

RESEARCH

Open Access



Persistence of *Escherichia coli* in the microbiomes of red Romaine lettuce (*Lactuca sativa* cv. 'Outredgeous') and mizuna mustard (*Brassica rapa* var. japonica) - does seed sanitization matter?

Anirudha R. Dixit¹, Christina L. M. Khodadad¹, Mary E. Hummerick¹, Cory J. Spem¹, LaShelle E. Spencer¹, Jason A. Fischer¹, Aaron B. Curry¹, Jennifer L. Gooden¹, Gretchen J. Maldonado Vazquez², Raymond M. Wheeler³, Gioia D. Massa³ and Matthew W. Romeyn^{3*}

Abstract

Background: Seed sanitization via chemical processes removes/reduces microbes from the external surfaces of the seed and thereby could have an impact on the plants' health or productivity. To determine the impact of seed sanitization on the plants' microbiome and pathogen persistence, sanitized and unsanitized seeds from two leafy green crops, red Romaine lettuce (*Lactuca sativa* cv. 'Outredgeous') and mizuna mustard (*Brassica rapa* var. japonica) were exposed to *Escherichia coli* and grown in controlled environment growth chambers simulating environmental conditions aboard the International Space Station. Plants were harvested at four intervals from 7 days post-germination to maturity. The bacterial communities of leaf and root were investigated using the 16S rRNA sequencing while quantitative polymerase chain reaction (qPCR) and heterotrophic plate counts were used to reveal the persistence of *E. coli*.

Result: *E. coli* was detectable for longer periods of time in plants from sanitized versus unsanitized seeds and was identified in root tissue more frequently than in leaf tissue. 16S rRNA sequencing showed dynamic changes in the abundance of members of the phylum *Proteobacteria*, *Firmicutes*, and *Bacteroidetes* in leaf and root samples of both leafy crops. We observed minimal changes in the microbial diversity of lettuce or mizuna leaf tissue with time or between sanitized and unsanitized seeds. Beta-diversity showed that time had more of an influence on all samples versus the *E. coli* treatment.

Conclusion: Our results indicated that the seed surface sanitization, a current requirement for sending seeds to space, could influence the microbiome. Insight into the changes in the crop microbiomes could lead to healthier plants and safer food supplementation.

Keywords: Microbiome, Phyllosphere, Rhizosphere, Red Romaine lettuce, Mizuna mustard, *E. coli*, ISS, Seed surface sanitization

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

*Correspondence: matthew.w.romeyn@nasa.gov

³ NASA UB, Kennedy Space Center, Merritt Island, FL, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

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 plant 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], 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]. Examples such as *Pseudomonas*, *Ralstonia*, and *Agrobacterium* pathovars can cause significant damage in crop plants and all have economic impacts [2]. 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].

National Aeronautics and Space Administration (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]. These initial studies utilized the 16S rRNA gene and the inter-spatial transcribed sequence (ITS) gene to identify the community players. However, the role of these microbes as members of the microbiome 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.

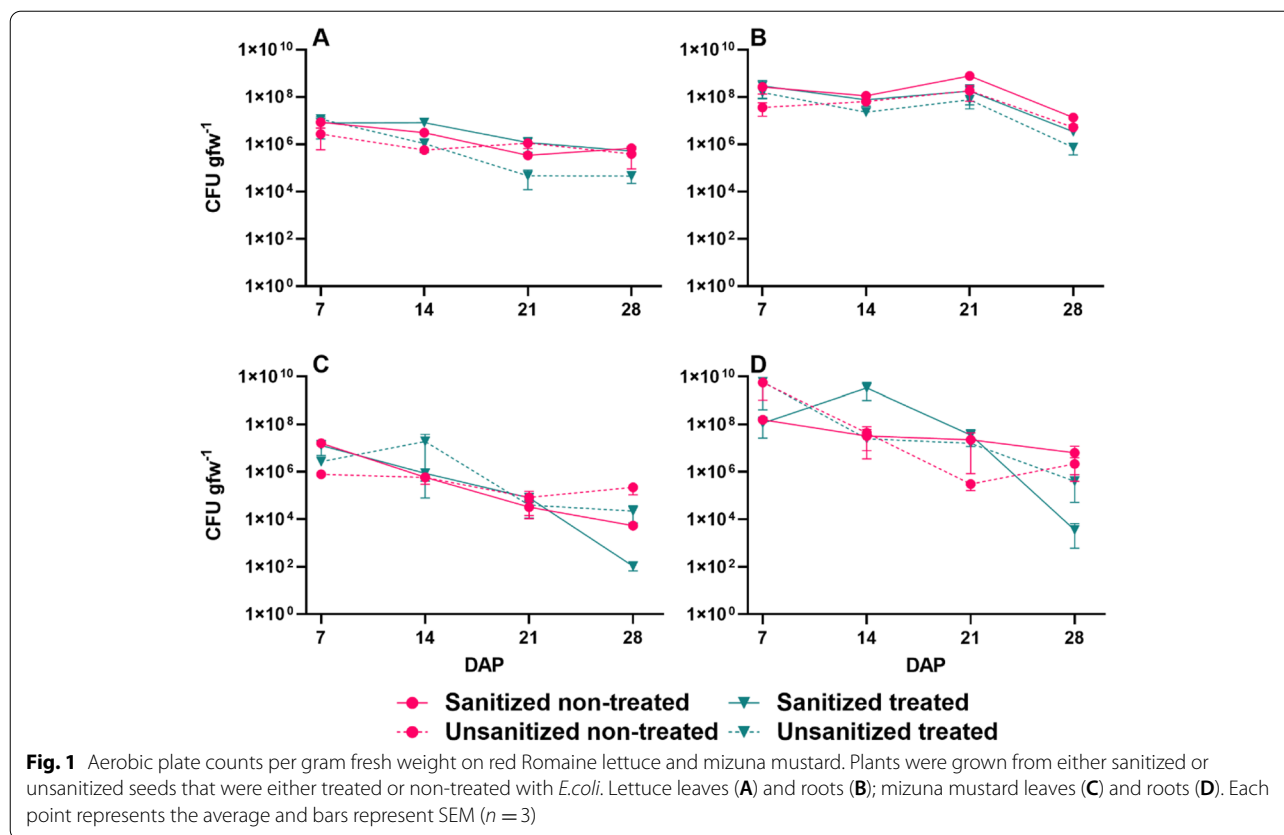
The aim of this study was to investigate the microbiomes of the plant's phyllosphere, rhizosphere, and respective endosphere when challenged with *E. coli*, strain ATCC 11775 as a surrogate for the pathogenic *E. coli* O157:H7 [17, 18] and grown under simulated ISS environmental conditions. Seeds from two leafy green crops that have previously grown on the ISS, specifically red Romaine lettuce, *Lactuca sativa* cv. 'Outredgeous,' and mizuna mustard, *Brassica rapa* var. *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 germinated 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.

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 (Trypticase Soy Agar) generally declined through the growth period from day 7 to day 28. On the lettuce leaves, the average APCs at day 7 were in the range of 2.69×10^6 – 1.15×10^7 colony forming units (CFU) gfw^{-1} for plants from sanitized and unsanitized seeds, both treated and non-treated with *E. coli*. Bacterial numbers declined approximately 1 order of magnitude on leaf samples from all the treatments except for the unsanitized seeds treated with *E. coli*. The counts on these samples were 1.15×10^7 CFU gfw^{-1} at day 7 declining to significantly lower levels ($P = 0.03$) when compared to the leaf tissue from sanitized seeds treated with *E. coli* reaching 4.62×10^4 CFU gfw^{-1} at 21 days and leveling off through day 28 (Fig. 1A).

Similar to the bacterial counts from the lettuce, average counts for mizuna leaf samples at day 7 ranged from 7.9×10^5 CFU gfw^{-1} on the non-treated seeds that were not sanitized to the highest level of 1.31×10^7 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. Leaves from unsanitized, *E. coli* non-treated seeds had the highest CFU counts, 2.23×10^5 after 28 days followed by the plants from seeds that were not sanitized but treated



with *E. coli* at 2.2×10^4 CFU gfw⁻¹. The lowest counts for the day 28 leaf samples were from seeds that were sanitized and treated with *E. coli* with average counts of 1.05×10^2 CFU gfw⁻¹, significantly lower ($P = 0.03$) than the leaves from the unsanitized *E.coli* treated seeds (Fig. 1C).

Bacterial counts from roots of both leafy greens were higher than leaves, ranging from 3.62×10^7 to 2.88×10^8 CFU gfw⁻¹ in lettuce and 1.14×10^8 to 6.07×10^9 CFU gfw⁻¹ in mizuna at day 7. 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. 1B). After 28 days of growth, the counts on mizuna roots from sanitized *E.coli* treated seeds was 3.62×10^3 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 treated sanitized and unsanitized seeds and plant tissues from treated sanitized and unsanitized seeds was enumerated using culture-independent quantitative PCR. *E. coli* cells per seed from each treatment and seed type after incubation in TSB (Trypticase Soy Broth) (T=0) were 5.3×10^4 for sanitized lettuce, 8.2×10^4 for

unsanitized lettuce seeds, 4.3×10^4 for sanitized mizuna and 2.5×10^4 for the unsanitized mizuna.

This assay demonstrated a decline in *E. coli* abundance over the 28-day growth period in all plant tissues tested. In lettuce leaves, at days 7–21 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 3 orders of magnitude from 3.2×10^5 at day 7 to 8.11×10^2 cells gfw⁻¹ at harvest (day 28), whereas *E. coli* levels for leaves from the unsanitized seeds remained undetectable after 21 days. At day 21, *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 day 28 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. Roots from sanitized and unsanitized seeds, showed a decline in average cell counts gfw⁻¹ from 1.36×10^6 to 1.86×10^3 and 7.28×10^5 to 4.39×10^2 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 4.77×10^7 cells

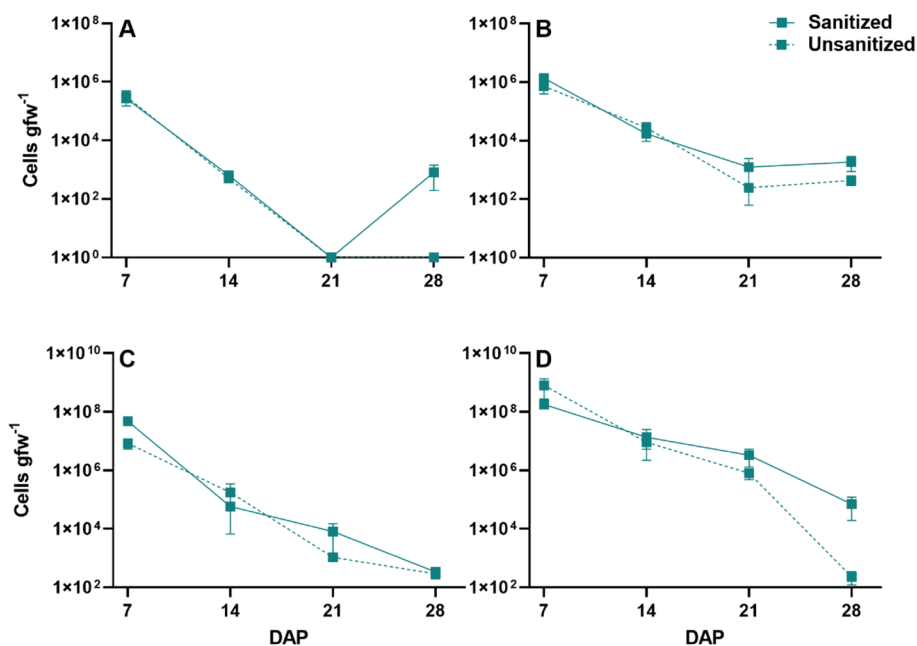


Fig. 2 *E. coli* counts per gram fresh weight for red Romaine lettuce and mizuna mustard as determined by qPCR. Plants were grown from either sanitized or unsanitized seeds that were treated with *E. coli*. Lettuce leaf (A) and roots (B); mizuna mustard leaves (C) and roots (D). Each point represents the average and bars represent SEM ($n = 3$)

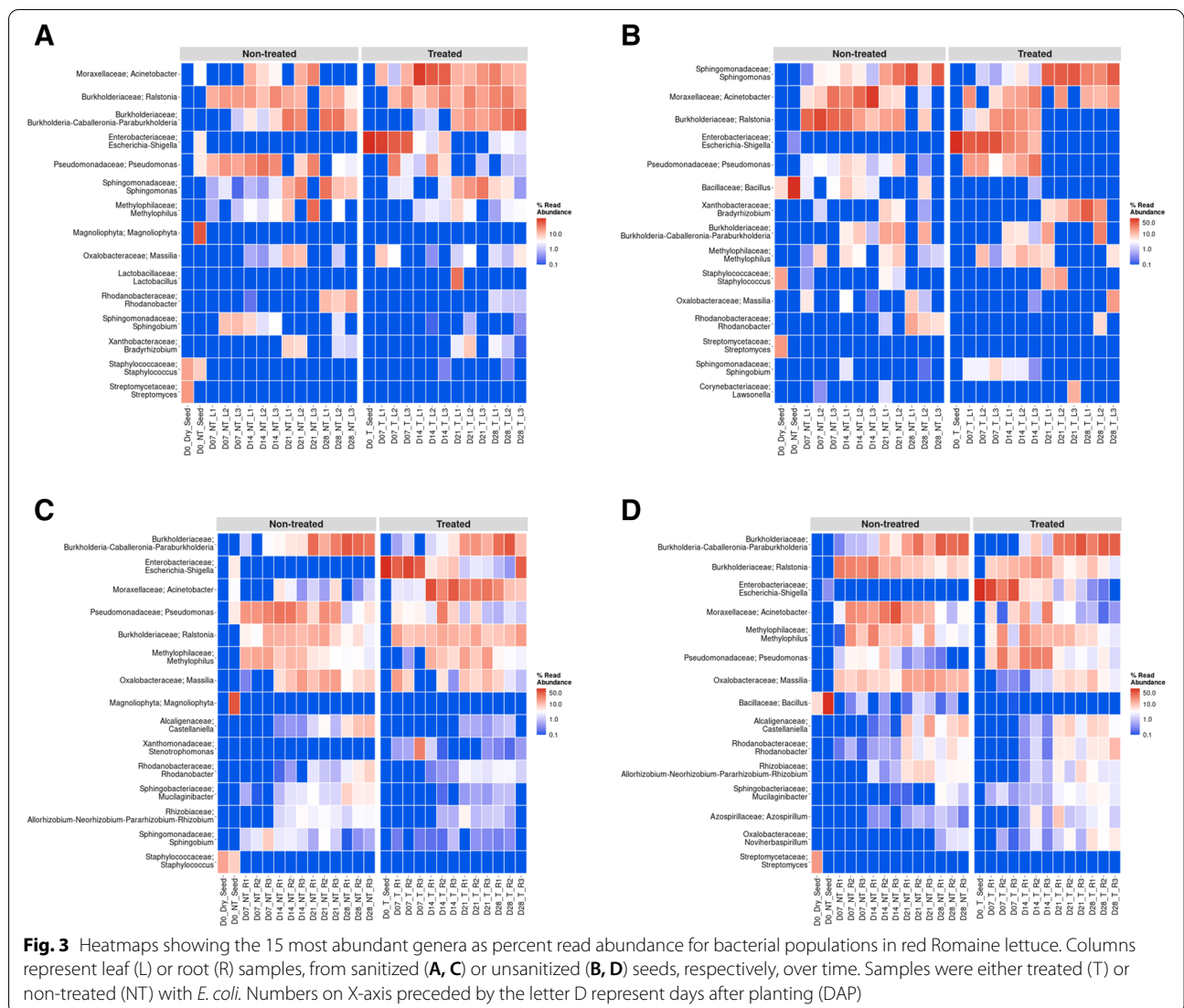
gfw^{-1} to 2.90×10^2 cells gfw^{-1} in the leaves from sanitized seeds and 8.34×10^6 cells gfw^{-1} to 2.9×10^2 cells gfw^{-1} in the leaves from unsanitized seeds. There were no differences in the qPCR data between seed sanitization treatments through the time course however, *E. coli* was detected in both mizuna leaf samples at day 28 (Fig. 2C). Mizuna roots from sanitized and unsanitized seeds showed no significant difference in *E. coli* counts between the treatments over time. For the roots from sanitized and unsanitized seeds, the average cell counts gfw^{-1} declined from 1.78×10^8 to 7.05×10^4 and 7.96×10^8 to 2.3×10^2 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. After removal of chloroplast and mitochondrial taxonomic assignments, there were 3,713,493 and 3,691,028 high-quality sequences for lettuce and mizuna, respectively.

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 ($\sim 10\%$) 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 & B). 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. 4A & B).

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 *Sphingobium* and *Stenotrophomonas*, which were present exclusively in the plants generated from sanitized seed whereas the genera *Bacillus* and *Azospirillum* were present only in the unsanitized samples (Fig. 3C & D). *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

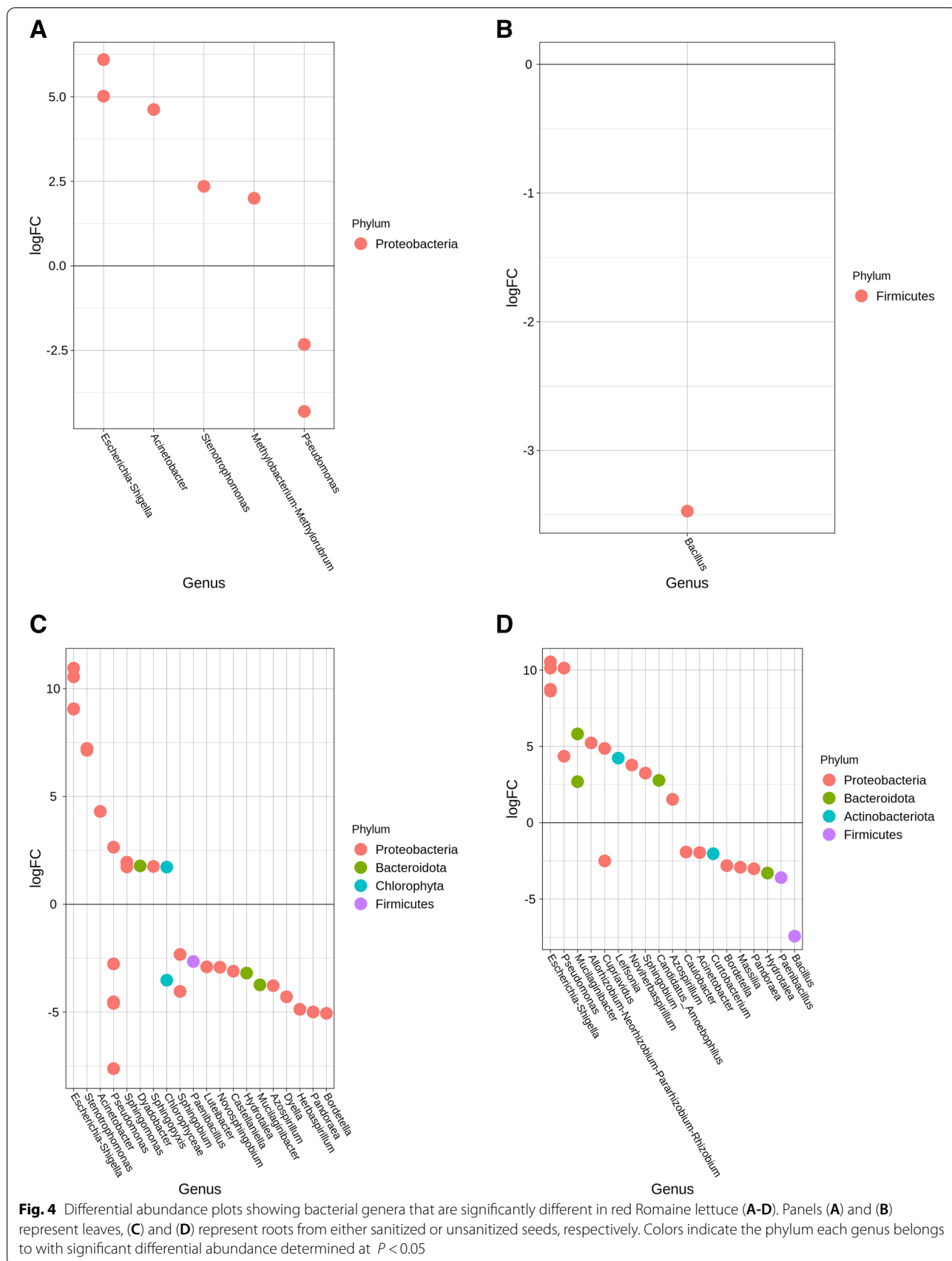


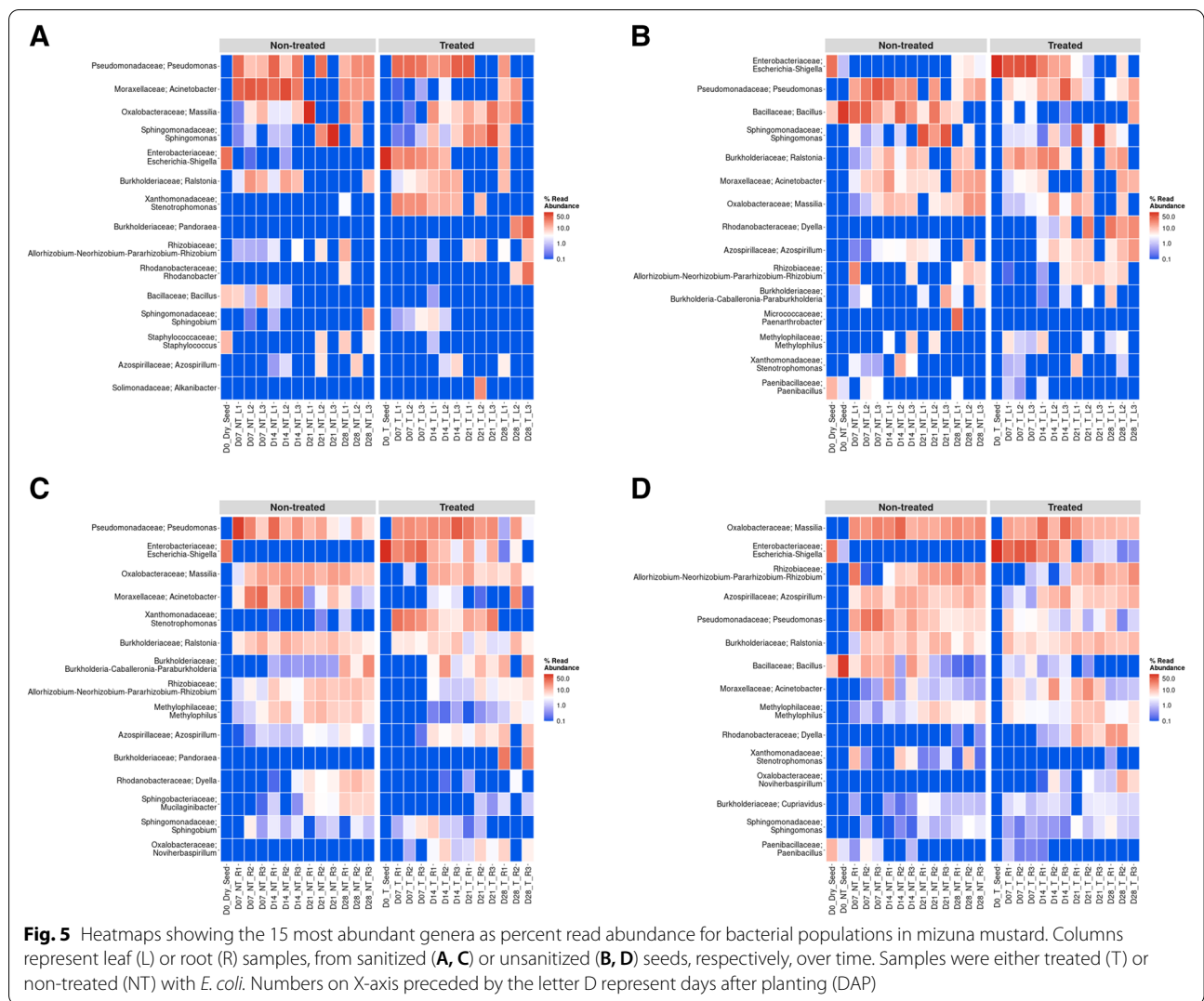
decreasing to less than 1% by day 28. Similar to leaves, genera such as *Burkholderia-Caballeronia-Paraburkholderia*, *Acinetobacter*, *Ralstonia*, *Methylophilus*, *Pseudomonas*, and *Massilia* exhibited dynamic changes in relative abundance, sometimes showing an inverse relationship to *E. coli* (Fig. 3C & D). Between lettuce roots from sanitized and unsanitized seeds, a dominance of phyla *Proteobacteria*, *Firmicutes*, and *Bacteroidetes* was observed. In particular, genera such as *Pantoea*, *Pseudomonas*, and *Sphingobium* showed an increased abundance whereas *Sphingomonas*, *Acinetobacter*, *Stenotrophomonas*, and *Escherichia-Shigella* showed an overall decrease in abundance throughout the study (Fig. 4C and D).

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 *Dyella*, and *Rhodanobacter* being exclusively present

in unsanitized samples with *Pandora* and *Sphingobium* present in sanitized 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. 5A & B). 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. 6A & B).

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 *Acinetobacter* showed dynamic changes in their relative abundance especially in day 7 and day 14 (Fig. 5C & D). 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*, *Bacteroidetes*, and *Actinobacteria* (Fig. 6C & D).

Diversity within bacterial communities

Alpha and beta diversity provided overall community characterization of abundance and distribution within and between samples over time. Alpha diversity, determined by the Shannon Index, revealed that diversity for both sanitized and unsanitized seed-generated plant tissues generally fell between 1.0 and 4.0 for both red Romaine lettuce and mizuna mustard, which were within normal range for environmental samples [17–19]. There was no significant difference in diversity between time points for all lettuce *E. coli* treated and non-treated samples for both sanitized and unsanitized seeds with the exception of between day 7 and day 28 in sanitized *E. coli* treated leaf tissue ($P = 0.033$) and unsanitized *E. coli* non-treated lettuce root tissue ($P = 0.034$). A detailed investigation into the lettuce samples indicated that there was an increase in the number of genera with time and nearly

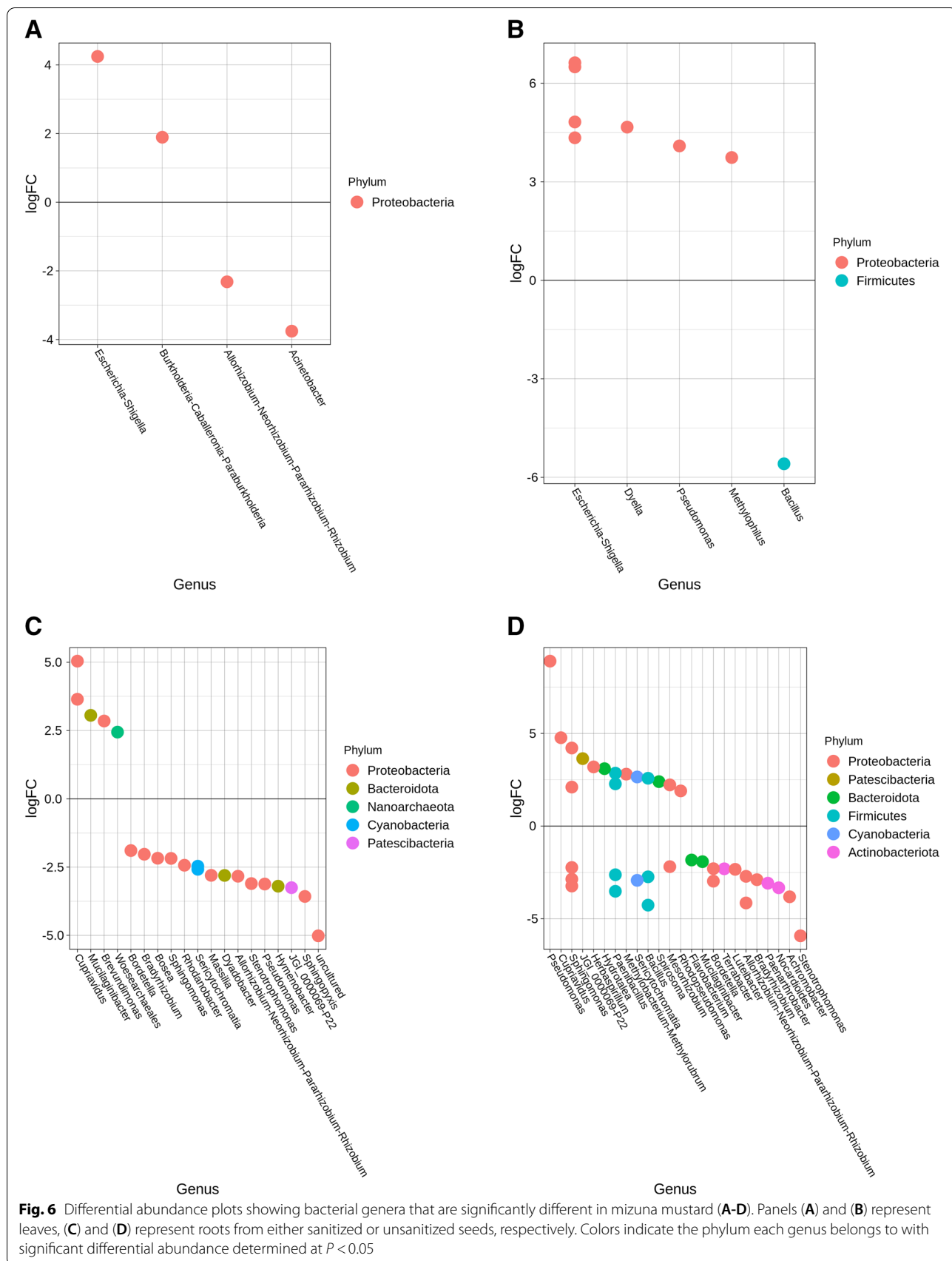


Fig. 6 Differential abundance plots showing bacterial genera that are significantly different in mizuna mustard (A-D). Panels (A) and (B) represent leaves, (C) and (D) represent roots from either sanitized or unsanitized seeds, respectively. Colors indicate the phylum each genus belongs to with significant differential abundance determined at $P < 0.05$

three times more genera in the day 28 replicates than in the day 7 replicates (Fig. S1).

The comparison between time points in each treatment type for mizuna mustard showed no significant differences in the overall alpha diversity values in leaf tissue ($P > 0.05$) with two exceptions in mizuna root tissue. A closer inspection of the alpha diversity of mizuna roots indicated that there was a similar increase in diversity 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 days 7 and 28 (Fig. S2).

Beta diversity determines the differentiation between samples. The Bray-Curtis statistic (R^2) 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 leaves grown from the sanitized seeds presented a significant difference ($P < 0.05$) (Fig. 7A). The differences explained in two axes (time and *E. coli* exposure) were low, at 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 microbiomes that are more similar. 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 ($P = 0.001$) (Table 1). 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. 7B). 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 right quadrants inferring that time may have a greater influence on the microbiome. The variation in the unsanitized seed-generated leaf tissue was not significant ($P = 0.086$) (Fig. 7B). 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% (Table 1). Both were significant ($P = 0.001$ and $P = 0.02$, respectively).

Figure 7C (sanitized) and 7D (unsanitized) illustrate the beta diversity of the lettuce root tissue. The PCoA in Fig. 7C and D accounted for 50 and 55% of the variation, respectively. The beta 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. 7C). The replicates from each of the non-treated tissues also grouped tightly together, but were separated by time, with day 7 and 14 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 ($P = 0.001$) (Table 1). The diversity within root tissues generated from unsanitized seeds showed less clustering, indicating more dissimilarity than the sanitized seed-generated root tissue with day 7 root tissue presenting the most dissimilarity (Fig. 7D). Days 7 and 14 *E. coli* treated plants did not show any clustering, and were therefore dissimilar, while days 21 and 28 grouped more closely to each other but differently from days 7 and 14 (Fig. 7D). This indicates a similarity to each other but more variation between time points, which could account for 46.2% of the variation while the *E. coli* treatment accounted for 13%. Both were significant ($P = 0.001$) (Table 1).

Beta diversity in the mizuna leaf tissue from sanitized seeds indicated little similarity between the *E. coli*-treated and non-treated leaves. The PCoA for sanitized and unsanitized seed generated leaf tissue explained 48.2 and 44.7% of the variation, respectively. (Fig. 8A and B). Mizuna samples from sanitized, *E. coli*-treated seeds showed some grouping 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% (Table 1). 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 little clustering between mizuna leaf tissue from *E.*

(See figure on next page.)

Fig. 7 PCoA plots for beta diversity (Bray-Curtis dissimilarity) in red Romaine lettuce, with time (DAP) versus *E. coli* treatment. Increasing distance in two-dimensional space represents increasingly dissimilar community structures. Leaves from sanitized (A) and unsanitized (B) seeds. Roots from sanitized (C) and unsanitized (D) seeds. The R^2 value represents percent variance that can be explained by the specified groups (DAP + *E. coli* treatment) as calculated using adonis, a multivariate PERMANOVA test implementation in QIIME2

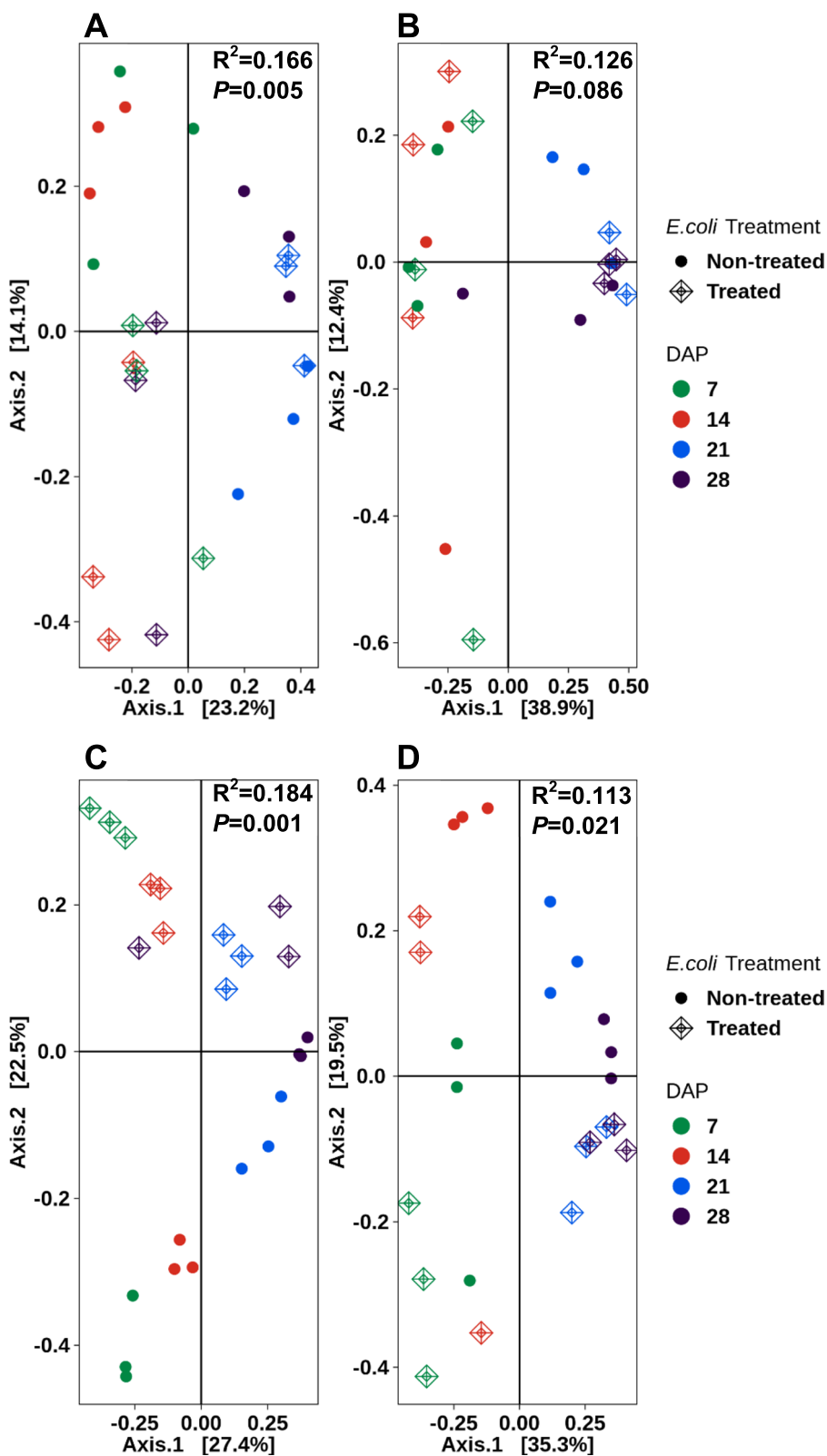


Fig. 7 (See legend on previous page.)

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

	TimePoint		<i>E. coli</i> Treatment		TimePoint + <i>E. coli</i> Treatment	
	R ²	P	R ²	P	R ²	P
Lettuce						
Leaves from sanitized seeds	0.301	0.001	0.093	0.001	0.166	0.005
Leaves from unsanitized seeds	0.401	0.001	0.076	0.020	0.126	0.086
Roots from sanitized seeds	0.348	0.001	0.198	0.001	0.184	0.001
Roots from unsanitized seeds	0.462	0.001	0.126	0.001	0.113	0.021
Mizuna Mustard						
Leaves from sanitized seeds	0.426	0.001	0.064	0.012	0.126	0.043
Leaves from unsanitized seeds	0.422	0.001	0.095	0.001	0.138	0.004
Roots from sanitized seeds	0.293	0.001	0.269	0.001	0.175	0.001
Roots from unsanitized seeds	0.293	0.001	0.160	0.001	0.218	0.001

coli-treated and non-treated seeds with the exception of day 28 indicating a similarity in the microbiomes (Fig. 8B). 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 combined axes for the sanitized seed-grown root tissue explains 55.3% of the variation ($P = 0.001$), while the unsanitized seed-grown root tissue explains 44% of the variation ($P = 0.001$) (Fig. 8C & D). There was some grouping 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 R² values of 0.175 and 0.218 support this outcome (Table 1). 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 [20] 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 [21–23], 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 [24–26]. Links et al. [23] 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 removing 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 the present study, both sanitized and unsanitized seeds of lettuce and mizuna were inoculated with *E. coli* to determine the impact of the seed surface 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, 27], which is consistent with our results for red Romaine lettuce and mizuna mustard through the 28-day growth period. At day 28, red Romaine lettuce from unsanitized, *E. coli* treated seeds

(See figure on next page.)

Fig. 8 PCoA plots for beta diversity (Bray-Curtis dissimilarity) in mizuna mustard, with time (DAP) versus *E. coli* treatment. Increasing distance in two-dimensional space represents increasingly dissimilar community structures. Leaves from sanitized (A) and unsanitized (B) seeds. Roots from sanitized (C) and unsanitized (D) seeds. The R² value represents percent variance that can be explained by the specified groups (DAP + *E. coli* treatment) as calculated using adonis, a multivariate PERMANOVA test implementation in QIIME2

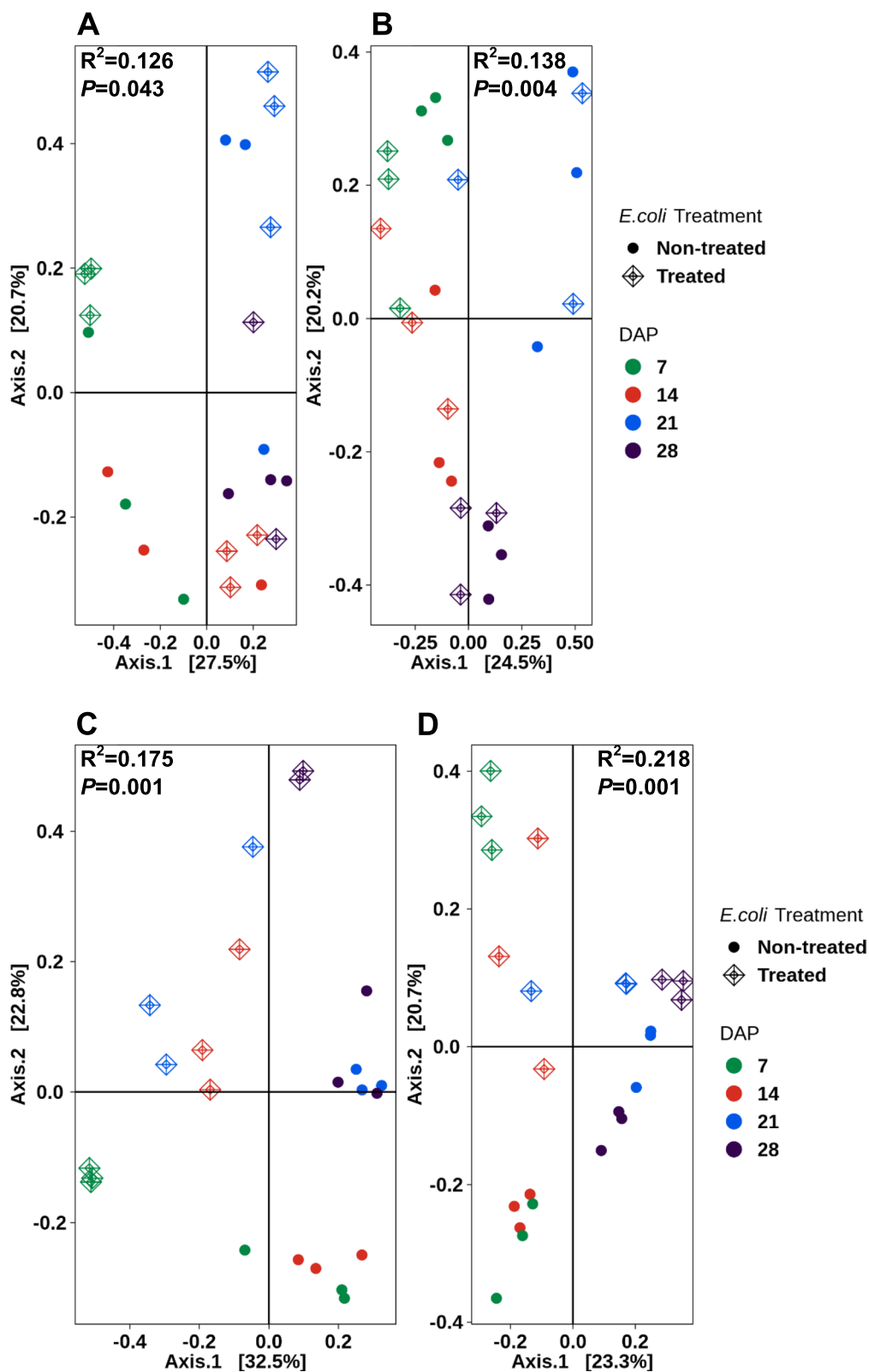


Fig. 8 (See legend on previous page.)

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 28-day 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 [28]. 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 day 28 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 [29, 30] or prevent pathogens from colonizing leaf surfaces [31, 32]. Microbial endophytes in healthy plants would have a positive influence on the host but could also play a role in microbe-microbe interactions [33, 34]. Early colonization would set up or establish varying community compositions in the phyllosphere [35]. 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*, *Bacteroidetes*, 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, 11, 14, 36, 37]. Our results showed that, in red Romaine lettuce roots, the relative abundance of *Actinobacteria* and *Bacteroidetes* gradually increased through different developmental stages, which was consistent with observations by Chaparro et al. [38].

In the case of mizuna mustard, we observed *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Cyanobacteria*, and *Bacteroidetes* as the most dominant phyla in leaves and roots, independent of seed sanitization. This is again consistent with what other studies describe [39]. Relative abundance of *Actinobacteria* and *Bacteroidetes* gradually increased through different developmental stages in mizuna, similar to red Romaine lettuce. 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 [40, 41].

Invasion by pathogenic microbes often correlates with shifts in microbial communities in different plant compartments, including leaves [42] and roots [43]. 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 parts of other plant microbiomes, either as indicators of plant health [44] or their role against invading pathogens [45–47] (Fig. S3). 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 [44], wheat stripe rust [48] and *Rhizoctonia solani* in lettuce plants [45].

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 (Operational Taxonomic Units) with time. Overall, richness was lower in leaf tissue than in root tissue. Studies with model organisms such as *Arabidopsis* have shown similar trends [32]. Additionally, root tissues are more diverse than leaf tissues. 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 days 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; however, these factors were maintained at steady settings in controlled environments, and therefore should 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 phyllosphere, especially at day 7, with diversity increasing with time. At day 7, approximately 10 genera were found in most leaf tissues but by 14–21 days, the number more than doubled with the *E. coli* 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 leaf tissue from non-treated, unsanitized seeds. *Ralstonia* does not

appear to persist in the leaf tissue from *E. coli*-treated, unsanitized seed, however as the *E. coli* declined, *Pseudomonas* and *Burkholderia* both increased, indicating a possible competition for resources. In the leaf tissue from sanitized seeds that are either treated or untreated with *E. coli*, we see *Ralstonia* present and increase with time. In addition, similar to the unsanitized seeds, *Pseudomonas* and *Sphingomonas* both increase in samples from the *E. coli*-treated seed until day 21. *Pseudomonas* parallels *E. coli* and begins to decline whereas *Sphingomonas* continues to increase through day 28. Innerebner et al. [49] have shown that some *Sphingomonas* species have the capacity to outcompete certain *Pseudomonas* species thereby serving as a nutrient-related beneficial microbe.

Root tissue had 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 *Hydrothalea*, *Bacillus*, and *Micrococcus*. *Massilia*, a root-colonizing bacterium 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 [32, 50]. We saw 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 studies could 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 classified as a pseudomonad, *Stenotrophomonas* has been isolated in soil, plant, water and human tissues. In soil and plants, it is known for production of plant hormones and is very persistent in the roots, and once established, is not easily out-competed [51, 52].

Conclusions

Microbial interactions with seeds can influence plant growth and development. Seed surface sanitization is commonly practiced in plant tissue culture and some commercial crop production operations and is a requirement for space crop agriculture for ISS. We investigated the effects of seed surface sanitization on plant-associated microbial communities of two leafy crops, red Romaine lettuce and mizuna mustard with

and without introduction of *E. coli* across four time points. Our results showed that *E. coli* persisted for a longer duration on plants from sanitized seeds compared to unsanitized seeds. Although we did not observe any growth and developmental constraints on plants, we did observe dynamic changes in the leaf- and root-associated microbial communities of both crops due to seed surface sanitization. Our results indicated that the seed surface sanitization, a requirement for sending seeds to space, might be altering the plant-associated microbiome, which could potentially influence the plants' growth and development.

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*) (Johnny's Select Seeds, Fairfield, Maine, United States). Effects of seed sanitization (sanitized vs. unsanitized) and *Escherichia coli*, ATCC 11775 (hereon referred as *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 previously [20, 53]. Approximately 75–100 seeds were batch sanitized by adding 0.5 ml concentrated HCl to 30 ml bleach (5% sodium hypochlorite solution) in a 500 ml wide-mouth jar 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 1 h 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. To verify effectiveness of sanitization, 20 seeds were placed on 2 plates of Trypticase Soy Agar (TSA) and 2 plates of Inhibitory Mold Agar (ISA) thus 5 seeds per plate. Enumeration of bacteria and fungi was also done by placing 20 seeds in 1 ml sterile phosphate buffered saline (PBS) with sterile glass beads (3 mm) vortexing for 2 min and plating in duplicate 100 μ l on TSA and IMA. All plates were incubated at 30 °C for 24–48 h for TSA and up to 5 days for IMA. Plates containing seeds were observed daily for any growth around the seed while colonies were counted at the end of the incubation period from those plated for enumeration. The same analysis was performed on unsanitized controls. CFUs per seed on sanitized seeds were below

detection limits (0.5 CFU/seed) for TSA plates. CFUs per unsanitized seed were 0.5 on red Romaine lettuce and 1.0 on mizuna mustard. 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 both crops were prescreened for the presence of *E. coli* prior to use by 16S rRNA sequencing and determined to be *E. coli* free. 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. Fifty seeds were inoculated by submersion in a 5 ml culture of 10^6 *E. coli* cells ml⁻¹, 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-in.) square pots containing (~500 ml) greens grade, Turface, calcined clay (arcillite) (Ewing Irrigation, St Petersburg, FL, United States) and 7.5 g liter⁻¹ 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 (50% relative humidity, 3000-ppm CO₂, temperature 23°C). The photoperiod was set at 16 h/8 h 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 Dickinson, 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 unique to *E. coli* (5'TAA GCTCCATGTCATTTGCC-3' Forward and 5'TTCCAT GGAGGGTATTCGG-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 and strain specific primers were available. The *ycfR* gene, though not specific to *E. coli*, produces a stress resistant protein, is a biofilm regulator in *E. coli*, and may serve as a surrogate to *E. coli* O157:H7 [54, 55]. A standard curve was generated with DNA isolated from *E. coli* with concentration ranging from 10⁰ to 10⁻⁷ ng of genomic DNA. The qPCR cycling conditions included a 10 min denaturation cycle followed by 45 amplification cycles (95°C for 10 s, 54°C for 1 min, and 72°C for 30 s). The melt curve was determined at 95°C for 5 s, 65°C for 1 min then 95°C). All standards and samples were run in triplicate. Copy number was converted to *E. coli* cell number gfw⁻¹ and log transformed.

16S PCR and sequencing

The 16S PCR was completed with custom barcoded 16S rRNA gene primers (V4 region) as described by Kozich et al. [56]. The final PCR master mix contained 1X PCR buffer, 2.25 mM MgCl₂, 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°C for 5 min followed by 30 cycles at 95°C for 1 min, 58°C for 1 min and 72°C for 2 min. A final elongation step at 72°C 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 the 16S rRNA amplicon sequencing of community samples were analyzed using QIIME2 (ver. 2020.2) [57]. Demultiplexed raw sequences were imported into the DADA2 de-noising algorithm to process the raw sequences into exact sequence variants (ESVs). The DADA2 pipeline performed filtering, de-replication, chimera identification and the merging of paired-end reads [58]. The taxonomies were assigned using SILVA 132 rRNA gene database available for use with QIIME2-feature-classifier plugin [59]. Exact sequence variants affiliated with chloroplasts and mitochondria were removed in order to keep only microbial sequences [60, 61]. We used the 'qiime feature-table filter-features' to select features that are present in a minimum of six samples ($--p\text{-min-samples}$ 6). All other analyses contained a full data set. A differential abundance analysis was carried out by importing essential QIIME2 artifacts using qiime2R [62] into a phyloseq [63] object, followed by the edgeR [64] analysis to identify taxa that are differentially abundant across the time course.

Statistical analysis

Differences between plate counts (CFU) and *E. coli* cell number 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 [65] and beta diversity represented by Bray-Curtis dissimilarity [66] were calculated using the diversity plugin in QIIME2

[57]. Both alpha and beta diversity were calculated using features (in this case OTUs) as frequencies as the default input to the 'qiime diversity alpha/beta' command. 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 [67]. 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 $--p\text{-formula}$ "treatment*time-course" was implemented [68]. 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 $P < 0.05$. Heatmaps were created using ampvis2 package in R [69].

Abbreviations

TSA: Trypticase Soy Agar; TSB: Trypticase Soy Broth; APC: Aerobic Plate Count; gfw: gram fresh weight; CFU: Colony-Forming Unit; ISS: International Space Station; NASA: National Aeronautics and Space Administration; QPCR: Quantitative Polymerase Chain Reaction; OTUs: Operational Taxonomic Units; PCoA: Principle Co-ordinate Analysis; ESVs: Exact Sequence Variants.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12866-021-02345-5>.

Additional file 1: Supplemental Figure S1. Boxplots representing alpha diversity in red Romaine lettuce. Top graphs (A-D) represent lettuce leaves while bottom graphs (E-H) represent lettuce root. The sequence of treatment across each row is Sanitized seed, *E. coli* treated (A, E); Sanitized seed, non-treated (B, F); Unsanitized seed, *E. coli* treated (C, G); and Unsanitized seed, non-treated (D, H) plant tissue. Alpha diversity was determined using the QIIME2 package on the 16S rRNA sequencing data. **Supplemental Figure S2.** Boxplots representing alpha diversity in mizuna mustard. (Top graphs (A-D) represent mizuna leaves while bottom graphs (E-H) represent mizuna root. The sequence of treatment across each row is Sanitized seed, *E. coli* treated (A, E); Sanitized seed, non-treated (B, F); Unsanitized seed, *E. coli* treated (C, G); and Unsanitized seed, non-treated (D, H) plant tissue. Alpha diversity was determined using the QIIME2 package on the 16S rRNA sequencing data. **Supplemental Figure S3.** Core microbiomes for red Romaine lettuce and mizuna mustard. (A) Venn diagram showing common genera between sanitized and unsanitized seed generated, leaf and root tissues of red Romaine lettuce. (B) Venn diagram showing common genera between sanitized and unsanitized seed generated, leaf and root tissues of mizuna mustard. Table shows genera represented by the "core microbiome" for red Romaine lettuce (27% from Venn diagram) and mizuna mustard (24.5% from Venn diagram).

Acknowledgements

The authors would like to thank Lawrence L. Koss for his expertise in setting up automated recording and monitoring of controlled environment parameters for plant growth studies.

Authors' contributions

ARD - Harvest, sample processing, and data analysis, manuscript development. CLMK - Design, harvest and sample processing, and data analysis, manuscript development. MEH - Design and sample processing, and data analysis, manuscript development. CJS - Harvest and sample processing, data collection. LES - Horticulture and crop management. JAF - sample processing, data collection. ABC - sample processing, data collection. JLG - sample processing, data collection. GJM - sample processing. Data collection. RMW - Project development and funding support. GDM - Project development and funding support. MWR - Project development and funding support. All authors read and approved the final manuscript.

Funding

Funding for this research and manuscript development was provided by NASA, Biological and Physical Sciences (BPS).

Availability of data and materials

The datasets generated and/or analyzed during the current study will be made available through the NASA GeneLab data repository, an open-access resource. The lettuce data can be accessed via the accession number GLDS-385 and <https://doi.org/10.26030/eseef-7r30> (<https://doi.org/10.26030/eseef-7r30>). The Mizuna data can be accessed via the accession number GLDS-386 and <https://doi.org/10.26030/thwa-cn80> (<https://doi.org/10.26030/thwa-cn80>).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹AMENTUM Services Inc., LASSO, Kennedy Space Center, Merritt Island, FL, USA. ²Southeastern Universities Research Association, LASSO, Kennedy Space Center, Merritt Island, FL, USA. ³NASA UB, Kennedy Space Center, Merritt Island, FL, USA.

Received: 5 May 2021 Accepted: 24 September 2021

Published online: 22 October 2021

References

- Nelson EB. The seed microbiome: origins, interactions, and impacts. *Plant Soil*. 2018;422(1–2):7–34.
- Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol*. 2012;13(6):614–29.
- Dewey-Mattia D, Manikonda K, Hall AJ, Wise ME, Crowe SJ. Surveillance for foodborne disease outbreaks - United States, 2009–2015. *MMWR Surveill Summ*. 2018;67(10):1998–2008.
- CDC. CDC Foodborne outbreaks. 2020 <http://www.cdc.gov/foodborneoutbreaks/>. Available from: <https://www.cdc.gov/foodsafety/outbreaks/multistate-outbreaks/index.html>
- Ibekwe AM, Grieve CM. Changes in developing plant microbial community structure as affected by contaminated water. *FEMS Microbiol Ecol*. 2004;48(2):239–48.
- Luna-Guevara JJ, Arenas-Hernandez MMP, Martínez De La Peña C, Silva JL, Luna-Guevara ML. The role of Pathogenic *E. coli* in fresh vegetables: behavior, contamination factors, and preventive measures. *Int J Microbiol*. 2019;2019.
- Tyler HL, Triplett EW. Plants as a habitat for beneficial and/or human pathogenic Bacteria. *Annu Rev Phytopathol*. 2008;46(1):53–73.
- Rastogi G, Tech JJ, Coaker GL, Leveau JHJ, Sbodio A, Suslow TV. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *ISME J*. 2012;6(10):1812–22.
- Williams TR, Marco ML. Transplantation on lettuce plants grown indoors. *MBio*. 2014;5(4):1–10.
- Williams TR, Moyne AL, Harris LJ, Marco ML. Season, irrigation, leaf age, and *Escherichia coli* inoculation influence the bacterial diversity in the lettuce Phyllosphere. *PLoS One*. 2013;8(7):1–14.
- Dees MW, Lysøe E, Nordskog B, Brurberg MB. Bacterial communities associated with surfaces of leafy greens: shift in composition and decrease in richness over time. *Appl Environ Microbiol*. 2015;81(4):1530–9.
- Merget B, Forbes KJ, Brennan F, Mcafeer S, Shepherd T, Strachan NJC, et al. Influence of plant species, tissue type, and temperature on the capacity of Shiga-Toxicogenic *Escherichia coli* To Colonize, Grow, and Be Internalized by Plants. *Appl Environ Microbiol*. 2019;85(11):1–16 Available from: <https://doi.org/10.1128/AEM.00123-19>.
- Santhanam R, Luu VT, Weinhold A, Goldberg J, Oh Y, Baldwin IT. Native root-associated bacteria rescue a plant from a sudden-wilt disease that emerged during continuous cropping. *Proc Natl Acad Sci U S A*. 2015;112(36):E5013–120.
- Bulgarelli D, Schlaeppi K, Spaepen S, van Themaat EVL, Schulze-Lefert P. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol*. 2013;64(1):807–38.
- Toju H, Peay KG, Yamamichi M, Narisawa K, Hiruma K, Naito K, et al. Core microbiomes for sustainable agroecosystems. *Nat Plants*. 2018;4(5):247–57. <https://doi.org/10.1038/s41477-018-0139-4>.
- Khodadad CLM, Hummerick ME, Spencer LSE, Dixit AR, Richards JT, Romeyn MW, et al. Microbiological and nutritional analysis of lettuce crops grown on the international Space Station. *Front Plant Sci*. 2020;11(March):1–15.
- Magurran AE. Measuring biological diversity. Hoboken: Wiley; 2003. p. 264. Available from: <http://www.wiley.com/WileyCDA/WileyTitle/productCd-0632056339.html>
- Begossi A. Use of ecological methods in ethnobotany: diversity indices. *Econ Bot*. 1996;50(3):280–9.
- Margalef R. Homage to Evelyn Hutchinson, or why there is an upper limit to diversity. *Trans Connecticut Acad Arts Sci*. 1972;44:211–35.
- Massa GD, Newsham G, Hummerick ME, Morrow RC, Wheeler RM. Plant pillow preparation for the veggie plant growth system on the international Space Station. *Gravitational Sp Res*. 2017;5(1):24–34 Available from: <http://gravitationalandspacebiology.org/index.php/journal/article/view/749/777>.
- Andrews WH, Wilson CR, Poelma PL, Romero A, Mislivec PB. Bacteriological survey of sixty health foods. *Appl Environ Microbiol*. 1979;37(3):559–66.
- Cankar K, Kraigher H, Ravnikar M, Rupnik M. Bacterial endophytes from seeds of Norway spruce (*Picea abies* L. Karst). *FEMS Microbiol Lett*. 2005;244(2):341–5.
- Links MG, Demeke T, Gräfenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ. Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on Triticum and Brassica seeds. *New Phytol*. 2014;202(2):542–53.
- Redford A, Bowers R, Knight R, Linhart Y, Fierer N. Variability in the distribution of Bacteria on tree leaves. *Environ Microbiol*. 2010;12(11):2885–93.
- van Overbeek LS, Saikkonen K. Impact of bacterial-fungal interactions on the colonization of the endosphere. *Trends Plant Sci*. 2016;21(3):230–42.
- Hardoim PR, Hardoim CCP, van Overbeek LS, van Elsas JD. Dynamics of seed-borne rice endophytes on early plant growth stages. *PLoS One*. 2012;7(2):e30438.
- Hunter PJ, Hand P, Pink D, Whipps JM, Bending GD. Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca species*) phyllosphere. *Appl Environ Microbiol*. 2010;76(24):8117–25.
- Matos A, Garland JL. Effects of community versus single strain inoculants on the biocontrol of Salmonella and microbial community dynamics in alfalfa sprouts. *J Food Prot*. 2005;68(1):40–8.
- Jones K. Nitrogen fixation in the phyllosphere of the Douglas fir, *Pseudotsuga douglasii*. *Ann Bot*. 1970;34(1):239–44. <https://doi.org/10.1093/oxfordjournals.aob.a084358>.
- Freiberg E. Microclimatic parameters influencing nitrogen fixation in the phyllosphere in a Costa Rican premontane rain forest. *Oecologia*. 1998;117(1–2):9–18.

31. Kishore GK, Pande S, Podile AR. Phylloplane bacteria increase seedling emergence, growth and yield of field-grown groundnut (*Arachis hypogaea* L.). *Lett Appl Microbiol*. 2005;40(4):260–8. <https://doi.org/10.1111/j.1472-765X.2005.11111>
32. Bodenhausen N, Horton MW, Bergelson J. Bacterial Communities Associated with the Leaves and the Roots of *Arabidopsis thaliana*. *PLoS One*. 2013;8(2):e56329.
33. Lebeis SL. The potential for give and take in plant-microbiome relationships. *Front Plant Sci*. 2014;5:1–6.
34. Weyens N, van der Lelie D, Taghavi S, Newman L, Vangronsveld J. Exploiting plant-microbe partnerships to improve biomass production and remediation. *Trends Biotechnol*. 2009;27(10):591–8 Available from: <http://www.sciencedirect.com/science/article/pii/S016777990900136X>.
35. Maignien L, DeForce EA, Chafee ME, Eren AM, Simmons SL. Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* Phyllosphere communities. *MBio*. 2014;5(1):e000682–13.
36. Lopez-Velasco G, Carder PA, Welbaum GE, Ponder MA. Diversity of the spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. *FEMS Microbiol Lett*. 2013;346(2):146–54.
37. Jackson CR, Randolph KC, Osborn SL, Tyler, Heather L. Culture dependent and independent analysis of bacterial communities associated with commercial salad leaf vegetables. *BMC Microbiol*. 2013;13(274).
38. Chaparro JM, Badri DV, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant development. *ISME J*. 2014;8(4):790–803.
39. Thapa S, Prasanna R. Prospecting the characteristics and significance of the phyllosphere microbiome. *Ann Microbiol*. 2018;68(5):229–45.
40. Herrmann M, Wegner CE, Taubert M, Geesink P, Lehmann K, Yan L, et al. Predominance of *Candida* Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. *Front Microbiol*. 2019;10:1–15.
41. Lemos LN, Medeiros JD, Dini-Andreote F, Fernandes GR, Varani AM, Oliveira G, et al. Genomic signatures and co-occurrence patterns of the ultra-small *Saccharimonadia* (phylum CPR/Patescibacteria) suggest a symbiotic lifestyle. *Mol Ecol*. 2019;28(18). <https://doi.org/10.1111/mec.15208>.
42. Agler MT, Ruhe J, Kroll S, Morhenn C, Kim ST, Weigel D, et al. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS Biol*. 2016;14(1):1–31.
43. Xue C, Penton CR, Shen Z, Zhang R, Huang Q, Li R, et al. Manipulating the banana rhizosphere microbiome for biological control of Panama disease. *Sci Rep*. 2015;5:1–11.
44. Köberl M, Dita M, Martinuz A, Staver C, Berg G. Members of Gammaproteobacteria as indicator species of healthy banana plants on Fusarium wilt-infested fields in Central America. *Sci Rep*. 2017;7:1–9.
45. Erlacher A, Cardinale M, Grosch R, Grube M, Berg G. The impact of the pathogen *Rhizoctonia solani* and its beneficial counterpart *Bacillus amyloliquefaciens* on the indigenous lettuce microbiome. *Front Microbiol*. 2014;5:1–8.
46. Fürnkranz M, Lukesch B, Müller H, Huss H, Grube M, Berg G. Microbial diversity inside pumpkins: microhabitat-specific communities display a high antagonistic potential against Phytopathogens. *Microb Ecol*. 2012;63(2):418–28.
47. Berg G, Mahner A, Moissl-Eichinger C. Beneficial effects of plant-associated microbes on indoor microbiomes and human health? *Front Microbiol*. 2014;5:1–5.
48. Dai Y, Yang F, Zhang L, Xu Z, Fan X, Tian Y, et al. Wheat-associated microbiota and their correlation with stripe rust reaction. *J Appl Microbiol*. 2020;128(2):544–55.
49. Innerebner G, Knief C, Vorholt JA. Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. *Appl Environ Microbiol*. 2011;77(10):3202–10.
50. Ofek M, Hadar Y, Minz D. Ecology of root colonizing Massilia (*Oxalobacteraceae*). *PLoS One*. 2012;7(7):e40117.
51. Denet E, Vasselon V, Burdin B, Nazaret S, Favre-Bonté S. Survival and growth of *Stenotrophomonas maltophilia* in free-living amoebae (FLA) and bacterial virulence properties. *PLoS One*. 2018;13(2):1–16.
52. Peralta KD, Araya T, Valenzuela S, Sossa K, Martínez M, Peña-Cortés H, et al. Production of phytohormones, siderophores and population fluctuation of two root-promoting rhizobacteria in *Eucalyptus globulus* cuttings. *World J Microbiol Biotechnol*. 2012;28(5):2003–14.
53. Clough SJ, Bent AE. Vapor-Phase Sterilization of *Arabidopsis* Seed. 2011 <http://www.plantpath.wisc.edu/fac/afb/vapster.html>. Available from: <http://www.plantpath.wisc.edu/fac/afb/vapster.html>
54. Deng K, Wang S, Rui X, Zhang W, Tortorello M, Lou. Functional analysis of *ycfR* and *ycfQ* in *Escherichia coli* O157:H7 linked to outbreaks of illness associated with fresh produce. *Appl Environ Microbiol*. 2011;77(12):3952–9.
55. Salazar JK, Deng K, Tortorello M, Lou, Brandl MT, Wang H, Zhang W. Genes *ycfR*, *sirA* and *yigG* Contribute to the Surface Attachment of *Salmonella enterica* Typhimurium and Saintpaul to Fresh Produce. *PLoS One*. 2013;8(2):e57272.
56. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl Environ Microbiol*. 2013;79(17):5112–20.
57. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37(8):852–7. <https://doi.org/10.7287/peerj.preprints.27295v2%7C>.
58. O'Callaghan M. Microbial inoculation of seed for improved crop performance: issues and opportunities. *Appl Microbiol Biotechnol*. 2016;100(13):5729–46.
59. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41(D1):590–6.
60. Beckers B, Op De Beeck M, Thijs S, Truyens S, Weyens N, Boerjan W, et al. Performance of 16S rDNA primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies. *Front Microbiol*. 2016;7:650.
61. Beirincx S, Viaene T, Haegeman A, Debode J, Amery F, Vandenabeele S, et al. Tapping into the maize root microbiome to identify bacteria that promote growth under chilling conditions. *Microbiome*. 2020;8(1):54.
62. Bisanz JE. qiime2R: Importing QIIME2 artifacts and associated data into R sessions. 2018. <https://github.com/jbisanz/qiime2R>. Available from: <https://github.com/jbisanz/qiime2R>
63. McMurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput Biol*. 2014;10(4):e1003531.
64. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2009;26(1):139–40.
65. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. 1948;27(4):623–56.
66. Bray JR, Curtis JT. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr*. 1957;27(4):325–49.
67. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc*. 1952;47(260):583–621.
68. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. 2001;26(1):32–46.
69. Andersen K, Kirkegaard R, Karst S, Albertsen M. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv*. 2018;299537.

Publisher's Note

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