Persistence of a potential pathogen on a plant microbiome-Does seed sanitization matter?

Anirudha R. Dixit, Christina LM. Khodadad, Mary E. Hummerick, Cory J. Spern, LaShelle E. Spencer, Jason A. Fischer, Aaron B. Curry, Jennifer L. Gooden, Gretchen J. Maldonado Vazquez, Raymond M. Wheeler, Gioia D. Massa, Matthew W. Romeyn

Introduction:

Microbial interactions on, in, and around the seeds can have profound effects on plant growth, development, and productivity. These interactions can be casual or intimate in nature, but ultimately they all contribute in varying degrees to an ever-evolving microbiome. Plant microbiomes have been investigated over decades and new data continue to reveal how microbiomes play an important role in the plant's success. Many eubacteria and fungi have been found to have a symbiotic relationship with plants and other microorganisms. The source of the plant microbiome on adult plants is provided by the seed via vertical transmission to leaf, root, flowers or fruit (1)(Nelson, 2018), as well as the surrounding environment.

Plant-pathogens can be associated with the plant microbiome, some broad ranging, infecting multiple plant species, and others being host specific. Damage to the hosts can be significant causing economic and yield impacts (2) (Mansfield, et al., 2012). Examples such as *Pseudomonas, Ralstonia* and *Agrobacterium* pathovars can cause significant damage in crop plants and all have economic impacts (2) (Mansfield et al., 2012). However, studies of these genera have provided scientific breakthroughs on plant diseases. Plant tissue also has the potential to carry non-plant pathogen associations.

In recent decades, leafy greens have been associated with large outbreaks of food-borne illness including enterohemorrhagic *Escherichia coli* (*E. coli)* infections (3, 4). Enteric human pathogens on produce are introduced through a variety of sources and production steps from farm to market. Studies have shown that *E. coli* can colonize and persist in low numbers on plant surfaces (5–7) and that colonization depends on a variety of abiotic and biotic factors (8, 9). The impact of indigenous plant microbial communities on the

persistence of human pathogens on plant surfaces and the inverse is unclear, and the subject of recent scientific investigations (10–12). It has been shown that native microflora may play a role in pathogen suppression on fruits and vegetables leading to studies on the development of biological control inoculants to inhibit the proliferation of human pathogens on produce (13–15) (Santhanam et al., 2015; Bulgarelli et al., 2013; Toju et al., 2018).

NASA has been growing plants on the International Space Station (ISS) to supplement the packaged astronaut diet since 2014, and food safety is critical. Plant tissue and rooting pillows returned from ISS have been compared to parallel ground controls to evaluate the presence of microbes from a food safety point of view (16) (Khodadad et al, 2020). These initial studies utilized the 16S rRNA gene and the interspatial transcribed sequence (ITS) gene to identify the community players. However, their role as members of these microbiomes remains a mystery. The degree to which crop cultivation in closed, isolated environments like the ISS may have on the development of the plant microbiome and consequent human pathogen proliferation is not understood.

The aim of this study was to investigate the microbiomes of the plant's phyllosphere and rhizosphere when challenged with a potential human pathogen, *Escherichia coli,* strain ATCC 11775 (*E. coli*) and grown under simulated ISS environmental conditions. Seeds from two leafy green crops that have previously grown on the ISS, specifically red romaine lettuce, *Lactuca sativa* cv 'Outredgeous,' and Mizuna mustard, *Brassica rapa* cv *japonica,* were exposed to *E*. *coli.* During a time-course study, using bacterial plate counts, quantitative polymerase chain reaction (qPCR), and 16S rRNA amplicon sequencing, survival of the pathogen and characterization of the plant community were traced from

seeds to the mature plants. Recent advances in sequencing technologies and parallel omics methods continue to increase our knowledge of community characterization as well as community function, and understanding the microbial communities of the plant-associated microbiome separated into above-ground (phyllosphere) and below-ground (rhizosphere) sections is vital to understanding plant response, impact on plant health, and food safety.

Methods:

Experimental design: The experimental design for this study was a multi-factorial design. It included two crops, red romaine lettuce (*Lactuca sativa* cv. Outredgeous) and mizuna mustard, (*Brassica rapa var. japonica*), effects of seed sanitization (sanitized vs. unsanitized) and *E.coli* treatment (treated vs. non-treated) on the microbial community in plant tissue (leaf and root) through four different time-points (7, 14, 21 and 28 days after planting (DAP)).

Seed sanitization: Crop seeds were sanitized using a bleach/HCl fuming method as described (17, 18). Seeds were batch sanitized by adding 0.5 mL concentrated HCl to 30 mL bleach (5% sodium hypochlorite solution) in a 500 mL container and placing an open petri dish containing the seeds into the bleach container, without submersion, then sealing the container to maintain the chlorine fumes around the seeds. Seeds were sanitized for one hour before being removed from the acid/bleach container for off-gassing overnight (approximately 24 h) in a laminar flow hood. Seed sanitization was confirmed for both bacterial and fungal contaminants as well as germination rate. Aerobic plate counts for

bacteria were confirmed on Trypticase Soy Agar (TSA) plates, while mold/fungi were confirmed on Inhibitory Mold Agar (IMA). Plates were incubated at 30ºC for 24-48 h. Germination rate was determined by the number of seeds (out of 10) that germinated on moist filter paper within that same period.

Seed inoculation: Seeds from each of the two crop plants were prepared in sufficient quantities with four treatments: 1) sanitized seeds exposed to *E. coli*; 2) sanitized seeds without exposure to *E. coli*; 3) unsanitized seeds exposed to *E. coli*; and 4) unsanitized seeds without exposure to *E. coli*. A cell culture of *E. coli* was prepared in trypticase soy broth (TSB) and incubated at 30 °C for 16 h in a shaking incubator. Seeds were inoculated by submersion in a 5 mL culture of 10^6 *E. coli cells,* overnight, at room temperature just prior to planting. Seeds not treated with *E. coli* were submerged in sterile TSB under the same conditions.

Plant cultivation and harvest: Four seeds were planted in 10 cm (4-inch) square pots containing (~500 mL) greens grade, Turface, calcined clay (arcillite) (Ewing Irrigation, St Petersburg, FL, United States) and 7.5 g/L of 18-6-8, type 180 day time-release fertilizer (Nutricote Florikan ESA, Sarasota, FL, United States). Seeds were germinated and grown in controlled environment growth chambers simulating environmental conditions aboard the ISS (RH-50 %, CO_2-3000 ppm, temperature 23 ºC). The photoperiod was set at 16h/8h light dark cycles with fluorescent lights providing 200-300 µmol m^{-2} s^{-1} photosynthetically active radiation (PAR). All plants of the same treatments were contained in a single tray and bottom watered with DI water using an automated watering

system and monitored daily. Pots were thinned to a single plant per pot at day 7 and additional harvests were completed every 7 days until day 28. Each harvest included three plants, chosen at random, separated into leaf and root tissue. All harvests were completed using aseptic techniques to minimize contamination. Plant tissues were separated into leaf and root and placed in pre-weighed 50 mL tubes containing sterile water and glass beads.

Cell count: Plants tissues were placed in a 50 mL sterile, centrifuge tube with 10-20 mL sterile water and 3 mm glass beads then placed in an OMNI BeadRupter for tissue disruption (OMNI International, Kennesaw, GA, United States). All sample weights were determined prior to disruption. Sample extracts were diluted into PBS and 100 µL of appropriate dilutions were plated in duplicate onto trypticase soy agar (TSA) (BBL, Difco, Becton, Dickenson, Franklin Lakes, NJ, United States) and incubated at 30 °C for 48 h before enumeration of colonies. Count data were determined for colony forming units/gram fresh weight (CFU gfw^{-1}) then log transformed.

DNA isolation, qPCR and 16S PCR with Sequencing: The remaining liquid was centrifuged at 13,000 x g and the pellet was processed for DNA isolation using the UltraClean Microbial DNA Isolation Kit per manufacturer's protocol (Qiagen, Carlsbad, CA, United States). DNA was quantified with the Qubit ds DNA assay (Invitrogen Inc, Grand Island, NY) for downstream qPCR, 16S PCR, and sequencing.

Quantitative PCR (qPCR) was performed on the Roche LightCycler 480 using the SYBR Green Master Kit (Roche Diagnostics, Indianapolis, IN, United States) with 200 nM each of the *ycfR* gene primers (5'TAAGCTCCATGTCATTTGCC-3' Forward and 5'TTCCATGGAGGGTATTCGG-3' Reverse) and 5 µL DNA template for a total of up to 2.5 ng of genomic DNA. The *ycfR* gene was selected, as it is present in one copy per cell in this *E. coli* strain. A standard curve was generated with DNA isolated from *E. coli* ATCC 11775 with concentration ranging from 10^{0} to 10^{-7} ng of genomic DNA. The qPCR cycling conditions included a 10 min denaturation cycle followed by 45 amplification cycles (95 ºC for 10 s, 54 ºC for 1 min, and 72 ºC for 30 s). The melt curve was determined at 95 ºC for 5 s, 65 ºC for 1 min then 95 ºC). All standards and samples were run in triplicate.

16S PCR and sequencing: The 16S PCR was completed with custom barcoded 16S rRNA gene primers (V4 region) as described by Kozich et al., 2013 (19). The final PCR master mix contained 1X PCR buffer, 2.25 mM $MgCl_2$, 300 nM dNTPs, 300 nM forward and reverse primers, and 0.25 units of Platinum Taq polymerase (Invitrogen, Grand Island, NY, United States). Each reaction was completed in duplicate with 1 ng of DNA per reaction. The samples were denatured at 95 ºC for 5 min followed by 30 cycles at 95 ºC for 1 min, 58 ºC for 1 min and 72 ºC for 2 min. A final elongation step at 72 ºC for 10 min completed the PCR. Replicates of each PCR reaction were combined and the reaction cleaned with the Min-Elute Cleanup System to remove excess primers and dNTPs (Qiagen, Carlsbad, CA, United States). Amplicon concentration was determined with the Qubit 2.0 broad range ds DNA assay (Invitrogen, Grand Island, NY, United States) and nanomolar concentration calculated for each sample. Samples were diluted

to 4 nM and an equimolar concentration library was created per Illumina's library preparation guidelines. The library was spiked with a 10% PhiX aliquot to add diversity to the library and sequenced on the Illumina MiSeq platform (Illumina Inc., San Diego, CA, United States) with a 500-cycle V2 kit with 250-bp paired ends and FASTQ reads at > 30 quality score.

Sequencing data analysis: Data generated from the16S rRNA amplicon sequencing of community samples were analyzed using QIIME2 (ver. 2020.2) (20) (Bolyen et al., 2019). Demultiplexed raw sequences were imported into the DADA2 denoising algorithm to process the raw sequences into exact sequence variants (ESVs). The DADA2 pipeline performed filtering, dereplication, chimera identification and the merging of paired-end reads (Callahan et al., 2016). The taxonomies were assigned using SILVA 132 rRNA gene database available for use with QIIME2-feature-classifier plugin (21)(Quast et al., 2013). Operational taxonomic units (OTU) not present in at least six samples and less than 50 observations in the entire dataset were excluded from further analysis. In addition, OTUs affiliated to chloroplasts and mitochondria were removed in order to keep only microbial sequences (22, 23) (Beckers et al., 2016; Beirinckx et al., 2020). A differential abundance analysis was carried out by importing essential QIIME2 artifacts using qiime2R (24) (Bisanz, 2018) into a phyloseq (25) (McMurdie and Holmes, 2013) object, followed by the edgeR (26) (Robinson et al., 2010) analysis to identify taxa that are differentially abundant across the time course .

Statistical Analysis: Differences between counts (CFU) from sanitized and unsanitized treatments in the treated and non-treated mizuna tissues ($P < 0.05$) were determined using a two-way ANOVA followed by Tukey's multiple comparison test. Data from lettuce were analyzed using a mixed model as implemented in GraphPad Prism 8. In the absence of missing values, this method gives the same P values and multiple comparisons tests as repeated measures ANOVA. In the presence of missing values (as in lettuce data set), the results can be interpreted like repeated measures ANOVA. Graphical representation was completed in GraphPad (GraphPad Prism version 8.0.0 for Windows. GraphPad Software, San Diego, CA, www.graphpad.com.)

Alpha diversity represented as Shannon Index (27)(Shannon, 1948) and beta diversity represented by Bray–Curtis dissimilarity (28)(Bray & Curtis, 1957) were calculated using the diversity plugin in QIIME2 (20)(Boylen et al., 2019). The Kruskal-Wallis pairwise test was implemented to compare alpha diversity values between each time-point for sanitized and unsanitized samples from leaf and root (29)(Kruskal & Wallis, 1952). A Bray-Curtis dissimilarity matrix was constructed to estimate the global differences in leaf and root samples due to 1) *E.coli* treatment vs. non-treatment, 2) sanitized vs. unsanitized seeds, across all four time-points and visualized via a Principal Coordinate Analysis (PCoA). A Permutational Multivariate Analysis of Variance Using Distance Matrices (adonis) was completed using "qiime diversity adonis". To assess global differences between treatments and the time-course (DAP), option --p-formula "treatment*time-course" was implemented (30)(Anderson, 2001). *P*-values on the PCoA plots indicate statistical significance of the R^2 value. Taxa represented in differential abundance plots represent a statistical significance of $P < 0.05$.

Results

Bacterial population dynamics

Bacterial abundance of the phyllosphere of lettuce and mizuna as indicated by aerobic plate counts per gram fresh weight (APC gfw^{-1}) on TSA generally declined through the growth period from 7 DAP to 28 DAP. On the lettuce leaves, the average APCs at 7 DAP were in the range of log 6-7 CFU gfw^{-1} for plants from sanitized and unsanitized seeds both treated and non-treated with *E.coli.* Bacterial numbers declined approximately 1 log on samples from all the treatments except for the unsanitized seeds treated with *E. coli.* The counts on these samples were log 7.0 CFU gfw^{-1} at 7 DAP declining to significantly lower levels (P=0.03) when compared to the plants from sanitized seeds treated with *E. coli* reaching log 4.4 CFU gfw^{-1} at 21 days and leveling off through day 28. (Fig. 1A)

Similar to the bacterial counts from the lettuce, average counts from mizuna leaf samples at 7 DAP ranged from log 6.0 CFU gfw^{-1} on the non-treated seeds that were not sanitized to the highest level of log 7.2 CFU gfw^{-1} on the non-treated sanitized seeds. This difference was significant (P=0.0095) but did not persist throughout the 28-day growth period. As was seen with lettuce counts, the APC gfw^{-1} declined over the 28-day period but the rate varied with each treatment. The unsanitized, non-treated plants had the highest CFU counts, log 4.8 after 28 days followed by the plants from seeds that were not sanitized but treated with *E. coli* at log 4.1 CFU gfw^{-1}. The lowest counts from the day 28 samples were from the plants from seeds that were sanitized and treated with *E. coli* with average counts of log 2 CFU gfw^{-1}, significantly lower (P=0.03) than the leaves from the unsanitized treated seeds. (Fig. 1B)

Bacterial counts from roots of both leafy greens were higher than leaves, ranging from log 7.4 to log 8.3 CFU gfw^{-1} in lettuce and log 7.8 to log 9.4 CFU gfw^{-1} in mizuna at 7 DAP. As with the leaf tissue, a decline in abundance was observed over time. There was no significant difference between the counts on lettuce root tissue with different seed treatments (Fig. 1C). After 28 days of growth, the counts on mizuna roots from treated sanitized seeds was log 3.2 CFU gfw^{-1}, a significantly lower number than those from roots from all other seed treatments. (Fig. 1D)

E. coli enumeration on leaf and root

E. coli on plant tissues from treated sanitized and unsanitized seeds was enumerated using culture-independent quantitative PCR. This assay demonstrated a decline in *E. coli* abundance over the 28-day growth period in all plant tissues tested. In lettuce leaves, at 7-21 DAP there was no significant difference in *E. coli* cell counts between the seed treatments. Over the 28-day growth period, cell counts from the lettuce leaves from sanitized seeds declined approximately log 2.8 cells gfw^{-1} from log 5.4 at 7 DAP to 2.6 cells gfw^{-1} at harvest (28 DAP), whereas *E. coli* levels from leaves from the unsanitized seeds remained undetectable after 21 DAP. At 21 DAP, *E. coli* was not detected in either treatment, but at day 28 *E. coli* was detected in the sanitized seed treatment only, indicating a significant difference at 28 DAP between lettuce leaves from sanitized and unsanitized seeds. (Fig. 2A)

Enumeration of *E. coli* on lettuce roots from sanitized and unsanitized seeds indicated no significant difference between the treatments over time. Average cell counts gfw^{-1}

declined in the roots from sanitized and unsanitized seeds from log 6.1 to log 3.2 and log 5.6 to log 2.6 respectively. Hence, *E. coli* was detected in the roots after 28 days of growth at harvest (Fig. 2B). Similar to lettuce leaves, *E. coli* from mizuna leaves declined over the course of 28 days of growth from log 7.7 CFU gfw^{-1} to log 2.4 CFU gfw^{-1} in the leaves from sanitized seeds and log 6.9 CFU gfw^{-1} to log 2.4 CFU gfw^{-1} in the leaves from unsanitized seeds. There were no differences in the qPCR data between seed sanitization treatments through the time course however, *E. coli* was detected in both mizuna leaf samples at 28 DAP (Fig. 2C). Mizuna roots from sanitized and unsanitized seeds showed no significant difference in *E*. *coli* counts between the treatments over time. Average cell counts gfw^{-1}declined in the roots from sanitized and unsanitized seeds from log 8.25 to log 3.6 and log 8.6 to log 2.3 respectively. As in the lettuce roots, *E. coli* was detected after 28 days of growth at harvest. (Fig. 2D)

Characterization and abundance of bacterial communities

To understand the combined effects of seed sanitization and *E. coli* treatment on plant-associated microbial communities, a 16S rRNA amplicon sequencing was completed comparing leaf and root microbial communities separately. The sequencing data contained, after removing chimeric and unassigned sequences, a total of 16,518,574 and 17,991,042 (leaf and root, forward and reverse) reads for lettuce and mizuna respectively. Sequences assigned to a chloroplast or mitochondrial taxonomy were removed from further analysis. In addition, OTUs not present in a minimum of six samples and a frequency of at least 50 were not considered in downstream processing.

Comparison of the relative abundances of the bacterial communities between lettuce leaves from sanitized and unsanitized seeds showed differences in the way *E. coli* persisted in treated samples. Between days 7 and 14, the average relative abundance for *E. coli* dropped from 40% to 15%, specifically, in the case of day 14, there was a notable change (~10%) in *E. coli* relative abundance between leaves from unsanitized and sanitized seeds. Genera such as *Pseudomonas*, *Ralstonia* and *Acinetobacter* also showed dynamic changes in relative abundance during days 7 and 14. *Sphingomonas, Burkholderia* and *Bradyrhizobium* exhibited dynamic changes in their relative abundances during days 21 and 28, when *E. coli* began to decline (Fig. 3A & 3B). The differential abundance analysis showed the above-mentioned genera to be common across all time points except for the presence of the genus *Bacillus,* whose abundance increased by more than 2 fold in unsanitized samples (Fig. 5A & 5B).

In the case of roots from sanitized and unsanitized seeds, 89% of the genera represented were common between both sample types, across all time points with the exception of *Caulobacter, Herbaspirillum,* and *Stenotrophomonas*, which were present exclusively in the sanitized seed plants whereas genus *Bacillus* was present only in the unsanitized samples. *E. coli* relative abundance followed a similar pattern for roots in plants from sanitized and unsanitized seeds, starting with the highest abundance at day 7 and gradually decreasing to less than 1% by day 28. Similar to leaves, genera such as *Burkholderia-Caballeronia-Paraburkholderia*, *Acinetobacter*, *Ralstonia*, *Methylophilus*, *Pseudomonas* and *Massilia* exhibited dynamic changes in relative abundance, sometimes showing an inverse relationship to *E.coli* (Fig. 3C & 3D). Between roots from sanitized and unsanitized seeds, a dominance of phyla *Proteobacteria*, *Firmicutes* and

Bacteriodetes was observed. In particular, genera such as *Pantoea*, *Pseudomonas* and *Sphingobium* showed an increased abundance whereas *Sphingomonas*, *Acinetobacter*, *Stentotrophomonas* and *Escherichia-Shigella* showed an overall decrease in abundance throughout the study (Fig. 5C and 5D).

Comparing mizuna leaves from sanitized and unsanitized seeds, more than 78% of the genera were common between the two types of samples, with the exception of *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium, Dyella and Escherichia-Shigella* being exclusively present in unsanitized samples. Leaves from sanitized and unsanitized seeds treated with *E. coli* did show a gradual decrease in *E. coli* abundance from an average of 44% at day 7 to less than 1% at day 28. At the same time, other genera such as *Pseudomonas*, *Acinetobacter*, *Ralstonia* and *Stenotrophomonas* exhibited dynamic shifts in their relative abundance at days 7 and 14, whereas *Sphingomonas* showed a gradual increase in abundance until day 21 (Fig. 4A & 4B). Differential abundance analysis of mizuna leaves from sanitized and unsanitized seeds showed an increased abundance for genera *Acinetobacter* and *Bacillus,* while *Pseudomonas* and *Escherichia-Shigella* showed a decrease in their abundance across all time points studied (Fig. 5E & 5F).

Mizuna roots on the other hand had approximately 73% of the taxa commonly represented between samples from sanitized and unsanitized seeds with each of them retaining 13% taxa exclusively. Similar to leaves, comparison of the roots from sanitized and unsanitized seeds showed a gradual decrease in relative abundance of *E. coli* from an average of 44% at day 7 to less the 1% at day 28. At the same time, some of the most abundant genera such as *Massilia, Allorhizobium-Neorhizobium-Pararhizobium-*

Rhizobium, Azospirillum, Pseudomonas and *Sphingomonas* showed dynamic changes in their relative abundance especially in day 7 and day 14 (Fig. 4C & 4D). These same genera were also found to be differentially abundant across all the time points studied representing some of the dominant phyla such as *Proteobacteria, Firmicutes, Bacteriodetes* and *Actinobacteria* (Fig. 5G & 5H).

Diversity within bacterial communities

Alpha and beta diversity provided overall community characterization of abundance and distribution within and between samples over time. Alpha diversity was determined by the Shannon Index. Sanitized and unsanitized seeds revealed that the Shannon Index generally fell between 1.0 and 4.0, which is within normal range for environmental samples (Fig. 6A-H) (Magurran, 2004, Begossi, 1996; Margalef, 1972).(31–33) There was no significant difference between time points for all lettuce *E. coli* treated and non-treated samples for both sanitized and unsanitized seeds with one exception. A significant difference between day 7 and day 28 was apparent in unsanitized *E. coli* non-treated lettuce root tissue (P=0.034) (Fig. 6 H). A detailed investigation into the Shannon Index and the genera present in the samples indicated that there was an increase in diversity with time and there was nearly three times more genera in the 28 DAP replicates than in the 7 DAP replicates. A close-up look at the abundance of these genera indicated an increase in abundance of several of the genera at 28 DAP including *Methylophilus* and *Acinetobacter*. However, similar trends appear apparent in other lettuce tissues regardless of treatment indicating additional influences on the plant microbiomes.

The Shannon Index for the mizuna was similar in value to the lettuce with index values for leaf tissue between 1.5 and 3, while root tissue index values were higher (between 2 and 4) (Fig. 7 A-H). The Kruskal Wallis comparison between time points in each treatment type showed no significant differences in these overall values in leaf tissue (P>0.05) with two exceptions in mizuna root tissue (Fig. 7 F and H). A closer inspection of the diversity of mizuna roots indicated that there was a similar increase in diversity. The diversity gradually increased with time and the number of species correspondingly increased as well. In the mizuna root of unsanitized non-treated samples, the number of genera nearly tripled while in the unsanitized *E. coli* treated samples the number of genera only doubled between 7 and 28 DAP.

Beta Diversity determines the differentiation between samples. The Bray-Curtis statistic was used to calculate the dissimilarity between *E. coli*-treated, and non-treated samples, along the time course. A higher dissimilarity exists at values closer to 1.0 and a similarity exists closer to zero. We determined that beta diversity in lettuce leaf grown from the sanitized seed presented a significant difference (P<0.05) (Fig. 8). The differences explained in two axes (DAP and *E. coli* exposure) was low, at only 37.3% and there appeared to be a significant variation in the sample distribution (P=0.005). Lettuce leaf tissue from the sanitized, non-treated seed generated samples clustered more closely for days 7, 14 and 28, indicating more similar microbiomes. Day 21 for this treatment diverged from this outcome and the three replicates were more dissimilar. The lettuce leaf tissue generated from sanitized, *E. coli*-treated seeds did not cluster indicating the samples were not similar to each other. The diversity comparisons indicated that time (R^2=0.301) had a greater effect than the *E. coli* treatment (R^2=0.093). Both were

significant (*P*=0.001). The leaf tissue generated from unsanitized seeds presented a similar clustering pattern; however, this pattern was translocated to a different quadrant and distributed across axes (Fig. 8B). The data from these two axes represented approximately 51% of the influence of the two parameters with cluster separation of days 7 and 14 from days 21 and 28. Days 21 and 28 clustered more tightly toward the left quadrants inferring that time may have less of an influence on the microbiome. The variation in the unsanitized seed-generated leaf tissue was not significant (*P*=0.086) (Fig.8 B). In addition, there was no clustering between the *E. coli* treated and non-treated tissue indicating dissimilarity between these samples. However, when time and *E. coli* treatment were considered independently, time could explain about 40% of the variation, and the *E. coli* treatment explained only 7.6% (S1). Both were significant (*P*=0.001 and *P*=0.02, respectively).

Figs. 8C (sanitized seed) and D (unsanitized seed) illustrate the beta diversity for the lettuce root tissue. The PCoA in Figs. 8C and D accounted for 50% and 54% of the variation, respectively. The diversity within root tissues indicates a near reversal of the information gleaned from the lettuce leaf tissue. The replicates from the sanitized seed-generated root tissue, which had been exposed to *E. coli*, clustered tightly with the exception of day 28 (Fig. 8 C). The replicates from each of the non-treated tissues also clustered tightly together, but were separated by time, with day 7 being influenced negatively by both axes, and day 28 influenced positively by both axes. Overall, time could explain approximately 35% of the variation, while the *E. coli* treatment could account for 20% with both being significant (*P*=0.001). The diversity within root tissues generated from unsanitized seeds showed less clustering, therefore indicating more

dissimilarity than the sanitized seed-generated root tissue. Day 7 root tissue showed the most dissimilarity, while the other three time points had greater clustering among the replicates. In addition, day 7 diverged from the other three time points. Days 7 and 14 *E. coli* treated plants did not show any clustering, and were therefore dissimilar, while days 21 and 28 clustered closely to each other but apart from days 7 and 14. This indicates a similarity to each other but more variation between time points. In fact, time could account for 46.2% of the variation while the *E. coli* treatment accounted for 13%. Both were significant (P=0.001).

Beta diversity in the mizuna leaf from sanitized seeds indicated little similarity between the *E. coli*-treated and non-treated plant leaf tissue. The PCoA for sanitized and unsanitized seed generated leaf tissue explained only 48.2% and 45.2% of the variation, respectively. (Fig. 9 A and B). Mizuna samples from sanitized, *E. coli*-treated seeds showed some clustering between samples (i.e. day 7 and 14) however, only 12% of the dissimilarity could be explained by both parameters combined. Time alone could explain approximately 42%, but the *E. coli* treatment explained only 6% (S1). The non-treated samples showed increased variation with the exception of day 28. There was some visible clustering between *E. coli* treated and non-treated samples but there was a clear separation of those samples with time. Similar trends were seen with the unsanitized seed-grown mizuna leaf tissue. There was some clustering between replicates but there was little clustering between mizuna leaf tissue from *E. coli*-treated and non-treated seeds with the exception of day 28. Both treated and non-treated day 28 samples clustered together indicating the close similarity in the microbiome (Fig. 9B). Approximately 42% of the variation was due to time, while the *E. coli* treatment provided only 9%. Together, the

influence dropped drastically to 13.8%. All were a significant influence (P<0.05). The mizuna root showed a different clustering pattern from the leaf tissue. The PCoA in Fig. 9 C and D represents the variation in the sanitized (C) and unsanitized (D) seed-grown root material. The combined axes for the sanitized seed-grown root tissue explains 55.2% of the variation (P=0.001), while the unsanitized seed-grown root tissue explains 44.1% of the variation (P=0.001). There was some clustering among the replicates for both treated and untreated samples with some similarity between day 7 and 14, and between day 21 and 28. However, there was little clustering between the *E. coli*-treated and non-treated samples. This indicated a dissimilarity in the microbiomes generated and the R^2 values of 0.175 and 0.218 support this outcome. In the case of mizuna root tissue, in both sanitized and unsanitized seed treatments, time played a greater role (29%) in the variation than did the *E. coli* exposure (26% and 16% respectively). All were significant (P=0.001)

Discussion

This study examined the dynamics of the culturable, indigenous, aerobic, bacterial population and the persistence of *E. coli* introduced on sanitized and unsanitized seeds in the phyllosphere and rhizosphere of mizuna and lettuce. To minimize the introduction of potential plant and human pathogens on food crops grown on ISS, seeds are routinely sanitized (34) (Massa et. al, 2017) consequently eliminating most of the seed microbial epiphytes. Microbial abundance on seeds is significant with culturable microbial counts ranging from 10^3 to 10^7 depending on the plant (35–37) (Andrews, 1979, Cankar, K, 2005,

Links 2014, therefore; surface sanitizing in an effort to eliminate potential pathogen transmission to the plant initiates a perturbation to the seed microbiome.

The influence of the seed microbiome on the development of the phyllosphere and rhizosphere microbial communities has been studied in a variety of plants. (38–40) (Redford, 2010, van Overbeek, 2011, ardoim, 2012). Links et al, (2014) (37) characterized the bacterial and fungal populations on the surface of wheat and canola seeds and identified the bacterium *Pantoea agglomerans* with antagonistic properties toward a fungal isolate on both seed types. The effect of the removal of seed surface microbes by sanitization on the development and function of succeeding plant microbial communities and pathogen proliferation, to our knowledge, has not been studied. In this study, both sanitized and unsanitized seeds of lettuce and mizuna were inoculated with *E. coli* to determine the impact of the sanitizing treatment, if any, on the growth and persistence of the pathogen on leaves and roots.

Aerobic bacterial counts on the phyllosphere of a variety of leafy greens have been reported in the range of 10^1 to 10^8 CFU g^{-1} (9, 41) (Hunter, 2010, Williams, 2014), which is consistent with our results for red romaine lettuce and mizuna mustard through the 28-day growth period. At 28 DAP, red romaine lettuce from unsanitized, *E. coli* treated seeds had a significantly lower CFU count gfw^{-1} and *E. coli* was not detected in the 21 and 28-day samples. Seeds that had been sanitized and treated with *E. coli* maintained a higher CFU count gfw^{-1} throughout the growth period, and *E. coli* was detected in the 28day samples. This trend suggests a seed sanitization effect on the abundance of bacteria and *E. coli* persistence on the red romaine lettuce leaves, possibly due to the sustained presence of competing bacterial species on the plants from unsanitized seeds. Seed

inoculation of native plant rhizosphere or antagonistic bacteria has proven to be an effective biocontrol to prevent the proliferation of pathogens in seed sprouts (42) (Matos, 2005). This trend was not evident in the mizuna mustard leaf bacterial counts with no difference between the treatments through 21 days of growth followed by a significant decrease in APC on the leaves from sanitized, *E. coli* treated seeds at 28 DAP and no significant difference in the *E. coli* counts. *E. coli* was detected after 28 days in all the mizuna leaf and root samples, although steadily declining through the 28-day growth period.

Diversity in these bacterial communities may play a role or serve as indicators of plant health. Phyllosphere bacteria play important roles as they provide a resource for nutrient cycling (43, 44) (Jones, 1970; Freiberg, 1998) or prevent pathogens from colonizing leaf surfaces (45, 46) (Kishore, et al., 2005; Bodenhausen, et al, 2013). Microbial endophytes in healthy plants would have a positive influence on the host but also play a role in microbe-microbe interactions (47, 48) (Lebeis, 2014; Weyens, et al., 2009). Early colonization would set up or establish varying community compositions in the phyllosphere (49) (Maignun, et al., 2014). By altering the microbial community of the seed, it is suspected that the microbiomes would vary, with a potential to affect plant health.

Through 16S rRNA gene amplicon sequencing, we identified *Proteobacteria, Actinobacteria, Bacteriodetes and Firmicutes* as the most prevalent phyla in red romaine lettuce leaves and roots, irrespective of seed sanitization status. This is consistent with previous studies of bacterial communities on leafy greens (8, 10, 14, 50–52)(Lopez-Velasco et al., 2011; Rastogi et al., 2012; Williams et al., 2013; Bulgarelli et al., 2013; Jackson et al., 2013; Dees et al., 2014). Our results showed that, in red romaine lettuce

roots, the relative abundance of *Actinobacteria* and *Bacteriodetes* gradually increased through different developmental stages, which was consistent with observations by (53) Chaparro et al. (2014).

In case of mizuna mustard, we observed *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Cyanobacteria*, *Bacteriodetes* as the most dominant phyla in leaves and roots, independent of seed sanitization. This is again consistent with what other studies describe (54) (Thapa & Prasanna, 2018). Similar to red romaine lettuce, relative abundance of *Actinobacteria* and *Bacteriodetes* gradually increased through different developmental stages. Comparing between mizuna mustard leaves and roots, *Firmicutes* were present throughout, and their abundance gradually decreased through the time course. Recently proposed super phylum *Patescibacteria* that encompasses mostly unculturable bacteria from habitats such as groundwater, deep-sea sediments, permafrost, and the continental deep subsurface was also one of the dominant phyla in roots. It has been suggested that members of phylum *Patescibacteria* to have a symbiotic life-style (55, 56) (Lemos et al., 2019; Herrmann et al., 2019).

Invasion by pathogenic microbes often correlates with shifts in microbial community in different plant compartments, including leaves (57) (Agler et al., 2016) and roots (Xue et al., 2015). Our results showed that, members of the class *Gammaproteobacteria*, such as *Pseudomonas*, *Ralstonia*, *Massilia* and *Acinetobacter* are some of the dominant genera present in each leafy crop. The same genera have been identified as crucial part of other plant microbiomes, either as indicators of plant health (58) (Köberl et al., 2017) or their role against invading pathogens (59–61)(Erlacher et al., 2014; Fürnkranz et al., 2012; Berg et al., 2014). Moreover, in our case, most of these core genera were found

to be significantly correlated with *E. coli* abundance, similar to what has been shown for *Fusarium* wilt of banana plants (58) (Köberl et al., 2017), wheat stripe rust (62) (Dai et al., 2020) and *Rhizoctonia solani* impact on lettuce plants (59) (Erlacher et al., 2014).

We found minimal changes in the diversity of lettuce or mizuna leaf tissue with time or between sanitized and unsanitized seeds, although there were changes in the number of OTUs with time. Overall, richness was lower in leaf tissue than in root tissue. Similar studies with model organisms such as *Arabidopsis* have shown similar trends (46) (Bodenhausen, 2013). Additionally, root tissues are more diverse than leaf tissue. Time (age) may also have a significant effect, a factor that we saw influencing diversity between samples. Time had more of an influence on all samples versus the *E. coli* introduction as shown in the beta diversity. Whereas the time course of 7 to 28 DAP accounted for over 50% of the variation in some samples, the *E. coli* treatment and the combined time-treatment accounted for much less of the diversity.

Bacteria on leaf surfaces are affected by environmental conditions such as temperature and humidity though these factors were maintained at steady settings and therefore would have had minimal impact on the bacterial community. The dynamic changes seen in the phyllosphere could not be explained by environmental factors. In general, there were lower diversities in the phyllospheres especially at 7 DAP with diversity increasing with time. At 7 DAP approximately 10 genera were found in most leaf tissue but by 14-21 DAP, the number more than doubled with the non-treated samples having more genera than the *E. coli*-treated samples. This could be attributed to genera such as *Bacillus, Ralstonia*, and *Pseudomonas*; however, *Pantoea* and *Massilia* persisted in the non-treated, unsanitized seed leaf tissue. *Ralstonia* does not appear to persist in the *E. coli*-treated

unsanitized seed leaf issue, however as we see the *E. coli* decline, *Pseudomonas* and *Burkholderia* both increase indicating a possible competition for resources. In the sanitized leaf, we see *Ralstonia* present and increasing with time in both *E. coli* treated and untreated plants. In addition, similar to the unsanitized seed, *Pseudomonas* and *Sphingomonas* both increase in samples from the *E. coli*-treated seed and arise about 21 DAP, again indicating possible competition with *E. coli.* Innerebner, et al., (2013) (63) has shown that some *Sphingomonas* species have the capacity to outcompete certain *Pseudomonas* species thereby serving as a nutrient-related beneficial microbe.

Root tissue has a higher diversity than leaf tissue as expected as the rhizospheres and roots possess a higher number of genera. Some species were present only in the roots of unsanitized seeds indicating they may be removed from the seed surface during the sanitization process. These include *Hydrotalea*, *Bacillus*, and *Micrococcus*. *Massilia*, a root colonizing bacteria was present in the roots from unsanitized seeds in high abundance. Ecologically a generalist, it has been found to be present in the early stages of plant development, proliferates on the seed coat and may be sensitive to surrounding media (46, 64) (Ofek, et al 2012; Bodenhausen 2013). We see either a marked decline or absence of *Massilia* in treated samples. This may be due to the competition with *E. coli* for nutrients or due to the media and fertilizers used in this study. Additional study would provide resolution to this observation. This genus, which can also inhabit the phyllosphere and spermosphere, was only present in the leaves and roots of; unsanitized, non-treated seeds indicating that sanitization of the seed had an additional effect on the bacterial community. One additional genus that warrants investigation is *Stenotrophomonas,* which was identified only in the root tissue of sanitized, *E. coli*-treated seeds. Previously

547 classified as a pseudomonad, *Stenotrophomonas* has been isolated in soil, plant, water
548 and human tissue. In soil and plants, it is known for production of plant hormones and is
549 very persistent in the roots, and once established, is not easily out-competed (65,
550 66)(Denet et al, 2018; Peralta et al, 2012).
551
552
553
554
555

References

1. Nelson EB. 2018. The seed microbiome: Origins, interactions, and impacts. Plant Soil 422:7–34.

2. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Beer S V., Machado MA, Toth I, Salmond G, Foster GD. 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. Mol Plant Pathol 13:614–629.

3. Dewey-Mattia D, Manikonda K, Hall AJ, Wise ME, Crowe SJ. 2018. Surveillance for Foodborne Disease Outbreaks - United States, 2009-2015. MMWR Surveill Summ 67:1998–2008.

4. CDC. 2020. CDC Foodborne outbreaks http://wwwn.cdc.gov/foodborneoutbreaks/.

5. Ibekwe AM, Grieve CM. 2004. Changes in developing plant microbial community structure as affected by contaminated water. FEMS Microbiol Ecol 48:239–248.

6. Luna-Guevara JJ, Arenas-Hernandez MMP, Martínez De La Peña C, Silva JL, Luna-Guevara ML. 2019. The Role of Pathogenic E. coli in Fresh Vegetables: Behavior, Contamination Factors, and Preventive Measures. Int J Microbiol 2019.

7. Tyler HL, Triplett EW. 2008. Plants as a Habitat for Beneficial and/or Human Pathogenic Bacteria. Annu Rev Phytopathol 46:53–73.

8. Rastogi G, Tech JJ, Coaker GL, Leveau JHJ, Sbodio A, Suslow T V. 2012. Leaf microbiota in an agroecosystem: Spatiotemporal variation in bacterial community

composition on field-grown lettuce. ISME J 6:1812–1822.

9. Williams TR, Marco ML. 2014. Transplantation on Lettuce Plants Grown Indoors. MBio 5:1–10.

10. Williams TR, Moyne AL, Harris LJ, Marco ML. 2013. Season, Irrigation, Leaf Age, and Escherichia coli Inoculation Influence the Bacterial Diversity in the Lettuce Phyllosphere. PLoS One 8:1–14.

11. Dees MW, Lysøe E, Nordskog B, Brurberg MB. 2015. Bacterial communities associated with surfaces of leafy greens: Shift in composition and decrease in richness over time. Appl Environ Microbiol 81:1530–1539.

12. Merget B, Forbes KJ, Brennan F, Mcateer S, Shepherd T, Strachan NJC, Holden NJ. 2019. Influence of plant species, tissue type, and temperature on the capacity of Shiga-Toxigenic Escherichia coli To Colonize, Grow , and Be Internalized by Plants. Appl Environ Microbiol 85:1–16.

13. Santhanam R, Luu VT, Weinhold A, Goldberg J, Oh Y, Baldwin IT. 2015. Native root-associated bacteria rescue a plant from a sudden-wilt disease that emerged during continuous cropping. Proc Natl Acad Sci U S A 112:E5013–E5120.

14. Bulgarelli D, Schlaeppi K, Spaepen S, van Themaat EVL, Schulze-Lefert P. 2013. Structure and Functions of the Bacterial Microbiota of Plants. Annu Rev Plant Biol 64:807–838.

15. Toju H, Peay KG, Yamamichi M, Narisawa K, Hiruma K, Naito K, Fukuda S, Ushio M, Nakaoka S, Onoda Y, Yoshida K, Schlaeppi K, Bai Y, Sugiura R, Ichihashi Y,

598 Minamisawa K, Kiers ET. 2018. Core microbiomes for sustainable
599 agroecosystems. Nat Plants 4:247–257.

600 16. Khodadad CLM, Hummerick ME, Spencer LSE, Dixit AR, Richards JT, Romeyn
601 MW, Smith TM, Wheeler RM, Massa GD. 2020. Microbiological and Nutritional
602 Analysis of Lettuce Crops Grown on the International Space Station. Front Plant
603 Sci 11:1–15.

604 17. Clough SJ, Bent AE. 2011. Vapor-Phase Sterilization of Arabidopsis Seed.
605 http://www.plantpath.wisc.edu/fac/afb/vapster.html.

606 18. Massa GD, Newsham G, Hummerick ME, Morrow RC, Wheeler RM. 2017.
607 Preparation for the Veggie Plant growth System on the international Space
608 Station. Gravitational Sp Res 5:24–34.

609 19. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013.
610 Development of a dual-index sequencing strategy and curation pipeline for
611 analyzing amplicon sequence data on the miseq illumina sequencing platform.
612 Appl Environ Microbiol 79:5112–5120.

613 20. Bolyen E, Dillon M, Bokulich N, Abnet C, Al-Ghalith G, Alexander H, Alm E,
614 Arumugam M, Asnicar F, Bai Y, Bisanz J, Bittinger K, Brejnrod A, Brislawn C,
615 Brown T, Callahan B, Chase J, Cope E, Dorrestein P, Douglas G, Durall D,
616 Duvallet C, Edwardson C, Ernst M, Estaki M, Fouquier J, Gauglitz J, Gibson D,
617 Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C,
618 Huttley G, Janssen S, Jarmusch A, Jiang L, Kaehler B, Keefe C, Keim P, Kelley
619 S, Knights D, Koester I, Kosciolek T, Kreps J, Lee J, Ley R, Liu Y-X, Loftfield E,

Lozupone C, Maher M, Marotz C, Martin B, McDonald D, McIver L, Melnik A, Metcalf J, Morgan S, Morton J, Navas-Molina J, Orchanian S, Pearson T, Peoples S, Petras D, Pruesse E, Rivers A, Robeson M, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Spear J, Swafford A, Thompson L, Torres P, Trinh P, Tripathi A, Turnbaugh P, Ul-Hasan S, Vargas F, Vogtmann E, Walters W, Wan Y, Wang M, Warren J, Weber K, Willis A, Zaneveld J, Zhang Y, Zhu Q, Knight R, Caporaso G. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ Prepr.

21. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:590–596.

22. Beckers B, Op De Beeck M, Thijs S, Truyens S, Weyens N, Boerjan W, Vangronsveld J. 2016. Performance of 16s rDNA primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies. Front Microbiol 7.

23. Beirinckx S, Viaene T, Haegeman A, Debode J, Amery F, Vandenabeele S, Nelissen H, Inzé D, Tito R, Raes J, De Tender C, Goormachtig S. 2020. Tapping into the maize root microbiome to identify bacteria that promote growth under chilling conditions. Microbiome 8:54.

24. Bisanz JE (2018). 2018. qiime2R: Importing QIIME2 artifacts and associated data into R sessions.

25. McMurdie PJ, Holmes S. 2014. Waste Not, Want Not: Why Rarefying Microbiome

642 Data Is Inadmissible. PLoS Comput Biol 10.

643 26. Robinson MD, McCarthy DJ, Smyth GK. 2009. edgeR: A Bioconductor package
644 for differential expression analysis of digital gene expression data. Bioinformatics
645 26:139–140.

646 27. Shannon CE. 1948. A Mathematical Theory of Communication. Bell Syst Tech J
647 27:623–656.

648 28. Bray JR, Curtis JT. 1957. An Ordination of the Upland Forest Communities of
649 Southern Wisconsin. Ecol Monogr 27:325–349.

650 29. Kruskal WH, Wallis WA. 1952. Use of Ranks in One-Criterion Variance Analysis.
651 J Am Stat Assoc 47:583–621.

652 30. Marti J. Anderson. 2001. A new method for non-parametric multivariate analysis
653 of variance. Austral Ecol 26:32–46.

654 31. Magurran AE. 2003. Measuring biological diversity. Wiley-Blackwell.

655 32. Begossi A. 1996. Use of ecological methods in ethnobotany: Diversity indices.
656 Econ Bot 50:280–289.

657 33. R. M. 1972. Homage to Evelyn Hutchinson, or Why There is an Upper Limit to
658 Diversity. Trans Conn Acad Arts Sci 44:211–235.

659 34. Massa GD, Newsham G, Hummerick ME, Morrow RC, Wheeler RM. 2017. Plant
660 Pillow Preparation for the Veggie Plant Growth System on the International Space
661 Station. Gravitational Sp Res 5:24–34.

662 35. Andrews WH, Wilson CR, Poelma PL, Romero A, Mislivec PB. 1979.

663 Bacteriological survey of sixty health foods. Appl Environ Microbiol 37:559–566.

664 36. Katarina Cankar 1, Hojka Kraigher, Maja Ravnikar MR. 2005. Bacterial
665 endophytes from seeds of Norway spruce (Picea Abies L. FEMS Microbiol Lett
666 244:341–345.

667 37. Links MG, Demeke T, Gräfenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ.
668 2014. Simultaneous profiling of seed-associated bacteria and fungi reveals
669 antagonistic interactions between microorganisms within a shared epiphytic
670 microbiome on Triticum and Brassica seeds. New Phytol 202:542–553.

671 38. Redford A, Bowers R, Knight R, Linhart Y, Fierer N. 2011. Variability in the
672 Distribution of Bacteria on Tree Leaves. Environ Microbiol 12:2885–2893.

673 39. van Overbeek LS, Saikkonen K. 2016. Impact of bacterial-fungal interactions on
674 the colonization of the endosphere. Trends Plant Sci 21:230–242.

675 40. Hardoim PR, Hardoim CCP, van Overbeek LS, van Elsas JD. 2012. Dynamics of
676 seed-borne rice endophytes on early plant growth stages. PLoS One 7.

677 41. Hunter PJ, Hand P, Pink D, Whipps JM, Bending GD. 2010. Both leaf properties
678 and microbe-microbe interactions influence within-species variation in bacterial
679 population diversity and structure in the lettuce (lactuca species) phyllosphere.
680 Appl Environ Microbiol 76:8117–8125.

681 42. Matos A, Garland JL. 2005. Effects of community versus single strain inoculants
682 on the biocontrol of Salmonella and microbial community dynamics in alfalfa
683 sprouts. J Food Prot 68:40–48.

684 43. Jones K. 1970. Nitrogen fixation in the phyllosphere of the douglas fir,
685 Pseudotsuga douglasii. Ann Bot 34:239–244.

686 44. Freiberg E. 1998. Microclimatic parameters influencing nitrogen fixation in the
687 phyllosphere in a Costa Rican premontane rain forest. Oecologia 117:9–18.

688 45. Kishore GK, Pande S, Podile AR. 2005. Phylloplane bacteria increase seedling
689 emergence, growth and yield of field-grown groundnut (Arachis hypogaea L.). Lett
690 Appl Microbiol 40.

691 46. Bodenhausen N, Horton MW, Bergelson J. 2013. Bacterial Communities
692 Associated with the Leaves and the Roots of Arabidopsis thaliana. PLoS One 8.

693 47. Lebeis SL. 2014. The potential for give and take in plant-microbiome
694 relationships. Front Plant Sci 5:1–6.

695 48. Weyens N, van der Lelie D, Taghavi S, Newman L, Vangronsveld J. 2009.
696 Exploiting plant–microbe partnerships to improve biomass production and
697 remediation. Trends Biotechnol 27:591–598.

698 49. Maignien L, DeForce EA, Chafee ME, Eren AM, Simmons SL. 2014. Ecological
699 Succession and Stochastic Variation in the Assembly of Arabidopsis thaliana
700 Phyllosphere Communities. MBio 5:e000682-13.

701 50. Lopez-Velasco G, Carder PA, Welbaum GE, Ponder MA. 2013. Diversity of the
702 spinach (Spinacia oleracea) spermosphere and phyllosphere bacterial
703 communities. FEMS Microbiol Lett 346:146–154.

704 51. Jackson CR, Randolph KC, Osborn SL, Tyler HL. 2013. Culture dependent and

independent analysis of bacterial communities associated with commercial salad leaf vegetables. BMC Microbiol 13.

52. Dees MW, Lysøe E, Nordskog B, Brurberg MB. 2015. Bacterial communities associated with surfaces of leafy greens: Shift in composition and decrease in richness over time. Appl Environ Microbiol 81:1530–1539.

53. Chaparro JM, Badri D V., Vivanco JM. 2014. Rhizosphere microbiome assemblage is affected by plant development. ISME J 8:790–803.

54. Thapa S, Prasanna R. 2018. Prospecting the characteristics and significance of the phyllosphere microbiome. Ann Microbiol 68:229–245.

55. Herrmann M, Wegner CE, Taubert M, Geesink P, Lehmann K, Yan L, Lehmann R, Totsche KU, Küsel K. 2019. Predominance of Cand. Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. Front Microbiol 10:1–15.

56. Lemos, Leandro N, Medeiros, Julliane D, Dini-Andreote F, Fernandes, Gabriel R, Varani, Alessandro M, Oliveira G, Pylro VS. 2019. Genomic signatures and co-occurrence patterns of the ultra-small Saccharimonadia (phylum CPR/Patescibacteria) suggest a symbiotic lifestyle. Mol Ecol 28.

57. Agler MT, Ruhe J, Kroll S, Morhenn C, Kim ST, Weigel D, Kemen EM. 2016. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. PLoS Biol 14:1–31.

58. Köberl M, Dita M, Martinuz A, Staver C, Berg G. 2017. Members of

726 Gammaproteobacteria as indicator species of healthy banana plants on Fusarium
727 wilt-infested fields in Central America. Sci Rep 7:1–9.

728 59. Erlacher A, Cardinale M, Grosch R, Grube M, Berg G. 2014. The impact of the
729 pathogen Rhizoctonia solani and its beneficial counterpart Bacillus
730 amyloliquefaciens on the indigenous lettuce microbiome. Front Microbiol 5:1–8.

731 60. Fürnkranz M, Lukesch B, Müller H, Huss H, Grube M, Berg G. 2012. Microbial
732 Diversity Inside Pumpkins: Microhabitat-Specific Communities Display a High
733 Antagonistic Potential Against Phytopathogens. Microb Ecol 63:418–428.

734 61. Berg G, Mahnert A, Moissl-Eichinger C. 2014. Beneficial effects of plant-
735 associated microbes on indoor microbiomes and human health? Front Microbiol
736 5:1–5.

737 62. Dai Y, Yang F, Zhang L, Xu Z, Fan X, Tian Y, Wang T. 2020. Wheat-associated
738 microbiota and their correlation with stripe rust reaction. J Appl Microbiol
739 128:544–555.

740 63. Innerebner G, Knief C, Vorholt JA. 2011. Protection of Arabidopsis thaliana
741 against leaf-pathogenic Pseudomonas syringae by Sphingomonas strains in a
742 controlled model system. Appl Environ Microbiol 77:3202–3210.

743 64. Ofek M, Hadar Y, Minz D. 2012. Ecology of root colonizing Massilia
744 (Oxalobacteraceae). PLoS One 7.

745 65. Denet E, Vasselon V, Burdin B, Nazaret S, Favre-Bonté S. 2018. Survival and
746 growth of stenotrophomonas maltophilia in free-living amoebae (FLA) and

bacterial virulence properties. PLoS One 13:1–16.

66. Peralta KD, Araya T, Valenzuela S, Sossa K, Martínez M, Peña-Cortés H, Sanfuentes E. 2012. Production of phytohormones, siderophores and population fluctuation of two root-promoting rhizobacteria in Eucalyptus globulus cuttings. World J Microbiol Biotechnol 28:2003–2014.

766

767

768

Acknowledgment

780

781

782

783

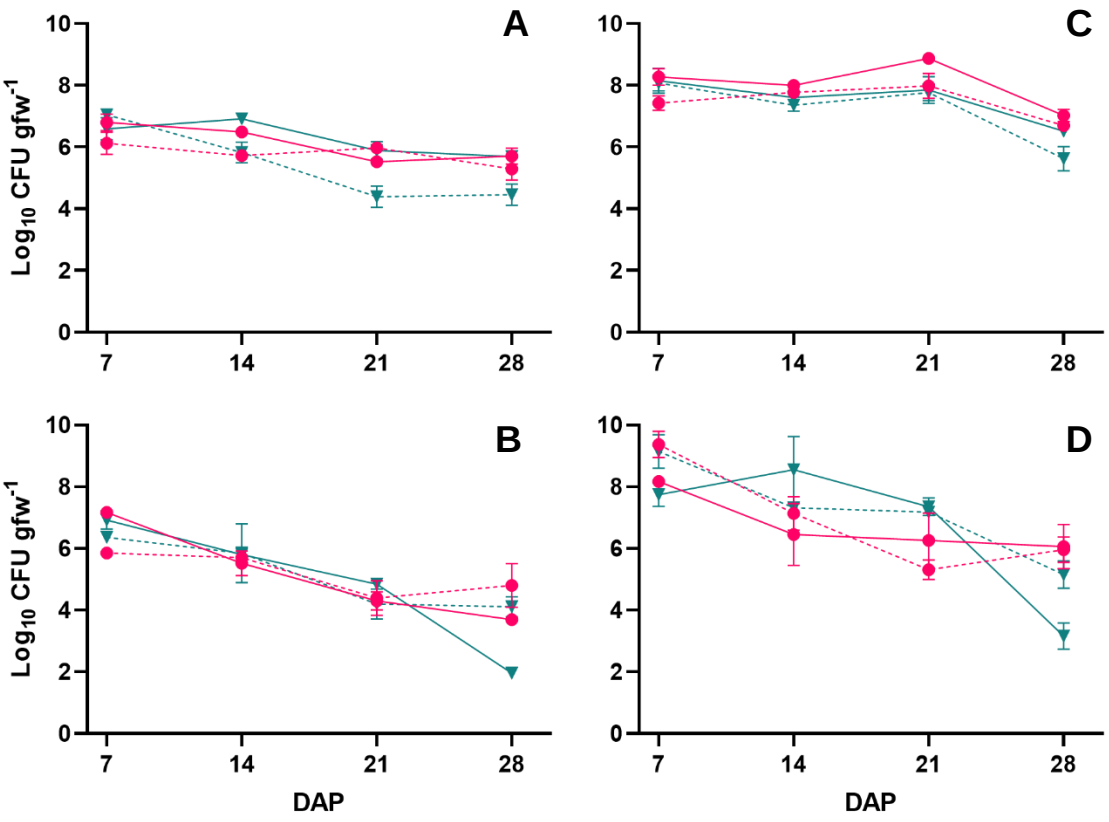


Figure 1. Aerobic plate counts per gram fresh weight on red romaine lettuce leaves (A) and roots (B) and mizuna mustard, leaves (C) and roots (D) from (●) sanitized (solid lines) or unsanitized (dashed lines) seeds treated with *E. coli* (▼) and non-treated . Each point represents the average and bars represent SEM (n=3).

784

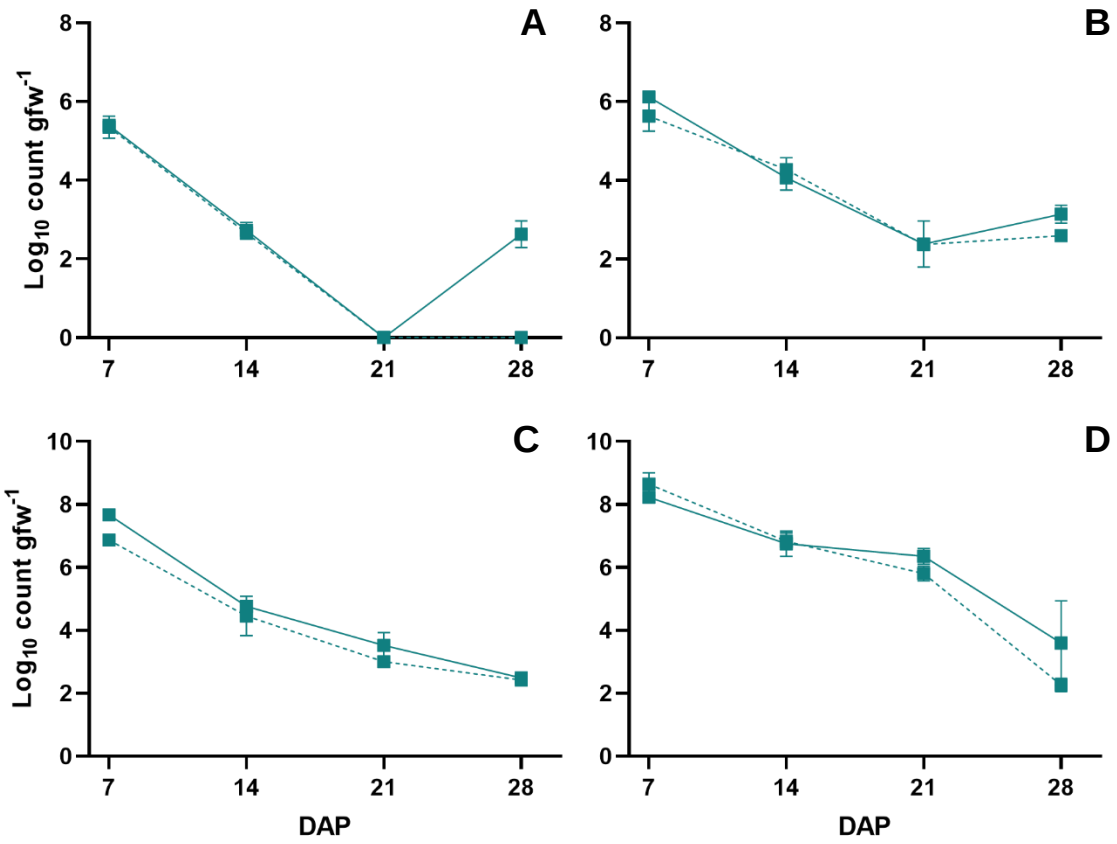
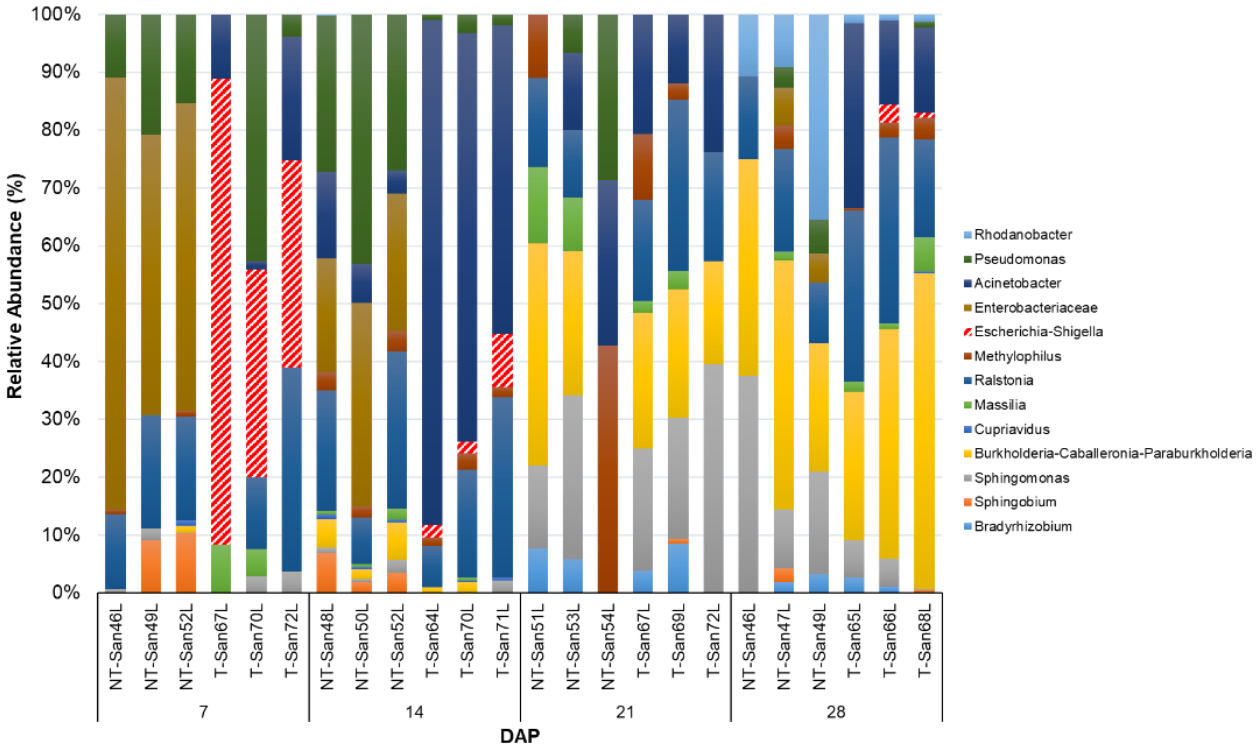
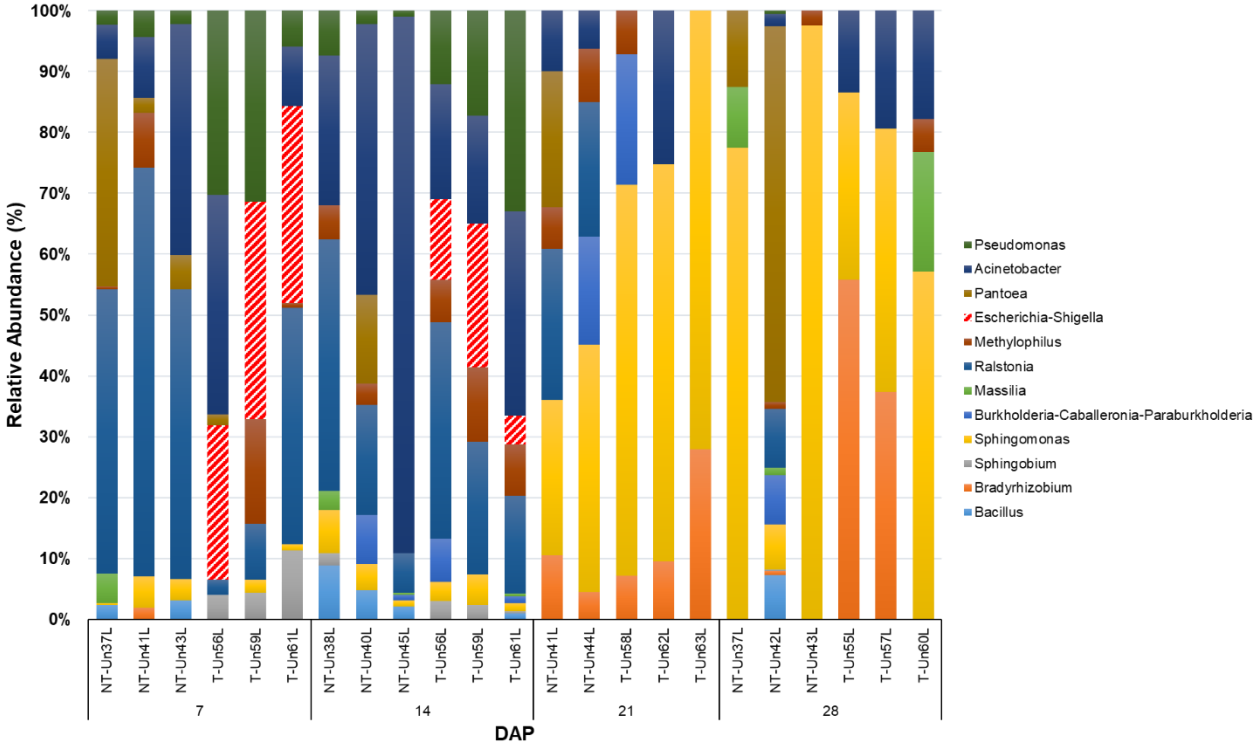


Figure 2. *E. coli* counts per gram fresh weight on red romaine lettuce leaves (A) and roots (B) and mizuna mustard, leaves (C) and roots (D) from *E. coli*-treated, sanitized (solid lines) or unsanitized (dashed lines) seeds as determined by qPCR (■). Each point represents the average and bars represent SEM (n=3).

(A)

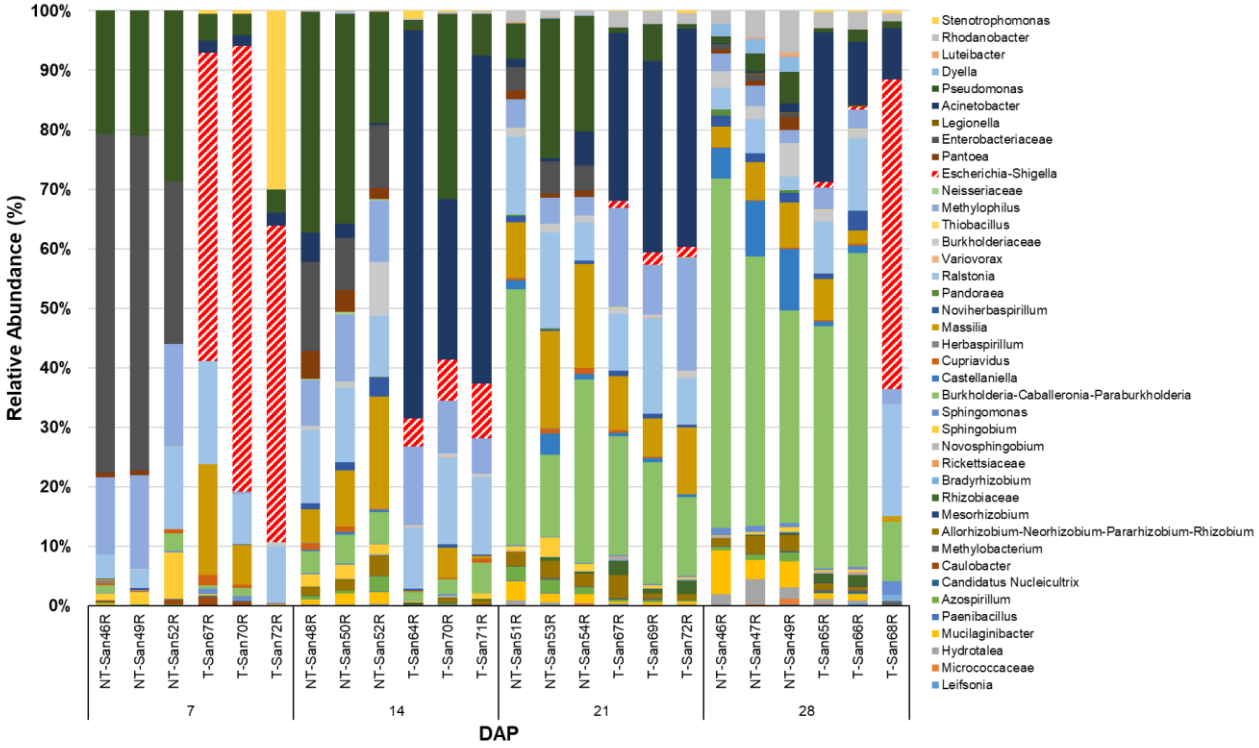

Relative Abundance (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NT-San46L
NT-San49L
NT-San52L
T-San67L
T-San70L
T-San72L
NT-San48L
NT-San50L
NT-San52L
T-San64L
T-San70L
T-San71L
NT-San51L
NT-San53L
NT-San54L
T-San67L
T-San69L
T-San72L
NT-San46L
NT-San47L
NT-San49L
T-San65L
T-San66L
T-San68L
7
14
21
28
DAP
Rhodanobacter
Pseudomonas
Acinetobacter
Enterobacteriaceae
Escherichia-Shigella
Methylophilus
Ralstonia
Massilia
Cupriavidus
Burkholderia-Caballeronia-Paraburkholderia
Sphingomonas
Sphingobium
Bradyrhizobium

(B)

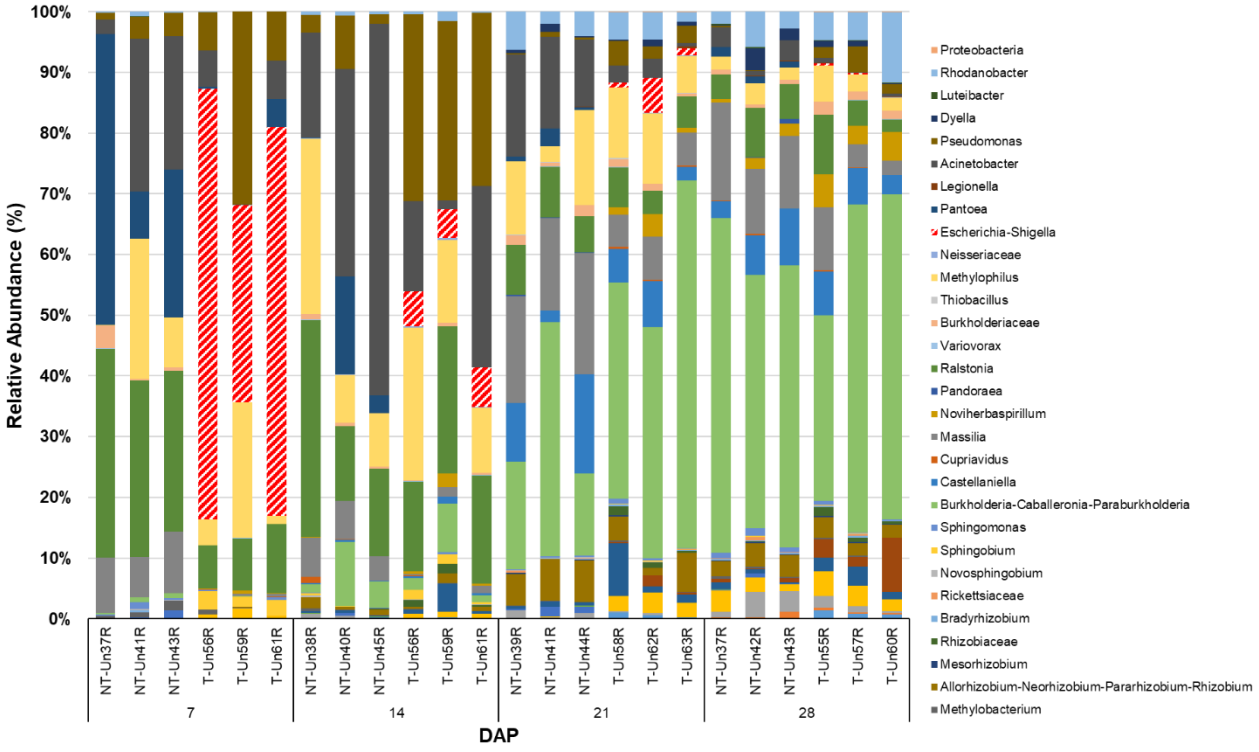

Relative Abundance (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NT-Un37L
NT-Un41L
NT-Un43L
T-Un56L
T-Un59L
T-Un61L
NT-Un38L
NT-Un40L
NT-Un45L
T-Un56L
T-Un59L
T-Un61L
NT-Un41L
NT-Un44L
T-Un58L
T-Un62L
T-Un63L
NT-Un37L
NT-Un42L
NT-Un43L
T-Un55L
T-Un57L
T-Un60L
7
14
21
28
DAP
Pseudomonas
Acinetobacter
Pantoea
Escherichia-Shigella
Methylophilus
Ralstonia
Massilia
Burkholderia-Caballeronia-Paraburkholderia
Sphingomonas
Sphingobium
Bradyrhizobium
Bacillus

785

(C)


Relative Abundance (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NT-San46R
NT-San49R
NT-San52R
T-San67R
T-San70R
T-San72R
NT-San48R
NT-San50R
NT-San52R
T-San64R
T-San70R
T-San71R
NT-San51R
NT-San53R
NT-San54R
T-San67R
T-San69R
T-San72R
NT-San46R
NT-San47R
NT-San49R
T-San65R
T-San66R
T-San68R
7
14
21
28
DAP
Stenotrophomonas
Rhodanobacter
Luteibacter
Dyella
Pseudomonas
Acinetobacter
Legionella
Enterobacteriaceae
Pantoea
Escherichia-Shigella
Neisseriaceae
Methylophilus
Thiobacillus
Burkholderiaceae
Variovorax
Ralstonia
Pandoraea
Noviherbaspirillum
Massilia
Herbaspirillum
Cupriavidus
Castellaniella
Burkholderia-Caballeronia-Paraburkholderia
Sphingomonas
Sphingobium
Novosphingobium
Rickettsiaceae
Bradyrhizobium
Rhizobiaceae
Mesorhizobium
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Methylobacterium
Caulobacter
Candidatus Nucleicultrix
Azospirillum
Paenibacillus
Mucilaginibacter
Hydrotalea
Micrococcaceae
Leifsonia

(D)


Relative Abundance (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NT-Un37R
NT-Un41R
NT-Un43R
T-Un56R
T-Un59R
T-Un61R
NT-Un38R
NT-Un40R
NT-Un45R
T-Un56R
T-Un59R
T-Un61R
NT-Un39R
NT-Un41R
NT-Un44R
T-Un58R
T-Un62R
T-Un63R
NT-Un37R
NT-Un42R
NT-Un43R
T-Un55R
T-Un57R
T-Un60R
7
14
21
28
DAP
Proteobacteria
Rhodanobacter
Luteibacter
Dyella
Pseudomonas
Acinetobacter
Legionella
Pantoea
Escherichia-Shigella
Neisseriaceae
Methylophilus
Thiobacillus
Burkholderiaceae
Variovorax
Ralstonia
Pandoraea
Noviherbaspirillum
Massilia
Cupriavidus
Castellaniella
Burkholderia-Caballeronia-Paraburkholderia
Sphingomonas
Sphingobium
Novosphingobium
Rickettsiaceae
Bradyrhizobium
Rhizobiaceae
Mesorhizobium
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Methylobacterium

786

Figure 3: Stacked bar plots showing percent relative abundance of bacterial populations in red romaine lettuce leaves (L) and roots (R), sanitized (A) and (C) or unsanitized (B) and (D) respectively over time. Samples were either treated (T) or non-treated (NT) with *E. coli.* Numbers on X-axis represent days after planting (DAP).

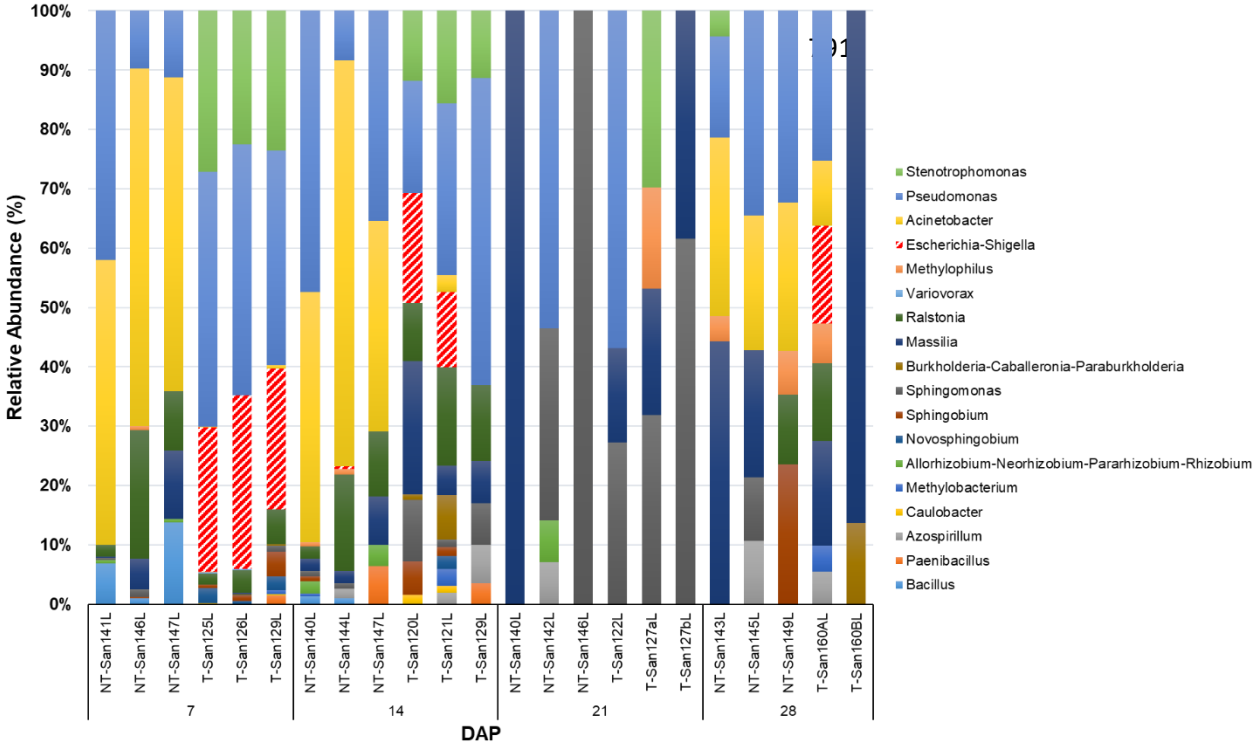
787

788

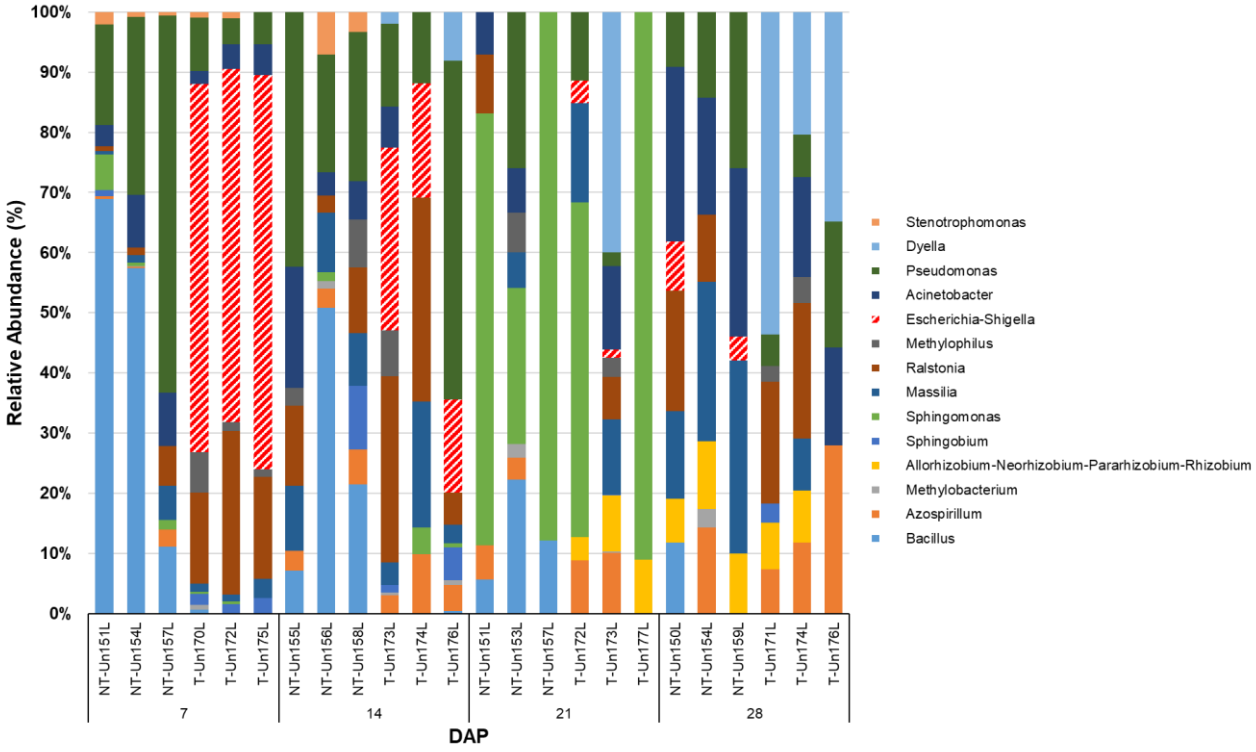
789

790

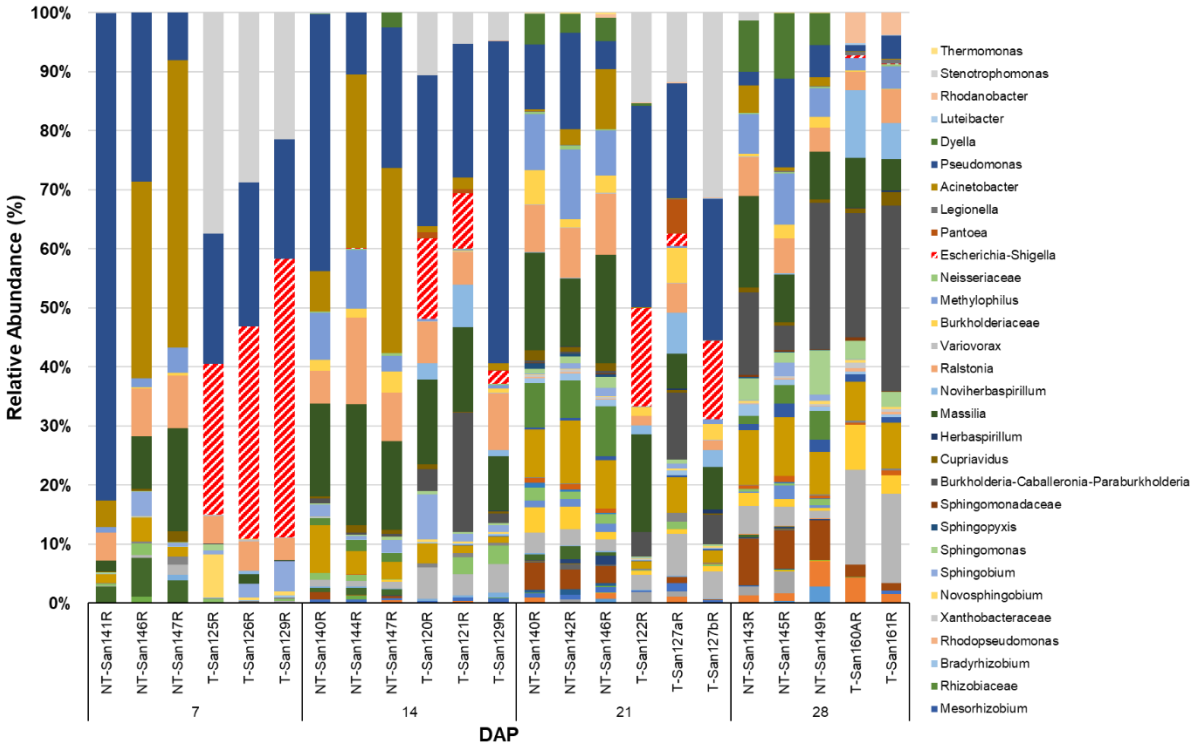
(A)


Relative Abundance (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NT-San141L
NT-San146L
NT-San147L
T-San125L
T-San126L
T-San129L
NT-San140L
NT-San144L
NT-San147L
T-San120L
T-San121L
T-San129L
NT-San140L
NT-San142L
NT-San146L
T-San122L
T-San127aL
T-San127bL
NT-San143L
NT-San145L
NT-San149L
T-San160AL
T-San160BL
7
14
21
28
DAP
Stenotrophomonas
Pseudomonas
Acinetobacter
Escherichia-Shigella
Methylophilus
Variovorax
Ralstonia
Massilia
Burkholderia-Caballeronia-Paraburkholderia
Sphingomonas
Sphingobium
Novosphingobium
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Methylobacterium
Caulobacter
Azospirillum
Paenibacillus
Bacillus

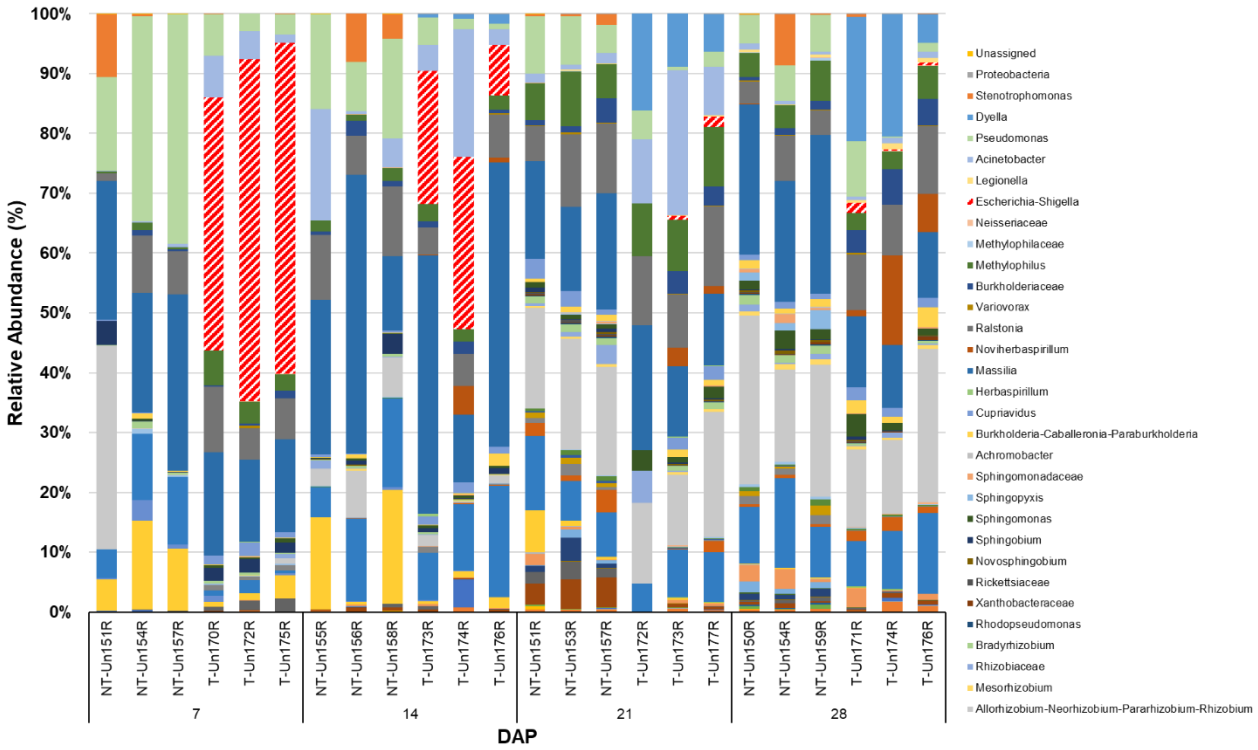
(B)


Relative Abundance (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NT-Un151L
NT-Un154L
NT-Un157L
T-Un170L
T-Un172L
T-Un175L
NT-Un155L
NT-Un156L
NT-Un158L
T-Un173L
T-Un174L
T-Un176L
NT-Un151L
NT-Un153L
NT-Un157L
T-Un172L
T-Un173L
T-Un177L
NT-Un150L
NT-Un154L
NT-Un159L
T-Un171L
T-Un174L
T-Un176L
7
14
21
28
DAP
Stenotrophomonas
Dyella
Pseudomonas
Acinetobacter
Escherichia-Shigella
Methylophilus
Ralstonia
Massilia
Sphingomonas
Sphingobium
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Methylobacterium
Azospirillum
Bacillus

(C)


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Relative Abundance (%)
NT-San141R
NT-San146R
NT-San147R
T-San125R
T-San126R
T-San129R
NT-San140R
NT-San144R
NT-San147R
T-San120R
T-San121R
T-San129R
NT-San140R
NT-San142R
NT-San146R
T-San122R
T-San127aR
T-San127bR
NT-San143R
NT-San145R
NT-San149R
T-San160AR
T-San161R
7
14
21
28
DAP
Thermomonas
Stenotrophomonas
Rhodanobacter
Luteibacter
Dyella
Pseudomonas
Acinetobacter
Legionella
Pantoea
Escherichia-Shigella
Neisseriaceae
Methylophilus
Burkholderiaceae
Variovorax
Ralstonia
Noviherbaspirillum
Massilia
Herbaspirillum
Cupriavidus
Burkholderia-Caballeronia-Paraburkholderia
Sphingomonadaceae
Sphingopyxis
Sphingomonas
Sphingobium
Novosphingobium
Xanthobacteraceae
Rhodopseudomonas
Bradyrhizobium
Rhizobiaceae
Mesorhizobium

(D)

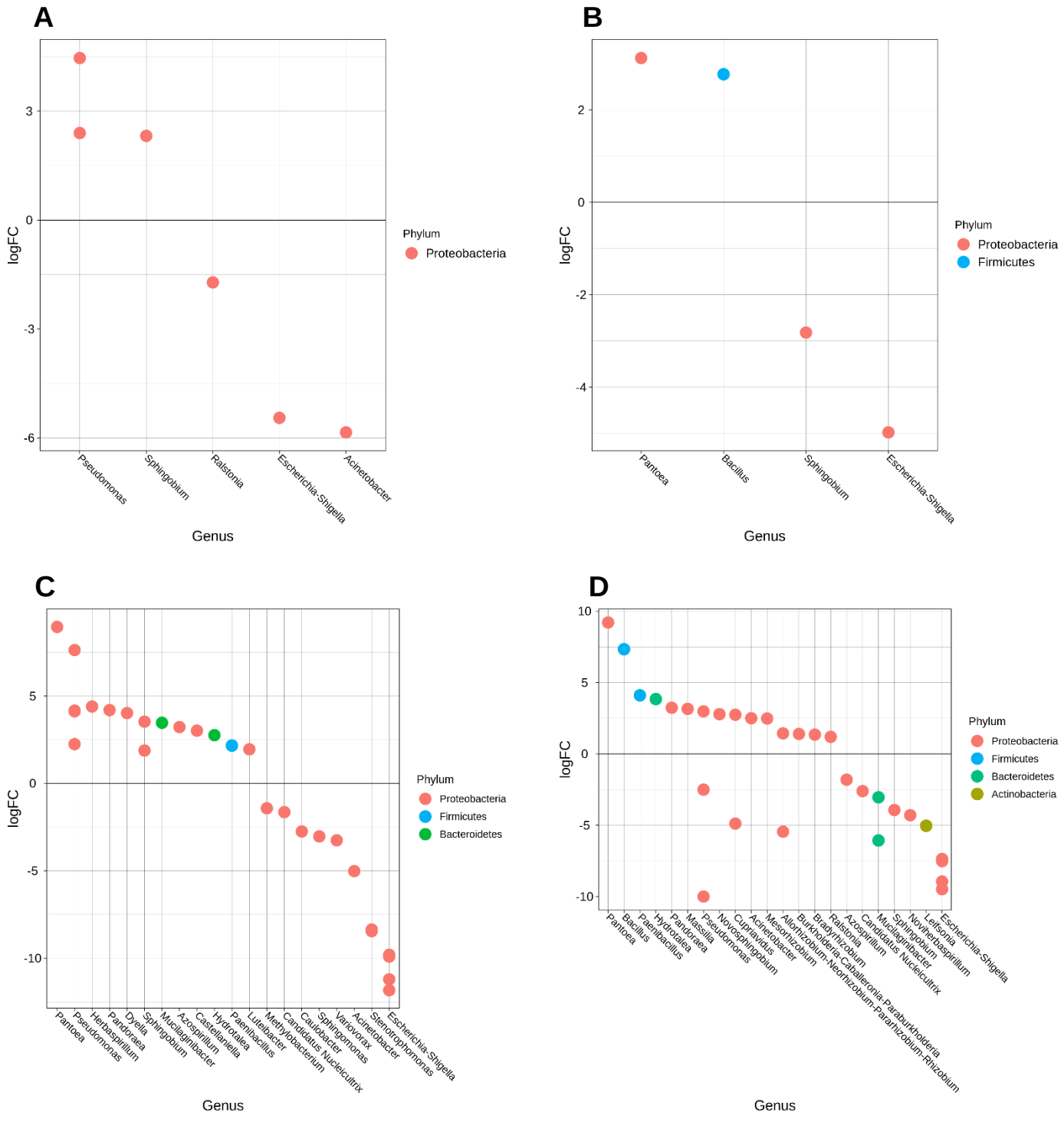

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Relative Abundance (%)
NT-Un151R
NT-Un154R
NT-Un157R
T-Un170R
T-Un172R
T-Un175R
NT-Un155R
NT-Un156R
NT-Un158R
T-Un173R
T-Un174R
T-Un176R
NT-Un151R
NT-Un153R
NT-Un157R
T-Un172R
T-Un173R
T-Un177R
NT-Un150R
NT-Un154R
NT-Un159R
T-Un171R
T-Un174R
T-Un176R
7
14
21
28
DAP
Unassigned
Proteobacteria
Stenotrophomonas
Dyella
Pseudomonas
Acinetobacter
Legionella
Escherichia-Shigella
Neisseriaceae
Methylophilaceae
Methylophilus
Burkholderiaceae
Variovorax
Ralstonia
Noviherbaspirillum
Massilia
Herbaspirillum
Cupriavidus
Burkholderia-Caballeronia-Paraburkholderia
Achromobacter
Sphingomonadaceae
Sphingopyxis
Sphingomonas
Sphingobium
Novosphingobium
Rickettsiaceae
Xanthobacteraceae
Rhodopseudomonas
Bradyrhizobium
Rhizobiaceae
Mesorhizobium
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium

792

Figure 4: Stacked bar plots showing percent relative abundance of bacterial populations in mizuna mustard, leaves (L) and roots (R), sanitized (A) and (C) or unsanitized (B) and (D) respectively over time. Samples were either treated (T) or non-treated (NT) with *E. coli.* Numbers on X-axis represent days after planting (DAP).

793

794


A
logFC
3
0
-3
-6
Pseudomonas
Sphingobium
Ralstonia
Escherichia-Shigella
Acinetobacter
Genus
Phylum
Proteobacteria
B
logFC
2
0
-2
-4
Pantoea
Bacillus
Sphingobium
Escherichia-Shigella
Genus
Phylum
Proteobacteria
Firmicutes
C
logFC
5
0
-5
-10
Pantoea
Pseudomonas
Herbaspirillum
Pandoraea
Dyella
Sphingobium
Mucilaginibacter
Azospirillum
Castellaniella
Hydrotalea
Paenibacillus
Lutelbacter
Methylobacterium
Candidatus Nucleicultrix
Caulobacter
Sphingomonas
Variovorax
Acinetobacter
Stenotrophomonas
Escherichia-Shigella
Genus
Phylum
Proteobacteria
Firmicutes
Bacteroidetes
D
logFC
10
5
0
-5
-10
Pantoea
Bacillus
Paenibacillus
Hydrotalea
Pandoraea
Massilia
Pseudomonas
Novosphingobium
Cupriavidus
Acinetobacter
Mesorhizobium
Allorhizobium
Burkholderia
Bradyrhizobium
Ralstonia
Azospirillum
Candidatus Nucleicultrix
Mucilaginibacter
Sphingobium
Novitherbaspirillum
Leifsonia
Escherichia-Shigella
Burkholderia-Caballeronia-Paraburkholderia
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Genus
Phylum
Proteobacteria
Firmicutes
Bacteroidetes
Actinobacteria

795

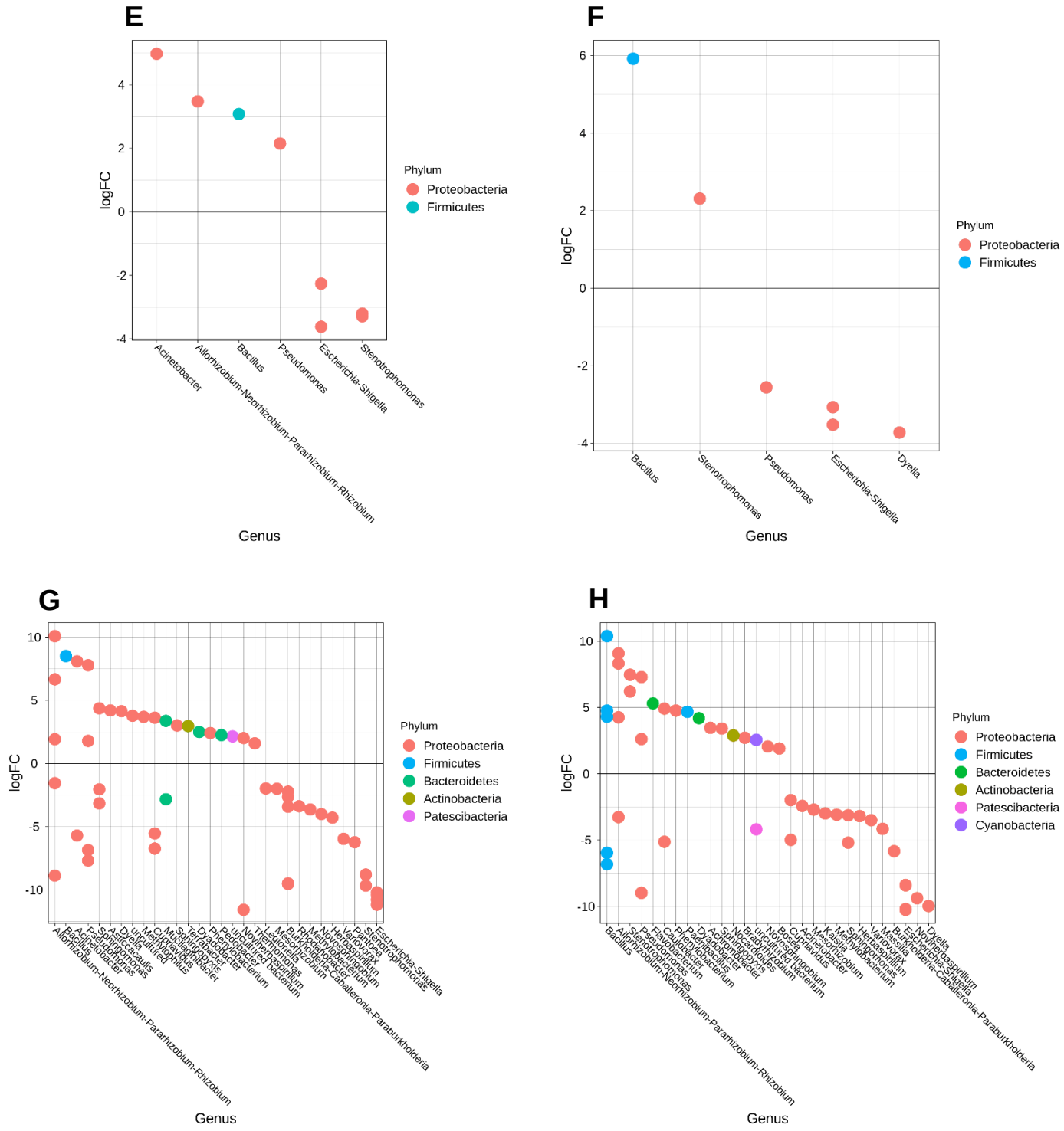

E
logFC
Genus
Phylum
Proteobacteria
Firmicutes
Acinetobacter
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Bacillus
Pseudomonas
Escherichia-Shigella
Stenotrophomonas
F
logFC
Genus
Phylum
Proteobacteria
Firmicutes
Bacillus
Stenotrophomonas
Pseudomonas
Escherichia-Shigella
Dyella
G
logFC
Genus
Phylum
Proteobacteria
Firmicutes
Bacteroidetes
Actinobacteria
Patescibacteria
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Escherichia-Shigella
Stenotrophomonas
Burkholderia-Caballeronia-Paraburkholderia
H
logFC
Genus
Phylum
Proteobacteria
Firmicutes
Bacteroidetes
Actinobacteria
Patescibacteria
Cyanobacteria
Bacillus
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium
Pseudomonas
Escherichia-Shigella
Burkholderia-Caballeronia-Paraburkholderia
Dyella

Figure 5 Differential abundance plots showing bacterial genera that are significantly different in red romaine lettuce (A-D) and mizuna mustard (E-H). Panels represents leaves from either sanitized (A, E) or unsanitized (B, D), and roots from, either sanitized (C, G) or unsanitized (F, H) seeds across the time course. The colors indicate the phylum each genus belongs to with significant differential abundance determined at the $P<0.05$

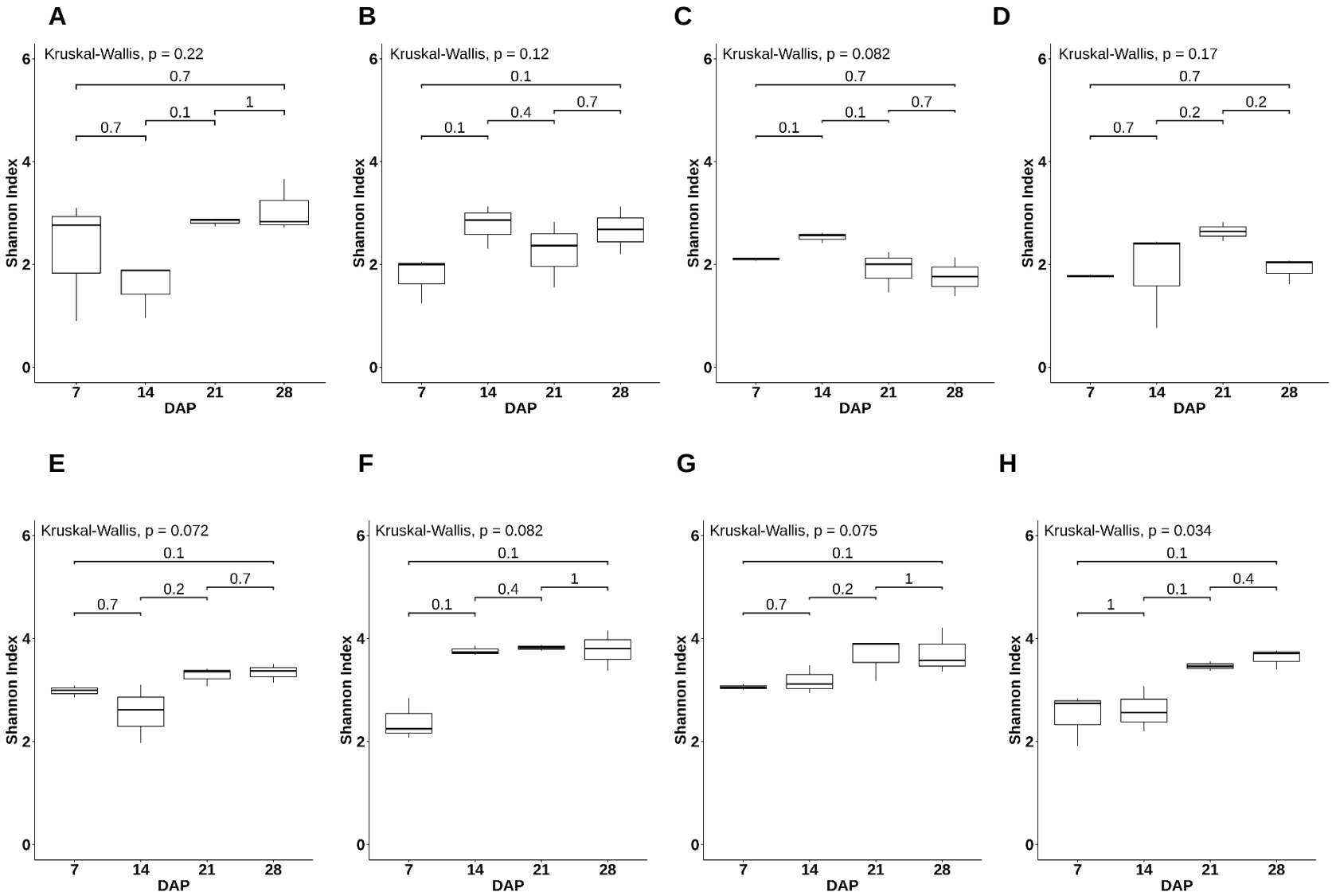


Figure 6. Boxplots representing alpha diversity in red romaine lettuce (*Lactuca sativa*). Top graphs (A-D) represent lettuce leaves while bottom graphs (E-H) represent lettuce root. The sequence of treatment across each row is Sanitized, *E. coli* treated (A, E); Sanitized, not treated (B, F); Unsanitized, *E. coli* treated (C, G); and Unsanitized, not treated (D, H) plant tissue. Alpha diversity was determined with 16S rRNA sequencing on the Illumina MiSeq NGS.

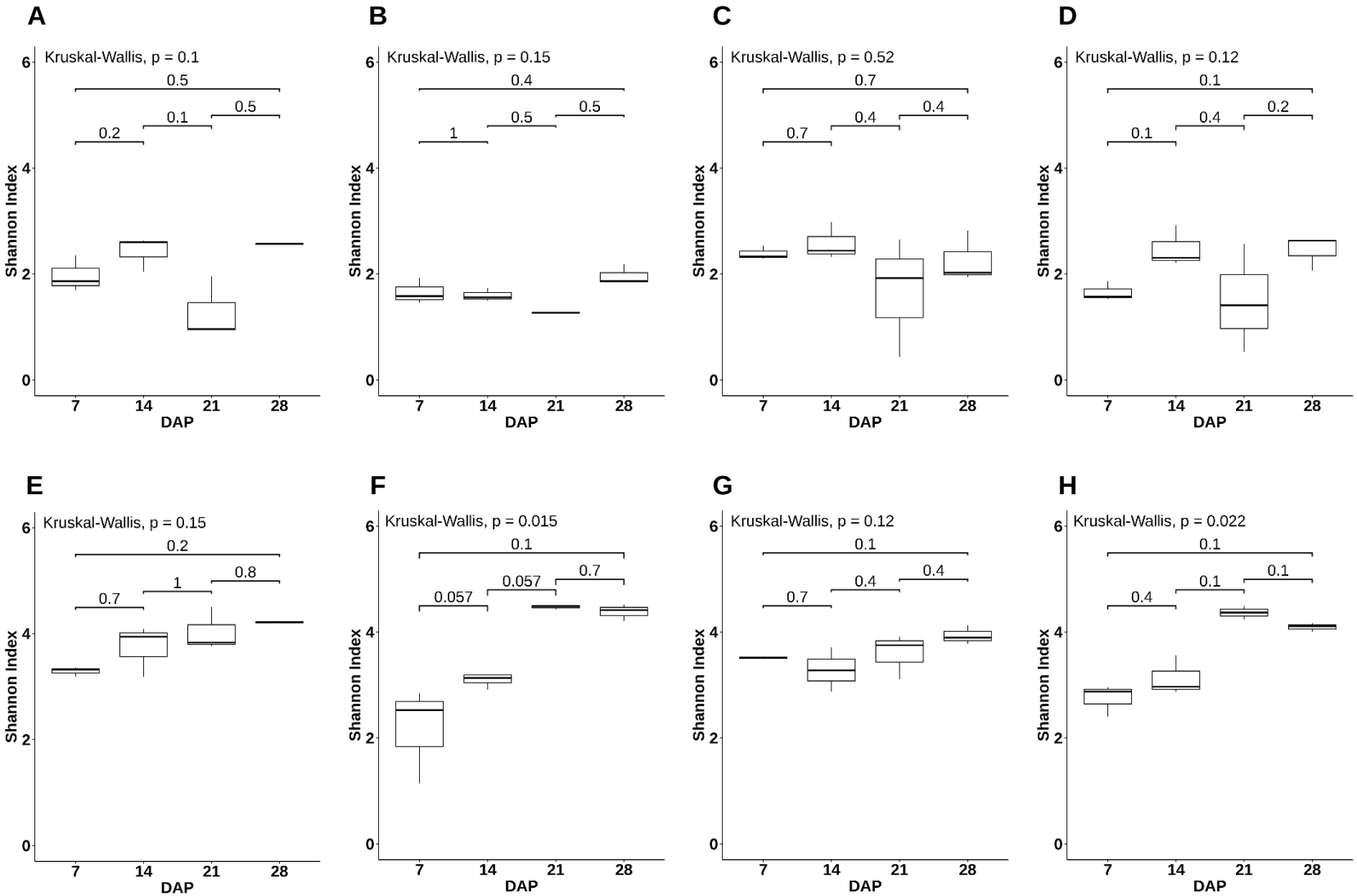


Figure 7: Boxplots representing alpha diversity in mizuna mustard. Top graphs (A-D) represent mizuna leaves while bottom graphs (E-H) represent mizuna roots. The sequence of treatment across each row is Sanitized, *E. coli*-treated (A, E); Sanitized, non-treated (B, F); Unsanitized, *E. coli*-treated (C, G); and Unsanitized, non-treated (D, H) plant tissue. Alpha diversity was determined with 16S rRNA sequencing on the Illumina MiSeq NGS.

802

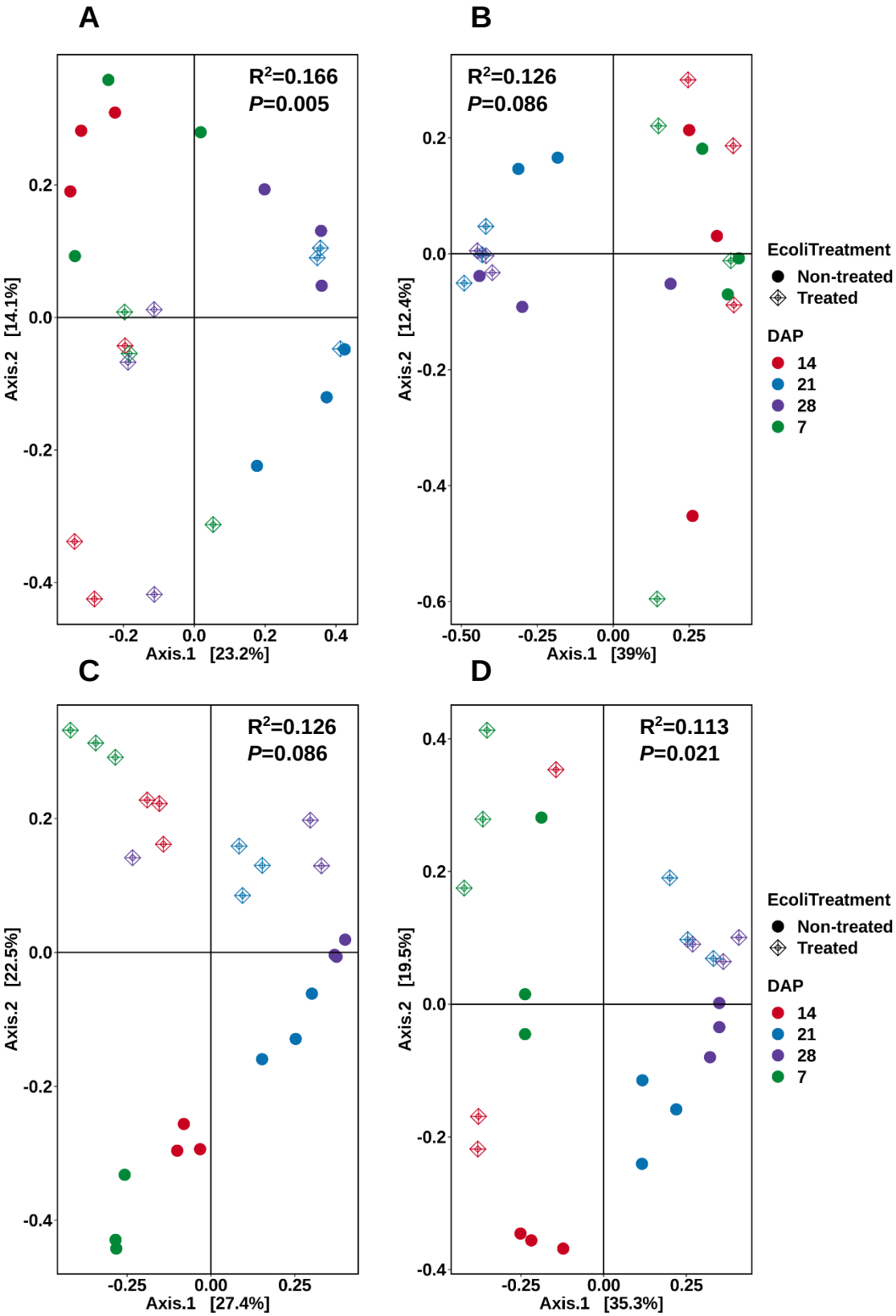


Figure 8: PCoA plots of Bray-Curtis beta diversity for red romaine lettuce by time point - *E. coli* treatment comparison. A=sanitized seed leaf; B=unsanitized seed leaf; C=sanitized seed root; D=unsanitized seed root.

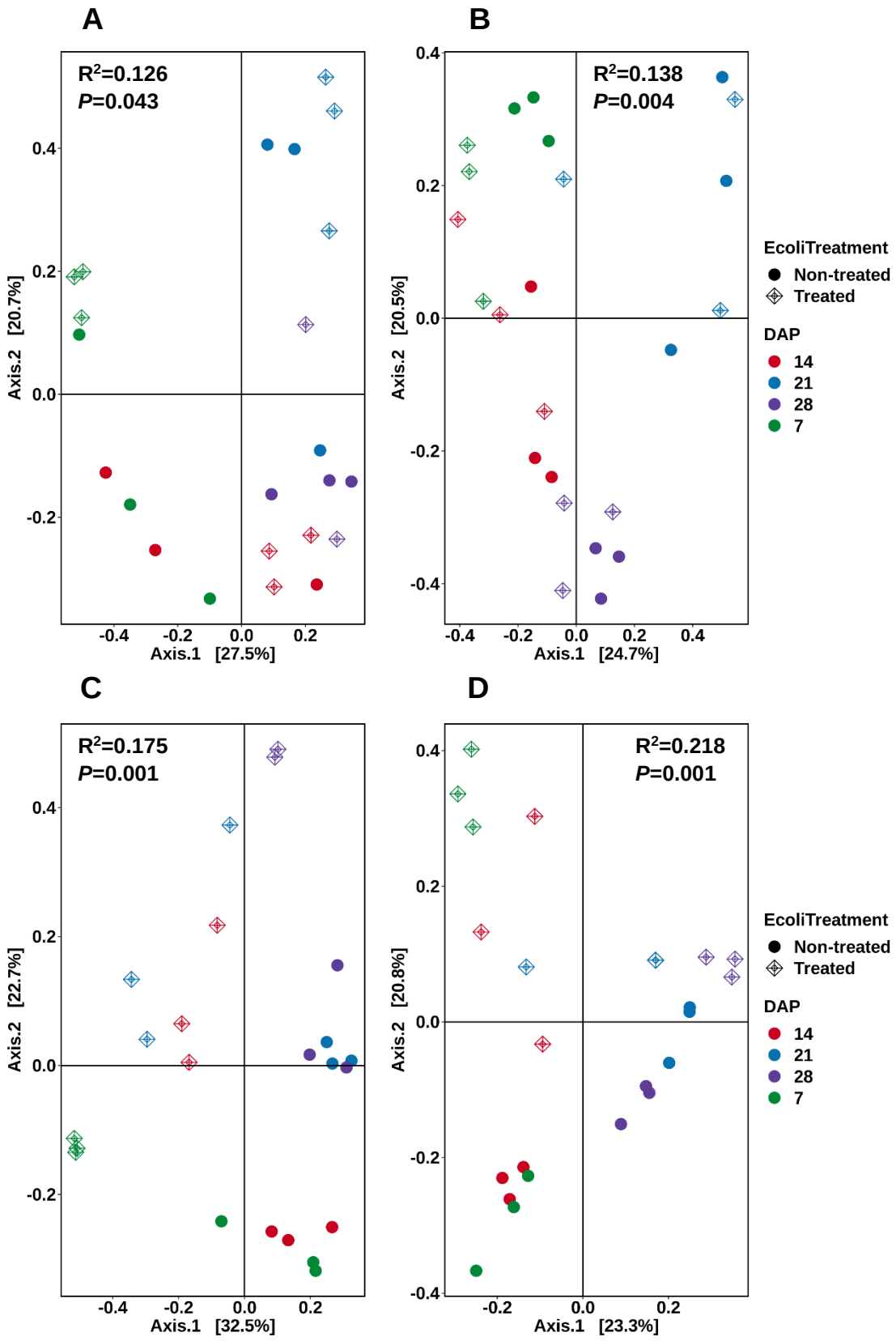
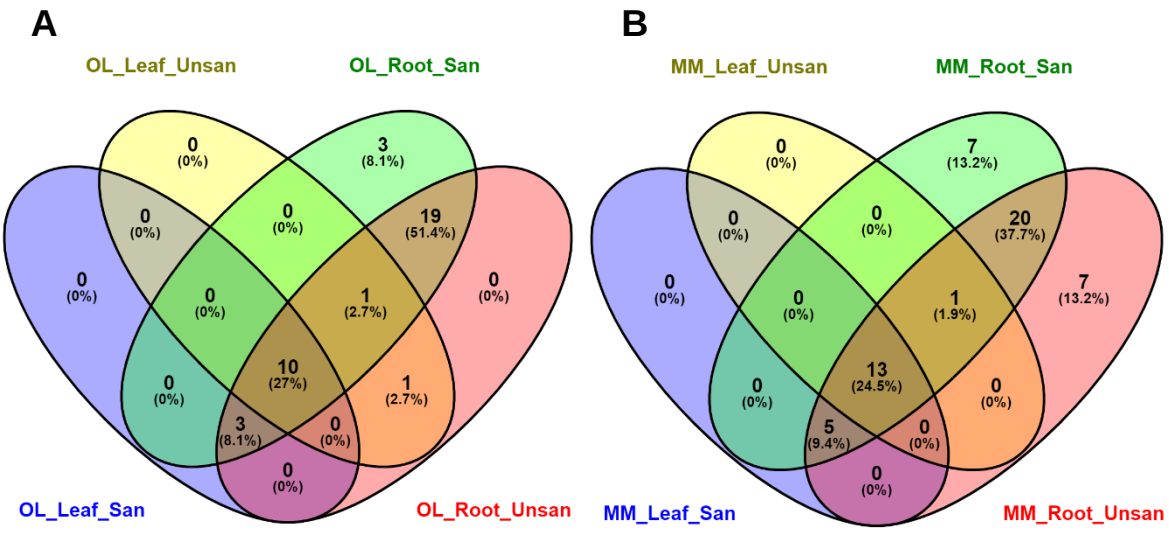


Figure 9: PCoA graphs for Beta diversity in mizuna mustard with time (DAP) versus *E. coli* treatment. A=sanitized seed leaf; B=unsanitized seed leaf; C=sanitized seed root; D=unsanitized seed root.



Core Microbiomes	
Red Romaine Lettuce	**Mizuna Mustard**
Acinetobacter	*Acinetobacter*
Bradyrhizobium	*Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*
Burkholderia-Caballeronia-Paraburkholderia	*Azospirillum*
Escherichia-Shigella	*Bacillus*
Massilia	*Escherichia-Shigella*
Methylophilus	*Massilia*
Pseudomonas	*Methylobacterium*
Ralstonia	*Methylophilus*
Sphingobium	*Ralstonia*
Sphingomonas	*Sphingobium*
	Sphingomonas

Supplemental Figure 1: Core microbiomes for red romaine lettuce and mizuna mustard. (A) Venn diagram showing common genera between sanitized and unsanitized, leaf and root tissues of read romaine lettuce. (B) Venn diagram showing common genera between sanitized and unsanitized, leaf and root tissues of mizuna mustard. Table shows genera representing “core microbiome” for red romaine lettuce (27% from Venn diagram) and mizuna mustard (20.8% from Venn diagram).

Supplemental Table 1: Adonis test of factors (TimePoint, *E.coli* treatment and combined effects) explaining distance variation in Bray-Curtis dissimilarity for lettuce and mizuna.

	TimePoint		*E. coli* Treatment		TimePoint + *E. coli* Treatment	
	R^2	P	R^2	P	R^2	P
LETTUCE						
Sanitized Leaf	0.301	0.001	0.093	0.001	0.166	0.005
Unsanitized Leaf	0.401	0.001	0.076	0.020	0.126	0.086
Sanitized Root	0.348	0.001	0.198	0.001	0.184	0.001
Unsanitized Root	0.462	0.001	0.126	0.001	0.113	0.021
MIZUNA MUSTARD						
Sanitized Leaf	0.426	0.001	0.064	0.012	0.126	0.043
Unsanitized Leaf	0.422	0.001	0.095	0.001	0.138	0.004
Sanitized Root	0.293	0.001	0.269	0.001	0.175	0.001
Unsanitized Root	0.293	0.001	0.160	0.001	0.218	0.001