Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 27:1-12

Comparative study of shared environments revealed an increased microbiome diversity of richness and abundance in open environments and evenness in closed environments

Angelina Ge, Tatiana Lau, Amos Fong, Felicia Liu-Fei

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Microbes are ubiquitous organisms that have been familiarly associated with human health and welfare, and human-bacteria interactions can shape the microbial makeup of the broader environment. While there has been extensive research conducted on the human microbiome, few studies have explored human microbiome dynamics as a function of open human social practices, and even fewer studies have explored how confinement parameters as extreme as space exploration simulation can impact the microbiome. Therefore, we aimed to compare two separate datasets modeling an open and confined environment, to investigate whether microbiome diversity differed between the two environments. Our study found that open environments have greater phylogenetic diversity and taxonomic richness at the genus level compared to confined environments, whereas confined environments have greater evenness indices compared to open environments. As a result, we also discovered that the top 20 differentially abundant genera between the datasets were all lower in abundance within our model for confined environments. These findings demonstrate that open and confined environments differ in microbiome diversity.

INTRODUCTION

Microbes have long been implicated in human health and well-being (1). Microbes are ubiquitous, having been found on animals, plants, building surfaces, and even the air (2). Moreover, bacterial colonization of humans can serve as a vector that shapes the microbial makeup of the spaces and other individuals they might encounter (3, 4). Thus, despite the ubiquity of microbes, the specific bacterial taxa that predominate a given space will exhibit differential biogeographical patterns. This is in part due to the fact that differences in bacterial communities are directly affected by contact with outdoor environmental factors (5).

Although the effects of open social interaction on microbial composition in various primates have been reported (6–9), few studies directly explore human microbiome dynamics as a function of migratory behaviors and human social practices (10). Even fewer studies have explored confinement parameters as stringent as in the case of space exploration simulation (11–13). As a result, direct comparisons between human microbial signatures of these confined (where individuals do not have any contact with the outside environment) and open (where individuals share regular contact with the outside environment) environments have not been explored. Our project thus aims to fill this gap by comparing the microbial communities of open and confined environments, modeled by a dorm dataset and a space stimulation (Hawai'i Space Exploration Analog and Stimulation; HI-SEAS) dataset, respectively.

Our open environment data is sourced from microbial communities on the surfaces of open shared dorm rooms and their inhabitants from a University of Chicago residence hall. Previous research on this data has investigated how individual microbial signatures impacted the microbiota of common spaces in a college dormitory (14). In this study, common hand-associated surfaces were found to be a central point of association, while individual shoe samples were more connected to each other than floors, indicating that hands were a primary means of transmission in comparison to shoes between common spaces (14). Our confined environment data is sourced from the HI-SEAS study(15), wherein the microbiota of the environmental surfaces and individuals in a shared dome were sampled. In this study, six

Published Online: September 2022

Citation: Angelina Ge, Tatiana Lau, Amos Fong, Felicia Liu-Fei. 2022. Comparative study of shared environments revealed an increased microbiome diversity of richness and abundance in open environments and evenness in closed environments. UJEMI 27:1-12

Editor: Andy An and Gara Dexter, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

individuals spent one year in isolation inside an 11-meter spherical dome in Hawaii to mimic the effects of space travel (15). The original dataset study utilized the HI-SEAS dataset to delineate the longitudinal dynamics of the human microbiomes under confined environments (15). This study observed significant differences in microbial diversity, abundance, and composition between abiotic and biotic surfaces sampled (15). Furthermore, this study provided evidence supporting the hypothesis that abiotic and biotic surfaces undergo longitudinal microbial homogenization over time (15).

By utilizing these two datasets as models for open and confined environments, we can novelly investigate the differences between the bacterial communities present, to ultimately better understand the microbial conditions that those with no outside contact endure. As such, the aims of this study were to: 1. determine whether the microbiome of abiotic and biotic surfaces within open environments are more or less diverse than abiotic and biotic surfaces within confined environments; 2. investigate the difference between taxonomic richness in the microbiomes of open and confined environments; and 3. investigate the difference in organism abundance between the microbiomes of open and confined environments in abiotic and biotic surfaces. To explore these experimental aims, we used alpha diversity analyses to compare microbiome diversity between open and confined environments in both abiotic and biotic surfaces, taxonomic analysis to determine unique and overlapping taxonomic groups between open and confined environments, and differential abundance testing to identify dominant bacteria in confined and open environments within abiotic and biotic surfaces. Quantitative Insights Into Microbial Ecology 2, (QIIME2) (25), a bioinformatics platform for the processing and analysis of amplicon library data, was used to conduct our alpha and beta diversity analyses and taxonomic analysis while R (22), a statistical computing and graphics programming language, was used for our differential abundance testing.

We hypothesized that both abiotic and biotic surfaces will be significantly different between confined and open environments as previous research has demonstrated that individuals sharing closed social networks harbor more similar profiles of microbiome diversity at strain-level resolution (16). We further hypothesized that both abiotic and biotic surfaces within confined environments will harbor less diverse microbiomes when compared to open environments, as it is well established that increasing social interaction contributes significantly to microbiome diversity in isogenic mouse and chimpanzee models (17, 18). After performing our analyses, we found that open environments exhibited greater phylogenetic diversity and taxonomic richness, while confined environments had greater evenness. We also found that our open environment had a higher abundance of the top 20 differentially abundant genera.

METHODS AND MATERIALS

Datasets. The HI-SEAS (model for confined environment) dataset was generated by Mahnert *et al.*, wherein 6 participants were enclosed in an 11-meter diameter dome for 1 year in Hawaii as a part of a study conducted by NASA and the University of Hawaii. Various surfaces in the dome including the toilet, kitchen floor, bedroom desk, and main desk, in addition to the torso of each crew member, were swabbed biweekly for the duration of the mission (15). The dorms (model for open environment) dataset was generated by Richardson *et al.*, who sampled 37 participants, in addition to their dormitory rooms (shared or single) and common areas, at four time points over the course of 3 months (14). Individual surfaces including hands, bed sheets, and shoes were swabbed from each participant, in addition to common surfaces including tables and bathrooms (14). For the HI-SEAS dataset, DNA extraction was performed using QIAGEN's DNeasy PowerSoil Kit(19). For the dorm's dataset, DNA extraction was performed using MO BIO's PowerSoil DNA Isolation Kit(20). Microbial sequences from both datasets were obtained using primer pairs F515-R806 specific for the V4 region of the 16S rRNA gene, hybridized to tags for Illumina Miseq sequencing(21).

Metadata manipulation. All data manipulation and analysis performed using R(22) are detailed in the supplemental R script (RScript). The dorms and HI-SEAS metadata were joined in R (version 4.1.2)(22) and RStudio(23) along with R packages: tidyverse (version 1.3.1) (24). A new metadata category "dataset" (with binary fields "dorms" or "HI-SEAS") was created to define the original dataset of a given sample after joining.

Data processing using the QIIME2 pipeline. All data analyses performed using QIIME2 (25)are detailed in the supplemental QIIME2 script (QIIME2Script). The modified metadata table and both the demultiplexed dorms and HI-SEAS data files were imported into QIIME2 (version 2022.2) (25). A read length of 150 base pairs was retained, given that all truncation lengths less than or equal to 150 have 25th percentile Phred quality scores over 30. A read length of 151 (maximum retained bases) has a 25th percentile Phred quality score of 16 indicating a significantly wider margin of sequencing error. Sequence quality control was performed by the QIIME2 compatible plugin Divisive Amplicon Denoising Algorithm 2 (q2-DADA2) (26), which corrects for potential sequencing errors and classifies unique amplicon sequence variants (ASVs) by clustering. After DADA2 denoising, dorms and HI-SEAS features tables and representative sequences files were merged in QIIME2.

Taxonomic classification. A pre-trained Naive Bayes classifier (27) (trained using scikit-learn 0.24.1 on the 99% Silva 138 reference database (28, 29) and 515F/908R primers) was used to assign the representative sequences taxonomic identities and confidence level in QIIME2. The taxonomic composition (classified representative sequences) was then visualized by the taxa barplot function in QIIME2.

Features table filtering. Eukaryotic ASVs (mitochondria and chloroplast DNA) were filtered out from the features table in QIIME2. The resulting features table was further filtered based on sample metadata for abiotic (defined as sample_type = 'surface') and biotic (defined as sample_type = 'skin') surfaces generating two unique features tables (abiotic and biotic) for downstream diversity and abundance analysis. We chose to retain a sampling depth of 7000 bp, which maximizes the number of samples and ASVs/features retained. The QIIME2-generated Alpha Rarefaction curves supported this sampling depth.

Alpha and beta diversity analysis. To facilitate downstream phylogenetic diversity analysis, a rooted phylogenetic tree was generated to relate features based on phylogenetic distance. The tree was made by the QIIME2 pipeline using Multiple Alignment using Fast Fourier Transform (MAFFT)(30, 31) to perform multiple sequence alignment of features, and the FastTree q2-phylogeny plugin (32)to generate a corresponding unrooted phylogenetic tree. Using the q2-diversity plugin, alpha diversity metrics (Shannon's diversity index (33), Observed Features, Faith's Phylogenetic Diversity (34), Pielou's evenness (35)) and beta diversity metrics (Jaccard distance (36), Bray-Curtis distance (37), unweighted UniFrac distance (38), weighted UniFrac distance (39)), were computed at the aforementioned sampling depth of 7000. Statistical significance for alpha diversity metrics was computed using Kruskal-Wallis one-way analysis of variance test (40). Statistical significance for beta diversity metrics was computed using the permutational multivariate analysis of variance (PERMANOVA) test (41). Visualizations for alpha diversity analysis boxplots and beta diversity principal coordinates analysis (PCoA) plots using Emperor (42, 43) were generated as an output of the QIIME2 q2-diversity plugin. QIIME2 generated outputs (features table, rooted tree, taxonomy, sample metadata) were imported into R. To facilitate downstream abundance analysis, the outputs were integrated into a phyloseq object using package Qiime2r (44) (converted into DESeq2 (Differential Expression Analysis For Sequence Count Data 2) object using package DESeq2) (45) and a microeco object using package File2meco (46).

Taxonomic abundance analysis. To identify unique taxonomic groups at the genus level, the level-6.csv taxonomic composition spreadsheet was extracted from the taxa bar plot export and converted to .xlsx format. From this spreadsheet, the number of unique taxonomic groups for the open dataset and confined dataset, as well as shared taxonomic groups between the datasets were calculated on Microsoft Excel (47). A venn diagram was made on Adobe Illustrator (48) to visualize the number of unique and shared taxonomic groups between datasets.

Differential abundance analysis. Differential abundance analysis was performed in R using the packages DESeq2 (45), Phyloseq (49), Tidyverse, Vegan (50) and Ape (51) detailed in the supplemental R Script (RScript). The confined dataset was defined as the reference group.

Significance was defined as differentially abundant genera with an adjusted FDR-corrected Wald Chi-Squared Test (52) p-value of <0.05. Top 20 hits for significant differentially abundant genera were plotted using ggplot2 (53).

Relative abundance analysis. Relative abundance analysis was performed in R using packages randomForest (54), Tidyverse (24), Microeco (46), Cowplot (55) and File2meco (46) detailed in the supplemental R Script (RScript). The microeco trans_diff function utilized the supervised machine learning algorithm randomForest analysis (54) to identify differentially abundant taxa. The sample group for comparison was the metadata category "dataset, and the significance cutoff was defined as average abundance <0.01%.

Figure formatting. All figures generated by QIIME2, Emperor, DESeq2-phyloseq pipeline and randomForest-microeco pipeline were reformatted and annotated using Adobe Illustrator.

Data availability. The HI-SEAS confined dataset is publicly available at QIITA under study ID 12858, and the European Nucleotide Archive (ENA) under accession number ERP118380. The dorms dataset is publicly available at QIITA under study ID 12470 and the European Bioinformatics Institute (EBI) under project number PRJEB33050/ERP115809.

RESULTS

Confined environments exhibit greater evenness while open environments exhibit greater phylogenetic diversity. To assess the microbiome diversity of abiotic and biotic surfaces in both open and confined environments, we performed alpha diversity analyses on the open and confined datasets using Pielou's evenness and Faith's phylogenetic diversity metrics in the QIIME2 pipeline. This analysis showed that Pielou's evenness was significantly higher in the confined dataset compared to the open dataset, while Faith's phylogenetic diversity was significantly higher in the open dataset compared to the confined dataset **(Figure 1)**. These results are consistent across both abiotic **(Figure 1A)** and biotic **(Figure 1B)** surfaces.

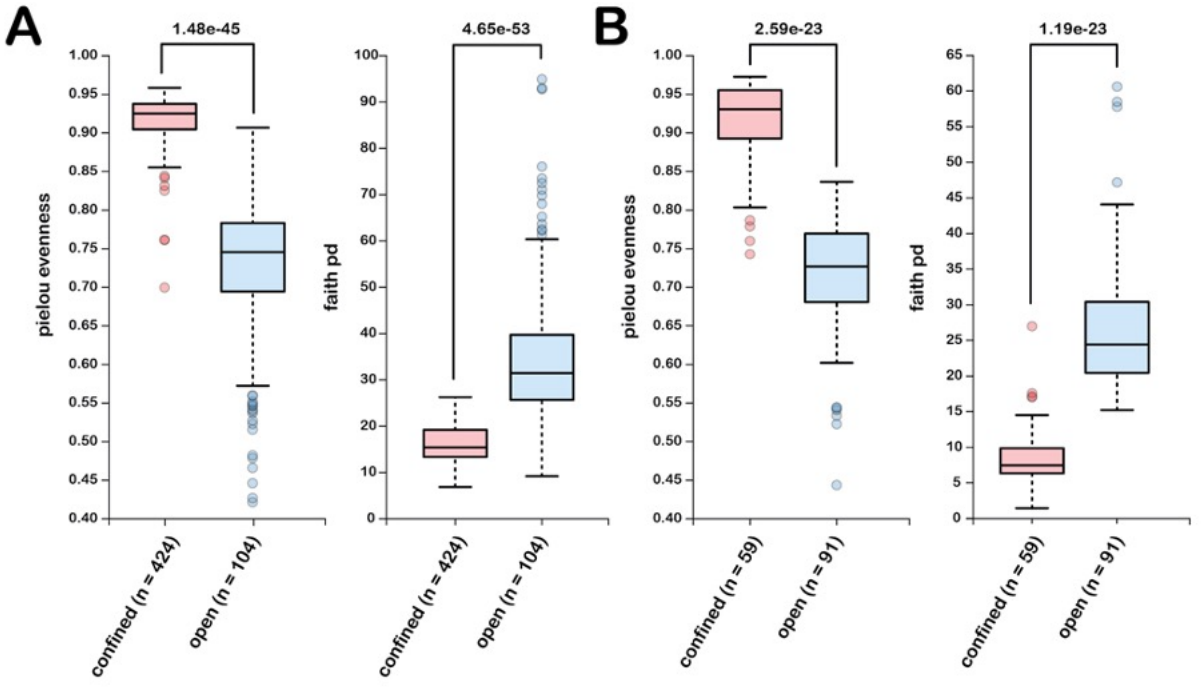

FIG. 1 Confined environments exhibit greater evenness while open environments exhibit greater phylogenetic diversity. Alpha diversity analyses of **(A)** abiotic or **(B)** biotic surfaces from the confined (red) and open (blue) datasets, box plots and error bars represent mean ± SEM. Statistical analysis was performed via Kruskal-Wallis one-way analysis of variance test. Q-values are denoted on the respective plots.

Additionally, to investigate the diversity between the two environments, we performed beta diversity analyses using the unweighted UniFrac metric in QIIME2. Here, we observed a similar trend to the alpha diversity analysis, wherein the two groups differed significantly in both abiotic **(Supplemental Figure 1A, 1C)** and biotic **(Supplemental Figure 1B, 1D)** surfaces.

Open environments exhibit greater taxonomic richness compared to confined environments. As we observed significant differences in microbiome diversity between the two datasets, we decided to investigate the taxonomic makeup of the confined and open datasets. Here, we found that across all surfaces, the confined dataset harboured only 146 unique genera, while the open dataset harboured 1160 unique genera **(Figure 2)**. Additionally, the two environments shared 1050 genera.

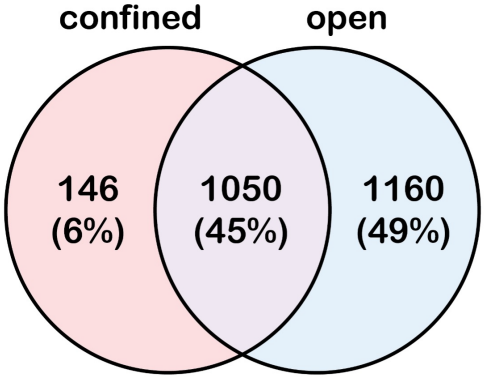

FIG. 2 Open environments exhibit greater taxonomic richness compared to confined environments. Taxonomic analysis of confined (red) and open (blue) datasets for both abiotic and biotic surfaces. The confined dataset had 146 unique microbial genera, while the open dataset had 1160 unique genera. Additionally, the two datasets shared 1050 genera.

The top 20 bacterial genera are more differentially abundant in open compared to confined environments. Upon exploring the taxonomic makeup of the confined and open datasets, we decided to take a closer look at the organism abundance and individual microorganisms using a differential abundance analysis with the DESeq2/Phyloseq pipeline. Here, we observed that the top 20 differentially abundant genera were more abundant in the open dataset compared to the confined dataset **(Figure 3)**. Upon analyzing abiotic surfaces, *Nesterenkonia* sp. was over 20 times more abundant in the open compared to confined environments **(Figure 3A)**. Similarly, on biotic surfaces, *Salinococcus* sp. and *Actinobacillus* sp. were more abundant in the open compared to confined dataset **(Figure 3B)**. In total, there were 426 and 221 differentially abundant genera between open and confined environments for abiotic and biotic surfaces, respectively.

Additionally, we looked at the relative abundance of specific genera using the RandomForest/Microeco (49, 42) pipeline **(Supplemental Figure 2)**. Here, we found that both the confined and open dataset displayed genera more relatively abundant in one dataset in comparison to the other, though dominantly abundant genera were present. The confined dataset was more abundant in *Pseudomonas* sp. for abiotic surfaces **(Supplemental Figure 2A)** and *Cutibacterium* sp. for biotic surfaces **(Supplemental Figure 2B)** relative to the open dataset. The open dataset was more abundant in *Streptococcus* sp. across both abiotic and biotic surfaces relative to the confined dataset. In total, there were 821 relatively abundant genera on abiotic surfaces, and 321 relatively abundant genera on biotic surfaces.

DISCUSSION

Few studies have compared different environments modeled by different datasets. In this study, we explored the human microbiome communities of an open and confined shared environment to determine whether or not microbiome diversity differed between the two. The need to address this knowledge gap is especially important to consider due to the increase in confined living conditions as a result of global lockdowns and quarantine guidelines during the COVID-19 pandemic. Using alpha and beta diversity analyses, taxonomic analysis, and

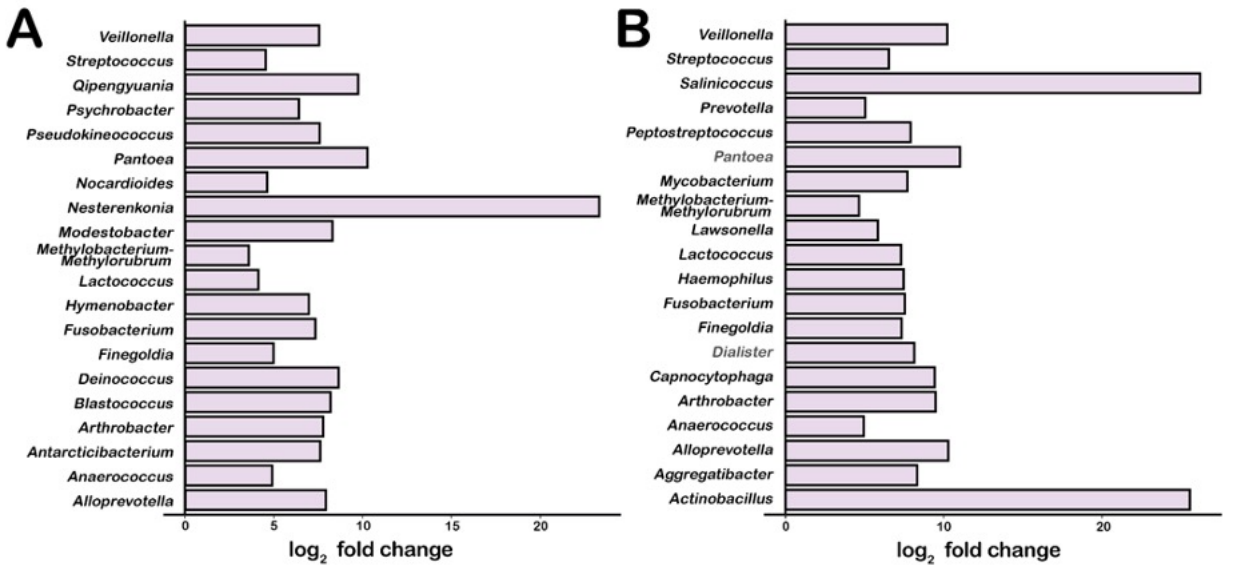

FIG. 3 *Nesterenkonia* sp., *Salinococcus* sp., and *Actinobacillus* sp. are differentially abundant in open environments compared to closed environments. Differentially abundant genera in the open dataset compared to the confined dataset was acquired through differential abundance analysis using the DESeq2/Phyloseq pipeline. The confined dataset was defined as the reference group. Top 20 genera were determined according to statistical significance via FDR-corrected Wald Chi-Squared Test. Both **(A)** abiotic and **(B)** biotic surfaces were analyzed with a >20 log2 fold change for *Nesterenkonia* sp. (abiotic), *Salinococcus* sp., and *Actinobacillus* sp. (biotic).

differential abundance testing, we found significant differences in microbiome diversity between open and confined shared environments.

Different metrics underlying diversity demonstrate that microbiome diversity differs between open and confined environments but is complex. We investigated whether the microbiome of abiotic and biotic surfaces within open environments are more or less diverse than abiotic and biotic surfaces within confined environments. Using alpha diversity analyses, we found that the confined environment exhibited significantly greater community evenness in comparison to the open environment for both abiotic and biotic surfaces, indicating that the confined environment had higher diversity in regards to this metric (**Figure 1**). This suggests that the lack of exogenous factors in a highly confined environment results in increased microbial uniformity. It is well known that the microbial community is highly dynamic, and it is possible that under confined conditions where there is a lack of external microbial input, the existing microbes adapt over time in the environment and interact with each other to form a homogenized community structure. It is also possible that the presence of exogenous factors favor certain microbial species over others, leading to decreased evenness in open environments.

On the other hand, our alpha diversity analyses showed that the open environment exhibited significantly greater phylogenetic diversity compared to the confined environment for both abiotic and biotic surfaces, demonstrating higher community richness and thus increased microbiome diversity when considering this metric (**Figure 1**). This was also supported by the taxonomic analysis, which showed higher taxonomic richness in the open environment when compared to the confined environment for both surface types (**Figure 2**). These data together suggest that exogenous factors can shape the microbiome of both abiotic and biotic surfaces, leading to increases in the absolute number of organisms in a given environment as well as the relatedness distance between organisms. It is likely that external environment and social interactions provide the opportunity for the introduction of new and more diverse microbes into a given space. Increased richness can favor community-level stability by ensuring the presence of various species that can tolerate different environmental fluctuations (56). That being said, pathogenic strains may also be introduced, and new interspecies interactions within the microbiome can also contribute to the complexity of the community. Our findings of increased diversity in open environments are consistent with the

existing literature, which demonstrates that individuals with larger social networks harbour more diverse microbiomes (10), and that microbial diversity declines over time in highly confined living situations (11). Animal studies have shown that social behavior can induce changes in the microbiome by influencing transmission and behavior patterns, but further investigation is needed to confirm these findings in the case of humans (57). Moreover, future studies are encouraged to explore whether it is the lack or presence of exogenous factors that drives factors measuring microbiome diversity.

Our results support our hypothesis which states that microbiome diversity within both abiotic and biotic surfaces differs significantly between confined and open environments. However, given the conflicting results from our diversity metrics, we were unable to deduce if microbiome diversity was comprehensively higher in either environment as measurements of diversity are more complex and dependent on the metric used. Therefore, our results partially supported our secondary hypothesis proposing reduced diversity in confined environments. This secondary hypothesis was supported by the phylogenetic diversity and taxonomic richness metrics, but refuted by the evenness metric. Hence, our study highlights the importance of considering various factors indicative of diversity, as different metrics may measure diversity contrastingly.

Dominant genera with possible human health implications were present amongst the differentially and relatively abundant genera. Next, we sought to investigate the difference in organism abundance between the microbiome of open and confined environments on abiotic and biotic surfaces. Changes in microbiome diversity and composition can affect the services provided to the host by the microbiome and impact host fitness (57). Thus, we aimed to identify dominant genera with possible implications for human health. Using a differential abundance analysis, we found a higher abundance of the top 20 differentially abundant genera in the open environment compared to the confined environment, suggesting that exogenous factors can lead to increases in microbial abundance (**Figure 3**). Specifically, the genus *Nesterenkonia* was found the most dominantly differentially abundant on abiotic surfaces (**Figure 3A**), while the genera *Salinococcus* and *Actinobacillus* were found to be the most dominantly differentially abundant on biotic surfaces (**Figure 3B**). Bacterial members of the genus *Nesterenkonia* are predominantly found in hypersaline habitats, and have been mainly reported as weak human pathogens that can cause asymptomatic bacteraemia (58, 59). That being said, certain species such as *Nesterenkonia jeotgali* and *N. massiliensis* have been proposed to be potentially clinically pathogenic (58, 59). *Salinococcus* sp. are also moderately halophilic but are poorly characterized, though no species have yet been reported as pathogenic (60). Lastly, the majority of *Actinobacillus* sp. are found on the alimentary, respiratory, and genital mucous membrane of various animals and humans, and act as commensals (61, 62). However, certain species have been discovered as pathogenic, and studies have also reported findings of commensal species in lesions (62). For instance, *A. actinomycetemcomitans* has been found to play a role in periodontal disease and endocarditis in humans (62). The variation in pathogenicity between the various species within a genus highlights the importance for further studies to explore these dominant genera at the species level, in order to elucidate their role in human health and well-being.

Using a relative abundance analysis, we were able to identify additional dominant genera more abundant in open or confined environments (**Supplemental Figure 2**). The genus *Streptococcus* was found more abundant in open environments relative to confined environments (**Supplemental Figure 2**), which was consistent with our findings from the top 20 differentially abundant genera (**Figure 3**). Many species of the *Streptococcus* genus are highly pathogenic, with as many as 35 species identified as causes of invasive infections in humans (63). The highest number of pneumonia cases worldwide have been caused by *S. pneumoniae*, and members of the human and animal flora have also been discovered to have pathogenic potential (63). Studies have also shown that streptococcal diseases are more common in developing countries, which may be associated with an increase in open environments in developing countries compared to developed countries (64). Interestingly, relative abundance analysis also displayed dominant genera that were more abundant in the confined environment relative to the open environment, with *Pseudomonas* sp. more abundant on abiotic surfaces and *Cutibacterium* sp. more abundant on biotic surfaces

(**Supplemental figure 2**). Most of the *Pseudomonas* sp. with implications on human health have been found to be associated with opportunistic infections (65), whereas *Cutibacterium* sp. are typically commensals of the skin, though some have also been discovered to have pathogenic potential and can be considered opportunistic (66). Given these findings, microbiome dynamics need to be monitored to prevent the spread of potential pathogens and to maintain a healthy, diverse microbiome.

Furthermore, from our differential and relative abundance analyses, we noted an increase in diversity on abiotic surfaces compared to biotic surfaces as we found 426 differentially abundant genera on abiotic surfaces and 221 on biotic surfaces. The relative abundance analysis produced the same trends, displaying 821 relatively abundant genera on abiotic surfaces and 321 on biotic surfaces. We propose that this can be attributed to the use of antibiotics by humans and an increase in antimicrobial compound production on biotic compared to abiotic surfaces.

Limitations A main limitation of this study was the use of two different datasets to conduct our analysis. As a result, the data was obtained from two different laboratories, and thus had inconsistent study procedures. This led to variability in factors including sample collection methods, sample collection time, and sample size. In regards to sample collection, biotic samples from the open dataset were obtained by swabbing a sterile cotton BD-Swube applicator (14), while biotic samples from the confined dataset were obtained via sterilized wipes and deionized water (15). Furthermore, the collection time of the sample varied between the groups. In the dorms study, samples were collected at 4 points during a 3-month period (14), while in the confined study, samples were collected biweekly for 11 months (15). Finally, the sample sizes between the open data (abiotic: 258, biotic: 114) (14), and confined data (abiotic: 104, biotic: 59) (15) used for this study differed slightly. However, QIIME2 normalizes sample size upon initial data processing, thus mitigating the issue of differential sample size for our study and likely did not greatly impact our results.

Another limitation of this project was the lack of species-specific data when evaluating taxonomic trends between confined and open environments. When looking at distinct species in the taxonomic and differential abundance analysis, many samples were lost as species-level taxonomic information was not available. As such, we decided to analyze our data at the genus level to retain a high level of organisms. Thus, distinct species were not able to be identified, diminishing the specificity of our results.

Conclusions Our study aimed to compare the microbiome diversity of abiotic and biotic surfaces in an open and confined shared environment. We found that microbiome diversity differed between open and confined environments. While the open environment exhibited significantly greater phylogenetic diversity and taxonomic richness, the confined environment had significantly greater evenness. We speculate that these results suggest that the presence or lack of exogenous factors impacts microbiome diversity, though further investigation is needed to confirm this and identify the specific factors. Furthermore, we observed a higher abundance of the top 20 differentially abundant genera in the open environment compared to the confined environment, and identified *Nesterenkonia* sp., *Salinococcus* sp., and *Actinobacillus* sp. as the most dominant genera. While potential pathogens were identified, further studies need to be conducted to explore the role of these dominant genera in open environments and their implications on human health.

Future Directions To address certain previous limitations mentioned, this experiment could be replicated with the same protocol for both confined and open living conditions. By utilizing a streamlined protocol, additional variations in sample collection methods, collection time, and sample size can be eliminated and the results obtained can harbour more significance. Furthermore, as discussed in the limitations section, future studies could investigate the individual species that arise from the dominant genera we have identified. Our current study was constrained by a lack of species-specific data, and as aforementioned, some species of the dominant genera are pathogenic whereas others are not. Thus, it would be pertinent to identify which species are more dominant in confined and open environments to determine clinical relevance.

Another future direction could be to evaluate which factors led to the observed differences in microbiome diversity between the two environments. While one environment was open and the other was confined, there are still various exogenous factors including outdoor exposure, movement patterns, and the number of social contacts, which may predominantly contribute to these differences in microbial diversity. Exposure to nature (67), increased social interaction (68), and higher levels of movement (69) are all characteristics of an open environment in comparison to a closed environment; these components have similarly been implicated in shaping the human microbiome, and thus are important to consider when investigating the differences between these environments. Altogether, the determination of the key factor(s) in differences between the microbial makeup of an open and confined environment can provide further rationale for our observed results. By studying these factors, we may be able to elucidate whether or not it is the presence or the lack of these factors that impact microbiome diversity, as well as the directionality of the connections between exogenous factors and the microbiome. For instance, in this study, we show that exogenous factors can affect the diversity of the microbiome, but other studies have also shown that the microbiome can influence host social behavior through the production of chemical signals and changes to the nervous system (57).

Lastly, an additional future direction is to explore other facets of the human microbiome in response to confined versus open environments. A currently understudied constituent of the microbiome is the virome, which mainly includes viruses that infect bacteria and plants, although also contains eukaryotic viruses (70). The composition of the human virome is similarly shaped by external factors including cohabitation and living environments (71), and thus may exhibit comparable trends to the bacterial constituent of the microbiome. Consequently, the virome can have major effects on both bacteria and humans and can be important to study in parallel with the bacterial microbiome.

ACKNOWLEDGEMENTS

We would like to thank Dr. Evelyn Sun, Divya Kriti, Zakhar Krekhno, and the rest of the MICB 447 teaching team for their guidance and support throughout this project. We would also like to thank the UBC Department of Microbiology and Immunology for providing the facilities, funding and resources for this project. Lastly, we would like to acknowledge Mahnert *et al.* and Richardson *et al.* for providing the study metadata.

CONTRIBUTIONS

Amos Fong (AF), Felicia Liu-Fei (FLF), Angelina Ge (AG), and Tatiana Lau (TL) wrote all scripts and performed data analysis in QIIME2, and AF performed data analysis on R. AF, FLF, AG, and TL contributed equally to the writing of the abstract, introduction, methods, results, and discussion of the manuscript, in addition to editing all sections.

REFERENCES

1. **Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, Creasy HH, Earl AM, FitzGerald MG, Fulton RS, Giglio MG, Hallsworth-Pepin K, Lobos EA, Madupu R, Magrini V, Martin JC, Mitreva M, Muzny DM, Sodergren EJ, Versalovic J, Wollam AM, Worley KC, Wortman JR, Young SK, Zeng Q, Aagaard KM, Abolude OO, Allen-Vercoe E, Alm EJ, Alvarado L, Andersen GL, Anderson S, Appelbaum E, Arachchi HM, Armitage G, Arze CA, Ayyaz T, Baker CC, Begg L, Belachew T, Bhonagiri V, Bihan M, Blaser MJ, Bloom T, Bonazzi V, Paul Brooks J, Buck GA, Buhay CJ, Busam DA, Campbell JL, Canon SR, Cantarel BL, Chain PSG, Chen I-MA, Chen L, Chhibba S, Chu K, Ciulla DM, Clemente JC, Clifton SW, Conlan S, Crabtree J, Cutting MA, Davidovics NJ, Davis CC, DeSantis TZ, Deal C, Delehaunty KD, Dewhirst FE, Deych E, Ding Y, Dooling DJ, Dugan SP, Michael Dunne W, Scott Durkin A, Edgar RC, Erlich RL, Farmer CN, Farrell RM, Faust K, Feldgarden M, Felix VM, Fisher S, Fodor AA, Forney LJ, Foster L, Di Francesco V, Friedman J, Friedrich DC, Fronick CC, Fulton LL, Gao H, Garcia N, Giannoukos G, Giblin C, Giovanni MY, Goldberg JM, Goll J, Gonzalez A, Griggs A, Gujja S, Kinder Haake S, Haas BJ, Hamilton HA, Harris EL, Hepburn TA, Herter B, Hoffmann DE, Holder ME, Howarth C, Huang KH, Huse SM, Izard J, Jansson JK, Jiang H, Jordan C, Joshi V, Katancik JA, Keitel WA, Kelley ST, Kells C, King NB, Knights D, Kong HH, Koren O, Koren S, Kota KC, Kovar CL, Kyrpides NC, La Rosa PS, Lee SL, Lemon KP, Lennon N, Lewis CM, Lewis L, Ley RE, Li K, Liolios K, Liu B,**

Liu Y, Lo C-C, Lozupone CA, Dwayne Lunsford R, Madden T, Mahurkar AA, Mannon PJ, Mardis ER, Markowitz VM, Mavromatis K, McCorrison JM, McDonald D, McEwen J, McGuire AL, McInnes P, Mehta T, Mihindukulasuriya KA, Miller JR, Minx PJ, Newsham I, Nusbaum C, O'Laughlin M, Orvis J, Pagani I, Palaniappan K, Patel SM, Pearson M, Peterson J, Podar M, Pohl C, Pollard KS, Pop M, Priest ME, Proctor LM, Qin X, Raes J, Ravel J, Reid JG, Rho M, Rhodes R, Riehle KP, Rivera MC, Rodriguez-Mueller B, Rogers Y-H, Ross MC, Russ C, Sanka RK, Sankar P, Fah Sathirapongsasuti J, Schloss JA, Schloss PD, Schmidt TM, Scholz M, Schriml L, Schubert AM, Segata N, Segre JA, Shannon WD, Sharp RR, Sharpton TJ, Shenoy N, Sheth NU, Simone GA, Singh I, Smillie CS, Sobel JD, Sommer DD, Spicer P, Sutton GG, Sykes SM, Tabbaa DG, Thiagarajan M, Tomlinson CM, Torralba M, Treangen TJ, Truty RM, Vishnivetskaya TA, Walker J, Wang L, Wang Z, Ward DV, Warren W, Watson MA, Wellington C, Wetterstrand KA, White JR, Wilczek-Boney K, Wu Y, Wylie KM, Wylie T, Yandava C, Ye L, Ye Y, Yooseph S, Youmans BP, Zhang L, Zhou Y, Zhu Y, Zoloth L, Zucker JD, Birren BW, Gibbs RA, Highlander SK, Methé BA, Nelson KE, Petrosino JF, Weinstock GM, Wilson RK, White O, The Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. 7402. Nature 486:207–214.

2. **Miletto M, Lindow SE**. 2015. Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. Microbiome 3:61.
3. **Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, Gibbons SM, Larsen P, Shogan BD, Weiss S, Metcalf JL, Ursell LK, Vázquez-Baeza Y, Van Treuren W, Hasan NA, Gibson MK, Colwell R, Dantas G, Knight R, Gilbert JA.** 2014. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science 345:1048–1052.
4. **Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, Caporaso JG, Knights D, Clemente JC, Nakielny S, Gordon JI, Fierer N, Knight R**. 2013. Cohabiting family members share microbiota with one another and with their dogs. eLife 2:e00458.
5. **Dunn RR, Fierer N, Henley JB, Leff JW, Menninger HL.** 2013. Home Life: Factors Structuring the Bacterial Diversity Found within and between Homes. PLOS ONE 8:e64133.
6. **Sarkar A, Harty S, Johnson KV-A, Moeller AH, Archie EA, Schell LD, Carmody RN, Clutton-Brock TH, Dunbar RIM, Burnet PWJ.** 2020. Microbial transmission in animal social networks and the social microbiome. Nat Ecol Evol 4:1020–1035.
7. **Raulo A, Ruokolainen L, Lane A, Amato K, Knight R, Leigh S, Stumpf R, White B, Nelson KE, Baden AL, Tecot SR.** 2018. Social behaviour and gut microbiota in red-bellied lemurs (Eulemur rubriventer): In search of the role of immunity in the evolution of sociality. J Anim Ecol 87:388–399.
8. **Gogarten JF, Davies TJ, Benjamino J, Gogarten JP, Graf J, Mielke A, Mundry R, Nelson MC, Wittig RM, Leendertz FH, Calvignac-Spencer S**. 2018. Factors influencing bacterial microbiome composition in a wild non-human primate community in Taï National Park, Côte d'Ivoire. ISME J 12:2559–2574.
9. **Tung J, Barreiro LB, Burns MB, Grenier J-C, Lynch J, Grieneisen LE, Altmann J, Alberts SC, Blekhman R, Archie EA**. Social networks predict gut microbiome composition in wild baboons. eLife 4:e05224.
10. **Johnson KV-A.** 2020. Gut microbiome composition and diversity are related to human personality traits. Hum Microbiome J 15:100069.
11. **Schwendner P, Mahnert A, Koskinen K, Moissl-Eichinger C, Barczyk S, Wirth R, Berg G, Rettberg P.** 2017. Preparing for the crewed Mars journey: microbiota dynamics in the confined Mars500 habitat during simulated Mars flight and landing. Microbiome 5:129.
12. **Van Houdt R, De Boever P, Coninx I, Le Calvez C, Dicasillati R, Mahillon J, Mergeay M, Leys N.** 2009. Evaluation of the Airborne Bacterial Population in the Periodically Confined Antarctic Base Concordia. Microb Ecol 57:640–648.
13. **Mayer T, Blachowicz A, Probst AJ, Vaishampayan P, Checinska A, Swarmer T, de Leon P, Venkateswaran K.** 2016. Microbial succession in an inflated lunar/Mars analog habitat during a 30-day human occupation. Microbiome 4:22.
14. **Richardson M, Gottel N, Gilbert JA, Lax S**. 2019. Microbial Similarity between Students in a Common Dormitory Environment Reveals the Forensic Potential of Individual Microbial Signatures. mBio 10:e01054-19.
15. **Mahnert A, Verseux C, Schwendner P, Koskinen K, Kumpitsch C, Blohs M, Wink L, Brunner D, Goessler T, Billi D, Moissl-Eichinger C.** 2021. Microbiome dynamics during the HI-SEAS IV mission, and implications for future crewed missions beyond Earth. Microbiome 9:27.
16. **Brito IL, Gurry T, Zhao S, Huang K, Young SK, Shea TP, Naisilisili W, Jenkins AP, Jupiter SD, Gevers D, Alm EJ.** 2019. Transmission of human-associated microbiota along family and social networks. Nat Microbiol 4:964–971.
17. **Rosshart SP, Vassallo BG, Angeletti D, Hutchinson DS, Morgan AP, Takeda K, Hickman HD, McCulloch JA, Badger JH, Ajami NJ, Trinchieri G, de Villena FP-M, Yewdell JW, Rehermann B**. 2017. Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance. Cell 171:1015-1028.e13.
18. **Moeller AH, Foerster S, Wilson ML, Pusey AE, Hahn BH, Ochman H**. 2016. Social behavior shapes the chimpanzee pan-microbiome. Sci Adv 2:e1500997.
19. **DNeasy PowerSoil Kit Handbook - QIAGEN. https://www.qiagen.com/ca/resources/resourcedetail?id=5a0517a7-711d-4085-8a28-**

2bb25fab828a&lang=en. Retrieved 23 April 2022.
20. **MO BIO's PowerSoil DNA Isolation Kit Handbook - QIAGEN. https://www.qiagen.com/ca/resources/resourcedetail?id=5c00f8e4-c9f5-4544-94fa-653a5b2a6373&lang=en.** Retrieved 23 April 2022.
21. **MiSeq System | Focused power for targeted gene and small genome sequencing. https://www.illumina.com/systems/sequencing-platforms/miseq.html**. Retrieved 23 April 2022.
22. **R: The R Project for Statistical Computing. https://www.r-project.org/.** Retrieved 15 April 2022.
23. **RStudio | Open source & professional software for data science teams. https://www.rstudio.com/.** Retrieved 17 April 2022.
24. **Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H**. 2019. Welcome to the Tidyverse. J Open Source Softw 4:1686.
25. **Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, vander Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG.** 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857.
26. **Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP.** 2016. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583.
27. **Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D.** Scikit-learn: Machine Learning in Python. Mach Learn PYTHON 6.
28. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196.
29. **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:D590–D596.
30. **Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability**. Mol Biol Evol 30:772–780.
31. **Katoh K, Misawa K, Kuma K, Miyata T.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.
32. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE 5:e9490.
33. **Shannon C, Weaver W.** The Mathematical Theory of Communication 131.
34. **Faith DP.** 1992. Conservation evaluation and phylogenetic diversity. Biol Conserv 61:1–10.
35. **Pielou EC.** 1966. The measurement of diversity in different types of biological collections. J Theor Biol 13:131–144.
36. **Jaccard P.** 1908. Nouvelles Recherches Sur la Distribution Florale. Bull Soc Vaudoise Sci Nat 44:223–70.
37. **Sørenson T.** 1948. A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons. I kommission hos E. Munksgaard.
38. **Lozupone C, Knight R.** 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. Appl Environ Microbiol 71:8228–8235.
39. **Lozupone CA, Hamady M, Kelley ST, Knight R.** 2007. Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. Appl Environ Microbiol 73:1576–1585.
40. **Kruskal WH, Wallis WA.** 1952. Use of Ranks in One-Criterion Variance Analysis. J Am Stat Assoc 47:583–621.
41. **Anderson MJ.** 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46.
42. **Vázquez-Baeza Y, Gonzalez A, Smarr L, McDonald D, Morton JT, Navas-Molina JA, Knight R.** 2017. Bringing the Dynamic Microbiome to Life with Animations. Cell Host Microbe 21:7–10.
43. **Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R.** 2013. EMPeror: a tool for visualizing high-

throughput microbial community data. GigaScience 2:16.
44. **jbisanz/qiime2R: qiime2R version 0.99.6 from GitHub. https://rdrr.io/github/jbisanz/qiime2R/.** Retrieved 17 April 2022.
45. **Love MI, Huber W, Anders S.** 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
46. **Liu C, Cui Y, Li X, Yao M.** 2021. microeco: an R package for data mining in microbial community ecology. FEMS Microbiol Ecol 97:fiaa255.
47. **Microsoft Excel Spreadsheet Software | Microsoft 365. https://www.microsoft.com/en-us/microsoft-365/excel.** Retrieved 17 April 2022.
48. **Industry-leading vector graphics software | Adobe Illustrator. https://www.adobe.com/ca/products/illustrator.html.** Retrieved 17 April 2022.
49. **McMurdie PJ, Holmes S.** 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.
50. **Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara B, Simpson G, Solymos P, Stevens H, Wagner H.** 2015. Vegan: Community Ecology Package. R Package Version 22-1 2:1–2.
51. **Paradis E, Schliep K.** 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.
52. **Wald A.** 1943. Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. Trans Am Math Soc 54:426–482.
53. **Wickham H.** 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.
54. **Liaw A, Wiener M**. 2002. Classification and Regression by randomForest 2:5.
55. **Streamlined Plot Theme and Plot Annotations for ggplot2. https://wilkelab.org/cowplot/.** Retrieved 17 April 2022.
56. **Ives A r., Klug J l., Gross K.** 2000. Stability and species richness in complex communities. Ecol Lett 3:399–411.
57. **Archie EA, Tung J.** 2015. Social behavior and the microbiome. Curr Opin Behav Sci 6:28–34.
58. **Edouard S, Sankar S, Dangui NPM, Lagier J-C, Michelle C, Raoult D, Fournier P-E.** 2014. Genome sequence and description of Nesterenkonia massiliensis sp. nov. strain NP1T. Stand Genomic Sci 9:866–882.
59. **Chander AM, Nair RG, Kaur G, Kochhar R, Dhawan DK, Bhadada SK, Mayilraj S.** 2017. Genome Insight and Comparative Pathogenomic Analysis of Nesterenkonia jeotgali Strain CD08_7 Isolated from Duodenal Mucosa of Celiac Disease Patient. Front Microbiol 8.
60. **Hyun D-W, Whon TW, Cho Y-J, Chun J, Kim M-S, Jung M-J, Shin N-R, Kim J-Y, Kim PS, Yun J-H, Lee J, Oh SJ, Bae J-W.** 2013. Genome sequence of the moderately halophilic bacterium Salinicoccus carnicancri type strain CrmT (= DSM 23852T). Stand Genomic Sci 8:255–263.
61. **Samanta I, Bandyopadhyay S.** 2020. Chapter 18 - Actinobacillus, p. 233–239. *In* Samanta, I, Bandyopadhyay, S (eds.), Antimicrobial Resistance in Agriculture. Academic Press.
62. **Phillips JE.** 1990. 12 - Actinobacillus, p. 143–149. *In* Carter, GR, Cole, JR (eds.), Diagnostic Procedure in Veterinary Bacteriology and Mycology (Fifth Edition). Academic Press, San Diego.
63. **Krzyściak W, Pluskwa KK, Jurczak A, Kościelniak D.** 2013. The pathogenicity of the Streptococcus genus. Eur J Clin Microbiol Infect Dis 32:1361–1376.
64. **Good MF.** 2020. Streptococcus: An organism causing diseases beyond neglect. PLoS Negl Trop Dis 14:e0008095.
65. **Iglewski BH.** 1996. PseudomonasMedical Microbiology. 4th edition. University of Texas Medical Branch at Galveston. https://www.ncbi.nlm.nih.gov/books/NBK8326/. Retrieved 17 April 2022.
66. **Corvec S**. 2018. Clinical and Biological Features of Cutibacterium (Formerly Propionibacterium) avidum, an Underrecognized Microorganism. Clin Microbiol Rev 31:e00064-17.
67. **Sobko T, Liang S, Cheng WHG, Tun HM.** 2020. Impact of outdoor nature-related activities on gut microbiota, fecal serotonin, and perceived stress in preschool children: the Play&Grow randomized controlled trial. Sci Rep 10:21993.
68. **Dill-McFarland KA, Tang Z-Z, Kemis JH, Kerby RL, Chen G, Palloni A, Sorenson T, Rey FE, Herd P.** 2019. Close social relationships correlate with human gut microbiota composition. 1. Sci Rep 9:703.
69. **Monda V, Villano I, Messina A, Valenzano A, Esposito T, Moscatelli F, Viggiano A, Cibelli G, Chieffi S, Monda M, Messina G.** 2017. Exercise Modifies the Gut Microbiota with Positive Health Effects. Oxid Med Cell Longev 2017:3831972.
70. **Robinson CM, Pfeiffer JK**. 2014. Viruses and the Microbiota. Annu Rev Virol 1:55–69.
71. **Liang G, Bushman FD.** 2021. The human virome: assembly, composition and host interactions. Nat Rev Microbiol 19:514–527.