1

MULTI-GENERATION SELECTION OF RHIZOSPHERE ECOSYSTEM ASSOCIATED

WITH PLANT GENOTYPE AND BIOMASS IN *ARABIDOPSIS THALIANA*

Nachiket Shankar[1*], Prateek Shetty[2], Rick Kesseli[1]

[1]Department of Biology, University of Massachusetts, Boston, MA 02125,

[2]Institute of Plant Biology, Hungarian Academy of Sciences, Biological Research Centre, 6726

Szeged, Hungary

Abstract

The role of the microbiome in shaping the host phenotype has emerged as a critical area of

investigation, with implications in ecology, evolution, and host health. The complex and dynamic

interactions involving plants and their diverse rhizosphere microbial communities are influenced

by a multitude of factors, including but not limited to soil type, environment, and plant genotype.

Understanding the impact of these factors on microbial community assembly is key to yielding

host-specific and robust benefits for plants, yet remains challenging. Here we ran an artificial

ecosystem selection experiment, over eight generations, in *Arabidopsis thaliana* L*er* and Cvi to

select soil microbiomes associated with higher or lower biomass of the host. This resulted in

divergent microbial communities, shaped by a complex interplay between random environmental

variations, plant genotypes, and biomass selection pressures. In the initial phases of the

experiment, the genotype and the biomass selection treatment have modest but significant impacts.

Over time, the plant genotype and biomass treatments gain more influence, explaining ~40% of

the variation in the microbial community composition. Furthermore, a genotype-specific

association of plant growth-promoting rhizobacterial taxa, *Labraceae* with L*er* and *Rhizobiaceae*

with Cvi, is observed under selection for high biomass. This study enhances our understanding of

both microbial community dynamics and plant genetics and can be applied to agricultural settings

to improve plant growth and soil restoration.

27   **Introduction**

The conventional understanding of the host phenotype involves genetics and environment shaping

observable traits. Yet, the last two decades have underscored the microbiome's significance in

shaping the host phenotype, driven by extensive research on its role in ecology, evolution, and host

health. The plant microbiome represents a rich source of functional diversity that is not encoded

within the host genome. The interactions between the plant host and its microbiome are dynamic

and reciprocal, with the plant shaping its immediate environment by exuding specific metabolites,

thereby promoting the growth of specific microbial taxa, while the microbiome in turn influences

plant health and growth (1). A multitude of biotic and abiotic factors influence the dynamic nature

of the microbiome, including but not limited to plant root exudates (2–4), soil type (5–9),

environment (10), and various aspects of the plant, including species (11), genotype (4,12–15) and

developmental stage (16,17).

Previous studies in *Arabidopsis thaliana* have shown that the microbial diversity of the soil

reduced with proximity to the endophytic compartment, and microbes found in the bulk soil were

not enriched in the endophytic compartment and vice-versa (18–20). Furthermore, the microbiome

of greenhouse-grown plants was similar to that of the field-grown plants, in the different plant

compartments – rhizosphere, woody stem, and endophytic compartment, indicating an active role

of the plant host in creating and maintaining an environment where certain microbes have

improved fitness (21). The soil environment was a stronger predictor of microbiome structure,

though plant genotype had a weak influence in some studies (20,22,23). A follow-up study

confirmed that variation in microbiome communities depends more on the environment than on

different *Arabidopsis* varieties and sister species. Though there were some differences between

different host species, the core microbiome interacting with the plant host at the root endophytic

zone was largely consistent and reproducible (24).

Unfortunately, most prior studies were conducted in a single plant growth cycle, precluding

an analysis of the temporal stability of the plant-microbiome association. Observations from plant-

soil feedback studies highlight that plant-microbiome interactions can be altered by successive

growth cycles of a given plant in the same soil (25,26). Prior studies have employed plant-mediated

selection on the soil ecosystem over multiple generations, and resulted in consistent effects on

plant characteristics such as biomass, flowering time, and germination (27,28). By selecting for

soils where hosts show the desired phenotype, such as high biomass or altered flowering time, it

is possible to enrich microbes that modulate host traits. Additionally, studies showed the plant's

response to abiotic stress, such as drought and salt stress, could be influenced in two ways. First,

through generational selection, where plants with desirable stress responses were chosen over

multiple generations (29–31). Second, by introducing beneficial microbes that were associated

with plants and had previously lived in similar drought or salt conditions (32). Mueller et al., (29)

used generational selection to create microbial communities that could improve plant seed

production by up to 205%. Soil inoculum obtained from drought-exposed soils improved wheat

biomass under drought conditions (32). Older studies, such as Swenson et al. (28), also have shown

similar results under optimal growing conditions without necessarily exploring the microbiome. A

recent study conducted on the rhizosphere microbiome of wild and domesticated tomato plants

over multiple generations demonstrated an escalating influence of host genotype on the

microbiome community (33). A separate study on the phyllosphere microbiome of various tomato

genotypes indicated a declining impact of host genotype across successive generations (34).

Despite several earlier efforts, the development of a robust and host-specific microbial

community with long-lasting beneficial effects on the plant host remains a challenge.

Using an artificial ecosystem selection experiment, we sought to advance our understanding of

how plant genotype, environment, and biomass selection treatment (henceforth referred to as

biomass treatment) impact the assembly of host microbial communities. Treatments selecting for

the microbial communities of high and low biomass plants did affect the growth of plants in

subsequent generations, though unknown environmental factors had substantial effects throughout

the experiment. In the first few generations of the experiment, the genotype and biomass treatments

played a modest role in shaping the microbial community. Over time, the plant genotype and

biomass treatments had increasing influence, explaining ~40% of the difference in the microbial

community composition. Moreover, we observe an enrichment of known plant growth-promoting

rhizobacteria (PGPR) in the high biomass treatment, indicating potentially beneficial host-specific

interactions.

84   **Material and Methods**

Multi-generation selection of soil ecosystem

Arabidopsis thaliana Cvi and L*er* accessions were cultivated using custom-made

"rhizotubes," (Stuart Morey, Univ. of Massachusetts, Boston, unpublished). The tubes were

equipped with a black polyethylene sleeve, which effectively blocked light penetration. This

design also facilitated the convenient removal of the plant from the pot, granting full access to

the root system. (Supplementary Figure S2.1).

The potting soil used was PRO-MIX PGX, a commercial mixture comprising 80-90%

sphagnum peat moss and small quantities of perlite. It was autoclaved twice for 40 minutes with

a 48-hour interval between each sterilization. The potting soil was then sifted through a 3mm

sieve and combined with field soil in a 6:1 ratio respectively. The field soil obtained from the

Center for Agricultural Research in Waltham, Massachusetts, was composed of 44% sand, 49%

silt, and 7% clay representative of agricultural and grassland ecosystems.

The mixture was homogenized using a custom-made cement mixer and attached with

sterile bins to avoid cross-contamination. Distilled water was gradually added to achieve a final

ratio of 1:2 (water to soil). The soil was incubated at room temperature for two days following

inoculation before planting the seeds. All seeds used in the experiment were obtained from a

single parent plant. Before planting, the seeds were treated with a solution of 50% bleach vol/vol

and a drop of tween 20 for 10 minutes and rinsed ten times with sterile distilled water. We placed

3-4 seeds in the center of each pot, and kept only one seedling per pot after emergence. All plants

were grown at 22°C day/ 18°C night with a 12/12 hour day night cycle in a controlled growth

chamber. Relative humidity ranged from 35-60 and the light intensity 96uE. No fertilization was

applied to plants.

In the first generation of the experiment, 100 plants of each accession (Cvi and L*er*) were

grown separately in individual pots. The plants were arranged in a randomized block design to

avoid any batch effects. All plants were harvested 35 days after germination. The above-ground

portion of the plant, the rhizosphere, and the bulk soil from each rhizotube were separated,

ensuring there was no cross-contamination (see Supplementary Figure S2.1). The above-ground

part of the plant, stem, and leaves were dried at 70°C for 4 days. All the plants were weighed

individually on a closed weighing scale accurate to 1 mg. The root-soil complex (comprising the

rhizosphere and the endosphere, henceforth referred to as the rhizosphere) was obtained by

shaking the excess soil off the root and placing the root and remaining attached soil in a sterile

5 ml tube. Bulk soil samples were moved to a sterile Ziplock bag. Tubes and bags were

immediately transferred to dry ice and then stored at -80°C for DNA analysis.

The inoculation process for the next generation of soil involved using the bulk soil of the

top five and bottom five plants by weight, resulting in two different biomass treatments for each

genotype: high biomass L*er*, low biomass L*er*, high biomass Cvi, and low biomass Cvi. Starting

from the second generation of the experiment, 50 plants are grown on each of the four soil

inoculants to maintain a manageable experiment size. This process is repeated for eight

generations. There was substantial stochastic variation in above-ground plant biomass and during

generation 6 all plants died for unknown reasons. That experiment was therefore restarted using

soil from generation 5. Plants in uninoculated sterile potting soil control were grown from

generation 3 onwards to act as a reference for random environmental variation (REV).

DNA extraction and 16S rRNA amplicon library prep

Microbial DNA was isolated from the frozen rhizosphere samples using the Machery-

Nagel Nucleospin Soil DNA extraction kit (MACHEREY-NAGEL Inc., Allentown, PA, USA).

Approximately 0.1g of rhizosphere soil sample was used for DNA isolation. All samples were

diluted to 5 ng ul-1 with PCR-grade water. 16S rRNA gene was amplified from the isolated

DNA samples in triplicate in 96-well PCR plates. The PCR primers used were for the 16S rRNA

V4 region, 515F (5'-GTGYCAGCMGCCGCGGTAA-3') (ref. 35) and 806R (5'-

GGACTACNVGGGTWTCTAAT-3') (ref. 36) for downstream paired-end Illumina (Illumina,

Inc., San Diego, CA, USA) barcoded sequencing (37). The PCR cycling conditions were as

follows: 94°C for 3 min; 25 cycles of 94°C for 45 s, 50°C for 60 s, 72°C for 90 s, and final

elongation at 72°C for 10 min. The triplicate amplified samples were pooled and then purified

and normalized using the SequalPrep™ Normalization Plate Kit (Invitrogen Corporation,

Carlsbad, Canada). Finally, multiplexed paired-end sequencing was carried out in the Illumina

MiSeq (SY-410-1003) platform using earth microbiome primers (38).

Sequence Data Analysis & Statistics

The paired-end sequences obtained from the Illumina MiSeq were demultiplexed with

QIIME 2 and converted into individual sequence fastq files for each sample (37,39). The rest of

the sequence processing was carried out in R (Version 4.2.2) using the DADA2 package

(Version 1.24.0) (ref. 40). The reads were processed in R using the following command in

DADA2; `filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen=c(145,145), minLen=50, maxN=0,

maxEE=c(2,2), truncQ=2, rm.phix=TRUE, compress=TRUE, multithread=TRUE). `

De novo OTU (Operational Taxonomic Units) picking was performed using DADA2 (37) which

resolves amplicon sequence variants (ASVs) to single nucleotide differences. Chimeras were

removed. The taxonomy was assigned using a 100% cut-off for species-level identification with

the (2022) GTDB 16S rRNA reference database (41). The phylogenetic tree was constructed

using IQTree (42), with 1000 ultrafast bootstraps and Modelfinder to identify the best model. In

this study, the best-fit model was SYM+R10 chosen according to BIC. The data comprising the

OTU table, phylogenetic tree, taxonomy table, and sample meta-data were then parsed through R

studio using the phyloseq package (43). Here, it was processed to remove: OTUs identified as

mitochondrial or chloroplast, which had less than five reads across all samples, and samples that

had fewer than 8000 reads. The packages phytools, phyloseq, microbiome, and ggplot2 were

used for further data analysis.

Diversity Analyses

Alpha diversity was computed using four metrics, Observed, Shannon, Chao1, and

Simpson diversity indices. Beta diversity was calculated using weighted and unweighted

UniFrac, and subsequently Principle Coordinate Analysis (PCoA) plots were constructed using

the first two axes that explain the most variance. Linear mixed effect models using, *lmerTest*

package in R were used to determine changes in alpha and beta diversity over time (lmer

(DiveristyMetric, Genotype*Generation*Biomass Treatment, (1|Generation)), with generation as

a fixed and random effect. Pairwise comparisons were determined with *emmeans* and adjusted p-

value obtained with Tukey.

Adonis

PERMANOVA was carried out for each generation using the adonis2 test. Both weighted

and unweighted UniFrac distance matrices were used to account for the abundance,

presence/absence, and phylogeny of the microbiome. The following model was used

`adonis2(distance.matrix~Biomass Treatment+Genotype, data = meta, permutations=999)`.

Neutral Model

To determine the impact of neutral processes like drift and dispersal or by deterministic

selective forces such as plant genotype and biomass treatments on microbial community

assembly, we carried out the analysis as described in Burns et al. (44), which fits the data to the

neutral model for Prokaryotes from Sloan et al. (45). The following design was used,

sncm.fit (spp=generation(n)_Biomass Treatment, pool = generation(n), stats=T), where n is the

nth generation and 'Biomass Treatment' is high or low biomass samples.

Differential Abundance

There are several challenges in estimating differentially abundant (DA) taxa in

microbiome data, these include high variability in abundance, zero-inflated data, and the

compositional nature of the data. Nearing et al. (46) compared several methods across 38

datasets and found that different methods often identified different sets of DA taxa. The DA taxa

are estimated using ANCOMBC2 (47), a conservative and robust approach. ANCOMBC2

incorporates bias correction, effectively addressing sampling-specific and sequencing biases

present in the data. This feature ensures that the analysis is not skewed by any systematic errors

introduced during the sampling or sequencing process. It also conducts a sensitivity analysis for

the pseudo-count addition to assess the impact of different pseudo-count values on zero counts

for each taxon.The analysis was run on L*er* and Cvi samples, from OTU to Family taxonomic

levels with (fixed effect = (Biomass Treatment + Genotype + Generation), group = Biomass

Treatment).

200   **Results**

Plant Biomass

The ecosystem selection experiment was carried out for the trait plant biomass on two

different accessions of *Arabidopsis thaliana,* L*er* and Cvi for eight generations. The two selected

treatments were high and low above-ground biomass. The absolute values for the mean biomass

of plants (n=50) from each genotype and treatment changed substantially through the course of

eight generations of the experiment. Despite the stochastic fluctuations in biomass caused by

REV, the high biomass selected lines were always the same or greater than those of the low

selected biomass for both genotypes (Figures 2.1a and 2.1b). Due to the significant drop in

biomass between generations 1 and 2, a sterile potting soil uninoculated control was planted in

all subsequent generations to serve as a reference for REV. The uninoculated control does show

that the drastic drop in biomass after generations 1 and 2, was not a result of the inoculated soil

but other largely unknown factors and growing conditions. Above-ground biomass in terms of

deviations from the mean removes some of these REVs and gives a clearer representation of the

difference in phenotype seen in every generation (Figures 2.2a and 2.2b). Significant differences

between the low and high biomass treatments become apparent from generation 4 onwards.

Microbial community Composition

After preprocessing in DADA 2, a total of 1,897,732 reads, with an average of 12822

reads per sample were obtained. Singletons and chimeras were removed during pre-processing.

The most well-represented phyla were *Proteobacteria* (36.2%), *Bacteroidetes* (10.3%),

Planctomycetes (7.8%), and *Actinobacteria* (7.6%) (Figure 3). The most well-represented

Classes identified in the dataset were, *Alphaproteobacteria* (24.7%), *Gammaproteobacteria*

(11.4%), *Bacteroidia (10.1%)*, and *Verrucomicrobiae* (6.2%) (Figure 4). The relative abundance

is typical of microbiome data, heavily weighted to a few abundant groups with a long tail of rare

taxa.

Alpha Diversity

We used four metrics to assess alpha diversity (Figure 5). Observed ($F_{7,160}$ =333.054,

P<0.001) and Chao1 ($F_{7,160}$ =195.54, P<0.001) estimate species richness, and both showed a

statistically significant decrease in OTUs (1477 to ~580) and richness across generations. The

Shannon diversity index ($F_{7,160}$ =82.23, P<0.001), which is more sensitive to the difference in

abundance, also showed this steady decrease across generations. The Simpson diversity ($F_{7,160}$

=14.46, P<0.001), which is more sensitive to evenness, exhibits fluctuations across successive

generations, albeit without any distinct trend. In addition, a linear mixed-effect model shows

significant interactions between genotype:biomass treatment, genotype:generation and

genotype:biomass treatment:generation, for both Shannon and Simpson diversity indices

(Supplementary Tables 2.7 and 2.8). No significant interaction terms were found for Observed

and Chao1, which do not consider the abundance of OTUs in their calculations.

Beta Diversity

Principal coordinate analyses (PCoA) of the unweighted UniFrac distances were plotted

with the first two axes that captured the most variance in the data, from 13.4% in generation 1 to

greater than 52.3% in generation 8 (Figure 6). UniFrac distance considers the evolutionary

relationships between taxa, which makes it more biologically meaningful than other distance

metrics that do not consider the evolutionary history. Unweighted UniFrac considers

presence/absence of taxa. The microbiomes associated with the treatments and genotypes are

very similar in generation 1, but begin to diverge by generation 2. By generation 3, the

microbiomes of the L*er* and Cvi genotypes have diverged and by generation 5 the high and low

biomass treatments appear to have further distinguished the microbial communities. The

divergence of the microbial community in generations 4 and 5 caused by the high/low biomass

treatments also aligns with the onset of notable differences in the above-ground plant biomass

within these generations.

The complex interplay between genotype and biomass treatments was modeled using a

PERMANOVA test with both weighted and unweighted UniFrac distances to account for

differences in abundance and presence-absence of OTUs respectively. The resulting R^2 values

for genotype and biomass treatment and residuals as a proxy for REVs were plotted (Figure 7a

and 7b). For both weighted and unweighted UniFrac, the variance in microbial community

explained by both genotype and biomass treatment increased significantly over the eight

generations. Despite observing a decrease in the variance accounted for by the residuals, which

could potentially signify a decline in the influence of REVs, they accounted for ~50 percent of

the variability within the microbial community in generation 8. For all metrics, weighted

UniFrac exhibits more fluctuations compared to the unweighted UniFrac. Weighted UniFrac

analyses found that genotype ($F_{1,39}$=1.55, P<0.05) and biomass treatment ($F_{1,39}$=1.33, P<0.001)

explained 5.7% and 13.7%, respectively, of the dissimilarity between microbiomes in generation

1. By generation 8, genotype ($F_{1,14}$=5.14, P<0.05) and biomass treatment ($F_{1,14}$=5.48, P<0.001)

explain 24% and 22%, respectively, of the dissimilarity between microbiomes (Supplementary

Tables 2.9 to 2.10). For unweighted UniFrac, the influence of genotype on the microbial

community increased from 3.3% in generation 1 ($F_{1,39}$=1.33, P<0.01) to 26% in generation 8

($F_{1,39}$=5.13, P<0.001), and for biomass increased from 3.8% in generation 1 ($F_{1,14}$=1.55 ,P<.001)

to 13% in generation 8 ($F_{1,14}$=2.58, P<0.05) (Supplementary Tables 2.11 and 2.12).

Sloan Neutral Model

Many forces can alter microbial community structure and dynamics and in our study

these can be divided into selective forces such as the plant genotype being colonized and the

biomass selection process, or neutral, stochastic forces innate in any experiment and

environment. The Sloan neutral model was fit to the data to assess the importance of selective

versus neutral drivers of change in the experiment. If neutral processes were the driving force in

the microbial community assembly, the null hypothesis would be that all generations of the

experiment would fit the model equally well.   The model fit to neutrality represented by the R^2

value decreases from 0.795 in generation 1 to 0.58 in generation 8 demonstrating the increasing

importance of selective drivers through the course of the experiment (Figure 8).

Differential Abundance

Differentially abundant (DA) taxa that distinguished between the high and low biomass

treatments were determined using ANCOMBC2 at multiple taxonomic levels, ranging from the

OTU to the Class. Taxa lacking classification at the Family level were denoted by the subsequent

identifiable taxonomic tier. The OTUs that were present in higher (red) and lower (blue)

abundance in the high biomass treatment are identified (Figure 9). Among the DA taxa in this

study, several are known to benefit plants. The Class *Bacilli* promote plant growth (48). The

Class *Gemmatimonadetes* are positively associated with vegetation restoration, plant richness,

and soil nutrients (49). *Cytophagacea* are chemoorganotrophs and important for remineralizing

organic materials into micronutrients. They could support both mycelial growth and plant

nutrition (50).

Enterobacteriaceae, *Paenibacillaceae*, and *JACDCH01* were all present in higher

abundance in the high biomass treatments of both L*er* and Cvi. Most *Paenibacillaceae* members

predominantly inhabited soil, frequently in close association with plant roots (51). These

rhizobacteria were known to play a significant role in enhancing plant growth and possessing

potential applications in agriculture and *Enterobacteriaceae* had been reported to enhance plant

growth (52–54).

Additionally, *Order_NS11-12g*, *BJHT01*, *JACDCH01*, and *UBA6156* were more

abundant in Cvi under high biomass conditions and less abundant in L*er* under high biomass

conditions. The majority of these were uncultivated or candidate taxa. Two alternative families

in the Order *Rhizobiales* appeared differentially associated with the two plant genotypes:

Labraceae with Ler and *Rhizobiaceae* with Cvi (55). These could indicate genotype-specific

interactions with different members of the microbial community.

Discussion

We applied artificial ecosystem selection for eight generations, in *Arabidopsis thaliana*

L*er* and Cvi to select soil microbiomes associated with higher or lower biomass of the host. In

contrast to some previous studies (27,33), we did not fertilize plants thus maintaining nutrient

limitation, and promoting the plant's interaction with the microbiome (56). Over the course of

eight generations, a response to the selection, most apparent after generations 4 and 5, was

observed both in the microbiome characteristics that shifted gradually and in the plant biomass.

Microbiome selection noticeably influenced plant biomass despite large phenotypic variation

from one generation to the next. Stochasticity in growth through generations may be attributed to

random environmental variations (REVs) as indicated by Swenson et al. (28). This was observed

in the uninoculated sterile potting soil reference that showed similar patterns of variability as the

inoculated treatments across generations (Figure 1). Ecological variability among generations

even under controlled growth chamber conditions is a common characteristic of studies (57) and

may be due to minute fluctuations in growth chamber conditions or batch and age of potting soil.

We observed a gradual decline in microbial species richness during the initial generations

of the experiment. This was evident from the observed OTU counts, as well as the Chao1 and

Shannon diversity indices. This pattern is consistent with findings from other studies (32–34),

highlighting the impact of selection pressures as the microbial communities adapted to the host

plant environment. However, in our study which continued for about twice the number of

generations of these earlier studies, we observed a stabilization of richness and alpha diversity in

the later generations.

To further understand the complex interplay between the forces of directed selection (plant

genotype and biomass treatment), we did a PCoA. Over successive generations, the results

demonstrate a strengthening effect of plant genotypes and biomass treatment on the microbial

community (Figure 6). This result was modeled using PERMANOVA , for both weighted and

unweighted UniFrac distances. Both weighted and unweighted UniFrac show marked increases in

the proportion of variance explained by genotype and biomass treatment, with unweighted UniFrac

exhibiting larger increases from generation to generation (Figure 7). Both metrics consider

phylogenetic relatedness. Weighted UniFrac considers the absolute abundance of OTUs and is

generally used to study changes in microbial community structure. Unweighted UniFrac considers

only the presence/absence of OTUs and is generally used to study changes in microbial community

composition. The results suggest that variance in microbiome composition as explained by

genotype consistently increases, whereas variance in microbiome structure as explained by

genotype, fluctuates more with changes in abundant taxa.

Here we would like to highlight that pronounced shifts in biomass between the high/low

biomass treatments, alpha diversity, and beta diversity all occur around generations 4 and 5 (Figure

1). This marks the point at which the changes in alpha diversity stabilize (Figure 5) and the

divergence between microbial communities by high/low biomass treatments in the PCoA plots

becomes more pronounced (Figure 6). It is possible that over the first four to five generations, the

starting microbial community underwent a period of restructuring before it stabilized and formed

four distinct communities under selection by high/low biomass treatments and by plant genotype.

We propose a complex interaction between plants and their associated microbiome ensued, where

differences in root exudation patterns between the two genotypes presumably established

associations with microbes in the early generations. Simultaneously, during this period, the

biomass treatment likely promoted host-specific microbe-mediated interactions that modulate

plant biomass.   This restructuring of the microbial community is driven by both structural

(abundance of OTUs) and compositional (presence/absence of OTUs) effects in the microbiome

(Figure 7). Differences in the abundance of taxa driven by the biomass and genotype selection

pressures, was a larger contributor to the generation to generation variation during the experiment.

Changes in the assembly of microbial communities can arise from either selective pressure,

like the plant genotype or biomass treatment observed in this experiment, or stochastic processes,

like minute changes in growth chamber humidity or microenvironment. To gain deeper insights

into the influence of selection on microbial community assembly, a neutral model was fit to the

data (44,45). Interestingly, this revealed a progressive decline in fit to neutrality, indicated by

decreasing R^2 values and increasing AIC values (Figure 8). This suggests an increasing influence

of selective forces such as biomass treatment and genotype over multiple generations of ecosystem

selection. This was also previously observed by Morella et al. (34).

Previous studies in *Arabidopsis thaliana* have often detected only a weak genotypic effect

on the rhizosphere microbiome (23,24,58). However, a majority of these studies are focused on a

single growth cycle. Insights from plant-soil feedback research emphasize that the interactions

between plants and their microbiomes can be modified by consecutive growth cycles of the same

plant species in its respective soil (25,26). Recent research has presented conflicting findings on

the influence of host genotype (33,34,59). The different experimental designs, communities being

assessed and the length of these experiments, alongside high stochastic environmental variation

that is clearly associated with these microbial studies likely account for the conflicting findings.

In our investigation, we find that during the initial stages, genotype and biomass treatments have

modest but significant impacts. Over time, the plant genotype and biomass treatments have an

increasing influence. Together these selective forces explained ~40% of the variation in the

microbial community composition in these later generations.

A key aim of microbiome engineering has been to develop host-specific microbial

communities that impart lasting beneficial effects on the plant host (27,29,32,60). Here we show

the enrichment of some common taxa in high biomass treatments but also some genotype specific

changes. Despite starting with the same soil, only three common families were enriched in the

high biomass treatment samples of both the L*er* and Cvi genotypes. In addition, well-known plant

growth promoting rhizobacteria were enriched in both L*er* and Cvi high biomass treatment

samples, however possible genotype-specific interactions are involved. In the Order *Rhizobiales*,

Labraceae was enriched in L*er,* while *Rhizobiaceae* was enriched in Cvi (55). Moreover, in high

biomass conditions, Cvi had increased levels of the taxa *Order_NS11-12g*, *BJHT01*, *JACDCH01*,

and *UBA6156*, while L*er* had decreased levels of these. These data suggest genotype specific plant

genetic control. Multiple studies have shown variations in the structure of the rhizosphere

microbiome, even within closely related plant genotypes, highlighting the importance of genotype-

specific root exudates in forming associations with the corresponding microbiome (61–65).

Further exploration must prioritize highlighting the mechanisms that uphold these interactions

through the utilization of comparative metagenome and metatranscriptome analyses. This will play

a part in decoding chemical crosstalk, which could eventually foster interactions between plants

and microbes, leading to an overall enhancement in plant fitness.

The relationship between plants and microorganisms in the rhizosphere involves complex

and diverse interactions, which influence crucial ecological and physiological processes. These

interactions can be mutualistic, competitive, or antagonistic in nature. Many biotic and abiotic

factors act in concert to influence the dynamic rhizosphere microbiome. Gaining insights into how

these factors influence the assembly of microbial communities is crucial for obtaining targeted and

long-lasting advantages for plants. Nevertheless, establishing a strong and host-specific microbial

community that consistently provides beneficial effects continues to be a persistent challenge. We

have shown that, despite stochastic fluctuation due to REVs, it is possible to select for microbial

communities that impact biomass in a host-specific manner within four generations, addressing

this key challenge in microbial community engineering (32,60). The rhizosphere microbiome that

evolves under plant-mediated selection offers better chances of survivability and efficacy when

applied as inoculum to the plant (66,67). This study enhances our understanding of the temporal

dynamics involved in microbial assembly in the rhizosphere and has implications for sustainable

agriculture, evolution, and ecology.

                                                    FIGURES

Figure 1. Results of the ecosystem selection on above-ground plant biomass for genotypes L*er* (a) and Cvi (b) of Arabidopsis thaliana. Dark lines represent the high biomass treatment, and light lines represent the low biomass treatment. Each point represents the mean of n=50 microcosms. The mean biomass of the selection lines differed for several generations of both genotypes (t.test p-value * < 0.05; **<0.01;*** <0.001).

428
429

Figure 2 By plotting deviations from the mean, the influence of REVs on above-ground biomass is mitigated for genotypes L*er* (a) and Cvi (b). The dark lines denote high biomass treatments, while the light lines represent low biomass treatments. Each data point is the average of n=50 microcosms. (t.test p-value * < 0.05;**<0.01; *** <0.001).

Figure 3. Taxonomic distribution of the microbial community represented in terms of the relative abundance at the Phylum level.   The plot shows eight generations (1-8) for the high and low biomass treatments of both the Cvi (left) and L*er* (right) genotypes of *Arabidopsis thaliana*. The microbial community is dominated by *Proteobacteria*, *Bacteroidetes*, *Verrucomicrobia, and Planctomycetes* which together comprise nearly 73% of the bacterial community.

Figure 4.   Taxonomic distribution of the microbial community is represented in terms of the relative abundance at the Class level. The plot shows eight generations (1-8) for the high and low biomass treatments of both the Cvi (left) and L*er* (right) genotypes of *Arabidopsis thaliana*. Ten samples were assayed for generation 1 experiments while five were done for all subsequent generations. The most represented classes across all samples were *Alphaproteobacteria* (24.7%), *Gammaproteobacteria* (11.4%), *Bacteroidia* (10.1%), and *Verrucomicrobiae* (6.2%).

Figure 5. Over the course of eight generations, four Alpha diversity indices were computed, including Observed, Chao1 estimate, Shannon, and Simpson indices. Results indicate a significant decline in both Observed and Chao1 estimate species richness. Additionally, the Shannon index was more responsive to species richness and exhibited a noteworthy decrease from generation 1 to 8. In contrast, the Simpson index was more sensitive to evenness and demonstrated fluctuations across successive generations without any distinct trend. (wilcox test *<0.05, **<0.01, *** <0.001).

Figure 7. The R^2 value obtained from PERMANOVA using adonis2 test on generations 1 through 8 for (a) weighted and (b) unweighted-UniFrac distances. The resulting R^2 and variance explained by residuals values for the biomass treatment (brown), genotype (blue) and residuals (black) are plotted. Residuals are plotted as a proxy for random environmental variations (REVs). There is a marked increase in the influence of both genotype and biomass treatment on differences in the microbial community from generations 1 through 8. Despite observing a decrease in the variance accounted for by the residuals, which could potentially signify a decline in the influence of REVs, the model elucidates more than 50 percent of the variability within the microbial community in the 8th generation. These differences are more pronounced in the weighted vs unweighted UniFrac. Model (distance.matrix~biomass treatment +genotype).

498
499
500
501

Figure 8. The occupancy (prevalence of OTU in samples) plotted against log(10) abundance is depicted using the Burns model. OTUs that are neutral, i.e., not selected for or against, are indicated in black. Green and blue-colored OTUs signify positive and negative selection, respectively. This plot highlights the presence of numerous microbes that undergo selection throughout the experiment.

Figure 9. ANCOMBC2 results are illustrated as log fold changes at the family level for high biomass vs low biomass samples. Each accession listed in the table represents a single OTU in either the L*er* or Cvi microbial community identified to the family level if possible. The scale indicates log fold change, red and blue for positive and negative fold change in the high biomass samples respectively. Well-known plant growth-promoting bacteria, such as *Paenibacillaceae, Bacilli, Labraceae, Rhizobiaceae* and *Bdellovibrio* are present in higher abundance. A vast majority of the other taxa identified are uncultivated and have little to no information on them.

References

1. Gouda S, Kerry RG, Das G, Paramithiotis S, Shin HS, Patra JK. Revitalization of plant
growth promoting rhizobacteria for sustainable development in agriculture. Microbiol Res.
2018;206.

2. Jones DL, Nguyen C, Finlay RD. Carbon flow in the rhizosphere: Carbon trading at the soil-
root interface. Plant Soil. 2009;321(1–2):5–33.

3. Dennis PG, Miller AJ, Hirsch PR. Are root exudates more important than other sources of
rhizodeposits in structuring rhizosphere bacterial communities? FEMS Microbiol Ecol.
2010;72(3):313–27.

4. Hartmann A, Schmid M, van Tuinen D, Berg G. Plant-driven selection of microbes. Plant
Soil. 2009;321(1–2):235–57.

5. Chaparro JM, Sheflin AM, Manter DK, Vivanco JM. Manipulating the soil microbiome to
increase soil health and plant fertility. Biol Fertil Soils. 2012;48(5):489–99.

6. Marschner P, Marschner P, Crowley D, Crowley D, Yang CH, Yang CH, et al. Development
of specific rhizosphere bacterial communities in relation to plant species, nutrition and soil
type. Plant Soil. 2004;261(1–2):199–208.

7. Garbeva P, van Veen J a, van Elsas JD. Microbial diversity in soil: selection microbial
populations by plant and soil type and implications for disease suppressiveness. Annu Rev
Phytopathol. 2004;42(29):243–70.

8. Tkacz A, Cheema J, Chandra G, Grant A, Poole PS. Stability and succession of the
rhizosphere microbiota depends upon plant type and soil composition. ISME J. 2015;1–11.

9. Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and function
of microbial communities in the rhizosphere. FEMS Microbiol Ecol. 2009;68(1):1–13.

10. Jin H, Yang X, Liu R, Yan Z, Li X, Li X, et al. Bacterial community structure associated
with the rhizosphere soils and roots of Stellera chamaejasme L. along a Tibetan elevation
gradient. Ann Microbiol. 2018 May 1;68(5):273–86.

11. Barbosa Lima A, Cannavan FS, Navarrete AA, Teixeira WG, Kuramae EE, Tsai SM.
Amazonian Dark Earth and Plant Species from the Amazon Region Contribute to Shape
Rhizosphere Bacterial Communities. Microb Ecol. 2015 May 1;69(4):855–66.

12. Marques JM, da Silva TF, Vollu RE, Blank AF, Ding GC, Seldin L, et al. Plant age and
genotype affect the bacterial community composition in the tuber rhizosphere of field-grown
sweet potato plants. FEMS Microbiol Ecol. 2014;88(2):424–35.

13. Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. A Synthetic Community
Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota. PLoS Genet.
2014;10(4).

14. Zancarini A, Mougel C, Voisin AS, Prudent M, Salon C, Munier-Jolain N. Soil Nitrogen
Availability and Plant Genotype Modify the Nutrition Strategies of M. truncatula and the
Associated Rhizosphere Microbial Communities. PLoS ONE. 2012;7(10).

15. Wei S, Liu B, Ni K, Ma L, Shi Y, Leng Y, et al. Rhizosphere Microbial Community Shows a
Greater Response Than Soil Properties to Tea (Camellia sinensis L.) Cultivars. Agronomy.
2023 Jan 1;13(1).

16. Yuan J, Chaparro JM, Manter DK, Zhang R, Vivanco JM, Shen Q. Roots from distinct plant
developmental stages are capable of rapidly selecting their own microbiome without the
influence of environmental and soil edaphic factors. Soil Biol Biochem.
2015;89(AUGUST):206–9.

17. Houlden A, Timms-Wilson TM, Day MJ, Bailey MJ. Influence of plant developmental stage
on microbial community structure and activity in the rhizosphere of three field crops. FEMS
Microbiol Ecol. 2008;65(2):193–201.

18. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the
core Arabidopsis thaliana root microbiome. Nature. 2013 Aug 1;488(7409):86–90.

19. Bulgarelli D, Rott M, Schlaeppi K, Ver E, Themaat LV, Ahmadinejad N, et al. Revealing
structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature.
2012;488(7409):91–5.

20. Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, Gross S, Clingenpeel S, Woyke T, et
al. Plant compartment and biogeography affect microbiome composition in cultivated and
native *Agave* species. New Phytol. 2016 Jan;209(2):798–811.

21. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, et al.
Structure, variation, and assembly of the root-associated microbiomes of rice. Proc Natl
Acad Sci U S A. 2015 Feb 24;112(8):E911–20.

22. Wagner MR, Lundberg DS, Del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T. Host
genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat
Commun. 2016 Jul 12;7.

23. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the
core Arabidopsis thaliana root microbiome. Nature. 2012;488(7409):86–90.

24. Schlaeppi K, Dombrowski N, Oter RG, Ver Loren Van Themaat E, Schulze-Lefert P.
Quantitative divergence of the bacterial root microbiota in Arabidopsis thaliana relatives.
Proc Natl Acad Sci U S A. 2014;111(2):585–92.

25. Hu L, Robert CAM, Cadot S, Zhang X, Ye M, Li B, et al. Root exudate metabolites drive
plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. Nat
Commun. 2018 Dec 1;9(1).

26. Van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, et al.
Plant-soil feedbacks: The past, the present and future challenges. J Ecol. 2013
Mar;101(2):265–76.

27. Panke-buisse K, Poole AC, Goodrich JK, Ley RE, Kao-kniffin J. Selection on soil
microbiomes reveals reproducible impacts on plant function. ISME J. 2014;1–10.

28. Swenson W, Wilson DS, Elias R. Artificial ecosystem selection. Proc Natl Acad Sci U S A.
2000 Aug 1;97(16):9110–4.

29. Mueller UG, Juenger TE, Kardish MR, Carlson AL, Burns KM, Edwards JA, et al. Artificial
Selection on Microbiomes To Breed Microbiomes That Confer Salt Tolerance to Plants
METHODS AND PROTOCOLS [Internet]. 2021. Available from:
https://journals.asm.org/journal/msystems

30. Lau JA, Lennon JT. Rapid responses of soil microorganisms improve plant fitness in novel
environments. Proc Natl Acad Sci. 2012;109(35):14058–62.

31. terHorst CP, Lennon JT, Lau JA. The relative importance of rapid evolution for plant-
microbe interactions depends on ecological context. Proc R Soc B Biol Sci. 2014 Jun
22;281(1785):20140028.

32. Jochum MD, McWilliams KL, Pierson EA, Jo YK. Host-mediated microbiome engineering
(HMME) of drought tolerance in the wheat rhizosphere. PLoS ONE. 2019 Dec 1;14(12).

33. Cordovez V, Rotoni C, Dini-Andreote F, Oyserman B, Carrión VJ, Raaijmakers JM.
Successive plant growth amplifies genotype-specific assembly of the tomato rhizosphere
microbiome. Sci Total Environ. 2021 Jun 10;772.

34. Norma M. Morella, Francis Cheng-Hsuan Weng, Pierre M. Joubert, C. Jessica E. Metcalf,
Steven Lindow, Britt Koskella. Successive passaging of a plant-associated microbiome
reveals robust habitat and host genotype-dependent selection. PNAS. 2020 Jan 14;117(2).

35. Parada AE, Needham DM, Fuhrman JA. Every base matters: Assessing small subunit rRNA
primers for marine microbiomes with mock communities, time series and global field
samples. Environ Microbiol. 2016 May 1;18(5):1403–14.

36. Apprill A, Mcnally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R
gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol.
2015 Jun 4;75(2):129–37.

37. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-
throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms.
ISME J. 2012 Aug;6(8):1621–4.

38. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal
catalogue reveals Earth's multiscale microbial diversity. Nature. 2017 Nov 1;551(7681):457.

39. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al.
626   Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.
627   Nat Biotechnol. 2019 Aug 1;37(8):852–7.

628   40. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-
629   resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jun
630   29;13(7):581–3.

631   41. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly
632   classification with the genome taxonomy database. Bioinforma Oxf Engl. 2022 Nov
633   30;38(23):5315–6.

634   42. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al.
635   IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic
636   Era. Mol Biol Evol. 2020 May 1;37(5):1530–4.

637   43. Bunn AG. A dendrochronology program library in R (dplR). Dendrochronologia. 2008 Oct
638   1;26(2):115–24.

639   44. Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, et al. Contribution
640   of neutral processes to the assembly of gut microbial communities in the zebrafish over host
641   development. ISME J. 2016;10(3):655–64.

642   45. Sloan WT, Woodcock S, Lunn M, Head IM, Curtis TP. Modeling taxa-abundance
643   distributions in microbial communities using environmental sequence data. In: Microbial
644   Ecology. 2007. p. 443–55.

645   46. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al.
646   Microbiome differential abundance methods produce different results across 38 datasets. Nat
647   Commun. 2022 Dec 1;13(1).

648   47. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat
649   Commun. 2020 Dec 1;11(1).

650   48. Radhakrishnan R, Hashem A, Abd Allah EF. Bacillus: A biological tool for crop
651   improvement through bio-molecular changes in adverse environments. Front Physiol. 2017
652   Sep 6;8(SEP).

653   49. Mujakić I, Piwosz K, Koblížek M. Phylum Gemmatimonadota and Its Role in the
654   Environment. Microorganisms. 2022 Jan 1;10(1).

655   50. Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E. The Prokaryotes: A
656   Handbook on the Biology of Bacteria: Archaea. Bacteria: Firmicutes, Actinomycetes. The
657   Prokaryotes: A Handbook on the Biology of Bacteria. 2006.

658   51. Grady EN, MacDonald J, Liu L, Richman A, Yuan ZC. Current knowledge and perspectives
659   of Paenibacillus: A review. Microb Cell Factories. 2016 Dec 1;15(1).

52. Ramesh A, Sharma SK, Sharma MP, Yadav N, Joshi OP. Plant Growth-Promoting Traits in
Enterobacter cloacae subsp. dissolvens MDSR9 Isolated from Soybean Rhizosphere and its
Impact on Growth and Nutrition of Soybean and Wheat Upon Inoculation. Agric Res.
2014;3(1):53–66.

53. Kämpfer P, Ruppel S, Remus R. Enterobacter radicincitans sp. nov., a plant growth
promoting species of the family Enterobacteriaceae. Syst Appl Microbiol. 2005 Apr
22;28(3):213–21.

54. Khalifa AYZ, Alsyeeh AM, Almalki MA, Saleh FA. Characterization of the plant growth
promoting bacterium, Enterobacter cloacae MSR1, isolated from roots of non-nodulating
Medicago sativa. Saudi J Biol Sci. 2016;23(1):79–86.

55. Harkes P, van Steenbrugge JJM, van den Elsen SJJ, Suleiman AKA, de Haan JJ, Holterman
MHM, et al. Shifts in the Active Rhizobiome Paralleling Low Meloidogyne chitwoodi
Densities in Fields Under Prolonged Organic Soil Management. Front Plant Sci. 2020 Jan
10;10.

56. Meier MA, Xu G, Lopez-Guerrero MG, Li G, Smith C, Sigmon B, et al. Association
analyses of host genetics, root-colonizing microbes, and plant phenotypes under different
nitrogen conditions in maize. eLife. 2022;11.

57. Porter AS, Evans-Fitz.Gerald C, McElwain JC, Yiotis C, Elliott-Kingston C. How well do
you know your growth chambers? Testing for chamber effect using plant traits. Plant
Methods. 2015;11(1).

58. Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y, et al. Structure and
function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe.
2015;17(3):392–403.

59. Brown SP, Grillo MA, Podowski JC, Heath KD. Soil origin and plant genotype structure
distinct microbiome compartments in the model legume Medicago truncatula. Microbiome.
2020 Sep 28;8(1).

60. Basu A, Prasad P, Das SN, Kalam S, Sayyed RZ, Reddy MS, et al. Plant Growth Promoting
Rhizobacteria (PGPR) as Green Bioinoculants: Recent Developments, Constraints, and
Prospects. Sustainability. 2021 Jan 22;13(3):1140.

61. Micallef S, Colón-Carmona A. Genetic and Developmental Control of Rhizosphere Bacterial
Communities. In: de Bruijn FJ, editor. Molecular Microbial Ecology of the Rhizosphere
[Internet]. 1st ed. Wiley; 2013 [cited 2023 Aug 20]. p. 257–63. Available from:
https://onlinelibrary.wiley.com/doi/10.1002/9781118297674.ch24

62. Micallef SA, Channer S, Shiaris MP, Colón-Carmona A. Plant age and genotype impact the
progression of bacterial community succession in the Arabidopsis rhizosphere. Plant Signal
Behav. 2009 Aug;4(8):777–80.

63. Lee S, I. Sergeeva L, Vreugdenhil D. Natural variation of hormone levels in *Arabidopsis*
roots and correlations with complex root architecture: Natural variation of phytohormones in
Arabidopsis roots. J Integr Plant Biol. 2018 Apr;60(4):292–309.

64. Matsushima C, Shenton M, Kitahara A, Wasaki J, Oikawa A, Cheng W, et al. Multiple
analysis of root exudates and microbiome in rice (Oryza sativa) under low P conditions.
Arch Microbiol. 2021 Nov 1;203(9):5599–611.

65. Huang XF, Chaparro JM, Reardon KF, Zhang R, Shen Q, Vivanco JM. Rhizosphere
interactions: Root exudates, microbes, and microbial communities1. Botany.
2014;92(4):267–75.

66. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant–microbiome interactions: from
community assembly to plant health. Nat Rev Microbiol. 2020;18(11).

67. Gu S, Wei Z, Shao Z, Friman VP, Cao K, Yang T, et al. Competition for iron drives
phytopathogen control by natural rhizosphere microbiomes. Nat Microbiol. 2020 Aug
1;5(8):1002–10.

710