in natural soils[19–24]. Yet, information about the extent to which differences in soil properties contribute to divergent selection and the maintenance of adaptive differentiation among plant populations is still limited beyond examples of adaptation to extreme soil conditions[25].

Here, we determined whether the roots of *A. thaliana* and co-occurring grasses establish stable associations with microbial communities across a latitudinal gradient in Europe, and disentangled the extent to which soil conditions and climate drive root microbiota assembly and adaptive differentiation between two *A. thaliana* populations in Sweden and Italy.

Results

Continental-scale survey of the *A. thaliana* root microbiota. We sampled natural *A. thaliana* populations at the flowering stage at 17 sites along a latitudinal gradient in Europe in 2015, 2016 and 2017. We harvested bulk soil (soil), rhizosphere, rhizoplane and root endosphere (root) compartments (Extended Data Fig. 1a,b) of *A. thaliana* and co-occurring grasses at four sites in Sweden (SW1–4), six in Germany (GE1–6), three in France (FR1–3), one in Italy (IT1) and three in Spain (SP1–3) (Fig. 1a and Supplementary Table 1). Bacterial, fungal and oomycetal communities were profiled using primer pairs targeting the V3–V4 and V5–V7 regions of the bacterial 16S ribosomal RNA (rRNA) gene, the internal transcribed spacer 1 and 2 (ITS1 and ITS2) segments of the fungal ITS, and the ITS1 segment of the oomycetal ITS (Supplementary Table 2). Given the limited primer bias observed (Extended Data Fig. 1c), the bacterial V5–V7, fungal ITS1 and oomycetal ITS1 regions were selected for later analyses. Sequences were clustered into operational taxonomic units (OTUs; 97% sequence similarity) and the robustness of

[1]Max Planck Institute for Plant Breeding Research, Cologne, Germany. [2]Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden. [3]Cluster of Excellence on Plant Sciences (CEPLAS), Max Planck Institute for Plant Breeding Research, Cologne, Germany. [4]Department of Microbial Interactions, IMIT/ZMBP, University of Tübingen, Tübingen, Germany. [5]LIPM, INRA, CNRS, Université de Toulouse, Castanet-Tolosan, France. [6]Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología (CNB), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain. [7]These authors contributed equally: Thorsten Thiergart, Paloma Durán. *e-mail: jon.agren@ebc.uu.se; schlef@mpipz.mpg.de; hacquard@mpipz.mpg.de

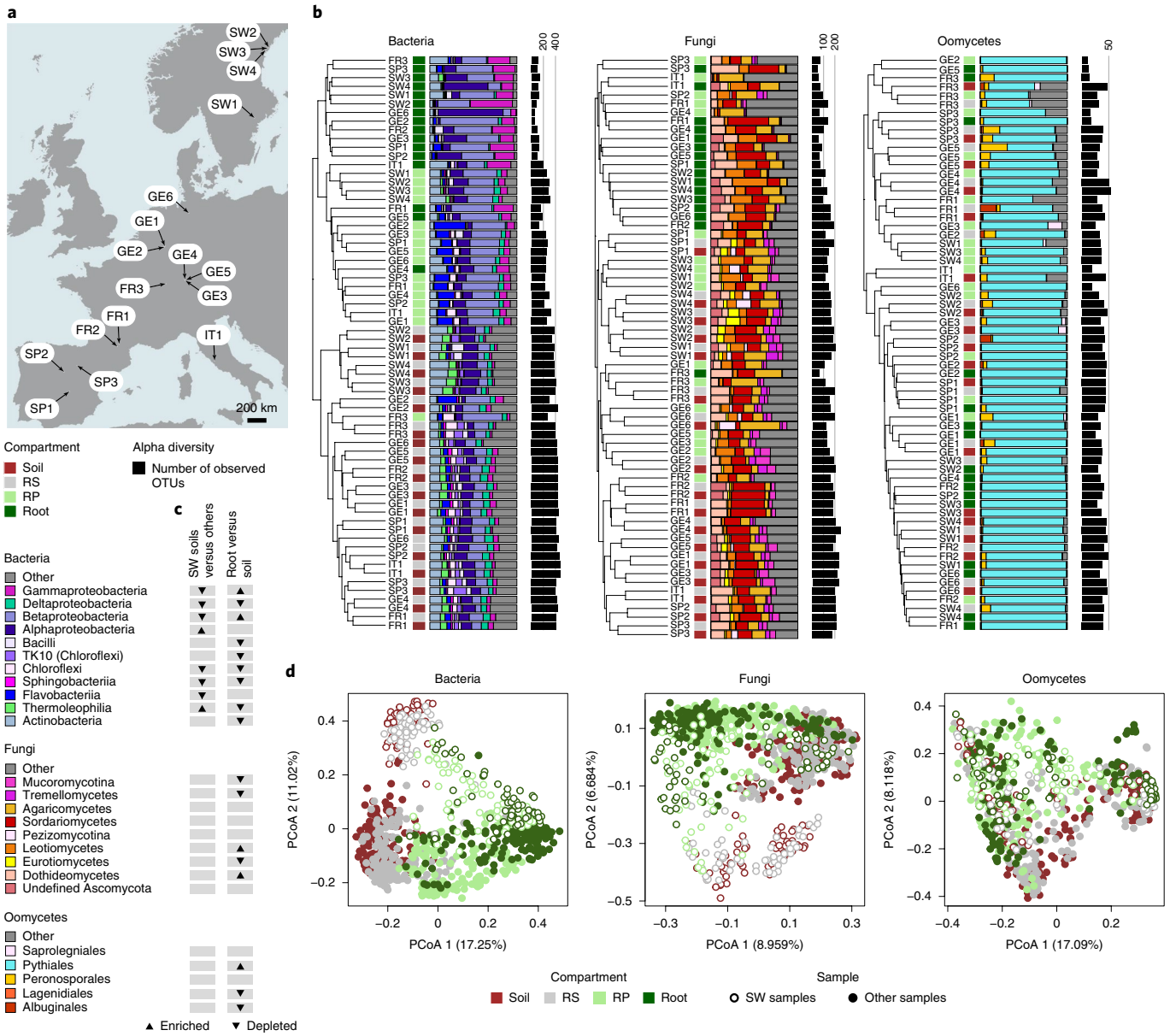


Fig. 1 | Microbial community structure in 17 European *A. thaliana* populations. a, European map showing the names and locations of the 17 *A. thaliana* populations. **b**, Bray–Curtis similarity-based dendrogram showing averaged bacterial (left), fungal (middle) and oomycetal (right) community composition for each compartment at each site across the 3 years. Only OTUs with RA > 0.1% were considered. The total number of processed samples was 896 and only those with more than 1,000 reads were used to calculate average Bray–Curtis distances. Compartments are indicated with coloured squares (dark red: soil; grey: rhizosphere (RS); light green: rhizoplane (RP); dark green: root). For each sample, community composition (class or order level) is indicated with bar plots, and microbial alpha diversity is represented with black bars according to the number of observed OTUs in the corresponding rarefied datasets (1,000 reads). **c**, Differential abundance analysis (class or order level) between the four Swedish soil samples and the other 13 soils, as well as between root and soil samples. Triangles depict statistically significant differences (Wilcoxon rank-sum test, FDR < 0.05). **d**, PCoA based on Bray–Curtis distances between samples harvested across 17 sites, four compartments and three successive years (2015, 2016 and 2017). Microbial communities are presented for the whole *A. thaliana* dataset for bacteria ($n = 881$), fungi ($n = 893$) and oomycetes ($n = 875$) and are colour coded according to the compartment. Open circles represent the Swedish samples and closed circles represent all of the other samples. OTUs with RA < 0.1% were excluded from the dataset. Map data in **a** adapted from Google Maps, 2018.

our conclusions was validated at amplicon sequence variant (ASV) resolution (UNOISE2; ref. [26]).

Convergence in root microbiota composition across *A. thaliana* populations. If plant roots establish stable associations with microbial communities across large geographical distances, we expect a strong host-filtering effect on root microbiota composition. Inspection of alpha-diversity indices across all 17 sites and 3 years revealed a gradual decrease of bacterial, fungal and oomycetal diversity from the soil to the root endosphere (Kruskal–Wallis with Dunn's post-hoc test, $P < 0.05$), with a stronger decrease for bacteria than for filamentous eukaryotes (Fig. 1b and Extended Data Fig. 2a,b). Analysis of microbial community structure based on average Bray–Curtis distances across replicates and years

revealed that bacterial communities in the root endosphere and rhizoplane cluster by compartments, and by site in the rhizosphere and soil (Fig. 1b). In contrast, no clear clustering by compartment was observed for fungi and oomycetes (Fig. 1b). This pattern was corroborated by permutational multivariate analysis of variance (PERMANOVA) with Bray–Curtis distances (Adonis function in the R library vegan; Supplementary Table 3), which indicated that the factor site explained far greater variation in fungal and oomycetal community composition (Adonis: degrees of freedom (d.f.) = 16; fungi: coefficient of determination (R^2) = 0.21; $P < 0.001$; oomycetes: $R^2 = 0.14$; $P < 0.001$) than did compartment (Adonis: d.f. = 3; fungi: $R^2 = 0.06$; $P < 0.001$; oomycetes: $R^2 = 0.03$; $P < 0.001$; Supplementary Table 3). In contrast, variation in bacterial community composition was similarly explained by site (Adonis: d.f. = 16; $R^2 = 0.18$; $P < 0.001$) and compartment (Adonis: d.f. = 3; $R^2 = 0.16$; $P < 0.001$) (Supplementary Table 3).

Compared with surrounding soil samples, microbial communities in plant roots showed a significant increase in relative abundance (RA; average across replicates and years; $n = 17$) for taxa belonging to the bacterial classes Betaproteobacteria and Gammaproteobacteria, the fungal classes Leotiomycetes and Dothideomycetes and the oomycetal order Pythiales (Wilcoxon rank-sum test, false discovery rate (FDR) < 0.05; Fig. 1c and Extended Data Fig. 2c). Enrichment tests based on taxa RA (class level; average across replicates and years; $n = 17$) and principal coordinate analysis (PCoA) ordination based on Bray–Curtis distances revealed marked differences in soil bacterial communities between Swedish soils and the other European soils (Fig. 1c,d and Extended Data Fig. 2d). Notably, these differences largely diminished in the corresponding root samples, suggesting convergence in bacterial community composition in plant endosphere (Fig. 1d). Calculation of pairwise distances between site centroids ($n = 17$; Bray–Curtis distances) for soil and root samples separately validated that between-site variation in bacterial community composition is significantly reduced in root samples compared with soil samples (t-test, d.f. = 518; $t = 5.5$; $P < 0.001$; Supplementary Fig. 3a). Furthermore, the percentage of shared abundant bacterial OTUs is significantly higher across root samples than across soil samples (t-test, d.f. = 18,328; $t = -63$; $P < 0.001$; Extended Data Fig. 3b,c). Importantly, all beta-diversity metrics reported above based on OTU clustering remain stable at ASV resolution (Supplementary Table 3 and Extended Data Fig. 4). Our results show that the root environment drives convergence in bacterial community composition across European sites.

Root endosphere bacteria and *A. thaliana* exhibit stable associations. We hypothesized that root colonization by geographically widespread taxa contributes to convergence in root microbiota composition. To identify microbial OTUs that are widely distributed in the roots of *A. thaliana*, we calculated their prevalence across all 17 sites (the average of all 3 years), and those detected in >80% of the sites were defined as geographically widespread (Fig. 2a and Extended Data Fig. 5). We observed a positive correlation (linear regression: $R^2 = 0.24$; $n = 401$; $P < 0.001$) between the prevalence of root-associated bacteria across sites and their RA in root samples (Fig. 2a). In contrast, no corresponding correlation was observed for root-associated fungi and oomycetes. By inspecting the abundance profiles of OTUs with restricted or widespread geographic distribution across compartments (Fig. 2b and Extended Data Fig. 5), we observed that geographically widespread bacterial OTUs, but not fungal and oomycetal OTUs, were significantly more abundant in rhizoplane and root samples than in the corresponding soil samples (Kruskal–Wallis with Dunn's post-hoc test, $P < 0.01$; Fig. 2b).

We identified 13 geographically widespread bacterial OTUs that were consistently detected in the root endosphere of *A. thaliana* across sites (Fig. 2a,c and Supplementary Table 4) and accounted for 38% of the total RA in this compartment (Fig. 2d). These 13 OTUs, which belong to diverse genera (for example, *Bradyrhizobium*, *Pseudomonas*, *Polaromonas*, *Acidovorax*, *Ralstonia*, *Massilia*, *Burkholderia*, *Kineospora* and *Flavobacterium*) were detected across all 3 years (Fig. 2a), and nine of them were significantly enriched in plant roots compared with soil (generalized linear model, FDR < 0.05; highlighted with an asterisk in Fig. 2c). Inspection of intra-OTU sequence variants for each of these OTUs revealed substantial sequence heterogeneity, ranging from one to eight ASVs per OTU. Sequence heterogeneity was the highest for OTU14 (*Burkholderia*; eight ASVs) and OTU5 (*Pseudomonas*; six ASVs) (Fig. 2c). In most cases, only one or two sequence variants showed both high prevalence (>80%) and high RA (>5%) across sites, suggesting that very few ASVs contribute to the ubiquitous nature of these OTUs. We also identified 14 geographically widespread OTUs of root-associated filamentous eukaryotes. In contrast with bacteria, these OTUs were not particularly root enriched and only five were detectable across all years (Fig. 2a,c and Supplementary Table 4). Few sequence variants were observed for these OTUs (ranging from one to two ASVs per OTU), probably due to the narrower phylogenetic breadth of the ITS versus 16S rRNA locus[27]. Our results suggest that the few geographically widespread bacteria that abundantly colonize the roots of *A. thaliana* drive convergence in bacterial community composition at a continental scale.

Conservation of geographically widespread OTUs in the roots of grass species. The taxa that consistently colonize the roots of *A. thaliana* might also associate with distantly related plant species. To test this hypothesis, we harvested co-occurring grasses at each of the 17 sites. Amplification of the plant *rbcL* and *MatK* sequences[28] revealed that 16 different grass species were harvested across the sites and years (Extended Data Fig. 6a). There was limited differentiation in microbial community composition observed between *A. thaliana* and grasses (data pooled across grass species; Adonis: d.f. = 1; root-associated bacteria: $R^2 = 0.01$; $P < 0.001$; root-associated fungi: $R^2 = 0.01$; $P < 0.001$; root-associated oomycetes: $R^2 = 0.01$; $P < 0.001$; Supplementary Table 5 and Extended Data Fig. 6b). Comparison of microbial OTU prevalence between roots of *A. thaliana* and grasses revealed overall consistency at a continental scale (Spearman's rank correlation: bacteria: $r_s = 0.61$; $P < 0.01$; fungi: $r_s = 0.58$; $P < 0.01$; oomycetes: $r_s = 0.55$; $P < 0.01$; Extended Data Fig. 6c). Further inspection of the widespread *A. thaliana* OTUs in the roots of grasses indicated high conservation of the bacterial OTUs (13/13 detected; representing 45% of the total RA), but not of the fungal OTUs (6/8 detected; 16% of the total RA; Extended Data Fig. 6a,d). These bacterial OTUs were also consistently detected in the roots of *Lotus japonicus* (Fabaceae) grown in a different soil type (Cologne agricultural soil[29]; 11/13 detected; 16% of the total RA; Extended Data Fig. 6d). The results suggest that a small number of geographically widespread bacteria can efficiently colonize the roots of distantly related plant species and establish stable associations with plant roots over evolutionary time.

Spatial and temporal variation in root microbiota differentiation. Despite limited between-site variation in endosphere communities across European sites, we observed spatial and temporal variation in their composition (Fig. 3a and Extended Data Fig. 7a,b).

Site was the primary factor explaining variation in root-associated microbial community composition (Adonis: d.f. = 16; bacteria: $R^2 = 0.17$; $P < 0.001$; fungi: $R^2 = 0.19$; $P < 0.001$; oomycetes: $R^2 = 0.17$; $P < 0.001$), but this effect was markedly lower compared with that observed in the corresponding soil samples (Adonis: d.f. = 16; bacteria: $R^2 = 0.53$; $P < 0.001$; fungi: $R^2 = 0.51$; $P < 0.001$; oomycetes: $R^2 = 0.27$; $P < 0.001$; Fig. 3b and Supplementary Table 5). We collected data at each of the 17 sites for a set of 18 environmental variables and identified seven groups of non-collinear variables (Extended Data Fig. 8a,b). Across European sites, the main drivers of bacterial community composition in the soil, rhizosphere

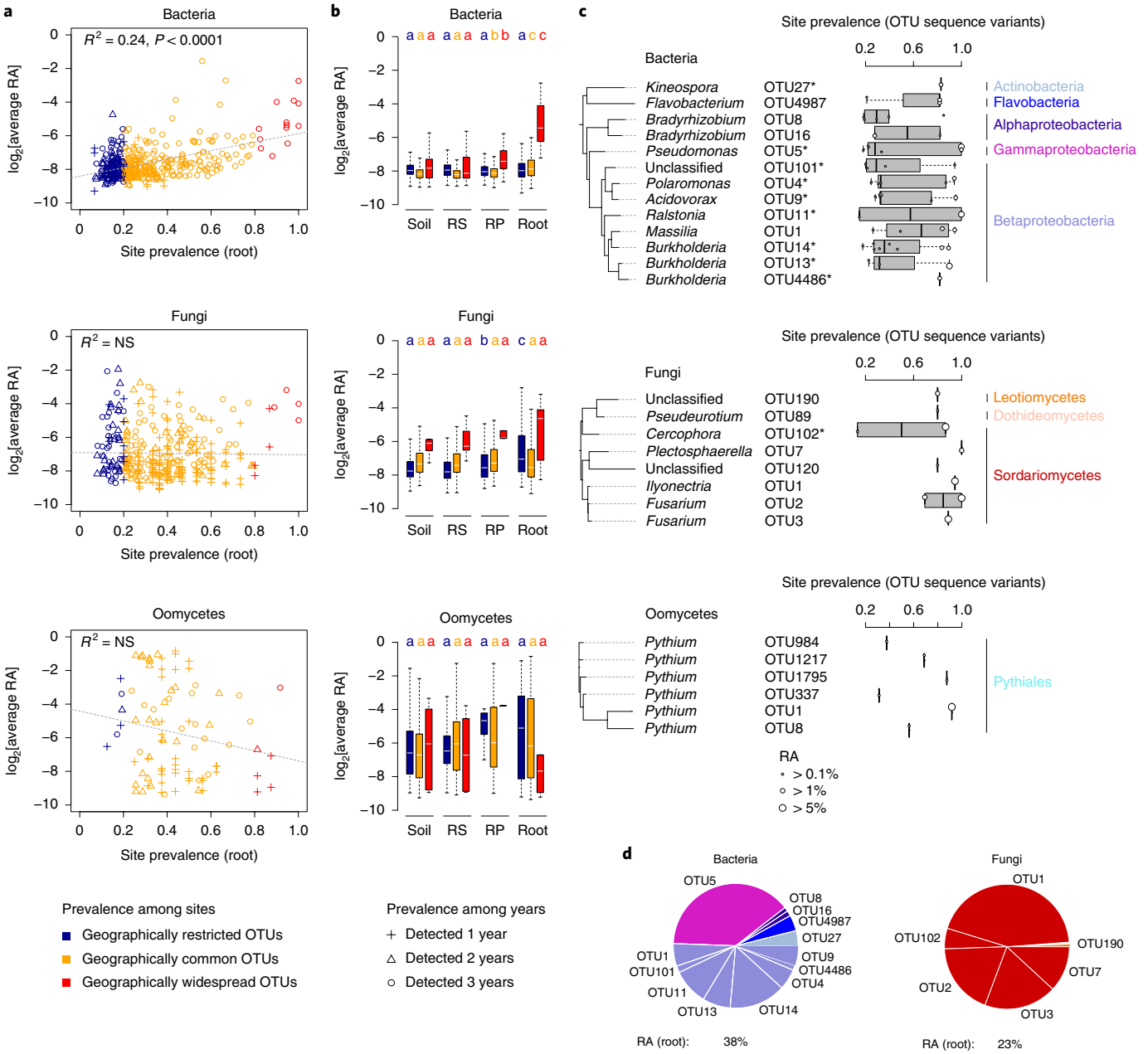


Fig. 2 | Geographically widespread taxa in the roots of *A. thaliana*. **a**, Correlation between OTU prevalence across sites (average of the 3 years) and OTU RA (log_2[average of the 3 years]) in plant roots, for bacteria (top), fungi (middle) and oomycetes (bottom). When calculating average RA, only samples for which the actual OTUs were present were considered (blue: geographically restricted OTUs (site prevalence: <20%); orange: geographically common OTUs (site prevalence: 20–80%); red: geographically widespread OTUs (site prevalence: >80%)). The different shapes highlight root-associated OTUs detected across 1, 2 or 3 years. RA and prevalence were averaged across the years for which one OTU was present. NS, not significant. **b**, Boxplots of the log_2[average RA] of geographically restricted, common and widespread OTUs detected in each of the four compartments, for bacteria (top), fungi (middle) and oomycetes (bottom). Letters depict significant differences across compartments (Kruskal–Wallis with Dunn's post-hoc test, $P < 0.01$). **c**, Phylogenetic trees, for bacteria (top), fungi (middle) and oomycetes (bottom), of geographically widespread root-associated OTUs (red symbols in **a**) constructed based on the 16S rRNA V5–V7 gene fragments (bacteria) and the ITS1 region (fungi and oomycetes). Asterisks indicate microbial OTUs significantly enriched in root compared with soil samples (generalized linear model, FDR < 0.05). The prevalence across sites (average of the 3 years) and average RA of sequence variants corresponding to each OTU are shown. In **b** and **c**, boxes represent the 1st and 3rd quartiles, the midline is the median and the whiskers extend to 1.5 of the interquartile range. **d**, Proportions of widespread bacterial (left) and fungal OTUs (right) detected in *A. thaliana* roots. OTUs with RA > 0.1% are shown. The overall RA of these OTUs in root samples, measured across sites and years, is indicated below each pie chart.

and rhizoplane were pH and/or environmental variables strongly correlated with pH, such as iron, mean air temperature and mean relative humidity (Adonis: d.f. = 1; soil: $R^2 = 0.35$; $P < 0.001$; rhizosphere: $R^2 = 0.32$; $P < 0.001$; rhizoplane: $R^2 = 0.25$; $P < 0.001$; Fig. 3c and Supplementary Table 6). Stronger correlations with environmental variables were observed for bacterial communities than for filamentous eukaryotes (Adonis: d.f. = 1; fungi: $R^2 < 0.17$; $P < 0.001$; oomycetes: $R^2 < 0.11$; $P < 0.001$; Fig. 3c and Supplementary Table 6), consistent with the prominent role of pH for spatial structuring of soil bacterial communities[10–12]. Notably, the explanatory power of

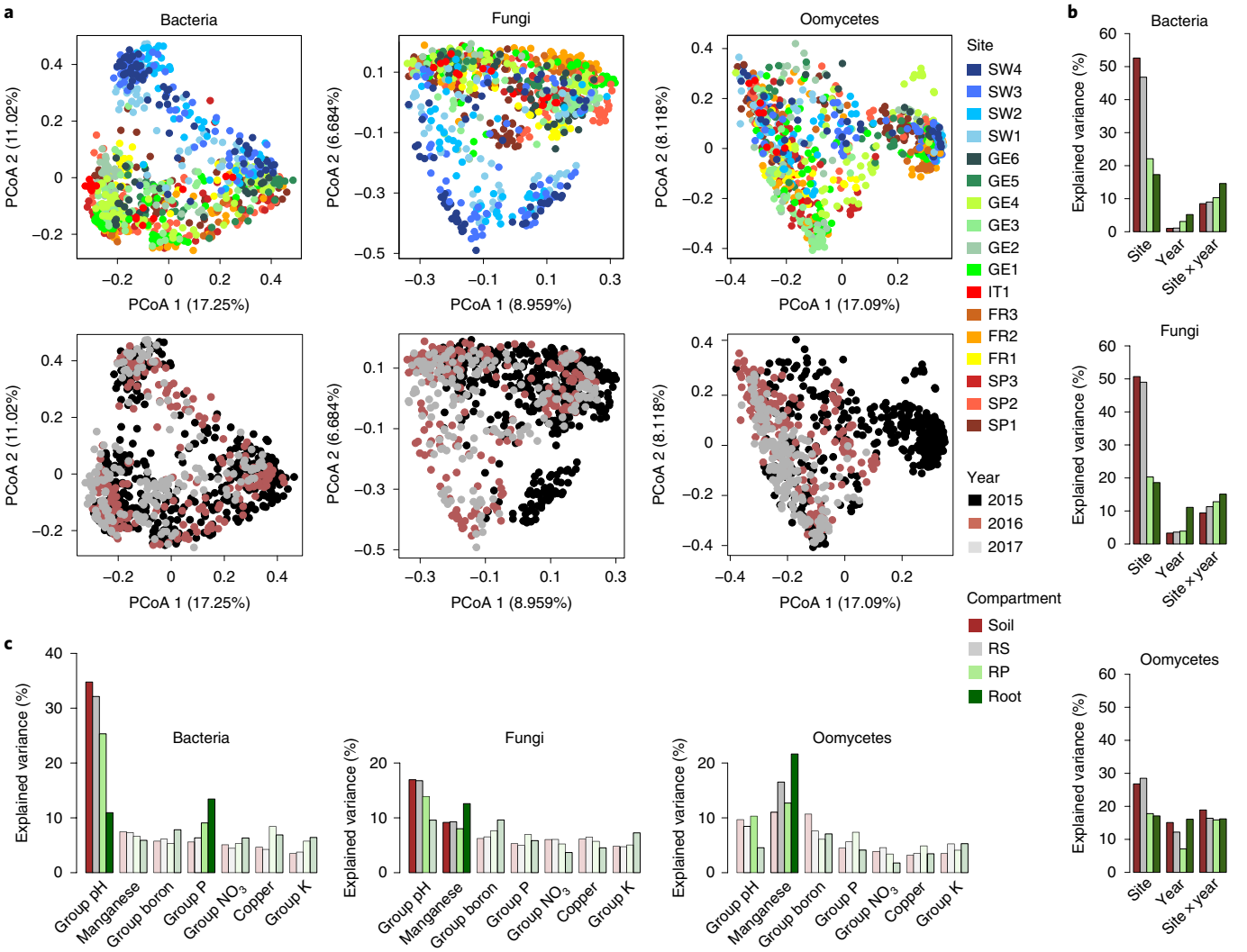


Fig. 3 | Factors shaping the *A. thaliana* root microbiota at the continental scale. a, PCoA based on Bray–Curtis distances between samples harvested across the 17 sites, four compartments and three successive years (2015, 2016 and 2017). Microbial communities are presented for the whole *A. thaliana* dataset for bacteria (top; $n = 881$ samples), fungi (middle; $n = 893$ samples) and oomycetes (bottom; $n = 875$) and are colour coded according to either the site or the harvesting year. OTUs with RA < 0.1% were excluded from the dataset. **b**, Effect of site, harvesting year and site × harvesting year on bacterial, fungal and oomycetal community composition. **c**, Effect of environmental variables or groups of collinear variables (see Extended Data Fig. 8) measured at each site on bacterial (left), fungal (middle) and oomycetal community composition (right). Group pH includes pH, iron, latitude, longitude, mean air temperature and mean relative humidity. Group boron includes the factors boron, available Ca and reserve Ca. Group P includes the factors reserve P and available P. Group NO_3 includes the factors available NO_3, reserve Mg and available Mg. Group K includes the factors available K and reserve K. Significant effects are highlighted with dark colours. In **b** and **c**, the explained variance for each factor is shown for the different compartments, and the percentage of explained variance for each parameter was calculated based on PERMANOVA ($P < 0.001$ (**b**) and $P < 0.01$ (**c**); Supplementary Tables 5 (**b**) and 6 (**c**)).

group pH was reduced in root samples, where variation in microbial community composition was primarily explained by the collinear soil variables reserve P and available P for bacteria (d.f. = 1; $R^2 = 0.13$; $P = 0.007$), and by manganese for fungi ($R^2 = 0.13$; $P = 0.015$) and oomycetes ($R^2 = 0.21$; $P = 0.017$) (Fig. 3c and Supplementary Table 6). This is consistent with recent observations indicating that the effect of soil mineral nutrients on root microbiota composition is partly driven by host responses[30,31].

The factor year explained more variation in root-associated microbial communities (Adonis: d.f. = 2; bacteria: $R^2 = 0.05$; $P < 0.001$; fungi: $R^2 = 0.11$; $P < 0.001$; oomycetes: $R^2 = 0.16$; $P < 0.001$) than did host species separated by >140 million years of reproductive isolation[32] (Adonis: d.f. = 1; $R^2 < 0.01$; $P < 0.001$). Notably, among-year variation was particularly strong for fungal and oomycetal communities (Fig. 3b, Extended Data Fig. 7b and Supplementary Table 5). Our results suggest that among-site variation in environmental conditions affects root-associated microbial communities more strongly than does temporal variation, and that differences in community composition between roots of *A. thaliana* and co-occurring grasses are small.

Differences in soil, climate and genotype drive root microbiota differentiation between *A. thaliana* populations. We hypothesized that continental-scale variation in the *A. thaliana* root microbiota is mediated by interactions between soil, climate and host genotype, and that these factors differentially influence bacteria and filamentous eukaryotes. To disentangle the contribution of these factors to microbial community composition, we conducted a transplant experiment between two geographically separated *A. thaliana* populations in Sweden and Italy (SW4 and IT1; Extended Data

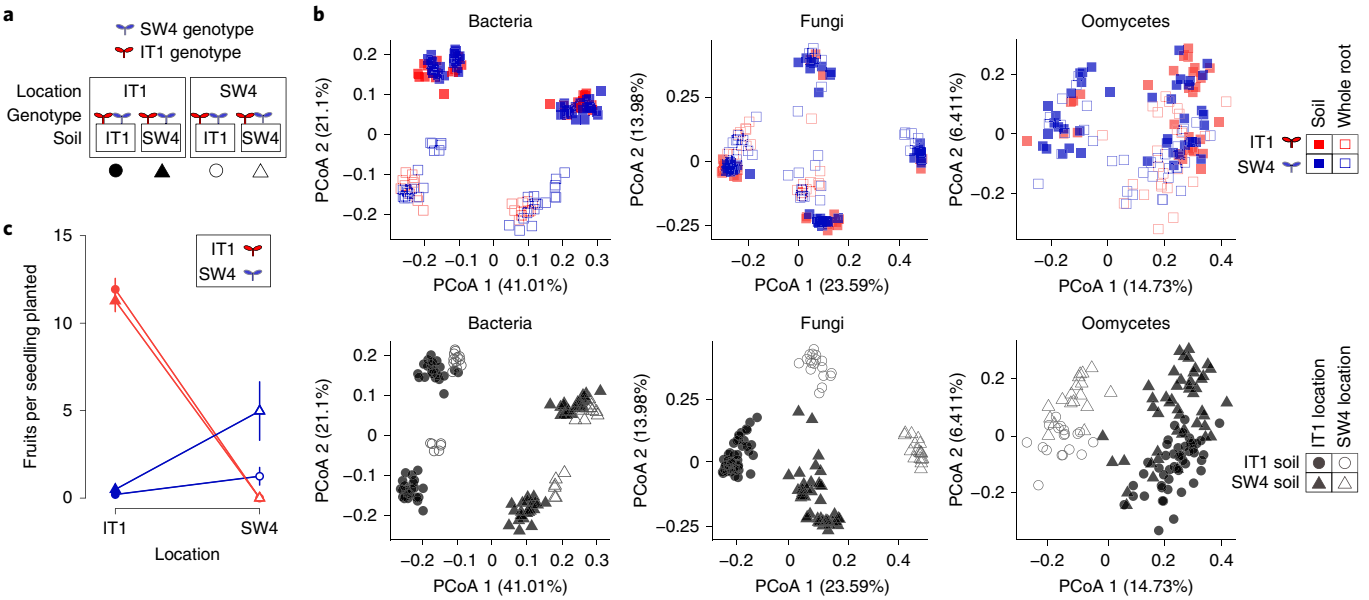


Fig. 4 | Reciprocal transplant between two *A. thaliana* populations in Sweden and Italy. a, Schematic of the reciprocal transplant experiment. Soils and plant genotypes from IT1 and SW4 sites were reciprocally transplanted in the two locations (eight different treatment combinations). The symbols below the schematic correspond to the symbols used in **b** and **c**. **b**, Community structures of bacteria (left), fungi (middle) and oomycetes (right) in the 131 samples were determined by PCoA. The first two dimensions of the PCoA are plotted based on Bray–Curtis distances. To facilitate visualization, the same PCoA plots are shown according to either the genotype and compartment (top) or the soil origin and location (bottom). Note that no Italian plant survived at the Swedish site. **c**, Overall fitness (number of fruits per seedling planted; mean ± s.e.) of Italian (red) and Swedish genotypes (blue) when reciprocally planted in Italian (circles) and Swedish soils (triangles) and grown at Italian (filled symbols) and Swedish locations (open symbols).

Fig. 9a,b). We transplanted seedlings of each plant genotype into soils from both SW4 and IT1 at each site in autumn, at the time of natural seedling establishment in the local populations (Fig. 4a). At fruit maturation, the two *A. thaliana* genotypes were harvested and community composition was defined for bacteria, fungi and oomycetes in the soil and whole-root (root- with rhizoplane-associated microbes) samples ($n = 131$; Fig. 4b).

Consistent with the European transect survey, we observed a larger compartment effect for bacteria than for filamentous eukaryotes (Adonis: d.f. = 1; bacteria: $R^2 = 0.22$; $P < 0.001$; fungi: $R^2 = 0.09$; $P < 0.001$; oomycetes: $R^2 = 0.04$; $P < 0.001$) (Fig. 4b (top panel) and Supplementary Table 7). Because no Italian plants survived at the Swedish site (Extended Data Fig. 9c), we used a canonical analysis of principal coordinates to assess the genotype effect among whole-root samples at the Italian site only (capscale function in R). Canonical analysis of principal coordinates, constrained by genotype, indicated a small but significant effect of the host genotype, which explained 2.4% (d.f. = 1; $P = 0.001$), 2.4% (d.f. = 1; $P = 0.026$) and 2.9% (d.f. = 1; $P = 0.002$) of the variance in bacterial, fungal and oomycetal community composition, respectively (Extended Data Fig. 9d).

PERMANOVA, conducted separately by soil and whole-root compartment, indicated that the origin of soil (Italy versus Sweden) explained substantially more variation in bacterial soil biota than did transplant location (Adonis: d.f. = 1; soil origin: $R^2 = 0.62$; $P < 0.001$; transplant location: d.f. = 1, $R^2 = 0.16$; $P < 0.001$; Supplementary Table 7). This was also the case for bacterial communities associated with whole-root samples (Adonis: d.f. = 1; soil origin: $R^2 = 0.47$; $P < 0.001$; transplant location: d.f. = 1; $R^2 = 0.17$; $P < 0.001$) (Fig. 4b (bottom panel) and Supplementary Table 7). Compared with bacteria, the percentage of variance explained by origin of soil was weaker for fungi and oomycetes in the soil fraction (Adonis: d.f. = 1; fungi, $R^2 = 0.30$; $P < 0.001$; oomycetes, $R^2 = 0.06$; $P < 0.001$). Instead, location was largely accounting for between-site variation in microbial eukaryotic assemblages in soil (Adonis: d.f. = 1; fungi, $R^2 = 0.27$; $P < 0.001$; oomycetes, $R^2 = 0.19$; $P < 0.001$). In whole-root samples, the effect of location was as strong as the effect of soil for fungi (d.f. = 1; location, $R^2 = 0.15$; $P < 0.001$; soil origin, $R^2 = 0.15$; $P < 0.001$) and stronger than the effect of the soil for oomycetes (d.f. = 1; location, $R^2 = 0.11$; $P < 0.001$; soil origin, $R^2 = 0.05$; $P < 0.001$) (Fig. 4b (lower panel) and Supplementary Table 7). These results were validated at ASV resolution (Supplementary Table 7).

Inspection of the abundance profiles of microbial taxa in whole-root samples showed stable prevalence of geographically widespread bacterial OTUs which are detectable across different soil × location × genotype combinations (Extended Data Fig. 10a–c), and pointed to the differential impact of site- and soil-specific factors on different taxonomic groups of the root microbiota (Extended Data Fig. 10d). For example, variation in root-associated Actinobacteria was almost exclusively explained by soil origin, whereas variation in root-associated Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria was explained by soil origin and location to similar degrees. Similarly, detection of root-associated fungi belonging to Dothideomycetes was location dependent, whereas detection of those belonging to Leotiomycetes was largely explained by soil origin. Overall, 74.4 and 86.3% of root-associated fungal and oomycetal OTUs, respectively, responded to a difference in location or location × soil, whereas this percentage dropped to 44.7% for root-associated bacterial OTUs (Extended Data Fig. 10d). Taken together, our data suggest that location-specific factors such as climatic conditions affect the differentiation of root-associated filamentous eukaryotic communities more than that of bacteria.

Differences in soil conditions weakly contribute to adaptive differentiation between *A. thaliana* populations. If adaptation to local soil conditions contributes to adaptive differentiation between *A. thaliana* populations, the advantage of the local over non-local genotype should be greater when plants are grown on local soil. To

test this hypothesis, we scored plant survival and fecundity (number of fruits per reproducing plant) and estimated overall fitness (number of fruits per seedling planted) of the plants grown in the reciprocal transplant experiments at the sites IT1 and SW4 ($n = 1{,}008$; Supplementary Table 8).

At the IT1 site, the relative fitness of the local genotype was higher on local than on non-local soil: the mean fitness of the Italian genotype was 58 times higher than that of the non-local Swedish genotype on Italian soil (corresponding to a selection coefficient of $s = 0.98$), and 23 times higher ($s = 0.96$) on Swedish soil (significant soil × genotype interaction; mixed-model analysis of variance: $F_{1,10} = 5.7$; $P = 0.038$) (Fig. 4c and Supplementary Table 9). The strong local advantage was a function of both higher survival and fecundity (Extended Data Fig. 9c). At the SW4 site, no Italian plant survived to reproduce (Fig. 4c); it was therefore not meaningful to compare adaptive differentiation on the two soil types. However, the Swedish genotype showed 3.5-fold higher survival and 4.0-fold higher overall fitness when planted in Swedish soil compared with in Italian soil (one-tailed test; survival: $t = 1.81$; $P = 0.060$; overall fitness: $t = 2.14$; $P = 0.039$) (Fig. 4c and Extended Data Fig. 9c). These results showed a strong selective advantage for the local *A. thaliana* genotype at both sites, consistent with previous studies on the same populations[15,33]. However, despite the marked differences in the geochemical properties and microbial communities of the soil, adaptation to local soil conditions explained only a small fraction of the adaptive differences between the two *A. thaliana* populations. This suggests that adaptation to climatic factors, and possibly other environmental variables not determined by soil characteristics, is the primary driver of adaptive differentiation between *A. thaliana* populations in northern and southern Europe.

Discussion

Our results indicate that environmental variables that explain the geographic distribution of microbes in soil have less predictive power in plant roots. The reduced between-site variation observed for root-associated bacterial communities is remarkable given the large geographical distances and contrasting environments between sites. The 13 ubiquitous root-associated bacterial OTUs comprise very few abundant sequence variants, implying that a small subset of conserved taxa have evolved mechanisms to dominate the bacterial root microbiota at a continental scale, irrespective of major differences in the surrounding bacterial soil biota. These taxa are often detected in the roots of divergent plant species across various environments[34–38], and probably explain the similarity in bacterial community composition observed across microbiome studies[39]. *Bradhyrhizobium* and *Burkholderia* (OTU8, OTU4987, OTU13, OTU14 and OTU4486 in this study) were the two most dominant of the 47 widespread genera in the roots of 31 diverse plant species across a 10-km transect in Australia[37]. The remarkable phylogenetic diversity among these taxa also suggests convergent evolution and metabolic adaptation to the root habitat in phylogenetically distant bacterial lineages[40].

A *Pseudomonas* OTU was the most dominant and prevalent in the roots of *A. thaliana* and grasses (OTU5; RA = 14.4%), pointing to *Pseudomonas* taxa as robust root colonizers. Inspection of sequence heterogeneity within OTU5 identified two variants (zero-radius OTU1 (zOTU1; RA = 11%; prevalence = 1; related to *Pseudomonas chlororaphis*) and zOTU4 (RA = 3.5%; prevalence = 1; related to *Pseudomonas simiae*)) that are phylogenetically distant from *Pseudomonas viridiflava*, the dominant taxon in the *A. thaliana* phyllosphere[41]. Although some of the widespread bacteria might represent potential pathogens[41], many are probably carrying important beneficial functions for *A. thaliana* survival, including pathogen protection[3,40,42] or plant growth promotion[43,44]. Whether this widespread association between plant roots and bacteria is evolutionarily ancient, and the extent to which it has contributed to plant adaptation to terrestrial ecosystems, remains to be seen. Irrespective of this, the 13 OTUs identified here in *A. thaliana* roots provide a framework for the future design of low-complexity synthetic bacterial communities from culture collections of root commensals, and co-culturing with gnotobiotic plants[45] to study their contribution to plant health in laboratory environments.

The results of the reciprocal transplant between two widely separated *A. thaliana* populations showed that bacterial and fungal assemblages are differentially controlled by edaphic and climatic conditions at the root interface. Our observation that transplant location affected the community composition of filamentous eukaryotes more strongly than that of bacteria suggests that the response to climate varies among microbial kingdoms[46,47]. Our data corroborate the hypothesis that climate is a key driver of among-site variation and geographic distribution of filamentous eukaryotes in soil[13,45,48,49], contrasting with the geographic distribution of soil bacteria, which is primarily controlled by edaphic factors[10]. Our data also suggest that the host genotype contributes only little to root microbiota differentiation (~2% at OTU resolution), consistent with previous studies[50–52].

The extent to which local adaptation in plants can be attributed to local soil conditions has been examined mainly in relation to high concentrations of heavy metals[53], serpentine[54] and high salinity[25,55]. We observed strong local adaptation between the two selected *A. thaliana* populations. However, the relative performance of local and non-local host genotypes was only weakly affected by soil origin, despite extensive differences in microbial community composition and physiochemical properties between the Swedish and Italian soils. Particularly, differences in available Ca, reserve K, available Mg, pH, iron and available K between the two soils are among the largest observed across all 17 soils (Supplementary Table 1). Previous work has shown that genetic differences in tolerance to freezing, and phenological traits such as the timing of germination and flowering, can explain a substantial proportion of selection against the non-local genotype at the two sites[15,56–59]. Taken together, our data indicate that differences in climate have been more important than differences in soil and indigenous microbiota for the adaptive divergence between the two study populations. Future studies should determine whether the environmental factors that affect root microbiota assembly and adaptive differentiation at large spatial scales are different from those acting at smaller geographical scales.

Methods

Harvesting of *A. thaliana* and grasses across 17 European sites. We selected 17 sites with variable soil characteristics across a climatic gradient from Sweden to Spain. At each site, *A. thaliana* and grasses occur naturally[1,3,15,60,61] (Fig. 1a and Supplementary Table 1). We harvested *A. thaliana* from February to May at the same developmental stage (bolting/flowering stage) for one (three sites), two (one site) and three (13 sites) consecutive years (Supplementary Table 1). Plants were harvested with their surrounding bulk soil with a hand shovel without disturbing the plant root system, transferred to 7 cm × 7 cm greenhouse pots and transported immediately to nearby laboratories for further processing. Four plant individuals pooled together were considered as one pooled-plant technical replicate (four technical replicates in total per site). In addition, during the first harvesting year, four plants per site were also kept individually, but negligible within-site variation in community composition was observed (data not shown). At each of the 17 sites, we also harvested and processed similarly three neighbouring grass plants growing within 50 cm of *A. thaliana*. In total, we collected 285 plants across 3 years.

Fractionation of soil and root samples. To distinguish and separate four microbial niches across the soil–root continuum, plants and respective roots were taken out of the pot. Samples from the bulk soil exempt from root or plant debris were taken, snap-frozen in liquid nitrogen and stored for further processing (soil compartment). Individual plants were manually separated from the main soil body, and non-tightly adhered soil particles were removed by gently shaking the roots. These roots were separated from the shoot and placed in a 15-ml falcon containing 10 ml deionized sterile water. After ten inversions, the roots were transferred to another falcon and further processed, while leftover wash-off (containing the rhizosphere fraction) was centrifuged at 4,000*g* for 10 min. The supernatant was discarded and the pellet was resuspended and transferred to a

new 2-ml screw-cap tube. This tube was centrifuged at 20,000 r.p.m. for 10 min, the supernatant was discarded and the pellet was snap-frozen in liquid nitrogen and stored for further processing (rhizosphere compartment). After rhizosphere removal, cleaned roots were placed in a 15-ml falcon with 6 ml detergent (1× TE buffer + 0.1% Triton X-100) and manually shaken for 2 min. This step was repeated for a total of three detergent washes, in between which the roots were transferred to a new 15-ml falcon with new detergent. After these washes in detergent, the roots were transferred to a new 15-ml falcon. The remaining washes (approximately 18 ml) were transferred to a 20-ml syringe and filtered through a 0.22-µM-pore membrane. The membrane was snap-frozen in liquid nitrogen until further processing (rhizoplane compartment). Lastly, roots that had been washed three times were subjected to a further surface sterilization step to further deplete the leftover microbes from the root surface. Roots were sterilized for 1 min in 70% ethanol, followed by 1 min in 3% NaClO, and rinsed five times with deionized sterile water. These roots were dried using sterile Whatman paper and snap-frozen in liquid nitrogen until further processing (root compartment) (Extended Data Fig. 1a).

In total, 1,125 samples were produced after fractionation. We validated the fractionation protocol by printing processed roots on 50% Tryptic Soy Agar (TSA) before fractionation, after each detergent wash and after surface sterilization. Washes produced after each fractionation step (without treatment, after each detergent step and after surface sterilization) were also plated on 50% TSA. In brief, treated roots were cut and placed on plates containing 50% TSA medium for 30 s and then removed. After 3 d at 25 °C, colony-forming units were counted. Similarly, 20 µl of remaining washes were spread onto 50% TSA medium-containing plates and colony-forming units were counted after 3 d of incubation at 25 °C (Extended Data Fig. 1b).

Reciprocal transplant experiment. We reciprocally transplanted both soil and local *A. thaliana* genotypes between sites in Italy (population IT4) and Sweden (population SW4) (Extended Data Fig. 9a). See ref. [15] for a description of the study sites and plant genotypes. Soil was collected at the two experimental sites in spring 2016 and stored at 6 °C until establishment of the experiments. Seeds were planted in Petri dishes on agar, cold stratified in the dark at 6 °C for 1 week, and then moved to a growth room (16-h days at 22 °C, 150 µE m^{-2} s^{-1} photosynthetically active radiation (PAR), and 8 h at 16 °C in dark) where the seeds germinated. Nine days after germination, seedlings were transplanted to 299-cell plug trays (cell size: 20 mm × 20 mm × 40 mm) with local and non-local soils in blocks of 6 × 7 cells in a checkerboard design (three blocks of each soil type per tray). At each site, we transplanted 20 local and 22 non-local seedlings into randomized positions in each of six blocks per soil type, with a total of 120 local and 132 non-local plants for each site × soil combination. During transplantation, plug trays were kept in a greenhouse at ~18 and ~12 °C for 16 h of day and 8 h of night, respectively. Within 6 d, the trays were transported to the field sites where they were sunk into the ground (on 9 September 2016 in Sweden and on 29 October 2016 in Italy). The transplanted seedlings were at the same stage of development as naturally germinating plants in the source population.

At fruit maturation in spring 2017, we scored survival to reproduction and the number of fruits per reproducing plant (an estimate of fecundity), and estimated total fitness as the number of fruits produced per seedling planted (following ref. [15]). Statistical analyses were based on block means for each genotype. No Italian plant survived at the Swedish site, and it was therefore not possible to fit a model examining the site × soil × genotype interaction. Instead, we conducted analyses separately by site. For the Italian site, we assessed differences in overall fitness due to soil, genotype and the soil × genotype association using a mixed-model analysis of variance, with block (random factor) nested within soil type. In Sweden, we used a one-tailed *t*-test to test the prediction that survival and overall fitness of the Swedish genotype are higher when planted in Swedish compared with Italian soil.

At the time of fitness assessment, we harvested plants and their surrounding soil by removing the whole soil plug. Soil was separated from the roots manually and a soil sample was taken. After removing loose soil particles by gentle shaking, roots were placed in a 15-ml falcon, washed by inverting with 10 ml deionized water, and surface sterilized by washing for 1 min with 70% ethanol, then washing with 3% NaClO for 1 min and rinsing several times with deionized sterile water. These whole-root samples correspond to the combination of root and rhizoplane compartments described for the European transect survey. Between six and 12 soil samples were harvested for each of the eight combinations of soil, location and genotype ($n = 72$), as well as three to 12 whole-root samples per condition ($n = 59$). Note that no Italian plants survived at the Swedish site. In total, 131 root and soil samples were harvested, stored in dry ice, and processed for DNA isolation and microbial community profiling.

Microbial community profiling. Total DNA was extracted from the aforementioned samples using the FastDNA SPIN Kit for Soil (MP Biomedicals). Samples were homogenized in Lysing Matrix E tubes using the Precellys 24 tissue lyzer (Bertin Technologies) at 6,200 r.p.m. for 30 s. DNA samples were eluted in 60 µl nuclease-free water and used for bacterial, fungal and oomycetal community profiling[1,3]. The concentration of DNA samples was fluorescently quantified, diluted to 3.5 ng $µl^{-1}$ and used in a two-step PCR amplification protocol. In the first step, V5–V7 of bacterial 16S rRNA (799F–1192R), V3–V4 of bacterial 16S rRNA (341F–806R), fungal ITS1 (ITS1F–ITS2), fungal ITS2 (fITS7–ITS4) and oomycetal ITS1 (ITS1-O–5.8 s-Rev-O) (Supplementary Table 2) were amplified. Under a sterile hood, each sample was amplified in triplicate in a 25-µl reaction volume containing 2 U DFS-Taq DNA polymerase, 1× incomplete buffer (both Bioron), 2 mM $MgCl_2$, 0.3% bovine serum albumin, 0.2 mM dNTPs (Life Technologies) and 0.3 µM forward and reverse primers. PCR was performed using the same parameters for all primer pairs (94 °C for 2 min, 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s and 72 °C for 10 min for 25 cycles). Afterwards, single-stranded DNA and proteins were digested by adding 1 µl of Antarctic phosphatase, 1 µl Exonuclease I and 2.44 µl Antarctic Phosphatase buffer (New England Biolabs) to 20 µl of the pooled PCR product. Samples were incubated at 37 °C for 30 min and enzymes were deactivated at 85 °C for 15 min. Samples were centrifuged for 10 min at 4,000 r.p.m., and 3 µl of this reaction was used for a second PCR. This was prepared in the same way as described above, using the same protocol, but with the number of cycles reduced to ten and with primers including barcodes and Illumina adaptors (Supplementary Table 2). The PCR quality was controlled by loading 5 µl of each reaction on a 1% agarose gel. Afterwards, the replicated reactions were combined and purified. First, bacterial amplicons were loaded on a 1.5% agarose gel and run for 2 h at 80 V; bands with the correct size of ~500 base pairs were cut out and purified using the QIAquick Gel Extraction Kit (Qiagen). Second, fungal and oomycetal amplicons were purified using Agencourt AMPure XP beads. The DNA concentration was again fluorescently determined, and 30 ng DNA of each of the barcoded amplicons was pooled in one library per microbial group. Each library was then purified and re-concentrated twice with Agencourt AMPure XP beads, and 100 ng of each library was pooled together. Paired-end Illumina sequencing was performed in-house using the MiSeq sequencer and custom sequencing primers (Supplementary Table 2).

16S rRNA gene and ITS read processing. All paired rRNA amplicon sequencing reads were analysed with a pipeline described recently[3]. The main parts include scripts from QIIME[62] and usearch[63]. OTUs were clustered at a 97% threshold (usearch–cluster otus). Bacterial reads were checked against the Greengenes database[64] to remove non-bacterial reads. Fungal and oomycetal reads were checked against an ITS sequences database (full-length ITS sequences from the National Center for Biotechnology Information (NCBI)) to remove unwanted reads. Taxonomic assignment was done via QIIME (assign_taxonomy) for bacterial OTUs, using the Greengenes database. Taxonomic assignment for fungal OTUs was done via RDP-classifier[65] against the Warcup database[66]. Assignment of oomycetal OTUs was also done via RDP-classifier but against a self-established database. For RDP-classifier based classification, a cut-off of 0.5 was used to filter out uncertain assignments. ASVs were computed using the UNOISE2 pipeline[26]. Quality-filtered reads were truncated to an equal length and abundance sorted. ASVs/zOTUs were calculated using the usearch --unoise3 command. ASVs/zOTUs were then further processed using the same steps as for the OTU data. To access intra-OTU diversity of widespread OTUs, all reads that were mapped to one particular OTU (using the results from the usearch --usearch_global) were used as an input for ASV/zOTU calculation (usearch --unoise3). ASVs/zOTUs were calculated for each widespread OTU individually

OTU tables and alpha- and beta-diversity metrics. All OTU tables obtained from the pipelines above were filtered for very low-abundance OTUs before any further analysis. For this, only OTUs with at least 0.1% RA in at least one sample were kept. These filtered tables were used for all further analyses. For OTU-based analysis of the bacterial data, OTUs assigned as chloroplast or mitochondria derived were removed before analysis. Alpha-diversity indices (Shannon index and observed OTUs) were calculated using OTU tables rarefied to 1,000 reads. Bray–Curtis distances between samples were calculated using OTU tables that were normalized using cumulative-sum scaling (CSS[67]). All distance-based analyses were performed using the respective Bray–Curtis distance matrices. PCoAs were done using cmdscale (R), and constrained analysis using capscale (vegan package in R).

Dendrogram and average distances. Bray–Curtis distances across replicates and years were calculated using the mean across all combinations between two sets of sites (for example, between all soil samples from site *x* and site *y*). These average distances were hierarchically clustered (hclust in R; method = 'complete') (see Fig. 1b). Similarly, averaged RAs of taxonomic groups were achieved by averaging across all samples from a particular site–compartment combination. These mean RA values were used to estimate the differential abundance of taxonomic groups across compartments and geographic locations (Wilcoxon rank-sum test; wilcox.test in R; FDR < 0.05) (Fig. 1c and Extended Data Fig. 2c,d).

Variance partitioning of microbial community composition. To test for the effects of site, year of harvesting, compartment and host species (*A. thaliana* versus other species) on the composition of the microbiota in samples collected across Europe, we analysed the Bray–Curtis distance matrix between pairs of samples with a permutation-based test using a PERMANOVA model with the Adonis function (R package vegan; see Supplementary Tables 3 and 5). The effect of environmental parameters on the microbiota composition at the European scale

was tested with the same approach (Supplementary Tables 6). To avoid pseudo-replication, and because environmental variables were measured only once per site, mean Bray–Curtis distances per pair of sites across years were used in this analysis. Collinearity between environmental variables was assessed with the Spearman's rank correlation (cor.test function in R), and correlated variables (Spearman's rank correlation, $r < -0.7$ and $r > 0.7$) were grouped into explanatory groups in the model (Fig. 3c and Extended Data Fig. 8). Within each group, only the best explanatory variable (that is, the variable with the highest R^2 in univariate analyses) was used in the model. Except for the general statistical model (Supplementary Table 3), PERMANOVA tests were computed separately for the different fractions. In the transplant experiment, the effects of soil, location and genotype were assessed with PERMANOVA separately on soil and whole-root samples (see Supplementary Table 7 for the models used), whereas the effect of compartment was assessed with a model that included both whole-root and soil samples. Unless noted differently, $P < 0.01$ was considered significant.

Calculation of OTU prevalence and abundance. To calculate the site prevalence of OTUs, OTU tables were restricted to samples with >1,000 reads. Count tables were transformed to RA by division of total column (sample) sums. An OTU was marked as present in a given sample if its RA was >0.1%. OTUs present in fewer than five samples were not considered. Prevalence is the proportion of sites sampled in a given year where an OTU is present (in at least one sample with RA > 0.1%). Prevalence was calculated separately for each year and then averaged across years. This procedure was carried out individually for each compartment. Site prevalence for OTUs present in co-occurring grass species was calculated in the same way. To calculate the average RA across sites and years for a given OTU, only those samples for which the OTU was actually present were considered. In this way, prevalence and mean RA are independent. OTUs present in >80% and <20% of the sites on average were considered as geographically widespread and geographically restricted, respectively (Fig. 2a and Extended Data Fig. 5). Abundances of OTUs in the three groups (geographically restricted, common and widespread) were compared between compartments using the Kruskal–Wallis test (kruskal.test with dunn.test in R; FDR < 0.05; Fig. 2b). Root-specific enrichment for widespread OTUs, compared with soil and rhizosphere samples, was tested with a generalized linear model, as described in ref. [22] (FDR < 0.05; Fig. 2c). To compare generalist OTUs from this study with OTUs found in the roots of *L. japonicus*[29], bacterial OTUs were directly compared using the representative sequences (BLASTn; 98% sequence identity). To compare fungal OTUs, the representative sequences from geographically widespread OTUs were blasted against full ITS sequences from the UNITE database[68]. The best hits received were then used to find similar OTUs in the *L. japonicus* data (98% sequence identity) (Extended Data Fig. 6d).

OTU enrichment signatures in the transplant experiment. Enrichment patterns of root-associated OTUs in the transplant experiment (Extended Data Fig. 10) were examined as follows. For bacterial and fungal datasets, samples with <1,000 reads were removed, whereas for the oomycetal dataset, samples with <250 reads were removed. OTUs with a mean RA > 0.01% across all root samples were kept for the analysis. RA entries of zero were replaced by 0.001%. All RA values were $\log_2$-transformed, and these data were used as input for generating a heatmap (heatmap.2 function in R; gplots library). Enrichment in one of the six tested conditions (Swedish location, Italian location, Swedish soil, Italian soil, Italian location and soil, and Swedish location and soil) was estimated by comparing the mean RA of each OTU across all samples with the mean RA in the respective sample combinations. For example, to be enriched at the Italian location, an OTU must be more abundant at the Italian location compared with the Swedish location, irrespective of the soil (see the caption of Extended Data Fig. 10). OTUs that were not enriched in any of the conditions but were prevalent across samples (mean RA > 0.1%) are marked in grey in the heatmap. OTUs that were not enriched in any condition are not shown. Geographically widespread OTUs were identified by comparing representative sequences from the transect data with those of the reciprocal transplant experiment (BLAST at 98% sequence identity).

Taxonomic assignment of grass species. We amplified the plant *rbcL* locus of DNA isolated from the roots of plants that co-occurred with *A. thaliana* at each of the 17 sites by PCR (94 °C for 40 s, then 94 °C for 30 s, 50 °C for 30 s and 72 °C for 60 s for 35 cycles, and finally 72 °C for 3 min) using the primers 1F (ATGTCACCACAAACAGAAAC) and 724R (TCGCATGTACCTGCAGTAGC)[28]. The PCR quality was controlled by loading 2 µl of each reaction on a 1% agarose gel. PCR products were purified using Agencourt AMPure XP beads, and the DNA concentration was quantified using NanoDrop. Amplification of the *rbcL* locus followed by Sanger sequencing was successful for 43 out of 61 samples. Amplification (94 °C for 40 s, then 94 °C for 30 s, 49 °C for 30 s and 72 °C for 60 s for 35 cycles, and finally 72 °C for 3 min) and sequencing of the *MatK* locus (3F_KIMf (CGTACAGTACTTTTGTGTTTACGAG) and 1R_KIMr (ACCCAGTCCATCTGGAAATCTTGGTTC)) could be determined for 11 of the 18 remaining samples, and *rbcL* sequences of corresponding species were retrieved at NCBI. Taxonomic assignment in different grass species was achieved by BLAST against the non-redundant database at NCBI. Four samples could not be assigned into Poacae and were excluded from all analyses. *RbcL* sequences were then aligned with MUSCLE using MEGA7 and trimmed at the same length (that is, 653 base pairs), and a neighbour-joining tree was inferred using evolutionary distances computed using the maximum-likelihood method (Extended Data Fig. 6a).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequencing reads of samples from the European transect experiment and reciprocal transplant experiment (MiSeq 16S rRNA and ITS reads) have been deposited in the European Nucleotide Archive under accession numbers ERP115101 and ERP115102, respectively.

Code availability

All scripts for computational analysis and the corresponding raw data are available at https://github.com/ththi/European-Root-Suppl.

Received: 16 May 2019; Accepted: 8 November 2019;
Published online: 23 December 2019

References

1. Agler, M. T. et al. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS Biol.* **14**, e1002352 (2016).
2. Hassani, M. A., Durán, P. & Hacquard, S. Microbial interactions within the plant holobiont. *Microbiome* **6**, 58 (2018).
3. Durán, P. et al. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. *Cell* **175**, 973–983.e14 (2018).
4. Berendsen, R. L., Pieterse, C. M. J. & Bakker, P. A. H. M. The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17**, 478–486 (2012).
5. Berendsen, R. L. et al. Disease-induced assemblage of a plant-beneficial bacterial consortium. *ISME J.* **12**, 1496–1507 (2018).
6. Lebeis, S. L. et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349**, 860–864 (2015).
7. Haichar, F. et al. Plant host habitat and root exudates shape soil bacterial community structure. *ISME J.* **2**, 1221–1230 (2008).
8. Stringlis, I. A. et al. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proc. Natl Acad. Sci. USA* **115**, E5213–E5222 (2018).
9. Bahram, M. et al. Structure and function of the global topsoil microbiome. *Nature* **560**, 233–237 (2018).
10. Fierer, N. & Jackson, R. B. The diversity and biogeography of soil bacterial communities. *Proc. Natl Acad. Sci. USA* **103**, 626–631 (2006).
11. Karimi, B. et al. Biogeography of soil bacteria and archaea across France. *Sci. Adv.* **4**, eaat1808 (2018).
12. Delgado-Baquerizo, M. et al. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. *Ecology* **99**, 583–596 (2018).
13. Tedersoo, L. et al. Global diversity and geography of soil fungi. *Science* **346**, 1256688 (2014).
14. Leimu, R. & Fischer, M. A meta-analysis of local adaptation in plants. *PLoS ONE* **3**, e4010 (2008).
15. Ågren, J. & Schemske, D. W. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. *New Phytol.* **194**, 1112–1122 (2012).
16. Fournier-Level, A. et al. A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86–89 (2011).
17. Wadgymar, S. M. et al. Identifying targets and agents of selection: innovative methods to evaluate the processes that contribute to local adaptation. *Methods Ecol. Evol.* **8**, 738–749 (2017).
18. Brachi, B. et al. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Mol. Ecol.* **22**, 4222–4240 (2013).
19. Wagner, M. R. et al. Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild *Arabidopsis* relative. *Ecol. Lett.* **17**, 717–726 (2014).
20. Panke-Buisse, K., Poole, A. C., Goodrich, J. K., Ley, R. E. & Kao-Kniffin, J. Selection on soil microbiomes reveals reproducible impacts on plant function. *ISME J.* **9**, 980–989 (2015).
21. Lu, T. et al. Rhizosphere microorganisms can influence the timing of plant flowering. *Microbiome* **6**, 231 (2018).
22. Robbins, C. et al. Root-associated bacterial and fungal community profiles of *Arabidopsis thaliana* are robust across contrasting soil P levels. *Phytobiomes J.* **2**, 24–34 (2018).
23. Fitzpatrick, C. R., Mustafa, Z. & Viliunas, J. Soil microbes alter plant fitness under competition and drought. *J. Evol. Biol.* **32**, 438–445 (2019).
24. Frachon, L. et al. Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. *Nat. Ecol. Evol.* **1**, 1551–1561 (2017).

25. Busoms, S. et al. Salinity is an agent of divergent selection driving local adaptation of *Arabidopsis* to coastal habitats. *Plant Physiol.* **168**, 915–929 (2015).
26. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/081257v1 (2016).
27. Glassman, S. I. & Martiny, J. B. H. Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. *MSphere* **3**, e00148-18 (2016).
28. CBOL Plant Working Groupet al. A DNA barcode for land plants. *Proc. Natl Acad. Sci. USA* **106**, 12794–12797 (2009).
29. Thiergart, T. et al. *Lotus japonicus* symbiosis genes impact microbial interactions between symbionts and multikingdom commensal communities. *mBio* **10**, e01833 (2019).
30. Castrillo, G. et al. Root microbiota drive direct integration of phosphate stress and immunity. *Nature* **543**, 513–518 (2017).
31. Finkel, O. M. et al. The effects of soil phosphorus content on plant microbiota are driven by the plant phosphate starvation response. *PLoS Biol.* **17**, e3000534 (2019).
32. Chaw, S. M., Chang, C. C., Chen, H. L. & Li, W. H. Dating the monocot–dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.* **58**, 424–441 (2004).
33. Ågren, J., Oakley, C. G., McKay, J. K., Lovell, J. T. & Schemske, D. W. Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **110**, 21077–21082 (2013).
34. Schlaeppi, K., Dombrowski, N., Oter, R. G., Ver Loren van Themaat, E. & Schulze-Lefert, P. Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc. Natl Acad. Sci. USA* **111**, 585–592 (2014).
35. Edwards, J. et al. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl Acad. Sci. USA* **112**, E911–E920 (2015).
36. Fitzpatrick, C. R. et al. Assembly and ecological function of the root microbiome across angiosperm plant species. *Proc. Natl Acad. Sci. USA* **115**, E1157–E1165 (2018).
37. Yeoh, Y. K. et al. Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. *Nat. Commun.* **8**, 215 (2017).
38. Xu, L. et al. Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proc. Natl Acad. Sci. USA* **115**, E4284–E4293 (2018).
39. Hacquard, S. et al. Microbiota and host nutrition across plant and animal kingdoms. *Cell Host Microbe* **17**, 603–616 (2015).
40. Levy, A. et al. Genomic features of bacterial adaptation to plants. *Nat. Genet.* **50**, 138–150 (2018).
41. Karasov, T. L. et al. *Arabidopsis thaliana* and *Pseudomonas* pathogens exhibit stable associations over evolutionary timescales. *Cell Host Microbe* **24**, 168–179.e4 (2018).
42. Kwak, M.-J. et al. Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nat. Biotechnol.* **36**, 1100–1109 (2018).
43. Garrido-Oter, R. et al. Modular traits of the rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. *Cell Host Microbe* **24**, 155–167.e5 (2018).
44. Poupin, M. J., Timmermann, T., Vega, A., Zuñiga, A. & González, B. Effects of the plant growth-promoting bacterium *Burkholderia phytofirmans* PsJN throughout the life cycle of *Arabidopsis thaliana*. *PLoS ONE* **8**, e69435 (2013).
45. Bai, Y. et al. Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* **528**, 364–369 (2015).
46. Anderson, L. C. et al. Bacteria and fungi respond differently to multifactorial climate change in a temperate heathland, traced with ^{13}C-glycine and FACE CO_2. *PLoS ONE* **9**, e85070 (2014).
47. Peay, K. G. et al. Convergence and contrast in the community structure of Bacteria, Fungi and Archaea along a tropical elevation–climate gradient. *FEMS Microbiol. Ecol.* **93**, 5 (2017).
48. Talbot, J. M. et al. Endemism and functional convergence across the North American soil mycobiome. *Proc. Natl Acad. Sci. USA* **111**, 6341–6346 (2014).
49. Coleman-Derr, D. et al. Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytol.* **209**, 798–811 (2016).
50. Lundberg, D. S. et al. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86–90 (2012).
51. Wagner, M. R. et al. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat. Commun.* **7**, 12151 (2016).
52. Cregger, M. A. et al. The Populus holobiont: dissecting the effects of plant niches and genotype on the microbiome. *Microbiome* **6**, 31 (2018).
53. Wright, K. M. et al. Adaptation to heavy-metal contaminated environments proceeds via selection on pre-existing genetic variation. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/029900v2 (2015).
54. Wright, J. W., Stanton, M. L. & Scherson, R. Local adaptation to serpentine and non-serpentine soils in *Collinsia sparsiflora*. *Evol. Ecol. Res.* **8**, 1–21 (2006).
55. Busoms, S. et al. Fluctuating selection on migrant adaptive sodium transporter alleles in coastal *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **115**, E12443–E12452 (2018).
56. Postma, F. M. & Ågren, J. Early life stages contribute strongly to local adaptation in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **113**, 7590–7595 (2016).
57. Postma, F. M. & Ågren, J. Among-year variation in selection during early life stages and the genetic basis of fitness in *Arabidopsis thaliana*. *Mol. Ecol.* **27**, 2498–2511 (2018).
58. Ågren, J., Oakley, C. G., Lundemo, S. & Schemske, D. W. Adaptive divergence in flowering time among natural populations of *Arabidopsis thaliana*: estimates of selection and QTL mapping. *Evolution* **71**, 550–564 (2017).
59. Oakley, C. G., Ågren, J., Atchison, R. A. & Schemske, D. W. QTL mapping of freezing tolerance: links to fitness and adaptive trade-offs. *Mol. Ecol.* **23**, 4304–4315 (2014).
60. Méndez-Vigo, B., Gomaa, N. H., Alonso-Blanco, C. & Picó, F. X. Among- and within-population variation in flowering time of Iberian *Arabidopsis thaliana* estimated in field and glasshouse conditions. *New Phytol.* **197**, 1332–1343 (2013).
61. Bartoli, C. et al. In situ relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. *ISME J.* **12**, 2024–2038 (2018).
62. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
63. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
64. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
65. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
66. Deshpande, V. et al. Fungal identification using a Bayesian classifier and the Warcup training set of internal transcribed spacer sequences. *Mycologia* **108**, 1–5 (2016).
67. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* **10**, 1200–1202 (2013).
68. Nilsson, R. H. et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* **47**, D259–D264 (2019).

Acknowledgements

This work was supported by funds from a European Research Council starting grant (MICRORULES) to S.H., a European Research Council advanced grant (ROOTMICROBIOTA) to P.S.-L. and grants from the Swedish Research Council to J.Å. S.H. and P.S.-L. were also supported by funds from the Max Planck Society, the 'Priority Programme 2125 DECRyPT' funded by the Deutsche Forschungsgemeinschaft and the 'Cluster of Excellence on Plant Sciences' programme funded by the Deutsche Forschungsgemeinschaft. The laboratory of C.A.-B. was funded by grant BIO2016-75754-P (AEI/FEDER). We thank N. Donnelly for scientific English editing.

Author contributions

S.H., P.S.-L. and J.Å. conceived the project. E.K., F.R., C.A.-B., J.Å. and S.H. selected natural *A. thaliana* populations. P.D. and S.H. collected the samples. P.D. prepared all of the samples and performed the microbial community profiling. P.D., T.T. and N.V. analysed the microbiota data. R.G.-O. provided bioinformatic tools. T.E. and J.Å. prepared the field reciprocal transplant experiment. J.Å., T.E. and P.D. analysed the plant fitness data. S.H. supervised the project. T.T., P.D., J.Å., P.S.-L. and S.H. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

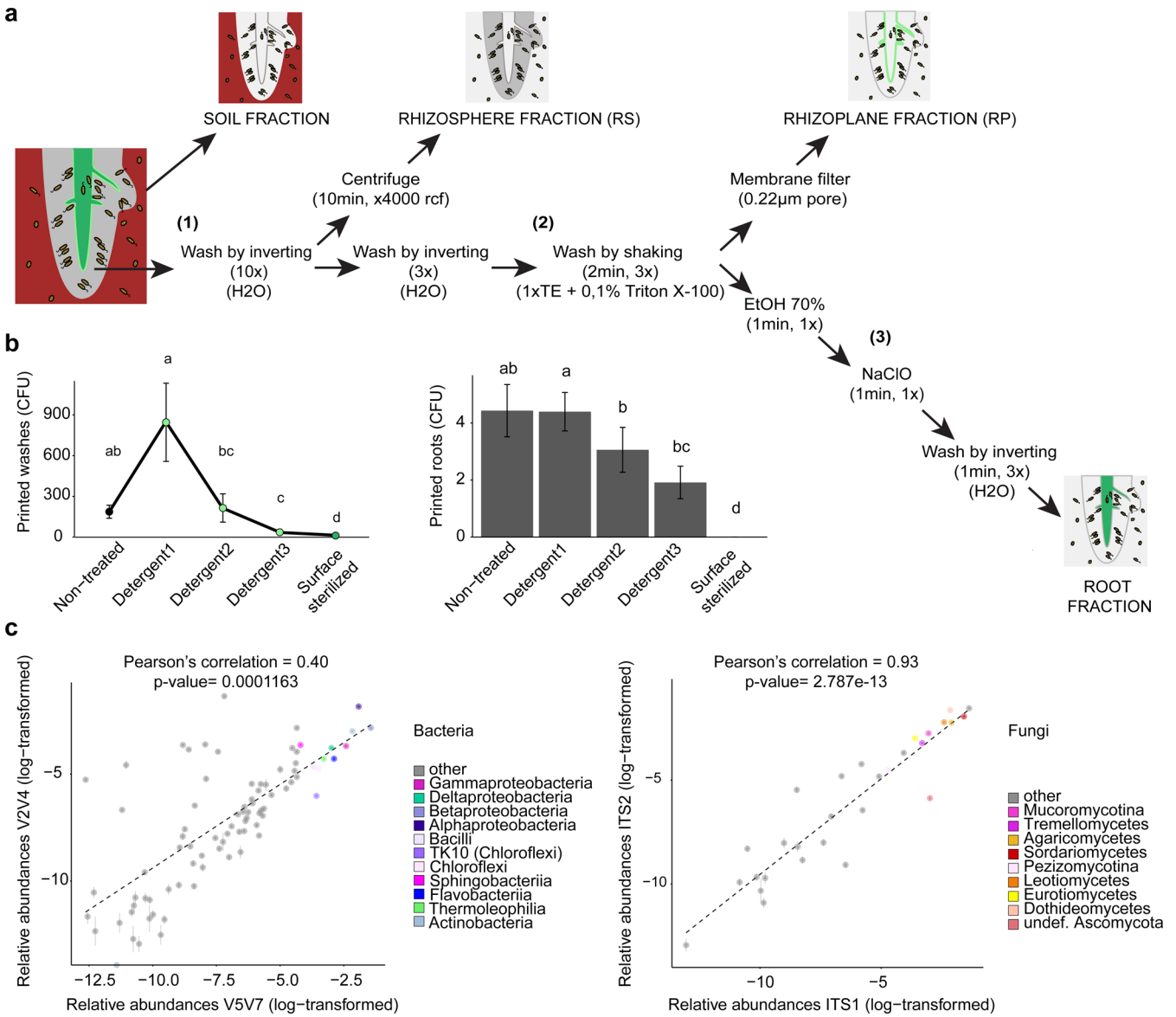
Extended data is available for this paper at https://doi.org/10.1038/s41559-019-1063-3.

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-1063-3.

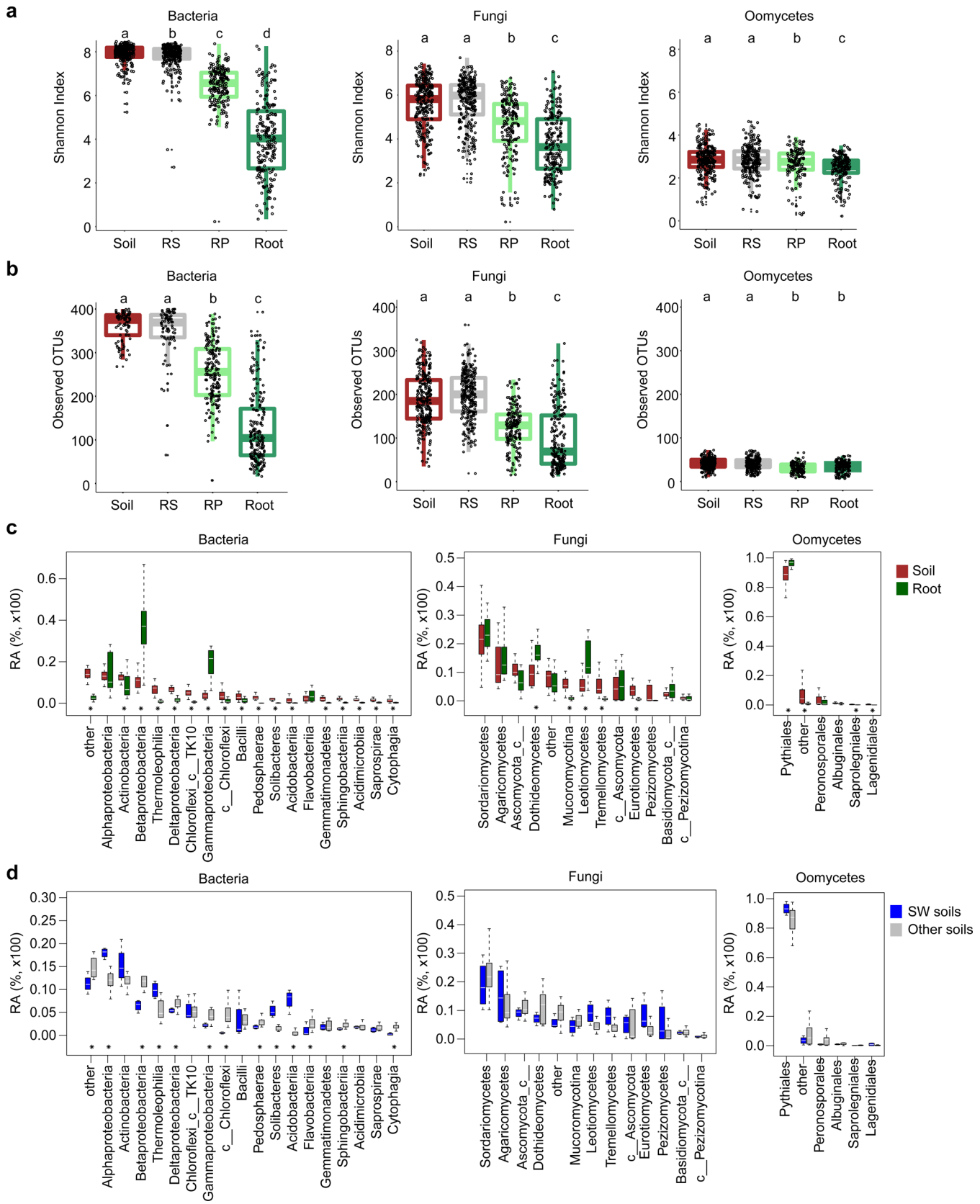
Correspondence and requests for materials should be addressed to J.Å., P.S.-L. or S.H.

Reprints and permissions information is available at www.nature.com/reprints.

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

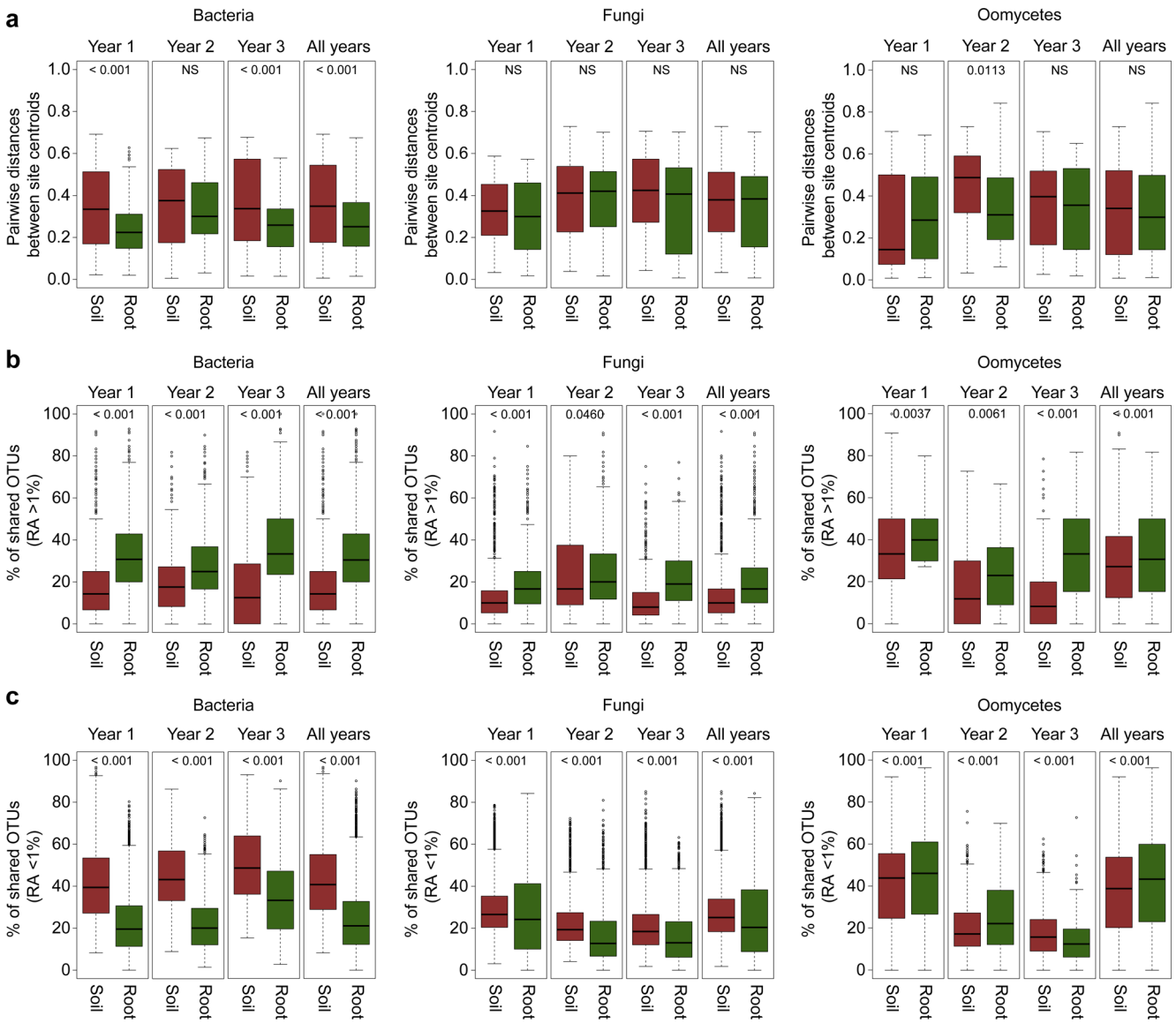


Extended Data Fig. 1 | Validation of the root fractionation protocol and assessment of primer amplification bias. a, Protocol to fractionate four microbial niches across a distance gradient from bulk soil to roots' interior. Roots of *A. thaliana* grown in their natural environments were briefly washed (1) to separate loosely attached soil particles from the root surface (rhizosphere, RS). After a second washing step, roots were vigorously washed with detergent (three times) to capture microbes that tightly adhere to the root surface (2). The resulting washes were then filtered through a 0.22 µM membrane (rhizoplane, RP). Finally, surface sterilization of root samples by consecutive EtOH and NaClO washes enriched the final root sample in microbial root endophytes (3). **b**, Validation of the fractionation protocol (depicted in panel a) was performed by printing root washes (left panel) and washed roots (right panel) on 50% Tryptic Soy Agar medium. Sequential detergent washes efficiently release microbes from the root surface and further root surface sterilization prevents the growth of rhizoplane-associated microbes (Wilcoxon rank sum test, $P < 0.01$). All three detergent steps were combined and filtered to prepare the RP fraction (light green dots, left panel). **c**, Comparison of bacterial (left panel) and fungal (right panel) classes profiled with the V3V4 and V5V7 regions of the bacterial 16 s rRNA gene, and the ITS1 and ITS2 of the fungal ITS. The correlation between the RA of each microbial class is shown (Pearson correlation: $P < 0.001$).

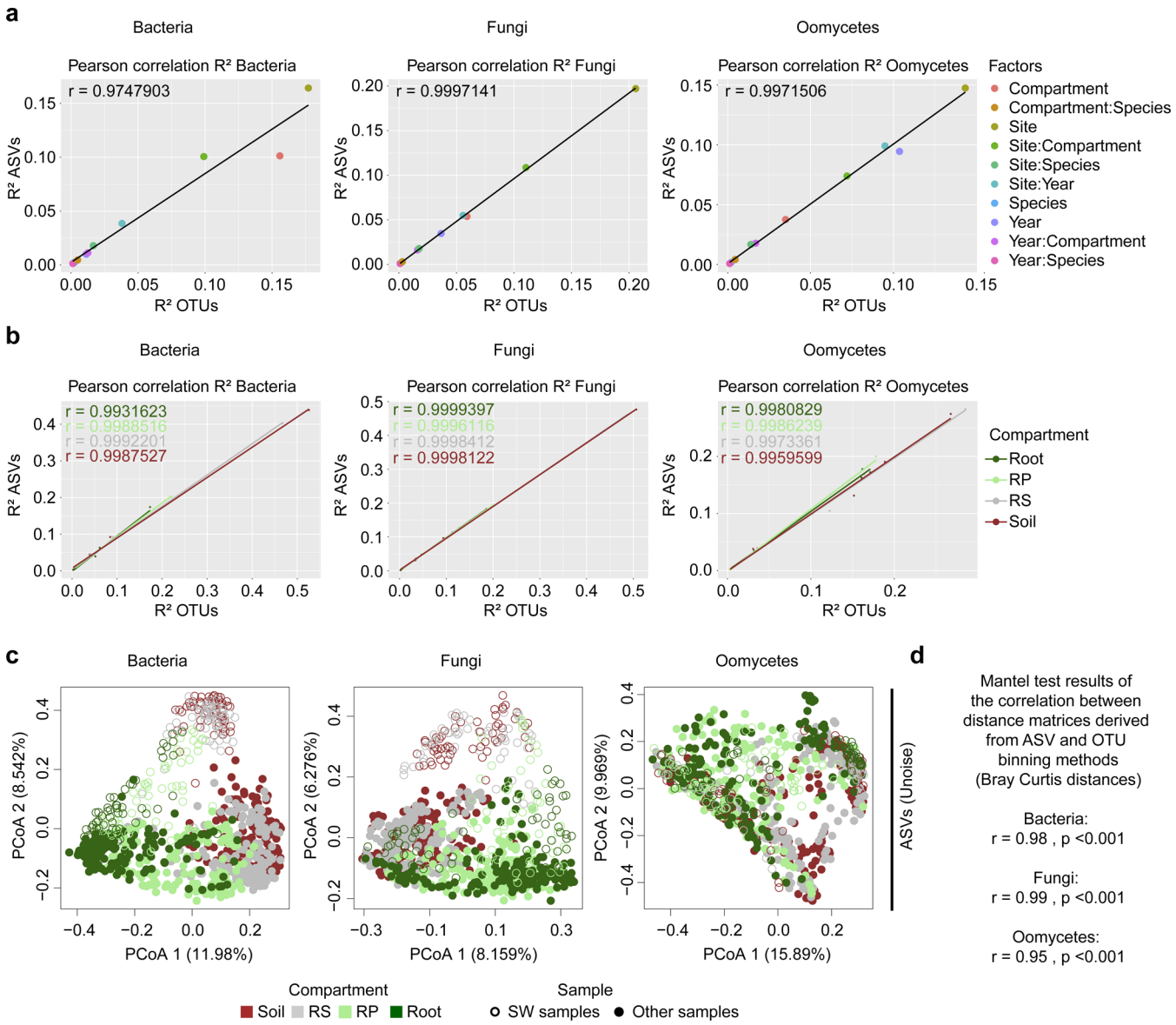


Extended Data Fig. 2 | See next page for caption.

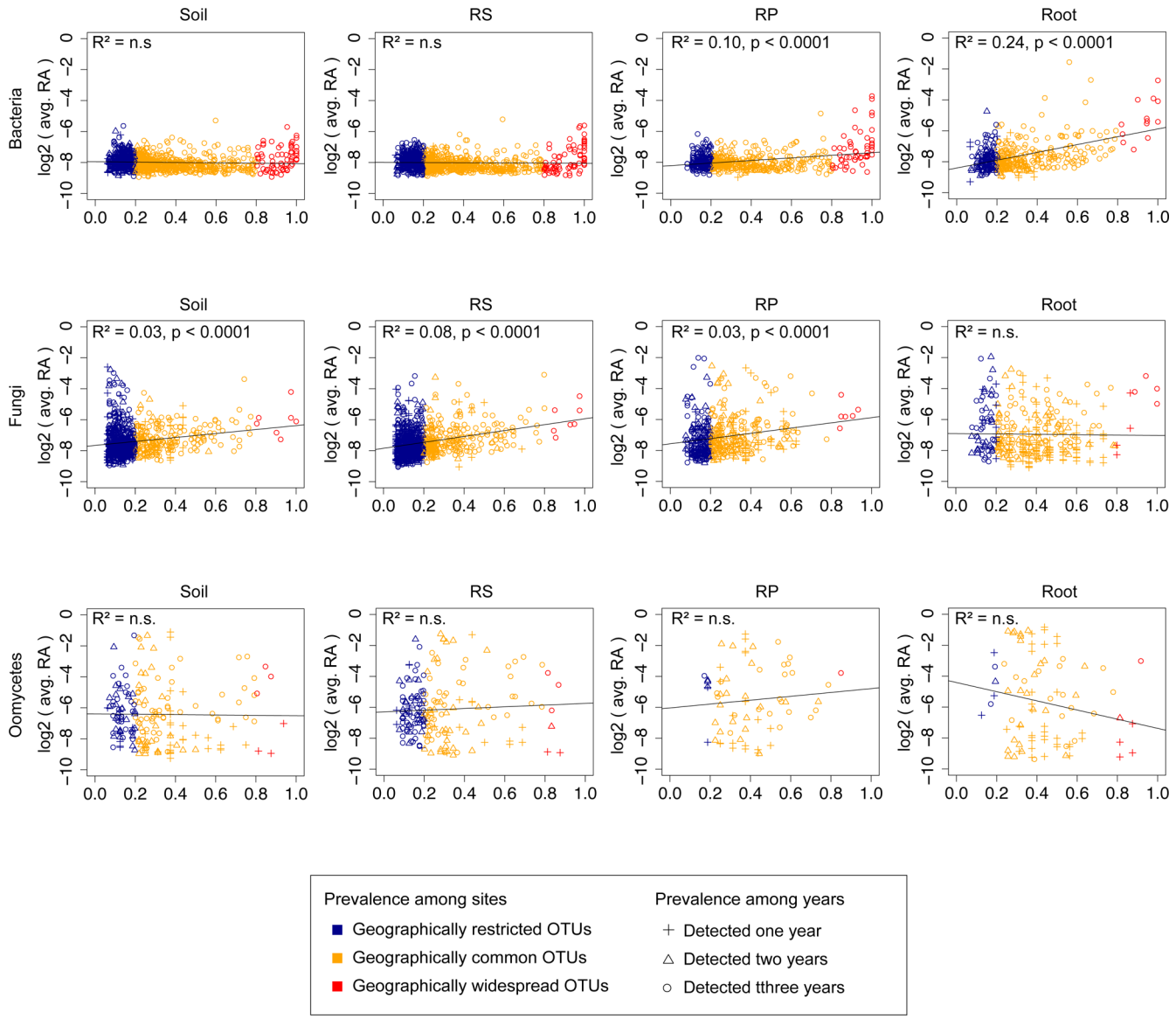
Extended Data Fig. 2 | Microbial alpha diversity and enrichment signatures in plant roots. a, Microbial alpha diversity measured across all 17 sites in soil, rhizosphere (RS), rhizoplane (RP), and root samples based on the Shannon index. All samples from a given site were taken into account and the datasets were rarefied to 1,000 reads. Individual data points within each box correspond to samples from the 17 natural sites (Kruskal-Wallis with Dunn's post hoc test: $P < 0.05$). **b**, same as **a**, but the alpha diversity metric observed OTUs is shown instead of the Shannon index. **c**, Comparison of taxa RA between soil (dark red) and root (dark green) samples for bacteria (left), fungi (middle), and oomycetes (right). RA measured in soil and root samples across the 17 *A. thaliana* populations were aggregated at the class (bacteria and fungi) and order (oomycetes) levels. Significant differences are marked with an asterisk (Wilcoxon rank sum test, FDR < 0.05). **d**, Comparison of taxa RA between Swedish soils (SW1-4, blue) and the other European soils (grey) for bacteria (left), fungi (middle), and oomycetes (right). RA is aggregated at the class (bacteria and fungi) and order (oomycetes) levels and significant differences are marked with an asterisk (Wilcoxon rank sum test: FDR < 0.05).



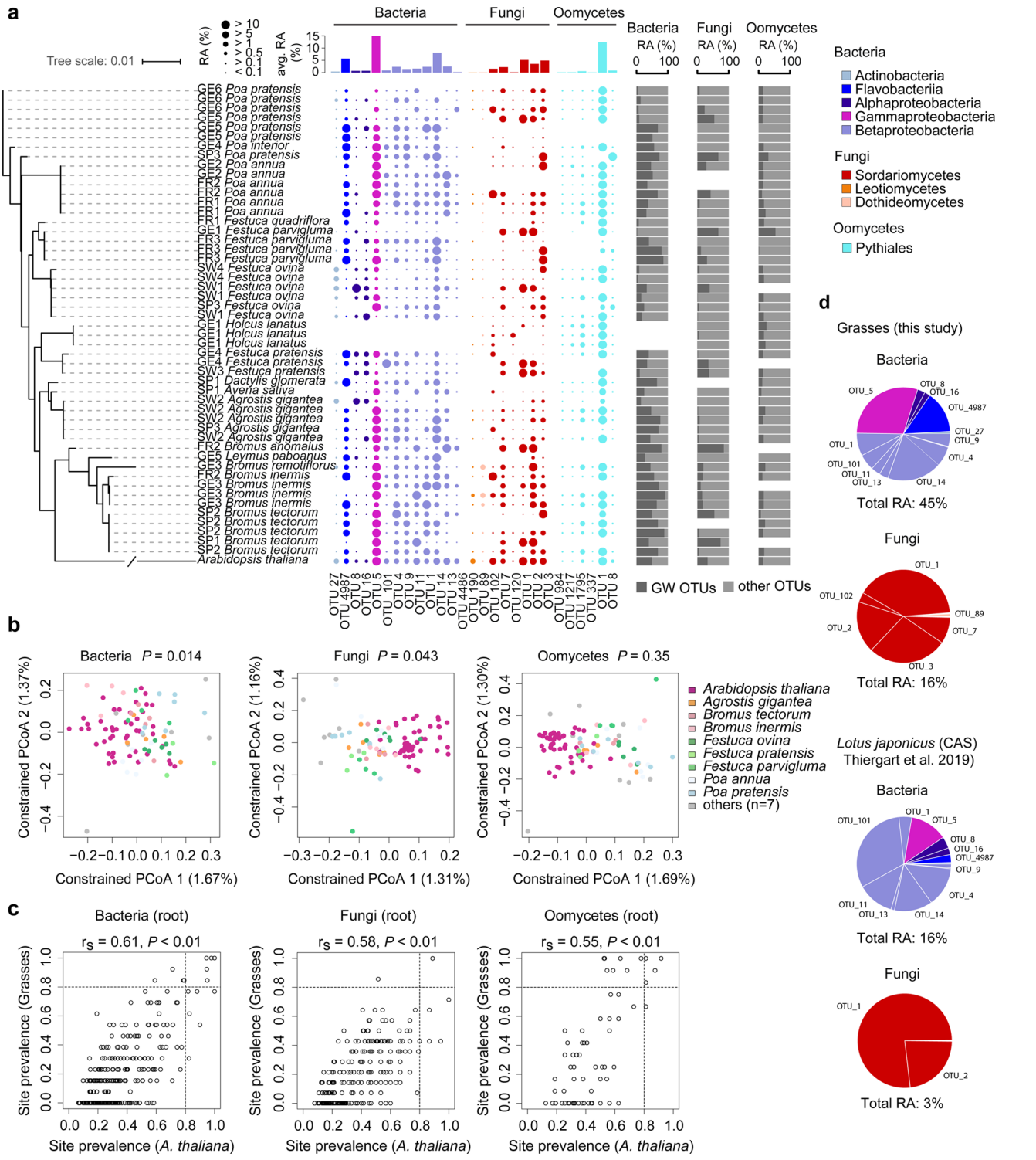
Extended Data Fig. 3 | Between-site variation and shared OTUs in soil and root samples. a, Between-site distances based on site centroids plotted for soil and root samples for each year and for all years combined together. Distances were calculated using site centroids of Bray-Curtis distances. Statistical significance was tested using a t-test. **b**, Percentage of shared high-abundant OTUs among soil and root samples. OTUs having RA > 1% were considered. Statistical significance was tested using a t-test. **c**, Percentage of shared low-abundant OTUs among soil and root samples. OTUs having RA < 1% were considered. Statistical significance was tested using t-test.



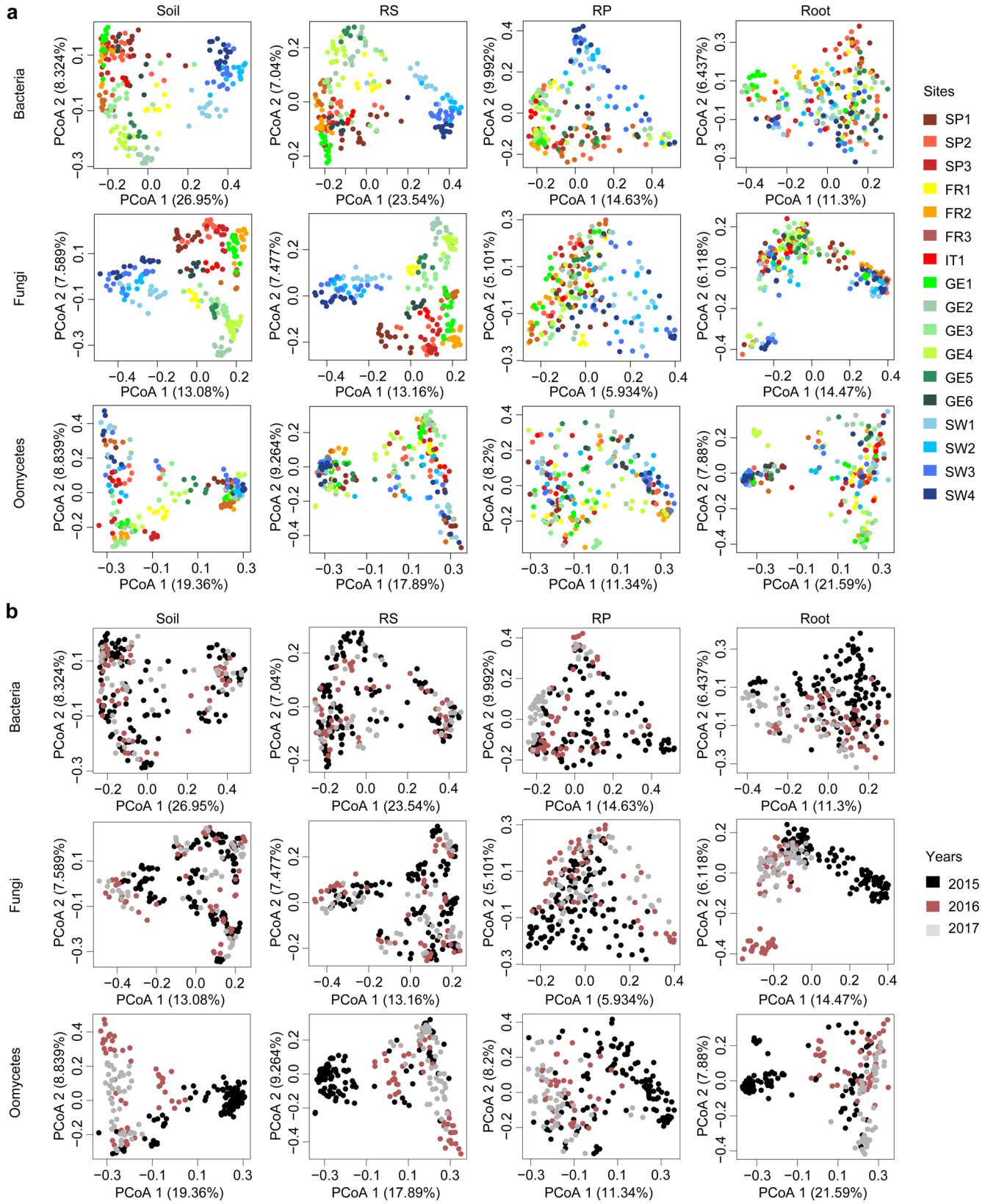
Extended Data Fig. 4 | Comparison of beta-diversity metrics deriving from OTU and ASV binning approaches. a, Pearson correlation between R^2 values of explanatory factors from PERMANOVA tests with ASV and OTU datasets for bacteria, fungi and oomycetes (n = 10, see Supplementary Table 3). Each dot represents an explanatory factor and Pearson's r values are indicated. **b**, Pearson correlation between R^2 values of explanatory factors from PERMANOVA tests with ASVs and OTUs datasets for bacteria, fungi and oomycetes (n = 6, see Supplementary Table 5). Each dot represents an explanatory factor and each line represents the correlation between R^2 values for a specific compartment. Pearson's r values for each compartment are indicated. **c**, Principal coordinate analysis (PCoA) of Bray-Curtis distances using ASVs data. Bacterial, fungal and oomycetal communities from soil, rhizosphere (RS), rhizoplane (RP) and root samples are shown with different colours. Samples from Swedish sites are indicated by open circles. **d**, Mantel test results of the correlation between Bray-Curtis distance matrices computed based on ASV and OTU datasets.



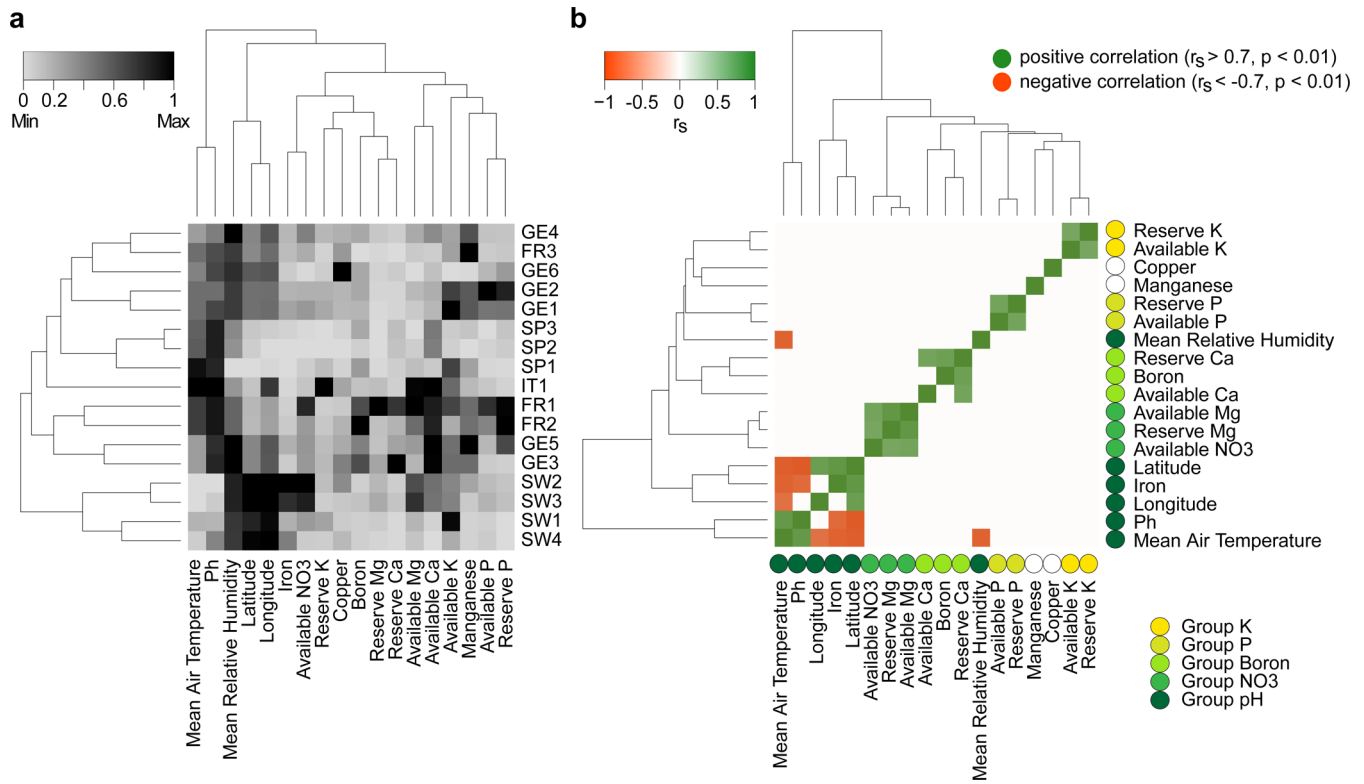
Extended Data Fig. 5 | Geographically widespread taxa at the soil-root interface. Correlation between OTUs prevalence across sites in soil, rhizosphere (RS), rhizoplane (RP), root and averaged OTUs RA (log2). Bacteria: upper panels. Fungi: middle panels. Oomycetes: lower panels. Blue: geographically restricted OTUs (site prevalence < 20%). Orange: geographically common OTUs (site prevalence 20-80%). Red: geographically widespread OTUs (site prevalence > 80%). For calculating averaged RA, only samples where the actual OTUs are present were considered. The different shapes highlight OTUs detected one year, or across two or three years. RA and prevalence were averaged across the years where one OTU is present. OTUs with RA < 0.1% were excluded from the datasets.



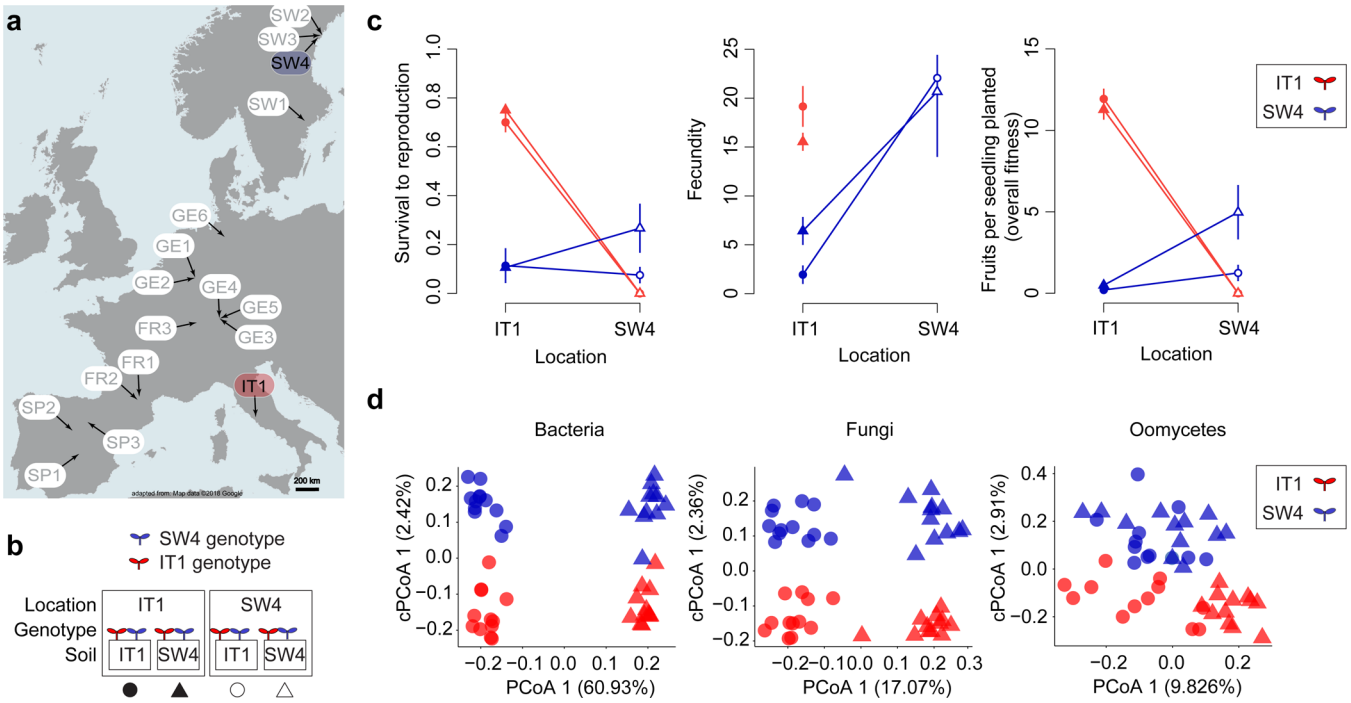
Extended Data Fig. 6 | Conservation of geographically widespread OTUs in roots of co-occurring grass species. a, Diversity of grass species and conservation profiles of geographically widespread OTUs in roots. The Neighbor-Joining tree was constructed based on the plant *rbcL* locus sequenced from 50 samples collected across 17 sites and consecutive years. Taxonomic assignment at the species level was determined by blast search against the nr database at NCBI. The RA (%) of each of the geographically widespread OTUs is shown for root endosphere samples, together with the cumulative abundance in root samples. **b**, Bray-Curtis distances constrained by species (n = 17) for bacterial, fungal, and oomycetal communities in root endosphere samples. All distinct plant species identified only once were grouped as others. The percentage of variation explained by plant species is shown for the first and second axis (bacteria: $P = 0.014$; fungi: $P = 0.043$; oomycetes: $P = 0.35$). **c**, For each microbial group (bacteria, fungi, and oomycetes), Spearman´s rank correlations ($P < 0.01$) were determined between OTUs prevalence in roots of *A. thaliana* and OTUs prevalence in roots of neighboring grasses. **d**, Conservation of geographically widespread OTUs in roots of co-occuring grasses and *Lotus japonicus*. The RA and proportion of widespread bacterial and fungal OTUs detected in *A. thaliana* roots are shown for co-occuring grasses (17 sites), as well as for *Lotus japonicus* grown in the Cologne Agricultural Soil (CAS). All shown OTUs have RA > 0.1%. The total RA of these OTUs in root samples is indicated below the circular diagrams.



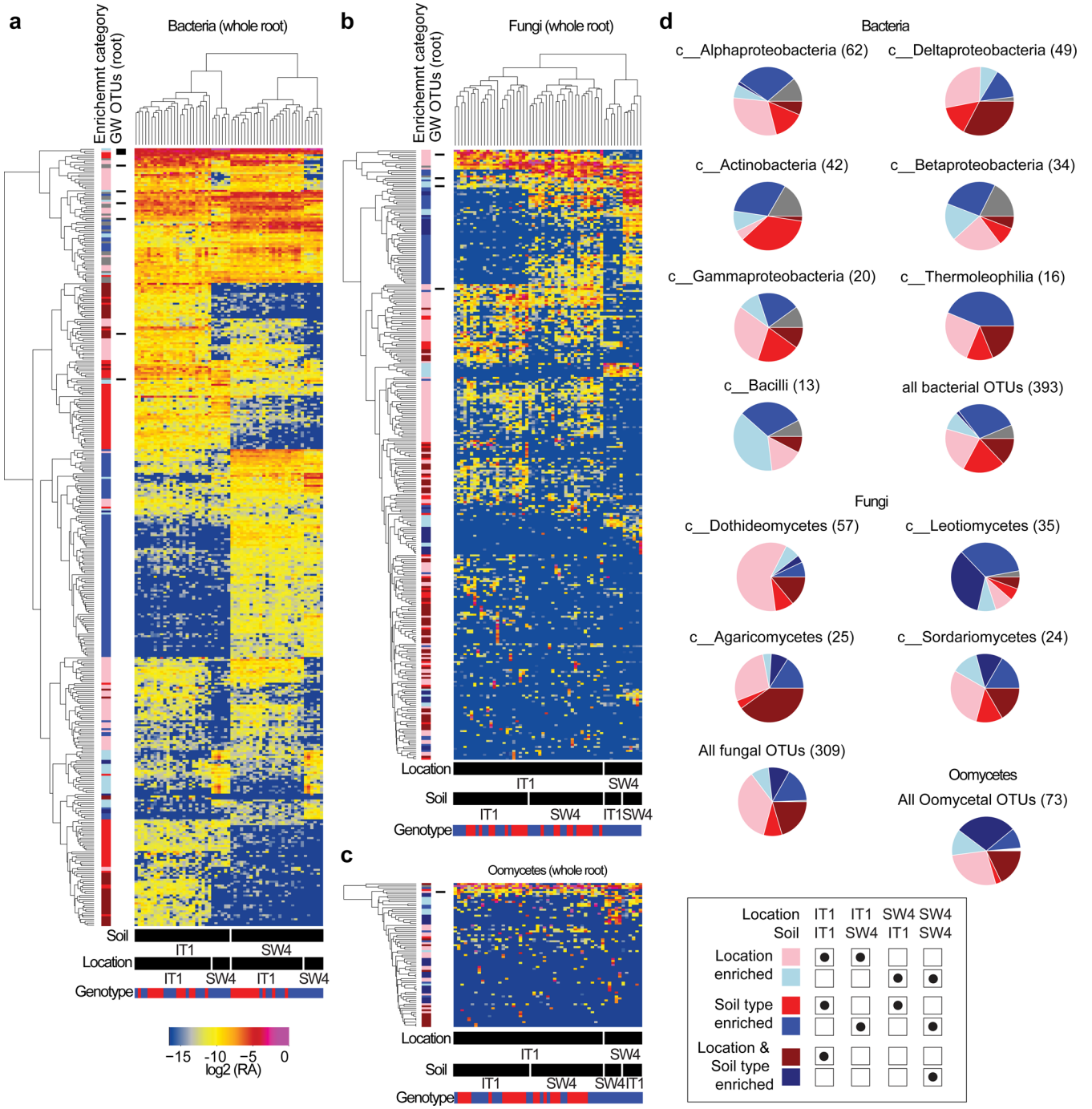
Extended Data Fig. 7 | Influence of site and harvesting year on microbial community structure in *A. thaliana* populations. a, Principal coordinate analysis (PCoA) based on Bray-Curtis distances for soil-, rhizosphere (RS), rhizoplane- (RP), and root-associated microbial communities detected in 17 sites across three successive years in European *A. thaliana* populations (across all compartments: n = 881 for bacteria, n = 893 for fungi, n = 875 for oomycetes). Microbial communities in each compartment are presented for bacteria, fungi, and oomycetes, and color-coded according to the site. **b**, Same PCoA plots as in panel a, but data points were color-coded according to the harvesting year (2015, 2016, 2017). OTUs with RA < 0.1% were excluded from the datasets.



Extended Data Fig. 8 | Environmental variables and multicollinearity analysis. a, Heatmap showing variation in 18 environmental variables measured across the 17 European sites. Values for each environmental property were normalized and the scale was adjusted to 1. 0 = lowest value measured value across sites, 1 = highest value measured value across sites). **b**, Heatmap depicting collinearity between the 18 environmental variables. Significant correlations detected between variables were assessed using Spearman´s rank correlation and variables showing correlation $r_s > 0.7$ and $r_s < -0.7$ ($n = 17$, $P < 0.01$) were considered as collinear. Two unique variables (Manganese, Copper) and five groups of collinear variables (group K, group P, group Boron, group NO3, group pH) were defined and used for PERMANOVA analyses.



Extended Data Fig. 9 | Reciprocal transplant between two *A. thaliana* populations in Sweden and Italy. a, European map showing names and locations of the 17 *A. thaliana* populations. The IT1 and SW4 sites selected for the reciprocal transplant experiment are highlighted in red and blue, respectively. **b**, Schematic overview of the reciprocal transplant experiment. Soils and plant genotypes from IT1 and SW4 sites were reciprocally transplanted in the two locations (eight different treatment combinations). The symbols below the schematic view correspond to the symbols used in panels c and d. **c**, Fitness of Italian and Swedish genotypes (red and blue color, respectively) when reciprocally planted in Italian and Swedish soils (circle and triangle symbols, respectively) and grown at Italian and Swedish locations (filled and open symbols, respectively). Plant survival, fecundity (number of fruits per reproducing plant), and overall fitness (number of fruits per seedling planted). Means based on block means ± SE are given. Note that no Italian plant survived to reproduce at the Swedish site. **d**, Bray-Curtis distances constrained by genotype for bacterial, fungal, and oomycetal communities in whole root samples (cPCoA, see axis 2). Results are shown for Italian and Swedish genotypes (red and blue color, respectively) planted in Italian and Swedish soils (circle and triangle symbols, respectively) at IT1 site only since no Italian plant survived at the SW4 site. The percentage of variation explained by the two genotypes is plotted along the second axis and refers to the fraction of the total variance of the data that is explained by the constrained factor (that is genotype; bacteria $P = 0.001$; fungi $P = 0.026$; oomycetes $P = 0.002$). Map data in **a** adapted from Google Maps, 2018.



Extended Data Fig. 10 | OTU distribution patterns across root samples in the transplant experiment. a, Heatmap depicting the RA (log2) of bacterial OTUs in roots of Italian and Swedish genotypes grown in Italian and Swedish soils at IT1 and SW4 locations. OTUs and samples are hierarchical clustered. Enrichment patterns of each OTU was estimated according to the categories described in the lower right side of the figure and highlighted with different colours next to the heatmap. The RA of OTUs falling into one of the six categories is always higher in that category compared to the mean RA measured across all samples. OTUs that are present in all samples (RA > 0.1%) and did not fall in any of the six categories are marked in grey. The heatmap is filtered for OTUs that have at least an average RA of 0.01% across all root samples. Samples have been filtered to contain at least 1,000 reads. Genotype of plants for each sample is indicated below each heatmap. Blue: Swedish genotype. Red: Italian genotype. Note that no Italian plant survived at the Swedish site. GW OTUs: geographically widespread OTUs. **b**, Heatmap depicting the RA (log2) of fungal OTUs in roots of Italian and Swedish genotypes grown in Italian and Swedish soils at IT1 and SW4 locations. **c**, Heatmap depicting the RA (log2) of oomycetal OTUs in roots of Italian and Swedish genotypes grown in Italian and Swedish soils at IT1 and SW4 locations. **d**, Percentage of OTUs falling into one of the six categories are presented as pie charts for each main taxonomic classes. The number of OTUs that belong to each microbial class is given in brackets.

nature research

Corresponding author(s): Jon Ågren, Paul Schulze-Lefert, Stéphane Hacquard

Last updated by author(s): 25/10/19

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect data in this study
Data analysis	Qiime: http://qiime.org/ Usearch: https://www.drive5.com/usearch/ RDP: https://rdp.cme.msu.edu/classifier/classifier.jsp R: https://www.r-project.org/ R, vegan package: https://www.rdocumentation.org/packages/vegan/versions/2.4-2 Code & raw data: https://github.com/ththi/European-Root-Suppl

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequencing data, european transect at ENA: ERP115101
Sequencing data, transplant experiment at ENA: ERP115102

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Spatial and temporal monitoring of the Arabidopsis thaliana root microbiota in 17 European sites across three successive years. Effect of soil origin and location were uncoupled in a reciprocal transplant experiment between two A. thaliana populations in northern and southern Europe to test the contribution of edaphic and climatic conditions to root microbiota variation and local plant adaptation at large spatial scale.
Research sample	Roots and associated soil fractions (soil, rhizosphere, rhizoplane, root) were harvested from flowering Arabidopsis thaliana and co-occurring grasses (Poaceae family) at 17 natural sites in spring 2015, 2016, and 2017 (n = 1,125 samples in total). For the reciprocal transplant experiment, soil and whole root samples (root with rhizoplane microbes) were harvested from A. thaliana at fruit maturation and include 8 different genotype-soil-location combinations (n = 131 samples in total). On the host side, we scored plant survival, fecundity (number of fruits per reproducing plant), and estimated overall fitness (number of fruits per seedling planted) across 1,008 plant individuals.
Sampling strategy	Plants were harvested with their surrounding soil, transferred in pots and transported to the laboratories of collaborators within less than 24 hours after harvesting. In the laboratory, plant compartments were separated using a custom protocol. No sample-size calculation was previously performed. Due to limited plants availability in some sites and to maintain population stability across years, a maximum of 16 A. thaliana individuals were harvested per site and year. Four pooled plants were considered as one technical replicate and four (A. thaliana) and three (for grasses) technical replicates were taken per site and year. For the reciprocal transplant experiment, six to twelve soil samples were harvested for each of the eight combinations of soil, location and genotype (n = 72), as well as three to twelve whole root samples per condition (n = 59).
Data collection	Isolation of DNA from soil and root samples, as well as library preparation are described in the Methods section. Bacterial, fungal and oomycetal communities were profiled using primer pairs targeting the V2V4 and V5V7 regions of the bacterial 16S rRNA gene, the ITS1 and ITS2 segments of the fungal ITS, and the ITS1 segment of the oomycetal ITS, resulting in the sequencing of 5,625 microbial community profiles for the European transect and 393 for the reciprocal transplant experiment. Sequencing was performed using an in-house MiSeq device. In the common garden experiment, A. thaliana fitness was obtained by scoring survival after winter, flowering time and number of siliques per individual (n = 1,008 plant individuals).
Timing and spatial scale	Plants were harvested along a latitudinal gradient (up to 3500 km) from Sweden to Spain. Across European sites (n = 17), all plants were harvested at the flowering stage from February to May in 2015, 2016 and 2017. The difference in harvesting time reflects that fact that not all populations flower at the same time across Europe. For example, plants at German sites were sampled in March, whereas those at Swedish sites were sampled in May. Note that plants at each site were harvested at one specific date and a similar time frame was selected every year to assess year-to-year variation. The reciprocal transplant include two geographically distant sites in Sweden and Italy in which plants were harvested at fruit maturation in April (Italy) and May (Sweden) 2016.
Data exclusions	Very few samples where 16S rRNA or ITS sequences did not meet predetermined quality checks (i.e. low quality or low read counts) were excluded.
Reproducibility	All attempts to repeat the same sampling strategies across years were successful.
Randomization	A. thaliana and grasses were randomly harvested across sites to maximize the representation of the samples taken. In the common garden experiment, seedlings were transplanted in blocks of 6x7 cells in a checkerboard design and randomly harvested for microbiota analysis.
Blinding	Blinding was not relevant for this study as plants were harvested regardless of their phenotypical state.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Soil chemical properties were determined for each site and yearly average air temperature and humidity data were obtained from public online sources (see Supplementary Table 1)
Location	GPS coordinates and information of natural sites and reciprocal transplant sites can be found in the Supplementary Table 1
Access and import/export	Collaborators of each country were contacted in order to identify natural sites and access them. Harvested samples were then transported to the laboratory and, once fractionated, shipped to the MPIPZ in Cologne in dry ice.

Disturbance	Soil was disturbed when plants and soil were harvested at each site. In order to minimize disturbance, holes made were covered with surrounding soil. To maintain plant population stability across years, a maximum of 16 A. thaliana individuals were harvested per site and year

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The study did not involve laboratory animals
Wild animals	The study did not involve wild animals
Field-collected samples	Plant and soil samples collected from natural sites
Ethics oversight	Since no animals were used in this study, no ethical approval or guidance was required

Note that full information on the approval of the study protocol must also be provided in the manuscript.