To Resistance and Beyond: Bioinformatic Investigation of Microbial Environmental Resistance Genes Using Previously Generated Metagenomic and Whole-Genome Sequencing

By

ALONNA WRIGHT
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Jonathan A. Eisen, Chair

Joanne B. Emerson

Maurice Pitesky

Karl A. Dawson

Committee in Charge

2022

Table of Contents

Acknowledgments

I'd like to express my deepest appreciation to my Principal Investigator, Dr. Jonathan Eisen, for his support, acceptance, accountability, and encouragement of my development as a whole person and not just as a training scientist. His kindness and compassion as a mentor, especially during the COVID-19 pandemic, have made me a better leader and a better person. I'd like to thank all the members of the Eisen lab for always being the first to jump right in to help solve a problem, for being the best cheerleaders for anyone's accomplishments – big or small, and for being the most welcoming group of people when I joined the lab at a tough time in my grad school career. The culture and the people of the Eisen lab are something I can never truly put into words all the ways I am thankful for.

I'd like to thank my dissertation committee – Dr. Jonathan Eisen, Dr. Joanne Emerson, Dr. Maurice Pitesky, and Dr. Karl Dawson – for their insightful and supportive input on my dissertation research. I appreciate the time and care that goes into serving on a dissertation committee and I'm very grateful for all of your support and contributions to my degree.

A very special thanks to my friends that I've met along this journey, who really have become more like family. You've adopted me into your lives and your own families, and you've shown up to celebrate my accomplishments, and to comfort me in some of my hardest times (even when I didn't want to accept it). I can never truly articulate how thankful I am for you all, and I can't wait to build our commune together.

I'd like to thank my undergraduate mentor turned friend, Dr. Kendra Steele, for showing me that badass women can show up unapologetically in science and industry, and for teaching me to never undervalue myself for the comfort of others. Thank you for nurturing me as a scientist and pushing me to pursue goals I truly thought were unattainable.

I'd like to thank my family, especially my parents, for supporting my decisions to get me to this point (even when you didn't fully understand why in the world I would make these choices).

You are the reason I've been able to pursue my dreams, and I hope you can celebrate this moment as yours too.

Most of all, I owe the deepest gratitude to my best friend and the best life partner I could ask for, Devin. Thank you for always encouraging me to follow my dreams, even if it means moving across the country for me to make minimum wage for 5 years. Thank you for listening to every practice talk, every excited rant about a result I found, every frustrated vent about my code not working (and then helping me fix the code), and every doubtful spiral where I was convinced I would never actually finish this degree. You have been the most unwavering constant in this journey and always my safe place to land. I could not have done this without you. Thank you for going on this journey with me.

Abstract

Environmental resistance genes in a microbial community influence the taxonomic persistence and subsequently the functional capacity of a microbial population. Understanding factors influencing the retention and dissemination of environmental resistance genes will elucidate avenues of treatment and mitigation of resistant microbes. Two data sets (one metagenomic set of 18 samples, and one whole genome sequencing set of 201 isolates) were analyzed for environmental resistance genes in the context of taxonomic and functional relationships. The bioinformatic analyses of these previously-generated datasets also highlights the importance of the adoption of FAIR data practices in microbial genomics to enable the continued development of reproducible and robust analyses answering important questions with real-world relevance.

Methods

Sequencing data from previously-generated metagenomic and whole genome sequencing projects were analyzed for environmental resistance genes in the context of associated environmental variables. Metagenomic data from samples of the residential built environment of backyard poultry (BYP) owners were analyzed for the presence and taxonomic origin of antimicrobial resistance (AMR) genes. Additionally, this data set was also analyzed to determine any significant correlations between AMR genes and genetically encoded elements that may have implications for future potential therapeutic solutions, such as antimicrobial peptide genes, bacteriophage lysins, and bacterio-phage encoded auxiliary genes. Bacterial isolates collected from spacecraft-associated hardware in the Spacecraft Assembly Facility (SAF) at NASA's Jet Propulsion Laboratory (JPL) were analyzed for genes implicated in the increased likelihood of microbial survival in harsh conditions. The evaluation of these environmental resistance genes is a critical and novel approach to understanding the risk of forward contamination to extraterrestrial environments.

Metagenomic analysis of BYP-associated built environments revealed ubiquitous AMR presence and significant relationships between AMR genes and other genomic features with clear differences between homes that administered antibiotics to their BYP flocks and homes that did not. WGS analysis of NASA JPL bacterial isolates revealed the presence of resistance genes in all isolates, with fluctuations in frequency between taxonomic groups and between NASA missions. These studies elucidate the ecological relationships surrounding environmental resistance gene retention and reveal associations with taxonomic and functional components that with further research may lead to measures preventing the retention and dissemination of environmental resistance genes.

Introduction

Microbial genomic data has been growing exponentially since the first bacterial genome was released in 1995, while sequencing cost per raw megabase of DNA sequence continues to remain well below the predicted Moore's Law, averaging $0.006 per megabase of DNA sequence in August 2021 (1, 2). The continued innovation of Next Generation Sequencing (NGS) technologies has enabled microbial sequencing to become more affordable and accessible, which in turn allows microbial scientists to continue to seek out answers to important biological questions using sequencing data. This positive feedback loop of data generation sparking further research questions has significantly increased the rates of sequencing data generation. At the time of this dissertation's publication, there are 3,039,101 results for "metagenome" search in the NCBI Sequence Read Archive (SRA), 459,976 prokaryotic genomes in the NCBI Genome Database, and it's predicted that genomic data will reach between 2 and 40 exabytes by 2025 (3, 4). Many of these DNA sequencing data sets are analyzed to answer a very specific set of research questions, however, there is valuable information contained in these datasets far beyond their initial analyses.

Emphasizing the importance of FAIR data – data that is Findable, Accessible, Interoperable, and Reusable (5), – in microbial genomics will pave the path for continued analyses of existing sequencing datasets to answer globally relevant questions in a way that is unprecedented in this field, enabling reproducible and robust analyses of cross-study datasets. There are multiple organizations working towards FAIR data standard adoption in microbiology research, including the National Microbiome Data Collaborative (NMDC). The NMDC approaches the adoption of FAIR data practices by educating the scientific community on best practices and benefits of FAIR data in microbial research, developing gold standard analysis pipelines, and development of a data portal for microbiome data to enable centralized access to FAIR data for

the field (6). I have had the pleasure of working with the NMDC as a Champion and as a member of their Scientific Advisory Board since 2019, and the work I have contributed to through the NMDC has been one of the driving motivations behind centering my dissertation analyses around previously generated datasets. All datasets analyzed in this dissertation were from samples that were collected and sequenced prior to my involvement with each of the projects. The analyses I performed for each of these sections include investigations of research questions surrounding environmental resistance genes, many of which go beyond the original scope of each project. Environmental resistance genes, as categorized within these analyses, are genes contributing to a microbe's increased likelihood of survival in harsh conditions, such as in the presence of antibiotics (Chapters 1 and 2) or survival in NASA JPL's spacecraft assembly Clean Rooms (Chapter 3). This dissertation serves as an example of the benefits of FAIR data adoption in the field of microbial genomics, highlighting the value that previously-generated data can bring to addressing novel research questions quickly and easily when empowered by FAIR data practices.

Chapter 1: Microbial Community Analysis of Backyard Poultry-Associated Homes Reveals Ubiquitous Presence and Correlations of Antimicrobial Resistance- and Antimicrobial Peptide-Encoding Genes

Introduction

Antimicrobial resistance (AMR) genes are genetic encoders of bacterial functions that contribute to resistance to specific antibiotic molecules, including the ability to inactivate the antibiotic compound, prevent the uptake or metabolization of the antibiotic, or increase biofilm production in the presence of antibiotics (7). Rates of AMR infections are rising and pose a threat to global health and agriculture (8, 9). AMR bacterial infections are a known global epidemic causing 50,000 deaths each year in the US and Europe and are predicted to rise to 10 million per year by 2050 (7). This increase in AMR is largely attributed to the overuse and misuse of antibiotics for treatment and prophylaxis of bacterial infections in medicine and agriculture and is amplified by the decline in the rate of antibiotic discovery to combat emerging resistance (8, 9). Environmental contributors to AMR can stem from the misuse of antibiotics in agricultural situations through water runoff, dust distribution, and intermediate vectors such as humans (10, 11). Understanding the mechanisms underlying the persistence and dissemination of AMR can aid in the prevention and mitigation of environmental AMR reservoirs.

Backyard poultry (BYP) ownership has many benefits to the humans in the household, including companionship and food production through egg and meat consumption. This coexistence of owners in close proximity to their BYP flocks also comes with potentially harmful risks as well, including the perpetuation of AMR in BYP-associated microbiomes due to improper disease mitigation and treatment by the BYP owners. Approximately 13 million rural, urban, and

suburban US residents reported owning backyard poultry in 2014, and interest in BYP ownership is nearly four times that amount (12). BYP ownership has risen recently due to product quality, public health, ethical, and animal welfare concerns of commercial operations. However, BYP ownership and disease treatment is largely under-regulated, unlike commercial poultry production. A 2014 survey of the 150 most populated urban jurisdictions found that only 3 required veterinary oversight for the treatment of ill birds, only 52 required any permitting to have a flock, and only 3 of the 52 required any linked educational component (13). Only 18.8% listed a general or avian veterinarian as a source of information, 1.7% mentioned a commercial poultry veterinarian as a source of information, and 89.1% reported no previous visits to a veterinarian for the care of their BYP flock (12). This lack of relationship with veterinary professionals regarding BYP flocks could be due to a lack of access or availability, but it still highlights a large gap in oversight for the care of urban BYP flocks.

Lack of regulation regarding disease prevention and treatment for urban BYP flocks poses public health concerns regarding the transmission of AMR bacteria, such as AMR strains of *Salmonella*, *Mycoplasma gallisepticum*, and *Escherichia coli* commonly associated with BYP. In a 2014 survey, BYP owners were largely uninformed about poultry diseases and treatments but were interested in learning more about disease management (12). Additionally, a 2010 USDA survey found that more than 50% of urban poultry owners were unaware that contact with poultry poses infectious disease risk, 25% reported not washing hands after handling live poultry, and 15.5% reported that chickens had been inside their home/living space in the past 3 months (14). The combination of a lack of regulation and public information warrants further research into the bacterial communities of BYP and their environments, including the built environments of their human companions.

Built environments, or structures built by humans, such as residential and occupational buildings and vehicles, are well established as major contributors to the microbial community of its occupants and have been implicated in the exacerbation and mitigation of human diseases

(15). The microbiome of the residential built environment can especially influence the microbial composition of its inhabitants' microbiomes since on average humans spend approximately 90% of their time indoors (15, 16). Considering the significant exposure time to the microbial community of the built environment, AMR frequency of built environments should be investigated further to understand the risk of exposure individuals face when indoors. Abiotic factors of the built environment can significantly influence the microbial community composition, auxiliary gene retention, and functional capacity. Such factors include pH, temperature, humidity, paint material, and antimicrobial compound concentration, among others (17). These pressures can also influence the frequency and capacity of microbial intra-community interactions, such as horizontal gene transfer events and exogenous DNA uptake, especially in the environmental context of the built environment (18). The cumulative effect of these selective pressures and potentials for microbial interactions associated with increased AMR prevalence poses the question: to what extent do human-inflicted variables significantly impact the frequency and composition of AMR genes in the built environment?

The intersection of the concerns of AMR infection rates, lack of disease prevention and treatment knowledge of urban BYP flock owners, and the implications of the built environment's influence on human microbiome composition and function, all contribute to the need to further characterize the microbiome of the homes of BYP owners to elucidate these interactions, particularly characterizing relationships between AMR and genomic features that have therapeutic potential, such as antimicrobial peptides (AMP). This study aimed to characterize the microbial communities of external and internal door frames of BYP flock owners and understand the relationship between the care provided to the BYP flock and associated changes in the built environment microbial community.

Materials and Methods

Sample Collection and Sequencing

A community science project was conducted where BYP owners were asked to provide swab samples for sequencing analysis from areas of interest to microbial community composition relative to BYP ownership. The purpose of this study was to (1) understand antibiotic resistance in backyard poultry flocks, (2) understand seasonal changes in antibiotic resistance, (3) provide information regarding the careful and proper use of antibiotics to backyard poultry owners. BYP owners self-reported their antibiotic administration to their flocks, in addition to other relevant metadata including the number of chickens in their flocks, frequency, and method of cleaning the outdoor BYP environment. A full table of this metadata can be found in Supplementary Tables S-Metadata 1A and 1B.

Participants were mailed sterile swabs and instructions on environmental sample collection. Swabs were brushed repeatedly against the environmental sample surface, then placed into a provided vial and sealed prior to being mailed back to the UC Davis laboratory for processing. This sample collection was part of a larger study that encompassed other animal and environmental samples. Samples were collected from the interior and exterior surfaces (henceforth referred to as Indoor and Outdoor, respectively) of the doorframes of each home. A description of the surfaces to be swabbed that was provided to the participants can be found in Table 1.1.

Table 1.1: Description Provided to Participants of Community Science Study of Swab Surfaces for Indoor and Outdoor Door Frames

Label on tube	**Description of surface to be swabbed**
#1: Exterior door frame – outside	The main entrance to your house or apartment building that is exposed to the outside environment. Sample the top of the door frame on the outside of the door (the small ledge on the top of the door frame where dust collects)
#2: Interior door frame - inside	An interior door on the main floor of your home. Sample the top of a door frame that is exposed to activity (and collects dust) in the main living area of your house (this could be the interior door frame of your front door, a closet, or the entrance to the kitchen, for example).

Homes were categorized into “Antibiotic” and “Antibiotic-Free” based on their self-reported antibiotic usage on their BYP flocks. The detail of reported antibiotic usage varied widely, and largely did not include concentration or frequency of antibiotic administration, however, treatment completion status and veterinary oversight were both reported in the metadata. This information can be found in the metadata table found in Supplementary Tables S-Metadata 1A and 1B. Table 1.2 describes all antibiotics reported to be used in the Antibiotic homes’ BYP flocks, including microbial targets and antibiotic class.

Table 1.2: Antibiotics Administered to BYP Flocks in Antibiotic Group of Sampled Homes. Information on administration was self-reported by BYP-owners and can be found in Table S-Metadata 1B.

Name	Description	Class	Effective Against
Enrofloxacin* (Baytril, Ciprofloxacin) (19, 20) **banned from use in poultry since 2005 (21)*	Broad spectrum fluoroquinolone	Fluoroquinolone	Mycoplasma and most Gram-negative bacteria, including: *Escherichia coli Enterobacter spp Klebsiella spp Pasteurella spp Proteus spp Salmonella Pseudomonas aeruginosa* (variably susceptible) And some Gram-positive bacteria: *Staphylococcus aureus, Staphylococcus intermedius*
Amoxicillin-Clavulanic Acid (Clavamox) (22)	Potentiated penicillin (broad spectrum amoxicillin + beta lactamase inhibiting clavulanate potassium)	Beta-lactam	Gram-positive bacteria, Gram-negative bacteria, beta-lactamase-producing strains
Amoxicillin (23)	Broad spectrum penicillin	Beta-lactam	Gram-positive bacteria, Gram-negative bacteria
Tylosin (Tylan) (24)	Broad-spectrum macrolide antibiotic, similar mechanism of action as erythromycin	Macrolide	Gram-positive bacteria (*Staphylococcus, Streptococcus, Listeria, Erysipelothrix, Enterococcus, Corynebacterium,* and *Clostridium*), Mycoplasma, Chlamydophila, and Pasteurella.

DNA from submitted swabs was extracted using PowerSoil DNA extraction kits (Qiagen, Cat. No. 47016), and sequencing was performed using the Illumina MiSeq platform by the UC Davis Genome Center Sequencing Core.

Sequencing Processing and Metagenomic Assembly

Illumina sequences were processed using BBDuk (25) to remove sequencing adapters and trim and filter for quality. Reads were bidirectionally trimmed to remove any regions with an average quality below Q10 using the Phred algorithm, and any reads shorter than 80 nucleotides were discarded (qtrim=rl trimq=10 minlength=80). Filtered reads were then co-assembled using MEGAHIT using the kmin-1pass option to optimize memory efficiency for ultra-low-depth datasets, and a minimum contig length of 300 nucleotides (26). To calculate the relative abundance of each contig reads from each sample were mapped back to the coassembly using bbmap (version 37.62) and reported by the scafstats option (25). Count Per Million (CPM) was calculated using EdgeR to normalize the relative abundance of each contig to the relative sequencing depth of each sample using the Trimmed Mean of M-values (TMM) method of normalization (27–29). While this package was developed with the intention of differential expression analysis of RNA sequencing data, these types of analyses have been co-opted with precedence for use of calculating differential relative abundance of metagenomic features (29). Sample A8_2 was discarded from further analysis due to the low relative abundance of reads meeting quality thresholds.

Functional and Taxonomic Annotation

Taxonomic assignment of contigs was performed using Kaiju utilizing the non-redundant database classifying bacteria, fungi, eukaryotes, and viruses (30). Taxonomic names were

acquired using the functions kaiju2table and kaiju-addTaxonNames to retrieve the taxonomic ranks and names using the NCBI ID provided in the primary kaiju output.

Annotation of known and putative AMR genes was performed using DeepArg predict pipeline for nucleotide analysis, using the LS parameter which specifies annotating AMR genes based on full gene length sequences, rather than shorter length sequences (31). DeepArg utilizes an independently curated database that encompasses a non-redundant representation of the CARD, ARDB, and UNIPROT AMR databases (31). AMP encoding genes were predicted using the Macrel software for the analysis of contigs (32).

Statistical analysis and Visualization

Dataframe manipulations were performed using tidyverse and base R functions to ensure tables were suitably formatted for downstream analysis and visualizations (33). Pearson correlations and statistical tests were performed using the rcorr function of the Hmisc package in R and only statistically significant Pearson correlations were graphically represented (34). Heatmaps were generated using the ggcorrplot R package (35), and pairwise Pearson correlations were hierarchically clustered by significance level using the hclust function of the stats base R package (36). Heatmaps displaying correlations between two unique sets of variables were unable to be clustered due to the technical limitations of the package. Associations of taxonomic and functional annotations on a per contig basis were performed through the joining of all contig-relative data frames using custom scripts built on tidyverse functions. All scripts can be found in the GitHub repository at https://github.com/alonnawright/backyardpoultry_metagenomic_analysis. All metadata and high-quality figure images can be found in the figshare repository for this project (https://doi.org/10.6084/m9.figshare.21585807) (37).

Venn diagrams of contig annotations were performed using the venn R package (38), with an input of a boolean table derived from the culmination of Kaiju, VIBRANT, VirSorter, Macrel,

and DeepArg analyses. Contigs were denoted as “phage” if the contig ID was in either VirSorter or VIBRANT packages, or if the Kaiju taxonomic assignment was any of the following at the Family or Order levels: "Caudovirales", "Myoviridae", "Siphoviridae", "Podoviridae", "Microviridae", "Corticoviridae", "Tectiviridae", "Leviviridae", "Cystoviridae", "Inoviridae", "Plasmaviridae". The superkingdom taxonomic rank provided by the Kaiju analysis determined the taxonomic category assignment, where any contig that was not Bacteria or Eukaryota and had not been determined to be of phage origin, was classified as “Other”. Any contig ID appearing in the DeepArg or Macrel results tables was classified as “AMR” and “AMP”, respectively.

A phyloseq object was generated using the following inputs: OTU table - read count data from bbmap scafstats, tax_table - kaiju taxonomic assignment, sample_data - manually curated metadata file. Stacked bar plots of relative taxonomic abundance were generated using the phyloseq function plot_bar, while stacked bar plots of functional annotations were generated using ggplot geom_bar and represented the CPM relative abundance of contigs associated with that function as determined by the individual functional analyses.

Shannon and Chao1 alpha diversity metrics were calculated and visualized using the estimate_richness and plot_richness functions in phyloseq, respectively (39). Bray Curtis beta diversity metrics were calculated and visualized as ordinations using the distance and ordinate functions within the phyloseq package, respectively (39).

A complete description of the R session info, including all package names and versions, can be found in the Supplementary section of this dissertation.

Results and Discussion

Alpha Diversity

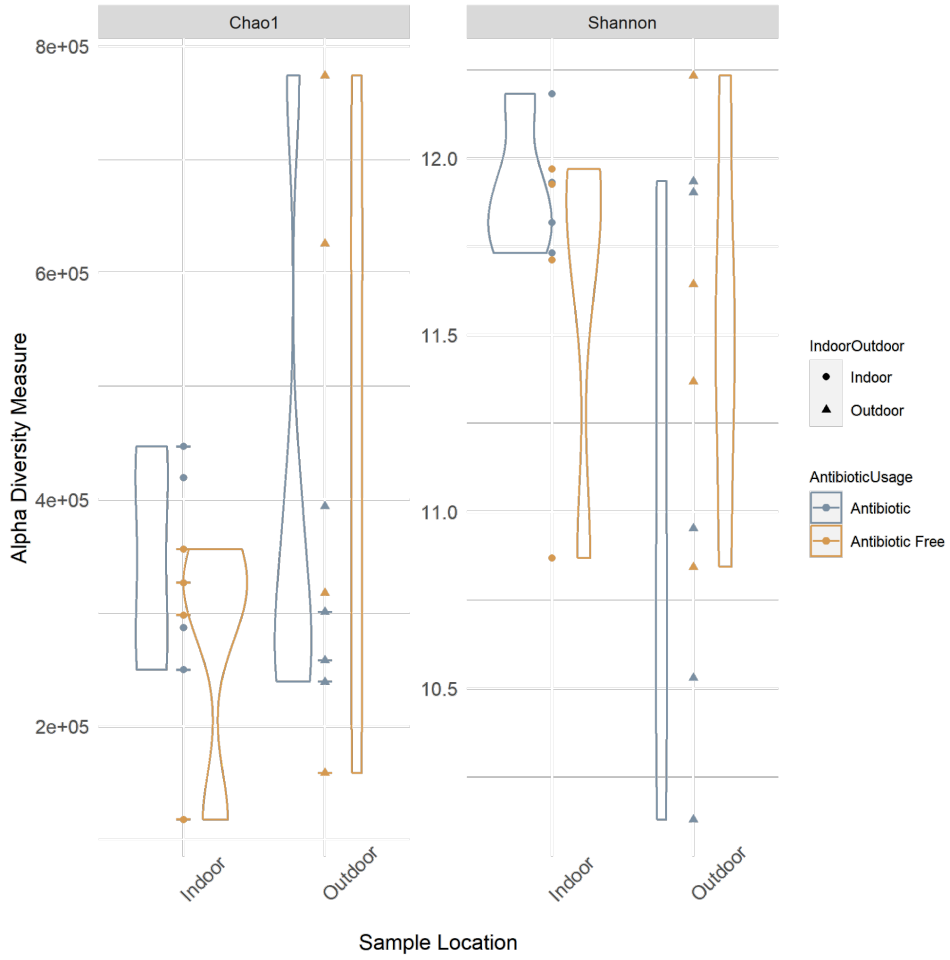


Figure 1.1: Alpha Diversity Did Not Differ Between Sample Groups. Chao1 and Shannon alpha diversity metrics of microbial communities of internal (indoor) and external (outdoor) door frames of BYP Owning Homes, significance determined by Kruskal-Wallis rank sum test at $p < 0.05$. Violin plots display mirrored density of continuous distributions.

Quantifying the alpha (within a sample) and beta (between samples) diversities can help contextualize the members of the community and the breadth of potential interactions which may influence functional characteristics in a microbial community (40). Alpha diversity captures the diversity of a community within a single sample, which is a helpful metric in understanding the

potential ecological interactions within a system (41). Shannon and Chao1 alpha diversity metrics, which measure abundance and richness respectively, were evaluated for each BYP-owning home sample. Shannon alpha diversity was not significantly different between Antibiotic and Antibiotic Free groups (Kruskal-Wallis rank sum test, p = 1), nor between Indoor and Outdoor groups (Kruskal-Wallis rank sum test, p = 0.1489) (Figure 1.1). Similarly, Chao1 alpha diversity was not significantly different between Antibiotic and Antibiotic Free groups (Kruskal-Wallis rank sum test, p = 1), nor between Indoor and Outdoor groups (Kruskal-Wallis rank sum test, p = 0.5637) (Figure 1.1). No significant differences in alpha diversity between any of the groups mean that microbial diversity within samples was of similar measurements between all evaluated variables.

Beta Diversity

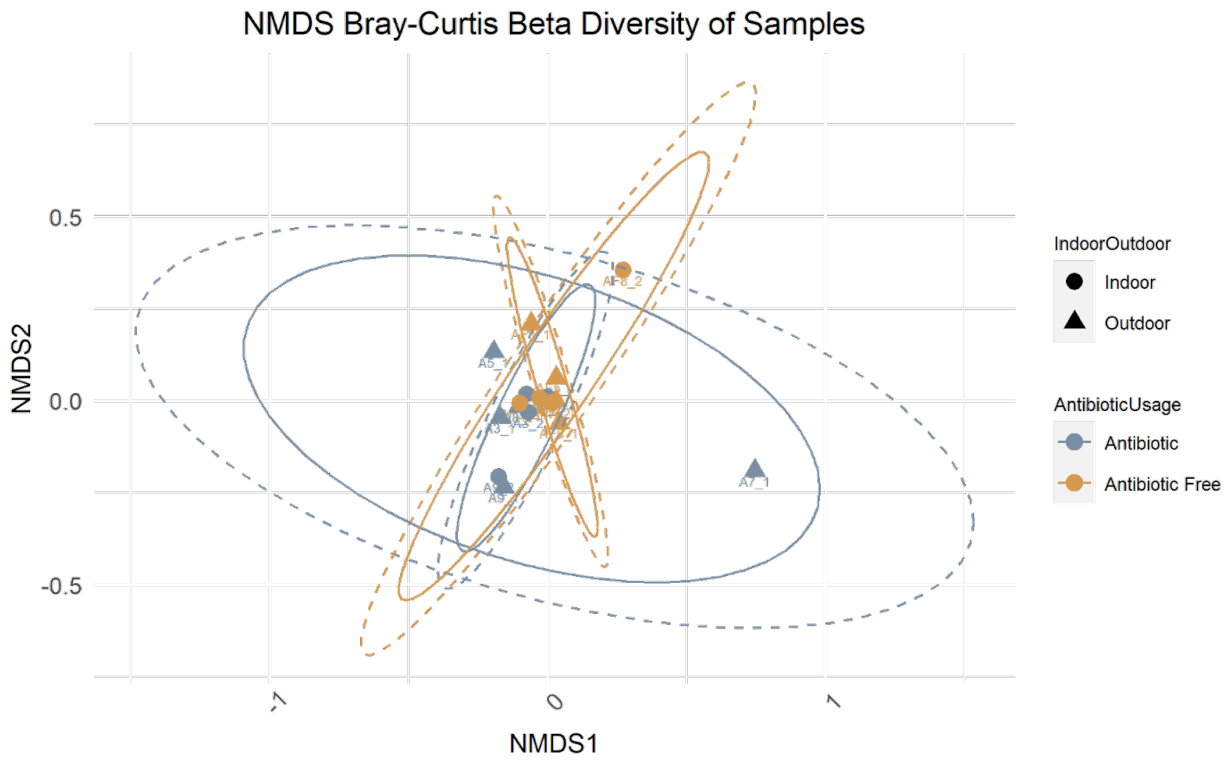


Figure 1.2: Non-metric Multidimensional Scaling (NMDS) of Bray-Curtis Beta Diversity Metrics of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes, significance determined by PERMANOVA significance test at p < 0.05. Solid-line

ellipses represent the multivariate t-distribution, while dotted-line ellipses represent the multivariate normal distribution. NMDS1 and NMDS2 axes represent the two most representative metrics of the calculated distances in multidimensional space.

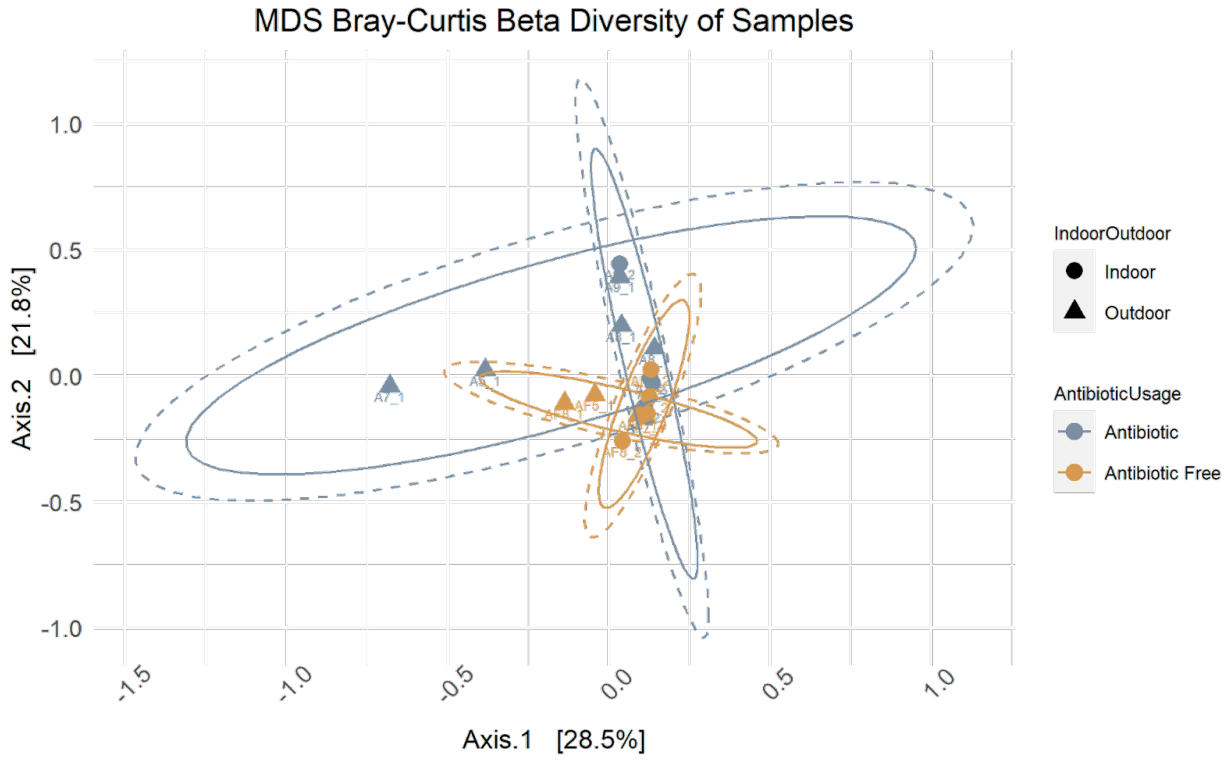


Figure 1.3: Antibiotic and Antibiotic-Free samples had significantly different beta diversity of microbial communities, while Indoor and Outdoor samples did not. Multidimensional Scaling (MDS) of Bray-Curtis beta diversity metrics of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes, significance determined by PERMANOVA significance test at $p < 0.05$. Solid-line ellipses represent the multivariate t-distribution, while dotted-line ellipses represent the multivariate normal distribution. Axes represent the two most explanatory eigenvalues, percentages represent the percent of variation attributed to each axis.

Beta diversity quantifies the differences in diversity between samples and serves a proxy metric to understand the impact of environmental variables on community composition (42). Bray Curtis distances are visualized in Figures 1.2 and 1.3 in Nonmetric Multidimensional Scaling (NMDS) and Multidimensional Scaling (MDS) plots, visualizing the values by their ranking of

dissimilarities based on Euclidean distances and by their calculated dissimilarities, respectively (43) Bray Curtis distances, a measure of taxonomic abundance dissimilarity between samples, were significantly different between Antibiotic and Antibiotic Free samples (PERMANOVA, adonis2, p=0.018), however, these distances were not statistically significant between Indoor and Outdoor samples (PERMANOVA, adonis2, p=0.189) nor pairwise between the combinations of these two groups (PERMANOVA, adonis2, p=0.911). Dispersion of beta diversity distances was also not significantly different for Antibiotic and Antibiotic Free samples (Tukey multiple comparisons of means, p=0.185) nor between Indoor and Outdoor samples (Tukey multiple comparisons of means, p=0.068). This means that microbial community compositions were more significantly influenced by the use of antibiotics than the location of the sample.

Phylum Relative Abundances

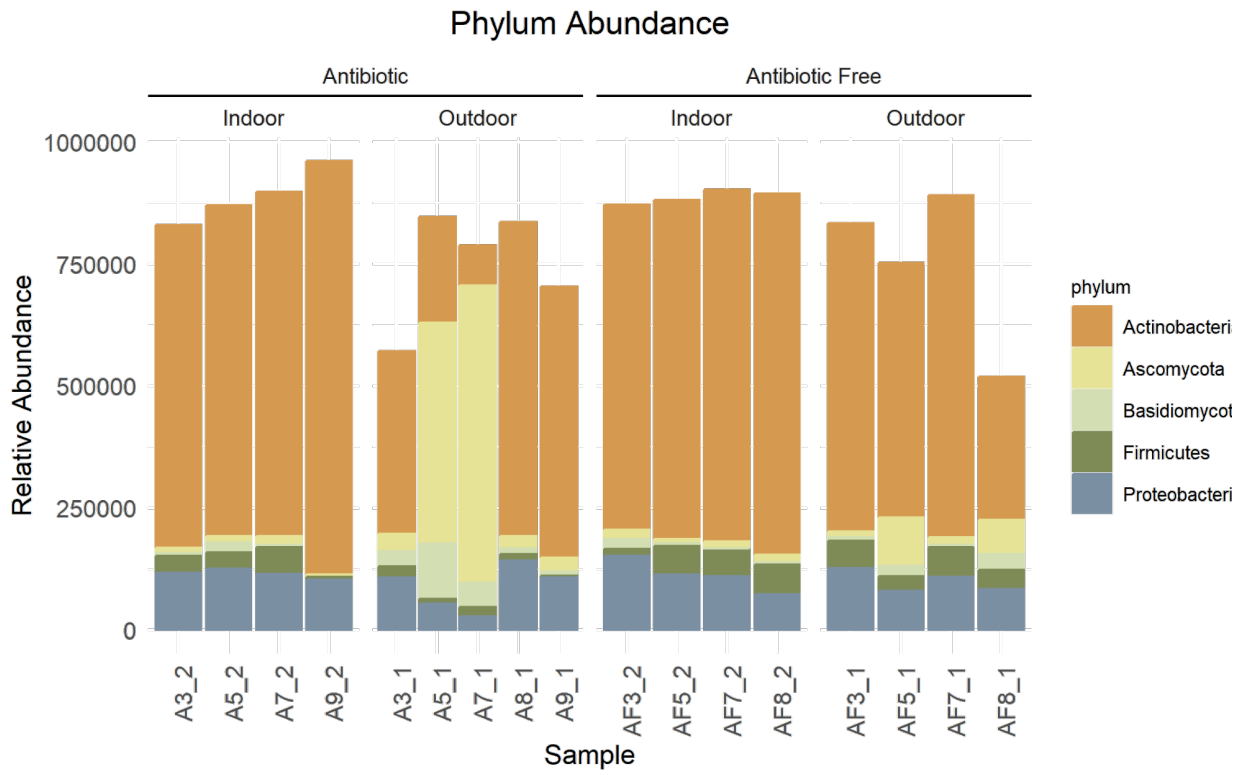


Figure 1.4: Top 5 relatively abundant microbial phyla for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized

antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

The top 5 most abundant phyla across all samples were quantified for each sample (Figure 1.4). All samples contained a similar cumulative relative abundance of these top 5 phyla. Outdoor samples contained greater amounts of fungal phyla, Ascomycota, and Basidiomycota, compared to the indoor samples, which also contained fungal contigs but at a much lower relative abundance. 15 of the 17 samples possessed Actinobacteria as the phylum with the highest relative abundance, followed by Proteobacteria, then Firmicutes, Ascomycota and Basidiomycota. However, the remaining two samples, which are both Outdoor samples from Antibiotic homes, have Ascomycota as their dominant phyla. These fungal phyla had higher relative abundances in outdoor samples than indoor samples, which is particularly interesting to contextualize the potential impacts that the difference in community composition may have on functional gene relative abundance.

Salmonella Relative Abundance

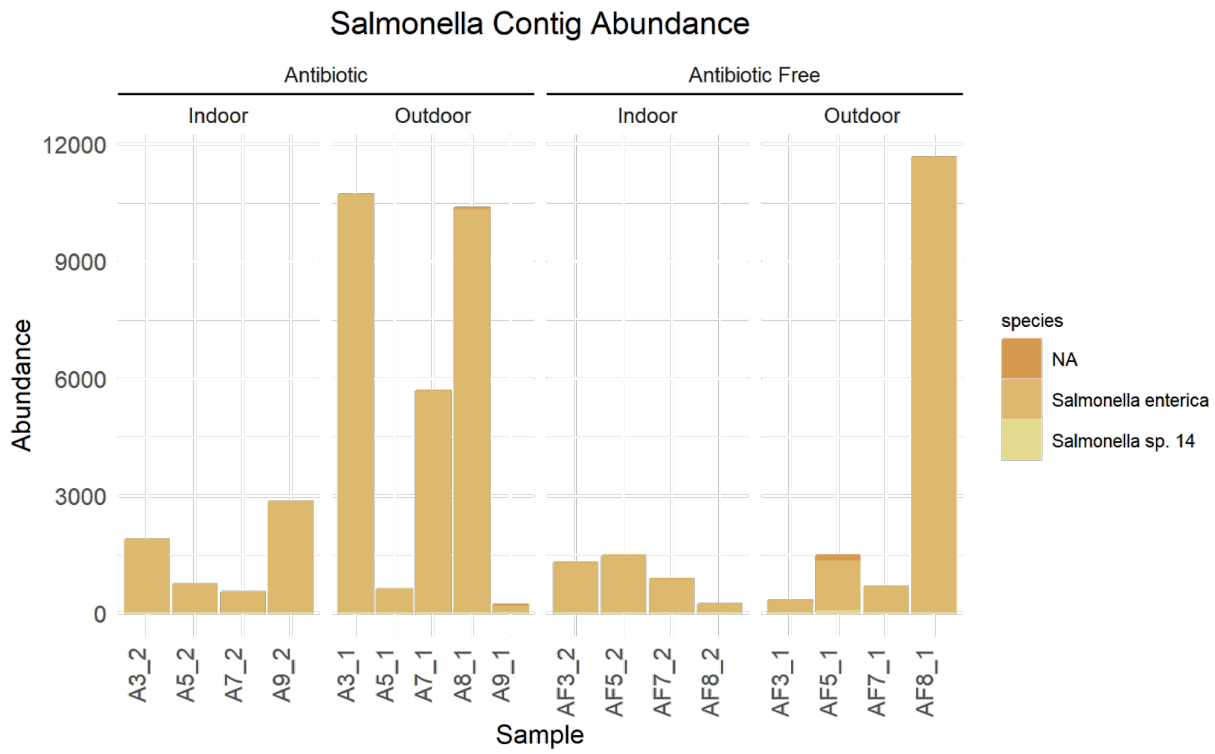


Figure 1.5: Relative abundance of contigs taxonomically identified as Salmonella for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

Contigs taxonomically assigned to the *Salmonella* genus were detected in every sample, with the overwhelming majority of Salmonella contigs being identified as *Salmonella enterica* (Figure 1.5). Relative abundances of contigs identified as *Salmonella* were generally higher in outdoor samples compared to indoor samples. Four samples reached over 3000 units of relative abundance, where three of those samples were Outdoor samples from Antibiotic homes and the remaining one sample from an Outdoor sample of an Antibiotic-Free home. Neither AMP nor AMR genes were detected in any of the contigs identified as Salmonella. Since Salmonella is one of the most commonly reported poultry-associated pathogens, understanding the relative abundance of Salmonella in these microbial communities and the influence of antibiotic usage

and built-environment location are important for the development of cleaning and care practices for BYP owners.

Functional Gene Relative Abundance and Distribution

Antimicrobial Resistance

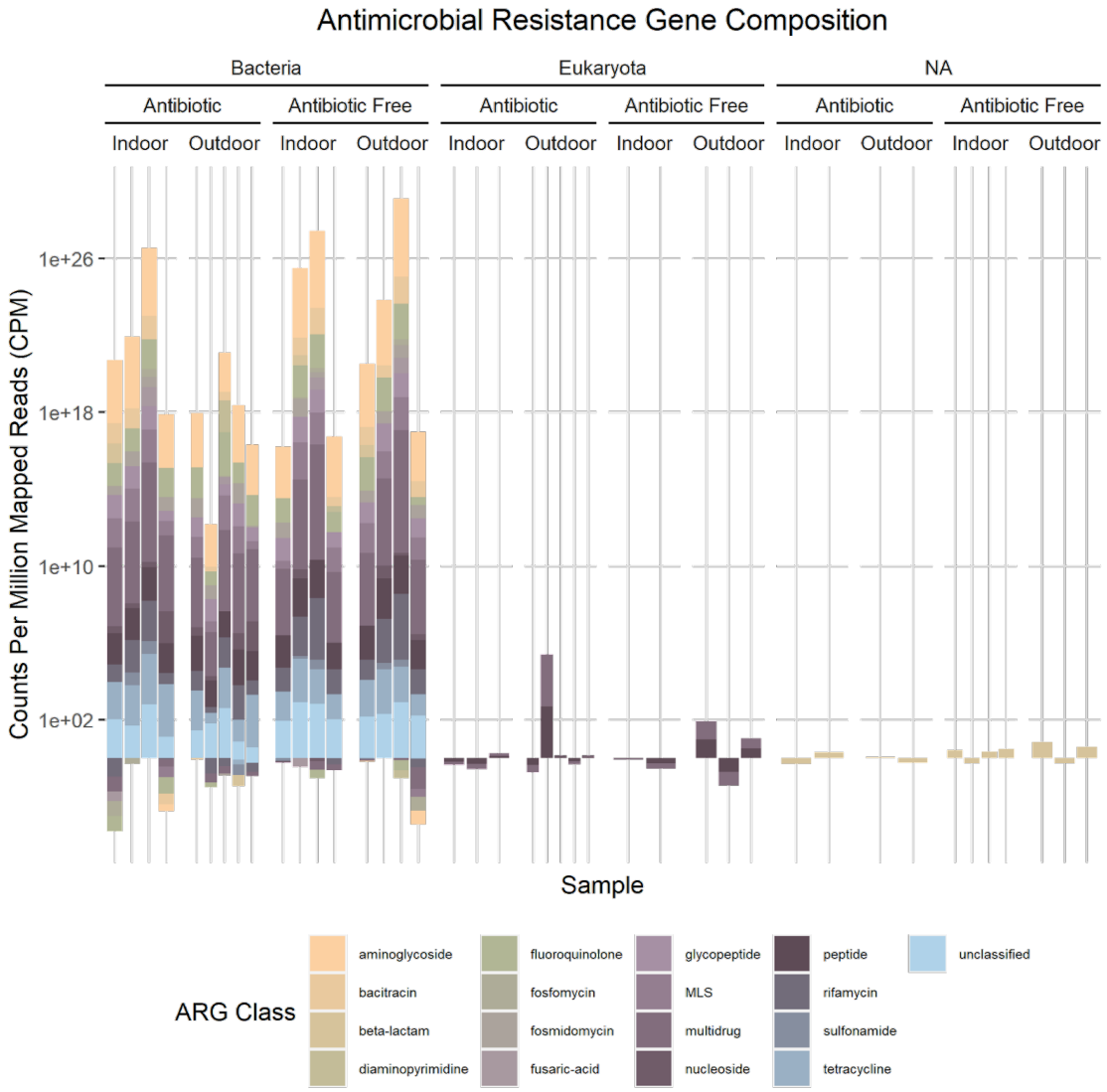


Figure 1.6: Antimicrobial Resistance Gene (ARG) Class Composition by Count Per Million Mapped Reads (CPM) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free). ARG Class CPM values are separated by identified taxonomy of the encoding contig sequence.

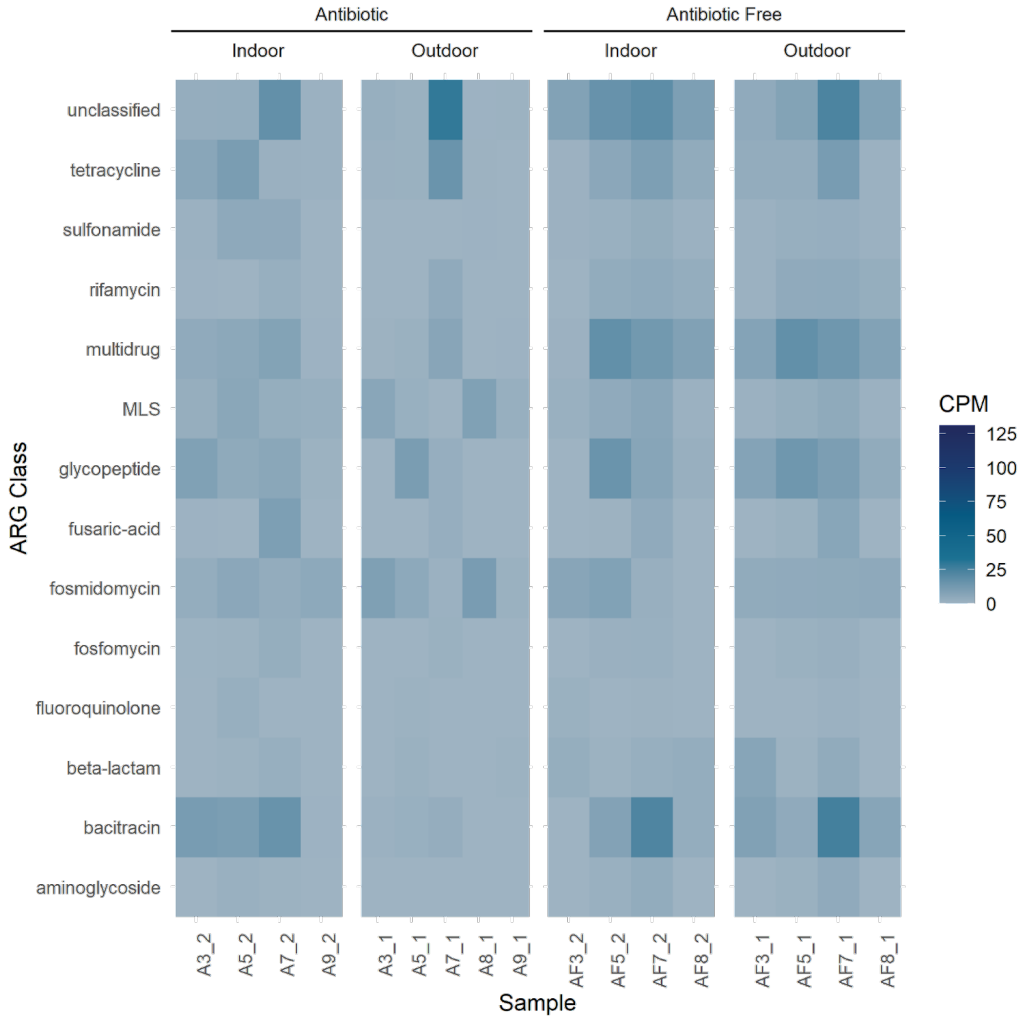


Figure 1.7: Antimicrobial Resistance Gene (ARG) class frequency in count per million mapped reads (CPM) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

The vast majority of AMR genes were contained within contigs identified as Bacteria (Figure 1.6). Eukaryota contigs contained small levels of nucleotide and multidrug AMR gene classes, and AMR genes associated with all other Kingdoms were in similarly low levels and of the aminoglycoside class. Contigs that contained AMR genes but were neither Bacteria nor Eukaryota contained only AMR genes of the beta-lactam class. ARG Class CPM values had a wide range of distributions between samples of all conditions (Figure 1.7). These values show

that there is not a visible trend of AMR gene class bias in any considered environmental condition or location.

Antimicrobial Peptides

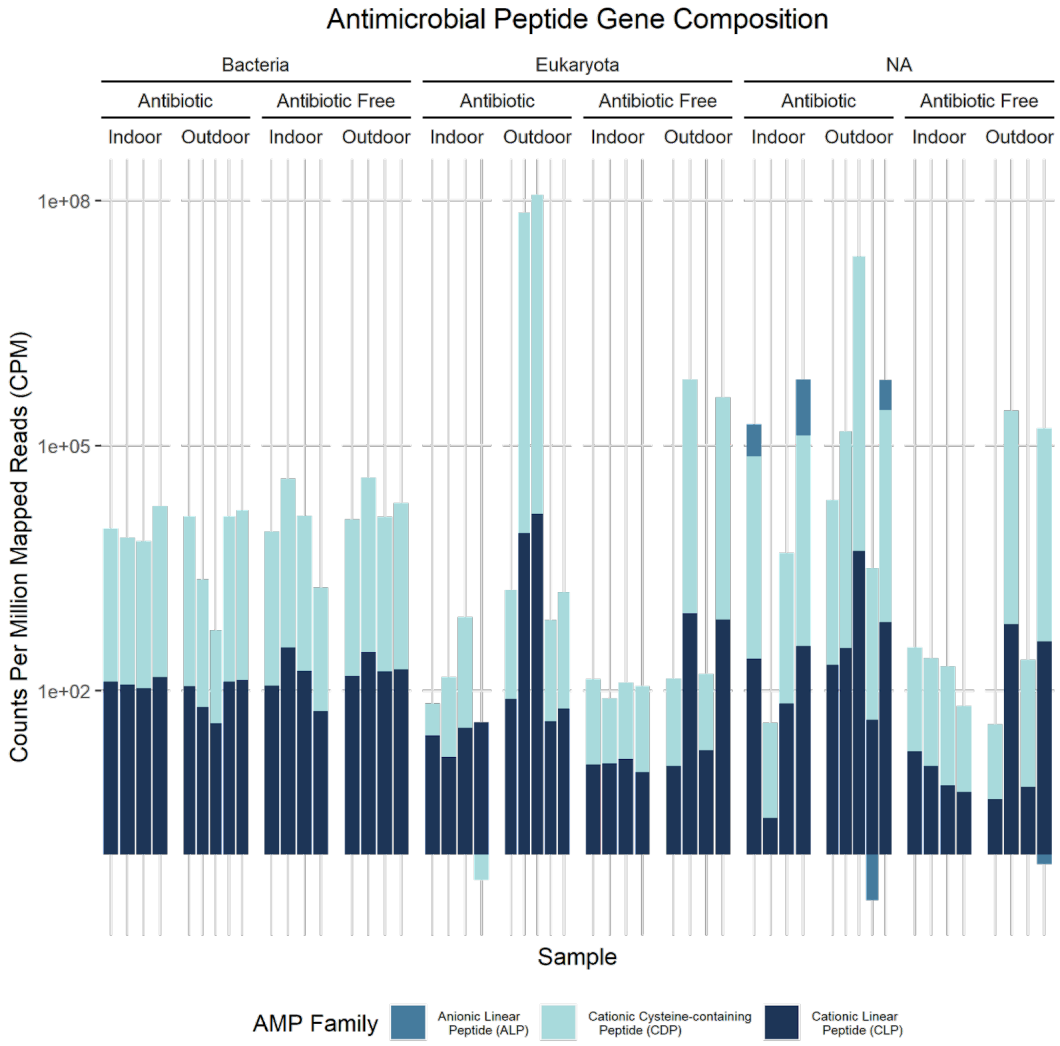


Figure 1.8: Antimicrobial Peptide (AMP) Gene Composition by Count Per Million Mapped Reads (CPM) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free). AMP CPM values are separated by identified taxonomy of the encoding contig sequence.

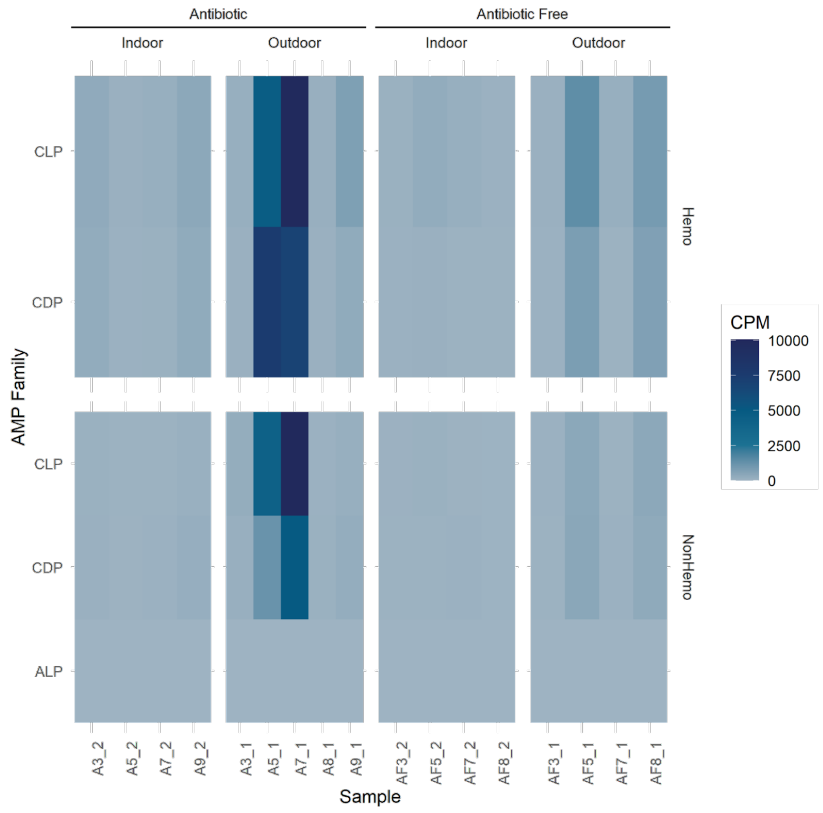


Figure 1.9: Frequency of AMP Class and Putative Hemolytic Activityclass frequency in count per million mapped reads (CPM) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

AMP content was generally evenly distributed between Cationic Cysteine-containing Peptide (CDP) and Cationic Linear Peptide (CLP), with only four samples containing Anionic Linear Peptide (ALP) AMP sequences (Figure 1.8). Outdoor samples from Antibiotic homes had a noticeably higher relative abundance of cationic AMPs, while all samples exhibited similarly low relative abundance of ALPs (Figure 1.8, 1.9). ALP sequences were detected in Indoor and Outdoor samples of Antibiotic homes and these contigs were not identified as bacterial or eukaryotic in origin. ALPs have been shown to attach to ribosomes or inhibit microbial

ribonuclease activity when in the cytoplasm, therefore further investigation of these ALP sequences is needed to determine the taxonomic origin and potential therapeutic capacity (44).

Contig Venn Diagrams

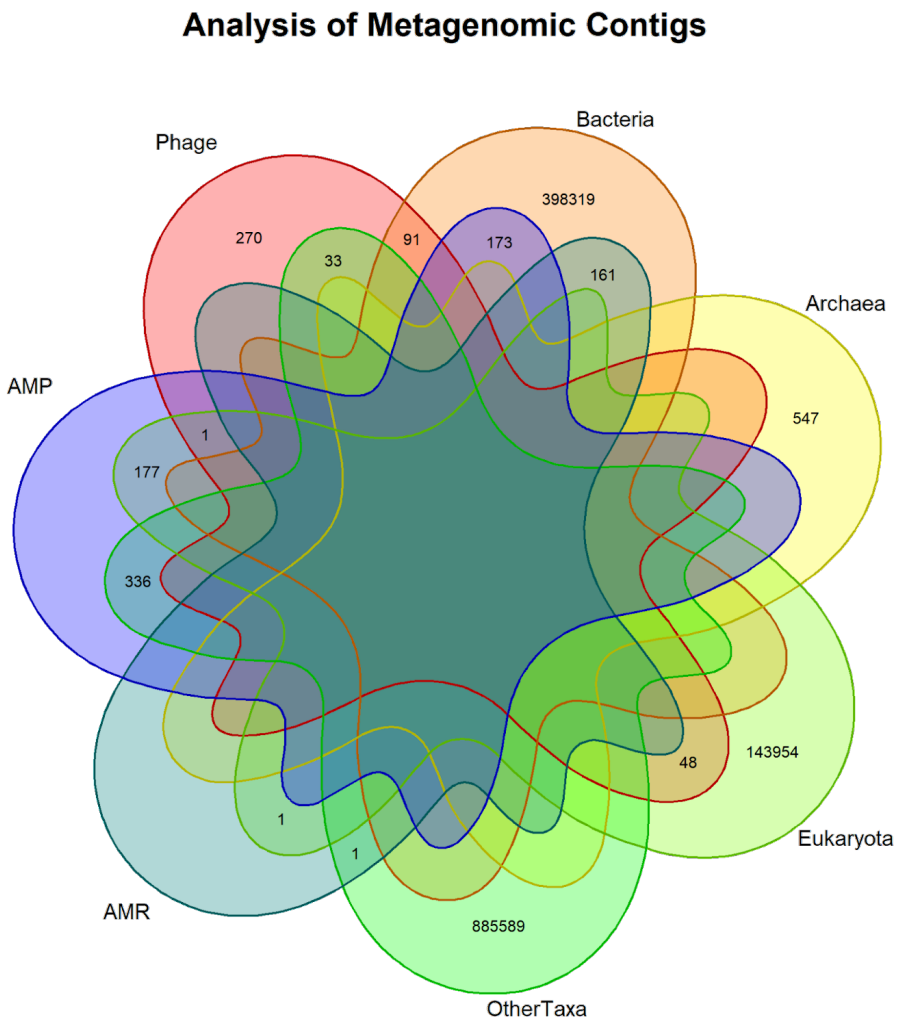


Figure 1.10: Venn diagram of frequency of individual contig functional and taxonomic assignment. Each contig was evaluated for all functional attributes and assigned a likely taxonomic origin, numbers displayed in the overlapping areas of the Venn diagram represent the instances of individual contigs (not relative abundance) that share the common attributes assigned to the overlapping section.

Contigs were cataloged to represent all taxonomic and functional assignments associated with each individual contig (Figure 1.10). Ninety-one contigs were assigned to both Bacterial and Phage taxonomies, 48 contigs were assigned to both Phage and Eukaryota, and 33 were assigned to both Phage and Other Taxa, indicating potential prophage detection or a contig population to be investigated for further refinement of phage taxonomic identification methods.

AMR-identified contigs were overwhelmingly identified as bacterial in origin, with contigs taxonomically identified as Eukaryota in one contig, Other Taxa in one contig, and as Bacteria for 161 contigs (Figure 1.10). This result is consistent with the evolutionary reasoning that AMR genes are most advantageous when retained in bacteria (45). Whereas AMP-containing contigs were most commonly from Other Taxa, finding 336 contigs originating from Other Taxa, 177 from Eukaryota, and 173 from Bacterial contigs (Figure 1.10). Archaea contigs did not contribute to any contigs identified as AMR or AMP-containing (Figure 1.10). This result is also consistent with previous findings that suggest AMPs are produced from a variety of taxonomic groups (44).

Pearson Correlation

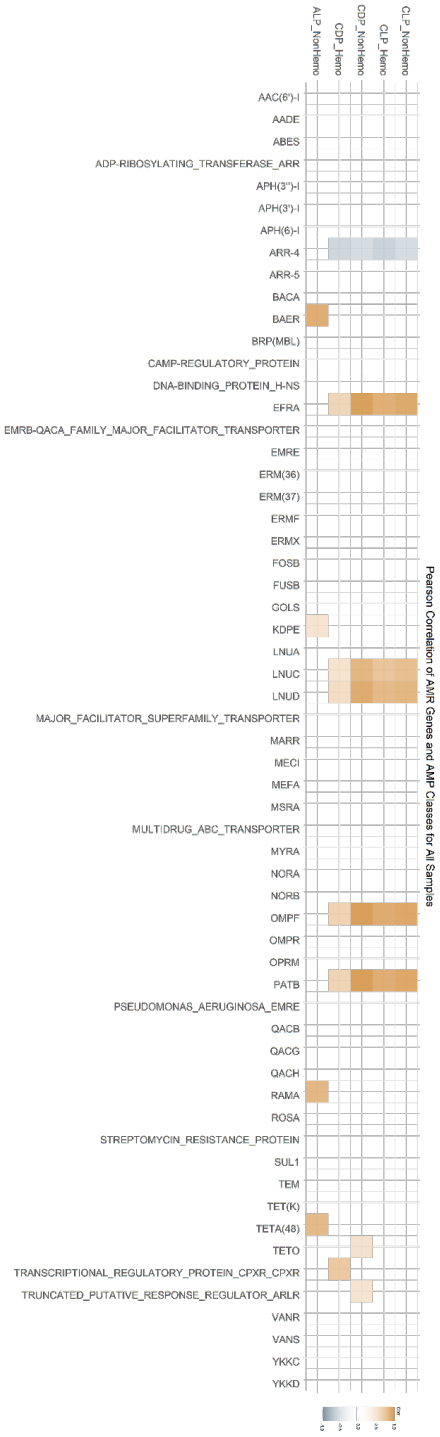


Figure 1.11: Pearson correlation between AMR genes (x-axis) and AMP classes (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did

not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

ARR-4, an integron-encoded ribosyltransferase found in *Pseudomonas aeruginosa* (46), was the only ARG that was negatively correlated with any AMP gene, and interestingly was similarly negatively correlated with all classes of AMP genes (Figure 1.11). This correlation may indicate an ecological interaction between ARR-4 and AMPs that drives this significantly negative correlation and should be investigated further.

AMR genes positively correlated with all Cationic classes of AMP genes including EFRA (efflux pump component), LNUC (transposon-mediated nucleotidyltransferase involved in antibiotic inactivation), LNUD (plasmid-mediated nucleotidyltransferase, involved in antibiotic inactivation), OMPF (porin), and PATB (efflux pump transporter component) (47–51) (Figure 1.11). ERFA and PATB are both individual components of two-part ABC efflux pumps where both components are required for functionality, so it is interesting to see a statistically significant association between each of these AMR genes and cationic AMP, but not their efflux pump counterparts.

While ALP AMPs are infrequent in comparison to cationic AMPs, there still is interest in understanding associated correlations for potential therapeutic purposes (52). AMR genes positively associated with Anionic AMP classes are BAER (response regulator of efflux pump), KDPE (transcriptional activator involved in pathogenic virulence), RAMA (regulator leading to high-level multidrug resistance), and TETA(48) (tetracycline efflux pump) (53–56).

No AMR genes had significant correlations with all classes of AMP genes.

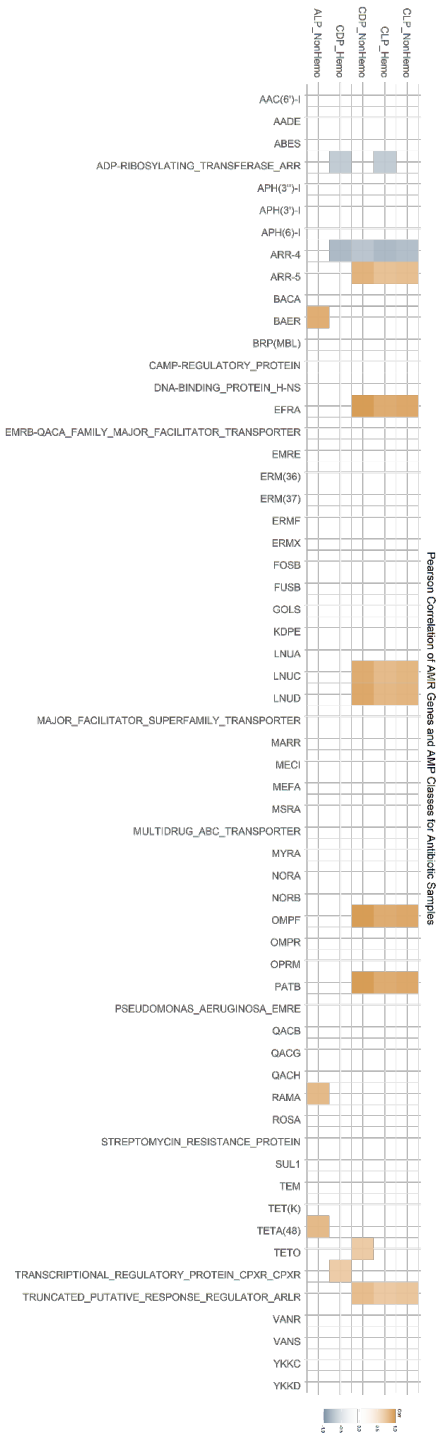


Figure 1.12: Pearson correlation between AMR genes (x-axis) and AMP classes (y-axis) for samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

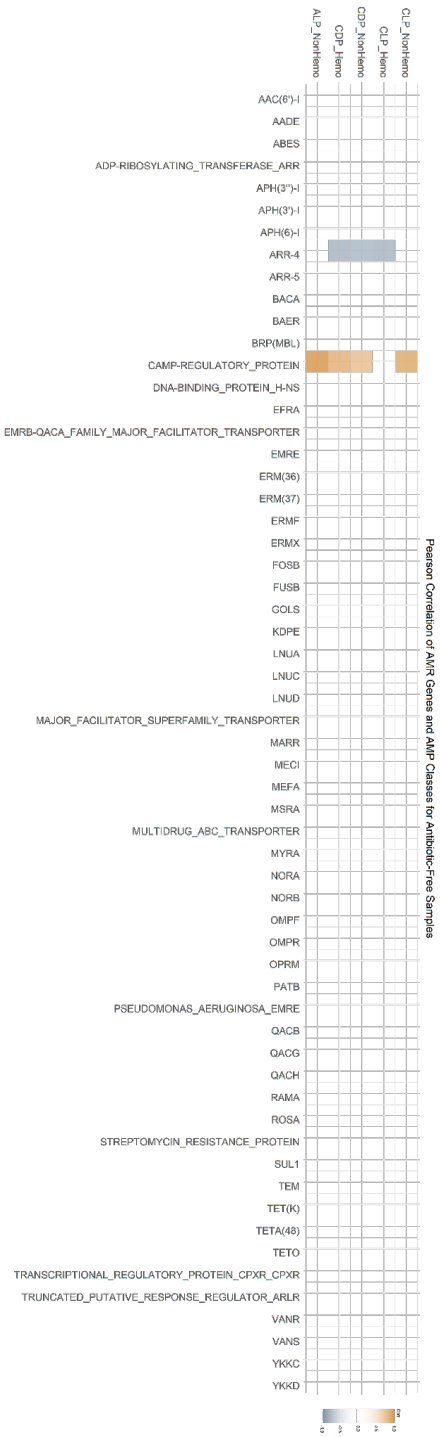


Figure 1.13: Pearson correlation between AMR genes (x-axis) and AMP classes (y-axis) for samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

More positive correlations are observed between AMR genes and AMP classes in the Antibiotic Samples (Figure 1.12), with an overall similar composition to the overall trends shown in Figure 1.11. Antibiotic samples also exhibited an additional set of negative correlations between the cationic hemolytic AMP classes and ADP-RIBOSYLATING_TRANSFERASE_ARR (ADP-ribosylation protein posttranslational modification) (57).

However, Antibiotic-Free samples only show two patterns of correlation, one of which is not also exhibited in the Antibiotic samples (Figure 1.13). Positive correlations were seen between the CAMP-REGULATORY_PROTEIN (global transcriptional regulator) and all AMP classes except CLP_Hemo (58). cAMP receptor protein (CRP) has been demonstrated to regulate over 490 genes in *E. coli*, which could implicate that these regulatory niches are being filled with other AMR genes in the Antibiotic samples, perhaps driven by the selective pressures increasing the diversity and abundance of AMR genes (59).

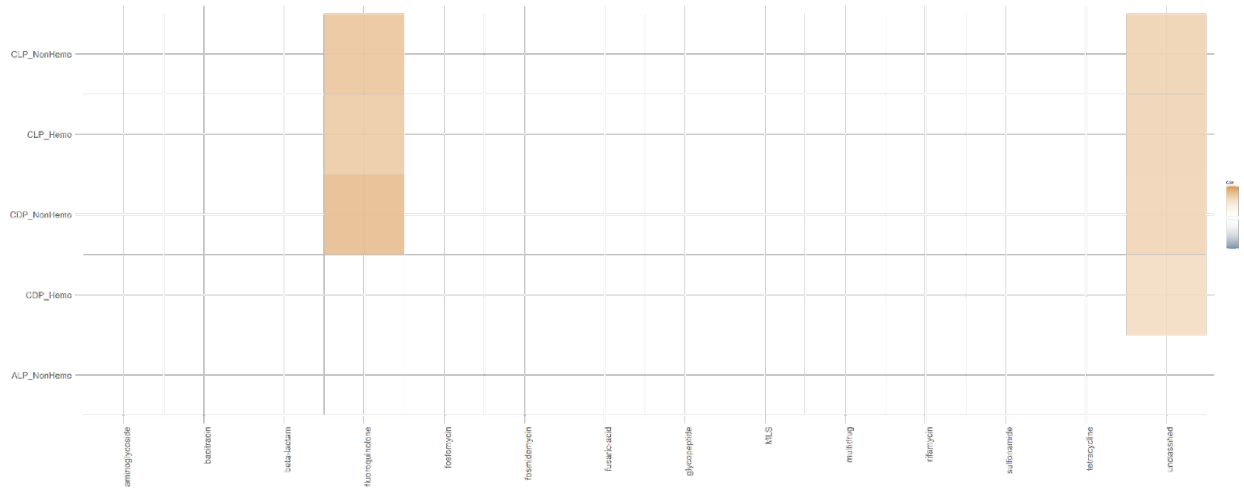


Figure 1.14: Pearson correlation between AMR gene classes (x-axis) and AMP classes (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames

of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Fluoroquinolone was the only AMR gene class, aside from unclassified AMR genes, that was statistically correlated with any AMP classes (Figure 1.14). Only cationic AMP classes were significantly associated with any AMR gene class, anionic AMP classes were not significantly associated with any AMR classes (Figure 1.14).

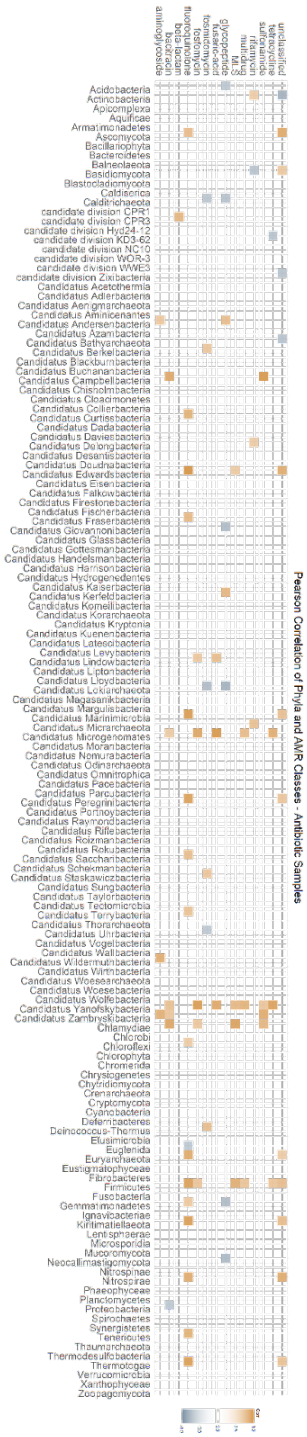


Figure 1.15: Pearson correlation between microbial phyla (y-axis) and AMR classes (x-axis) for samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

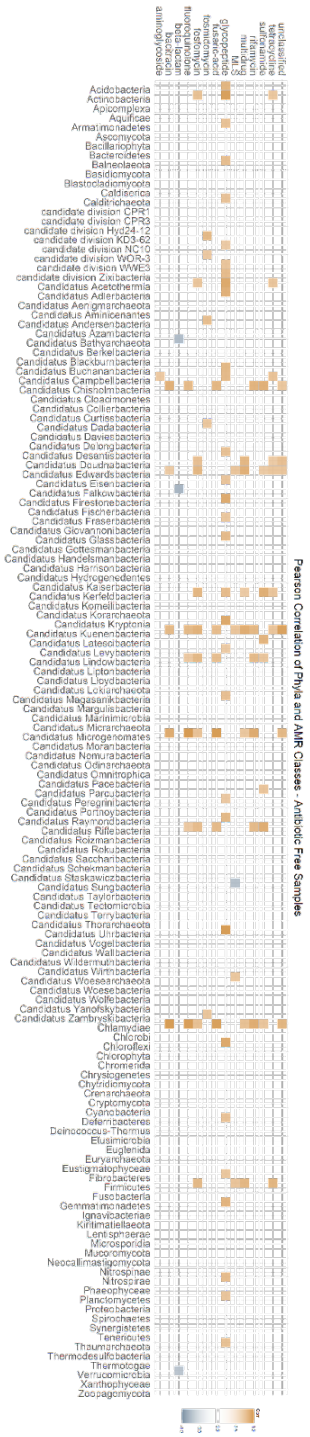


Figure 1.16: Pearson correlation between microbial phyla (y-axis) and AMR classes (x-axis) for samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Pearson correlations between AMR gene classes and phyla revealed that in antibiotic samples, negative correlations were only observed in unclassified, rifamycin, fosmidomycin, bacitracin, and glycopeptide (Figure 1.15). Glycopeptide had the largest number of significant associations, being significantly associated with 6 phyla: Acidobacteria, Calditrichaeota, Candidatus Giovannonibacteria, Candidatus Lokiarchaeota, Gemmatimonadetes, and Neocallimastigomycota. In Antibiotic-Free samples, negative correlations were only associated with MLS and beta-lactam AMR genes, where beta-lactam had the highest frequency of negative correlations, being associated with three phyla: Verrucomicobia, Candidatus Eisenbacteria, Candidatus Azambacteria (Figure 1.16).

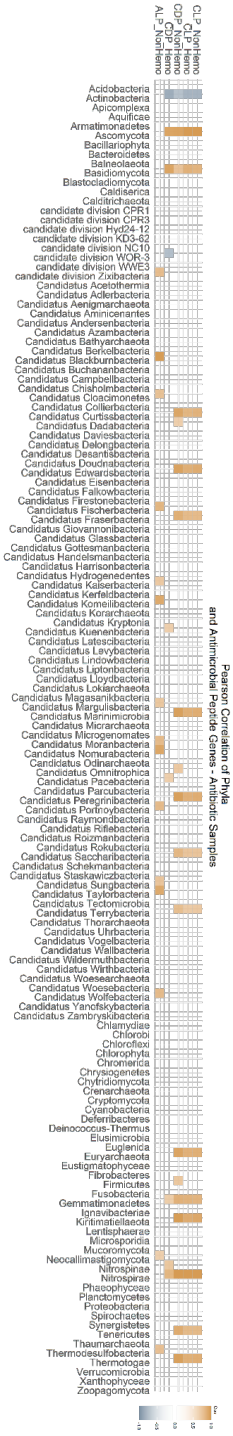


Figure 1.17: Pearson correlation between microbial phyla (y-axis) and AMP classes (x-axis) for samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

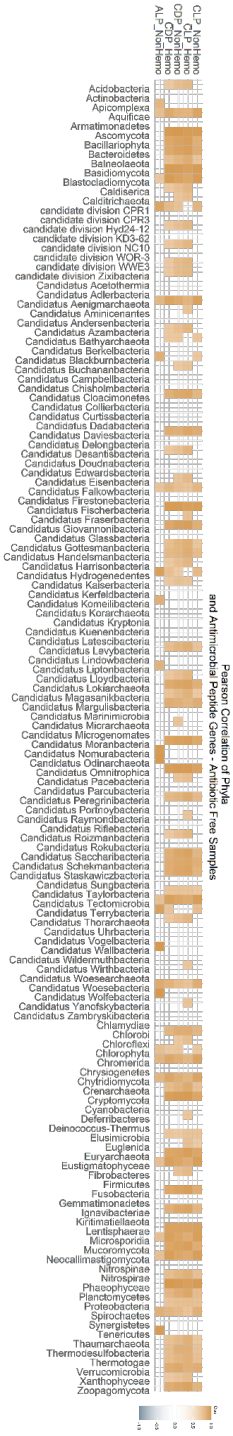


Figure 1.18: Pearson correlation between microbial phyla (y-axis) and AMP classes (x-axis) for samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Acidobacteria was the only phylum that was negatively correlated with multiple AMP classes in the Antibiotic samples, showing negative correlations with all cationic classes, but not anionic classes (Figure 1.17). Members of the Acidobacteria phyla are common in built environments, as they generally thrive in moderately acidic environments and often possess genetic features that enable a competitive lifestyle in scarce ecological niches, like soil and built environments (60). However, to my knowledge, there are no published ecological studies presenting such a stark negative correlation between Acidobacteria and cationic AMPs. This presents an interesting future course of study to determine the molecular basis of the ecological interactions responsible for these correlations.

In contrast, the Antibiotic-Free samples had a much higher instance of positive correlations at higher correlation strengths between Phyla and AMP classes, across both cationic and anionic AMPs (Figure 1.18). This association may suggest that built environment microbiomes of homes that administered antibiotics to their BYP exhibit similar living conditions, which may be related to the administration of antibiotics to BYP, which discourages the tight association between AMP classes and specific phyla. It has been shown that built environment microbiomes are influenced by a myriad of factors. Therefore, there are likely factors beyond the scope of this study influencing this microbial community in addition to the described factors (17).

Limitations

Community science surveys are often subjected to bias and inconsistencies (61). For this study, the exact antibiotic dosage, frequency, and administration methods were not adequately captured in a way that would allow conclusions to be drawn surrounding the impacts of antibiotic usage on the microbial community. However, there are important implications in the differences in microbial communities between the Antibiotic and the Antibiotic Free groups as a proxy for differences in lifestyle and the animal husbandry trends that are found within both groups.

Functional gene analyses were not subjected to any limit of detection threshold filtering, normally implemented to preserve only levels of detection likely to be true positives (62), in the interest of preserving any ecological interactions that, while sparse, may have been of interest.

Conclusion

Analysis of bacterial taxonomic and functional distributions among built-environment samples of BYP-owning homes showed that sample location and antibiotic administration to BYP flocks did not significantly impact microbial community diversity within samples, but antibiotic administration was significantly associated with microbial community diversity between samples. AMR- and AMP-encoding genes are ubiquitous across sample locations and antibiotic usage and were distributed across a wide taxonomic range of hosts. Pearson correlation analysis revealed a variety of significant positive and negative correlations between functional and taxonomic groups of interest that provide an abundance of future research investigation paths to pursue in the continuation of research in the prevention and mitigation of AMR bacteria in BYP environments.

Chapter 2: Investigation of Ecological Relationships between Bacteriophage-Encoded Functional Genes and Antimicrobial Resistance Genes in Backyard Poultry-Associated Built Environments Elucidates Potential Targets for Development of Bacteriophage-Derived Antimicrobials

Introduction

Bacteriophage (phage) are known to influence the community dynamics of their bacterial hosts, including the modulation of taxonomic and functional composition within the community (63–68). Phage have been investigated as an alternative strategy to antibiotics as a way to circumvent the evolutionary pressures that exacerbate AMR proliferation (69–72). Applications of phage-derived bacterial control strategies have been demonstrated in medicine, agriculture, and biotechnology, showcasing the inherent value and vast potential for understanding the ecological and evolutionary relationships between phage and microbial communities in their native environments (73–75).

There have been conflicting reports on whether phage directly and significantly contribute to the dissemination of AMR and other virulence genes (65, 69). However, it is undeniable that phage have the means and the mechanisms required for the modulation of bacterial communities. These ecological implications suggest potential previously uncharacterized relationships between phage features and AMR genes or classes that are not directly related to phage dissemination that may be advantageous in the development of phage-derived therapeutics for antibiotic-resistant bacterial infections. Phage-encoded lysin genes, encoding lysin proteins that facilitate the lysis of bacterial hosts, are of particular interest for understanding their relationships to AMR

genes in microbial communities, as they are already being utilized as a promising antibiotic alternative (74, 76–81)

Findings from Chapter 1 demonstrated that AMR gene presence is ubiquitous in the BYP samples analyzed, encompassing a wide range of AMR molecular modes of action. The breadth of resistance mechanisms present in this environment provides an ideal system to investigate correlations between AMR gene presence and bacteriophage features. This analysis aims to understand the significance of relationships between bacteriophage and AMR in BYP environments with the goal of elucidating a foundational understanding of potential bacteriophage-derived targets for biotechnology applications of prevention and mitigation of AMR.

Materials and Methods

Sample Collection and Sequencing

Sequence data for this analysis is a continuation of the analysis described in Chapter 1, for detailed information on sample collection and sequencing refer to Materials and Methods described in Chapter 1.

Functional and Taxonomic Annotation

Bacteriophage contigs were identified using VirSorter (82) and VIBRANT (83), and additional phage contigs were identified from metagenomic co-assemblies performed using Kaiju, as described in Chapter 1. Contigs were denoted as phage-originating if any of the following taxa were within the assigned Family or Order taxonomic ranks: "Caudovirales", "Myoviridae", "Siphoviridae", "Podoviridae", "Microviridae", "Corticoviridae", "Tectiviridae", "Leviviridae", "Cystoviridae", "Inoviridae", "Plasmaviridae".

Total Pfam annotation of genes within bacteriophage contigs identified by VIBRANT were also described, and potential auxiliary metabolic genes were further subset from these gene annotations (84). Genes classified as Auxillary Metabolic Genes (AMG) were defined within the VIBRANT analysis pipeline, which they define as metabolic genes that phage will “steal” from their hosts for expression during active infections for fitness advantages (83). Lysin orthologous groups were identified by searching the annotations of the VIBRANT co-assembly for “lysin” within the “VOG name” variable, and any contig with a matching identification was classified as a Lysin Viral Orthologous Group (LVOG).

Taxonomic classification of bacteriophage with the existing bioinformatic tools is a tedious process that lacks the foundational reference database architecture needed for accurate and automated taxonomic assignments comparable to bacterial taxonomy analyses (68, 85, 86). Since phage lack a universal marker gene and have high mutation rates, accurate bioinformatic taxonomic classification of phage from metagenomic microbial community sequencing is inherently difficult and error-prone. Therefore, putative hosts of the identified phage were used as proxies for diversity metrics in this analysis. All contigs taxonomically assigned as phage, through Kaiju, VirSorter, and VIBRANT, were used as input to predict putative bacterial host pairings using VirHostMatcher-Net (87). Putative viral-host interactions are calculated using virus-virus similarity, virus-host alignment-free similarity using k-mer comparisons, virus-host alignment-based matches, and virus-host shared CRISPR spacers. The combination of these methods allows the classification of these interactions at higher confidence than any of these methods alone (87). Relative abundances of the phage contigs grouped by their respective putative host taxa were used as proxies to evaluate the alpha and beta diversity of the phage population within the samples.

AMP and AMR genes were identified using the same methods described in the Materials and Methods section of Chapter 1.

Statistical analysis and Visualization

Statistical analysis and visualization were performed using the same methods described in the Materials and Methods section of Chapter 1. All scripts can be found in the GitHub repository at https://github.com/alonnawright/backyardpoultry_metagenomic_analysis. All metadata and high-quality figure images can be found in the figshare repository for this project (https://doi.org/10.6084/m9.figshare.21585807) (37).

Results and Discussion

Alpha Diversity

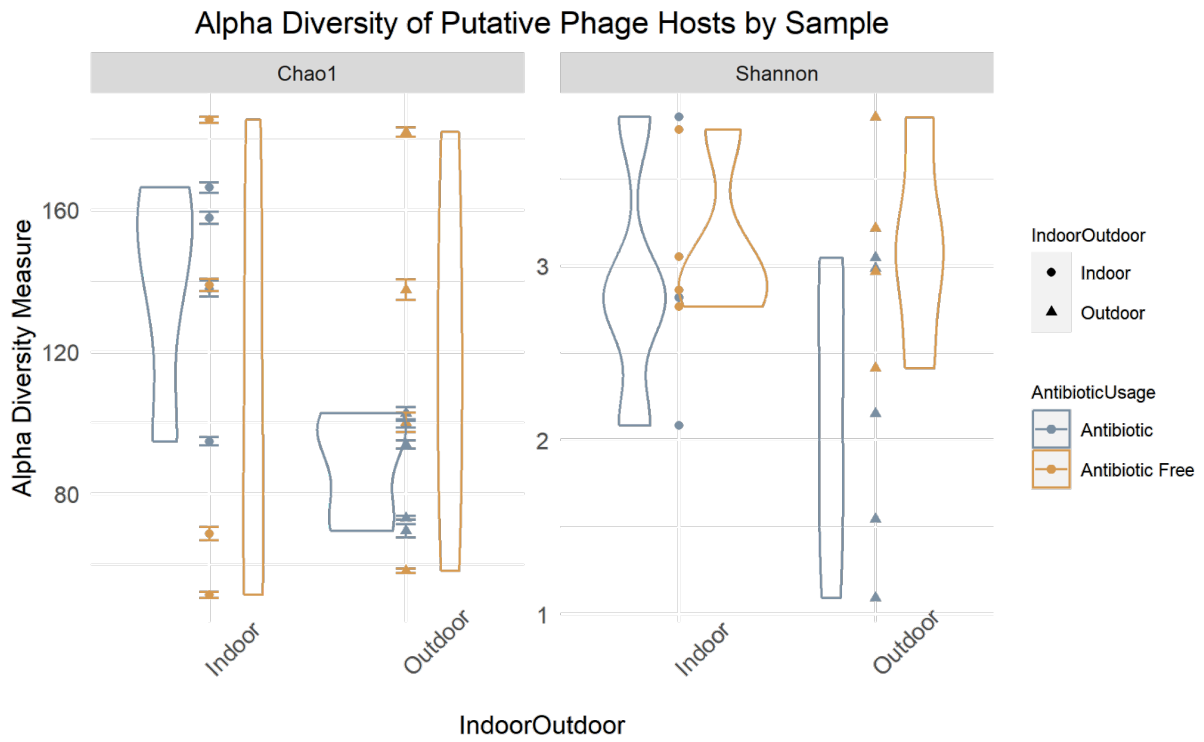


Figure 2.1: Alpha Diversity of Phage Microbial Communities Did Not Differ Between Sample Groups. Chao1 and Shannon alpha diversity metrics of phage microbial communities of internal (indoor) and external (outdoor) door frames of BYP Owning Homes, significance determined by

Kruskal-Wallis rank sum test at $p < 0.05$. Violin plots display mirrored density of continuous distributions

Quantifying the alpha (within a sample) and beta (between samples) diversities can help contextualize the members of the community and the breadth of potential interactions which may influence functional characteristics in a microbial community, which especially relevant when examining the phage community since the potential for AMR retention and dissemination via phage is of particular interest (40). Alpha diversity captures the diversity of a community within a single sample, which is a helpful metric in understanding the potential ecological interactions within a system (41). Shannon and Chao1 alpha diversity metrics, which measure abundance and richness respectively, were evaluated for each BYP-owning home sample. Shannon Alpha Diversity of putative phage hosts was not significantly different between Antibiotic and Antibiotic Free groups (Kruskal-Wallis rank sum test, $p = 0.1237$), nor between Indoor and Outdoor groups (Kruskal-Wallis rank sum test, $p = 0.5637$). Similarly, Chao1 Alpha Diversity was not significantly different between Antibiotic and Antibiotic Free groups (Kruskal-Wallis rank sum test, $p = 1$), nor between Indoor and Outdoor groups (Kruskal-Wallis rank sum test, $p = 0.3359$).

Beta Diversity

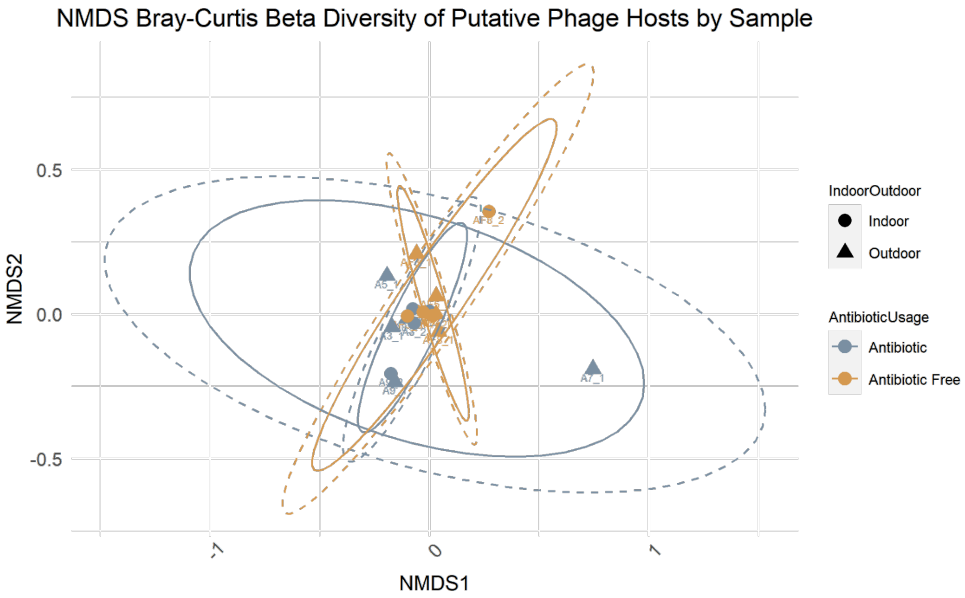


Figure 2.2: Non-metric Multidimensional Scaling (NMDS) of Bray-Curtis Beta Diversity Metrics of putative phage hosts of internal (indoor) and external (outdoor) door frames of BYP-owning homes, significance determined by PERMANOVA significance test at $p < 0.05$. Solid-line ellipses represent the multivariate t-distribution, while dotted-line ellipses represent the multivariate normal distribution. NMDS1 and NMDS2 axes represent the two most representative metrics of the calculated distances in multidimensional space.

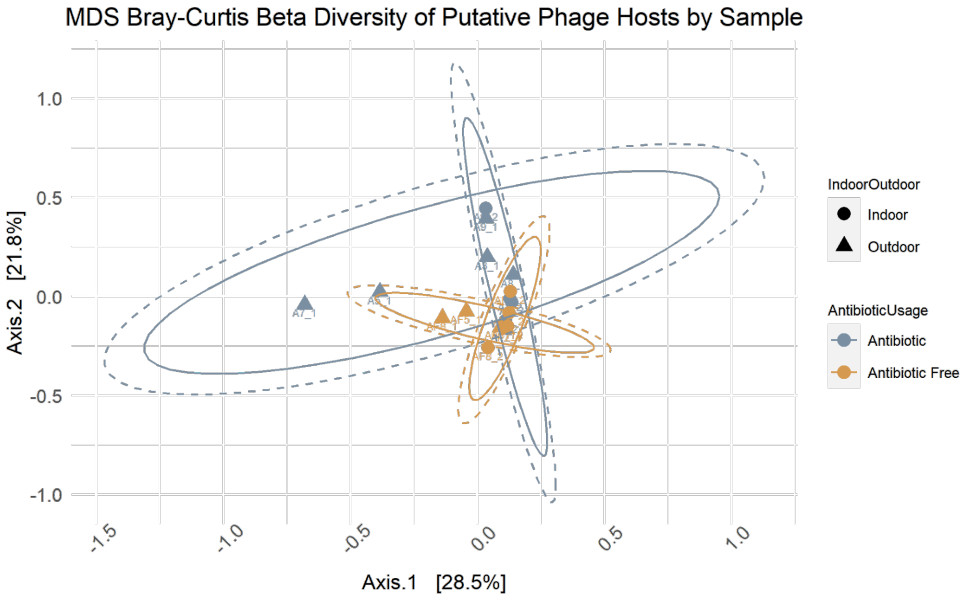


Figure 2.3: Multidimensional Scaling (MDS) of Bray-Curtis beta diversity metrics of putative phage hosts of internal (indoor) and external (outdoor) door frames of BYP-owning homes, significance determined by PERMANOVA significance test at $p < 0.05$. Solid-line ellipses represent the multivariate t-distribution, while dotted-line ellipses represent the multivariate normal distribution. Axes represent the two most explanatory eigenvalues, percentages represent the percent of variation attributed to each axis.

Bray Curtis distances were not significantly different between Antibiotic and Antibiotic Free samples (Figure 2.2, 2.3; PERMANOVA, adonis2, p=0.118), in comparison to beta diversity metrics analyzed in Chapter 1 that did show significant differences between these groups (Figures 1.2, 1.3). However, similar to overall microbial community trends found in Chapter 1, these distances were not statistically significant between Indoor and Outdoor samples (PERMANOVA, adonis2, p=0.241) nor pairwise between the combinations of these two groups (PERMANOVA, adonis2, p=0.978). Dispersion of beta diversity distances was also not significantly different for Antibiotic and Antibiotic Free samples (Tukey multiple comparisons of means, p=0.185) nor between Indoor and Outdoor samples (Tukey multiple comparisons of means, p=0.068). Since the evaluated contigs are merely a subset of the originally analyzed microbial community

presented in Chapter 1, the overall similarity between the two results is expected but is still interesting to note that the significance between Antibiotic and Antibiotic Free groups is lost when subsetting to putative phage hosts.

Phylum Abundances

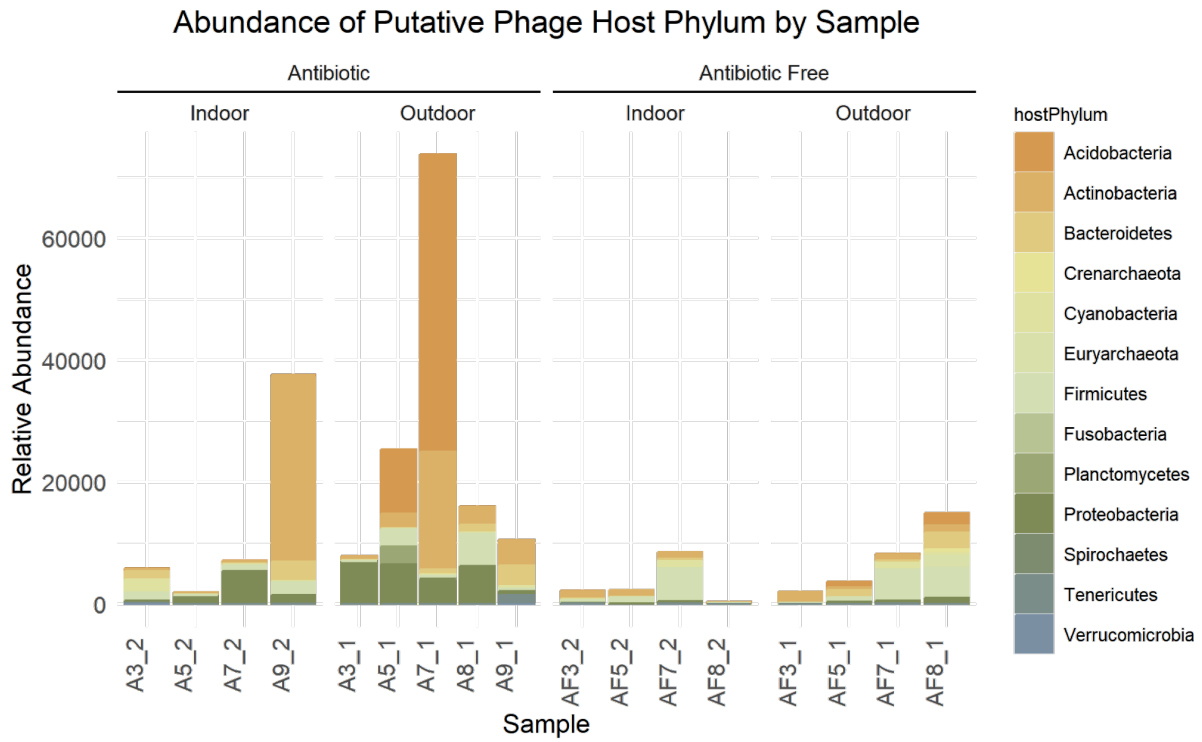


Figure 2.4: Relative abundance of putative phage host phylum for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

Outdoor samples from Antibiotic homes had higher overall relative abundances of any putative phage host, and notably higher relative abundances of Acidobacteria in A5_1 and A7_1 samples, as well as a higher relative abundance of Spirochaetes across the Outdoor Antibiotic

samples. Antibiotic-Free samples had overall much lower relative abundances of putative host phage contigs across most phyla.

Functional Gene Relative Abundances

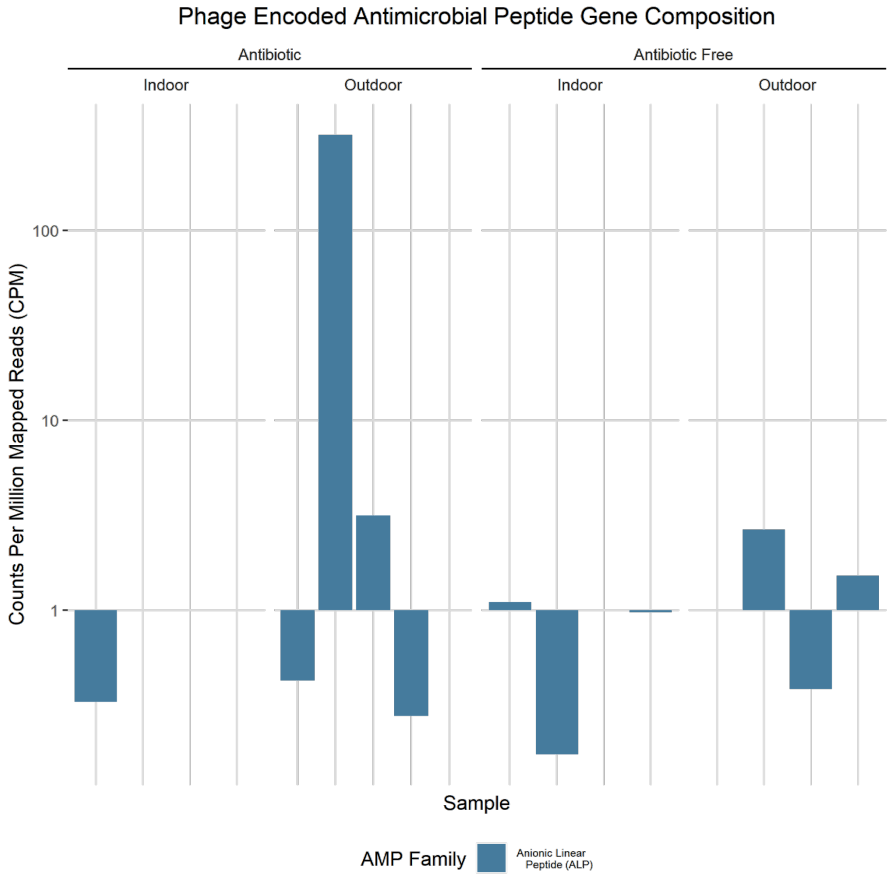


Figure 2.5: Frequency of AMP-encoding Contigs in Count Per Million for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

Phage contigs did not contain any detected AMR genes, but AMP genes were detected in phage contigs of 11 of the 18 samples. Interestingly, phage-encoded AMPs were only of the Anionic Linear Peptide (ALP) class (Figure 2.5). ALP AMPs are much less common than their cationic counterparts, as seen in results from Chapter 1 and previously stated in the literature

(44). Notably, it has been shown that in certain instances the expression of a functional cationic AMP is necessary for a phage to successfully lyse its bacterial host (88). However, to my knowledge, there are no reports of anionic AMPs serving as similarly necessary accessories in host lysis. The exclusive presence of ALP AMPs in these phage contigs presents an interesting avenue for further research of the potential ecological role these AMP sequences play in phage microbial community modulation.

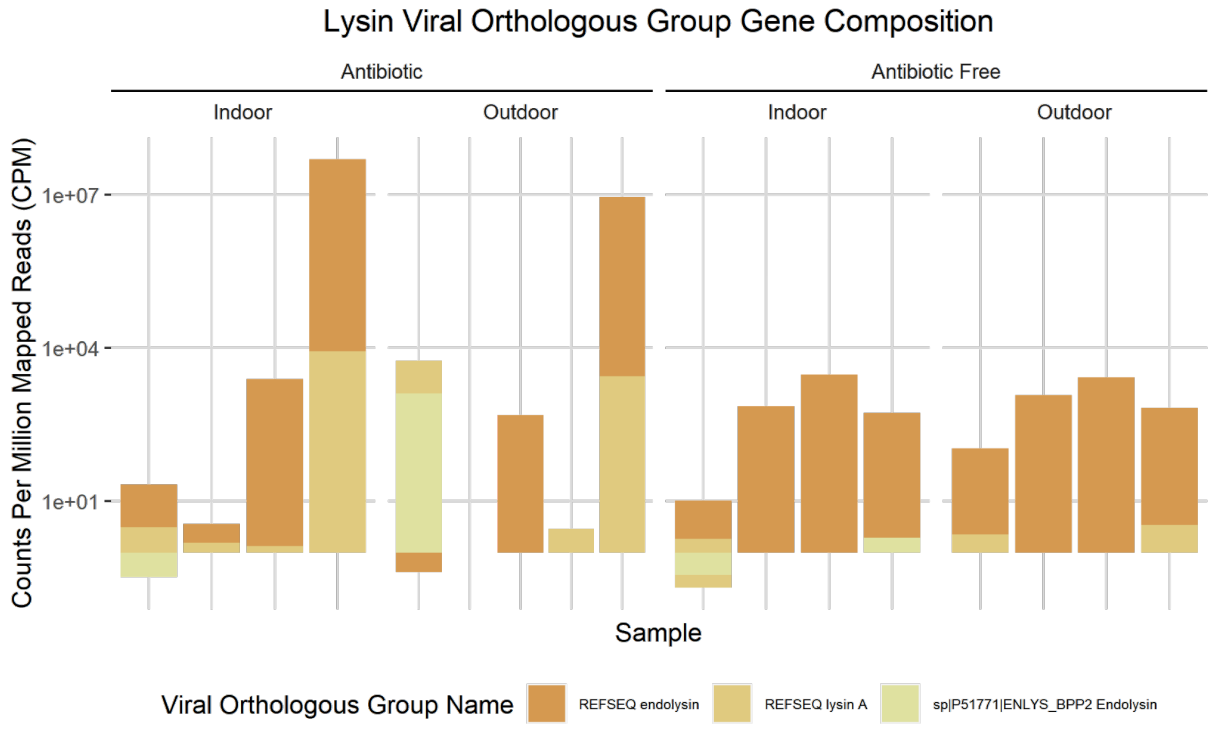


Figure 2.6: Relative abundance of Lysin Viral Orthologous Groups (VOGs) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

Three lysin VOGs were identified in the phage contigs (Figure 2.6). Two were broadly classified as “REFSEQ endolysin” and “REFSEQ lysin A”, while one had a more specific protein designation of “sp|P51771|ENLYS_BPP2 Endolysin”. This ENLYS_BPP2 Endolysin contains

transglycosylase activity to degrade host peptidoglycan for the purpose of releasing mature viral particles (89). REFSEQ endolysin was detected in nearly every sample, with the exception of one Outdoor Antibiotic sample, and was the most commonly detected lysin VOG across all groups. Antibiotic samples had a much larger variability in lysin relative abundance, representing both the highest and lowest relative abundance detection levels across all samples, while Antibiotic-Free samples contained more consistent levels of lysin VOGs, with the vast majority being classified as REFSEQ endolysin.

Pearson Correlations

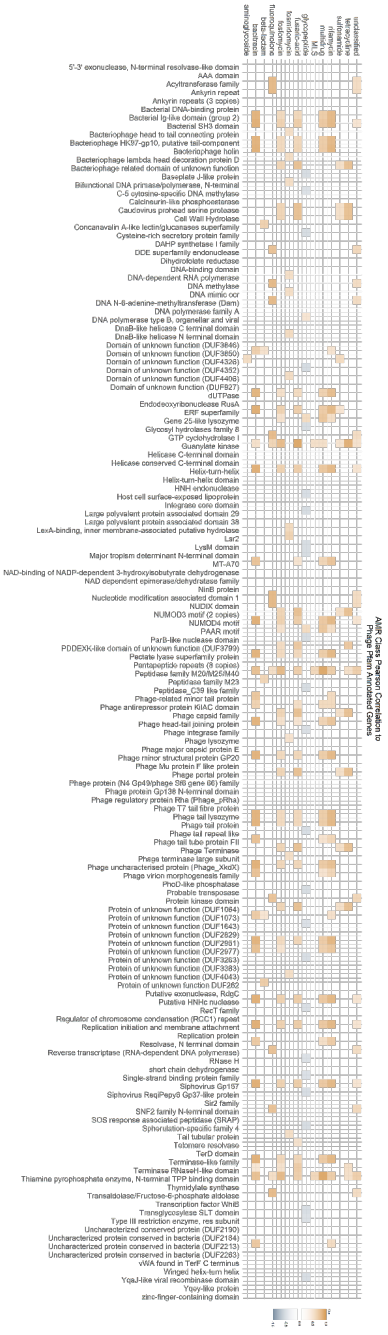


Figure 2.7: Pearson correlation between AMR class (x-axis) and phage features as annotated by the Pfam database (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Glycopeptide was the only AMR class that exhibited any negative correlation with Phage Pfam annotated features, and interestingly, exhibited negative correlations across a variety of the Phage Pfam annotated features (Figure 2.7). One hypothesis for this negative correlation trend is that the expression of some bacterial cell wall components may decrease in response to the selective pressure of phage who utilize these cell wall components as receptors. The decrease in the expression of the cell wall components may increase the susceptibility to antibiotics, such as glycopeptide antibiotics (70, 90–92).

Bacitracin, fosfomycin, fusaric-acid, multidrug, and rifamycin AMR classes all exhibited multiple positive correlations with Phage Pfam annotated features, often with strong correlation scores (Figure 2.7). These associations may suggest unknown underlying mechanisms that enable a synergistic effect between phage and these AMR classes that are not directly related to AMR gene dissemination via phage infections.

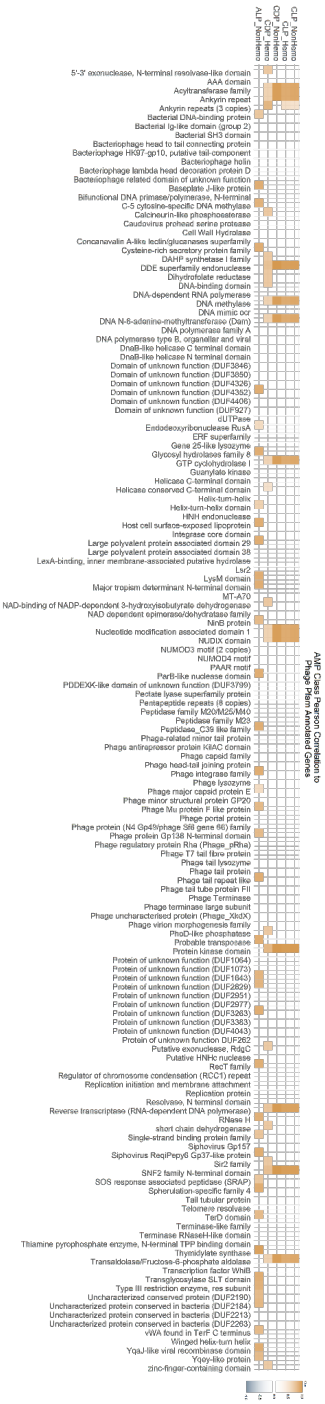


Figure 2.8: Pearson correlation between AMP class (x-axis) and phage features as annotated by the Pfam database (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Cationic AMPs showed similar patterns of significant correlations with Phage Pfam annotated features, however, anionic AMPs exhibited a higher abundance of correlations with Phage Pfam annotated features (Figure 2.8). This suggests that charge, more than hemolytic capacity or linearity, likely influences these correlations. All correlations between AMPs and Phage Pfam annotated features were positively correlated, with no significant negative correlations observed.

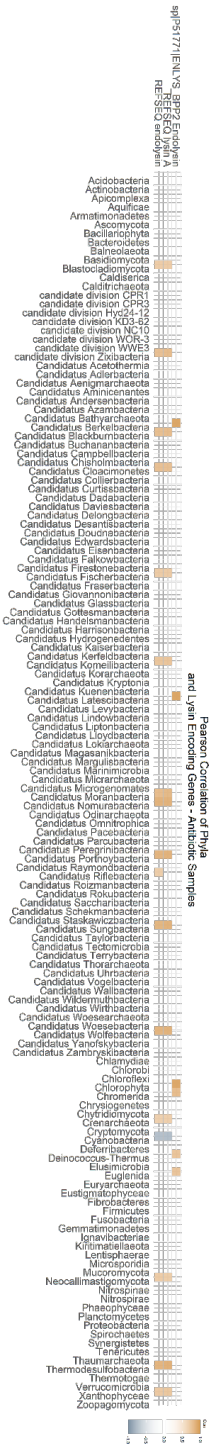


Figure 2.9: Pearson correlation between lysin viral orthologous groups (VOGs) (x-axis) and phyla of the total microbial community (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value

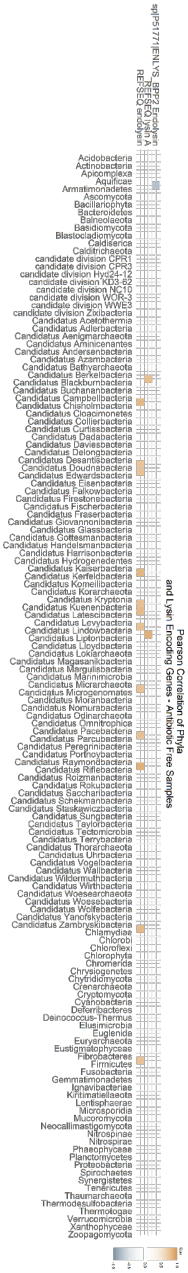


Figure 2.10: Pearson correlation between lysin viral orthologous groups (VOGs) (x-axis) and phyla of the total microbial community (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

In the Antibiotic Samples, Cyanobacteria was the only phyla that were significantly negatively correlated with any lysin VOG, and was negatively correlated to REFSEQ lysin A and REFSEQ endolysin (Figure 2.9). However, in the Antibiotic-Free samples Cyanobacteria had no significant positive or negative correlations (Figure 2.10). These two lysin VOGs exhibited clear similarities in correlated phyla. Further analysis of the amino acid sequence charge of these lysin VOGs in comparison with the Gram stain of associated phyla may reveal a molecular mechanistic explanation between the association of these putative lysins and these taxonomic groups.

Armalimonadetes was the only phyla in the Antibiotic-Free samples to exhibit a negative correlation, where it was significantly negatively correlated with sp|P51771|ENLYS_BPP2 Endolysin (Figure 2.10). This was the only significant correlation for this lysin VOG in the Antibiotic-Free samples.

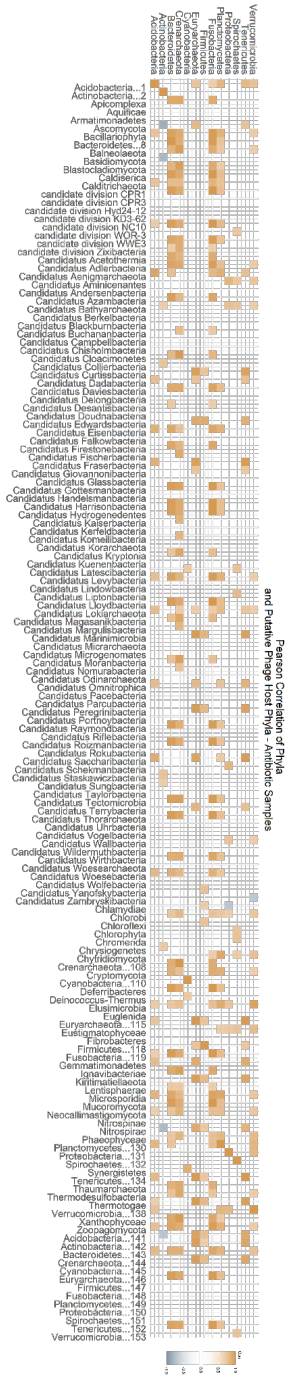


Figure 2.11: Pearson correlation between phage contigs as categorized by their putative hosts as a taxonomic proxy (x-axis) and phyla of the total microbial community (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

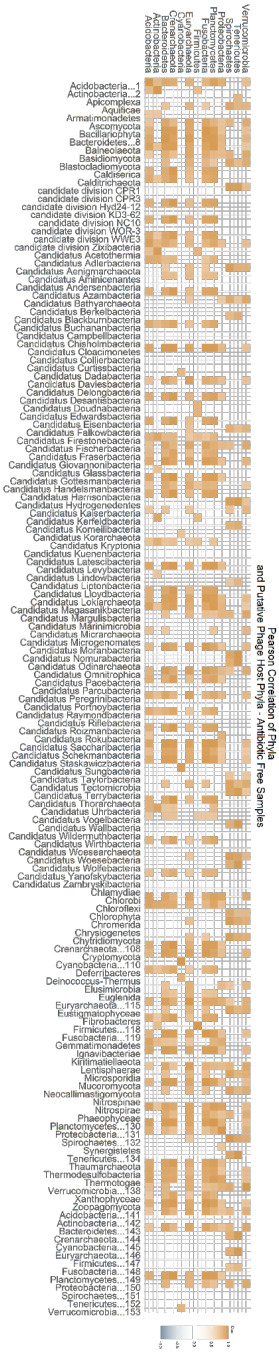


Figure 2.12: Pearson correlation between phage contigs as categorized by their putative hosts as a taxonomic proxy (x-axis) and phyla of the total microbial community (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Phage contigs associated with hosts in the Planctomycetes, Fusobacteria, Crenarchaeota, and Bacteroidetes phyla all showed similar patterns of positive correlations with phyla of all contigs in the Antibiotic samples (Figure 2.11). Similar patterns for these putative phage host phyla were exhibited in the Antibiotic Free samples, but additionally, Verrucomicrobia, Proteobacteria, Euryarchaeota, and Acidobacteria also displayed very similar positive correlation patterns (Figure 2.12). This loss of positive correlation between putative phage host phyla and overall phyla in the Antibiotic samples may be an indication of host-specific microbial community modulation facilitated through bacteria-phage host-prey dynamics (63, 64).

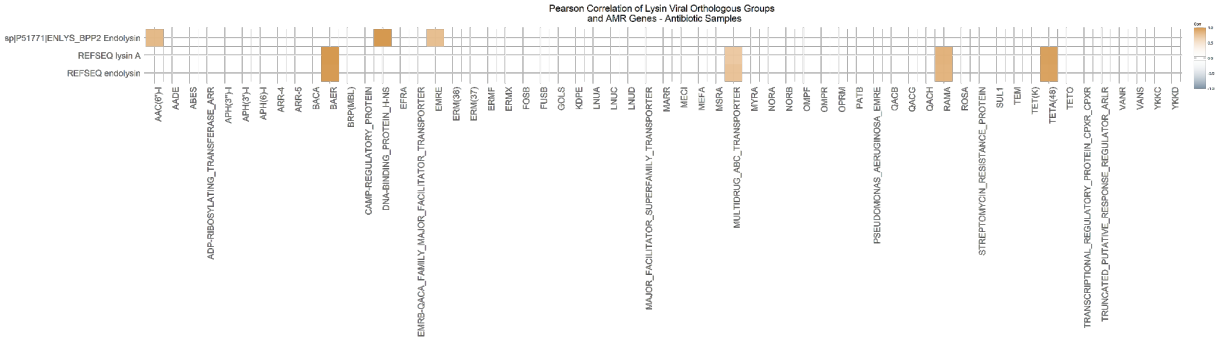


Figure 2.13: Pearson correlation between AMR genes (x-axis) and lysin viral orthologous groups (VOGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

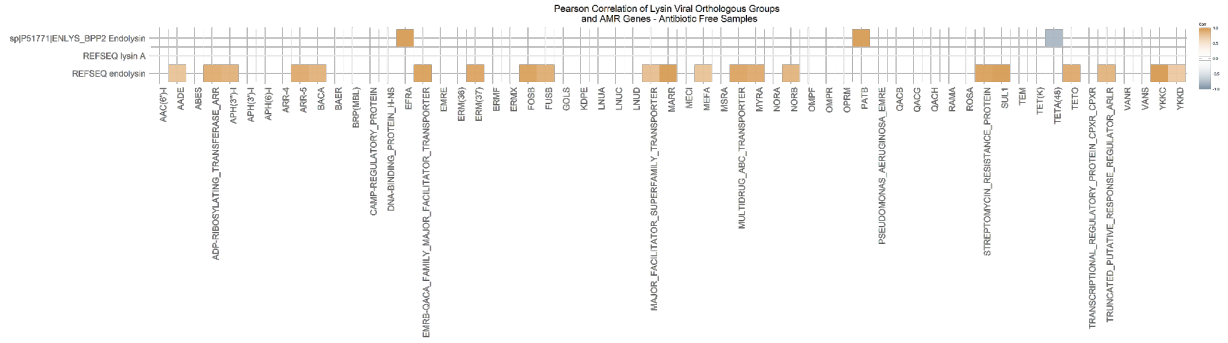


Figure 2.14: Pearson correlation between AMR genes (x-axis) and lysin viral orthologous groups (VOGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Significant correlations with any of the three lysin VOGs were only identified in 7 AMR genes in the Antibiotic Samples: AAC(6')-I (aminoglycoside acetyltransferase), BAER (response regulator of efflux pump), DNA-BINDING_PROTEIN_H-NS (DNA binding protein modulating RNA stability), EMRE (multidrug transporter), MULTIDRUG_ABC_TRANSPORTER (efflux pump), RAMA (regulator leading to high-level multidrug resistance), and TETA(48) (tetracycline efflux pump) (53, 55, 56, 93–96) (Figure 2.13). Notably, three of these AMR genes (BAER, RAMA, TETA(48)) were also positively correlated with ALP AMPs as shown in Figure 1.11 in Chapter 1.

Antibiotic-Free samples had a much higher instance of positive correlations, with the vast majority of correlations being with the REFSEQ endolysin VOG. There were no significant correlations with REFSEQ lysin A, and only three with sp|P51771|ENLYS_HPP2 Endolysin (two

positive, and one negative) (Figure 2.14). The only negative correlation exhibited in these samples was between the sp|P51771|ENLYS_HPP2 Endolysin VOG and TETA(48).

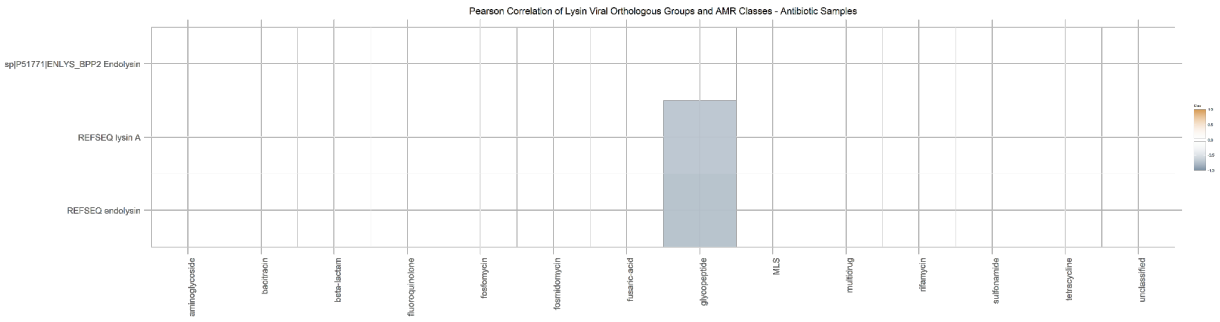


Figure 2.15: Pearson correlation between AMR gene classes (x-axis) and lysin viral orthologous groups (VOGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

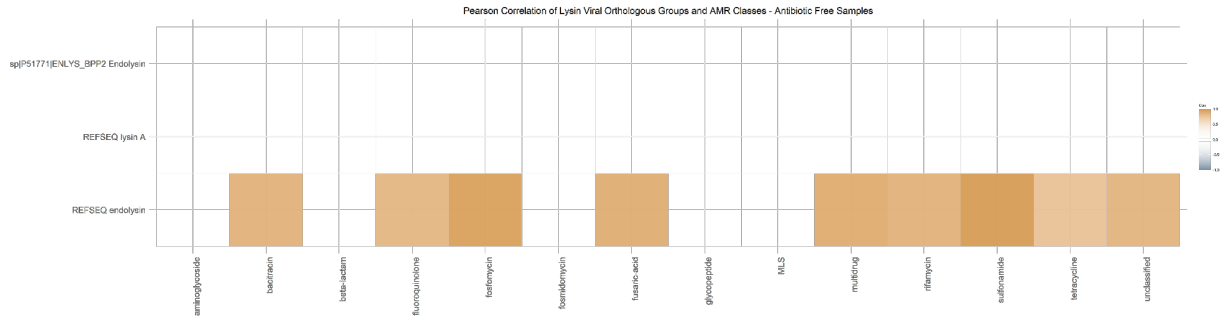


Figure 2.16: Pearson correlation between AMR gene classes (x-axis) and lysin viral orthologous groups (VOGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Pearson correlations between lysin VOGs and AMR classes in Antibiotic samples revealed only 2 significant correlations, glycopeptide AMR class genes were negatively correlated with REFSEQ lysin A and REFSEQ endolysin (Figure 2.15). Antibiotic-Free samples, however, showed no negative correlations (Figure 2.16). Still, REFSEQ endolysin was positively correlated with 9 or the 14 AMR classes, while no other lysin VOG exhibited any significant correlation for these samples (Figure 2.16). Speculatively, this may indicate that selective pressures imposed on Antibiotic samples were selective against bacterial and phage communities involving the REFSEQ endolysin gene.

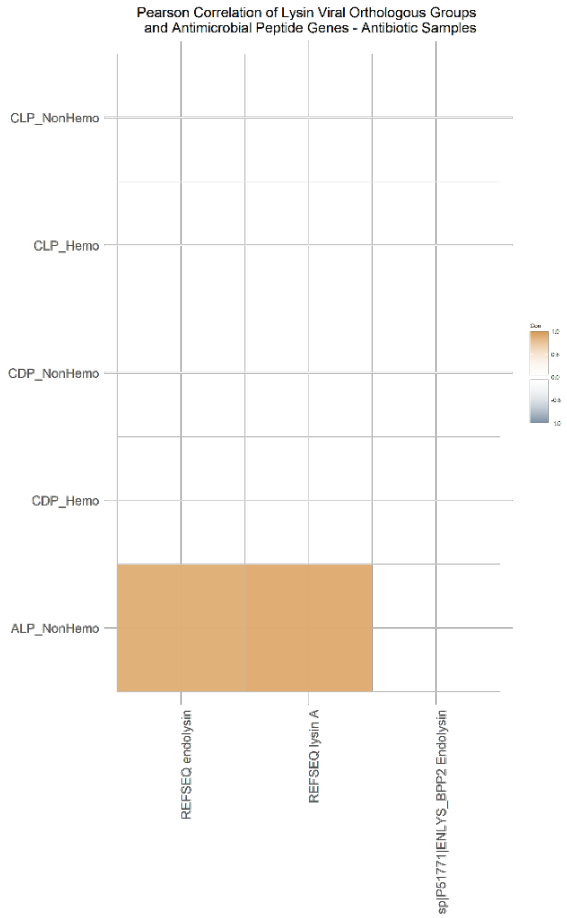


Figure 2.17: Pearson correlation between lysin viral orthologous groups (VOGs) (x-axis) and AMP classes (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Antibiotic samples showed positive correlations between the ALP AMP class and REFSEQ lysin A and REFSEQ endolysin (Figure 2.17), while Antibiotic-Free samples showed no correlations between any lysin VOG and AMP class (not shown).

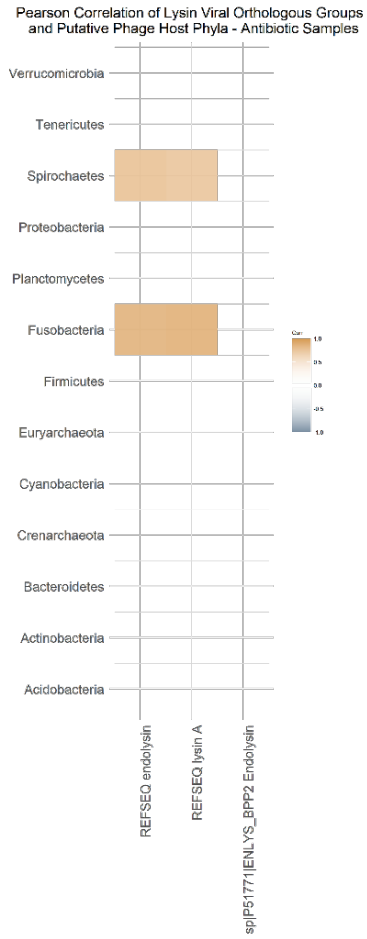


Figure 2.18: Pearson correlation between lysin viral orthologous groups (VOGs) (x-axis) and phage contigs as categorized by their putative hosts as a taxonomic proxy (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

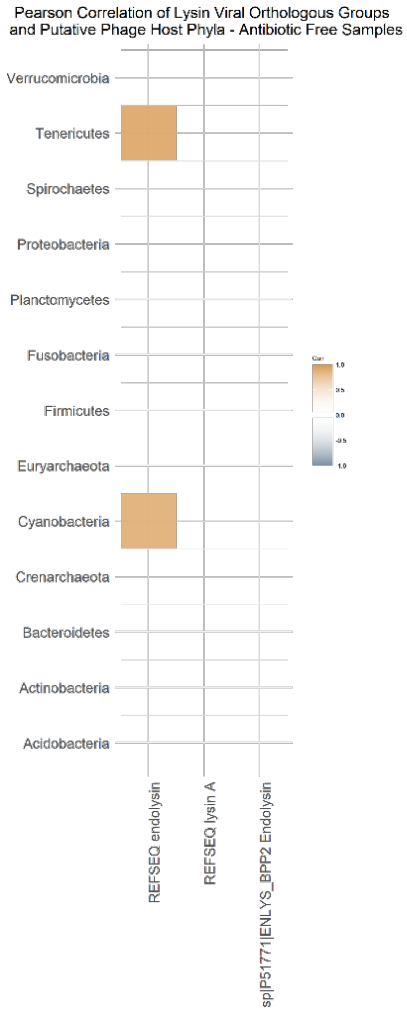


Figure 2.19: Pearson correlation between lysin viral orthologous groups (VOGs) (x-axis) and phage contigs as categorized by their putative hosts as a taxonomic proxy (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Antibiotic and Antibiotic-Free samples both showed significant correlations with two phyla of putative phage hosts, however, neither set of samples shared a phyla of significance (Figures 2.18, 2.19). Antibiotic samples exhibited associations between REFSEQ endolysin and REFSEQ

lysin A and Spirochaetes and Fusobacteria (Figure 2.18), while Antibiotic-Free samples showed correlations between REFSEQ endolysin with Tenericutes and Cyanobacteria (Figure 2.19).

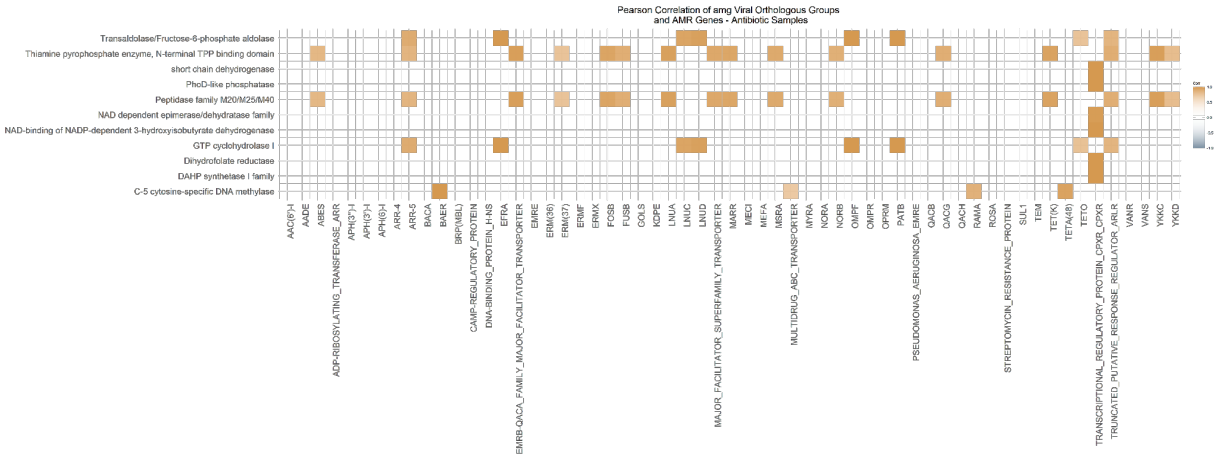


Figure 2.20: Pearson correlation between AMR genes (x-axis) and phage-encoded auxiliary metabolic genes (AMGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

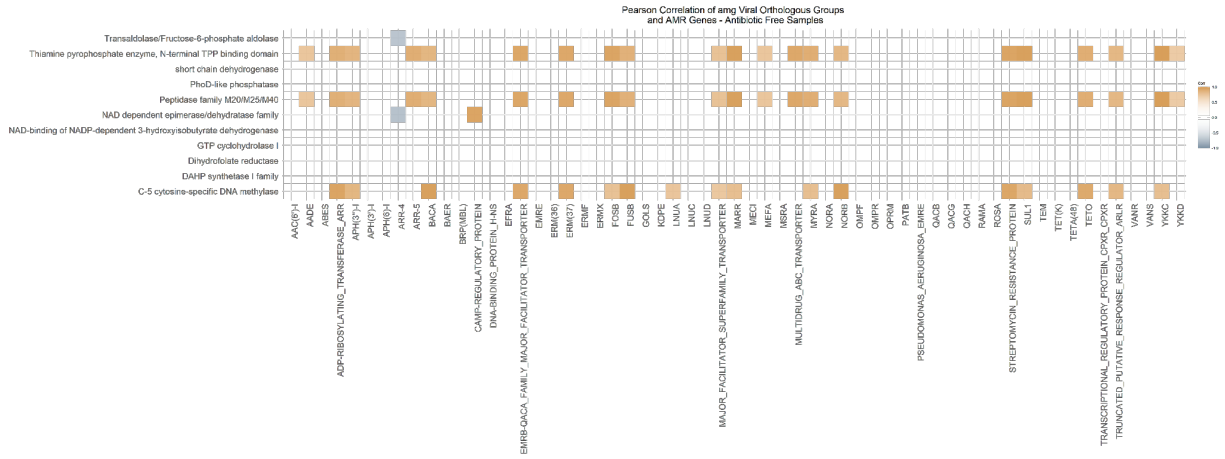


Figure 2.21: Pearson correlation between AMR genes (x-axis) and phage-encoded auxiliary metabolic genes (AMGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

In Antibiotic samples, C-5 cytosine-specific DNA methylase showed correlations with AMR genes BAER, MULTLIDRUG_ABC_TRANSPORTER, RAMA, and TETA(48), the same set of AMR genes seen in other Pearson correlations discussed in previous figures (Figure 2.20).

Thiamine pyrophosphate enzyme, N-terminal TPP binding domain and Peptidase family M20/M35/M40 exhibited almost identical Pearson correlation trends between Antibiotic and Antibiotic-Free samples (Figures 2.20, 2.21), with C-5 cytosine-specific DNA methylase exhibiting similar patterns of correlation in the Antibiotic-Free samples (Figure 2.21). None of the AMR genes associated with C-5 cytosine-specific DNA methylase in Antibiotic samples was also correlated with this gene in the Antibiotic-Free samples (Figures 2.20, 2.21).

ARR-4 exhibited significant negative correlations in Antibiotic-Free samples with NAD dependent epimerase/dehydralase family and Transaldolase/Fructose-6-phosphate aldolase (Figure 2.21). TRANSCRIPTIONAL_REGULATROY_PROTEIN_CPXR_CPXR exhibited strong positive correlations with 6 AMGS in Antibiotic Samples (DAHP synthelaste I family, Dihydrofolate reductase, NAD-binding of NADP-dependent 3-hydroxyisobutyrate dehydrogenase, NAD dependent epimerase/dehydralase family, PhoD-like phosphatase, short chain dehydrogenase) (Figure 2.20). However, TRANSCRIPTIONAL_REGULATROY_PROTEIN_CPXR_CPXR did not exhibit any significant correlations in Antibiotic-Free samples (Figure 2.21).

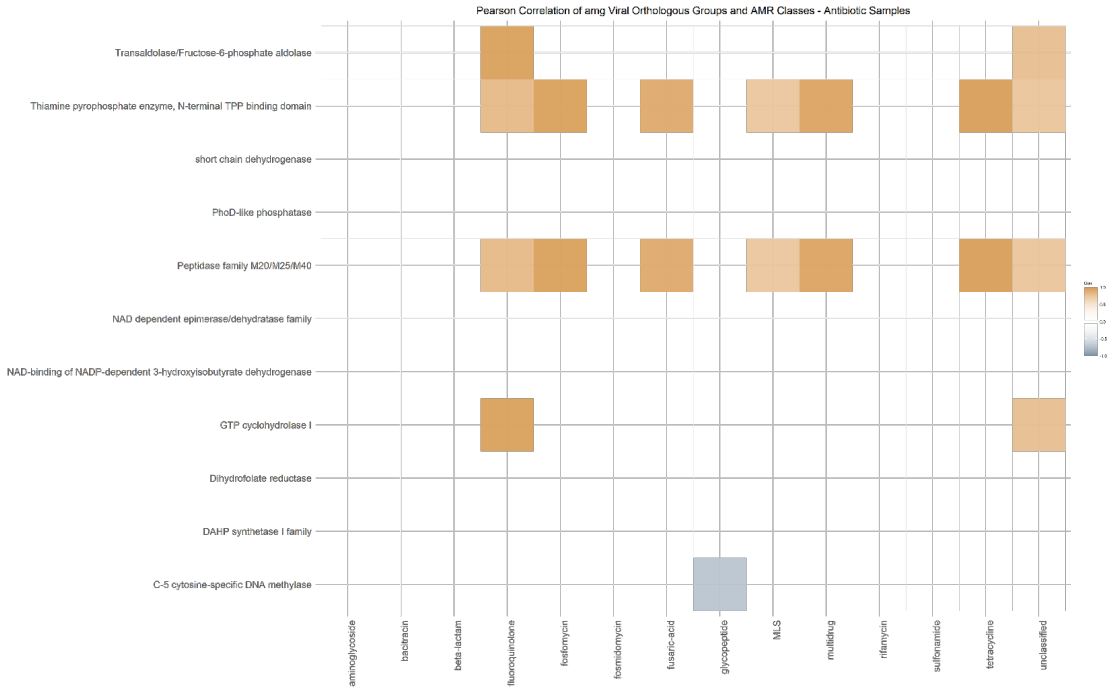


Figure 2.22: Pearson correlation between AMR gene classes (x-axis) and phage-encoded auxiliary metabolic genes (AMGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

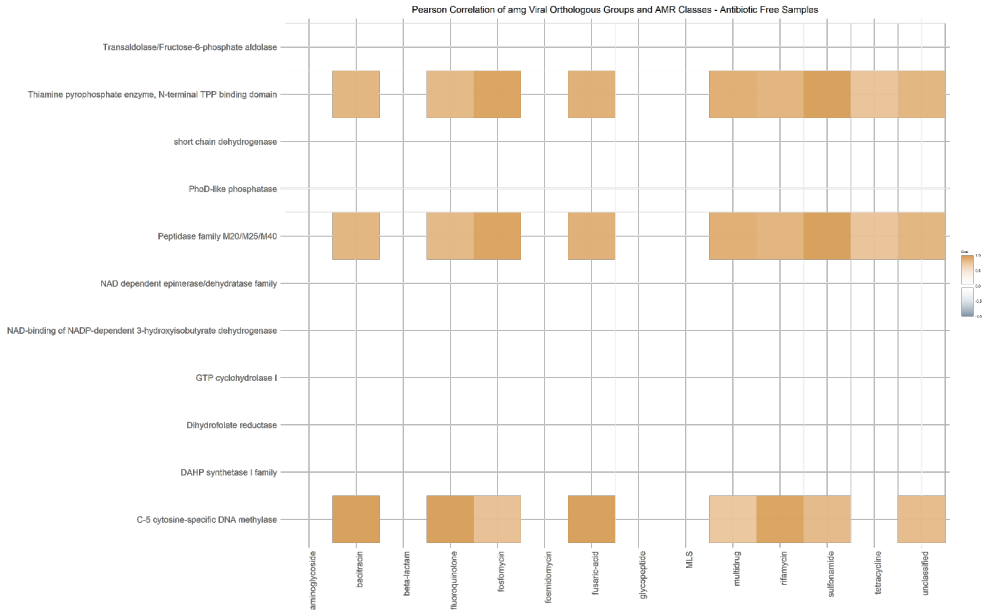


Figure 2.23: Pearson correlation between AMR gene classes (x-axis) and phage-encoded auxiliary metabolic genes (AMGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Understandably, similar to in trends of Pearson correlations of AMR genes shown in Figures 2.22 and 2.23, Thiamine pyrophosphate enzyme, N-terminal TPP binding domain and Peptidase family M20/M35/M40 exhibited almost identical Pearson correlation trends between Antibiotic and Antibiotic-Free samples, with C-5 cytosine-specific DNA methylase exhibiting similar patterns of correlation in the Antibiotic-Free samples. C-5 cytosine-specific DNA methylase only exhibited a single significant Pearson correlation in Antibiotic samples, a negative correlation with glycopeptide AMR class (Figure 2.22).

Fluoroquinolone and Unclassified AMR classes were correlated with the same 4 AMGs in Antibiotic samples (Figure 2.22), Transalsolase/Fructose-6-phosphate aldolase, Thiamine pyrophosphate enzyme N-terminal TPP binding domain, Peptidase family M20/M35/M40, and GTP cyclohydrolase I.

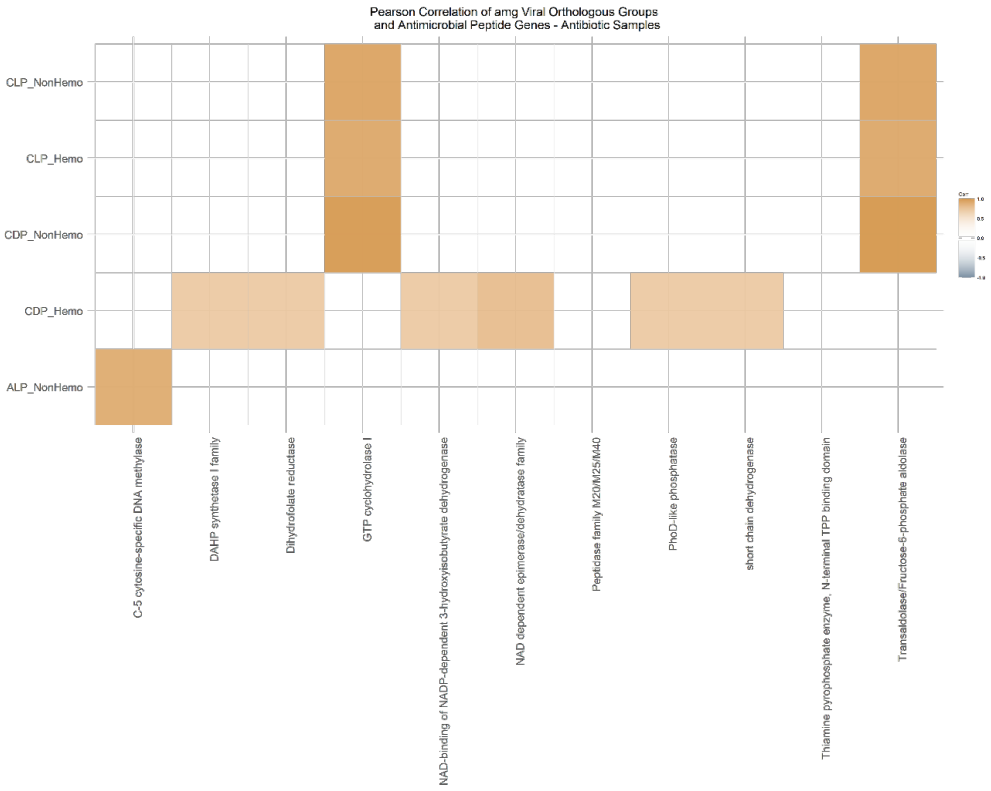


Figure 2.24: Pearson correlation between phage-encoded auxiliary metabolic genes (AMGs) (x-axis) and AMP classes (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

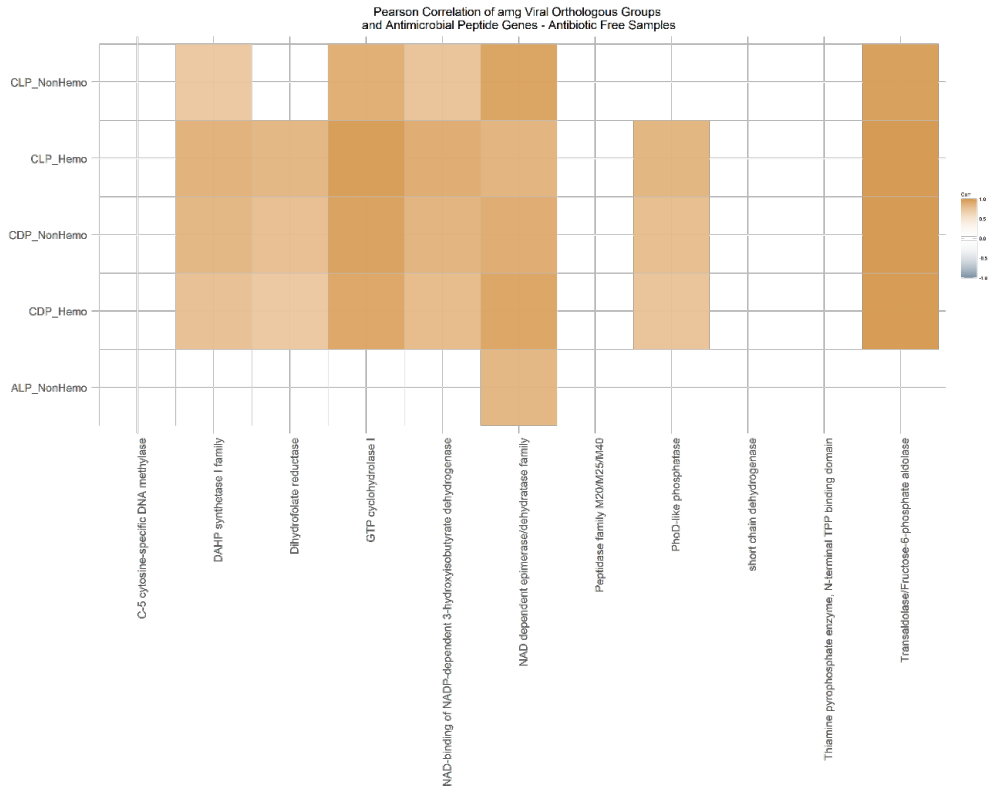


Figure 2.25: Pearson correlation between phage-encoded auxiliary metabolic genes (AMGs) (x-axis) and AMP classes (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations ($p < 0.05$) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

In Antibiotic samples, GTP cyclohydrolase I and Transaldolase/Fructose-6-phosphate alsolase were both strongly positively correlated with hemolytic and non-hemolytic cationic linear peptides (CLP), as well as non-hemolytic cationic cysteine-containing peptides (CDP) (Figure 2.24). Additionally, C-5 cytosine-specific DNA methylase was positively correlated with non-hemolytic ALPs in Antibiotic samples.

Five AMGs were positively correlated with all categories of cationic AMPs in Antibiotic-Free samples: CAHP synthetase I family, GTP cyclohydrolase I, NAD-binding of NADP-

dependent 3-hydroxisobutryate dehydrogenase, NAD dependent epimerase/dehydralase family, and Transalsolase/Fructose-6-phosphate aldolase (Figure 2.25).

NAD dependent epimerase/dehydralase family was positively correlated with all classes of AMPs in Antibiotic-Free samples and was the only significant correlation with anionic AMPs in these samples (Figure 2.25).

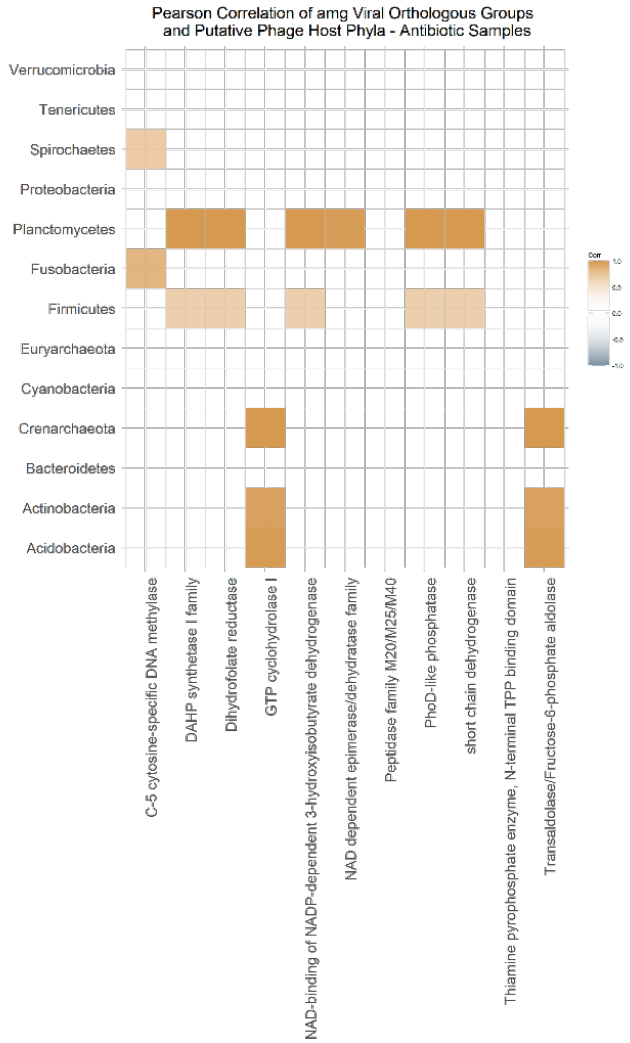


Figure 2.26: Pearson correlation between phage-encoded auxiliary metabolic genes (AMGs) (x-axis) and phage contigs as categorized by their putative hosts as a taxonomic proxy (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only

significant Pearson correlations (p < 0.05) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

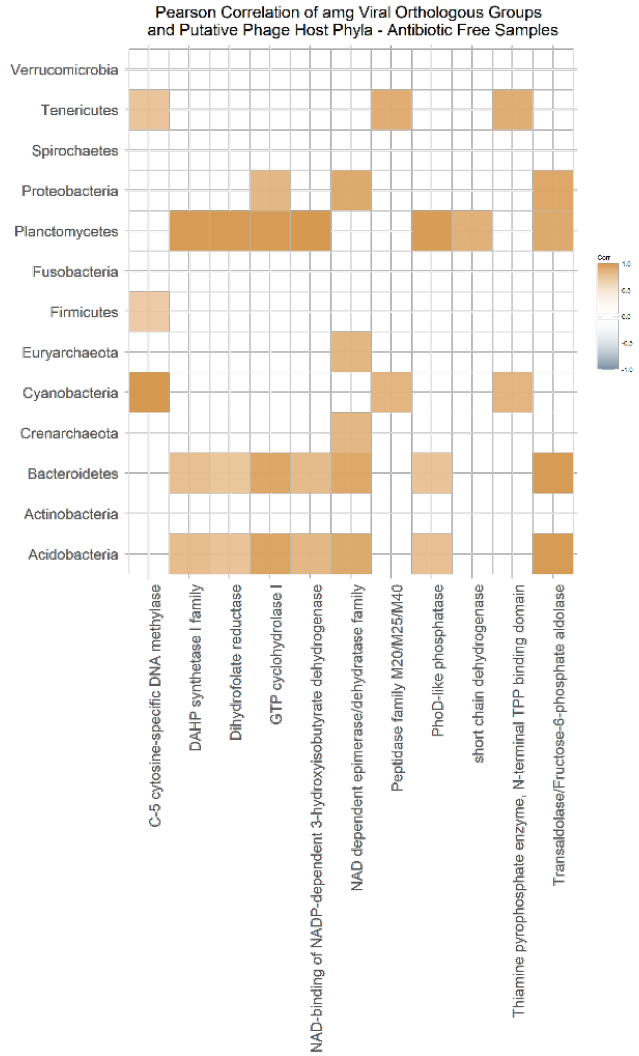


Figure 2.27: Pearson correlation between phage-encoded auxiliary metabolic genes (AMGs) (x-axis) and phage contigs as categorized by their putative hosts as a taxonomic proxy (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations (p < 0.05) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

Planctomycetes and Firmicutes were the most highly correlated putative host phyla in the Antibiotic samples (Figure 2.26), showing positive correlations with 6 and 5 AMGs respectively, with Planctomycetes exhibiting stronger positive correlations than Firmicutes with these AMGs. GTP cyclohydrolase I and Transalsolase/Fructose-6-phosphate aldolase were the most highly correlated AMGs in Antibiotic samples, both being correlated with the same three putative phage host phyla, Acidobacteria, Actinobacteria, and Crenarchaeota (Figure 2.26).

In Antibiotic-Free samples, three putative phage host phyla, Acidobacteria, Bacteriodetes, and Planctomycetes, were positively correlated with 7 AMGs, with positive correlations between these phyla and 6 AMGs being shared: DAHP synthetase I family, Dihydrofolate reductase, GTP cyclohydrolase I, NAD-binding of NADP-dependent 3-hydroxisobutryate dehydrogenase, PhoD-like phosphatase, and Transalsolase/Fructose-6-phosphate aldolase (Figure 2.27).

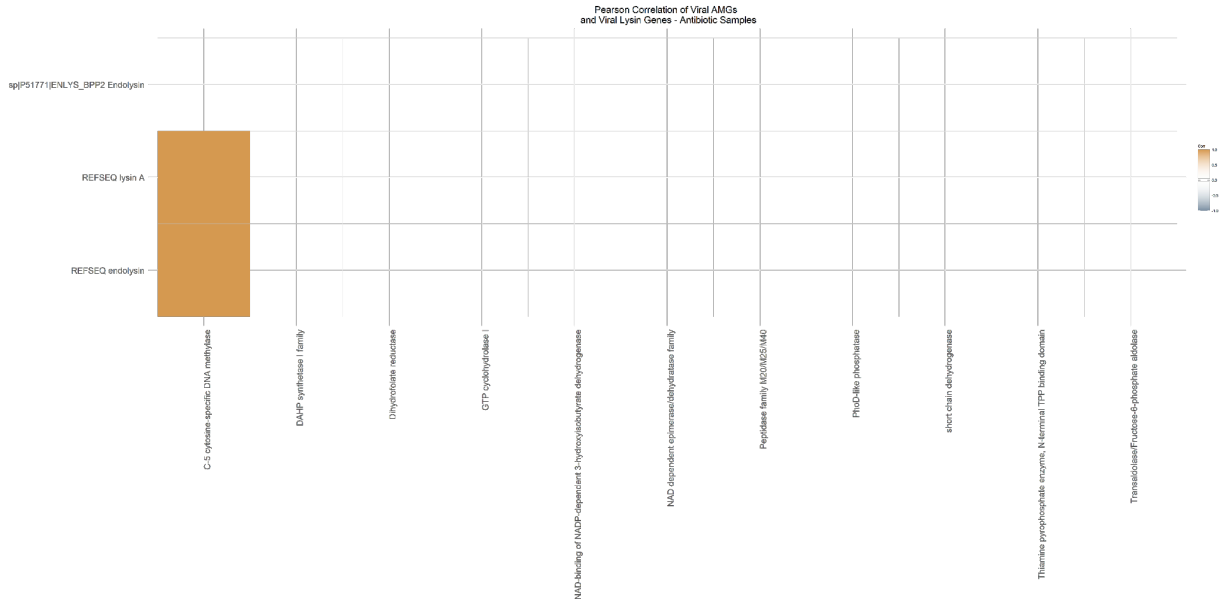


Figure 2.28: Pearson correlation between phage-encoded auxiliary metabolic genes (AMGs) (x-axis) and lysin viral orthologous groups (VOGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who

utilized antibiotics for treatment of their BYP flocks (Antibiotic). Only significant Pearson correlations (p < 0.05) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

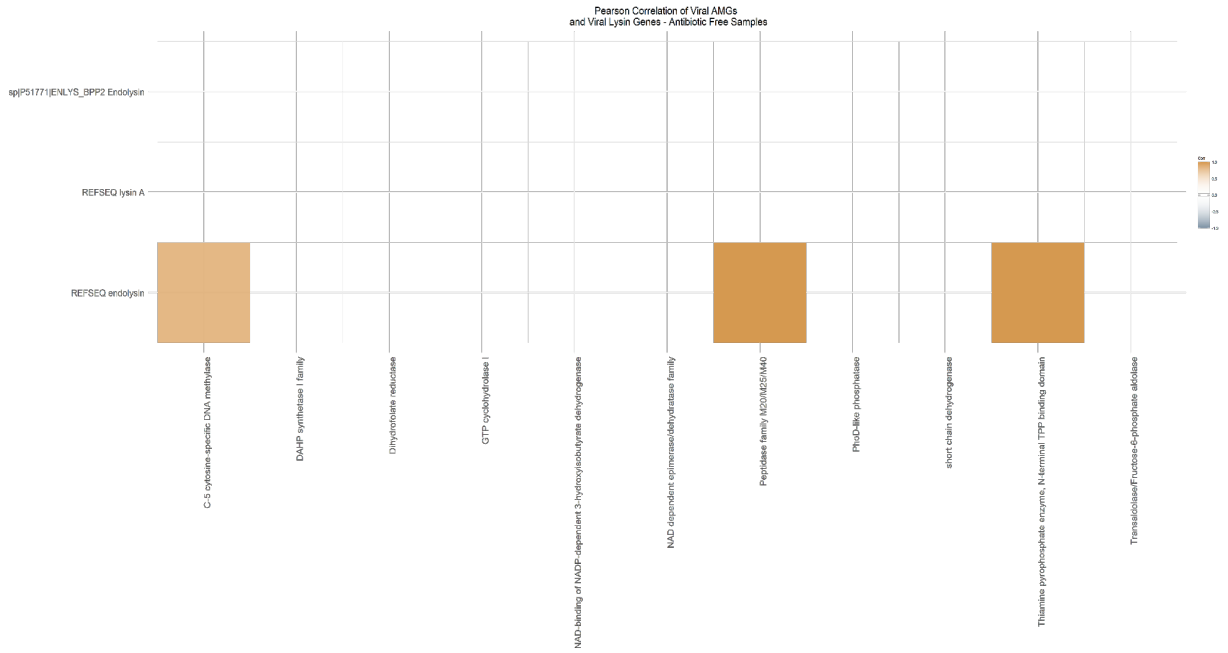


Figure 2.29: Pearson correlation between phage-encoded auxiliary metabolic genes (AMGs) (x-axis) and lysin viral orthologous groups (VOGs) (y-axis) for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free). Only significant Pearson correlations (p < 0.05) are displayed, with color indicating the directionality and intensity of the Pearson r-value.

C-5 cytosine-specific DNA methylase was the only AMG correlated with any lysin VOG in Antibiotic samples, exhibiting a positive correlation with REFSEQ lysin A and REFSEQ endolysin (Figure 2.28). The positive correlation between C-5 cytosine-specific DNA methylase and REFSEQ endolysin was also seen in Antibiotic-Free samples (Figure 2.29). REFSEQ endolysin was also positively correlated with Thiamine pyrophosphate enzyme N-terminal TPP binding domain and Peptidase family M20/M35/M40 in Antibiotic-Free samples (Figure 2.29).

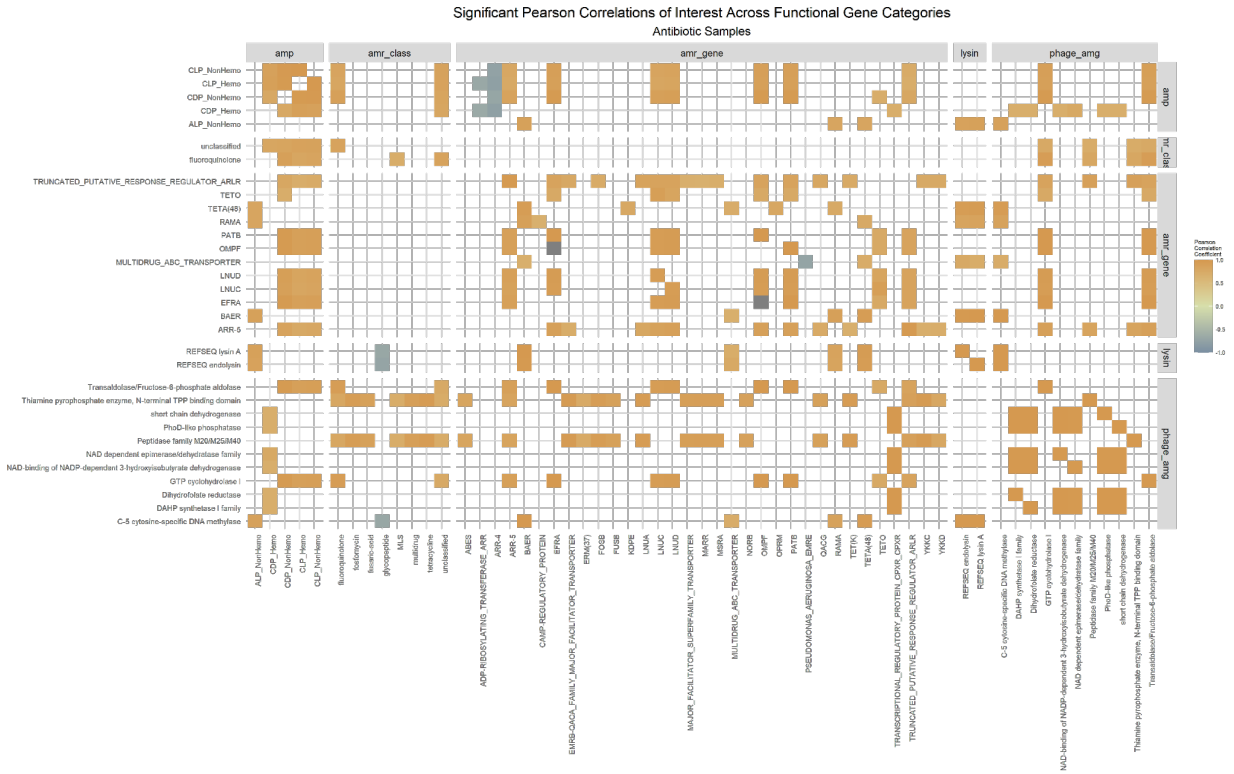


Figure 2.30: Summary figure of significant Pearson correlations ***for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).*** *Functional attributes were included in this summary if they had significant Pearson correlation values with other attributes across at least three functional categories (AMR gene, AMR class, AMP class, Lysin VOG, Phage AMG).*

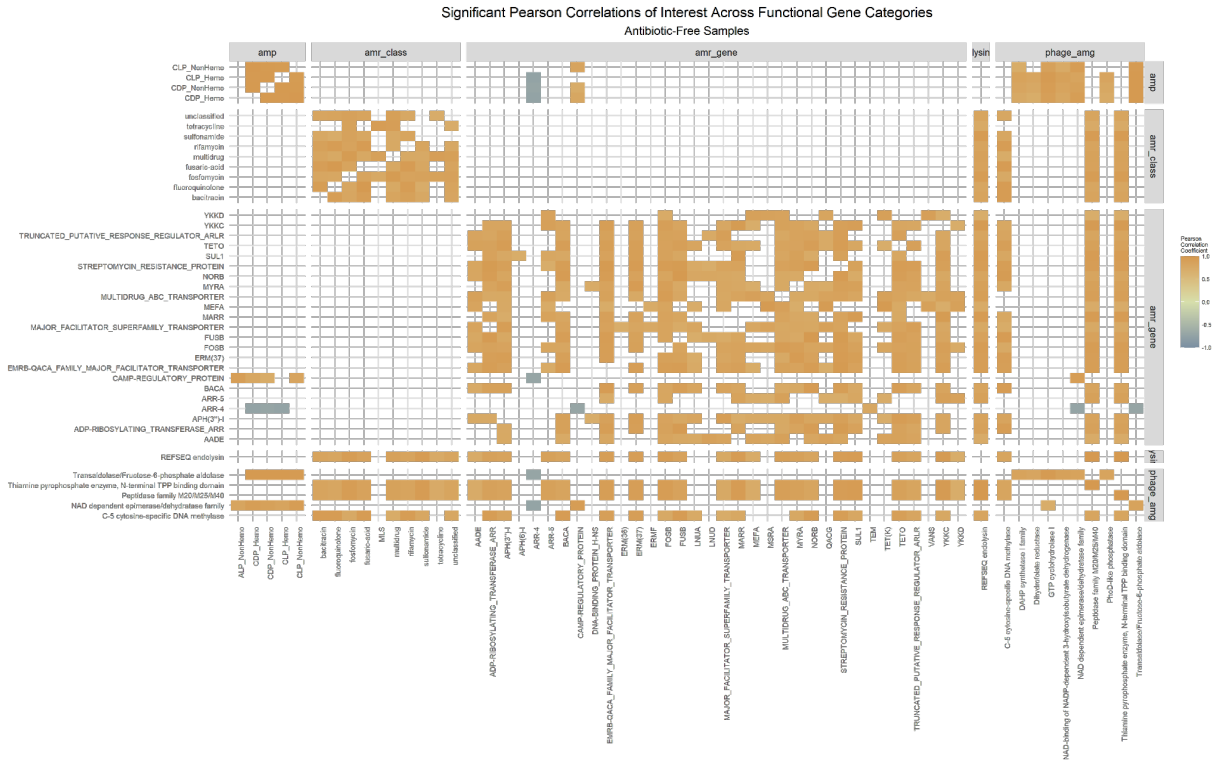


Figure 2.31: Summary figure of significant Pearson correlations ***for all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).*** *Functional attributes were included in this summary if they had significant Pearson correlation values with other attributes across at least three functional categories (AMR gene, AMR class, AMP class, Lysin VOG, Phage AMG).*

Figures 2.30 and 2.31 show all associations of interest described in Figures 2.7 - 2.29 and Supplementary figures S2.1 - S2.25 to summarize all Pearson correlations that may be of interest for further investigation for biotech application of AMR mitigation and control. Correlations shown in Figures 2.30 and 2.31 were filtered to only include individual functional items that exhibited significant correlations across three or more functional categories. Self-correlations between individual functional items were not included in the criteria for inclusion in this summary heatmap.

Correlations displayed in these figures provide potential avenues of interest for research aiming to establish a foundational understanding of these naturally occurring, synergistic, ecological relationships for the development of antimicrobial treatments.

Limitations

Identification of lysin VOGs was performed by searching Pfam annotations of identified phage-originating contigs. However, this method likely does not capture the entirety of the abundance and diversity of lysin sequences in these microbial communities. Lysin VOGs identified in this study are likely a small subset of the lysins in these microbial communities since identification of these genes was not through an approach catered specifically to the discovery and identification of lysin sequences. There may be many more of lysins of interest that would be uncovered if more rigorous bioinformatic approaches were used here for lysin-specific discoveries. Methods such as those described in Fernandez-Ruiz et al. are more comprehensive in their identification and characterization of known and putative lysin sequences (97).

This analysis categorized phage-originating contigs by their putative phage-host taxonomy rather than the taxonomy of the phage itself, since taxonomic identification of phage from metagenomic sequences is difficult and error-prone. Manual curation of phage contacts using vContact2 may yield more accurate taxonomic classification than the phage-host method chosen for this study, but is much more laborious and time-consuming (98).

Conclusion

Investigation of the taxonomic and functional distribution of the bacteriophage community in BYP environments revealed that (1) phage are not major contributors to AMR retention or

dissemination in these environments and (2) phage in these environments encode only anionic, linear peptide AMPs. Pearson correlation analysis revealed that while phage do not directly contribute to the dissemination of AMR genes in this environment, there are many significant correlations between phage-encoded features, such as lysins, and AMR classes and genes. These newly elucidated relationships are of particular interest for further research as leveraging these native relationships may be critical in the development of phage-derived therapeutic strategies for antibiotic-resistant bacterial infections.

Chapter 3: Taxonomic and Functional Characterization of NASA JPL GESAM Bacterial Isolates

Introduction

The National Aeronautics and Space Administration (NASA) Department of Planetary Protection (PP) is responsible for the monitoring and mitigation of forward contamination risks, i.e. the risk of microorganisms to extraterrestrial environments (99). This consideration is especially important when considering the sterilization protocols of Spacecraft Assembly Facilities (SAF), which undergo extremely stringent sterilization processes to minimize the risk of any microbial contamination of spacecraft hardware (100). Since complete sterilization of an entire SAF is unattainable, the continuous sterilization of these facilities likely selects for extremely resistant microbes that can survive under extraordinarily harsh conditions and therefore pose an additional risk of forward contamination. Extensive efforts have been made in the development of the NASA Standardized Assay, a standardized microbial examination assay used to monitor the "cleanliness" of spacecraft surfaces, monitor microbial burden, and calculate relative forward planetary protection risk (101). This microbial monitoring assay has been utilized since the Viking I mission in 1975 and consequently has enabled the cultivation and storage of 5,494 microbial isolates collected from the spacecraft hardware and associated surfaces from eight Mars missions, creating a unique "microbial time capsule" of preserved microbial organisms. Genomic characterization of this microbial collection through whole genome sequencing (WGS) analysis was undertaken by the Genome Encyclopedia of Spacecraft Associated Microbes (GESAM)

project at the NASA Jet Propulsion Laboratory (JPL). The goal of the GESAM project was to perform comprehensive taxonomic and functional analyses of this microbial archive and to curate reference taxonomic and functional genomic databases for metagenomic and WGS analysis of microbes from ongoing and future missions.

Genes that may enable microbes to survive extreme environmental conditions, such as spacecraft flight or other extreme planetary environmental conditions, are of particular interest for investigation in the GESAM collection. Environmental resistance genes may enable microbial organisms to survive in otherwise inhospitable conditions, increasing the likelihood of contributing to forward contamination risk. Investigation into these environmental resistance genes may elucidate trends in functional gene retention, such as longitudinal or taxonomic trends, that can be utilized in the development and execution of future forward planetary protection efforts. This study aims to characterize the taxonomic assignment and functional characterization of microbial isolates collected from spacecraft of NASA missions, Viking I, Mars Pathfinder, Odyssey, Mars Exploration Rover (MER), Phoenix, and Mars Science Laboratory (MSL), between 1975 and 2012.

Materials and Methods

Sample Collection and Sequencing

The NASA Standardized Assay is a cultivation-dependent, bacterial spore-based method detecting only aerobic, mesophilic, heterotrophic spore-forming organisms from spacecraft surfaces. Although this assay does not provide a comprehensive view of the microbial community due to the explicit and inherent bias towards culturable, spore-forming organisms, it does provide a view of the taxonomy and functional capacity of spore-forming microbes, a microbial demographic most likely to pose a forward contamination threat.

Microbial isolates of interest, based on previous attempts at taxonomic classification using MALDI-TOF and partial 16S rRNA amplicon sequencing, were revived from glycerol stocks in triplicate and incubated at 32°C on tryptic soy agar (TSA) plates for 72 hours. Isolates forming isolated colonies within 72 hours were visually inspected for purity and colony morphology. Resulting pure isolate cultures were assigned a barcode and pursued for WGS analysis.

DNA from pure isolate cultures was extracted using the Promega Maxwell RSC automated instrument and cultured cells DNA extraction kit (Fisher Scientific, Cat. No. PRAS1620), and was quantified using the Promega Quantus fluorometer to ensure a minimum of 100 ml of 100 ng/ml of DNA was obtained. DNA purity was additionally quantified on the Nanodrop 1000 spectrophotometer to confirm a 260/280 of >1.8 and 260/230 >2.0, metrics which indicate adequate purity of DNA in the solution, free from contamination or excess protein.

Sequencing library preparation was performed using Nextera XT adapters, modified for large inserts. Paired-end 250 base-pair sequencing was performed on an Illumina HiSeq platform by MicrobesNG's WGS service, using standard parameters, and sequenced to 30x coverage.

Genomic Assembly

201 WGS isolates were selected for further analysis based on distribution between missions and microbial characteristics of interest. Illumina reads were assessed for quality and assembled through the MicrobesNG web portal using the default workflow parameters. Code for the nextflow workflow utilized by MicrobesNG can be found at https://github.com/MicrobesNG/process-run. Defined Illumina adapters were removed and reads were quality trimmed by trimmomatic bidirectionally with a minimum quality score of phred33, a sliding window of 4:15, and a minimum length of 36 bases (102). Trimmed reads were used as input to SPAdes assembler in "careful" mode, which attempts to reduce the number of mismatches and short indels. This mode also incorporates MismatchCorrector, a SPAdes native software that utilizes the BWA tool to improve mismatch and short indel rates resulting in contigs

and scaffolds (103–105). Assemblies were evaluated for quality and coverage through manual inspection of QUAST quality metrics of N50 and total contig numbers (106).

Taxonomic Assignment

WGS were assigned taxonomy using the GTDB-Tk analysis pipeline, which leverages single-copy genes to infer phylogenetic relationships and assign taxonomic groups (107). This analysis was performed using the Classify workflow ("classify_wf") with default parameters.

In the event taxonomy was unable to be determined by GTDB-TK analysis (defined as less than 97% match to species level), 16S rRNA gene sequences were extracted from the assembly using Anvio function anvi-get-sequences-for-hmm-hits and queried against the SILVA LTP database (LTP_09_2021) for a higher confidence match using blastn with default parameters (4, 108–110). Full scientific names of the taxonomic groups were retrieved using the R package taxonomizr.

Resistance Gene Annotation

Gene Selection

Categories relevant to resistance types of interest were defined as genes related to the following functional categories: Cold Shock, Oxidative Damage Resistance, Repair and Recombination, and Sporulation. Miscellaneous genes related to environmental resistance processes of interest but otherwise unrelated to one another were categorized as "WTF Processes". Genes names were consolidated from specific genes of interest related to and literature reviews of functions of interest, and accession numbers were collected for each gene (Table S-Metadata 3). Although sporulation is one of the major selection criteria for the NASA Standardized Assay, understanding the diversity of sporulation genes within the category was also of interest and therefore included as a Resistance Gene Category for this study.

Gene Query to WGS

Accession numbers were used to retrieve the gene's hidden markov model (HMM) using anvio's function anvi-script-pfam-accessions-to-hmms-directory for Pfam accession numbers, and NCBI Entrez API for TIGRFAM accession numbers. HMM querying is a more comprehensive method of querying sequencing data than traditional alignment-based approaches, since HMM-based querying approaches utilize statistical models that account for the probability of differences in a collection of sequences used to build reference HMMs that are not directly observable, and therefore unable to be accounted for in alignment-based approaches (111). Each WGS was then subjected to HMM query for each resistance gene HMM using the anvio function anvi-run-hmms. All metadata and raw data tables can be found in the figshare repository for this project (112).

Data Analysis and Visualization

Dataframe manipulations of taxonomic and functional annotations described above were performed in R using tidyverse. Visualizations were generated using ggplot2 in R. Statistical calculations were performed using base R functions. A complete description of the R session info, including all package names and versions, can be found in the Supplementary section of this dissertation. All scripts for analyses described in this chapter can be found in the GitHub repository at https://github.com/alonnawright/2021-jpl_gesam. All metadata and high-quality figure images can be found in the figshare repository for this project (https://doi.org/10.6084/m9.figshare.21555225) (112).

Results and Discussion

Table 3.1: Frequency of Genus of Isolates From Each MIssion of bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly. Pink cells represent the highest taxonomic frequency among isolates collected from the respective mission.

Genus	Mars Pathfinder	MER	MSL	Odyssey	Phoenix	Viking
Bacillus	6	20	8	0	7	0
Solibacillus	1	0	0	0	0	0
Staphylococcus	1	8	4	3	1	0
NA	1	2	3	0	5	0
Agromyces	0	1	0	0	0	0
Alkalihalobacillus	0	2	0	2	0	0
Bhargavaea	0	1	0	0	0	0
Brevibacillus	0	1	3	0	0	2
Carnobacterium	0	1	0	0	0	0
Curtobacterium	0	1	0	1	0	0
Cytobacillus	0	3	1	1	4	0
Exiguobacterium	0	1	0	0	0	0
Fictibacillus	0	1	0	0	2	0
Kocuria	0	1	1	1	0	0
Mesobacillus	0	5	1	1	0	0
Metabacillus	0	3	1	0	0	0
Microbacterium	0	1	0	3	0	0
Neobacillus	0	5	1	0	0	0
Niallia	0	3	1	0	0	0
Nocardioides	0	1	0	0	0	0
Oceanobacillus	0	1	0	0	0	0
Paenibacillus	0	11	4	0	2	0
Peribacillus	0	1	1	0	0	0
Planococcus	0	1	0	0	0	0

Table 3.1: Frequency of Genus of Isolates From Each MIssion of bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly. Pink cells represent the highest taxonomic frequency among isolates collected from the respective mission.

Genus	Mars Pathfinder	MER	MSL	Odyssey	Phoenix	Viking
Priestia	0	8	3	2	2	0
Psychrobacillus	0	1	1	0	0	0
Rummeliibacillus	0	1	0	0	0	0
Sphingomonas	0	1	0	0	0	0
Sphingopyxis	0	1	0	0	0	0
Sporosarcina	0	3	0	0	0	0
Stenotrophomonas	0	1	0	0	0	0
Streptococcus	0	1	0	0	0	0
Streptomyces	0	2	0	0	0	0
Terribacillus	0	1	0	0	0	0
Ureibacillus	0	1	0	0	0	0
Caldibacillus	0	0	2	0	0	1
Cellulosimicrobium	0	0	1	0	0	0
Cupriavidus	0	0	1	0	0	0
Domibacillus	0	0	1	0	0	0
Heyndrickxia	0	0	1	0	1	0
Lederbergia	0	0	1	0	0	0
Micrococcus	0	0	3	0	0	0
Ralstonia	0	0	1	0	0	0
Weizmannia	0	0	1	0	0	0
Alkalihalophilus	0	0	0	1	0	0
Georgenia	0	0	0	1	0	0
Rothia	0	0	0	1	1	0
Sutcliffiella	0	0	0	1	0	0
Hydrogenophaga	0	0	0	0	1	0

Bacillus genus was the most abundant taxonomic assignment to isolates from the Mars Pathfinder, MER, MSL, and Phoenix missions, at 6, 20, 8, and 7 *Bacillus* identified isolates respectively. Isolates from the Odyssey mission had a tied highest frequency of identified isolate genus between *Staphylococcus* and *Microbacterium* at three isolates each, while the most abundant identified genus of the Viking mission was *Brevibacillus* at two identified isolates. Since the initial isolate culturing procedure was facilitated through a spore-forming assay, the abundance and diversity of these isolates is not representative of the original microbial community. However, accurate taxonomic identification of these isolates is critical for future metagenomic and WGS analysis and therefore the concession of capturing microbial abundance and diversity for these missions is a worthwhile endeavor.

Interestingly, some of the bacterial isolates were taxonomically classified as non-spore-forming genera, such as *Staphylococcus.* There is precedent for *Staphylococcus* being able to survive sterilization methods in clinical settings in virulent strains of methicillin-resistant *Staphylococcus aureus* (MRSA) and coagulase-negative Staphylococci (CoNS) (113). Therefore while the NASA Standardized Assay is intended to only select for “spore-forming” individuals, it is not unreasonable to assume that other non-spore-forming organisms encoding environmental resistance genes may also be able to survive in the spacecraft assembly clean rooms and be cultivated through this assay.

Table 3.2: Summary of HMM Hit Abundances Per Mission of bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly.

				Resistance Gene HMM Hits per HMM per Genome			
Mission	**Genomes per Mission**	**Total HMM Hits**	**Average HMM Hits per Genome**	**Minimum**	**Maximum**	**Mean**	**Median**
Mars Pathfinder	9	2560	284.44	0	20	1.07	1

MER	96	28304	294.83	0	45	1.10	1
MSL	45	13298	295.51	0	24	1.12	1
Odyssey	18	4110	228.33	0	17	0.86	0
Phoenix	26	7853	302.04	0	20	1.14	1
Viking	3	1043	347.67	0	19	1.31	1

Table 3.2 shows statistical summaries of the frequency of HMM hits for all resistance gene categories for each mission. Isolates from the Viking mission had the highest frequency of resistance gene HMM hits among the missions, with 347.67 resistance gene HMM hits per genome, while Odyssey isolates showed the lowest resistance gene frequencies at an average of 228.33 resistance gene HMM hits per genome. All missions showed similar mean HMM hits per genome for individual resistance genes, with a range of the means being from 0.86 for Odyssey isolates to 1.31 for Viking isolates.

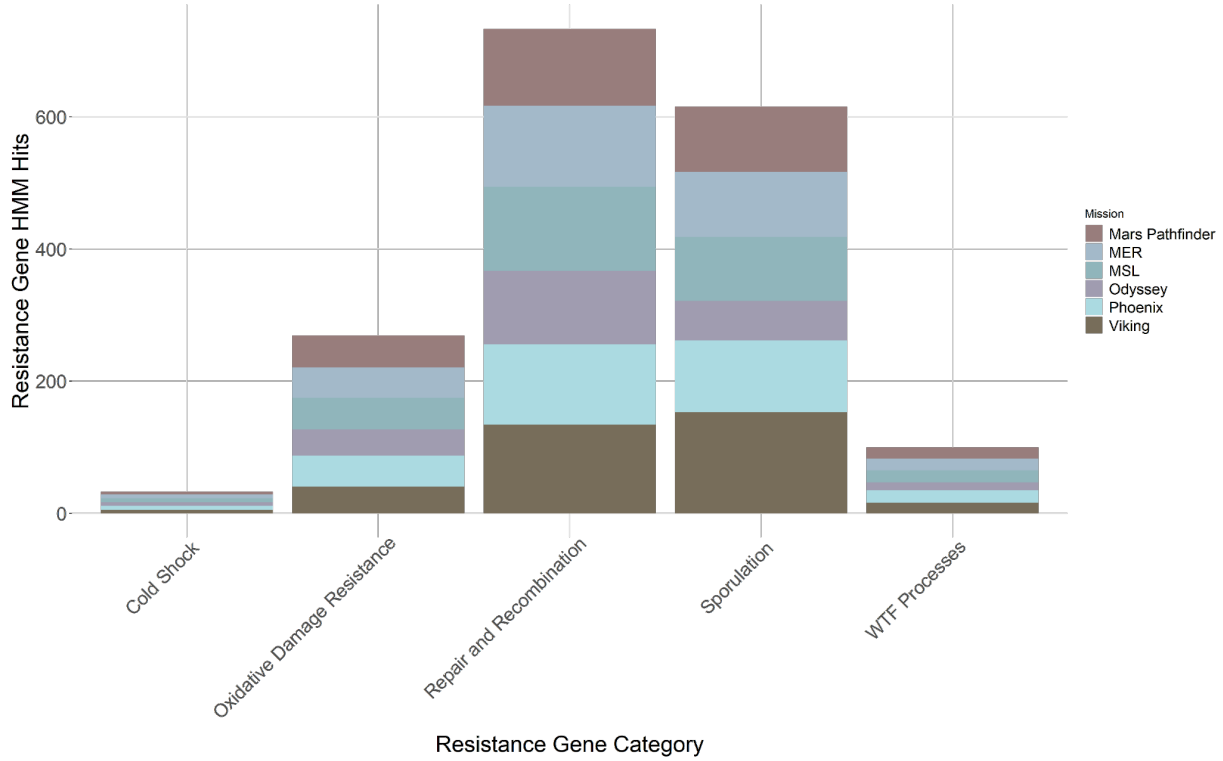


Figure 3.1: Total HMM Hit Abundance of Resistance Gene Categories for Missions of GESAM bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA

JPL during spacecraft assembly.

Table 3.3: Summary of Resistance Gene HMM Hits Per Mission, Normalized by Number of Genomes Sequencedof GESAM bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly.

Mission	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
Mars Pathfinder	4.11	48.11	116.44	98.67	17.11
MER	5.97	46.80	124.20	99.83	18.03
MSL	6.18	47.67	126.93	96.69	18.04
Odyssey	5.50	39.61	111.33	59.50	12.39
Phoenix	5.92	47.04	121.38	109.04	18.65
Viking	5.00	40.33	134.00	152.67	15.67

Table 3.3 and Figure 3.1 summarize the distribution of resistance gene category frequencies across missions when normalized to the number of genomes sequenced for each mission. Cold shock tolerance genes were the least frequently encoded category, ranging from 4.11 HMM hits per genome in the Mars Pathfinder mission isolates to 6.18 HMM hits per genome in the MSL mission isolates. Repair and recombination was the overall most frequently detected resistance gene category, ranging from 111.33 HMM hits per genome in the Odyssey mission isolates to 134.00 HMM hits per genome in the Viking mission isolates. As many of the identified resistance genes are essential for bacterial homeostasis and survival, it is not surprising to see a high frequency of these genes, but rather the interesting aspect lies in the copy number retention between missions. Bacterial retention of multiple genes within a category, including paralogs and

xenologs, indicates that there is a potential ecological advantage to the retention of functionally redundant genes in order to survive in a particularly harsh environment, such as the Clean Room these microbes were originally isolated from (114).

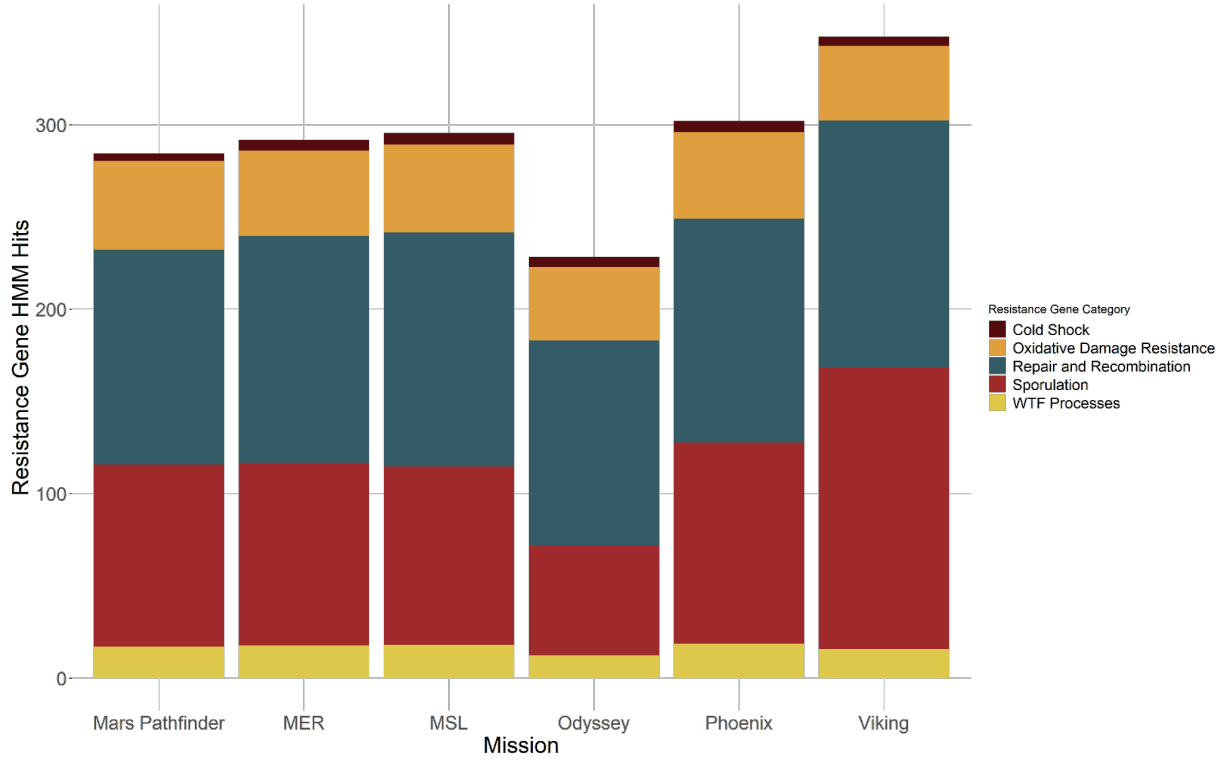


Figure 3.1: Total HMM Hit Abundance Within Missions by Resistance Gene Categories of GESAM bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly.

Figure 3.1 shows the distribution of resistance gene category frequencies for each mission, normalized by the number of genomes sequenced per mission. All missions show similar frequencies of each resistance gene category, while Odyssey mission isolates exhibited overall lower rates of HMM hits and Viking mission isolates showed overall higher rates of HMM hits.

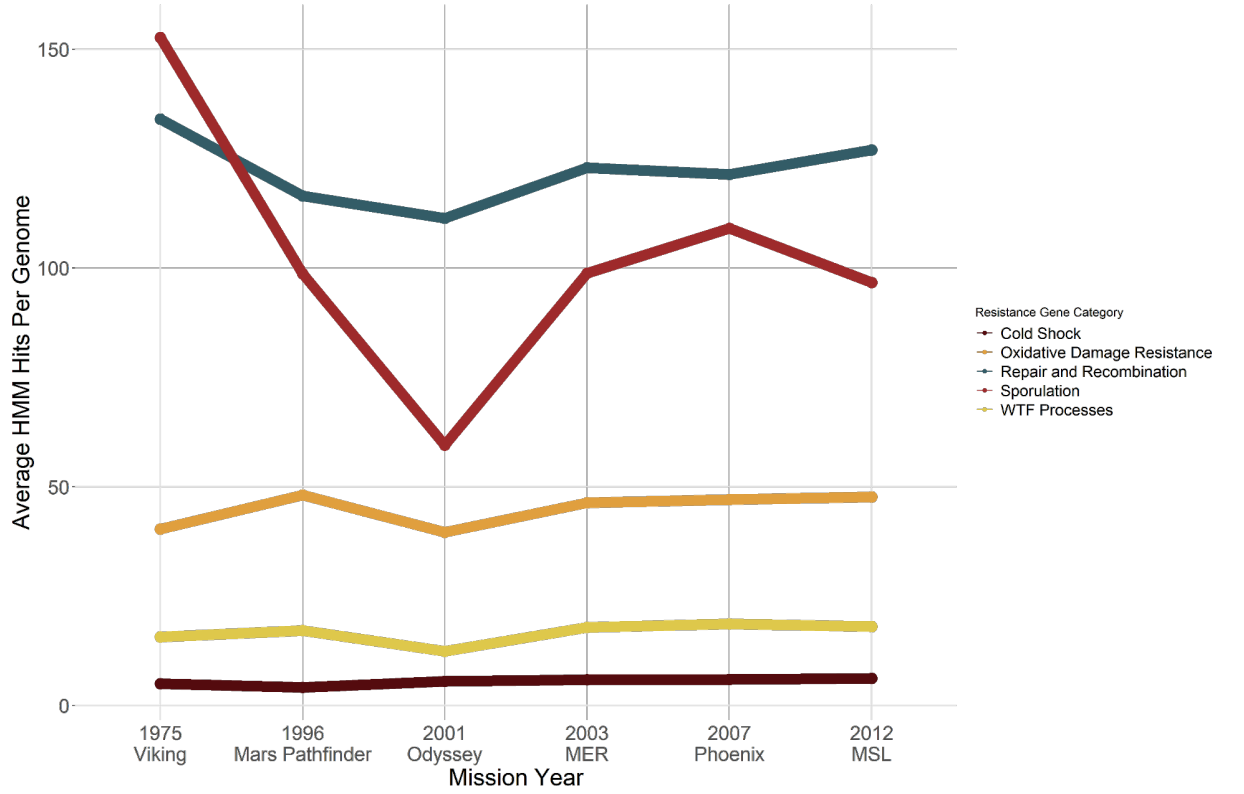


Figure 3.3: Average HMM Hits of Resistance Gene Categories per Genome by Mission Year of GESAM bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly.

Understanding longitudinal fluctuations of resistance gene frequencies are of particular interest in the characterization of the GESAM project collection. Changes to resistance gene abundances over time may indicate environmental changes of interest or a potential increase in acquired resistance genes perpetuated by inadequate sterilization methods. Characterization efforts of the individual resistance genes, and the overall resistance gene categories, provide an avenue for monitoring resistance potential in cultivated microbes and is essential in the calculation of forward contamination risk. Figure 3.3 shows the resistance gene category abundances over each mission as a proxy for time. Sporulation genes exhibited the most notable fluctuation of all

resistance gene categories. The frequency of sporulation genes dropped from 152.67 to 59.50 HMM hits per genome between the Viking mission in 1975 and the Odyssey mission in 2001, respectively (Figure 3.3, Table 3.3). The abundance of sporulation genes recovered to 109.09 by 2007 in the isolates from the Phoenix mission spacecraft.

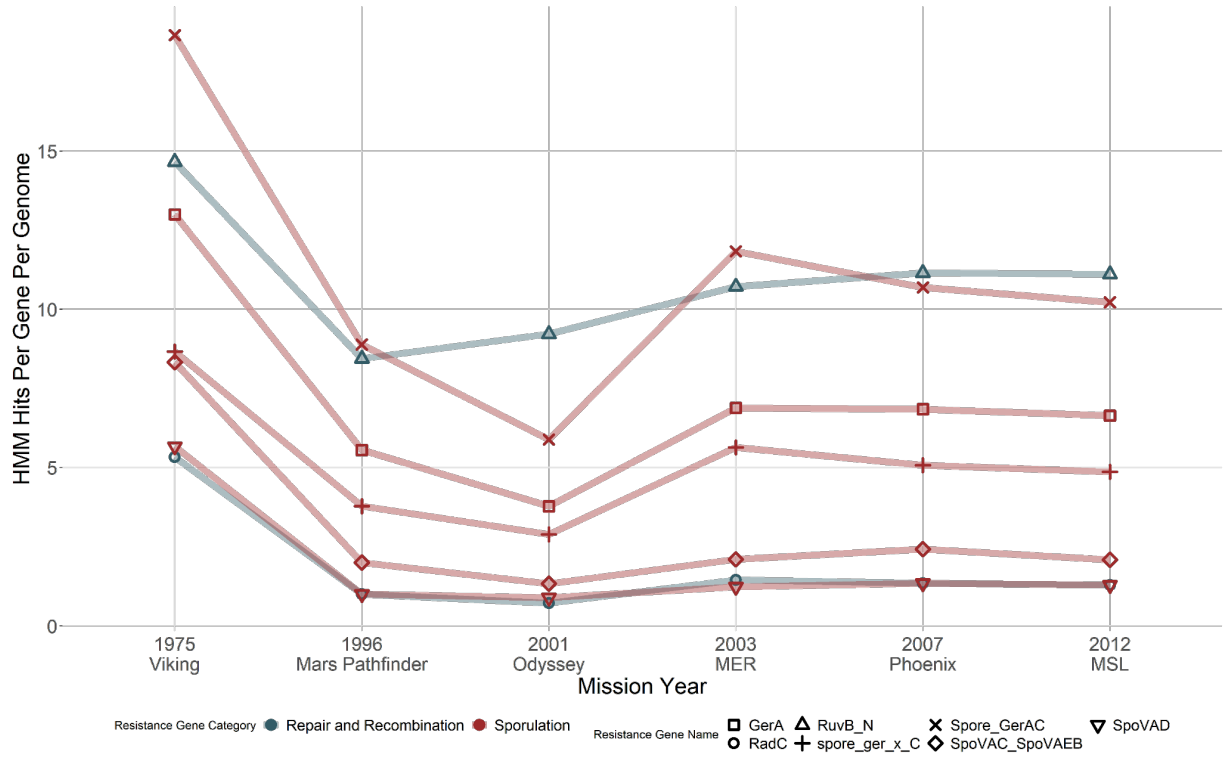


Figure 3.4: Average HMM Hits of Resistance Genes per Genome by Mission Year of GESAM bacterial isolates collected using NASA Standardized Assay from spacecraft hardware at NASA JPL during spacecraft assembly.

Visualization of genes exhibiting large frequency changes is shown in Figure 3.4, displaying the average frequency of each gene, normalized to the number of genomes sequenced for each mission. Genes for this visualization were filtered by overall change in frequency over time, selecting resistance genes whose range of average HMM hits per genome between all missions was at least four. Interestingly, genes exhibiting large change in gene frequency were only from the Repair and Recombination, and the Sporulation categories – five Sporulation genes

(Spore_GerAC, GerA, spore_ger_x_C, SpoVAX_SpoVAEB, SpoVAD) and two Repair and Recombination genes (RuvB_N, RadC). These genes exhibited their highest frequencies in isolates from the Viking mission in 1975 and dropped starkly by the Mars Pathfinder mission in 1996.

Limitations

Since initial isolate culturing procedures were facilitated through the standardized spore-forming assay, the abundance and diversity of this subset of isolates do not adequately capture the true microbial community composition of the environment, since non-spore formers and unculturable microbes are inherently excluded from this selection. However, accurate taxonomic identification of these isolates is critical for future metagenomic and WGS analysis and therefore the concession of capturing microbial abundance and diversity for these missions is a worthwhile endeavor.

The total number of genomes sequenced per mission is not a proxy for microbial activity associated with each mission, but rather these sequences are a subset of interest from the larger collection of 5,494 isolates associated with the GESAM project.

Individual resistance genes analyzed in this study were manually curated, other influential environmental resistance genes not evaluated in this analysis could contribute to environmental resistance in ways not captured in this study.

Conclusion

Microbial isolates from the NASA GESAM project were analyzed for taxonomic identification and functional capacity of environmental resistance genes of interest. Isolates from each mission contained genes from all of the resistance gene categories, with Repair and Recombination being the most abundant and Cold Shock being the least abundant across all

missions. Sporulation genes exhibited the largest fluctuation of abundance over time of all resistance gene categories. The taxonomic and functional analysis of these spacecraft-associated microbes will be critical in the development of bioinformatic tools for screening metagenomic and WGS data from other spacecraft-associated microbial samples.

Bibliography

1. Koonin EV, Makarova KS, Wolf YI. 2021. Evolution of Microbial Genomics: Conceptual Shifts over a Quarter Century. Trends Microbiol 29:582–592.
2. Wetterstrand K. DNA Sequencing Costs: Data. Genome.gov. https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data. Retrieved 13 November 2022.
3. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, Iyer R, Schatz MC, Sinha S, Robinson GE. 2015. Big Data: Astronomical or Genomical? PLoS Biol 13:e1002195.
4. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50:D20–D26.
5. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3:160018.
6. Wood-Charlson EM, Anubhav null, Auberry D, Blanco H, Borkum MI, Corilo YE, Davenport KW, Deshpande S, Devarakonda R, Drake M, Duncan WD, Flynn MC, Hays D, Hu B, Huntemann M, Li P-E, Lipton M, Lo C-C, Millard D, Miller K, Piehowski PD, Purvine S,

Reddy TBK, Shakya M, Sundaramurthi JC, Vangay P, Wei Y, Wilson BE, Canon S, Chain PSG, Fagnan K, Martin S, McCue LA, Mungall CJ, Mouncey NJ, Maxon ME, Eloe-Fadrosh EA. 2020. The National Microbiome Data Collaborative: enabling microbiome science. Nat Rev Microbiol 18:313–314.

7. Ali J, Rafiq QA, Ratcliffe E. 2018. Antimicrobial resistance mechanisms and potential synthetic treatments. Future Sci OA 4:FSO290.
8. Ahmad M, Khan AU. 2019. Global economic impact of antibiotic resistance: A review. J Glob Antimicrob Resist 19:313–316.
9. Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S, Rasool MH, Nisar MA, Alvi RF, Aslam MA, Qamar MU, Salamat MKF, Baloch Z. 2018. Antibiotic resistance: a rundown of a global crisis. Infect Drug Resist. Dove Press. https://www.dovepress.com/antibiotic-resistance-a-rundown-of-a-global-crisis-peer-reviewed-article-IDR. Retrieved 8 August 2020.
10. Agyare C, Boamah VE, Osei CNZ and FB. 2018. Antibiotic Use in Poultry Production and Its Effects on Bacterial Resistance. Antimicrob Resist - Glob Threat https://doi.org/10.5772/intechopen.79371.
11. Centers for Disease Control and Prevention (U.S.). 2019. Antibiotic resistance threats in the United States, 2019. Centers for Disease Control and Prevention (U.S.).
12. Elkhoraibi C, Blatchford RA, Pitesky ME, Mench JA. 2014. Backyard chickens in the United States: A survey of flock owners. Poult Sci 93:2920–2931.
13. Tobin MR, Goldshear JL, Price LB, Graham JP, Leibler JH. 2015. A Framework to Reduce Infectious Disease Risk from Urban Poultry in the United States. Public Health Rep 130:380–391.
14. USDA. 2011. Poultry10_dr_Urban_Chicken_1.pdf. 592.0511. USDA-APHIS-VA, CEAH, Fort Collins, CO.
15. Gilbert JA, Stephens B. 2018. Microbiology of the built environment. 11. Nat Rev Microbiol

16:661–670.

16. Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P, Behar JV, Hern SC, Engelmann WH. 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. 3. J Expo Sci Environ Epidemiol 11:231–252.
17. Rai S, Singh DK, Kumar A. 2021. Microbial, environmental and anthropogenic factors influencing the indoor microbiome of the built environment. J Basic Microbiol 61:267–292.
18. Ben Maamar S, Hu J, Hartmann EM. 2020. Implications of indoor microbial ecology and evolution on antibiotic resistance. 1. J Expo Sci Environ Epidemiol 30:1–15.
19. Grabowski Ł, Gaffke L, Pierzynowska K, Cyske Z, Choszcz M, Węgrzyn G, Węgrzyn A. 2022. Enrofloxacin—The Ruthless Killer of Eukaryotic Cells or the Last Hope in the Fight against Bacterial Infections? Int J Mol Sci 23:3648.
20. Mitchell MA. 2006. Enrofloxacin. J Exot Pet Med 15:66–69.
21. Medicine C for V. 2021. Withdrawal of Enrofloxacin for Poultry. FDA.
22. Shannon L, Cox SK, Bailey J, Fortner C, Davis R, Gerhardt L, Souza MJ. 2020. Pharmacokinetics and Drug Residue in Eggs After Multiple-Day Oral Dosing of Amoxicillin-Clavulanic Acid in Domestic Chickens. J Avian Med Surg 34:3.
23. Krasucka D, Kowalski CJ. PHARMACOKINETIC PARAMETERS OF AMOXICILLIN IN PIGS AND POULTRY 4.
24. Mcguire JM, Boniece WS, Higgens CE, Hoehn MM, Stark WM, Westhead J, Wolfe RN. 1961. Tylosin, a New Antibiotic : I. Microbiological Studies. Antibiot Amp Chemother 11:320–27.
25. BBMap. SourceForge. https://sourceforge.net/projects/bbmap/. Retrieved 5 August 2020.
26. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676.

27. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.
28. Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11:R25.
29. Jonsson V, Österlund T, Nerman O, Kristiansson E. 2016. Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. BMC Genomics 17:78.
30. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257.
31. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. 2018. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome 6:23.
32. Santos-Júnior CD, Pan S, Zhao X-M, Coelho LP. 2020. Macrel: antimicrobial peptide screening in genomes and metagenomes. PeerJ 8:e10555.
33. Wickham H. 2021. tidyverse: Easily Install and Load the Tidyverse. https://CRAN.R-project.org/package=tidyverse.
34. Harrell FE Jr. 2021. Hmisc: Harrell Miscellaneous. https://hbiostat.org/R/Hmisc/.
35. Kassambara A. 2019. ggcorrplot: Visualization of a Correlation Matrix using ggplot2. http://www.sthda.com/english/wiki/ggcorrplot.
36. R Core Team. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
37. Wright, Alonna. 2022. BYP Community Science - Home Metadata. figshare.
38. Dusa A. 2021. venn: Draw Venn Diagrams. https://github.com/dusadrian/venn.
39. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.
40. Shade A. 2017. Diversity is the question, not the answer. 1. ISME J 11:1–6.

41. Thukral A. 2017. A review on measurement of Alpha diversity in biology. Agric Res J 54:1.
42. Lozupone CA, Knight R. 2008. Species divergence and the measurement of microbial diversity. FEMS Microbiol Rev 32:557–578.
43. Xia Y. 2020. Chapter Eleven - Correlation and association analyses in microbiome study integrating multiomics in health and disease, p. 309–491. *In* Sun, J (ed.), Progress in Molecular Biology and Translational Science. Academic Press.
44. Frontiers | Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields. https://www.frontiersin.org/articles/10.3389/fmicb.2020.582779/full. Retrieved 15 October 2022.
45. van Hoek AHAM, Mevius D, Guerra B, Mullany P, Roberts AP, Aarts HJM. 2011. Acquired Antibiotic Resistance Genes: An Overview. Front Microbiol 2:203.
46. Fonseca ÉL da, Freitas F dos S, Amorim JC de, Vicente ACP. 2008. Detection of New *arr-4* and *arr-5* Gene Cassettes in Clinical *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* Strains from Brazil. Antimicrob Agents Chemother 52:1865–1867.
47. Lee E-W, Huda MN, Kuroda T, Mizushima T, Tsuchiya T. 2003. EfrAB, an ABC multidrug efflux pump in Enterococcus faecalis. Antimicrob Agents Chemother 47:3733–3738.
48. Achard A, Villers C, Pichereau V, Leclercq R. 2005. New lnu(C) gene conferring resistance to lincomycin by nucleotidylation in Streptococcus agalactiae UCN36. Antimicrob Agents Chemother 49:2716–2719.
49. Petinaki E, Guérin-Faublée V, Pichereau V, Villers C, Achard A, Malbruny B, Leclercq R. 2008. Lincomycin resistance gene lnu(D) in Streptococcus uberis. Antimicrob Agents Chemother 52:626–630.
50. Cohen SP, McMurry LM, Levy SB. 1988. marA locus causes decreased expression of OmpF porin in multiple-antibiotic-resistant (Mar) mutants of Escherichia coli. J Bacteriol 170:5416–5422.

51. El Garch F, Lismond A, Piddock LJV, Courvalin P, Tulkens PM, Van Bambeke F. 2010. Fluoroquinolones induce the expression of patA and patB, which encode ABC efflux pumps in Streptococcus pneumoniae. J Antimicrob Chemother 65:2076–2082.
52. Zhang Q-Y, Yan Z-B, Meng Y-M, Hong X-Y, Shao G, Ma J-J, Cheng X-R, Liu J, Kang J, Fu C-Y. 2021. Antimicrobial peptides: mechanism of action, activity and clinical potential. Mil Med Res 8:48.
53. Nagakubo S, Nishino K, Hirata T, Yamaguchi A. 2002. The putative response regulator BaeR stimulates multidrug resistance of Escherichia coli via a novel multidrug exporter system, MdtABC. J Bacteriol 184:4161–4167.
54. Freeman ZN, Dorus S, Waterfield NR. 2013. The KdpD/KdpE two-component system: integrating K^+ homeostasis and virulence. PLoS Pathog 9:e1003201.
55. Chollet R, Chevalier J, Bollet C, Pages J-M, Davin-Regli A. 2004. RamA is an alternate activator of the multidrug resistance cascade in Enterobacter aerogenes. Antimicrob Agents Chemother 48:2518–2523.
56. Roberts MC. 2005. Update on acquired tetracycline resistance genes. FEMS Microbiol Lett 245:195–203.
57. Baysarowich J, Koteva K, Hughes DW, Ejim L, Griffiths E, Zhang K, Junop M, Wright GD. 2008. Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. Proc Natl Acad Sci 105:4886–4891.
58. Gao H, Zhang Y, Yang L, Liu X, Guo Z, Tan Y, Han Y, Huang X, Zhou D, Yang R. 2011. Regulatory effects of cAMP receptor protein (CRP) on porin genes and its own gene in Yersinia pestis. BMC Microbiol 11:40.
59. Geng H, Jiang R. 2015. cAMP receptor protein (CRP)-mediated resistance/tolerance in bacteria: mechanism and utilization in biotechnology. Appl Microbiol Biotechnol 99:4533–4543.
60. Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE. 2016. The Ecology of

Acidobacteria: Moving beyond Genes and Genomes. Front Microbiol 7.

61. Thornhill I, Loiselle S, Lind K, Ophof D. 2016. The Citizen Science Opportunity for Researchers and Agencies. BioScience 66:720–721.
62. Ebinger A, Fischer S, Höper D. 2021. A theoretical and generalized approach for the assessment of the sample-specific limit of detection for clinical metagenomics. Comput Struct Biotechnol J 19:732–742.
63. Flores CO, Meyer JR, Valverde S, Farr L, Weitz JS. 2011. Statistical structure of host-phage interactions. Proc Natl Acad Sci 108:E288–E297.
64. Haerter JO, Mitarai N, Sneppen K. 2014. Phage and bacteria support mutual diversity in a narrowing staircase of coexistence. 11. ISME J 8:2317–2326.
65. Arya S, Todman H, Baker M, Hooton S, Millard A, Kreft J-U, Hobman JL, Stekel DJ. 2020. A generalised model for generalised transduction: the importance of co-evolution and stochasticity in phage mediated antimicrobial resistance transfer. FEMS Microbiol Ecol 96:fiaa100.
66. Batinovic, Wassef, Knowler, Rice, Stanton, Rose, Tucci, Nittami, Vinh, Drummond, Sobey, Chan, Seviour, Petrovski, Franks. 2019. Bacteriophages in Natural and Artificial Environments. Pathogens 8:100.
67. Borin JM, Avrani S, Barrick JE, Petrie KL, Meyer JR. 2021. Coevolutionary phage training leads to greater bacterial suppression and delays the evolution of phage resistance. Proc Natl Acad Sci 118:e2104592118.
68. Chopyk J, Nasko DJ, Sakowski EG. 2021. Bacteriophage and Viral Ecology in the “Omics Age,” p. 49–87. *In* Studies in Viral Ecology. John Wiley & Sons, Ltd.
69. Mohan Raj JR, Karunasagar I. 2019. Phages amid antimicrobial resistance. Crit Rev Microbiol 45:701–711.
70. North OI, Brown ED. 2021. Phage–antibiotic combinations: a promising approach to constrain resistance evolution in bacteria. Ann N Y Acad Sci 1496:23–34.

71. Am O, M J. 2013. Phage therapy: Should bacterial resistance to phages be a concern, even in the long run? Bacteriophage. Bacteriophage. https://pubmed.ncbi.nlm.nih.gov/23819105/. Retrieved 1 September 2020.
72. Chan BK, Sistrom M, Wertz JE, Kortright KE, Narayan D, Turner PE. 2016. Phage selection restores antibiotic sensitivity in MDR Pseudomonas aeruginosa. 1. Sci Rep 6:26717.
73. Abubakar S, Suleiman BH, Abbagana BA, Mustafa IA, Musa IA. 2016. Novel Uses of Bacteriophages in the Treatment of Human Infections and Antibiotic Resistance. 3. Am J Biosci 4:34.
74. Chang Y. 2020. Bacteriophage-Derived Endolysins Applied as Potent Biocontrol Agents to Enhance Food Safety. Microorganisms 8.
75. Endersen L, O'Mahony J, Hill C, Ross RP, McAuliffe O, Coffey A. 2014. Phage Therapy in the Food Industry. Annu Rev Food Sci Technol 5:327–349.
76. Abdelrahman F, Easwaran M, Daramola OI, Ragab S, Lynch S, Oduselu TJ, Khan FM, Ayobami A, Adnan F, Torrents E, Sanmukh S, El-Shibiny A. 2021. Phage-Encoded Endolysins. Antibiotics 10:124.
77. Gervasi T, Horn N, Wegmann U, Dugo G, Narbad A, Mayer MJ. 2014. Expression and delivery of an endolysin to combat Clostridium perfringens. Appl Microbiol Biotechnol 98:2495–2505.
78. Hathaway H, Milo S, Sutton JM, Jenkins TA. 2017. Recent advances in therapeutic delivery systems of bacteriophage and bacteriophage-encoded endolysins. Ther Deliv 8:543–556.
79. Nelson DC, Schmelcher M, Rodriguez-Rubio L, Klumpp J, Pritchard DG, Dong S, Donovan DM. 2012. Chapter 7 - Endolysins as Antimicrobials, p. 299–365. *In* Łobocka, M, Szybalski, W (eds.), Advances in Virus Research. Academic Press.
80. Oliveira H, Azeredo J, Lavigne R, Kluskens LeonD. 2012. Bacteriophage endolysins as a response to emerging foodborne pathogens. Trends Food Sci Technol 28:103–115.
81. Bacteriophage endolysins as novel antimicrobials | Future Microbiology.

https://www.futuremedicine.com/doi/10.2217/fmb.12.97. Retrieved 8 August 2020.

82. Roux S, Enault F, Hurwitz BL, Sullivan MB. 2015. VirSorter: mining viral signal from microbial genomic data. PeerJ 3:e985.
83. Kieft K, Zhou Z, Anantharaman K. 2020. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome 8:90.
84. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A. 2021. Pfam: The protein families database in 2021. Nucleic Acids Res 49:D412–D419.
85. Benler S, Koonin EV. 2021. Fishing for phages in metagenomes: what do we catch, what do we miss? Curr Opin Virol 49:142–150.
86. Coclet C, Roux S. 2021. Global overview and major challenges of host prediction methods for uncultivated phages. Curr Opin Virol 49:117–126.
87. Wang W, Ren J, Tang K, Dart E, Ignacio-Espinoza JC, Fuhrman JA, Braun J, Sun F, Ahlgren NA. 2020. A network-based integrated framework for predicting virus–prokaryote interactions. NAR Genomics Bioinforma 2.
88. Holt A, Cahill J, Ramsey J, Martin C, O'Leary C, Moreland R, Maddox LT, Galbadage T, Sharan R, Sule P, Cirillo JD, Young R. 2022. Phage-Encoded Cationic Antimicrobial Peptide Required for Lysis. J Bacteriol 204:e00214-21.
89. InterPro protein families and domains database: 20 years on | Nucleic Acids Research | Oxford Academic. https://academic.oup.com/nar/article/49/D1/D344/5958491?login=false. Retrieved 20 October 2022.
90. Collins LV, Kristian SA, Weidenmaier C, Faigle M, van Kessel KPM, van Strijp JAG, Götz F, Neumeister B, Peschel A. 2002. Staphylococcus aureus Strains Lacking d-Alanine Modifications of Teichoic Acids Are Highly Susceptible to Human Neutrophil Killing and Are Virulence Attenuated in Mice. J Infect Dis 186:214–219.

91. Vadyvaloo V, Arous S, Gravesen A, Héchard Y, Chauhan-Haubrock R, Hastings JW, Rautenbach M 2004. Cell-surface alterations in class IIa bacteriocin-resistant Listeria monocytogenes strains. Microbiology 150:3025–3033.
92. Walter J, Loach DM, Alqumber M, Rockel C, Hermann C, Pfitzenmaier M, Tannock GW. 2007. d-Alanyl ester depletion of teichoic acids in Lactobacillus reuteri 100-23 results in impaired colonization of the mouse gastrointestinal tract. Environ Microbiol 9:1750–1760.
93. Smith CA, Bhattacharya M, Toth M, Stewart NK, Vakulenko SB. 2017. Aminoglycoside resistance profile and structural architecture of the aminoglycoside acetyltransferase AAC(6')-Im. Microb Cell 4:402–410.
94. Brescia CC, Kaw MK, Sledjeski DD. 2004. The DNA Binding Protein H-NS Binds to and Alters the Stability of RNA in vitro and in vivo. J Mol Biol 339:505–514.
95. Schuldiner S, Granot D, Mordoch SS, Ninio S, Rotem D, Soskin M, Tate CG, Yerushalmi H. 2001. Small is mighty: EmrE, a multidrug transporter as an experimental paradigm. News Physiol Sci Int J Physiol Prod Jointly Int Union Physiol Sci Am Physiol Soc 16:130–134.
96. Orelle C, Mathieu K, Jault J-M. 2019. Multidrug ABC transporters in bacteria. Res Microbiol 170:381–391.
97. Fernández-Ruiz I, Coutinho FH, Rodriguez-Valera F. 2018. Thousands of Novel Endolysins Discovered in Uncultured Phage Genomes. Front Microbiol 9:1033.
98. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R, Turner D, Sullivan MB. 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol 37:632–639.
99. 2006. Preventing the Forward Contamination of Mars. National Academies Press, Washington, D.C. http://www.nap.edu/catalog/11381. Retrieved 25 October 2022.
100. Hendrickson R, Urbaniak C, Minich JJ, Aronson HS, Martino C, Stepanauskas R, Knight R, Venkateswaran K. 2021. Clean room microbiome complexity impacts planetary

protection bioburden. Microbiome 9:238.

101. NASA Technical Handbook: Handbook For The Microbial Examination of Space Hardware. National Aeronautics and Space Administration Approved: 08-17-2010 Washington, DC 20546-0001.

102. Trimmomatic: a flexible trimmer for Illumina sequence data | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/30/15/2114/2390096?login=true. Retrieved 25 October 2022.

103. Vasilinetc I, Prjibelski AD, Gurevich A, Korobeynikov A, Pevzner PA. 2015. Assembling short reads from jumping libraries with large insert sizes. Bioinformatics 31:3262–3268.

104. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads, p. 158–170. *In* Deng, M, Jiang, R, Sun, F, Zhang, X (eds.), Research in Computational Molecular Biology. Springer, Berlin, Heidelberg.

105. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl 25:1754–1760.

106. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

107. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/36/6/1925/5626182. Retrieved 25 October 2022.

108. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319.

109. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glöckner FO. 2014. The SILVA and "All-species Living Tree Project (LTP)"

taxonomic frameworks. Nucleic Acids Res 42:D643–D648.

110. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

111. Yoon B-J. 2009. Hidden Markov Models and their Applications in Biological Sequence Analysis. Curr Genomics 10:402–415.

112. Wright, Alonna. 2022. NASA GESAM JPL Results Tables. figshare.

113. Taheri N, Ardebili A, Amouzandeh-Nobaveh A, Ghaznavi-Rad E. 2016. Frequency of Antiseptic Resistance Among Staphylococcus aureus and Coagulase-Negative Staphylococci Isolated From a University Hospital in Central Iran. Oman Med J 31:426–432.

114. Gevers D, Vandepoele K, Simillion C, Van de Peer Y. 2004. Gene duplication and biased functional retention of paralogs in bacterial genomes. Trends Microbiol 12:148–154.

Supplementary

Sample Metadata

Chapters 1 and 2

Table S-Metadata 1A: Metadata Collected from BYP Owners in Community Science Project Survey

Sample_ID	House_ID	AntibioticUsage	IndoorOutdoor	ZipCode	#Chicks	#pullets	#adults	Total
A3_1	A3	Antibiotic	Outdoor	94087	0	0	7	7
A3_2	A3	Antibiotic	Indoor	94087	0	0	7	7
A5_1	A5	Antibiotic	Outdoor	95747	0	0	35	35
A5_2	A5	Antibiotic	Indoor	95747	0	0	35	35
A7_1	A7	Antibiotic	Outdoor	94803	0	0	17	17
A7_2	A7	Antibiotic	Indoor	94803	0	0	17	17
A8_1	A8	Antibiotic	Outdoor	94063	0	0	5	5
A8_2	A8	Antibiotic	Indoor	94063	0	0	5	5

Table S-Metadata 1A: Metadata Collected from BYP Owners in Community Science Project Survey

Sample_ID	House_ID	AntibioticUsage	IndoorOutdoor	ZipCode	#Chicks	#pullets	#adults	Total
A9_1	A9	Antibiotic	Outdoor	94301	0	0	7	7
A9_2	A9	Antibiotic	Indoor	94301	0	0	7	7
AF3_1	AF3	Antibiotic Free	Outdoor	95148	0	0	3	3
AF3_2	AF3	Antibiotic Free	Indoor	95148	0	0	3	3
AF5_1	AF5	Antibiotic Free	Outdoor	95618	0	0	8	8
AF5_2	AF5	Antibiotic Free	Indoor	95618	0	0	8	8
AF7_1	AF7	Antibiotic Free	Outdoor	94591	4	0	11	15
AF7_2	AF7	Antibiotic Free	Indoor	94591	4	0	11	15
AF8_1	AF8	Antibiotic Free	Outdoor	94061	0	0	3	3
AF8_2	AF8	Antibiotic Free	Indoor	94061	0	0	3	3

Table S-Metadata 1B: Metadata Collected from BYP Owners in Community Science Project Survey

Sample_ID	Clean?	Disenfectants?	Bathe?	soap	other animals	Vet ?	Antibio?	Which one?	completed tx?
A3_1	Once every six months	None	Yes	vet soap		Yes	Yes	Baytril, Clavamox and there might have been a third	Yes
A3_2	Once every six months	None	Yes	vet soap		Yes	Yes	Baytril, Clavamox and there might have been a third	Yes
A5_1	Once every six months		No			Yes	Yes	amoxicillin	Yes
A5_2	Once every six months		No			Yes	Yes	amoxicillin	Yes
A7_1	Once a month	None	No	Non e		No	Yes	Tylon 50 injectable	Yes

Table S-Metadata 1B: Metadata Collected from BYP Owners in Community Science Project Survey

Sample_ID	Clean?	Disenfectants?	Bathe?	soap	other animals	Vet ?	Antibio?	Which one?	completed tx?
A7_2	Once a month	None	No	Non e		No	Yes	Tylon 50 injectable	Yes
A8_1	Once a year	None	No			Yes	Yes	enfloxacin, ciprofloxacin	Yes
A8_2	Once a year	None	No			Yes	Yes	enfloxacin, ciprofloxacin	Yes
A9_1	top to bottom every 6mos., daily coop poop scoop, change coop bedding 4x yr, change run straw 4-6x yr	vinegar	Yes	dish soap		Yes	Yes	Amoxicillin	Yes

Table S-Metadata 1B: Metadata Collected from BYP Owners in Community Science Project Survey

Sample_ID	Clean?	Disenfectants?	Bathe?	soap	other animals	Vet ?	Antibio?	Which one?	completed tx?
A9_2	top to bottom every 6mos., daily coop poop scoop, change coop bedding 4x yr, change run straw 4-6x yr	vinegar	Yes	dish soap		Yes	Yes	Amoxicillin	Yes
AF3_1	Once a week	None	No		Cats	No	No		
AF3_2	Once a week	None	No		Cats	No	No		
AF5_1	Once every six months	None	No			No	No		

Table S-Metadata 1B: Metadata Collected from BYP Owners in Community Science Project Survey

Sample_ID	Clean?	Disenfectants?	Bathe?	soap	other animals	Vet ?	Antibio?	Which one?	completed tx?
AF5_2	Once every six months	None	No			No	No		
AF7_1	Once a month	None	No		Dog	No	No		
AF7_2	Once a month	None	No		Dog	No	No		
AF8_1	Once a year	Bleach, White wash	No		Cat	No	No		
AF8_2	Once a year	Bleach, White wash	No		Cat	No	No		

Chapter 3

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52039wB5	MSL 321.1	MSL	2012	6/28/2010	GCA_023712705.1	SAMN27922062	Micrococcus luteus
52039wC10	AMY 19.1.2	Phoenix	2007	6/5/2007	GCA_023715845.1	SAMN28071792	Fictibacillus phosphorivorans
52039wC5	MSL 314.1	MSL	2012	6/28/2010	GCA_023712685.1	SAMN27922060	Staphylococcus hominis
52039wD1	197.1	MSL	2012	12/12/2008	GCA_023715005.1	SAMN27921941	Staphylococcus warneri
52039wD10	AMY 6.1.1	Phoenix	2007	5/20/2007	GCA_023714845.1	SAMN27921948	Bacillus licheniformis
52039wD11	MSL 107	MSL	2012	10/8/2008	GCA_023712865.1	SAMN27922050	Staphylococcus epidermidis
52039wE3	MER135A	MER	2003	NA	NA	NA	NA
52039wF10	AMY 32.2	Phoenix	2007	6/1/2007	NA	NA	NA
52039wF8	AMY 28.1.2	Phoenix	2007	6/1/2007	GCA_023714745.1	SAMN27921954	Bacillus safensis
52039wH10	MSL 016.1	MSL	2012	3/27/2008	GCA_023713005.1	SAMN27922044	Staphylococcus epidermidis
52039wH11	TPS 14-3.1	Viking	1975	4/26/2006	GCA_023712185.1	SAMN27922091	Brevibacillus borstelensis
52040wF2	P97	Odyssey	2001	NA	GCA_023712445.1	SAMN27922075	Curtobacterium sp. P97
52040wH6	PF4F.2.1	Phoenix	2007	8/28/2007	GCA_023712325.1	SAMN27922081	Cytobacillus oceanisediminis
52040wD1	P7	Odyssey	2001	NA	GCA_023715365.1	SAMN28071810	Georgenia satyanarayanai
52039wA1	MER 33	MER	2003	4/9/2003	GCA_023714285.1	SAMN27921980	Mesobacillus sp. MER 33
52039wA2	MER 53-1	MER	2003	4/10/2003	GCA_023714185.1	SAMN27921985	Priestia megaterium

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52039wA6	MSL 185.1	MSL	2012	12/8/2008	GCA_023712765.1	SAMN27922056	Caldibacillus thermoamylovorans
52039wB1	MER 37	MER	2003	4/10/2003	GCA_023715585.1	SAMN28071799	Sporosarcina luteola
52039wB4	MER 128	MER	2003	5/28/2004	GCA_023713925.1	SAMN27921999	Priestia megaterium
52039wB6	MSL 200.1	MSL	2012	12/4/2008	GCA_023715925.1	SAMN27922058	Priestia megaterium
52039wB7	MER 46	MER	2003	4/9/2003	GCA_023714225.1	SAMN27921983	Mesobacillus sp. MER 48
52039wC1	MER 13	MER	2003	4/9/2003	GCA_023714325.1	SAMN27921978	Priestia koreensis
52039wC4	MER 145	MER	2003	5/28/2004	GCA_023715685.1	SAMN28071793	Priestia aryabhattai
52039wD4	MER 153	MER	2003	5/28/2004	GCA_023713865.1	SAMN27922002	Priestia aryabhattai
52039wD8	AMY 5.2	Phoenix	2007	5/11/2007	NA	NA	NA
52039wE2	251.1	MSL	2012	1/5/2009	GCA_023714945.1	SAMN27921944	Bacillus altitudinis
52039wE4	MER 110	MER	2003	7/18/2014	GCA_023713965.1	SAMN27921995	Priestia megaterium
52039wE5	FAIRING 12A-4	Phoenix	2007	1/23/2007	GCA_023714665.1	SAMN27921960	Heyndrickxia oleronia
52039wE7	MER 20	MER	2003	4/9/2003	GCA_023715615.1	SAMN28071797	Sporosarcina aquimarina
52039wE9	P107	Odyssey	2001	NA	GCA_023712365.1	SAMN27922077	Priestia aryabhattai
52039wF1	206.2.2	MSL	2012	12/27/2008	GCA_023714985.1	SAMN27921942	Bacillus amyloliquefaciens
52039wF4	MER 65	MER	2003	3/13/2003	GCA_023714125.1	SAMN27921988	Paenibacillus sp. MER 78
52039wF6	MER 74	MER	2003	3/11/2003	GCA_023714065.1	SAMN27921990	Bacillus subtilis

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52039wF7	MER 6	MER	2003	4/9/2003	GCA_023714145.1	SAMN27921986	Peribacillus frigoritolerans
52039wF9	P121	Odyssey	2001	NA	GCA_023712385.1	SAMN27922079	Priestia megaterium
52039wG8	AMY 5.1.1	Phoenix	2007	5/11/2007	GCA_023715565.1	SAMN28071800	Cytobacillus firmus
52039wG9	CFPSW 5.3	MSL	2012	4/10/2008	GCA_023714715.1	SAMN27921957	Cytobacillus oceanisediminis
52039wH3	MER 100	MER	2003	4/20/2004	GCA_023714025.1	SAMN27921992	Niallia taxi
52039wH8	PF 3F.1.2	Phoenix	2007	8/20/2007	NA	NA	NA
52040wA11	MER TA 168	MER	2003	7/18/2013	GCA_023713305.1	SAMN27922031	Psychrobacillus sp. MER TA 171
52040wA2	TA 76	MER	2003	2/17/2004	GCA_023715505.1	SAMN28071802	Neobacillus cucumis
52040wA7	MER TA 137-5	MER	2003	7/3/2013	GCA_023713265.1	SAMN27922029	Oceanobacillus profundus
52040wB1	AMY 7.1	Phoenix	2007	5/19/2007	GCA_023714855.1	SAMN27921949	Priestia flexa
52040wB11	MER TA 176	MER	2003	7/17/2013	GCA_023713585.1	SAMN27922019	Paenibacillus sp. MER TA 81-3
52040wB6	MER TA 106	MER	2003	7/30/2013	GCA_023713415.1	SAMN27922023	Streptococcus oralis
52040wB8	KSC 645	MSL	2012	NA	NA	NA	NA
52040wC2	MER 54.2	MER	2003	7/13/2012	NA	NA	NA
52040wC6	MER TA 87	MER	2003	7/30/2013	GCA_023715185.1	SAMN28071819	Planococcus sp. MERTA32b
52040wD4	MER TA 138-2	MER	2003	7/17/2013	GCA_023715205.1	SAMN28071818	Mesobacillus maritimus
52040wD6	AMY 13.1.2	Phoenix	2007	6/1/2007	GCA_023714795.1	SAMN27921951	Priestia megaterium

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52040wD8	KSC 657	MSL	2012	NA	GCA_023714385.1	SAMN27921975	Niallia circulans
52040wE6	AMY 15.2	Phoenix	2007	6/1/2007	GCA_023714805.1	SAMN27921952	Cytobacillus sp. AMY 15.2
52040wE8	KSC 351	MSL	2012	NA	GCA_023714505.1	SAMN27921967	Priestia aryabhattai
52040wE9	34.1	MSL	2012	5/14/2008	NA	NA	NA
52040wF10	MER 170	MER	2003	6/2/2014	GCA_023713705.1	SAMN27922009	Priestia megaterium
52040wF5	TA 149	MER	2003	7/18/2013	GCA_023712225.1	SAMN27922086	Alkalihalobacillus clausii
52040wF7	MER TA 17	MER	2003	8/9/2013	GCA_023713235.1	SAMN27922032	Ureibacillus chungkukjangi
52040wF8	KSC 283	MSL	2012	NA	GCA_023714565.1	SAMN27921965	Priestia megaterium
52040wF9	128.1.2	MSL	2012	NA	GCA_023715265.1	SAMN28071815	Mesobacillus maritimus
52040wG1	TA 170	MER	2003	7/18/2013	GCA_023715255.1	SAMN28071816	Mesobacillus subterraneus
52040wG11	MSL 179.1	MSL	2012	12/4/2008	GCA_023712805.1	SAMN27922055	Heyndrickxia oleronia
52040wG6	PF24B.2	Phoenix	2007	8/24/2007	NA	NA	NA
52040wH11	P67	Odyssey	2001	7/18/2014	GCA_023712515.1	SAMN27922071	Sutcliffiella horikoshii
52040wH2	MER TA 32b	MER	2003	8/12/2013	GCA_023715415.1	SAMN28071808	Neobacillus niacini
52040wH4	FAIRING 3B-1.2	Phoenix	2007	8/19/2013	GCA_023714705.1	SAMN27921958	Bacillus velezensis
52040wH7	KSC 432	MSL	2012	NA	GCA_023714455.1	SAMN27921971	Weizmannia ginsengihumi
52040wH9	41.1	MSL	2012	NA	GCA_023715085.1	SAMN27921937	Brevibacillus invocatus

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52041	MPF 8	Mars Pathfinder	1996	7/17/2013	GCA_023713165.1	SAMN27922035	Bacillus licheniformis
52045	258.1A	MSL	2012	NA	GCA_023714885.1	SAMN27921945	Lederbergia lenta
52052	MPF 2	Mars Pathfinder	1996	7/17/2013	GCA_023713085.1	SAMN27922042	Bacillus licheniformis
52039wA10	P30	Odyssey	2001	NA	GCA_023715885.1	SAMN28071785	Microbacterium hydrocarbonoxydans
52039wA11	MSL 060.1.1	MSL	2012	7/21/2008	GCA_023712925.1	SAMN27922049	Bacillus atrophaeus
52039wA12	TPS 11-9.1	Viking	1975	8/24/2006	GCA_023715945.1	SAMN27922090	Caldibacillus thermoamylovorans
52039wA3	MER 132	MER	2003	5/26/2004	GCA_023713815.1	SAMN27922001	Bacillus cereus
52039wA4	MER 99-2	MER	2003	4/21/2004	GCA_023715445.1	SAMN28071806	Fictibacillus nanhaiensis
52039wA5	MSL 359	MSL	2012	5/6/2010	GCA_023712595.1	SAMN27922065	Kocuria rosea
52039wA7	MER 78	MER	2003	3/11/2003	GCA_023715535.1	SAMN28071801	Sporosarcina luteola
52039wA8	MER 107	MER	2003	4/21/2004	GCA_023715865.1	SAMN28071787	Neobacillus mesonae
52039wA9	P10	Odyssey	2001	NA	GCA_023715745.1	SAMN28071788	Microbacterium oleivorans
52039wB10	AMY 19.1.2 vial 1	Phoenix	2007	6/5/2007	GCA_023715765.1	SAMN28071789	Fictibacillus phosphorivorans
52039wB11	MSL 140.1	MSL	2012	11/8/2008	GCA_023712885.1	SAMN27922051	Brevibacillus borstelensis
52039wB12	MPF 24	Mars Pathfinder	1996	5/6/2004	GCA_023713145.1	SAMN27922037	Bacillus subtilis

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52039wB2	MER 36	MER	2003	4/10/2003	GCA_023714245.1	SAMN27921981	Paenibacillus lutimineralis
52039wB3	MER 116	MER	2003	4/20/2004	GCA_023713935.1	SAMN27921997	Bacillus cereus
52039wB8	MER 196A	MER	2003	3/30/2004	GCA_023713605.1	SAMN27922014	Streptomyces thermoviolaceus
52039wB9	P106	Odyssey	2001	7/18/2004	GCA_023715705.1	SAMN28071791	Alkalihalobacillus oceani
52039wC11	TPS 8-13.1	Viking	1975	9/20/2006	GCA_023712145.1	SAMN27922088	Brevibacillus borstelensis
52039wC12	AMY 2.1.4	Phoenix	2007	8/31/2007	GCA_023714915.1	SAMN27921947	Staphylococcus equorum
52039wC2	MER 26	MER	2003	4/9/2003	GCA_023714295.1	SAMN27921979	Paenibacillus polysaccharolyticus
52039wC3	MER 156	MER	2003	5/26/2004	GCA_023713845.1	SAMN27922003	Bacillus subtilis
52039wC6	MSL 160.2	MSL	2012	11/12/2008	GCA_023712905.1	SAMN27922052	Bacillus velezensis
52039wC7	MER 89	MER	2003	4/3/2003	GCA_023713405.1	SAMN27922022	Cytobacillus oceanisediminis
52039wC8	P25	Odyssey	2001	NA	GCA_023712565.1	SAMN27922068	Microbacterium enclense
52039wC9	P83	Odyssey	2001	NA	GCA_023712485.1	SAMN27922073	Staphylococcus capitis
52039wD12	MPF 38	Mars Pathfinder	1996	5/6/2004	GCA_023713125.1	SAMN27922038	Bacillus paralicheniformis
52039wD2	183.1	MSL	2012	12/8/2008	GCA_023715065.1	SAMN27921940	Caldibacillus thermoamylovorans
52039wD3	MER 112	MER	2003	4/20/2004	GCA_023714045.1	SAMN27921996	Bacillus altitudinis

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52039wD5	FAIRING 10M-2.2	Phoenix	2007	1/29/2007	GCA_023715645.1	SAMN28071794	Bacillus cytotoxicus
52039wD6	MSL 225.1.2	MSL	2012	1/15/2009	GCA_023715895.1	SAMN28071795	Domibacillus indicus
52039wD7	MER 10	MER	2003	NA	GCA_023714345.1	SAMN27921977	Bacillus safensis
52039wD9	P75	Odyssey	2001	NA	GCA_023712505.1	SAMN27922072	Staphylococcus capitis
52039wE1	MER 172A	MER	2003	NA	GCA_023713685.1	SAMN27922010	Streptomyces pseudogriseolus
52039wE10	P100	Odyssey	2001	NA	GCA_023712425.1	SAMN27922076	Rothia sp. P100
52039wE11	MSL 036.1	MSL	2012	5/14/2008	GCA_023712985.1	SAMN27922045	Brevibacillus invocatus
52039wE6	MER 189	Mars Pathfinder	1996	6/28/2004	GCA_023713625.1	SAMN27922012	Staphylococcus capitis
52039wE8	AMY 1.1.1	MER	2003	5/20/2007	GCA_023714905.1	SAMN27921946	Bhargavaea ginsengi
52039wF11	MSL 004.1.2	Phoenix	2007	3/17/2008	GCA_023713025.1	SAMN27922043	Rothia dentocariosa
52039wF2	236.1.1	MSL	2012	1/3/2009	GCA_023714965.1	SAMN27921943	Bacillus altitudinis
52039wF3	MER 108	MER	2003	4/20/2004	GCA_023713985.1	SAMN27921994	Terribacillus saccharophilus
52039wF5	FAIRING 19B-1.2	MER	2003	1/27/2007	GCA_023714585.1	SAMN27921961	Staphylococcus warneri
52039wG1	MER 166	Phoenix	2007	6/2/2014	GCA_023713745.1	SAMN27922007	Bacillus velezensis
52039wG10	MSL 058.1.2	MER	2003	10/3/2008	GCA_023712945.1	SAMN27922048	Metabacillus litoralis
52039wG11	P112	MSL	2012	NA	GCA_023712395.1	SAMN27922078	Cellulosimicrobium funkei
52039wG2	214.1.1	Odyssey	2001	1/3/2009	GCA_023715605.1	SAMN28071798	Alkalihalobacillus oceani

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52039wG3	MER 73	MER	2003	3/13/2003	GCA_023714085.1	SAMN27921989	Cytobacillus horneckiae
52039wG4	MSL 259.1	MER	2003	4/6/2008	GCA_023712745.1	SAMN27922059	Staphylococcus lugdunensis
52039wG5	MSL 172.1.2	MSL	2012	12/2/2008	GCA_023712845.1	SAMN27922053	Bacillus safensis
52039wG6	MER 82	MSL	2012	4/17/2003	GCA_023714095.1	SAMN27921991	Paenibacillus sp. MER 99-2
52039wG7	MER 47	MER	2003	4/9/2003	GCA_023715465.1	SAMN28071805	Paenibacillus macerans
52039wH1	MER 118	MER	2003	4/20/2004	GCA_023713905.1	SAMN27921998	Bacillus cereus
52039wH2	MER 193	MER	2003	6/28/2004	GCA_023713645.1	SAMN27922013	Exiguobacterium sp. MER 193
52039wH4	MSL 316.2	MER	2003	6/28/2010	GCA_023712675.1	SAMN27922061	Staphylococcus warneri
52039wH5	MSL 173.2.2	MSL	2012	12/2/2008	GCA_023712815.1	SAMN27922054	Peribacillus simplex
52039wH6	MER 9	MSL	2012	4/9/2003	GCA_023715155.1	SAMN28071820	Neobacillus mesonae
52039wH7	MER 101	MER	2003	7/3/2012	GCA_023714005.1	SAMN27921993	Paenibacillus illinoisensis
52039wH9	PF3F.2	MER	2003	8/20/2007	GCA_023712345.1	SAMN27922080	Priestia endophytica
52040wA1	MER TA 114	Phoenix	2007	8/1/2013	GCA_023713435.1	SAMN27922026	Bacillus subtilis
52040wA10	68.1	MER	2003	10/3/2008	GCA_023715035.1	SAMN27921938	Bacillus safensis
52040wA12	P42	Odyssey	2001	7/18/2004	GCA_023712545.1	SAMN27922070	Alkalihalophilus marmarensis
52040wA3	TA 28	Odyssey	2001	2/25/2004	GCA_023715525.1	SAMN28071803	Kocuria rosea

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52040wA4	MER TA 138-1	MER	2003	6/27/2013	GCA_023713525.1	SAMN27922016	Kocuria palustris
52040wA5	FAIRING W8B-1	MER	2003	8/19/2013	GCA_023714625.1	SAMN27921962	Stenotrophomonas maltophilia
52040wA6	MER TA 82	Phoenix	2007	7/23/2014	GCA_023715725.1	SAMN28071790	Paenibacillus pasadenensis
52040wA8	MSL 3003	MER	2003	7/21/2014	GCA_023715825.1	SAMN27922066	Staphylococcus xylosus
52040wA9	KSC 422	MSL	2012	NA	NA	NA	NA
52040wB10	MER TA 170	MER	2003	7/18/2013	GCA_023713505.1	SAMN27922018	Staphylococcus warneri
52040wB12	KSC 114	MER	2003	NA	GCA_023714595.1	SAMN27921963	Bacillus altitudinis
52040wB2	MER 50.2	MER	2003	7/3/2012	GCA_023714395.1	SAMN27921976	Niallia sp. MER 6
52040wB3	MER 157	MER	2003	6/15/2004	GCA_023713805.1	SAMN27922004	Paenibacillus elgii
52040wB4	MER TA 111	MER	2003	7/17/2013	GCA_023713565.1	SAMN27922015	Rummeliibacillus stabekisii
52040wB5	MPF 76A	MER	2003	5/11/2004	GCA_023713065.1	SAMN27922041	Bacillus licheniformis
52040wB7	MER TA 35	Mars Pathfinder	1996	7/18/2013	NA	NA	NA
52040wB9	KSC 386	MER	2003	NA	GCA_023714485.1	SAMN27921968	Bacillus subtilis
52040wC1	AMY 31.2	MSL	2012	6/1/2007	GCA_023714675.1	SAMN27921955	Bacillus pumilus
52040wC10	MER 180	Phoenix	2007	6/3/2014	NA	NA	NA
52040wC11	MER TA 154	MER	2003	7/17/2013	GCA_023715485.1	SAMN28071804	Mesobacillus subterraneus

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52040wC12	MER 48	MER	2003	4/9/2003	GCA_023714195.1	SAMN27921984	Brevibacillus sp. MER 51
52040wC3	TA 104	MER	2003	2/6/2004	GCA_023712285.1	SAMN27922083	Bacillus subtilis
52040wC4	MER TA 13	MER	2003	6/27/2013	GCA_023713335.1	SAMN27922027	Alkalihalobacillus rhizosphaerae
52040wC5	MPF 57	MER	2003	5/19/2004	GCA_023713105.1	SAMN27922039	Bacillus intestinalis
52040wC7	MER TA 139-2	Mars Pathfinder	1996	7/18/2013	GCA_023715405.1	SAMN28071807	Solibacillus isronensis
52040wC8	KSC 640	MER	2003	NA	GCA_023714365.1	SAMN27921973	Bacillus pumilus
52040wC9	KSC 155	MSL	2012	NA	GCA_023714545.1	SAMN27921964	Bacillus safensis
52040wD10	MER TA 171	MSL	2012	7/18/2013	GCA_023713225.1	SAMN27922033	Metabacillus halosaccharovorans
52040wD11	154.2	MER	2003	NA	GCA_023715025.1	SAMN27921939	Bacillus subtilis
52039wB2	MER 36	MER	2003	4/10/2003	GCA_023714245.1	SAMN27921981	Paenibacillus lutimineralis
52040wD2	MSL 047.1	MER	2003	5/14/2008	GCA_023712965.1	SAMN27922046	Sphingopyxis alaskensis
52040wD3	TA 29	MSL	2012	7/30/2012	GCA_023715345.1	SAMN28071811	Paenibacillus camelliae
52040wD5	RA 14A.10	MER	2003	4/6/2007	GCA_023715325.1	SAMN28071812	Agromyces mediolanus
52040wD7	MER TA 18	Phoenix	2007	6/27/2013	GCA_023713445.1	SAMN27922020	Cytobacillus kochii
52040wD9	MER TA 181	MER	2003	7/17/2013	GCA_023715305.1	SAMN28071813	Metabacillus litoralis
52040wE1	ODYSSEY 48 V2	MER	2003	1/29/2007	GCA_023712625.1	SAMN27922067	Curtobacterium sp.

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
							ODYSSEY 48 V2
52040wE10	MER 165	MER	2003	6/2/2014	GCA_023713725.1	SAMN27922006	Bacillus licheniformis
52040wE11	AMY 24.1.1	MER	2003	6/1/2007	GCA_023714765.1	SAMN27921953	Bacillus safensis
52040wE12	MER 51	Phoenix	2007	4/9/2003	GCA_023715105.1	SAMN28071822	Hydrogenophaga intermedia
52040wE2	MSL 047.2	MER	2003	5/14/2008	GCA_023715805.1	SAMN27922047	Sphingomonas paucimobilis
52040wE3	TA 121-4	MSL	2012	3/8/2004	GCA_023712295.1	SAMN27922084	Micrococcus luteus
52040wE4	MER TA 86	MER	2003	7/30/2013	GCA_023715785.1	SAMN28071786	Neobacillus niacini
52040wE5	TA 33-2	MER	2003	7/23/2013	GCA_023712235.1	SAMN27922082	Carnobacterium inhibens
52040wE7	MER TA 97	MER	2003	8/9/2013	GCA_023715145.1	SAMN28071821	Paenibacillus glycanilyticus
52040wF1	P26	MER	2003	NA	GCA_023712585.1	SAMN27922069	Microbacterium sp. P26
52040wF11	KSC 339	Odyssey	2001	NA	GCA_023714515.1	SAMN27921966	Staphylococcus saprophyticus
52040wF3	TA 127	MSL	2012	7/2/2013	GCA_023712255.1	SAMN27922085	Micrococcus luteus
52040wF4	MER TA 48	MER	2003	7/23/2012	GCA_023713665.1	SAMN27922011	Paenibacillus sp. MER_180
52040wF6	AMY 11.1.2	MER	2003	5/31/2007	GCA_023714785.1	SAMN27921950	Staphylococcus warneri
52040wG10	MER 169	Phoenix	2007	6/2/2014	GCA_023713735.1	SAMN27922008	Bacillus velezensis

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52040wG2	P86	MER	2003	NA	GCA_023712465.1	SAMN27922074	Nocardioides sp. P86
52040wG3	MER TA 136-3-2	Odyssey	2001	7/2/2013	GCA_023713325.1	SAMN27922028	Cytobacillus oceanisediminis
52040wG4	FAIRING 4G-1.1	MER	2003	8/20/2013	GCA_023714645.1	SAMN27921959	Staphylococcus epidermidis
52040wG5	MER TA 81-3	Phoenix	2007	7/23/2013	GCA_023715385.1	SAMN28071809	Paenibacillus motobuensis
52040wG7	MER TA 112	MER	2003	8/15/2013	GCA_023713365.1	SAMN27922025	Bacillus halotolerans
52040wG8	KSC 591	MER	2003	NA	GCA_023714405.1	SAMN27921972	Bacillus subtilis
52040wG9	MER TA 152	MSL	2012	7/17/2013	GCA_023713535.1	SAMN27922017	Psychrobacillus sp. MER TA 17
52040wH1	TA 172	MER	2003	7/18/2013	GCA_023712125.1	SAMN27922087	Metabacillus idriensis
52040wH10	MER 161	MER	2003	6/2/2014	GCA_023713785.1	SAMN27922005	Bacillus licheniformis
52040wH3	MER TA 14	MER	2003	7/23/2013	GCA_023713275.1	SAMN27922030	Niallia sp. MER TA 168
52040wH5	MER TA 107	MER	2003	8/9/2013	GCA_023713375.1	SAMN27922024	Cytobacillus kochii
52040wH8	MSL 187.1	MSL	2012	12/8/2008	GCA_023712785.1	SAMN27922057	Paenibacillus lactis
52042	MPF 67	Mars Pathfinder	1996	8/21/2013	GCA_023713045.1	SAMN27922040	Bacillus licheniformis
52043	MER TA 38	MER	2003	7/24/2013	GCA_023713485.1	SAMN27922021	Staphylococcus capitis
52044	KSC 418	MSL	2012	NA	GCA_023714445.1	SAMN27921969	Paenibacillus cellulositrophicus

Table S-Metadata 2: Metadata of NASA JPL GESAM Bacterial Isolate Assemblies Submitted to NCBI SRA

barcode	microbesNG_ref	mission	mission_year	isolation_date	# Assembly	BioSample	Taxonomy
52047	MER 131	MER	2003	5/28/2004	GCA_023713885.1	SAMN27922000	Paenibacillus lautus
52048	MER 62	MER	2003	3/5/2003	GCA_023714155.1	SAMN27921987	Neobacillus sp. MER 74
52049	MSL 336.2	MSL	2012	6/28/2010	GCA_023712645.1	SAMN27922063	Ralstonia pickettii
52050	MSL 348	MSL	2012	4/20/2010	GCA_023712665.1	SAMN27922064	Cupriavidus pauculus
52054	P18	Odyssey	2001	7/18/2014	GCA_023715125.1	SAMN28071823	Mesobacillus subterraneus
52039wE12	MPF 19	Mars Pathfinder	1996	NA	GCA_023713175.1	SAMN27922036	Bacillus velezensis
52040wD12	MER 36	MER	2003	4/10/2003	GCA_023714245.1	SAMN27921981	Paenibacillus lutimineralis

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF01257	2Fe-2S_thioredx	Thioredoxin-like [2Fe-2S] ferredoxin	Oxidative Damage Resistance
TIGR00567	3mg	DNA-3-methyladenine glycosylase	Repair and Recombination
PF17864	AAA_lid_4	RuvB AAA lid domain	Repair and Recombination
PF05145	AbrB	Transition state regulatory protein AbrB	WTF Processes
PF14250	AbrB-like	AbrB-like transcriptional regulator	WTF Processes
PF17981	ADD_ATRX	Cysteine Rich ADD domain	WTF Processes
PF17980	ADD_DNMT3	Cysteine rich ADD domain in DNMT3	WTF Processes
TIGR02784	addA_alphas	double-strand break repair helicase AddA	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02785	addA_Gpos	helicase-exonuclease AddAB, AddA subunit	Repair and Recombination
TIGR02786	addB_alphas	double-strand break repair protein AddB	Repair and Recombination
TIGR02773	addB_Gpos	helicase-exonuclease AddAB, AddB subunit	Repair and Recombination
PF03352	Adenine_glyco	Methyladenine glycosylase	Repair and Recombination
PF09171	AGOG	N-glycosylase/DNA lyase	Repair and Recombination
TIGR00777	ahpD	alkylhydroperoxidase, AhpD family	Oxidative Damage Resistance
TIGR00778	ahpD_dom	alkylhydroperoxidase AhpD family core domain	Oxidative Damage Resistance
PF06029	AlkA_N	AlkA N-terminal domain	Repair and Recombination
TIGR00568	alkb	alkylated DNA repair protein AlkB	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02055	APS_reductase	adenylylsulfate reductase, thioredoxin dependent	Oxidative Damage Resistance
TIGR00432	arcsn_tRNA_tgt	tRNA-guanine(15) transglycosylase	WTF Processes
TIGR02691	arsC_pI258_fam	arsenate reductase (thioredoxin)	Oxidative Damage Resistance
PF09501	Bac_small_YrzI	Probable sporulation protein (Bac_small_yrzI)	Sporulation
TIGR00198	cat_per_HPI	catalase/peroxidase HPI	Oxidative Damage Resistance
PF18011	Catalase_C	C-terminal domain found in long catalases	Oxidative Damage Resistance
PF06628	Catalase-rel	Catalase-related immune-responsive	Oxidative Damage Resistance
PF03150	CCP_MauG	Di-haem cytochrome c peroxidase	Repair and Recombination
PF00313	CSD	Cold-shock' DNA-binding domain	Cold Shock

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF17876	CSD2	Cold shock domain	Cold Shock
TIGR02381	cspD	cold shock domain protein CspD	Cold Shock
PF00875	DNA_photolyase	DNA photolyase	Repair and Recombination
PF04244	DPRP	Deoxyribodipyrimidine photo-lyase-related protein	Repair and Recombination
PF01323	DSBA	DSBA-like thioredoxin domain	Repair and Recombination
PF09895	DUF2122	RecB-family nuclease (DUF2122)	Repair and Recombination
TIGR01413	Dyp_perox_fam	Dyp-type peroxidase family	Repair and Recombination
PF03441	FAD_binding_7	FAD binding domain of DNA photolyase	Repair and Recombination
PF01149	Fapy_DNA_glyco	Formamidopyrimidine-DNA glycosylase N-terminal domain	Repair and Recombination
PF02941	FeThRed_A	Ferredoxin thioredoxin reductase variable alpha chain	Oxidative Damage Resistance

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF02943	FeThRed_B	Ferredoxin thioredoxin reductase catalytic beta chain	Oxidative Damage Resistance
TIGR00577	fpg	DNA-formamidopyrimidine glycosylase	Repair and Recombination
PF03323	GerA	Bacillus/Clostridium GerA spore germination protein	Sporulation
PF10646	Germane	Sporulation and spore germination	Sporulation
TIGR03082	Gneg_AbrB_dup	membrane protein AbrB duplication	WTF Processes
TIGR02540	gpx7	putative glutathione peroxidase Gpx7	Oxidative Damage Resistance
PF06831	H2TH	Formamidopyrimidine-DNA glycosylase H2TH domain	Repair and Recombination
PF13749	HATPase_c_4	Putative ATP-dependent DNA helicase recG C-terminal	Repair and Recombination
PF05127	Helicase_RecD	Helicase	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF11408	Helicase_Sgs1	Sgs1 RecQ helicase	Repair and Recombination
PF00730	HhH-GPD	HhH-GPD superfamily base excision DNA repair protein	Repair and Recombination
PF00730	HhH-GPD	HhH-GPD superfamily base excision DNA repair protein	Repair and Recombination
PF07475	Hpr_kinase_C	HPr Serine kinase C-terminal domain	WTF Processes
PF02603	Hpr_kinase_N	HPr Serine kinase N terminus	WTF Processes
TIGR04352	HprK_rel_A	HprK-related kinase A	WTF Processes
TIGR04355	HprK_rel_B	HprK-related kinase B	WTF Processes
TIGR04274	hypoxanDNAglyco	DNA-deoxyinosine glycosylase	Repair and Recombination
PF14089	KbaA	KinB-signalling pathway activation in sporulation	Sporulation
PF02735	Ku	Ku70/Ku80 beta-barrel domain	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02772	Ku_bact	Ku protein	Repair and Recombination
PF03730	Ku_C	Ku70/Ku80 C-terminal arm	Repair and Recombination
PF03731	Ku_N	Ku70/Ku80 N-terminal alpha/beta domain	Repair and Recombination
PF08785	Ku_PK_bind	Ku C terminal domain like	Repair and Recombination
TIGR00578	ku70	ATP-dependent DNA helicase II, 70 kDa subunit (ku70)	Repair and Recombination
TIGR00498	lexA	repressor LexA	Repair and Recombination
PF01726	LexA_DNA_bind	LexA DNA binding domain	Repair and Recombination
PF13298	LigD_N	DNA polymerase Ligase (LigD)	Repair and Recombination
TIGR02777	LigD_PE_dom	DNA ligase D, 3'-phosphoesterase domain	Repair and Recombination
TIGR02778	ligD_pol	DNA ligase D, polymerase domain	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR01439	lp_hng_hel_AbrB	transcriptional regulator, AbrB family	WTF Processes
PF05067	Mn_catalase	Manganese containing catalase	Oxidative Damage Resistance
TIGR02070	mono_pep_trsgly	monofunctional biosynthetic peptidoglycan transglycosylase	WTF Processes
TIGR00401	msrA	peptide-methionine (S)-S-oxide reductase	Oxidative Damage Resistance
PF01624	MutS_I	MutS domain I	Repair and Recombination
PF05188	MutS_II	MutS domain II	Repair and Recombination
PF05192	MutS_III	MutS domain III	Repair and Recombination
PF05190	MutS_IV	MutS family domain IV	Repair and Recombination
PF00488	MutS_V	MutS domain V	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR01070	mutS1	DNA mismatch repair protein MutS	Repair and Recombination
TIGR01084	mutY	A/G-specific adenine glycosylase	Repair and Recombination
TIGR00588	ogg	8-oxoguanine DNA-glycosylase (ogg)	Repair and Recombination
PF07934	OGG_N	8-oxoguanine DNA glycosylase, N-terminal domain	Repair and Recombination
PF03419	Peptidase_U4	Sporulation factor SpoIIGA	Sporulation
TIGR04030	perox_Avi_7169	alkylhydroperoxidase domain protein, Avi_7169 family	Oxidative Damage Resistance
TIGR04169	perox_w_seleSAM	alkylhydroperoxidase/carboxymuconolactone decarboxylase family protein	Oxidative Damage Resistance
TIGR01926	peroxid_rel	uncharacterized peroxidase-related enzyme	Oxidative Damage Resistance
PF00141	peroxidase	Peroxidase	Oxidative Damage Resistance

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF01328	Peroxidase_2	Peroxidase, family 2	Oxidative Damage Resistance
PF16773	Phage_SSB	Lactococcus phage single-stranded DNA binding protein	Repair and Recombination
TIGR03556	photolyase_8HDF	deoxyribodipyrimidine photo-lyase, 8-HDF type	Repair and Recombination
TIGR00591	phr2	deoxyribodipyrimidine photolyase	Repair and Recombination
PF01625	PMSR	Peptide methionine sulfoxide reductase	Oxidative Damage Resistance
PF11565	PorB	Alpha helical Porin B	WTF Processes
PF02245	Pur_DNA_glyco	Methylpurine-DNA glycosylase (MPG)	Repair and Recombination
PF03013	Pyr_excise	Pyrimidine dimer DNA glycosylase	Repair and Recombination
TIGR00430	Q_tRNA_tgt	tRNA-guanine transglycosylase	WTF Processes

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR00608	radc	DNA repair protein RadC	Repair and Recombination
PF04002	RadC	RadC-like JAB domain	Repair and Recombination
PF00154	RecA	recA bacterial DNA recombination protein	Repair and Recombination
PF16786	RecA_dep_nuc	Recombination enhancement, RecA-dependent nuclease	Repair and Recombination
TIGR00609	recB	exodeoxyribonuclease V, beta subunit	Repair and Recombination
TIGR01450	recC	exodeoxyribonuclease V, gamma subunit	Repair and Recombination
PF17946	RecC_C	RecC C-terminal domain	Repair and Recombination
TIGR01447	recD	exodeoxyribonuclease V, alpha subunit	Repair and Recombination
TIGR01448	recD_rel	helicase, RecD/TraA family	Repair and Recombination
TIGR00611	recf	DNA replication and repair protein RecF	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR00643	recG	ATP-dependent DNA helicase RecG	Repair and Recombination
TIGR00643	recG	ATP-dependent DNA helicase RecG	Repair and Recombination
PF17190	RecG_N	RecG N-terminal helical domain	Repair and Recombination
PF17190	RecG_N	RecG N-terminal helical domain	Repair and Recombination
TIGR00644	recJ	single-stranded-DNA-specific exonuclease RecJ	Repair and Recombination
PF17768	RecJ_OB	RecJ OB domain	Repair and Recombination
TIGR00634	recN	DNA repair protein RecN	Repair and Recombination
TIGR00613	reco	DNA repair protein RecO	Repair and Recombination
PF02565	RecO_C	Recombination protein O C terminal	Repair and Recombination
PF11967	RecO_N	Recombination protein O N terminal	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF13114	RecO_N_2	RecO N terminal	Repair and Recombination
TIGR01389	recQ	ATP-dependent DNA helicase RecQ	Repair and Recombination
TIGR00614	recQ_fam	ATP-dependent DNA helicase, RecQ family	Repair and Recombination
PF06959	RecQ5	RecQ helicase protein-like 5 (RecQ5)	Repair and Recombination
PF16099	RMI1_C	Recq-mediated genome instability protein 1, C-terminal OB-fold	Repair and Recombination
PF08585	RMI1_N	RecQ mediated genome instability protein	Repair and Recombination
PF16100	RMI2	RecQ-mediated genome instability protein 2	Repair and Recombination
TIGR00084	ruvA	Holliday junction DNA helicase RuvA	Repair and Recombination
PF07499	RuvA_C	RuvA, C-terminal domain	Repair and Recombination
PF07499	RuvA_C	RuvA, C-terminal domain	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF01330	RuvA_N	RuvA N terminal domain	Repair and Recombination
PF01330	RuvA_N	RuvA N terminal domain	Repair and Recombination
TIGR00635	ruvB	Holliday junction DNA helicase RuvB	Repair and Recombination
TIGR00635	ruvB	Holliday junction DNA helicase RuvB	Repair and Recombination
PF05491	RuvB_C	RuvB C-terminal winged helix domain	Repair and Recombination
PF05491	RuvB_C	RuvB C-terminal winged helix domain	Repair and Recombination
PF05496	RuvB_N	Holliday junction DNA helicase RuvB P-loop domain	Repair and Recombination
PF05496	RuvB_N	Holliday junction DNA helicase RuvB P-loop domain	Repair and Recombination
PF02075	RuvC	Crossover junction endodeoxyribonuclease RuvC	Repair and Recombination
TIGR00228	ruvC	crossover junction endodeoxyribonuclease RuvC	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF18516	RuvC_1	RuvC nuclease domain	Repair and Recombination
PF18541	RuvC_III	RuvC endonuclease subdomain 3	Repair and Recombination
TIGR00618	sbcc	exonuclease SbcC	Repair and Recombination
PF13558	SbcCD_C	Putative exonuclease SbcCD, C subunit	Repair and Recombination
TIGR00619	sbcd	exonuclease SbcCD, D subunit	Repair and Recombination
PF12320	SbcD_C	Type 5 capsule protein repressor C-terminal domain	WTF Processes
TIGR00024	SbcD_rel_arch	putative phosphoesterase	Repair and Recombination
PF08970	Sda	Sporulation inhibitor A	Sporulation
PF17418	SdpA	Sporulation delaying protein SdpA	Repair and Recombination
PF18335	SH3_13	ATP-dependent RecD-like DNA helicase SH3 domain	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF10747	SirA	Sporulation inhibitor of replication protein SirA	Sporulation
PF02463	SMC_N	RecF/RecN/SMC N terminal domain	Repair and Recombination
PF00080	Sod_Cu	Copper/zinc superoxide dismutase (SODC)	Oxidative Damage Resistance
PF02777	Sod_Fe_C	Iron/manganese superoxide dismutases, C-terminal domain	Oxidative Damage Resistance
PF00081	Sod_Fe_N	Iron/manganese superoxide dismutases, alpha-hairpin domain	Oxidative Damage Resistance
PF09055	Sod_Ni	Nickel-containing superoxide dismutase	Oxidative Damage Resistance
TIGR02754	sod_Ni_protease	nickel-type superoxide dismutase maturation protease	Oxidative Damage Resistance
TIGR02753	sodN	superoxide dismutase, Ni	Oxidative Damage Resistance

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02831	spo_II_M	stage II sporulation protein M	Sporulation
TIGR02834	spo_ytxC	putative sporulation protein YtxC	Sporulation
TIGR02832	spo_yunB	sporulation protein YunB	Sporulation
PF08769	Spo0A_C	Sporulation initiation factor Spo0A C terminal	Sporulation
PF07070	Spo0M	SpoOM protein	Sporulation
PF08631	SPO22	Meiosis protein SPO22/ZIP4 like	WTF Processes
PF15407	Spo7_2_N	Sporulation protein family 7	Sporulation
PF14682	SPOB_ab	Sporulation initiation phospho-transferase B, C-terminal	Sporulation
PF08486	SpoIID	Stage II sporulation protein	Sporulation
PF07228	SpoIIE	Stage II sporulation protein E (SpoIIE)	Sporulation
PF06686	SpoIIIAC	Stage III sporulation protein AC/AD protein family	Sporulation
PF12116	SpoIIID	Stage III sporulation protein D	Sporulation
PF01944	SpoIIM	Stage II sporulation protein M	Sporulation

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF07454	SpoIIP	Stage II sporulation protein P (SpoIIP)	Sporulation
PF09388	SpoOE-like	Spo0E like sporulation regulatory protein	Sporulation
PF05036	SPOR	Sporulation related domain	Sporulation
PF10957	Spore_Cse60	Sporulation protein Cse60	Sporulation
TIGR02887	spore_ger_x_C	germination protein, Ger(x)C family	Sporulation
PF05504	Spore_GerAC	Spore germination B3/ GerAC like, C-terminal	Sporulation
PF05504	Spore_GerAC	Spore germination B3/ GerAC like, C-terminal	Sporulation
TIGR02728	spore_gerQ	spore coat protein GerQ	Sporulation
TIGR02870	spore_II_D	stage II sporulation protein D	Sporulation
TIGR02865	spore_II_E	stage II sporulation protein E	Sporulation
TIGR02867	spore_II_P	stage II sporulation protein P	Sporulation
TIGR02837	spore_II_R	stage II sporulation protein R	Sporulation
PF09551	Spore_II_R	Stage II sporulation protein R (spore_II_R)	Sporulation
TIGR02858	spore_III_AA	stage III sporulation protein AA	Sporulation

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02833	spore_III_AB	stage III sporulation protein AB	Sporulation
PF09548	Spore_III_AB	Stage III sporulation protein AB (spore_III_AB)	Sporulation
TIGR02848	spore_III_AC	stage III sporulation protein AC	Sporulation
TIGR02849	spore_III_AD	stage III sporulation protein AD	Sporulation
TIGR02829	spore_III_AE	stage III sporulation protein AE	Sporulation
PF09546	Spore_III_AE	Stage III sporulation protein AE (spore_III_AE)	Sporulation
TIGR02896	spore_III_AF	stage III sporulation protein AF	Sporulation
PF09581	Spore_III_AF	Stage III sporulation protein AF (Spore_III_AF)	Sporulation
TIGR02830	spore_III_AG	stage III sporulation protein AG	Sporulation
TIGR02844	spore_III_D	sporulation transcriptional regulator SpoIIID	Sporulation
TIGR02836	spore_IV_A	stage IV sporulation protein A	Sporulation
PF09547	Spore_IV_A	Stage IV sporulation protein A (spore_IV_A)	Sporulation
TIGR02860	spore_IV_B	stage IV sporulation protein B	Sporulation

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02838	spore_V_AC	stage V sporulation protein AC	Sporulation
TIGR02845	spore_V_AD	stage V sporulation protein AD	Sporulation
TIGR02839	spore_V_AE	stage V sporulation protein AE	Sporulation
TIGR02900	spore_V_B	stage V sporulation protein B	Sporulation
TIGR02881	spore_V_K	stage V sporulation protein K	Sporulation
TIGR02851	spore_V_T	stage V sporulation protein T	Sporulation
TIGR02907	spore_VI_D	stage VI sporulation protein D	Sporulation
TIGR02892	spore_yabP	sporulation protein YabP	Sporulation
PF09578	Spore_YabQ	Spore cortex protein YabQ (Spore_YabQ)	Sporulation
PF14147	Spore_YhaL	Sporulation protein YhaL	Sporulation
TIGR02877	spore_yhbH	sporulation protein YhbH	Sporulation
PF09580	Spore_YhcN_YlaJ	Sporulation lipoprotein YhcN/YlaJ (Spore_YhcN_YlaJ)	Sporulation
TIGR02888	spore_YlmC_YmxH	sporulation protein, YlmC/YmxH family	Sporulation
TIGR02873	spore_ylxY	probable sporulation protein, polysaccharide deacetylase family	Sporulation

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR02878	spore_ypjB	sporulation protein YpjB	Sporulation
PF09577	Spore_YpjB	Sporulation protein YpjB (SpoYpjB)	Sporulation
TIGR02856	spore_yqfC	sporulation protein YqfC	Sporulation
TIGR02876	spore_yqfD	sporulation protein YqfD	Sporulation
TIGR02840	spore_YtaF	putative sporulation protein YtaF	Sporulation
TIGR02874	spore_ytfJ	sporulation protein YtfJ	Sporulation
PF09579	Spore_YtfJ	Sporulation protein YtfJ (Spore_YtfJ)	Sporulation
PF14034	Spore_YtrH	Sporulation protein YtrH	Sporulation
PF09560	Spore_YunB	Sporulation protein YunB (Spo_YunB)	Sporulation
TIGR02841	spore_YyaC	putative sporulation protein YyaC	Sporulation
PF12164	SporV_AA	Stage V sporulation protein AA	Sporulation
PF08183	SpoV	Stage V sporulation protein family	Sporulation
PF13782	SpoVAB	Stage V sporulation protein AB	Sporulation
PF03862	SpoVAC_SpoVAEB	SpoVAC/SpoVAEB sporulation membrane protein	Sporulation

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF07451	SpoVAD	Stage V sporulation protein AD (SpoVAD)	Sporulation
PF14097	SpoVAE	Stage V sporulation protein AE1	Sporulation
TIGR02214	spoVD_pbp	stage V sporulation protein D	Sporulation
TIGR02615	spoVE	stage V sporulation protein E	Sporulation
PF04026	SpoVG	SpoVG	Sporulation
PF14069	SpoVIF	Stage VI sporulation protein F	Sporulation
PF14069	SpoVIF	Stage VI sporulation protein F	Sporulation
PF04232	SpoVS	Stage V sporulation protein S (SpoVS)	Sporulation
PF15714	SpoVT_C	Stage V sporulation protein T C-terminal, transcription factor	Sporulation
PF00436	SSB	Single-strand binding protein family	Repair and Recombination
TIGR00621	ssb	single-stranded DNA-binding protein	Repair and Recombination
PF04686	SsgA	Streptomyces sporulation and cell division protein, SsgA	Sporulation

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
TIGR00624	tag	DNA-3-methyladenine glycosylase I	Repair and Recombination
TIGR01438	TGR	thioredoxin and glutathione reductase	Oxidative Damage Resistance
TIGR00449	tgt_general	tRNA-guanine family transglycosylase	WTF Processes
PF00085	Thioredoxin	Thioredoxin	Oxidative Damage Resistance
TIGR01068	thioredoxin	thioredoxin	Oxidative Damage Resistance
PF17991	Thioredoxin_10	Thioredoxin like C-terminal domain	Oxidative Damage Resistance
PF13098	Thioredoxin_2	Thioredoxin-like domain	Oxidative Damage Resistance
PF13192	Thioredoxin_3	Thioredoxin domain	Oxidative Damage Resistance

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF13462	Thioredoxin_4	Thioredoxin	Oxidative Damage Resistance
PF13743	Thioredoxin_5	Thioredoxin	Oxidative Damage Resistance
PF13848	Thioredoxin_6	Thioredoxin-like domain	Oxidative Damage Resistance
PF13899	Thioredoxin_7	Thioredoxin-like	Oxidative Damage Resistance
PF13905	Thioredoxin_8	Thioredoxin-like	Oxidative Damage Resistance
PF14595	Thioredoxin_9	Thioredoxin	Oxidative Damage Resistance
TIGR03491	TIGR03491	putative RecB family nuclease, TM0106 family	Repair and Recombination
TIGR02012	tigrfam_recA	protein RecA	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF00912	Transgly	Transglycosylase	WTF Processes
TIGR01292	TRX_reduct	thioredoxin-disulfide reductase	Oxidative Damage Resistance
PF03167	UDG	Uracil DNA glycosylase superfamily	Repair and Recombination
TIGR03914	UDG_fam_dom	uracil-DNA glycosylase family domain	Repair and Recombination
TIGR00758	UDG_fam4	uracil-DNA glycosylase, family 4	Repair and Recombination
TIGR00628	ung	uracil-DNA glycosylase	Repair and Recombination
PF02151	UVR	UvrB/uvrC motif	Repair and Recombination
TIGR00630	uvra	excinuclease ABC subunit A	Repair and Recombination
PF17755	UvrA_DNA-bind	UvrA DNA-binding domain	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF17760	UvrA_inter	UvrA interaction domain	Repair and Recombination
TIGR00631	uvrb	excinuclease ABC subunit B	Repair and Recombination
PF12344	UvrB	Ultra-violet resistance protein B	Repair and Recombination
PF17757	UvrB_inter	UvrB interaction domain	Repair and Recombination
TIGR00194	uvrC	excinuclease ABC subunit C	Repair and Recombination
PF08459	UvrC_HhH_N	UvrC Helix-hairpin-helix N-terminal	Repair and Recombination
TIGR01075	uvrD	DNA helicase II	Repair and Recombination
PF13361	UvrD_C	UvrD-like helicase C-terminal domain	Repair and Recombination
PF13538	UvrD_C_2	UvrD-like helicase C-terminal domain	Repair and Recombination
PF00580	UvrD-helicase	UvrD/REP helicase N-terminal domain	Repair and Recombination

Table S-Metadata 3: Resistance Gene Descriptions and Accession Numbers

ACCESSION	NAME	FUNCTION	CATEGORY
PF09680	YjcZ_2	Family of unknown function	WTF Processes
PF14620	YPEB	YpeB sporulation	Sporulation
PF06898	YqfD	Putative stage IV sporulation protein YqfD	Sporulation

R session info

Chapters 1 and 2

R version 4.1.0 (2021-05-18)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 19043)

Matrix products: default

locale:
[1] LC_COLLATE=English_United States.1252 LC_CTYPE=English_United States.1252
LC_MONETARY=English_United States.1252
[4] LC_NUMERIC=C LC_TIME=English_United States.1252

attached base packages:
[1] grid stats4 stats graphics grDevices utils datasets methods base

other attached packages:
[1] ragp_0.3.5.9000 ggcorrplot_0.1.3 ggdendro_0.1.23 broom_0.7.11
[5] glue_1.4.2 cowplot_1.1.1 ggtext_0.1.1 apeglm_1.14.0
[9] Hmisc_4.6-0 Formula_1.2-4 survival_3.2-11 circlize_0.4.13
[13] ComplexHeatmap_2.8.0 hciR_1.5 Rcpp_1.0.8.2 eulerr_6.1.1
[17] taxonomizr_0.9.3 UpSetR_1.4.0 pheatmap_1.0.12 ggh4x_0.2.1.9000
[21] ellipse_0.4.2 ggpolypath_0.1.0 venn_1.10 ggVennDiagram_1.2.1
[25] DESeq2_1.32.0 SummarizedExperiment_1.22.0 MatrixGenerics_1.4.3 matrixStats_0.61.0
[29] GenomicRanges_1.44.0 GenomeInfoDb_1.30.1 IRanges_2.26.0 S4Vectors_0.30.2
[33] DescTools_0.99.44 EcolUtils_0.1 vegan_2.5-7 lattice_0.20-44

[37] permute_0.9-5 extrafont_0.17 ggfortify_0.4.13 gridExtra_2.3
[41] dichromat_2.0-0.1 nationalparkcolors_0.1.0 ape_5.5 prodlim_2019.11.13
[45] phyloseq_1.36.0 forcats_0.5.1 stringr_1.4.0 purrr_0.3.4
[49] tidyr_1.1.3 tibble_3.1.2 tidyverse_1.3.1 reshape_0.8.8
[53] ggplot2_3.3.5.9000 data.table_1.14.2 dplyr_1.0.6 readr_2.1.1
[57] plyr_1.8.6 knitr_1.38 metagenomeSeq_1.34.0 RColorBrewer_1.1-2
[61] glmnet_4.1-3 Matrix_1.3-3 Biobase_2.52.0 BiocGenerics_0.40.0
[65] edgeR_3.34.1 limma_3.48.3

loaded via a namespace (and not attached):
[1] R.methodsS3_1.8.1 coda_0.19-4 bit64_4.0.5 DelayedArray_0.18.0 R.utils_2.11.0
[6] rpart_4.1-15 KEGGREST_1.32.0 RCurl_1.98-1.3 doParallel_1.0.16 generics_0.1.1
[11] RSQLite_2.2.10 proxy_0.4-26 future_1.23.0 bit_4.0.4 tzdb_0.2.0
[16] xml2_1.3.3 lubridate_1.8.0 assertthat_0.2.1 xfun_0.29 hms_1.1.1
[21] evaluate_0.15 fansi_0.5.0 progress_1.2.2 caTools_1.18.2 dbplyr_2.1.1
[26] readxl_1.3.1 igraph_1.2.6 DBI_1.1.2 geneplotter_1.70.0 htmlwidgets_1.5.4
[31] ellipsis_0.3.2 backports_1.4.1 annotate_1.70.0 vctrs_0.3.8 Cairo_1.5-14
[36] cachem_1.0.5 withr_2.5.0 RVenn_1.1.0 bdsmatrix_1.3-4 checkmate_2.0.0
[41] prettyunits_1.1.1 cluster_2.1.2 crayon_1.5.1 genefilter_1.74.1 pkgconfig_2.0.3
[46] nlme_3.1-152 nnet_7.3-16 rlang_1.0.2 globals_0.14.0 lifecycle_1.0.1
[51] seqinr_4.2-16 extrafontdb_1.0 modelr_0.1.8 cellranger_1.1.0 Rhdf5lib_1.14.2
[56] boot_1.3-28 reprex_2.0.1 base64enc_0.1-3 GlobalOptions_0.1.2 png_0.1-7
[61] rjson_0.2.20 rootSolve_1.8.2.3 bitops_1.0-7 R.oo_1.24.0 KernSmooth_2.23-20
[66] rhdf5filters_1.4.0 Biostrings_2.60.2 blob_1.2.2 shape_1.4.6 parallelly_1.30.0
[71] jpeg_0.1-9 scales_1.2.0 memoise_2.0.1 magrittr_2.0.1 gplots_3.1.1
[76] zlibbioc_1.38.0 compiler_4.1.0 bbmle_1.0.24 clue_0.3-60 cli_3.1.0
[81] ade4_1.7-16 XVector_0.32.0 listenv_0.8.0 htmlTable_2.4.0 MASS_7.3-54
[86] mgcv_1.8-35 tidyselect_1.1.1 stringi_1.6.1 emdbook_1.3.12 yaml_2.2.1
[91] locfit_1.5-9.4 latticeExtra_0.6-29 tools_4.1.0 lmom_2.8 future.apply_1.8.1
[96] parallel_4.1.0 rstudioapi_0.13 foreach_1.5.1 foreign_0.8-81 gld_2.6.3
[101] digest_0.6.27 lava_1.6.10 gridtext_0.1.4 httr_1.4.3 AnnotationDbi_1.56.2
[106] Wrench_1.10.0 colorspace_2.0-1 rvest_1.0.2 XML_3.99-0.6 fs_1.5.2
[111] splines_4.1.0 expm_0.999-6 xgboost_1.6.0.1 multtest_2.48.0 Exact_3.1
[116] xtable_1.8-4 jsonlite_1.7.2 R6_2.5.1 pillar_1.7.0 htmltools_0.5.2
[121] fastmap_1.1.0 BiocParallel_1.26.2 class_7.3-19 codetools_0.2-18 mvtnorm_1.1-3
[126] utf8_1.2.1 numDeriv_2016.8-1.1 gtools_3.9.2 Rttf2pt1_1.3.9 admisc_0.16
[131] rmarkdown_2.11 biomformat_1.20.0 munsell_0.5.0 e1071_1.7-9 GetoptLong_1.0.5
[136] rhdf5_2.36.0 GenomeInfoDbData_1.2.7 iterators_1.0.13 haven_2.4.3 reshape2_1.4.4
[141] gtable_0.3.0

Chapter 3

R version 4.1.0 (2021-05-18)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 19043)

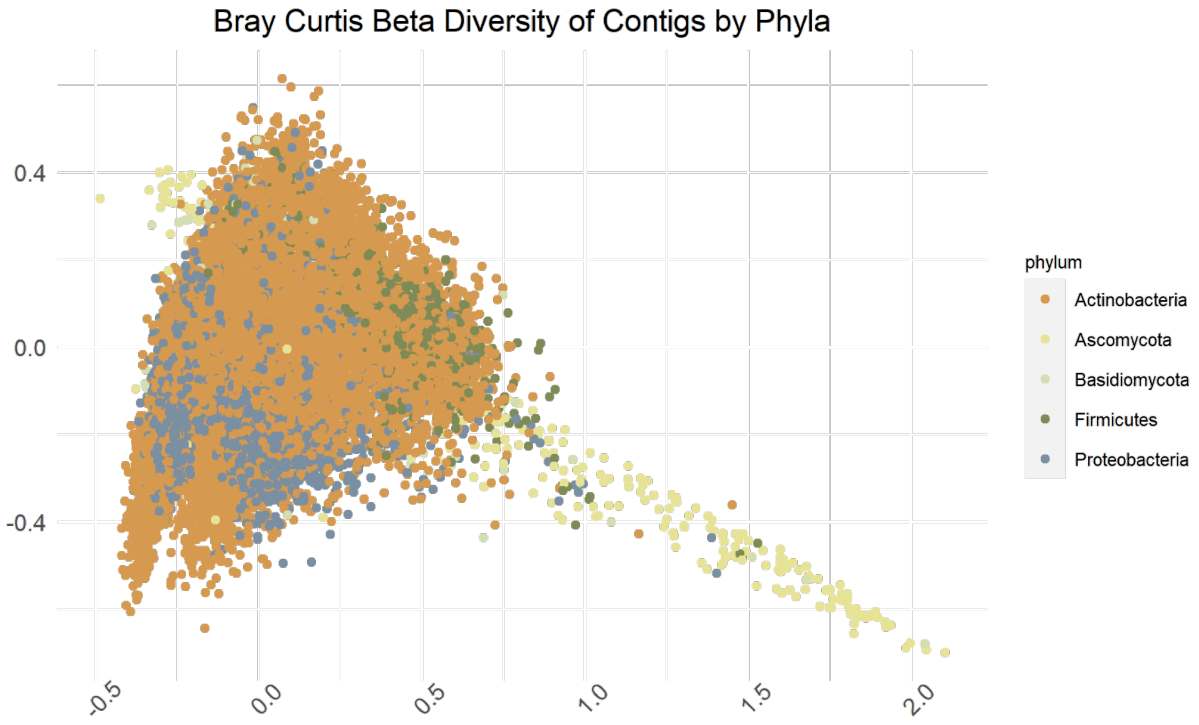
Matrix products: default

locale:

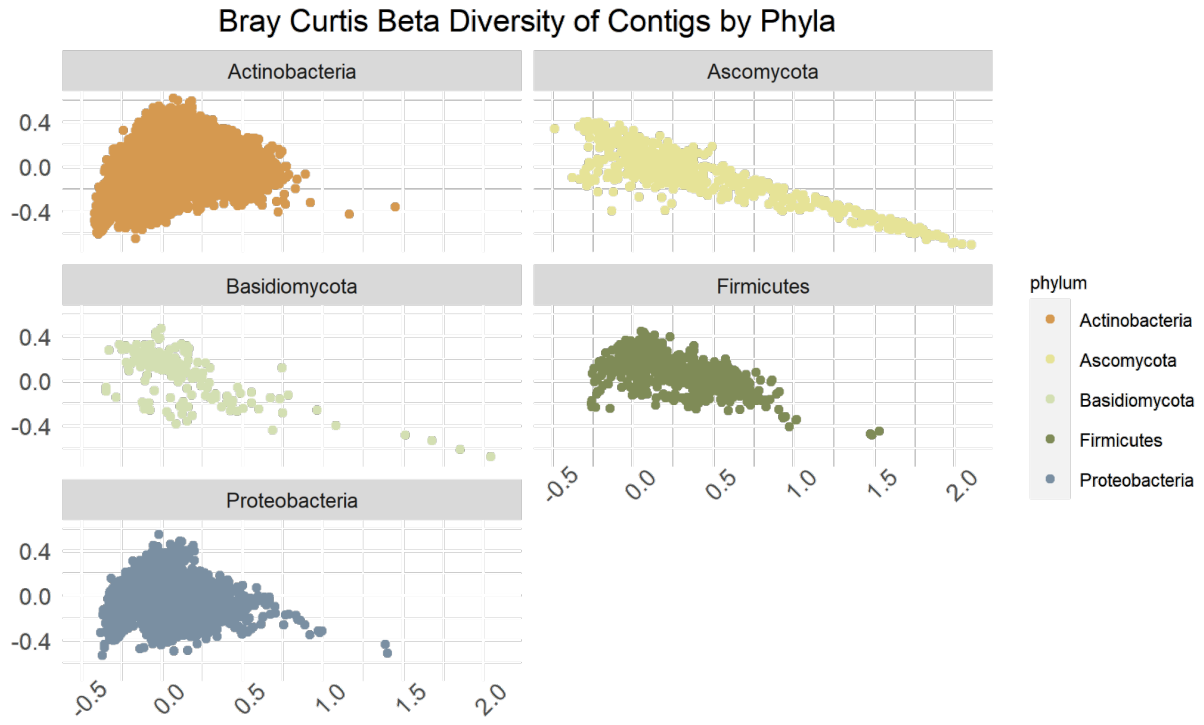
```
[1] LC_COLLATE=English_United States.1252  LC_CTYPE=English_United States.1252
LC_MONETARY=English_United States.1252
[4] LC_NUMERIC=C                           LC_TIME=English_United States.1252

attached base packages:
[1] stats4    stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
 [1] BiocManager_1.30.16  dichromat_2.0-0.1    data.table_1.14.2    rentrez_1.2.3        ragp_0.3.5.9000
 [6] AnnotationDbi_1.56.2 IRanges_2.26.0       S4Vectors_0.30.2     Biobase_2.52.0       BiocGenerics_0.40.0
[11] forcats_0.5.1        stringr_1.4.0        dplyr_1.0.6          purrr_0.3.4          readr_2.1.1
[16] tidyr_1.1.3          tibble_3.1.2         ggplot2_3.3.5.9000   tidyverse_1.3.1      taxonomizr_0.9.3
[21] rlang_1.0.2

loaded via a namespace (and not attached):
  [1] colorspace_2.0-1       seqinr_4.2-16          ellipsis_0.3.2         rprojroot_2.0.2        htmlTable_2.4.0
  [6] XVector_0.32.0         base64enc_0.1-3        fs_1.5.2               rstudioapi_0.13        farver_2.1.0
 [11] remotes_2.4.2          bit64_4.0.5            fansi_0.5.0            lubridate_1.8.0        xml2_1.3.3
 [16] splines_4.1.0          cachem_1.0.5           knitr_1.38             pkgload_1.2.4          ade4_1.7-16
 [21] Formula_1.2-4          jsonlite_1.7.2         broom_0.7.11           cluster_2.1.2          dbplyr_2.1.1
 [26] png_0.1-7              compiler_4.1.0         httr_1.4.3             backports_1.4.1        assertthat_0.2.1
 [31] Matrix_1.3-3           fastmap_1.1.0          cli_3.1.0              htmltools_0.5.2        prettyunits_1.1.1
 [36] tools_4.1.0            gtable_0.3.0           glue_1.4.2             GenomeInfoDbData_1.2.7 Rcpp_1.0.8.2
 [41] cellranger_1.1.0       vctrs_0.3.8            Biostrings_2.60.2      xfun_0.29              ps_1.6.0
 [46] testthat_3.1.1         rvest_1.0.2            lifecycle_1.0.1        devtools_2.4.3         XML_3.99-0.6
 [51] zlibbioc_1.38.0        MASS_7.3-54            scales_1.2.0           vroom_1.5.7            hms_1.1.1
 [56] parallel_4.1.0         RColorBrewer_1.1-2     yaml_2.2.1             curl_4.3.2             gridExtra_2.3
 [61] memoise_2.0.1          rpart_4.1-15           latticeExtra_0.6-29    stringi_1.6.1          RSQLite_2.2.10
 [66] desc_1.4.0             checkmate_2.0.0        pkgbuild_1.3.1         GenomeInfoDb_1.30.1    pkgconfig_2.0.3
 [71] bitops_1.0-7           evaluate_0.15          lattice_0.20-44        htmlwidgets_1.5.4      labeling_0.4.2
 [76] bit_4.0.4              processx_3.5.2         tidyselect_1.1.1       magrittr_2.0.1         R6_2.5.1
 [81] generics_0.1.1         Hmisc_4.6-0            DBI_1.1.2              foreign_0.8-81         pillar_1.7.0
 [86] haven_2.4.3            withr_2.5.0            nnet_7.3-16            survival_3.2-11        KEGGREST_1.32.0
 [91] RCurl_1.98-1.3         modelr_0.1.8           crayon_1.5.1           xgboost_1.6.0.1        utf8_1.2.1
 [96] tzdb_0.2.0             rmarkdown_2.11         jpeg_0.1-9             progress_1.2.2         usethis_2.1.5
[101] grid_4.1.0             readxl_1.3.1           blob_1.2.2             callr_3.7.0            reprex_2.0.1
[106] digest_0.6.27          munsell_0.5.0          sessioninfo_1.2.2
```

Figures and Tables

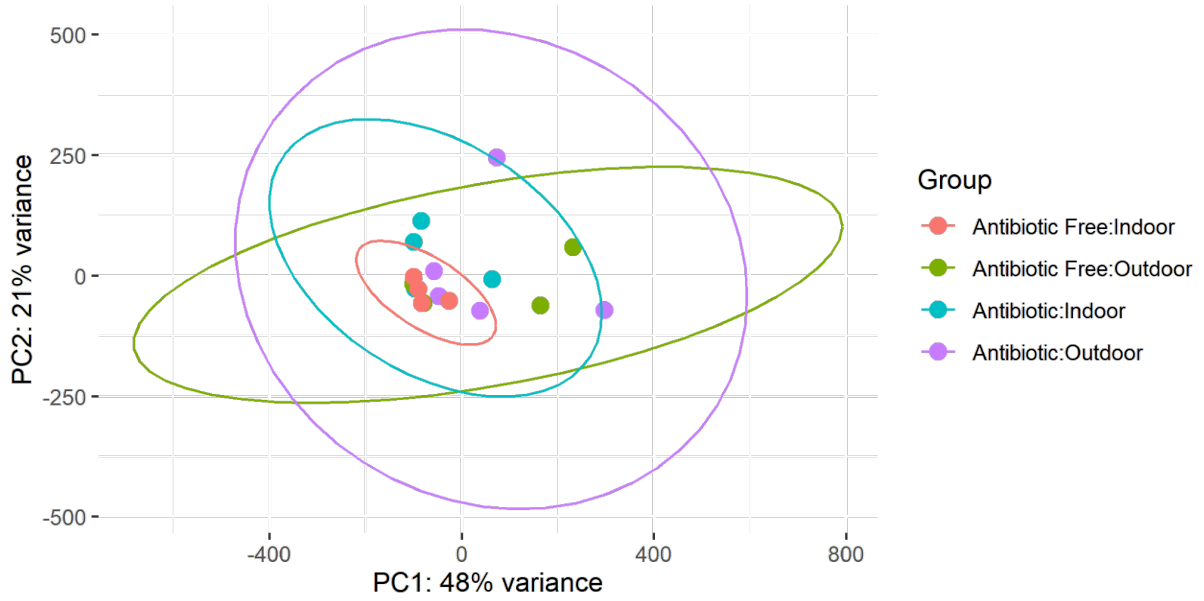
Chapter 1



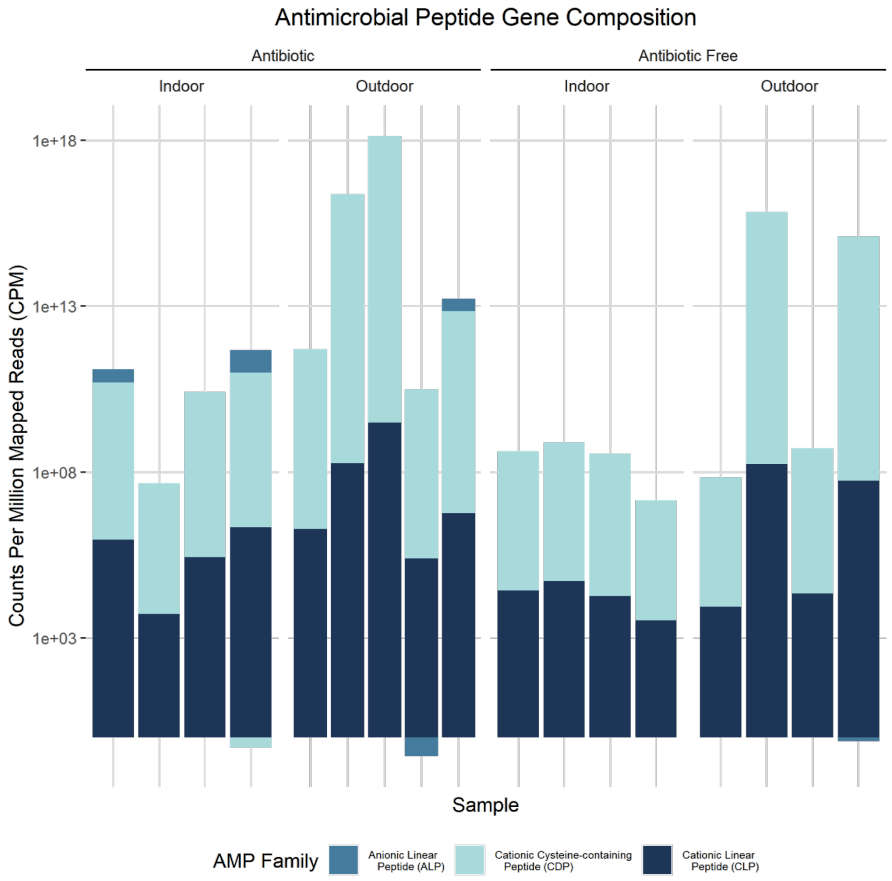
S1.1: Bray Curtis Beta Diversity of Contigs by Top 5 Most Abundant Phyla of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



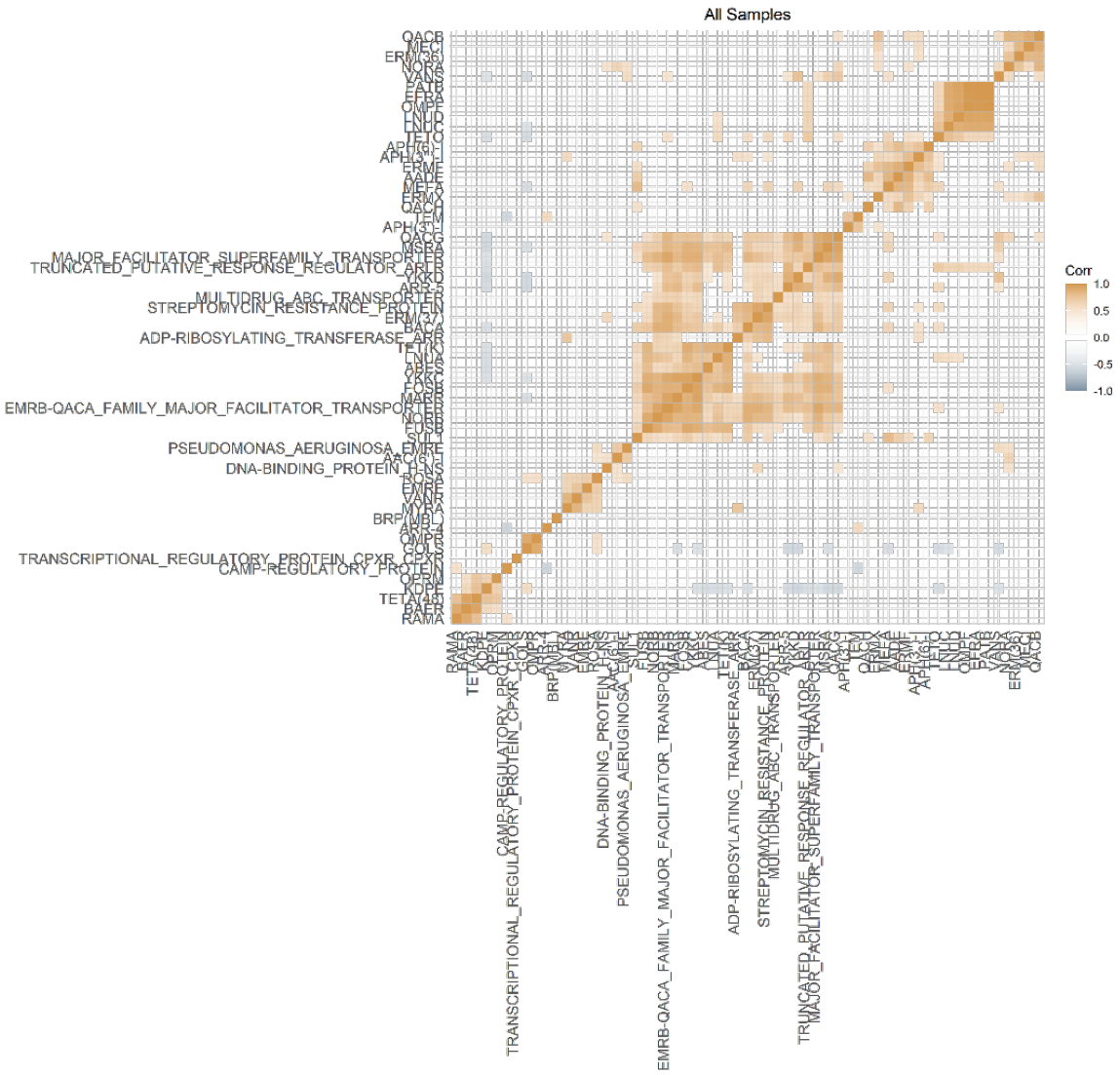
S1.2: Bray Curtis Beta Diversity of Contigs Separated by Top 5 Most Abundant Phyla of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



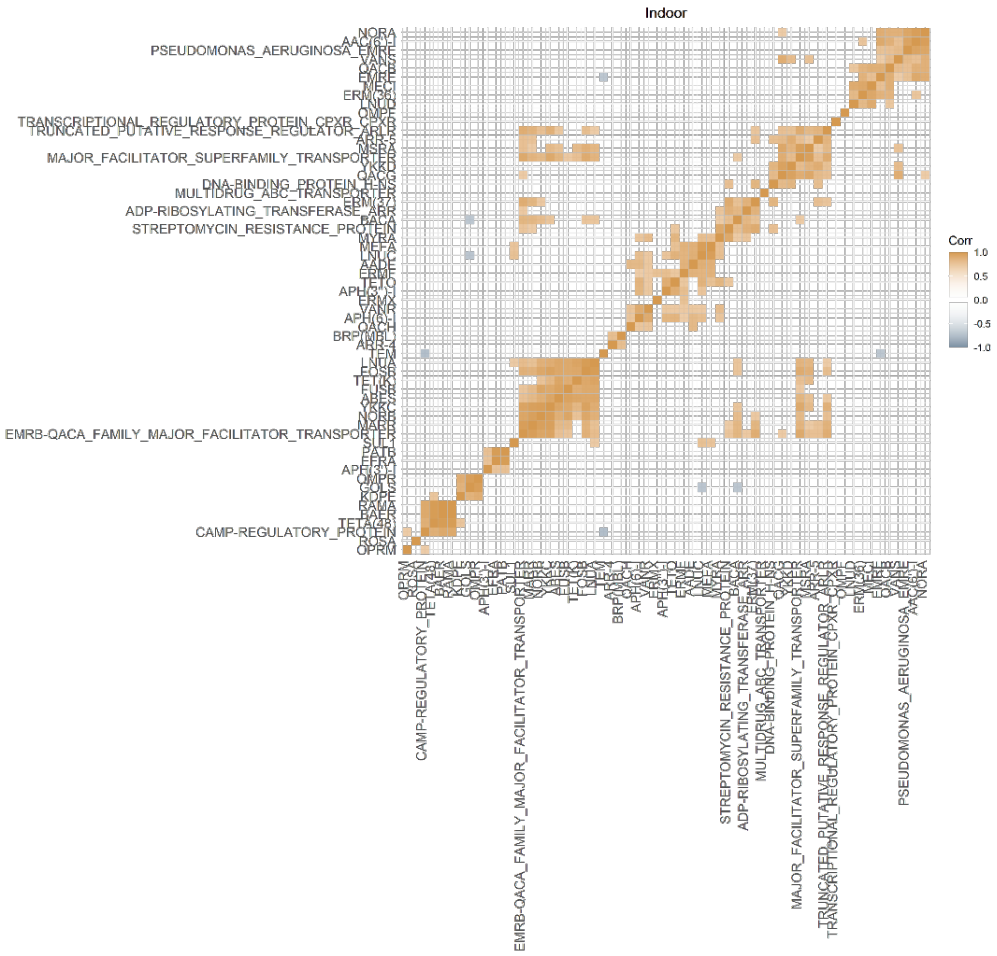
S1.3: PCA Plot of samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



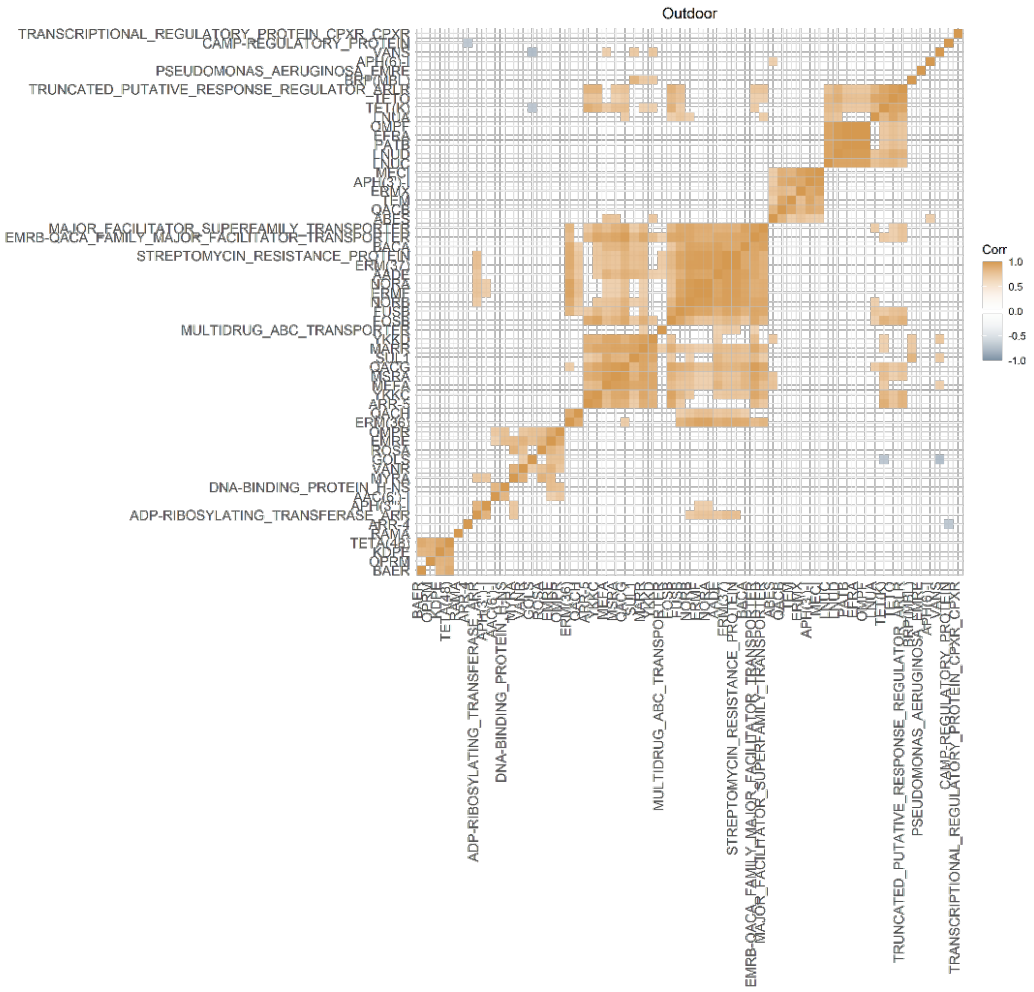
S1.4: AMP Class Composition by Sample of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



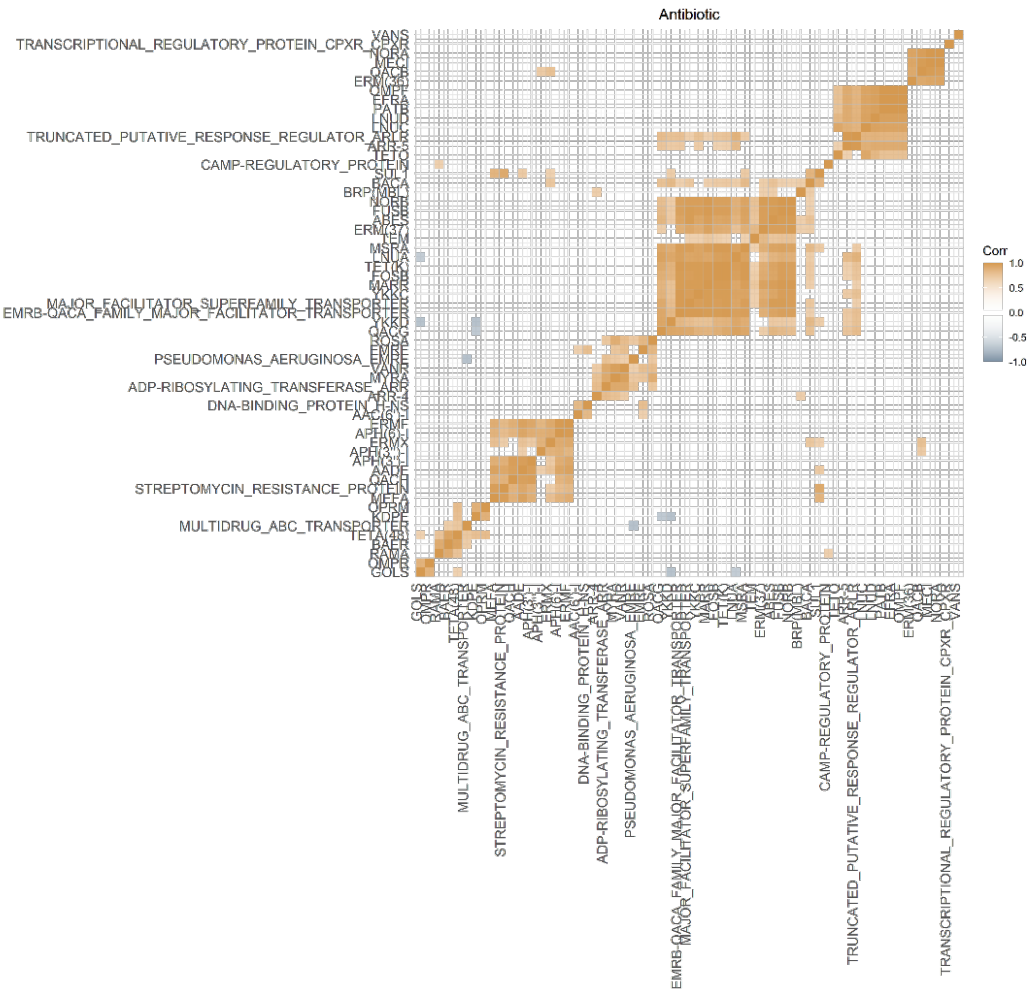
S1.5: Pairwise Pearson Correlation of AMR genes of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



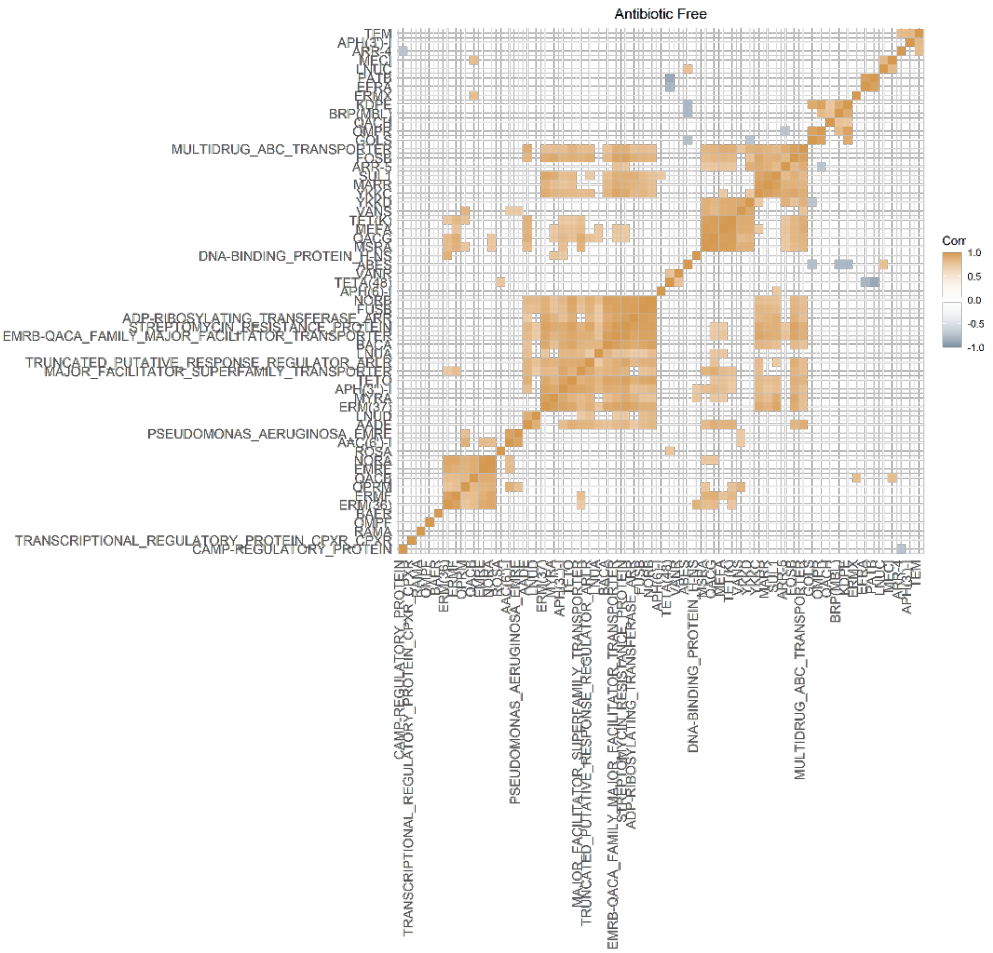
S1.6: Pairwise Pearson Correlation of AMR genes for Indoor Samples of microbial communities of internal (indoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



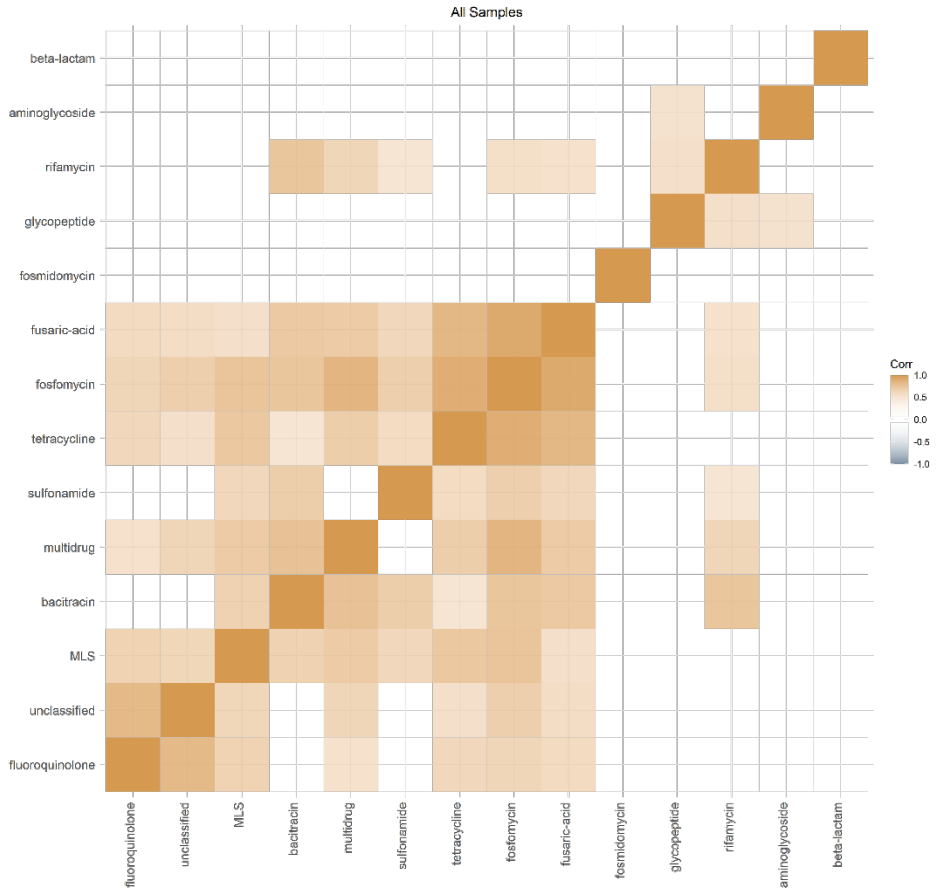
S1.7: Pairwise Pearson Correlation of AMR genes for Outdoor Samples of microbial communities of external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



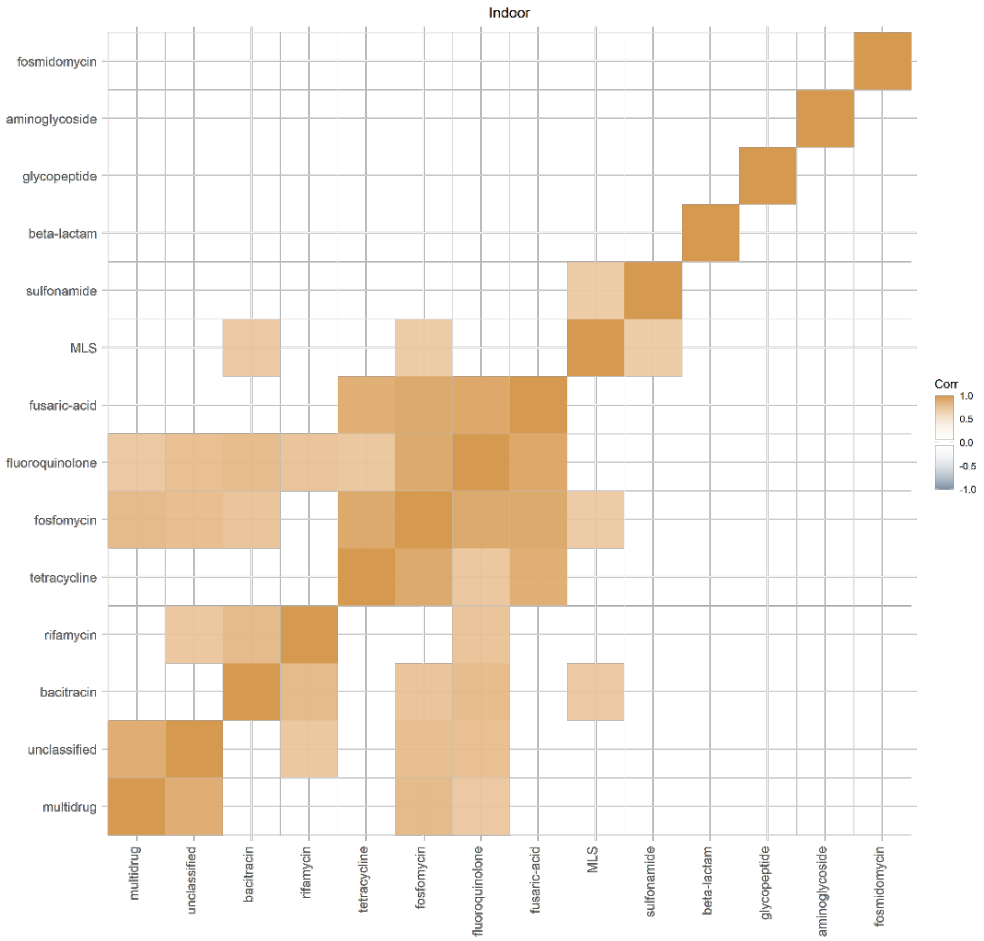
S1.8: Pairwise Pearson Correlation of AMR genes for Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).



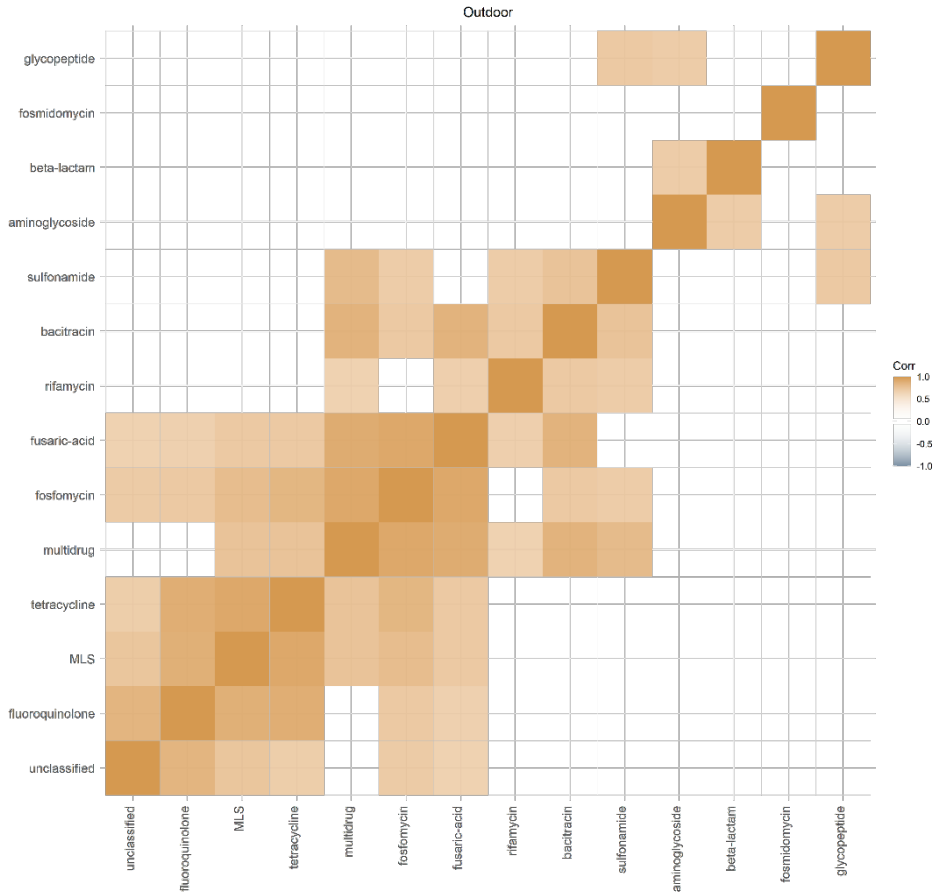
S1.9: Pairwise Pearson Correlation of AMR genes for Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).



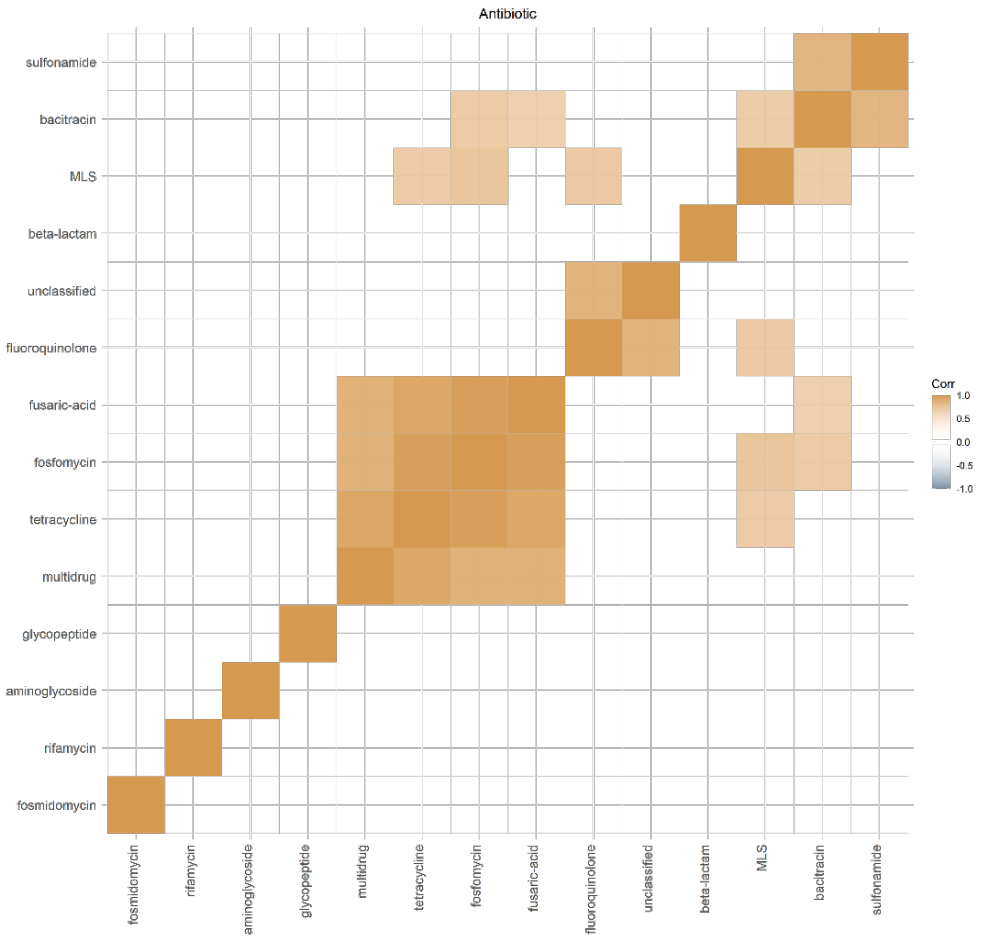
S1.10: Pairwise Pearson Correlation of AMR Classes Across All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



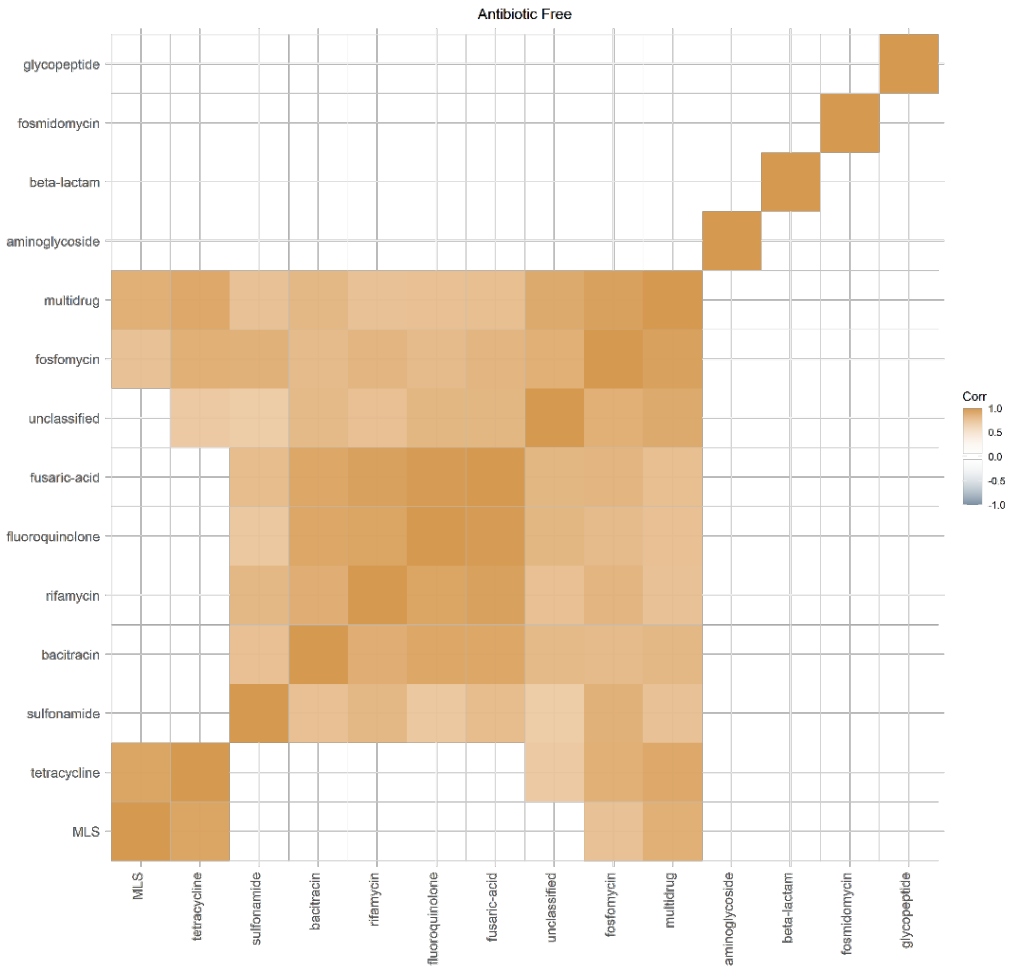
S1.11: Pairwise Pearson Correlation of AMR Classes of Indoor Samples of microbial communities of internal (indoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



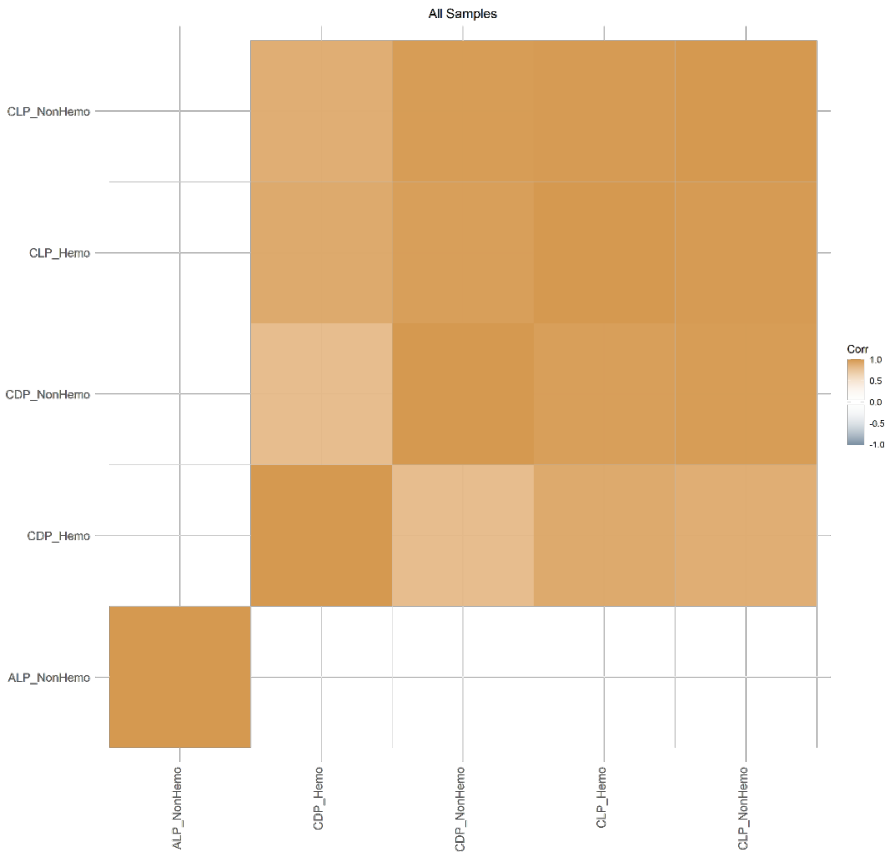
S1.12: Pairwise Pearson Correlation of AMR Classes of Outdoor Samples of microbial communities of external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



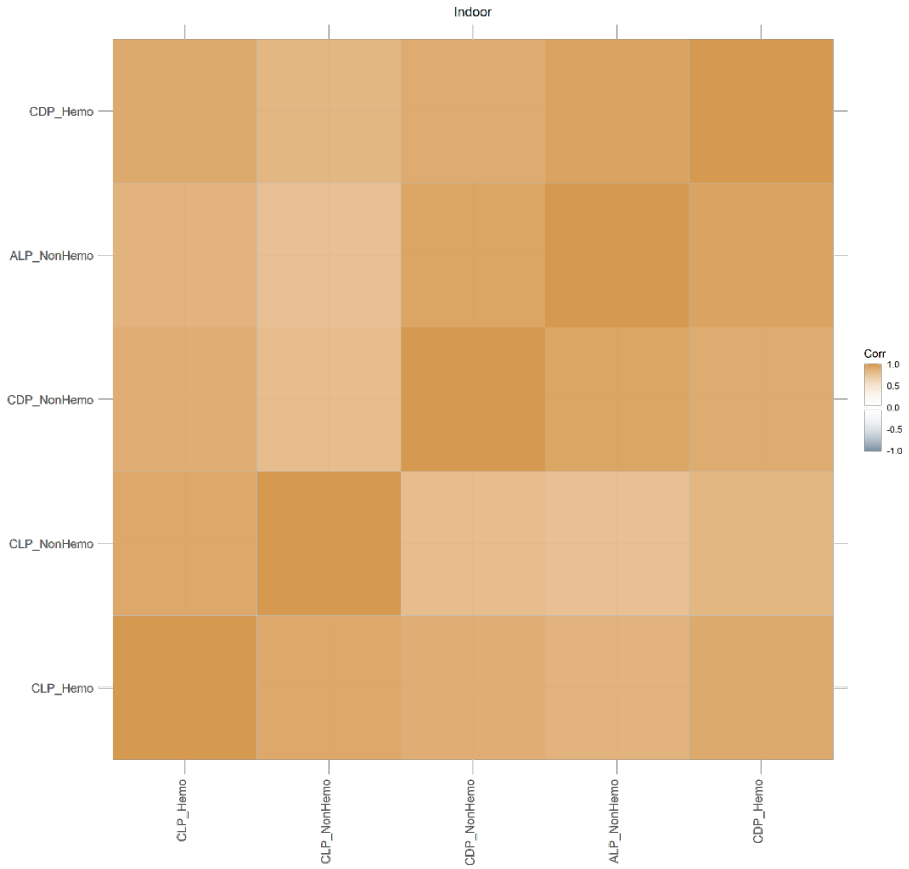
S1.13: Pairwise Pearson Correlation of AMR Classes of Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).



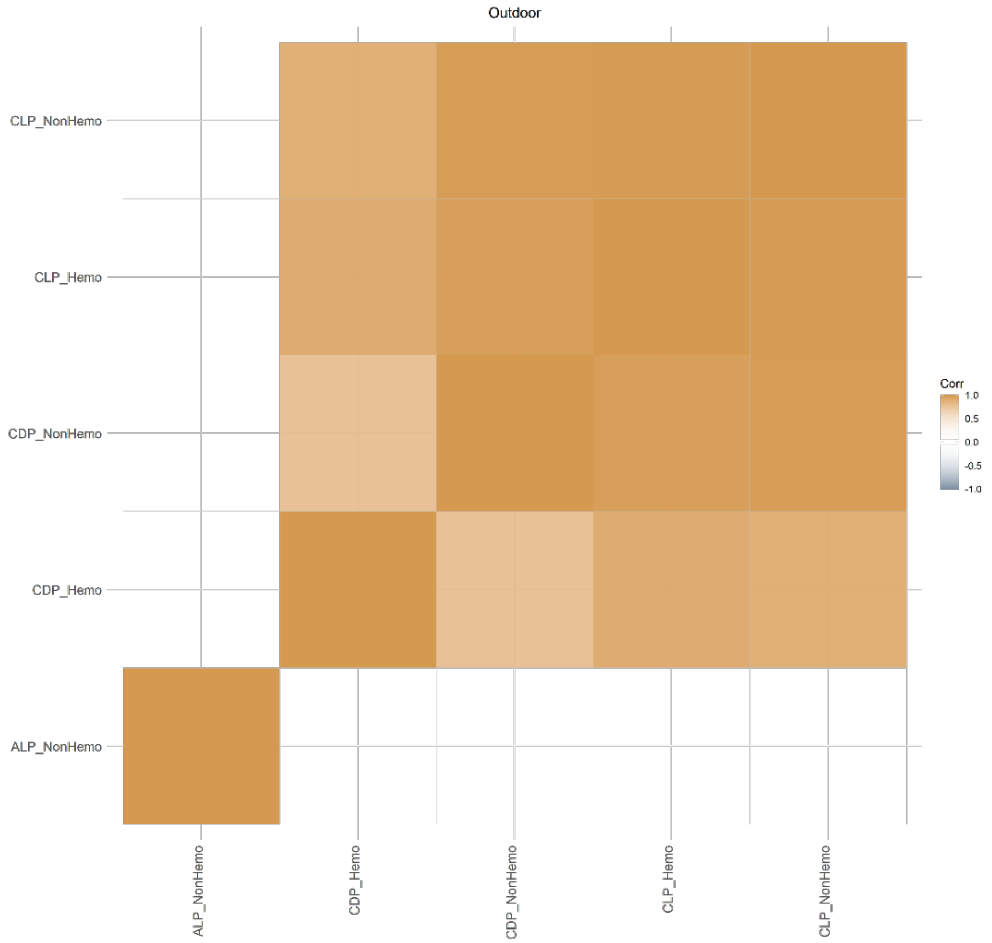
S1.14: Pairwise Pearson Correlation of AMR Classes of Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).



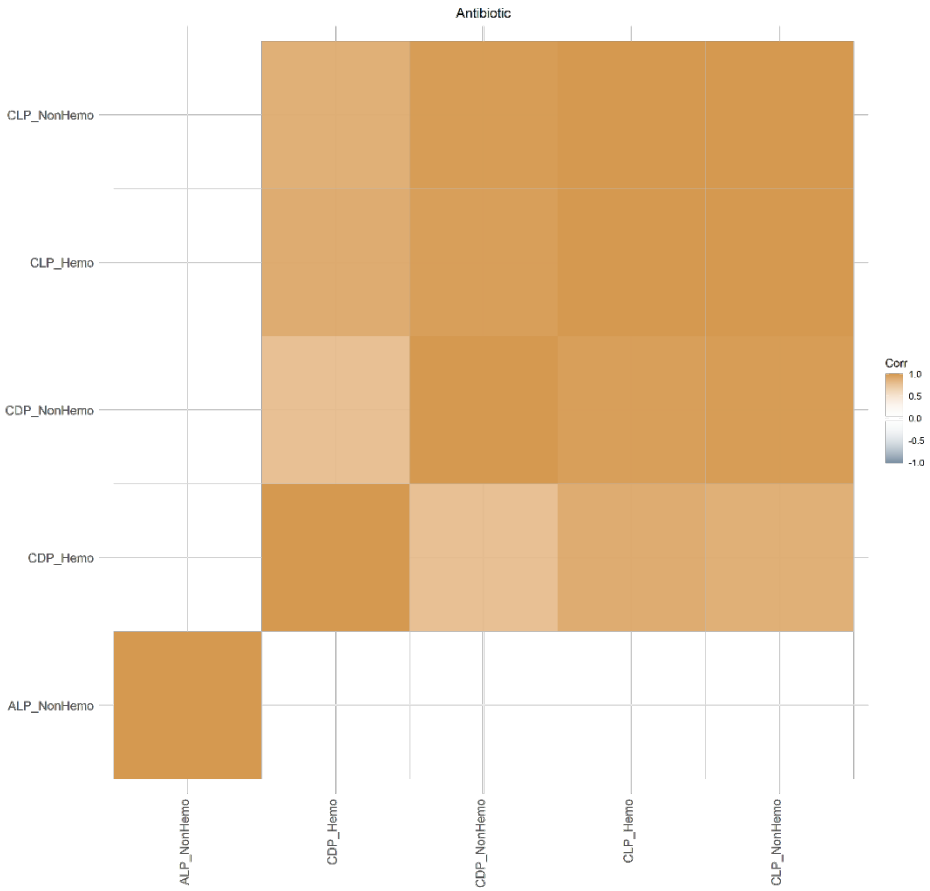
S1.15: Pairwise Pearson Correlation of AMP Classes Across All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



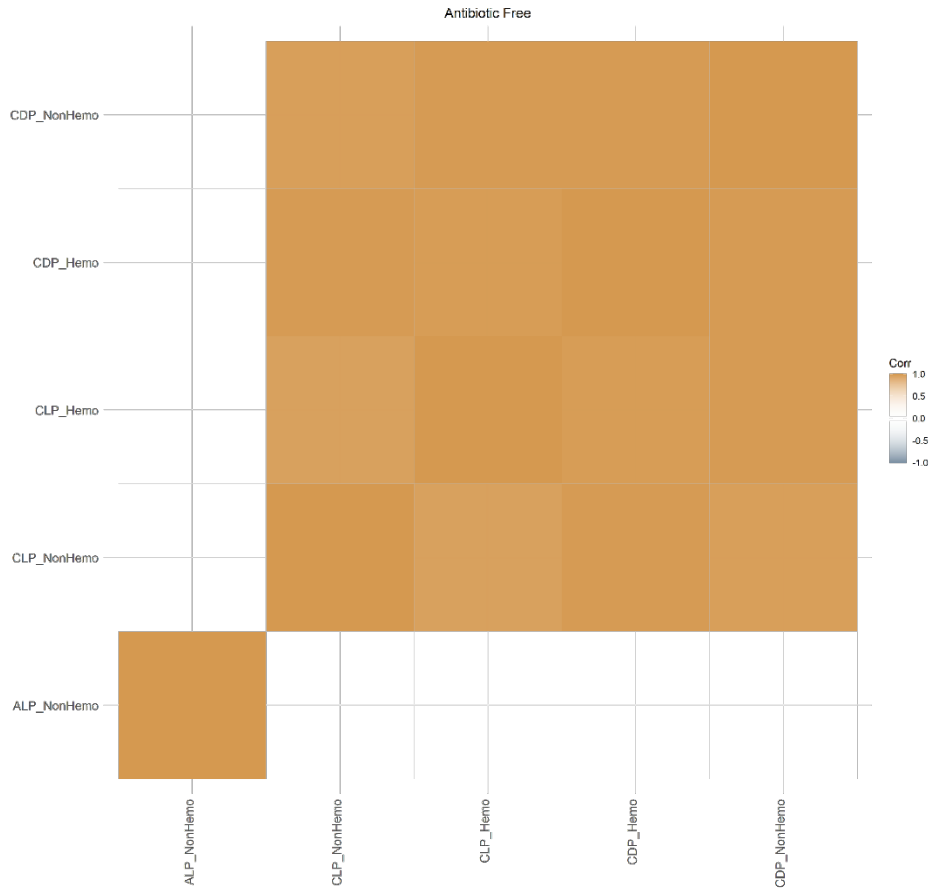
S1.16: Pairwise Pearson Correlation of AMP Classes for Indoor Samples of microbial communities of internal (indoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



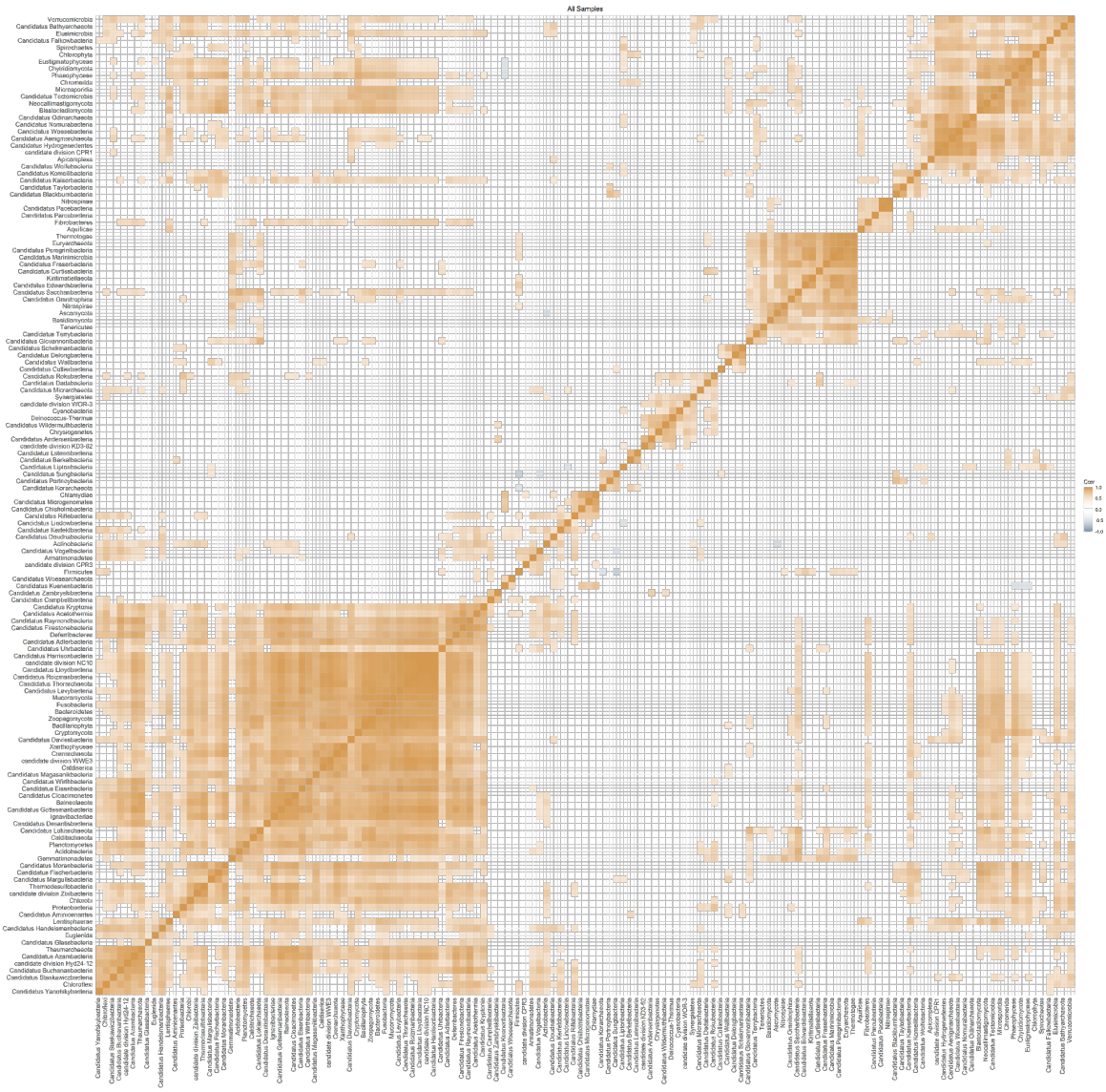
S1.17: Pairwise Pearson Correlation of AMP Classes for Outdoor Samples of microbial communities of external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



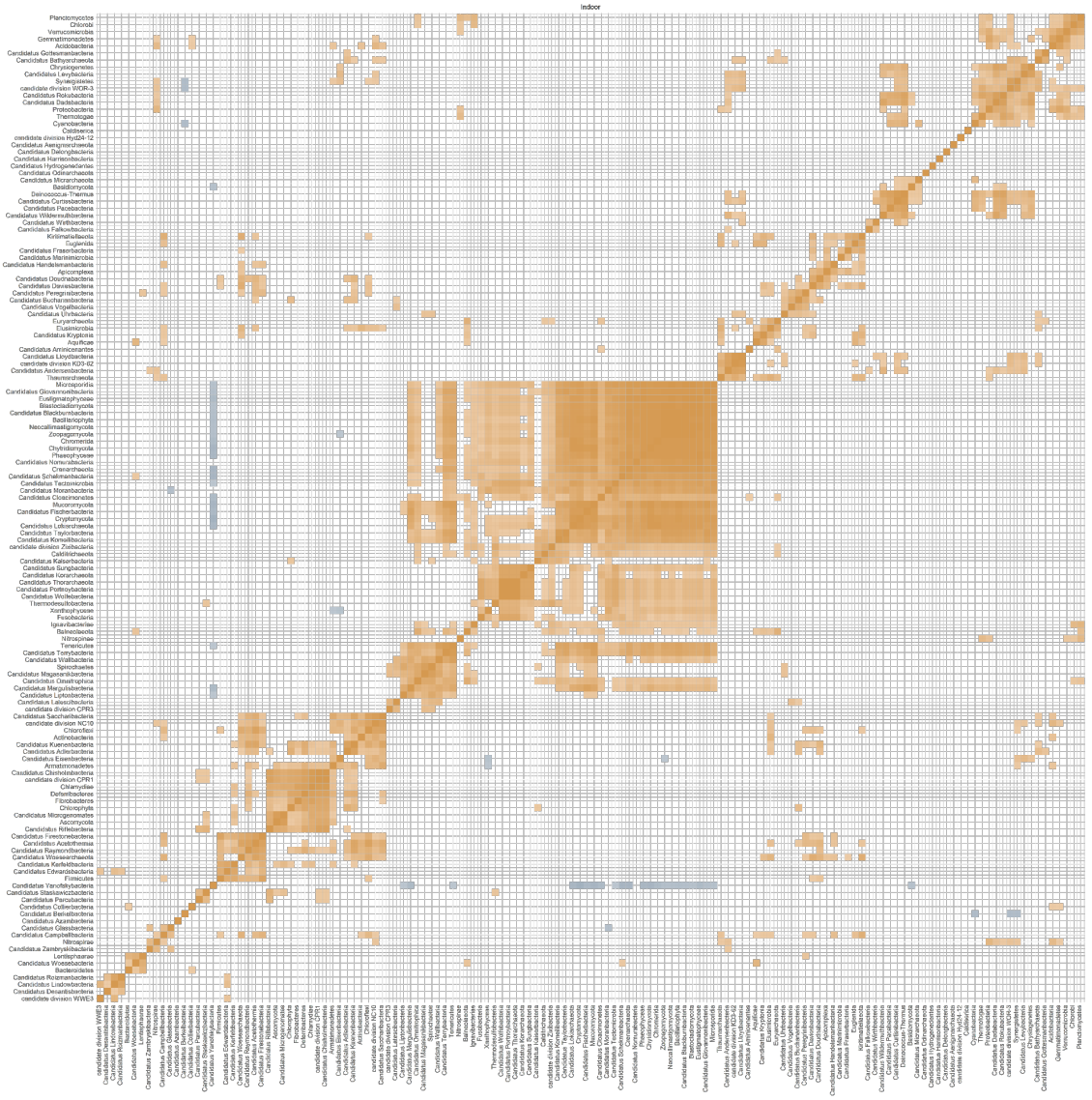
S1.18: Pairwise Pearson Correlation of AMP Classes for Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).



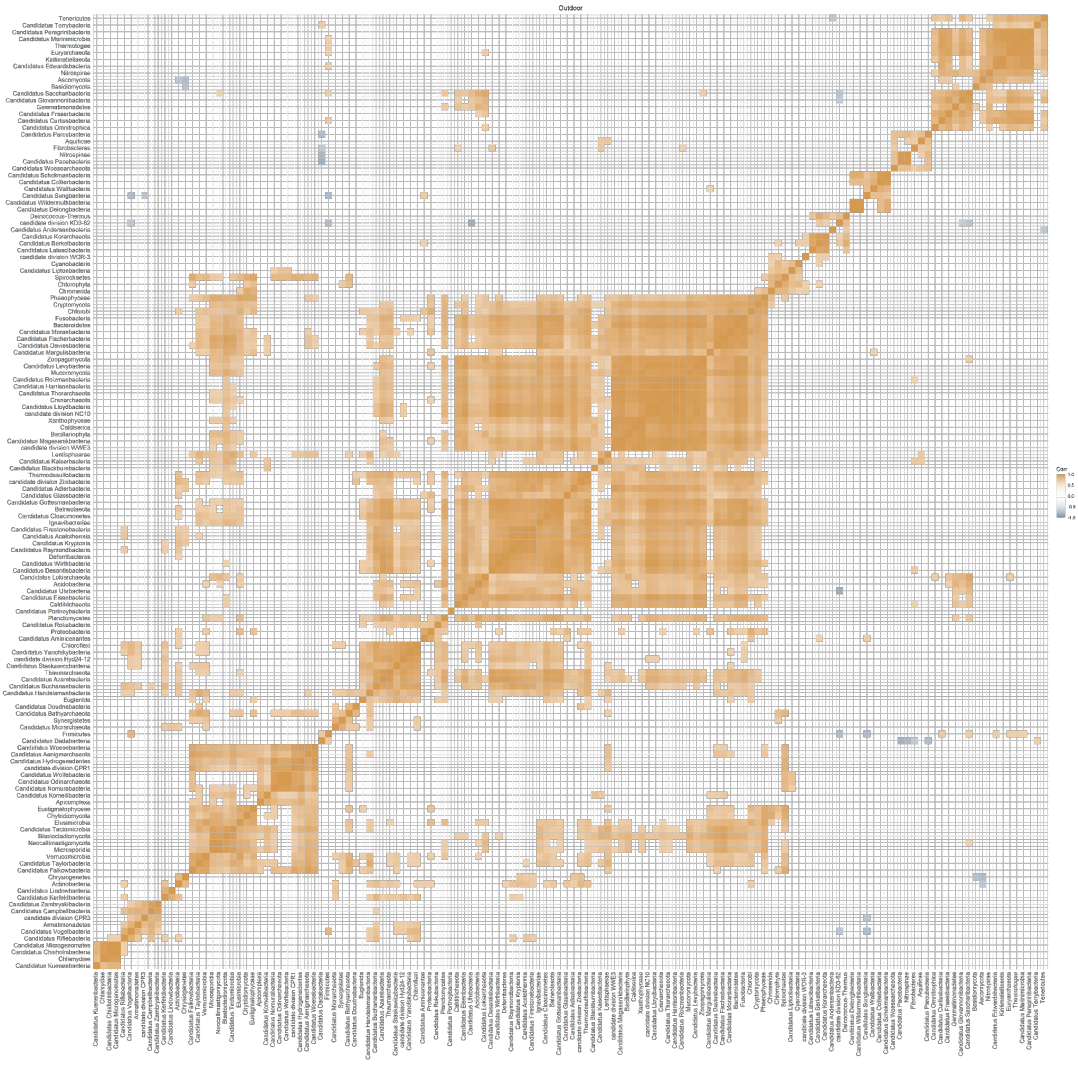
S1.19: Pairwise Pearson Correlation of AMP Classes for Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).



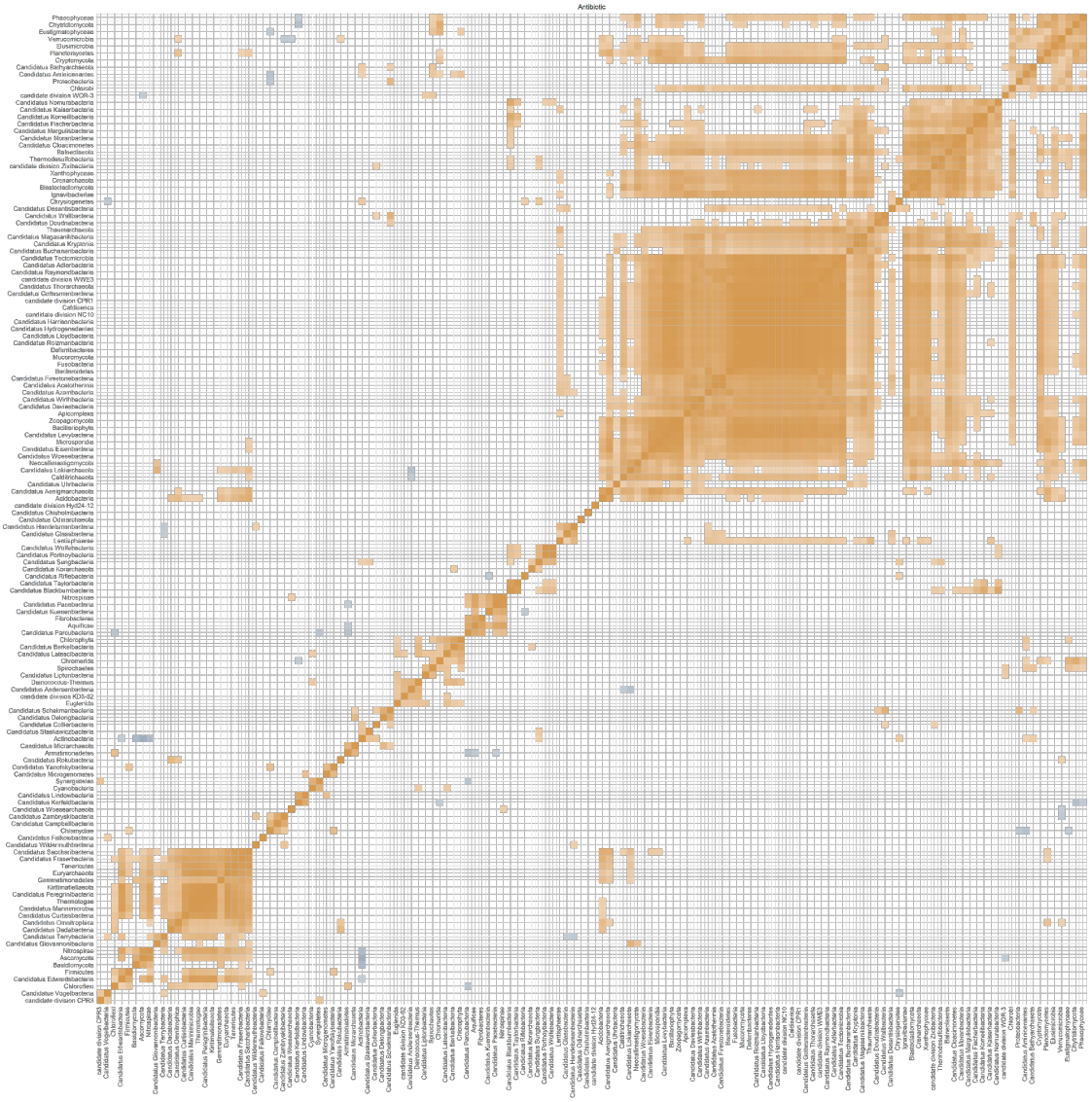
S1.20: Pairwise Pearson Correlation of microbial community phyla across all samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



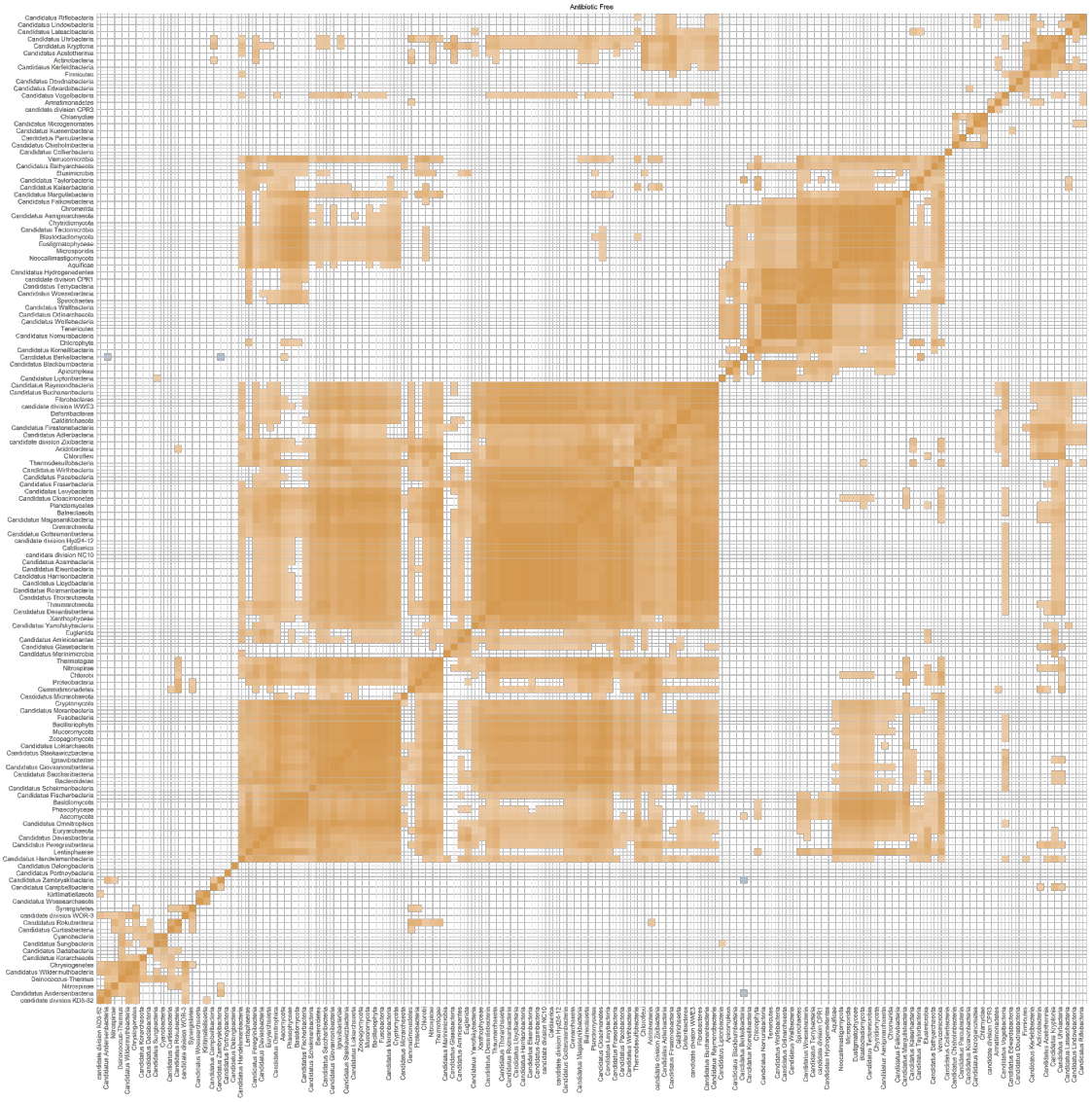
S1.21: Pairwise Pearson Correlation of Phyla for Indoor Samples of microbial communities of internal (indoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



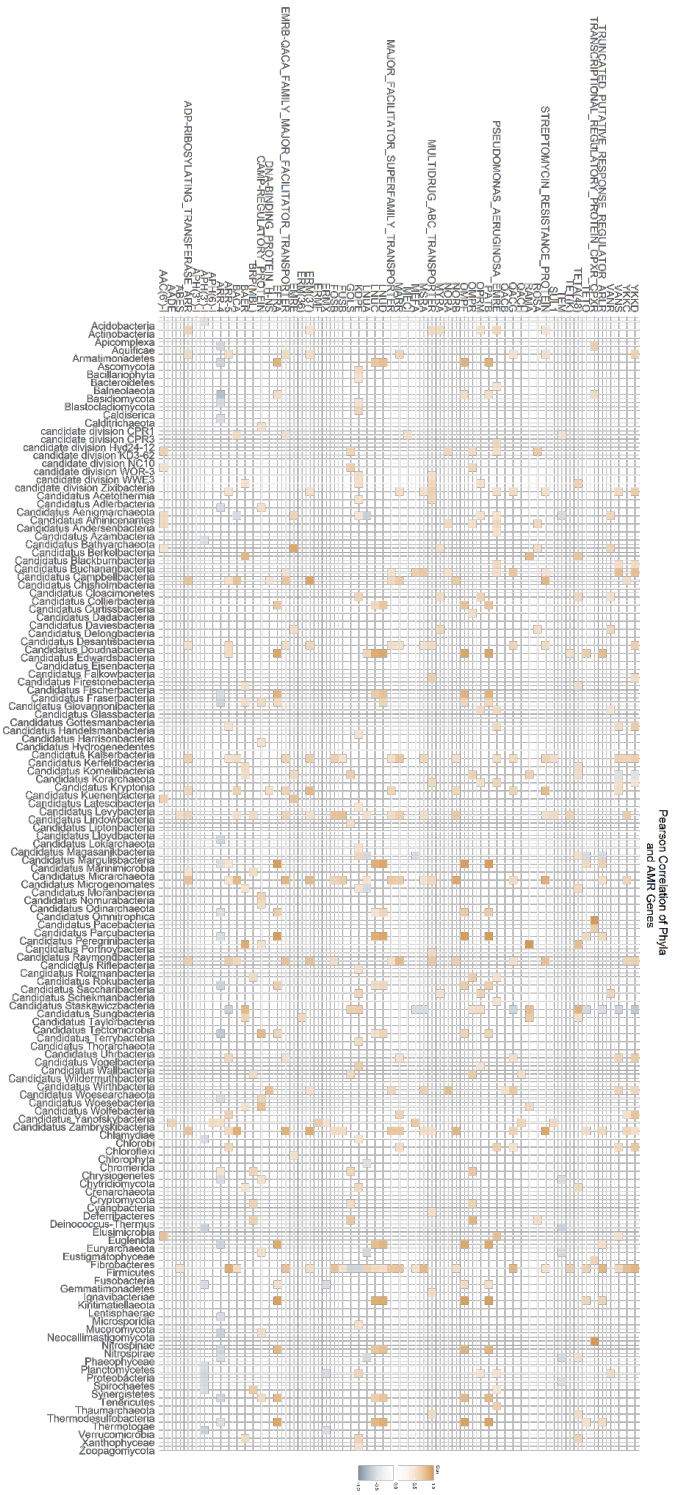
S1.22: Pairwise Pearson Correlation of Phyla for Outdoor Samples of microbial communities of external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



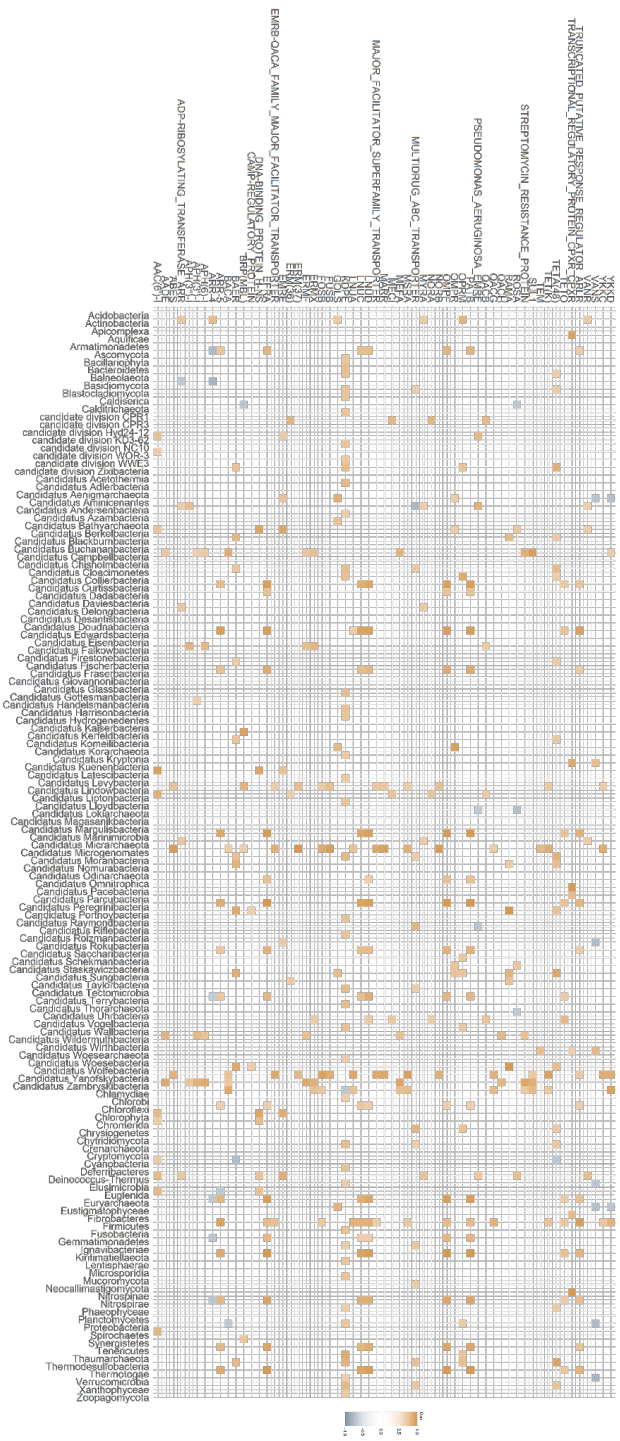
S1.23: Pairwise Pearson Correlation of Phyla for Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).



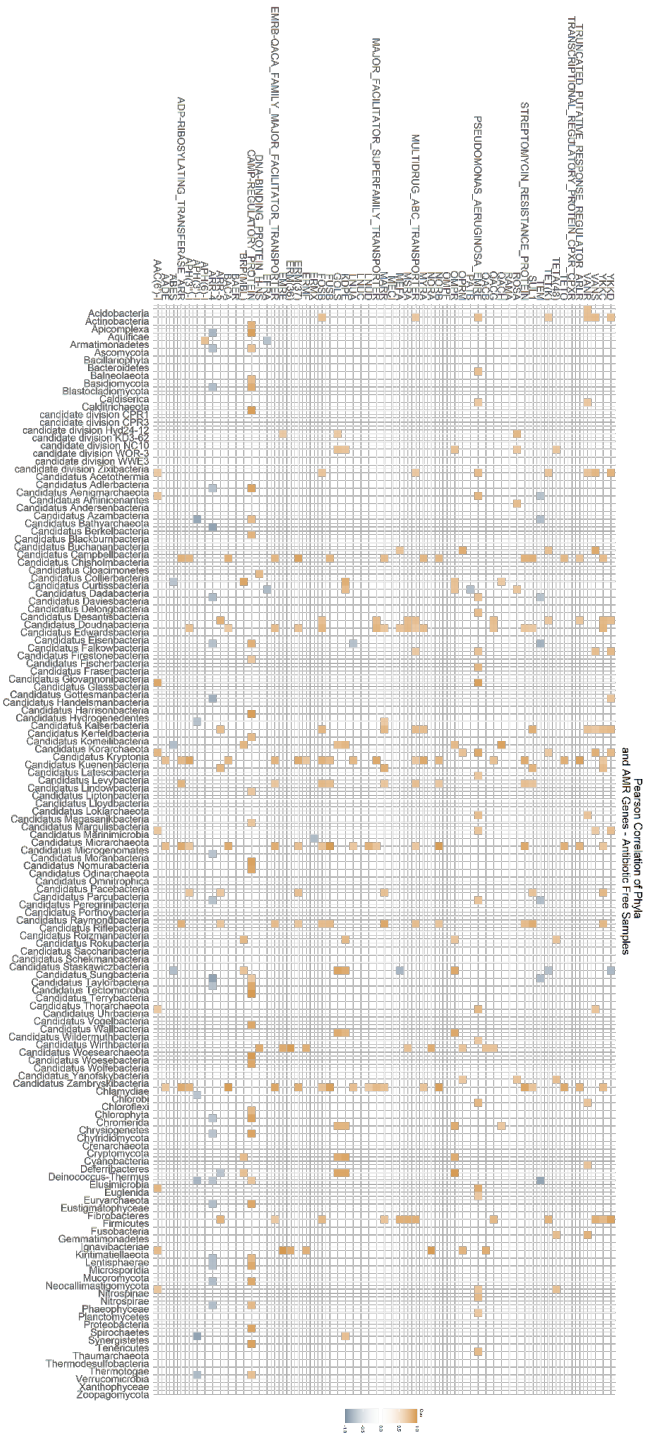
S1.24: Pairwise Pearson Correlation of Phyla for Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).



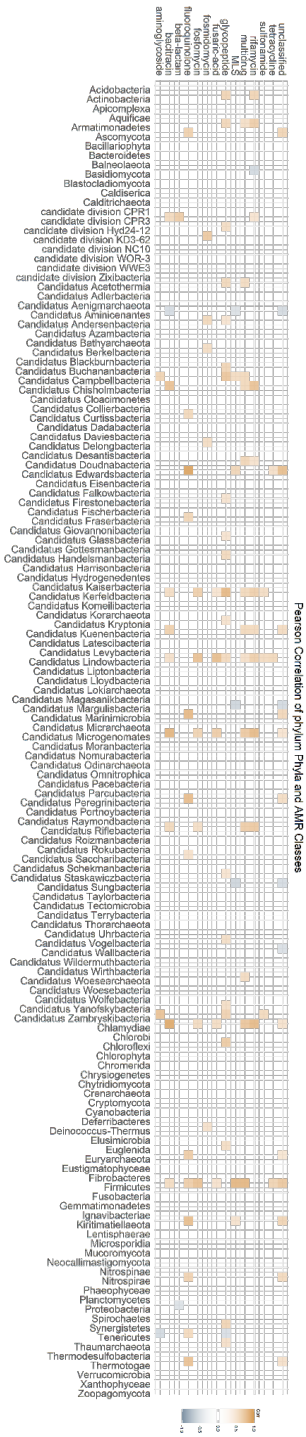
S1.25: Pearson Correlation of Phyla and AMR Genes Across All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



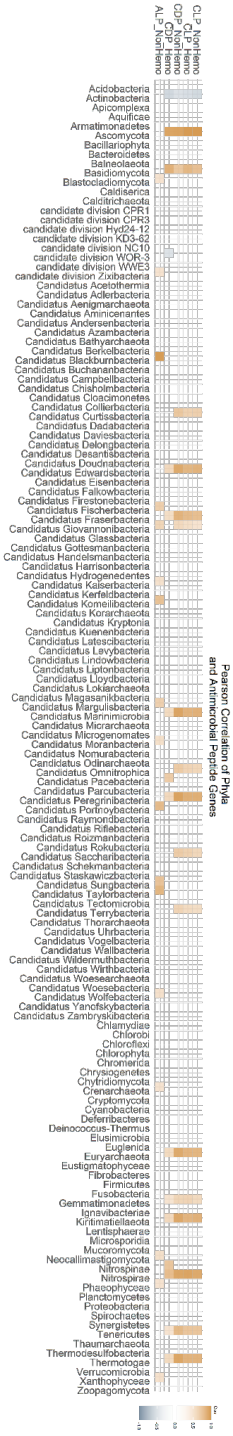
S1.26: Pearson Correlation of Phyla and AMR Genes for Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).



S1.27: Pearson Correlation of Phyla and AMR Genes for Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).



S1.28: Pearson Correlation of Phyla and AMR Class Across All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).



S1.29: Pearson Correlation of Phyla and AMP Class Across All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

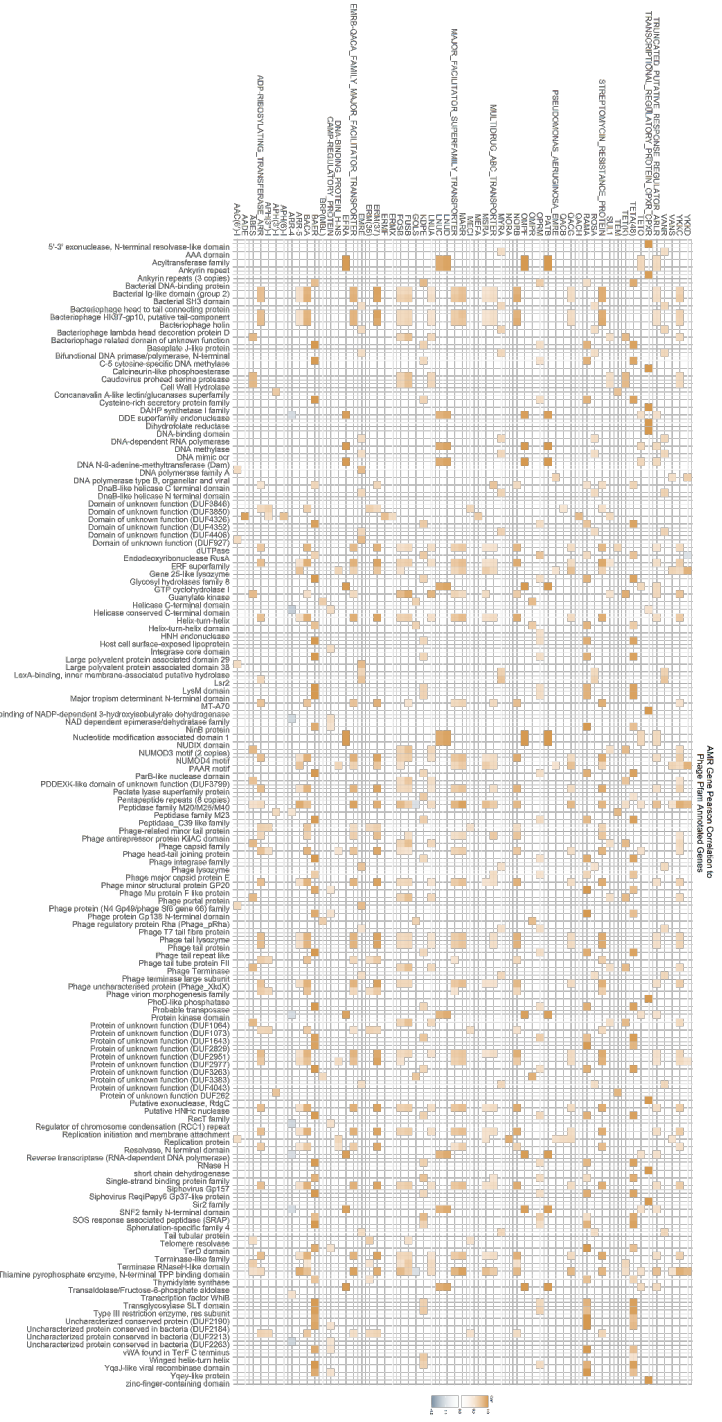


Figure S2.1: Pearson Correlation of AMR genes and Phage Pfam Annotations of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

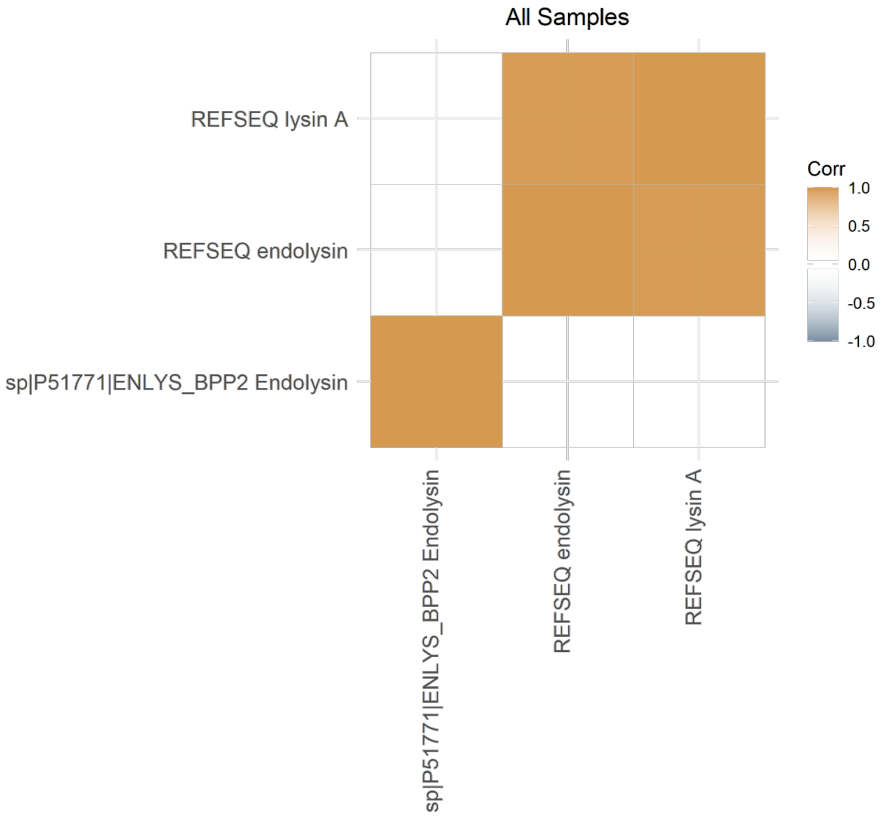


Figure S2.2: Significant Pairwise Pearson Correlation of Lysin VIral Orthologous Groups of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

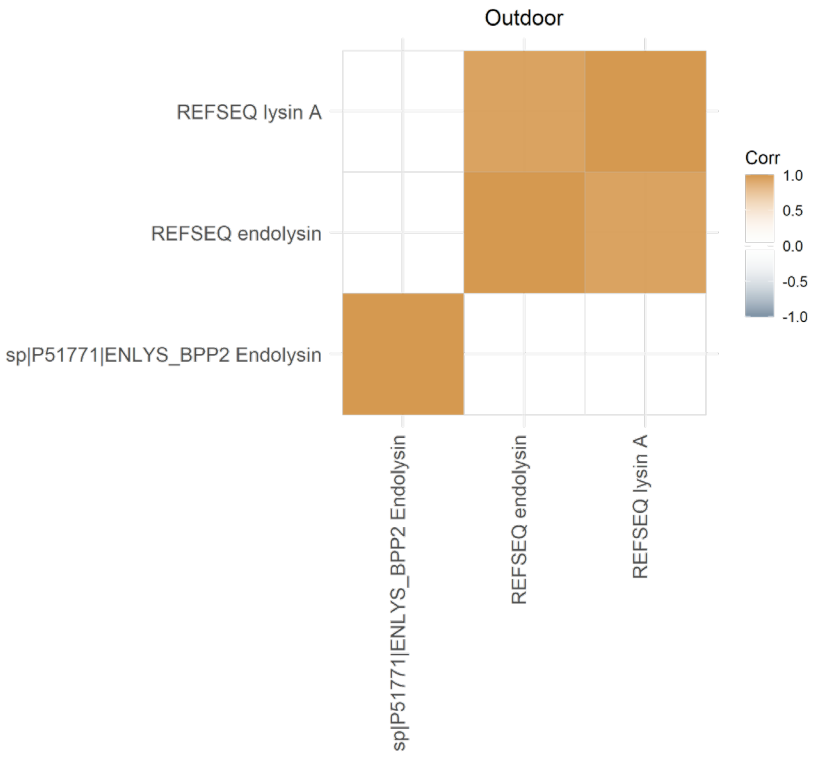


Figure S2.3: Significant Pairwise Pearson Correlation of Lysin VIral Orthologous Groups for Outdoor Samples of microbial communities of external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

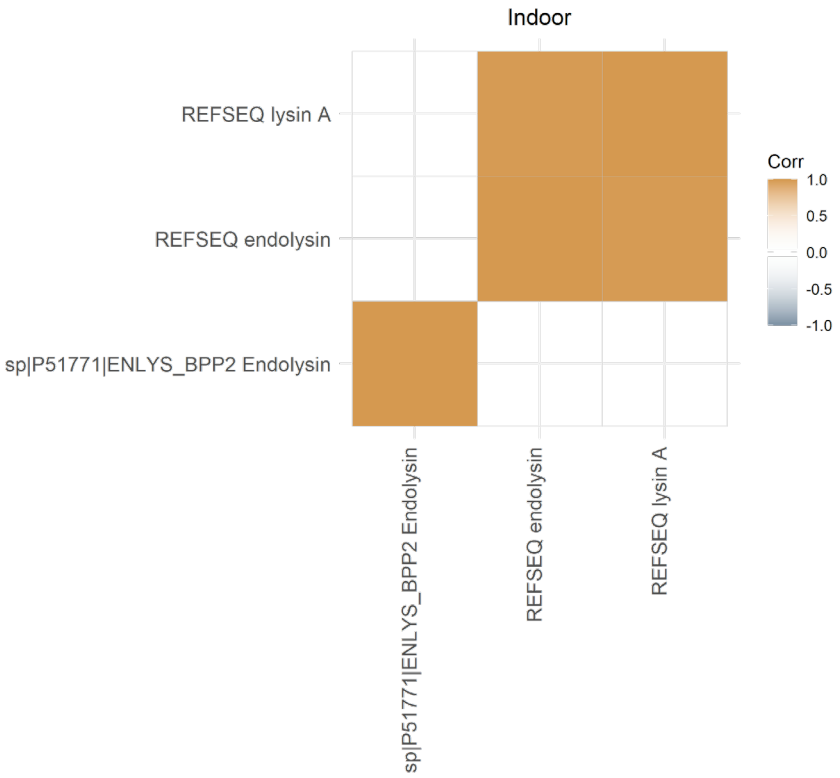


Figure S2.4: Significant Pairwise Pearson Correlation of Lysin VIral Orthologous Groups for Indoor Samples

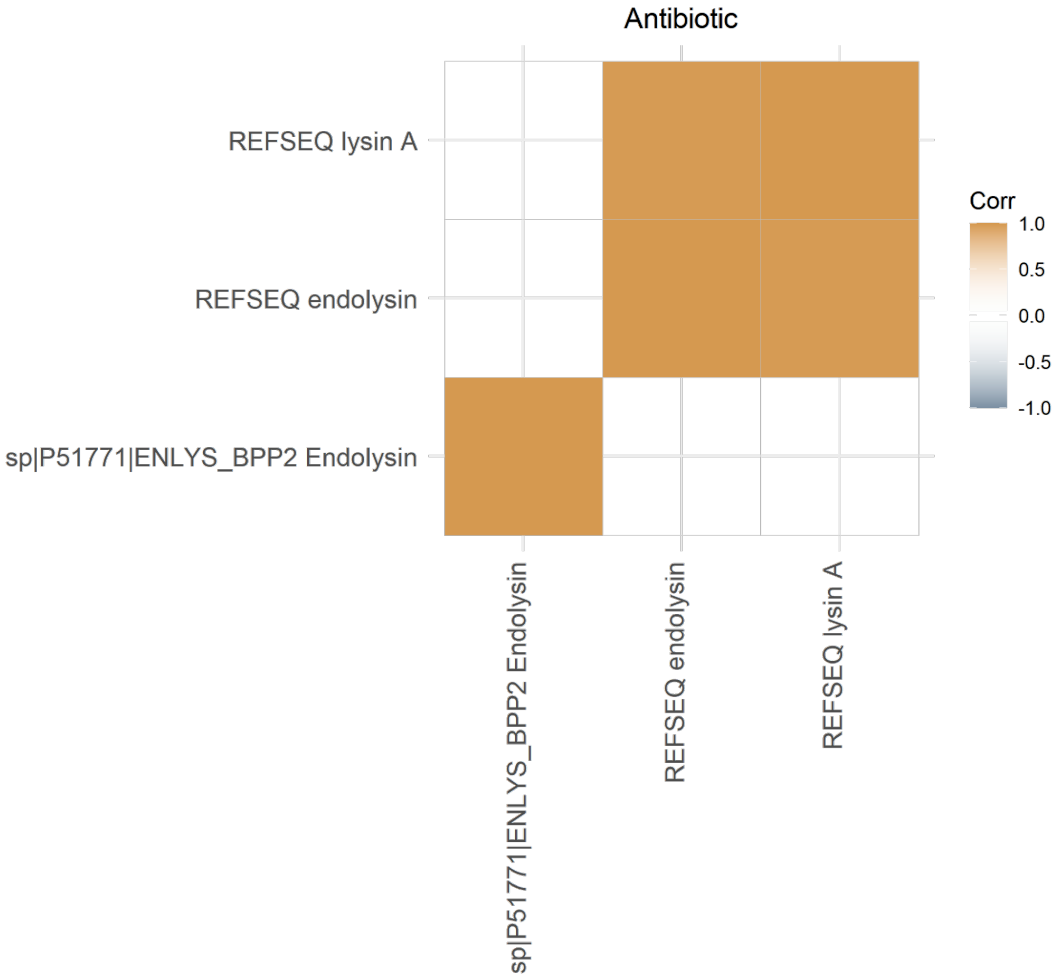


Figure S2.5: Significant Pairwise Pearson Correlation of Lysin VIral Orthologous Groups for Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).

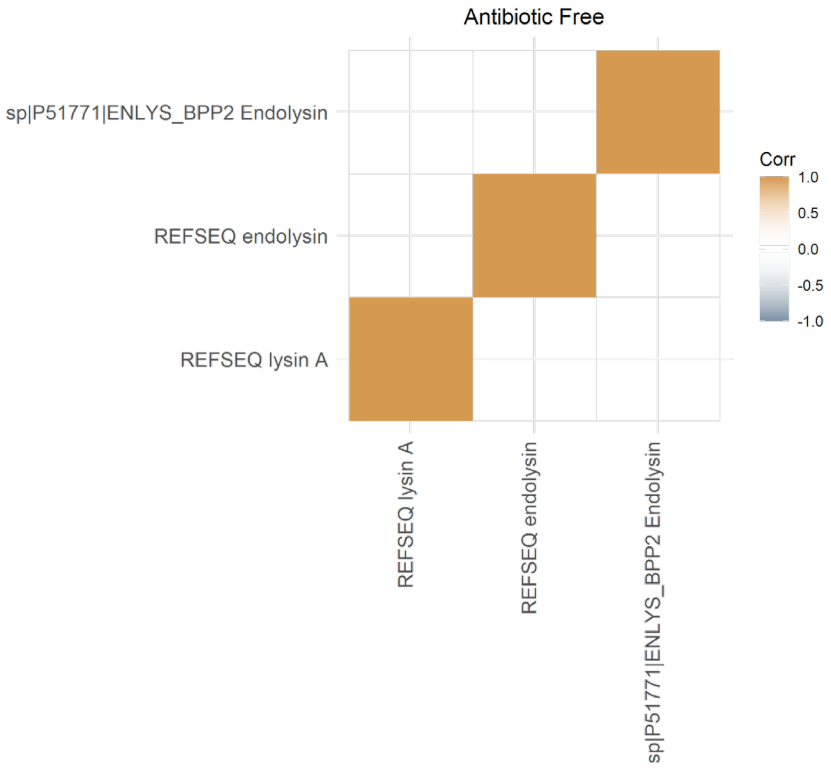


Figure S2.6: Significant Pairwise Pearson Correlation of Lysin VIral Orthologous Groups for Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).

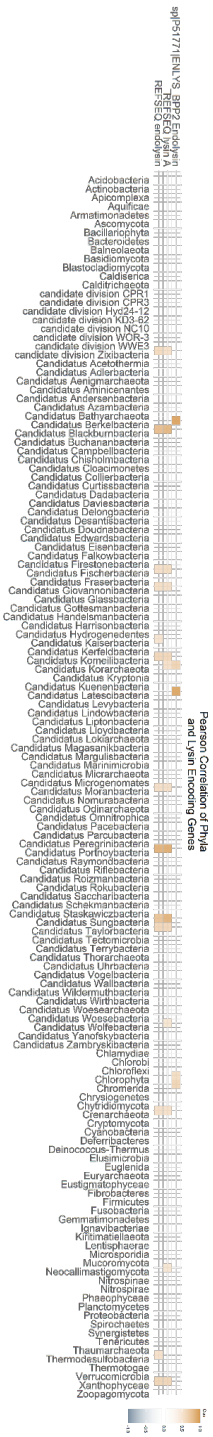


Figure S2.7: Significant Pearson Correlation of Lysin VIral Orthologous Groups and Phyla for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

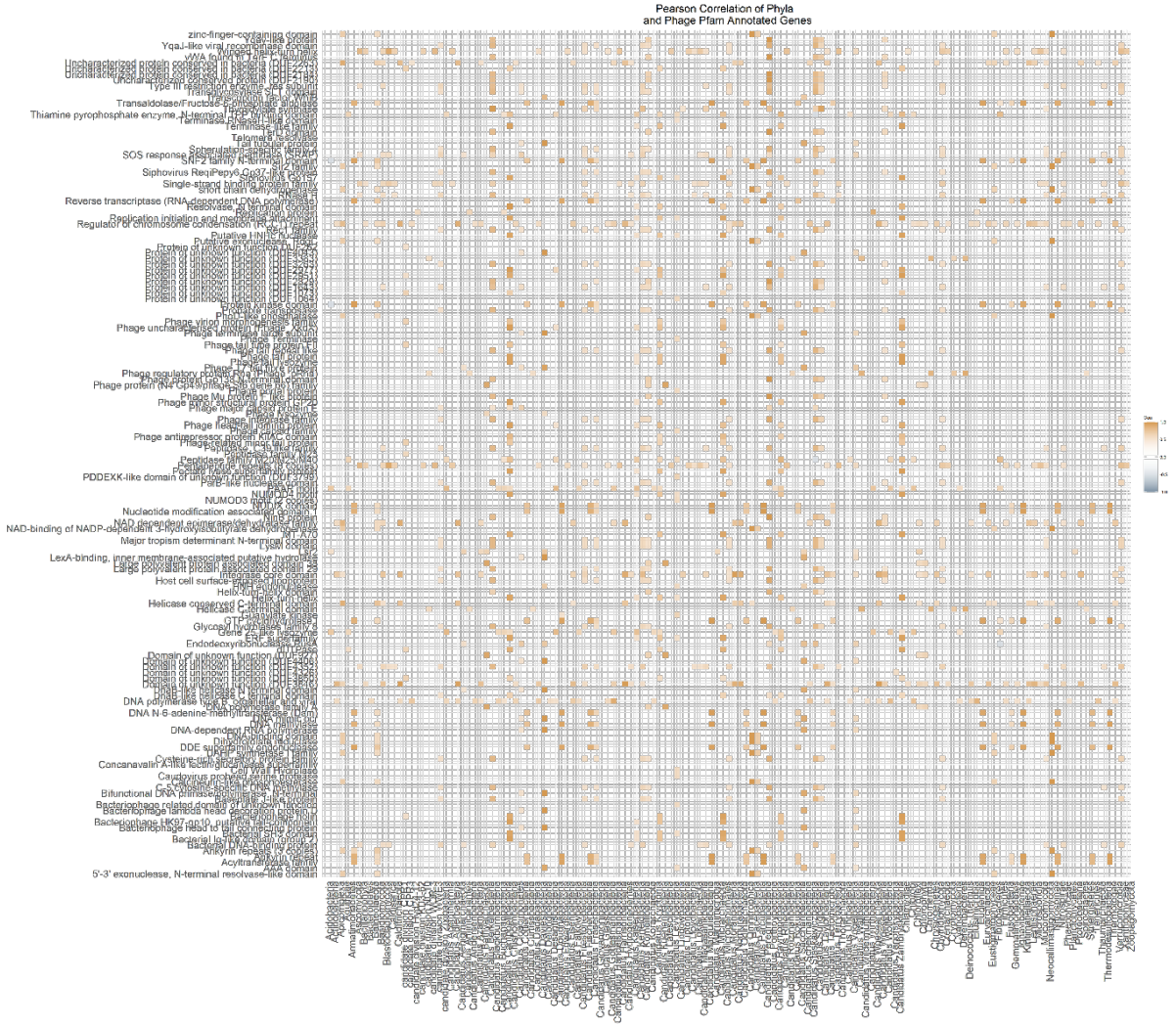


Figure S2.8: Significant Pearson Correlation of Phage Pfam Annotations and Phyla for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

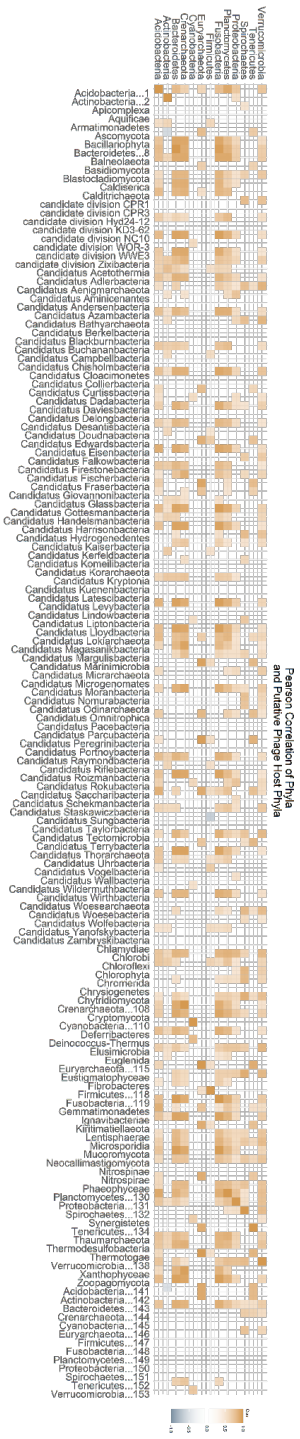


Figure S2.9: Significant Pearson Correlation of Phage Pfam Annotations and Putative Phage-Host Phyla for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

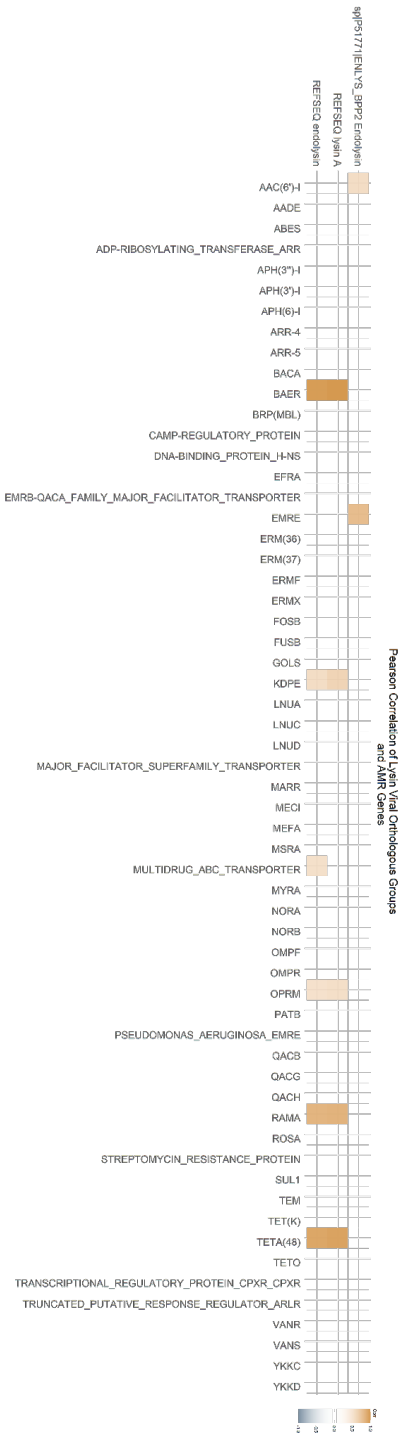


Figure S2.10: Significant Pearson Correlation of Lysin Viral Orthologous Groups and AMR Genes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

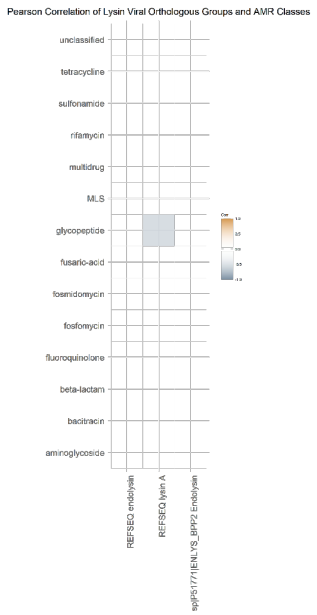


Figure S2.11: Significant Pearson Correlation of Lysin Viral Orthologous Groups and AMR Classes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

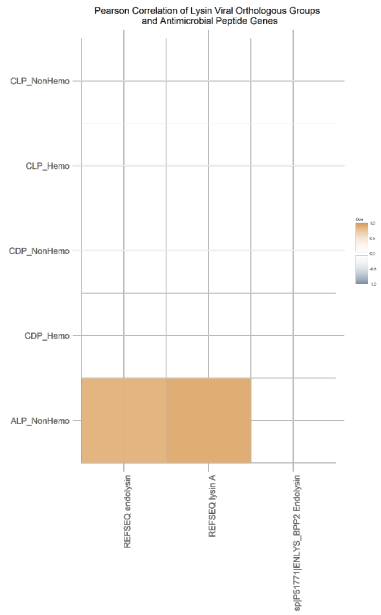


Figure S2.12: Significant Pearson Correlation of Lysin Viral Orthologous Groups and AMP Classes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

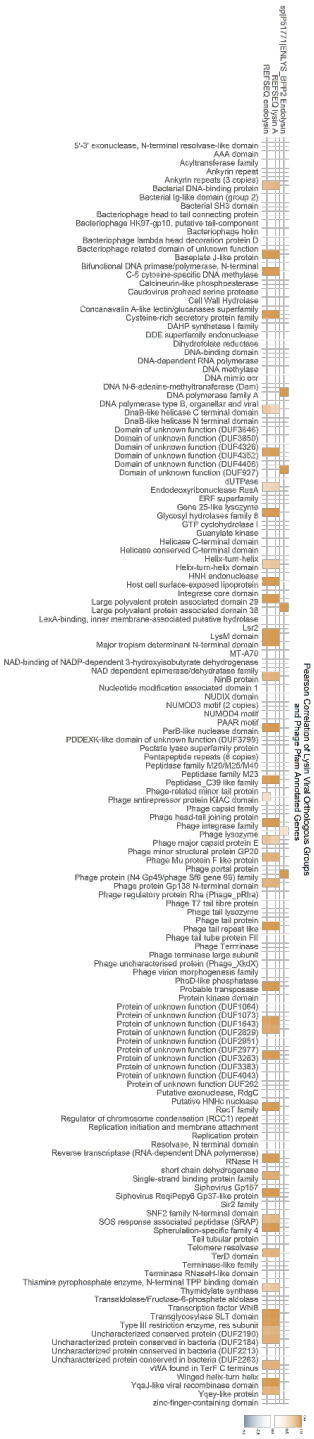


Figure S2.13: Significant Pearson Correlation of Lysin Viral Orthologous Groups and Pfam Annotated Phage Features for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

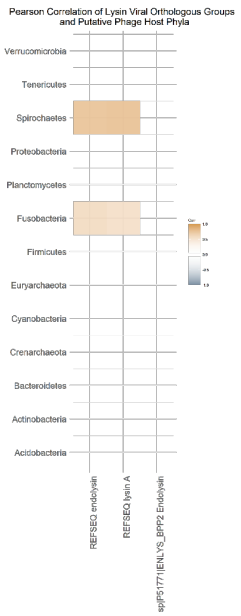


Figure S2.14: Significant Pearson Correlation of Lysin Viral Orthologous Groups and Putative Phage-Host Phyla for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

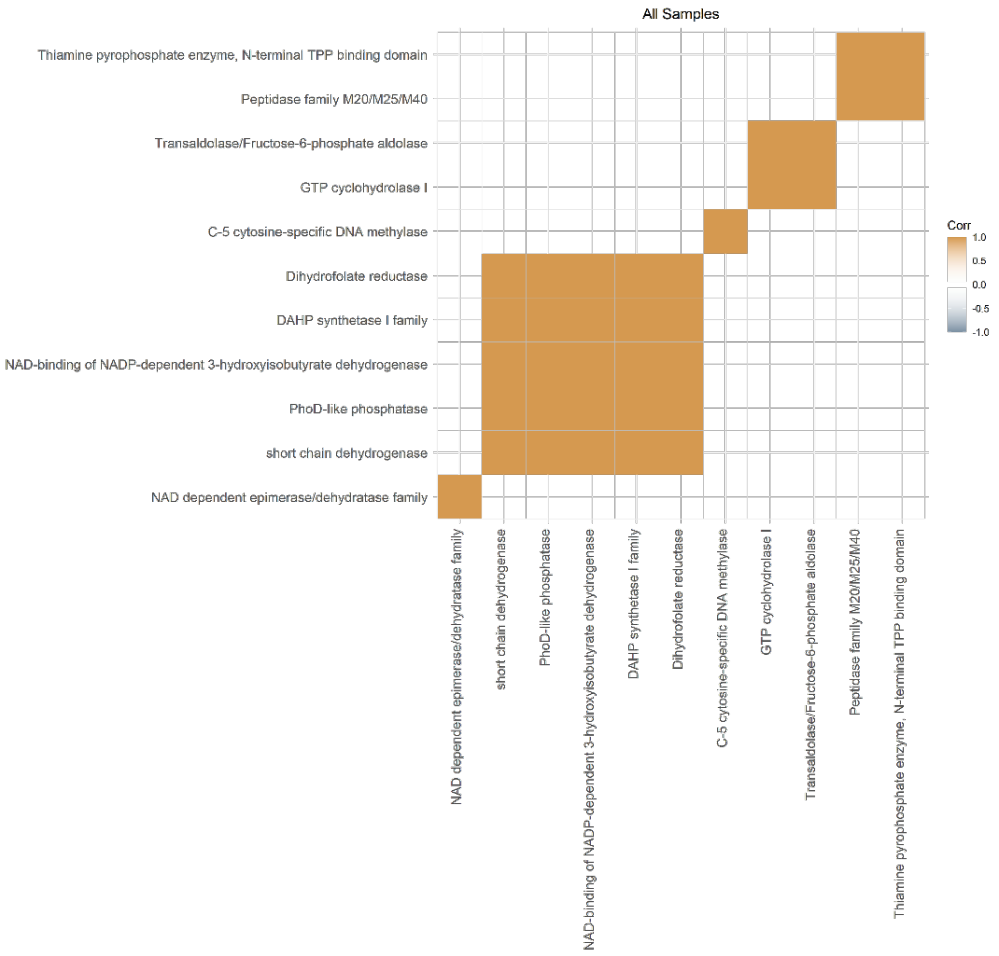


Figure S2.15: Significant Pairwise Pearson Correlation of Auxiliary Metabolic Genes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

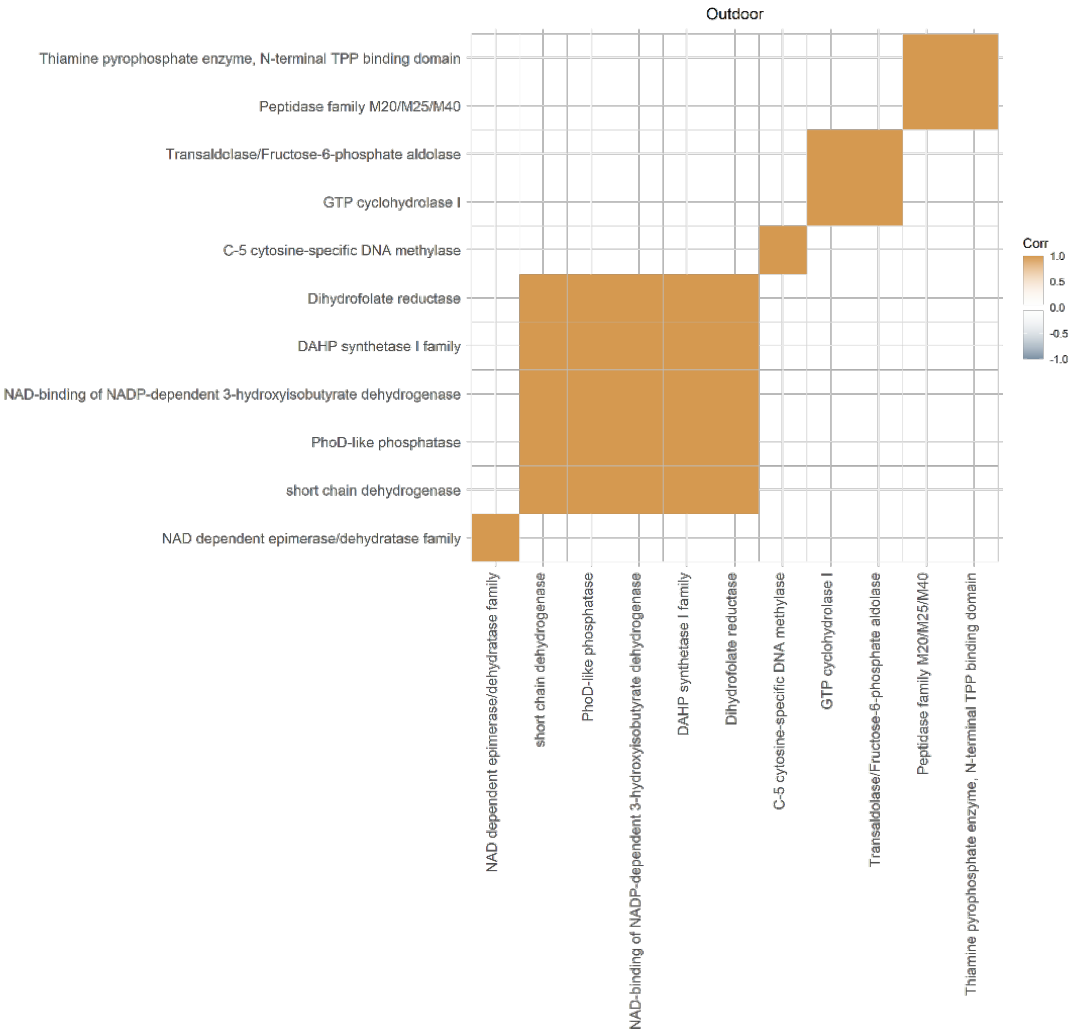


Figure S2.16: Significant Pairwise Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes for Outdoor Samples of microbial communities of external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

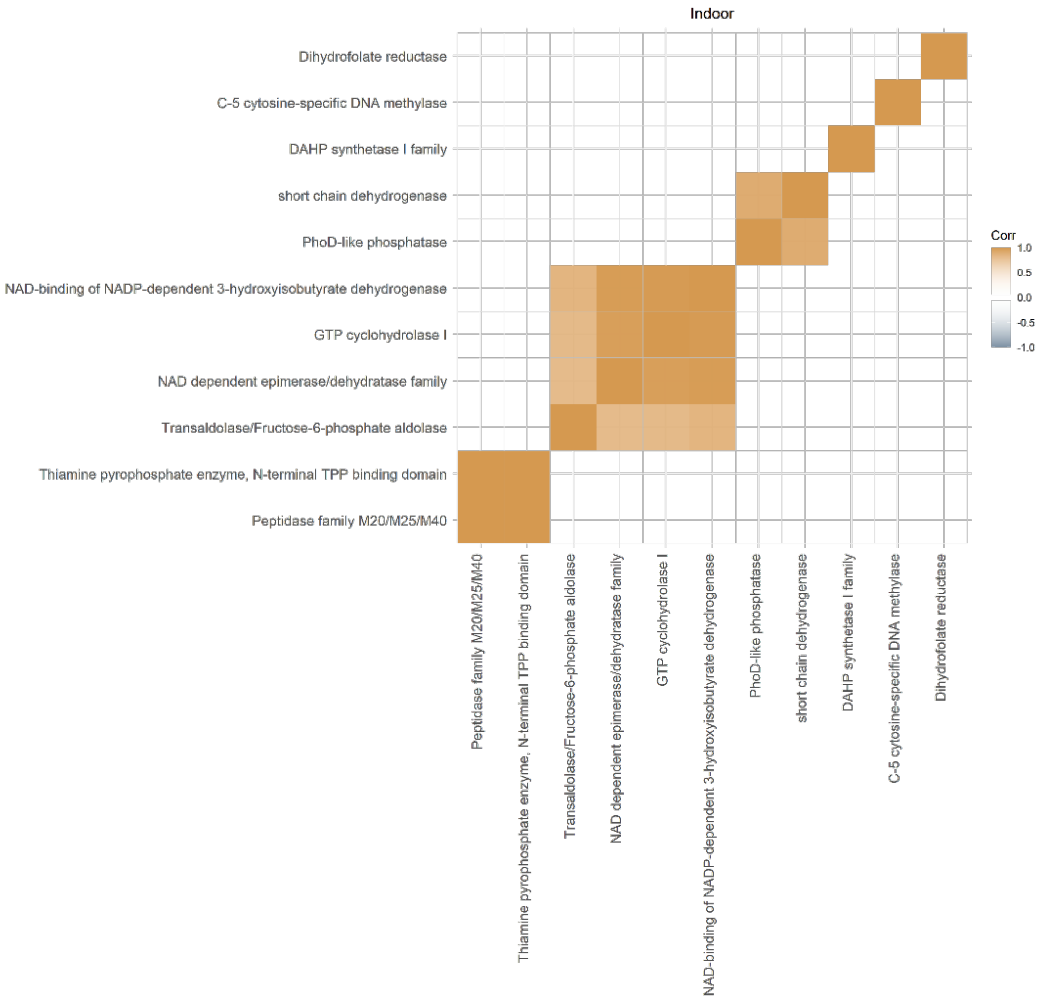


Figure S2.17: Significant Pairwise Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes for Indoor Samples of microbial communities of internal (indoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

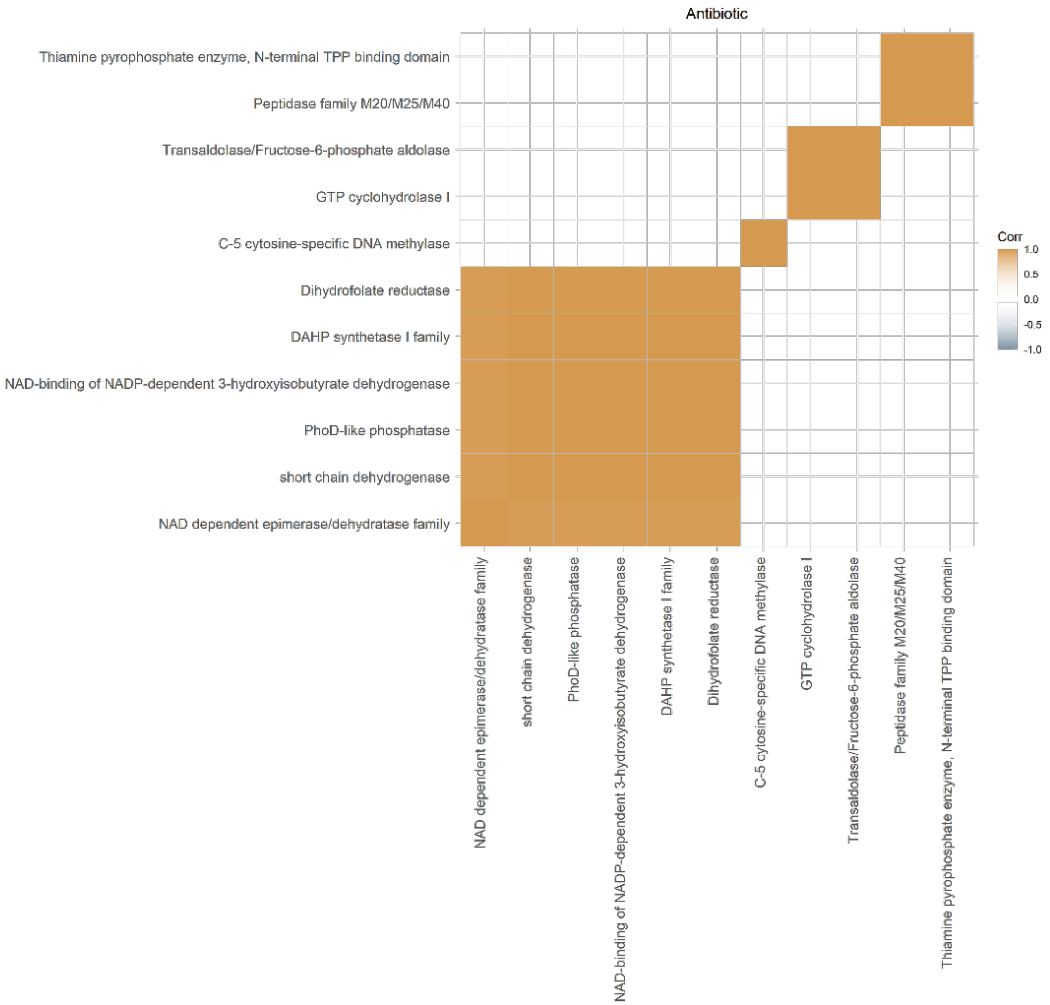


Figure S2.18: Significant Pairwise Pearson Correlation of Auxiliary Metabolic Genes for Antibiotic Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic).

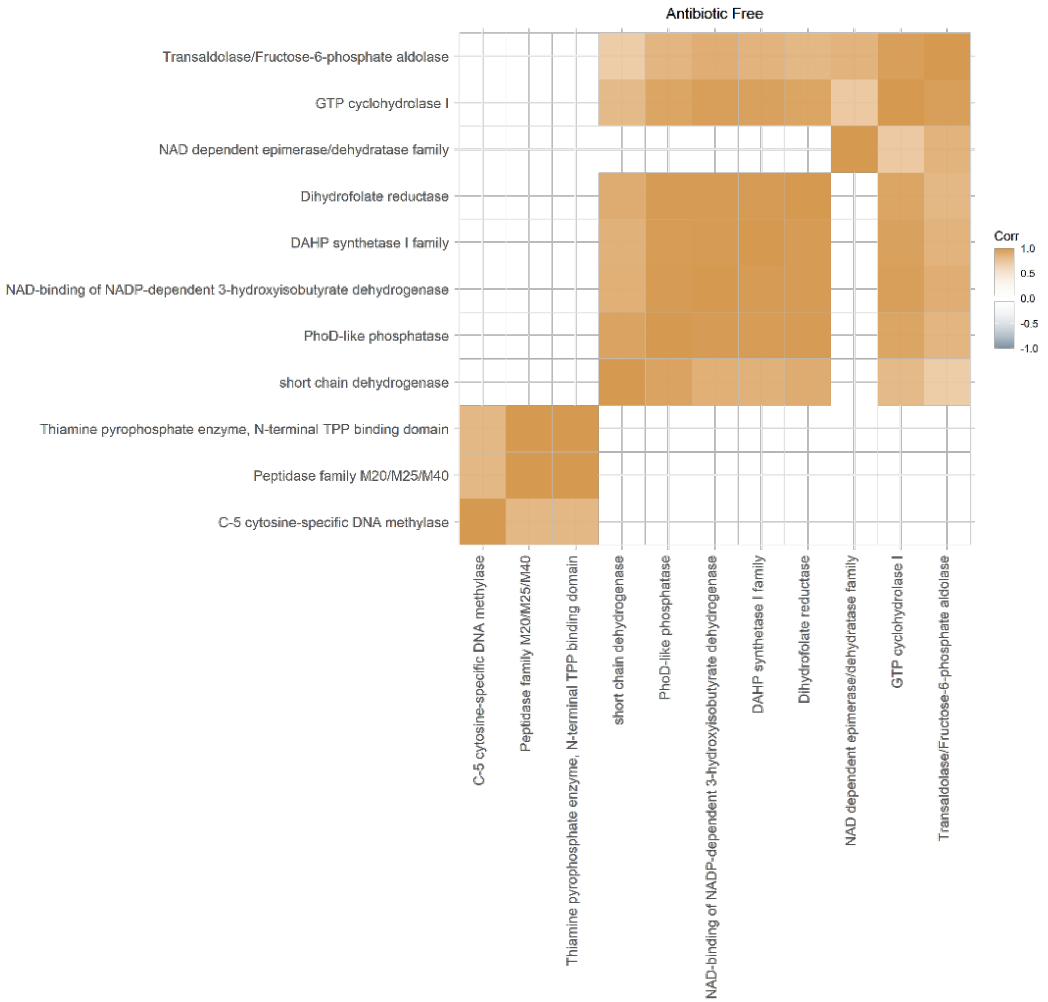


Figure S2.19: Significant Pairwise Pearson Correlation of Auxiliary Metabolic Genes for Antibiotic-Free Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who did not administer antibiotics to their BYP flocks (Antibiotic-Free).

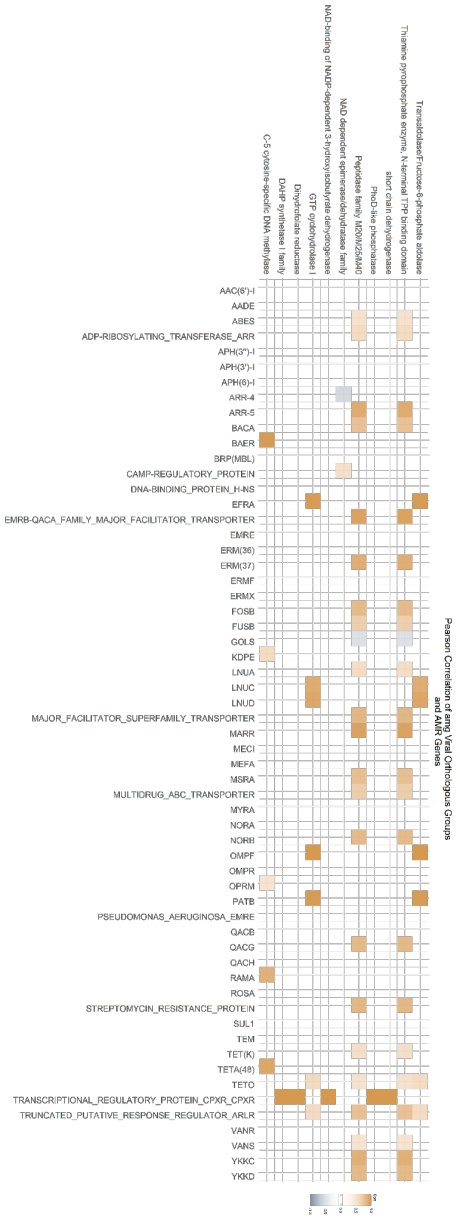


Figure S2.20: Significant Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes and AMR Genes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

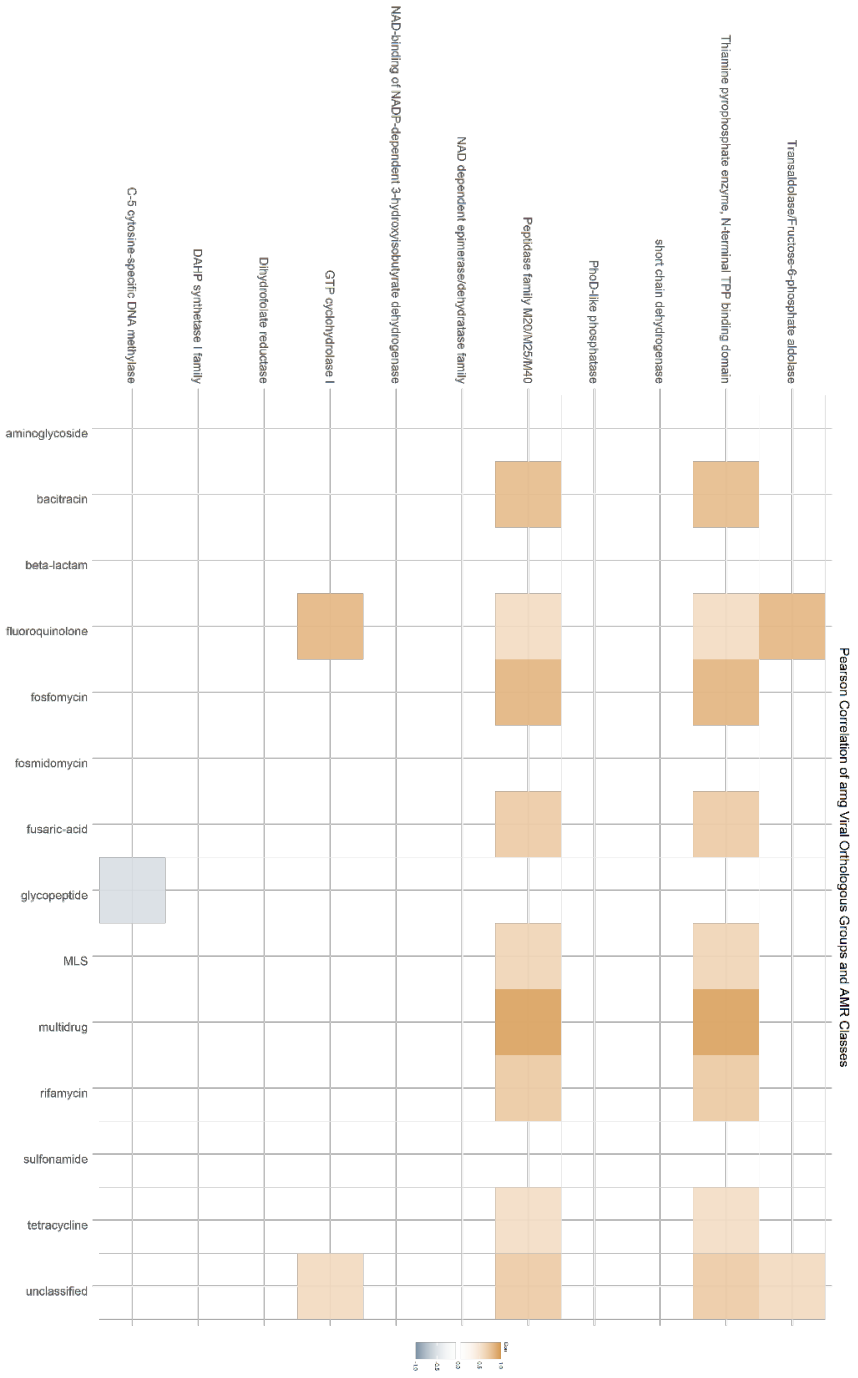


Figure S2.21: Significant Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes and AMR Classes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

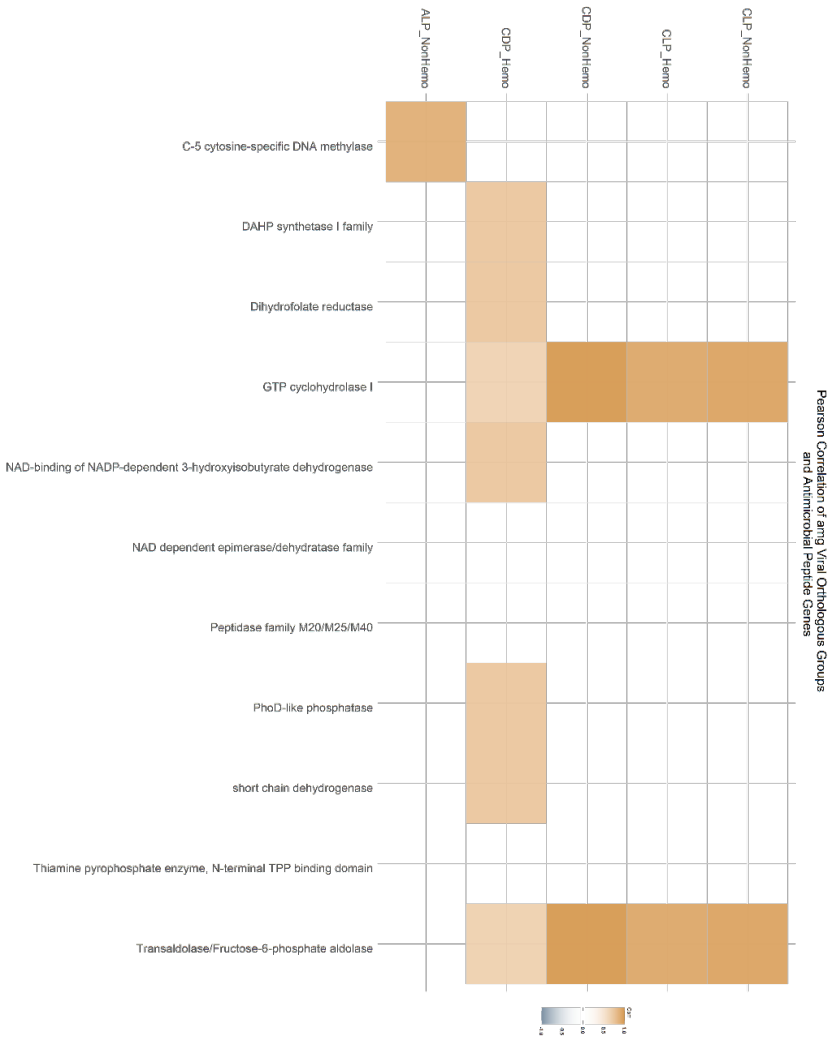


Figure S2.22: Significant Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes and AMP Classes for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

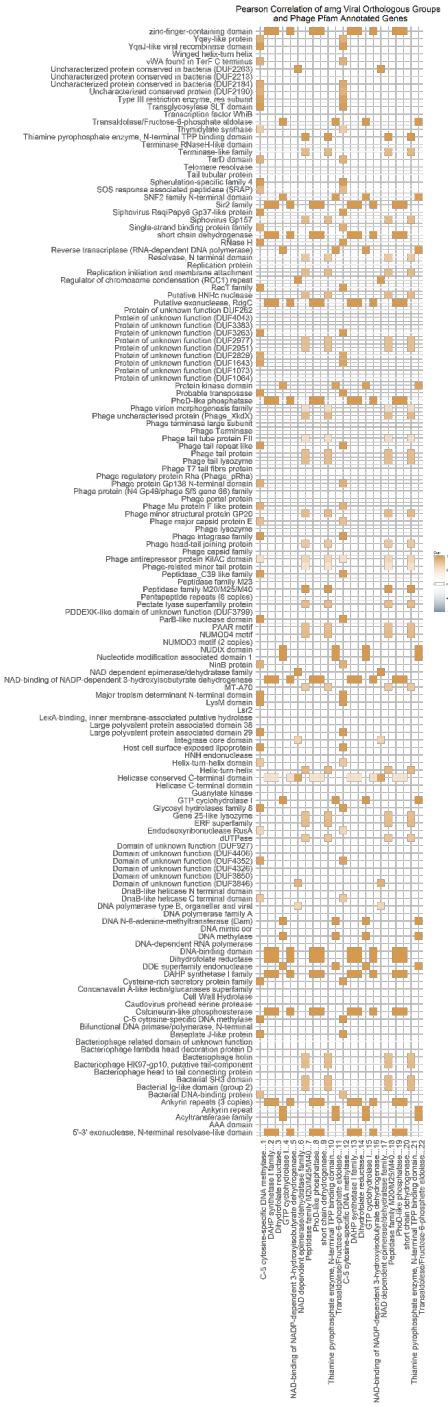


Figure S2.23: Significant Pearson Correlation of Auxiliary Metabolic Genes and Pfam Annotated Phage Features for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

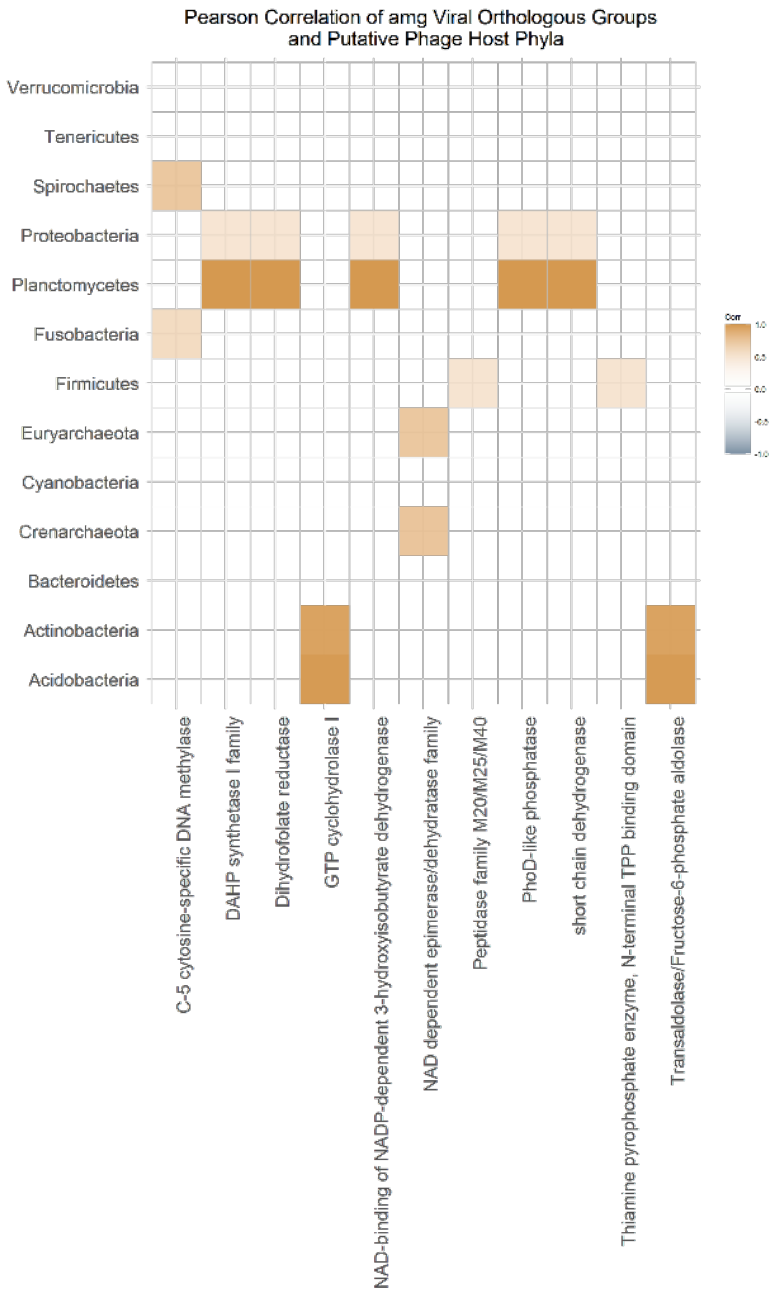


Figure S2.24: Significant Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes and Putative Phage-Host Phyla for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

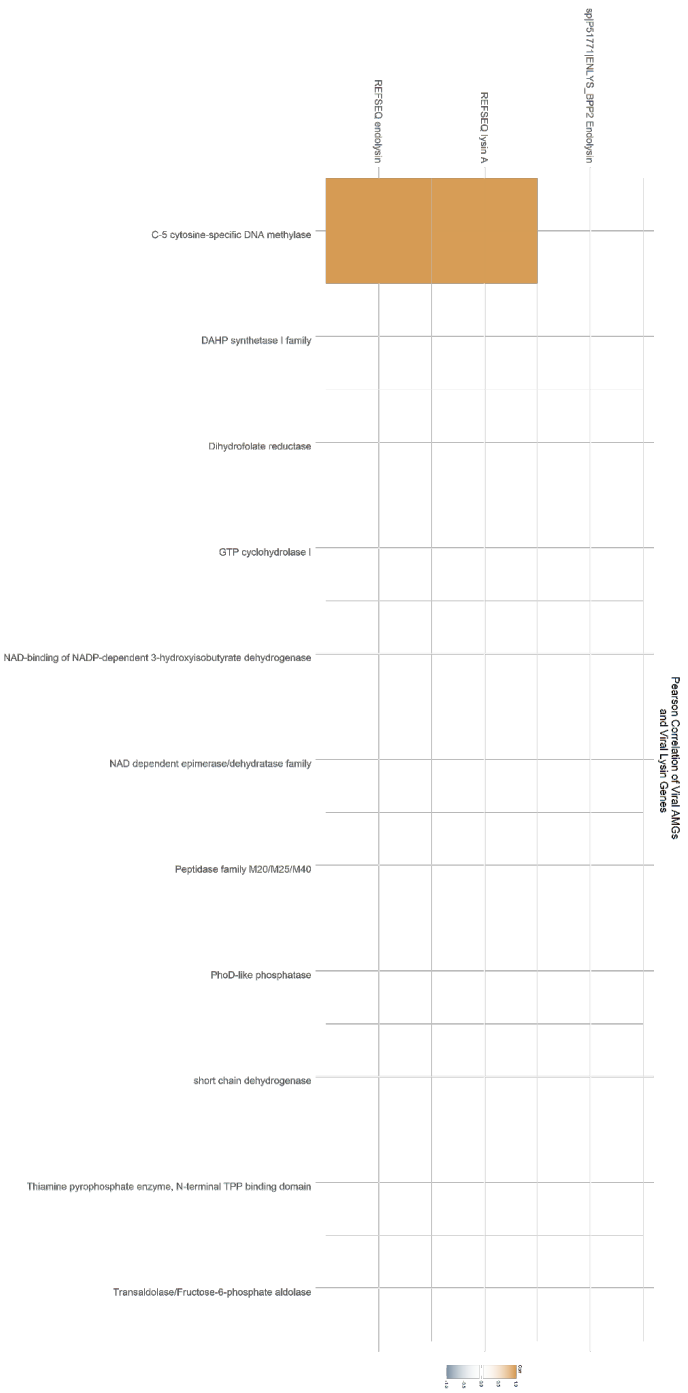


Figure S2.25: Significant Pearson Correlation of Phage-Encoded Auxiliary Metabolic Genes and Lysin Viral Orthologous Groups for All Samples of microbial communities of internal (indoor) and external (outdoor) door frames of BYP-owning homes who utilized antibiotics for treatment of their BYP flocks (Antibiotic) or did not administer antibiotics to their BYP flocks (Antibiotic-Free).

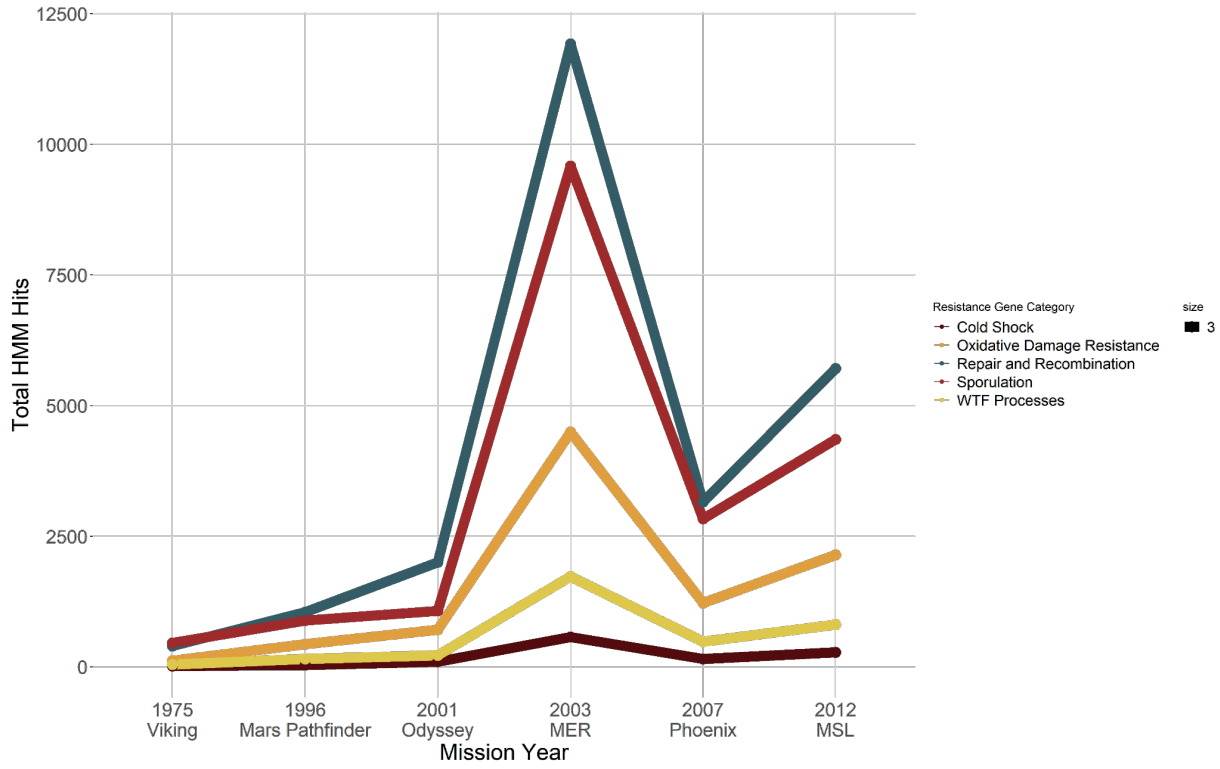


Figure S3.1: Total Cumulative HMM Hits of Resistance Gene Categories by Mission Year of bacterial isolates collected using NASA Standard Assay from spacecraft hardware at NASA JPL during spacecraft assembly.

Table S3.1: Total HMM Hit Abundances Per Mission of Resistance Gene Categories

Mission	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
Mars Pathfinder	37	433	1048	888	154
MER	573	4493	11923	9584	1731
MSL	278	2145	5712	4351	812

Odyssey	99	713	2004	1071	223
Phoenix	154	1223	3156	2835	485
Viking	15	121	402	458	47
Total	1156	9128	24245	19187	3452

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52039wB12_MPF24	4	51	116	113	17
JPL_52039wD12_MPF38	4	54	116	107	19
JPL_52039wE12_MPF19	4	47	119	107	20
JPL_52039wE6_MER189	3	26	101	3	10
JPL_52040wB7_MERTA35	6	43	127	110	14
JPL_52040wC7_MERTA1392	4	56	115	109	17
JPL_52041_MPF8	3	53	120	113	19
JPL_52042_MPF67	4	52	116	113	19
JPL_52052_MPF2	5	51	118	113	19
JPL_52039wA1_MER33	6	47	132	138	20
JPL_52039wA2_MER531	7	54	153	145	30
JPL_52039wA3_MER132	7	55	133	144	21
JPL_52039wA4_MER992	6	41	127	113	15
JPL_52039wA7_MER78	5	46	130	106	16
JPL_52039wA8_MER107	4	45	131	127	16
JPL_52039wB1_MER37	4	34	132	124	15
JPL_52039wB2_MER36	10	106	262	220	34
JPL_52039wB3_MER116	7	54	133	139	20
JPL_52039wB4_MER128	17	52	123	150	25
JPL_52039wB7_MER46	7	60	121	147	33
JPL_52039wB8_MER196A	3	32	123	32	6

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52039wC1_MER13	8	41	131	103	17
JPL_52039wC2_MER26	4	43	134	117	14
JPL_52039wC3_MER156	4	48	120	108	18
JPL_52039wC4_MER145	16	53	119	130	28
JPL_52039wC7_MER89	4	49	118	114	17
JPL_52039wD3_MER112	4	45	110	108	20
JPL_52039wD4_MER153	15	54	119	124	23
JPL_52039wD7_MER10	4	42	109	107	21
JPL_52039wE1_MER172A	4	44	126	59	5
JPL_52039wE3_MER135A	4	42	109	110	20
JPL_52039wE4_MER110	16	54	119	147	24
JPL_52039wE7_MER20	5	50	117	49	17
JPL_52039wE8_AMY111	4	51	117	53	16
JPL_52039wF3_MER108	4	41	110	94	20
JPL_52039wF4_MER65	15	53	119	129	28
JPL_52039wF5_FAIRING19B12	5	27	101	5	10
JPL_52039wF6_MER74	7	42	141	157	20
JPL_52039wF7_MER6	6	53	127	126	24
JPL_52039wG10_MSL05812	6	53	130	157	33
JPL_52039wG3_MER73	7	66	136	158	25
JPL_52039wG4_MSL2591	3	22	112	6	10
JPL_52039wG7_MER47	6	49	125	51	18
JPL_52039wH1_MER118	7	56	133	144	22
JPL_52039wH2_MER193	3	39	112	10	9
JPL_52039wH3_MER100	6	48	128	139	23
JPL_52039wH4_MSL3162	5	27	102	4	10

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52039wH7_MER101	5	40	128	123	15
JPL_52039wH9_PF3F2	12	50	120	112	29
JPL_52040wA10_681	4	46	110	110	19
JPL_52040wA11_MERTA168	6	49	143	120	24
JPL_52040wA2_TA76	11	50	142	153	20
JPL_52040wA4_MERTA1381	3	43	119	122	19
JPL_52040wA5_FAIRINGW8B1	4	61	136	6	10
JPL_52040wA7_MERTA1375	4	48	117	111	17
JPL_52040wA8_MSL3003	3	31	106	3	10
JPL_52040wB10_MERTA170	7	48	128	117	24
JPL_52040wB11_MERTA176	5	56	115	107	16
JPL_52040wB12_KSC114	4	46	113	108	20
JPL_52040wB2_MER502	4	44	139	136	13
JPL_52040wB3_MER157	5	61	143	153	18
JPL_52040wB4_MERTA111	6	60	116	147	21
JPL_52040wB5_MPF76A	4	53	123	113	20
JPL_52040wB6_MERTA106	7	47	136	138	22
JPL_52040wB9_KSC386	4	53	126	113	17
JPL_52040wC11_MERTA154	7	50	119	67	19
JPL_52040wC12_MER48	6	47	132	138	20
JPL_52040wC2_MER542	16	88	275	321	41
JPL_52040wC3_TA104	4	53	112	113	17
JPL_52040wC4_MERTA13	8	44	127	66	12
JPL_52040wC5_MPF57	4	49	111	112	18
JPL_52040wC6_MERTA87	8	47	152	163	25
JPL_52040wC8_KSC640	5	43	111	108	18

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52040wD11_1542	5	51	119	113	18
JPL_52040wD12_MER36	5	53	131	110	17
JPL_52040wD2_MSL0471	3	47	91	5	12
JPL_52040wD4_MERTA1382	6	53	134	165	22
JPL_52040wD5_RA14A10	3	32	114	4	2
JPL_52040wD9_MERTA181	8	67	125	159	30
JPL_52040wE1_ODYSSEY48V2	3	27	103	7	4
JPL_52040wE10_MER165	4	49	116	122	19
JPL_52040wE11_AMY2411	4	42	114	109	21
JPL_52040wE2_MSL0472	4	59	111	4	12
JPL_52040wE4_MERTA86	8	50	117	141	29
JPL_52040wE5_TA332	11	28	99	0	8
JPL_52040wE7_MERTA97	3	25	97	2	10
JPL_52040wF1_P26	4	20	106	3	2
JPL_52040wF10_MER170	21	53	117	123	26
JPL_52040wF4_MERTA48	6	52	133	63	19
JPL_52040wF5_TA149	3	44	119	122	19
JPL_52040wF6_AMY1112	4	26	105	5	10
JPL_52040wF7_MERTA17	3	42	119	121	18
JPL_52040wG1_TA170	7	48	129	117	24
JPL_52040wG2_P86	2	25	129	12	5
JPL_52040wG4_FAIRING4G11	4	24	100	2	10
JPL_52040wG7_MERTA112	1	22	93	0	7
JPL_52040wG8_KSC591	4	55	115	113	17
JPL_52040wH1_TA172	6	68	126	155	26

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52040wH10_MER161	4	49	116	122	19
JPL_52040wH2_MERTA32b	5	43	109	8	11
JPL_52040wH3_MERTA14	2	26	91	1	3
JPL_52040wH5_MERTA107	8	47	135	146	25
JPL_52043_MERTA38	5	56	128	211	16
JPL_52047_MER131	5	45	137	139	13
JPL_52048_MER62	1	40	122	6	10
JPL_52039wA11_MSL06011	5	48	111	108	18
JPL_52039wA5_MSL359	5	28	98	5	3
JPL_52039wA6_MSL1851	4	39	122	94	15
JPL_52039wB11_MSL1401	6	41	145	168	14
JPL_52039wB5_MSL3211	2	37	85	2	4
JPL_52039wB6_MSL2001	14	51	118	134	24
JPL_52039wC5_MSL3141	3	28	96	4	10
JPL_52039wC6_MSL1602	4	47	113	109	18
JPL_52039wD1_1971	5	26	101	5	10
JPL_52039wD11_MSL107	3	24	100	2	10
JPL_52039wD2_1831	4	37	119	93	17
JPL_52039wD6_MSL22512	7	54	113	97	24
JPL_52039wE11_MSL0361	8	46	125	129	14
JPL_52039wE2_2511	4	43	113	111	21
JPL_52039wF1_20622	4	45	118	115	19
JPL_52039wF2_23611	4	44	113	111	21
JPL_52039wG11_P112	7	28	118	14	3
JPL_52039wG5_MSL17212	4	43	112	110	19
JPL_52039wG6_MER82	6	53	135	127	23
JPL_52039wG9_CFPSW53	6	64	137	160	22

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52039wH10_MSL0161	3	24	97	3	10
JPL_52039wH5_MSL17322	7	62	129	142	24
JPL_52039wH6_MER9	7	50	130	61	19
JPL_52040wA9_KSC422	8	104	236	243	39
JPL_52040wB8_KSC645	18	90	227	231	37
JPL_52040wC1_AMY312	4	45	112	108	19
JPL_52040wC9_KSC155	4	44	110	108	22
JPL_52040wD10_MERTA171	6	48	122	58	15
JPL_52040wD3_TA29	6	42	124	108	17
JPL_52040wD8_KSC657	6	50	130	103	21
JPL_52040wE3_TA1214	2	21	87	2	4
JPL_52040wE8_KSC351	15	51	119	130	25
JPL_52040wE9_341	12	103	237	167	43
JPL_52040wF3_TA127	2	23	83	2	4
JPL_52040wF8_KSC283	16	53	119	135	23
JPL_52040wF9_12812	6	56	141	140	27
JPL_52040wG11_MSL1791	6	49	128	116	27
JPL_52040wG9_MERTA152	7	50	134	148	26
JPL_52040wH7_KSC432	5	42	122	100	16
JPL_52040wH8_MSL1871	6	45	153	139	15
JPL_52040wH9_411	8	49	125	127	14
JPL_52044_KSC418	4	47	141	157	16
JPL_52045_2581A	4	56	129	114	17
JPL_52049_MSL3362	5	48	138	5	8
JPL_52050_MSL348	6	67	147	6	15
JPL_52039wA10_P30	3	28	97	3	2
JPL_52039wA9_P10	3	25	105	3	2

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52039wB9_P106	4	49	111	133	23
JPL_52039wC8_P25	4	27	104	4	2
JPL_52039wC9_P83	4	30	98	2	10
JPL_52039wD9_P75	4	25	99	3	10
JPL_52039wE10_P100	0	32	89	1	3
JPL_52039wE9_P107	16	54	121	127	24
JPL_52039wF9_P121	14	53	120	128	26
JPL_52039wG2_21411	4	59	117	134	25
JPL_52040wA12_P42	6	53	123	131	15
JPL_52040wA3_TA28	9	23	91	8	4
JPL_52040wD1_P7	3	32	119	11	4
JPL_52040wF11_KSC339	3	28	108	2	10
JPL_52040wF2_P97	3	30	110	7	4
JPL_52040wG3_MERTA13632	5	52	120	105	17
JPL_52040wH11_P67	6	56	133	138	19
JPL_52054_P18	8	57	139	131	23
JPL_52039wB10_AMY1912vial1	6	43	113	114	16
JPL_52039wC10_AMY1912	6	43	112	114	16
JPL_52039wC12_AMY214	5	31	103	3	10
JPL_52039wD10_AMY611	4	49	112	107	19
JPL_52039wD5_FAIRING10M22	6	44	119	112	20
JPL_52039wD8_AMY52	5	54	110	105	18
JPL_52039wE5_FAIRING12A4	6	54	129	111	27
JPL_52039wF10_AMY322	5	48	132	121	24

Table S3.2: HMM Hit Abundances of Each Isolate Genome for Resistance Gene Categories

genome_or_bin	Cold Shock	Oxidative Damage Resistance	Repair and Recombination	Sporulation	WTF Processes
JPL_52039wF11_MSL00412	0	29	89	1	3
JPL_52039wF8_AMY2812	4	44	109	111	20
JPL_52039wG1_MER166	4	57	115	110	19
JPL_52039wG8_AMY511	6	51	130	149	20
JPL_52039wH8_PF3F12	12	55	146	143	23
JPL_52040wA1_MERTA114	7	53	127	115	24
JPL_52040wA6_MERTA82	7	53	127	115	24
JPL_52040wB1_AMY71	6	37	121	104	18
JPL_52040wC10_MER180	4	46	134	117	15
JPL_52040wD6_AMY1312	18	54	119	129	30
JPL_52040wD7_MERTA18	4	26	101	6	10
JPL_52040wE12_MER51	8	48	137	153	14
JPL_52040wE6_AMY152	6	53	136	169	19
JPL_52040wG10_MER169	4	57	118	110	19
JPL_52040wG5_MERTA813	4	46	142	117	17
JPL_52040wG6_PF24B2	7	52	125	147	21
JPL_52040wH4_FAIRING3B12	4	49	122	109	18
JPL_52040wH6_PF4F21	6	47	128	143	21
JPL_52039wA12_TPS1191	4	38	117	114	19
JPL_52039wC11_TPS8131	5	42	145	165	14
JPL_52039wH11_TPS1431	6	41	140	179	14

Table S3.3: Summary Statistics of Resistance Gene Category HMM Hits for Each Mission

mission	CATEGORY	total_hmm_hits_per_category	min_hmm_hits_per_category	max_hmm_hits_per_category	mean_hmm_hits_per_category	median_hmm_hits_per_category	hmm_hits_per_genome_per_category_per_mission
Mars Pathfinder	Cold Shock	37	0	5.00	1.37	1.0	0.24
Mars Pathfinder	Oxidative Damage Resistance	433	0	11.00	1.27	1.0	0.02
Mars Pathfinder	Repair and Recombination	1048	0	20.00	0.97	1.0	0.01
Mars Pathfinder	Sporulation	888	0	8.00	1.16	1.0	0.01
Mars Pathfinder	WTF Processes	154	0	5.00	0.90	0.0	0.06
MER	Cold Shock	573	0	20.00	1.97	1.0	0.17
MER	Oxidative Damage Resistance	4493	0	13.00	1.22	0.0	0.02
MER	Repair and Recombination	11923	0	28.00	1.02	1.0	0.01
MER	Sporulation	9584	0	45.00	1.16	1.0	0.01
MER	WTF Processes	1731	0	13.00	0.94	0.0	0.06
MSL	Cold Shock	278	0	16.00	2.06	1.0	0.16
MSL	Oxidative Damage Resistance	2145	0	18.00	1.25	0.0	0.02
MSL	Repair and Recombination	5712	0	24.00	1.06	1.0	0.01
MSL	Sporulation	4351	0	22.00	1.14	1.0	0.01
MSL	WTF Processes	812	0	13.00	0.95	0.0	0.06
Odyssey	Cold Shock	99	0	15.00	1.83	1.0	0.18
Odyssey	Oxidative Damage Resistance	713	0	11.00	1.04	0.0	0.03

Table S3.3: Summary Statistics of Resistance Gene Category HMM Hits for Each Mission

mission	CATEGORY	total_hmm_hits_per_category	min_hmm_hits_per_category	max_hmm_hits_per_category	mean_hmm_hits_per_category	median_hmm_hits_per_category	hmm_hits_per_genome_per_category_per_mission
Odyssey	Repair and Recombination	2004	0	17.00	0.93	1.0	0.01
Odyssey	Sporulation	1071	0	11.00	0.70	0.0	0.02
Odyssey	WTF Processes	223	0	10.00	0.65	0.0	0.08
Phoenix	Cold Shock	154	0	17.00	1.97	1.0	0.17
Phoenix	Oxidative Damage Resistance	1223	0	12.00	1.24	1.0	0.02
Phoenix	Repair and Recombination	3156	0	20.00	1.01	1.0	0.01
Phoenix	Sporulation	2835	0	16.00	1.28	1.0	0.01
Phoenix	WTF Processes	485	0	11.00	0.98	0.0	0.05
Viking	Cold Shock	15	0	5.00	1.67	1.0	0.20
Viking	Oxidative Damage Resistance	121	0	9.00	1.06	0.5	0.02
Viking	Repair and Recombination	402	0	14.00	1.12	1.0	0.01
Viking	Sporulation	458	0	19.00	1.80	1.0	0.01
Viking	WTF Processes	47	0	6.00	0.82	0.0	0.06

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Cold Shock	Bacillus	181	0	6	1.47	1
Cold Shock	Solibacillus	4	0	3	1.33	1
Cold Shock	Staphylococcus	69	0	6	1.35	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Cold Shock	Agromyces	3	0	3	1.00	0
Cold Shock	Alkalihalobacillus	19	0	7	1.58	1
Cold Shock	Bhargavaea	4	0	3	1.33	1
Cold Shock	Brevibacillus	39	0	7	2.17	1
Cold Shock	Carnobacterium	11	0	10	3.67	1
Cold Shock	Curtobacterium	6	0	3	1.00	0
Cold Shock	Cytobacillus	52	0	7	1.93	1
Cold Shock	Exiguobacterium	3	0	2	1.00	1
Cold Shock	Fictibacillus	18	0	5	2.00	1
Cold Shock	Kocuria	17	0	9	1.89	0
Cold Shock	Mesobacillus	47	0	7	2.24	1
Cold Shock	Metabacillus	26	0	7	2.17	1
Cold Shock	Microbacterium	14	0	4	1.17	0
Cold Shock	Neobacillus	36	0	10	2.00	1
Cold Shock	Niallia	18	0	5	1.50	1
Cold Shock	Nocardioides	2	0	2	0.67	0
Cold Shock	Oceanobacillus	4	0	3	1.33	1
Cold Shock	Paenibacillus	102	0	14	1.89	1
Cold Shock	Peribacillus	13	0	6	2.17	1
Cold Shock	Planococcus	8	0	7	2.67	1
Cold Shock	Priestia	211	0	20	4.69	1
Cold Shock	Psychrobacillus	13	0	6	2.17	1
Cold Shock	Rummeliibacillus	6	0	5	2.00	1
Cold Shock	Sphingomonas	4	0	3	1.33	1
Cold Shock	Sphingopyxis	3	0	2	1.00	1
Cold Shock	Sporosarcina	14	0	4	1.56	1
Cold Shock	Stenotrophomonas	4	0	3	1.33	1
Cold Shock	Streptococcus	7	0	6	2.33	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Cold Shock	Streptomyces	7	0	4	1.17	0
Cold Shock	Terribacillus	4	0	3	1.33	1
Cold Shock	Ureibacillus	3	0	2	1.00	1
Cold Shock	Caldibacillus	12	0	3	1.33	1
Cold Shock	Cellulosimicrobium	7	0	6	2.33	1
Cold Shock	Cupriavidus	6	0	5	2.00	1
Cold Shock	Domibacillus	7	0	6	2.33	1
Cold Shock	Heyndrickxia	12	0	5	2.00	1
Cold Shock	Lederbergia	4	0	3	1.33	1
Cold Shock	Micrococcus	6	0	2	0.67	0
Cold Shock	Ralstonia	5	0	4	1.67	1
Cold Shock	Weizmannia	5	0	4	1.67	1
Cold Shock	Alkalihalophilus	6	0	5	2.00	1
Cold Shock	Georgenia	3	0	3	1.00	0
Cold Shock	Rothia	0	0	0	0.00	0
Cold Shock	Sutcliffiella	6	0	5	2.00	1
Cold Shock	Hydrogenophaga	8	0	7	2.67	1
Cold Shock	NA	97	0	16	2.94	1
Oxidative Damage Resistance	Bacillus	1973	0	12	1.27	1
Oxidative Damage Resistance	Solibacillus	56	0	11	1.47	1
Oxidative Damage Resistance	Staphylococcus	503	0	12	0.78	0
Oxidative Damage Resistance	Agromyces	32	0	7	0.84	0

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Oxidative Damage Resistance	Alkalihalobacillus	196	0	11	1.29	1
Oxidative Damage Resistance	Bhargavaea	51	0	12	1.34	1
Oxidative Damage Resistance	Brevibacillus	266	0	10	1.17	1
Oxidative Damage Resistance	Carnobacterium	28	0	4	0.74	0
Oxidative Damage Resistance	Curtobacterium	57	0	5	0.75	0
Oxidative Damage Resistance	Cytobacillus	455	0	11	1.33	1
Oxidative Damage Resistance	Exiguobacterium	39	0	9	1.03	0
Oxidative Damage Resistance	Fictibacillus	127	0	9	1.11	0
Oxidative Damage Resistance	Kocuria	94	0	8	0.82	0
Oxidative Damage Resistance	Mesobacillus	371	0	12	1.39	1
Oxidative Damage Resistance	Metabacillus	236	0	13	1.55	1
Oxidative Damage Resistance	Microbacterium	100	0	5	0.66	0
Oxidative Damage Resistance	Neobacillus	278	0	11	1.22	0

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Oxidative Damage Resistance	Niallia	168	0	9	1.11	0
Oxidative Damage Resistance	Nocardioides	25	0	4	0.66	0
Oxidative Damage Resistance	Oceanobacillus	48	0	9	1.26	1
Oxidative Damage Resistance	Paenibacillus	869	0	12	1.27	0
Oxidative Damage Resistance	Peribacillus	115	0	10	1.51	1
Oxidative Damage Resistance	Planococcus	47	0	8	1.24	1
Oxidative Damage Resistance	Priestia	764	0	9	1.34	1
Oxidative Damage Resistance	Psychrobacillus	99	0	9	1.30	0.5
Oxidative Damage Resistance	Rummeliibacillus	60	0	10	1.58	1
Oxidative Damage Resistance	Sphingomonas	59	0	7	1.55	1
Oxidative Damage Resistance	Sphingopyxis	47	0	6	1.24	0
Oxidative Damage Resistance	Sporosarcina	130	0	11	1.14	0
Oxidative Damage Resistance	Stenotrophomonas	61	0	9	1.61	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Oxidative Damage Resistance	Streptococcus	47	0	7	1.24	1
Oxidative Damage Resistance	Streptomyces	76	0	6	1.00	1
Oxidative Damage Resistance	Terribacillus	41	0	6	1.08	0
Oxidative Damage Resistance	Ureibacillus	42	0	8	1.11	0
Oxidative Damage Resistance	Caldibacillus	114	0	7	1.00	1
Oxidative Damage Resistance	Cellulosimicrobium	28	0	4	0.74	0
Oxidative Damage Resistance	Cupriavidus	67	0	12	1.76	0.5
Oxidative Damage Resistance	Domibacillus	54	0	9	1.42	0.5
Oxidative Damage Resistance	Heyndrickxia	103	0	10	1.36	0
Oxidative Damage Resistance	Lederbergia	56	0	10	1.47	0.5
Oxidative Damage Resistance	Micrococcus	81	0	5	0.71	0
Oxidative Damage Resistance	Ralstonia	48	0	9	1.26	1
Oxidative Damage Resistance	Weizmannia	42	0	8	1.11	0

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Oxidative Damage Resistance	Alkalihalophilus	53	0	10	1.39	1
Oxidative Damage Resistance	Georgenia	32	0	5	0.84	0
Oxidative Damage Resistance	Rothia	61	0	6	0.80	0
Oxidative Damage Resistance	Sutcliffiella	56	0	9	1.47	1
Oxidative Damage Resistance	Hydrogenophaga	48	0	9	1.26	1
Oxidative Damage Resistance	NA	725	0	18	1.73	1
Repair and Recombination	Bacillus	4787	0	18	0.97	1
Repair and Recombination	Solibacillus	115	0	13	0.96	1
Repair and Recombination	Staphylococcus	1785	0	18	0.88	1
Repair and Recombination	Agromyces	114	0	12	0.95	1
Repair and Recombination	Alkalihalobacillus	474	0	16	0.99	1
Repair and Recombination	Bhargavaea	117	0	8	0.98	1
Repair and Recombination	Brevibacillus	812	0	14	1.13	1
Repair and Recombination	Carnobacterium	99	0	11	0.83	1
Repair and Recombination	Curtobacterium	213	0	10	0.89	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Repair and Recombination	Cytobacillus	1141	0	14	1.06	1
Repair and Recombination	Exiguobacterium	112	0	13	0.93	1
Repair and Recombination	Fictibacillus	352	0	19	0.98	1
Repair and Recombination	Kocuria	308	0	9	0.86	0
Repair and Recombination	Mesobacillus	915	0	15	1.09	1
Repair and Recombination	Metabacillus	503	0	15	1.05	1
Repair and Recombination	Microbacterium	412	0	8	0.86	1
Repair and Recombination	Neobacillus	751	0	17	1.04	1
Repair and Recombination	Niallia	488	0	16	1.02	1
Repair and Recombination	Nocardioides	129	0	9	1.08	1
Repair and Recombination	Oceanobacillus	117	0	9	0.98	1
Repair and Recombination	Paenibacillus	2346	0	20	1.09	1
Repair and Recombination	Peribacillus	256	0	13	1.07	1
Repair and Recombination	Planococcus	152	0	16	1.27	1
Repair and Recombination	Priestia	1838	0	17	1.02	1
Repair and Recombination	Psychrobacillus	277	0	20	1.15	1
Repair and Recombination	Rummeliibacillus	116	0	12	0.97	1
Repair and Recombination	Sphingomonas	111	0	7	0.93	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Repair and Recombination	Sphingopyxis	91	0	6	0.76	0
Repair and Recombination	Sporosarcina	379	0	15	1.05	1
Repair and Recombination	Stenotrophomonas	136	0	7	1.13	1
Repair and Recombination	Streptococcus	136	0	14	1.13	1
Repair and Recombination	Streptomyces	249	0	11	1.04	1
Repair and Recombination	Terribacillus	110	0	11	0.92	1
Repair and Recombination	Ureibacillus	119	0	9	0.99	1
Repair and Recombination	Caldibacillus	358	0	17	0.99	1
Repair and Recombination	Cellulosimicrobium	118	0	9	0.98	1
Repair and Recombination	Cupriavidus	147	0	9	1.23	1
Repair and Recombination	Domibacillus	113	0	8	0.94	1
Repair and Recombination	Heyndrickxia	257	0	14	1.07	1
Repair and Recombination	Lederbergia	129	0	14	1.08	1
Repair and Recombination	Micrococcus	255	0	6	0.71	0
Repair and Recombination	Ralstonia	138	0	10	1.15	1
Repair and Recombination	Weizmannia	122	0	11	1.02	1
Repair and Recombination	Alkalihalophilus	123	0	10	1.03	1
Repair and Recombination	Georgenia	119	0	7	0.99	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Repair and Recombination	Rothia	178	0	9	0.74	0
Repair and Recombination	Sutcliffiella	133	0	17	1.11	1
Repair and Recombination	Hydrogenophaga	137	0	16	1.14	1
Repair and Recombination	NA	1858	0	28	1.41	1
Sporulation	Bacillus	4583	0	15	1.32	1
Sporulation	Solibacillus	109	0	7	1.28	1
Sporulation	Staphylococcus	380	0	32	0.26	0
Sporulation	Agromyces	4	0	1	0.05	0
Sporulation	Alkalihalobacillus	455	0	10	1.34	1
Sporulation	Bhargavaea	53	0	4	0.62	0
Sporulation	Brevibacillus	906	0	19	1.78	1
Sporulation	Carnobacterium	0	0	0	0.00	0
Sporulation	Curtobacterium	14	0	5	0.08	0
Sporulation	Cytobacillus	1150	0	16	1.50	1
Sporulation	Exiguobacterium	10	0	2	0.12	0
Sporulation	Fictibacillus	341	0	8	1.34	1
Sporulation	Kocuria	135	0	9	0.53	0
Sporulation	Mesobacillus	905	0	14	1.52	1
Sporulation	Metabacillus	529	0	12	1.56	1
Sporulation	Microbacterium	13	0	3	0.04	0
Sporulation	Neobacillus	496	0	15	0.97	1
Sporulation	Niallia	379	0	13	1.11	1
Sporulation	Nocardioides	12	0	7	0.14	0
Sporulation	Oceanobacillus	111	0	5	1.31	1
Sporulation	Paenibacillus	1977	0	19	1.29	1
Sporulation	Peribacillus	268	0	11	1.58	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
Sporulation	Planococcus	163	0	16	1.92	1
Sporulation	Priestia	1921	0	15	1.51	1
Sporulation	Psychrobacillus	268	0	12	1.58	1
Sporulation	Rummeliibacillus	147	0	15	1.73	1
Sporulation	Sphingomonas	4	0	4	0.05	0
Sporulation	Sphingopyxis	5	0	4	0.06	0
Sporulation	Sporosarcina	279	0	10	1.09	1
Sporulation	Stenotrophomonas	6	0	4	0.07	0
Sporulation	Streptococcus	138	0	10	1.62	1
Sporulation	Streptomyces	91	0	45	0.54	0
Sporulation	Terribacillus	94	0	4	1.11	1
Sporulation	Ureibacillus	121	0	9	1.42	1
Sporulation	Caldibacillus	301	0	10	1.18	1
Sporulation	Cellulosimicrobium	14	0	12	0.16	0
Sporulation	Cupriavidus	6	0	3	0.07	0
Sporulation	Domibacillus	97	0	13	1.14	1
Sporulation	Heyndrickxia	227	0	6	1.34	1
Sporulation	Lederbergia	114	0	8	1.34	1
Sporulation	Micrococcus	6	0	1	0.02	0
Sporulation	Ralstonia	5	0	3	0.06	0
Sporulation	Weizmannia	100	0	4	1.18	1
Sporulation	Alkalihalophilus	131	0	8	1.54	1
Sporulation	Georgenia	11	0	5	0.13	0
Sporulation	Rothia	2	0	1	0.01	0
Sporulation	Sutcliffiella	138	0	9	1.62	1
Sporulation	Hydrogenophaga	153	0	14	1.80	1
Sporulation	NA	1815	0	36	1.94	1

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
WTF Processes	Bacillus	781	0	7	1.00	1
WTF Processes	Solibacillus	17	0	4	0.89	0
WTF Processes	Staphylococcus	190	0	6	0.59	0
WTF Processes	Agromyces	2	0	2	0.11	0
WTF Processes	Alkalihalobacillus	79	0	6	1.04	0
WTF Processes	Bhargavaea	16	0	3	0.84	1
WTF Processes	Brevibacillus	90	0	6	0.79	0
WTF Processes	Carnobacterium	8	0	3	0.42	0
WTF Processes	Curtobacterium	8	0	3	0.21	0
WTF Processes	Cytobacillus	176	0	8	1.03	1
WTF Processes	Exiguobacterium	9	0	3	0.47	0
WTF Processes	Fictibacillus	47	0	5	0.82	0
WTF Processes	Kocuria	26	0	5	0.46	0
WTF Processes	Mesobacillus	168	0	13	1.26	1
WTF Processes	Metabacillus	104	0	12	1.37	0.5
WTF Processes	Microbacterium	8	0	1	0.11	0
WTF Processes	Neobacillus	105	0	8	0.92	0
WTF Processes	Niallia	60	0	9	0.79	0

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
WTF Processes	Nocardioides	5	0	2	0.26	0
WTF Processes	Oceanobacillus	17	0	4	0.89	1
WTF Processes	Paenibacillus	314	0	9	0.92	0
WTF Processes	Peribacillus	48	0	8	1.26	1
WTF Processes	Planococcus	25	0	6	1.32	1
WTF Processes	Priestia	372	0	12	1.31	1
WTF Processes	Psychrobacillus	50	0	8	1.32	1
WTF Processes	Rummeliibacillus	21	0	6	1.11	1
WTF Processes	Sphingomonas	12	0	5	0.63	0
WTF Processes	Sphingopyxis	12	0	5	0.63	0
WTF Processes	Sporosarcina	48	0	5	0.84	0
WTF Processes	Stenotrophomonas	10	0	4	0.53	0
WTF Processes	Streptococcus	22	0	6	1.16	1
WTF Processes	Streptomyces	11	0	5	0.29	0
WTF Processes	Terribacillus	20	0	5	1.05	1
WTF Processes	Ureibacillus	18	0	5	0.95	0
WTF Processes	Caldibacillus	51	0	6	0.89	0
WTF Processes	Cellulosimicrobium	3	0	2	0.16	0

Table S3.4: Summary Statistics of HMM Hits for Resistance Gene Categories by Isolate Genus

CATEGORY	genus	total_hmm_hits_per_genus	min_hmm_hits_per_genus	max_hmm_hits_per_genus	mean_hmm_hits_per_genus	median_hmm_hits_per_genus
WTF Processes	Cupriavidus	15	0	7	0.79	0
WTF Processes	Domibacillus	24	0	5	1.26	1
WTF Processes	Heyndrickxia	54	0	11	1.42	1
WTF Processes	Lederbergia	17	0	5	0.89	0
WTF Processes	Micrococcus	12	0	1	0.21	0
WTF Processes	Ralstonia	8	0	3	0.42	0
WTF Processes	Weizmannia	16	0	4	0.84	1
WTF Processes	Alkalihalophilus	15	0	6	0.79	0
WTF Processes	Georgenia	4	0	2	0.21	0
WTF Processes	Rothia	6	0	1	0.16	0
WTF Processes	Sutcliffiella	19	0	5	1.00	0
WTF Processes	Hydrogenophaga	14	0	4	0.74	1
WTF Processes	NA	295	0	13	1.41	0

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeof Genus_per_category_per_mission	median_hmm_hits
Cold Shock	Bacillus	Mars Pathfinder	24	6	0	4	1.33	4.00	1
Cold Shock	Bacillus	MER	91	20	0	6	1.52	4.55	1
Cold Shock	Bacillus	MSL	33	8	0	4	1.38	4.13	1
Cold Shock	Bacillus	Phoenix	33	7	0	6	1.57	4.71	1
Cold Shock	Solibacillus	Mars Pathfinder	4	1	0	3	1.33	4.00	1
Cold Shock	Staphylococcus	Mars Pathfinder	3	1	0	2	1.00	3.00	1
Cold Shock	Staphylococcus	MER	36	8	0	6	1.50	4.50	1
Cold Shock	Staphylococcus	MSL	14	4	0	4	1.17	3.50	1
Cold Shock	Staphylococcus	Odyssey	11	3	0	3	1.22	3.67	1
Cold Shock	Staphylococcus	Phoenix	5	1	0	4	1.67	5.00	1
Cold Shock	Agromyces	MER	3	1	0	3	1.00	3.00	0
Cold Shock	Alkalihalobacillus	MER	11	2	0	7	1.83	5.50	1
Cold Shock	Alkalihalobacillus	Odyssey	8	2	0	3	1.33	4.00	1
Cold Shock	Bhargavaea	MER	4	1	0	3	1.33	4.00	1
Cold Shock	Brevibacillus	MER	6	1	0	5	2.00	6.00	1
Cold Shock	Brevibacillus	MSL	22	3	0	7	2.44	7.33	1
Cold Shock	Brevibacillus	Viking	11	2	0	5	1.83	5.50	1
Cold Shock	Carnobacterium	MER	11	1	0	10	3.67	11.00	1
Cold Shock	Curtobacterium	MER	3	1	0	3	1.00	3.00	0
Cold Shock	Curtobacterium	Odyssey	3	1	0	3	1.00	3.00	0
Cold Shock	Cytobacillus	MER	19	3	0	7	2.11	6.33	1
Cold Shock	Cytobacillus	MSL	6	1	0	5	2.00	6.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Cold Shock	Cytobacillus	Odyssey	5	1	0	4	1.67	5.00	1
Cold Shock	Cytobacillus	Phoenix	22	4	0	5	1.83	5.50	1
Cold Shock	Exiguobacterium	MER	3	1	0	2	1.00	3.00	1
Cold Shock	Fictibacillus	MER	6	1	0	5	2.00	6.00	1
Cold Shock	Fictibacillus	Phoenix	12	2	0	5	2.00	6.00	1
Cold Shock	Kocuria	MER	3	1	0	2	1.00	3.00	1
Cold Shock	Kocuria	MSL	5	1	0	5	1.67	5.00	0
Cold Shock	Kocuria	Odyssey	9	1	0	9	3.00	9.00	0
Cold Shock	Mesobacillus	MER	33	5	0	6	2.20	6.60	1
Cold Shock	Mesobacillus	MSL	6	1	0	5	2.00	6.00	1
Cold Shock	Mesobacillus	Odyssey	8	1	0	7	2.67	8.00	1
Cold Shock	Metabacillus	MER	20	3	0	7	2.22	6.67	1
Cold Shock	Metabacillus	MSL	6	1	0	5	2.00	6.00	1
Cold Shock	Microbacterium	MER	4	1	0	4	1.33	4.00	0
Cold Shock	Microbacterium	Odyssey	10	3	0	4	1.11	3.33	0
Cold Shock	Neobacillus	MER	29	5	0	10	1.93	5.80	1
Cold Shock	Neobacillus	MSL	7	1	0	6	2.33	7.00	1
Cold Shock	Niallia	MER	12	3	0	5	1.33	4.00	1
Cold Shock	Niallia	MSL	6	1	0	5	2.00	6.00	1
Cold Shock	Nocardioides	MER	2	1	0	2	0.67	2.00	0
Cold Shock	Oceanobacillus	MER	4	1	0	3	1.33	4.00	1
Cold Shock	Paenibacillus	MER	69	11	0	14	1.92	6.27	1
Cold Shock	Paenibacillus	MSL	22	4	0	5	1.83	5.50	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Cold Shock	Paenibacillus	Phoenix	11	2	0	6	1.83	5.50	1
Cold Shock	Peribacillus	MER	6	1	0	5	2.00	6.00	1
Cold Shock	Peribacillus	MSL	7	1	0	6	2.33	7.00	1
Cold Shock	Planococcus	MER	8	1	0	7	2.67	8.00	1
Cold Shock	Priestia	MER	112	8	0	20	4.67	14.00	1
Cold Shock	Priestia	MSL	45	3	0	15	5.00	15.00	1
Cold Shock	Priestia	Odyssey	30	2	0	15	5.00	15.00	1
Cold Shock	Priestia	Phoenix	24	2	0	17	4.00	12.00	1
Cold Shock	Psychrobacillus	MER	6	1	0	5	2.00	6.00	1
Cold Shock	Psychrobacillus	MSL	7	1	0	6	2.33	7.00	1
Cold Shock	Rummeliibacillus	MER	6	1	0	5	2.00	6.00	1
Cold Shock	Sphingomonas	MER	4	1	0	3	1.33	4.00	1
Cold Shock	Sphingopyxis	MER	3	1	0	2	1.00	3.00	1
Cold Shock	Sporosarcina	MER	14	3	0	4	1.56	4.67	1
Cold Shock	Stenotrophomonas	MER	4	1	0	3	1.33	4.00	1
Cold Shock	Streptococcus	MER	7	1	0	6	2.33	7.00	1
Cold Shock	Streptomyces	MER	7	2	0	4	1.17	3.50	0
Cold Shock	Terribacillus	MER	4	1	0	3	1.33	4.00	1
Cold Shock	Ureibacillus	MER	3	1	0	2	1.00	3.00	1
Cold Shock	Caldibacillus	MSL	8	2	0	3	1.33	4.00	1
Cold Shock	Caldibacillus	Viking	4	1	0	3	1.33	4.00	1
Cold Shock	Cellulosimicrobium	MSL	7	1	0	6	2.33	7.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeof Genus_per_category_per_mission	median_hmm_hits
Cold Shock	Cupriavidus	MSL	6	1	0	5	2.00	6.00	1
Cold Shock	Domibacillus	MSL	7	1	0	6	2.33	7.00	1
Cold Shock	Heyndrickxia	MSL	6	1	0	5	2.00	6.00	1
Cold Shock	Heyndrickxia	Phoenix	6	1	0	5	2.00	6.00	1
Cold Shock	Lederbergia	MSL	4	1	0	3	1.33	4.00	1
Cold Shock	Micrococcus	MSL	6	3	0	2	0.67	2.00	0
Cold Shock	Ralstonia	MSL	5	1	0	4	1.67	5.00	1
Cold Shock	Weizmannia	MSL	5	1	0	4	1.67	5.00	1
Cold Shock	Alkalihalophilus	Odyssey	6	1	0	5	2.00	6.00	1
Cold Shock	Georgenia	Odyssey	3	1	0	3	1.00	3.00	0
Cold Shock	Rothia	Odyssey	0	1	0	0	0.00	0.00	0
Cold Shock	Rothia	Phoenix	0	1	0	0	0.00	0.00	0
Cold Shock	Sutcliffiella	Odyssey	6	1	0	5	2.00	6.00	1
Cold Shock	Hydrogenophaga	Phoenix	8	1	0	7	2.67	8.00	1
Cold Shock	NA	Mars Pathfinder	6	1	0	5	2.00	6.00	1
Cold Shock	NA	MER	20	2	0	14	3.33	10.00	1.5
Cold Shock	NA	MSL	38	3	0	16	4.22	12.67	2
Cold Shock	NA	Phoenix	33	5	0	11	2.20	6.60	1
Oxidative Damage Resistance	Bacillus	Mars Pathfinder	308	6	0	10	1.35	51.33	1
Oxidative Damage Resistance	Bacillus	MER	953	20	0	11	1.25	47.65	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Bacillus	MSL	359	8	0	9	1.18	44.88	1
Oxidative Damage Resistance	Bacillus	Phoenix	353	7	0	12	1.33	50.43	1
Oxidative Damage Resistance	Solibacillus	Mars Pathfinder	56	1	0	11	1.47	56.00	1
Oxidative Damage Resistance	Staphylococcus	Mars Pathfinder	26	1	0	5	0.68	26.00	0
Oxidative Damage Resistance	Staphylococcus	MER	261	8	0	12	0.86	32.63	0
Oxidative Damage Resistance	Staphylococcus	MSL	102	4	0	5	0.67	25.50	0
Oxidative Damage Resistance	Staphylococcus	Odyssey	83	3	0	6	0.73	27.67	0
Oxidative Damage Resistance	Staphylococcus	Phoenix	31	1	0	5	0.82	31.00	0
Oxidative Damage Resistance	Agromyces	MER	32	1	0	7	0.84	32.00	0
Oxidative Damage Resistance	Alkalihalobacillus	MER	88	2	0	8	1.16	44.00	0
Oxidative Damage Resistance	Alkalihalobacillus	Odyssey	108	2	0	11	1.42	54.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Bhargavaea	MER	51	1	0	12	1.34	51.00	1
Oxidative Damage Resistance	Brevibacillus	MER	47	1	0	9	1.24	47.00	1
Oxidative Damage Resistance	Brevibacillus	MSL	136	3	0	10	1.19	45.33	1
Oxidative Damage Resistance	Brevibacillus	Viking	83	2	0	9	1.09	41.50	0
Oxidative Damage Resistance	Carnobacterium	MER	28	1	0	4	0.74	28.00	0
Oxidative Damage Resistance	Curtobacterium	MER	27	1	0	5	0.71	27.00	0
Oxidative Damage Resistance	Curtobacterium	Odyssey	30	1	0	5	0.79	30.00	0
Oxidative Damage Resistance	Cytobacillus	MER	162	3	0	10	1.42	54.00	1
Oxidative Damage Resistance	Cytobacillus	MSL	64	1	0	11	1.68	64.00	1
Oxidative Damage Resistance	Cytobacillus	Odyssey	52	1	0	9	1.37	52.00	1
Oxidative Damage Resistance	Cytobacillus	Phoenix	177	4	0	8	1.16	44.25	0.5

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Exiguobacterium	MER	39	1	0	9	1.03	39.00	0
Oxidative Damage Resistance	Fictibacillus	MER	41	1	0	9	1.08	41.00	0.5
Oxidative Damage Resistance	Fictibacillus	Phoenix	86	2	0	9	1.13	43.00	0
Oxidative Damage Resistance	Kocuria	MER	43	1	0	8	1.13	43.00	0
Oxidative Damage Resistance	Kocuria	MSL	28	1	0	5	0.74	28.00	0
Oxidative Damage Resistance	Kocuria	Odyssey	23	1	0	3	0.61	23.00	0
Oxidative Damage Resistance	Mesobacillus	MER	258	5	0	12	1.36	51.60	1
Oxidative Damage Resistance	Mesobacillus	MSL	56	1	0	10	1.47	56.00	0
Oxidative Damage Resistance	Mesobacillus	Odyssey	57	1	0	10	1.50	57.00	1
Oxidative Damage Resistance	Metabacillus	MER	188	3	0	13	1.65	62.67	1
Oxidative Damage Resistance	Metabacillus	MSL	48	1	0	10	1.26	48.00	0.5

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Microbacterium	MER	20	1	0	3	0.53	20.00	0
Oxidative Damage Resistance	Microbacterium	Odyssey	80	3	0	5	0.70	26.67	0
Oxidative Damage Resistance	Neobacillus	MER	228	5	0	10	1.20	45.60	0
Oxidative Damage Resistance	Neobacillus	MSL	50	1	0	11	1.32	50.00	1
Oxidative Damage Resistance	Niallia	MER	118	3	0	9	1.04	39.33	0
Oxidative Damage Resistance	Niallia	MSL	50	1	0	8	1.32	50.00	0.5
Oxidative Damage Resistance	Nocardioides	MER	25	1	0	4	0.66	25.00	0
Oxidative Damage Resistance	Oceanobacillus	MER	48	1	0	9	1.26	48.00	1
Oxidative Damage Resistance	Paenibacillus	MER	583	11	0	12	1.28	53.00	0
Oxidative Damage Resistance	Paenibacillus	MSL	187	4	0	10	1.23	46.75	1
Oxidative Damage Resistance	Paenibacillus	Phoenix	99	2	0	10	1.30	49.50	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Peribacillus	MER	53	1	0	10	1.39	53.00	0
Oxidative Damage Resistance	Peribacillus	MSL	62	1	0	8	1.63	62.00	1
Oxidative Damage Resistance	Planococcus	MER	47	1	0	8	1.24	47.00	1
Oxidative Damage Resistance	Priestia	MER	411	8	0	9	1.35	51.38	1
Oxidative Damage Resistance	Priestia	MSL	155	3	0	8	1.36	51.67	1
Oxidative Damage Resistance	Priestia	Odyssey	107	2	0	9	1.41	53.50	1
Oxidative Damage Resistance	Priestia	Phoenix	91	2	0	8	1.20	45.50	1
Oxidative Damage Resistance	Psychrobacillus	MER	49	1	0	8	1.29	49.00	1
Oxidative Damage Resistance	Psychrobacillus	MSL	50	1	0	9	1.32	50.00	0
Oxidative Damage Resistance	Rummeliibacillus	MER	60	1	0	10	1.58	60.00	1
Oxidative Damage Resistance	Sphingomonas	MER	59	1	0	7	1.55	59.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Sphingopyxis	MER	47	1	0	6	1.24	47.00	0
Oxidative Damage Resistance	Sporosarcina	MER	130	3	0	11	1.14	43.33	0
Oxidative Damage Resistance	Stenotrophomonas	MER	61	1	0	9	1.61	61.00	1
Oxidative Damage Resistance	Streptococcus	MER	47	1	0	7	1.24	47.00	1
Oxidative Damage Resistance	Streptomyces	MER	76	2	0	6	1.00	38.00	1
Oxidative Damage Resistance	Terribacillus	MER	41	1	0	6	1.08	41.00	0
Oxidative Damage Resistance	Ureibacillus	MER	42	1	0	8	1.11	42.00	0
Oxidative Damage Resistance	Caldibacillus	MSL	76	2	0	6	1.00	38.00	1
Oxidative Damage Resistance	Caldibacillus	Viking	38	1	0	7	1.00	38.00	1
Oxidative Damage Resistance	Cellulosimicrobium	MSL	28	1	0	4	0.74	28.00	0
Oxidative Damage Resistance	Cupriavidus	MSL	67	1	0	12	1.76	67.00	0.5

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Domibacillus	MSL	54	1	0	9	1.42	54.00	0.5
Oxidative Damage Resistance	Heyndrickxia	MSL	49	1	0	8	1.29	49.00	0
Oxidative Damage Resistance	Heyndrickxia	Phoenix	54	1	0	10	1.42	54.00	0
Oxidative Damage Resistance	Lederbergia	MSL	56	1	0	10	1.47	56.00	0.5
Oxidative Damage Resistance	Micrococcus	MSL	81	3	0	5	0.71	27.00	0
Oxidative Damage Resistance	Ralstonia	MSL	48	1	0	9	1.26	48.00	1
Oxidative Damage Resistance	Weizmannia	MSL	42	1	0	8	1.11	42.00	0
Oxidative Damage Resistance	Alkalihalophilus	Odyssey	53	1	0	10	1.39	53.00	1
Oxidative Damage Resistance	Georgenia	Odyssey	32	1	0	5	0.84	32.00	0
Oxidative Damage Resistance	Rothia	Odyssey	32	1	0	6	0.84	32.00	0
Oxidative Damage Resistance	Rothia	Phoenix	29	1	0	5	0.76	29.00	0

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Oxidative Damage Resistance	Sutcliffiella	Odyssey	56	1	0	9	1.47	56.00	1
Oxidative Damage Resistance	Hydrogenophaga	Phoenix	48	1	0	9	1.26	48.00	1
Oxidative Damage Resistance	NA	Mars Pathfinder	43	1	0	10	1.13	43.00	0.5
Oxidative Damage Resistance	NA	MER	130	2	0	13	1.71	65.00	1
Oxidative Damage Resistance	NA	MSL	297	3	0	18	2.61	99.00	2
Oxidative Damage Resistance	NA	Phoenix	255	5	0	9	1.34	51.00	1
Repair and Recombination	Bacillus	Mars Pathfinder	705	6	0	14	0.98	117.50	1
Repair and Recombination	Bacillus	MER	2358	20	0	18	0.98	117.90	1
Repair and Recombination	Bacillus	MSL	902	8	0	13	0.94	112.75	1
Repair and Recombination	Bacillus	Phoenix	822	7	0	14	0.98	117.43	1
Repair and Recombination	Solibacillus	Mars Pathfinder	115	1	0	13	0.96	115.00	1
Repair and Recombination	Staphylococcus	Mars Pathfinder	101	1	0	8	0.84	101.00	1
Repair and Recombination	Staphylococcus	MER	882	8	0	18	0.92	110.25	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Repair and Recombination	Staphylococcus	MSL	394	4	0	12	0.82	98.50	1
Repair and Recombination	Staphylococcus	Odyssey	305	3	0	11	0.85	101.67	1
Repair and Recombination	Staphylococcus	Phoenix	103	1	0	9	0.86	103.00	1
Repair and Recombination	Agromyces	MER	114	1	0	12	0.95	114.00	1
Repair and Recombination	Alkalihalobacillus	MER	246	2	0	16	1.03	123.00	1
Repair and Recombination	Alkalihalobacillus	Odyssey	228	2	0	9	0.95	114.00	1
Repair and Recombination	Bhargavaea	MER	117	1	0	8	0.98	117.00	1
Repair and Recombination	Brevibacillus	MER	132	1	0	14	1.10	132.00	1
Repair and Recombination	Brevibacillus	MSL	395	3	0	13	1.10	131.67	1
Repair and Recombination	Brevibacillus	Viking	285	2	0	12	1.19	142.50	1
Repair and Recombination	Carnobacterium	MER	99	1	0	11	0.83	99.00	1
Repair and Recombination	Curtobacterium	MER	103	1	0	10	0.86	103.00	1
Repair and Recombination	Curtobacterium	Odyssey	110	1	0	10	0.92	110.00	1
Repair and Recombination	Cytobacillus	MER	389	3	0	14	1.08	129.67	1
Repair and Recombination	Cytobacillus	MSL	137	1	0	12	1.14	137.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Repair and Recombination	Cytobacillus	Odyssey	120	1	0	13	1.00	120.00	1
Repair and Recombination	Cytobacillus	Phoenix	495	4	0	12	1.03	123.75	1
Repair and Recombination	Exiguobacterium	MER	112	1	0	13	0.93	112.00	1
Repair and Recombination	Fictibacillus	MER	127	1	0	19	1.06	127.00	1
Repair and Recombination	Fictibacillus	Phoenix	225	2	0	10	0.94	112.50	1
Repair and Recombination	Kocuria	MER	119	1	0	9	0.99	119.00	1
Repair and Recombination	Kocuria	MSL	98	1	0	6	0.82	98.00	0
Repair and Recombination	Kocuria	Odyssey	91	1	0	5	0.76	91.00	0
Repair and Recombination	Mesobacillus	MER	635	5	0	15	1.06	127.00	1
Repair and Recombination	Mesobacillus	MSL	141	1	0	12	1.18	141.00	1
Repair and Recombination	Mesobacillus	Odyssey	139	1	0	14	1.16	139.00	1
Repair and Recombination	Metabacillus	MER	381	3	0	14	1.06	127.00	1
Repair and Recombination	Metabacillus	MSL	122	1	0	15	1.02	122.00	1
Repair and Recombination	Microbacterium	MER	106	1	0	8	0.88	106.00	1
Repair and Recombination	Microbacterium	Odyssey	306	3	0	8	0.85	102.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Repair and Recombination	Neobacillus	MER	621	5	0	17	1.04	124.20	1
Repair and Recombination	Neobacillus	MSL	130	1	0	9	1.08	130.00	1
Repair and Recombination	Niallia	MER	358	3	0	16	0.99	119.33	1
Repair and Recombination	Niallia	MSL	130	1	0	16	1.08	130.00	1
Repair and Recombination	Nocardioides	MER	129	1	0	9	1.08	129.00	1
Repair and Recombination	Oceanobacillus	MER	117	1	0	9	0.98	117.00	1
Repair and Recombination	Paenibacillus	MER	1524	11	0	20	1.06	138.55	1
Repair and Recombination	Paenibacillus	MSL	553	4	0	18	1.15	138.25	1
Repair and Recombination	Paenibacillus	Phoenix	269	2	0	20	1.12	134.50	1
Repair and Recombination	Peribacillus	MER	127	1	0	13	1.06	127.00	1
Repair and Recombination	Peribacillus	MSL	129	1	0	12	1.08	129.00	1
Repair and Recombination	Planococcus	MER	152	1	0	16	1.27	152.00	1
Repair and Recombination	Priestia	MER	1001	8	0	17	1.04	125.13	1
Repair and Recombination	Priestia	MSL	356	3	0	10	0.99	118.67	1
Repair and Recombination	Priestia	Odyssey	241	2	0	10	1.00	120.50	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Repair and Recombination	Priestia	Phoenix	240	2	0	13	1.00	120.00	1
Repair and Recombination	Psychrobacillus	MER	143	1	0	20	1.19	143.00	1
Repair and Recombination	Psychrobacillus	MSL	134	1	0	13	1.12	134.00	1
Repair and Recombination	Rummeliibacillus	MER	116	1	0	12	0.97	116.00	1
Repair and Recombination	Sphingomonas	MER	111	1	0	7	0.93	111.00	1
Repair and Recombination	Sphingopyxis	MER	91	1	0	6	0.76	91.00	0
Repair and Recombination	Sporosarcina	MER	379	3	0	15	1.05	126.33	1
Repair and Recombination	Stenotrophomonas	MER	136	1	0	7	1.13	136.00	1
Repair and Recombination	Streptococcus	MER	136	1	0	14	1.13	136.00	1
Repair and Recombination	Streptomyces	MER	249	2	0	11	1.04	124.50	1
Repair and Recombination	Terribacillus	MER	110	1	0	11	0.92	110.00	1
Repair and Recombination	Ureibacillus	MER	119	1	0	9	0.99	119.00	1
Repair and Recombination	Caldibacillus	MSL	241	2	0	17	1.00	120.50	1
Repair and Recombination	Caldibacillus	Viking	117	1	0	14	0.98	117.00	1
Repair and Recombination	Cellulosimicrobium	MSL	118	1	0	9	0.98	118.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Repair and Recombination	Cupriavidus	MSL	147	1	0	9	1.23	147.00	1
Repair and Recombination	Domibacillus	MSL	113	1	0	8	0.94	113.00	1
Repair and Recombination	Heyndrickxia	MSL	128	1	0	14	1.07	128.00	1
Repair and Recombination	Heyndrickxia	Phoenix	129	1	0	14	1.08	129.00	1
Repair and Recombination	Lederbergia	MSL	129	1	0	14	1.08	129.00	1
Repair and Recombination	Micrococcus	MSL	255	3	0	6	0.71	85.00	0
Repair and Recombination	Ralstonia	MSL	138	1	0	10	1.15	138.00	1
Repair and Recombination	Weizmannia	MSL	122	1	0	11	1.02	122.00	1
Repair and Recombination	Alkalihalophilus	Odyssey	123	1	0	10	1.03	123.00	1
Repair and Recombination	Georgenia	Odyssey	119	1	0	7	0.99	119.00	1
Repair and Recombination	Rothia	Odyssey	89	1	0	8	0.74	89.00	0
Repair and Recombination	Rothia	Phoenix	89	1	0	9	0.74	89.00	0
Repair and Recombination	Sutcliffiella	Odyssey	133	1	0	17	1.11	133.00	1
Repair and Recombination	Hydrogenophaga	Phoenix	137	1	0	16	1.14	137.00	1
Repair and Recombination	NA	Mars Pathfinder	127	1	0	20	1.06	127.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Repair and Recombination	NA	MER	384	2	0	28	1.60	192.00	1
Repair and Recombination	NA	MSL	700	3	0	24	1.94	233.33	1
Repair and Recombination	NA	Phoenix	647	5	0	19	1.08	129.40	1
Sporulation	Bacillus	Mars Pathfinder	666	6	0	6	1.31	111.00	1
Sporulation	Bacillus	MER	2263	20	0	15	1.33	113.15	1
Sporulation	Bacillus	MSL	880	8	0	6	1.29	110.00	1
Sporulation	Bacillus	Phoenix	774	7	0	8	1.30	110.57	1
Sporulation	Solibacillus	Mars Pathfinder	109	1	0	7	1.28	109.00	1
Sporulation	Staphylococcus	Mars Pathfinder	3	1	0	1	0.04	3.00	0
Sporulation	Staphylococcus	MER	353	8	0	32	0.52	44.13	0
Sporulation	Staphylococcus	MSL	14	4	0	2	0.04	3.50	0
Sporulation	Staphylococcus	Odyssey	7	3	0	1	0.03	2.33	0
Sporulation	Staphylococcus	Phoenix	3	1	0	2	0.04	3.00	0
Sporulation	Agromyces	MER	4	1	0	1	0.05	4.00	0
Sporulation	Alkalihalobacillus	MER	188	2	0	9	1.11	94.00	1
Sporulation	Alkalihalobacillus	Odyssey	267	2	0	10	1.57	133.50	1
Sporulation	Bhargavaea	MER	53	1	0	4	0.62	53.00	0
Sporulation	Brevibacillus	MER	138	1	0	9	1.62	138.00	1
Sporulation	Brevibacillus	MSL	424	3	0	16	1.66	141.33	1
Sporulation	Brevibacillus	Viking	344	2	0	19	2.02	172.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Sporulation	Carnobacterium	MER	0	1	0	0	0.00	0.00	0
Sporulation	Curtobacterium	MER	7	1	0	5	0.08	7.00	0
Sporulation	Curtobacterium	Odyssey	7	1	0	5	0.08	7.00	0
Sporulation	Cytobacillus	MER	418	3	0	12	1.64	139.33	1
Sporulation	Cytobacillus	MSL	160	1	0	12	1.88	160.00	1
Sporulation	Cytobacillus	Odyssey	105	1	0	5	1.24	105.00	1
Sporulation	Cytobacillus	Phoenix	467	4	0	16	1.37	116.75	1
Sporulation	Exiguobacterium	MER	10	1	0	2	0.12	10.00	0
Sporulation	Fictibacillus	MER	113	1	0	8	1.33	113.00	1
Sporulation	Fictibacillus	Phoenix	228	2	0	8	1.34	114.00	1
Sporulation	Kocuria	MER	122	1	0	9	1.44	122.00	1
Sporulation	Kocuria	MSL	5	1	0	2	0.06	5.00	0
Sporulation	Kocuria	Odyssey	8	1	0	5	0.09	8.00	0
Sporulation	Mesobacillus	MER	634	5	0	14	1.49	126.80	1
Sporulation	Mesobacillus	MSL	140	1	0	11	1.65	140.00	1
Sporulation	Mesobacillus	Odyssey	131	1	0	8	1.54	131.00	1
Sporulation	Metabacillus	MER	471	3	0	12	1.85	157.00	1
Sporulation	Metabacillus	MSL	58	1	0	5	0.68	58.00	0
Sporulation	Microbacterium	MER	3	1	0	1	0.04	3.00	0
Sporulation	Microbacterium	Odyssey	10	3	0	3	0.04	3.33	0
Sporulation	Neobacillus	MER	435	5	0	15	1.02	87.00	1
Sporulation	Neobacillus	MSL	61	1	0	4	0.72	61.00	0
Sporulation	Niallia	MER	276	3	0	13	1.08	92.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Sporulation	Niallia	MSL	103	1	0	6	1.21	103.00	1
Sporulation	Nocardioides	MER	12	1	0	7	0.14	12.00	0
Sporulation	Oceanobacillus	MER	111	1	0	5	1.31	111.00	1
Sporulation	Paenibacillus	MER	1214	11	0	15	1.19	110.36	1
Sporulation	Paenibacillus	MSL	531	4	0	19	1.56	132.75	1
Sporulation	Paenibacillus	Phoenix	232	2	0	8	1.36	116.00	1
Sporulation	Peribacillus	MER	126	1	0	9	1.48	126.00	1
Sporulation	Peribacillus	MSL	142	1	0	11	1.67	142.00	1
Sporulation	Planococcus	MER	163	1	0	16	1.92	163.00	1
Sporulation	Priestia	MER	1034	8	0	15	1.52	129.25	1
Sporulation	Priestia	MSL	399	3	0	11	1.56	133.00	1
Sporulation	Priestia	Odyssey	255	2	0	11	1.50	127.50	1
Sporulation	Priestia	Phoenix	233	2	0	7	1.37	116.50	1
Sporulation	Psychrobacillus	MER	120	1	0	6	1.41	120.00	1
Sporulation	Psychrobacillus	MSL	148	1	0	12	1.74	148.00	1
Sporulation	Rummeliibacillus	MER	147	1	0	15	1.73	147.00	1
Sporulation	Sphingomonas	MER	4	1	0	4	0.05	4.00	0
Sporulation	Sphingopyxis	MER	5	1	0	4	0.06	5.00	0
Sporulation	Sporosarcina	MER	279	3	0	10	1.09	93.00	1
Sporulation	Stenotrophomonas	MER	6	1	0	4	0.07	6.00	0
Sporulation	Streptococcus	MER	138	1	0	10	1.62	138.00	1
Sporulation	Streptomyces	MER	91	2	0	45	0.54	45.50	0
Sporulation	Terribacillus	MER	94	1	0	4	1.11	94.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
Sporulation	Ureibacillus	MER	121	1	0	9	1.42	121.00	1
Sporulation	Caldibacillus	MSL	187	2	0	4	1.10	93.50	1
Sporulation	Caldibacillus	Viking	114	1	0	10	1.34	114.00	1
Sporulation	Cellulosimicrobium	MSL	14	1	0	12	0.16	14.00	0
Sporulation	Cupriavidus	MSL	6	1	0	3	0.07	6.00	0
Sporulation	Domibacillus	MSL	97	1	0	13	1.14	97.00	1
Sporulation	Heyndrickxia	MSL	116	1	0	6	1.36	116.00	1
Sporulation	Heyndrickxia	Phoenix	111	1	0	5	1.31	111.00	1
Sporulation	Lederbergia	MSL	114	1	0	8	1.34	114.00	1
Sporulation	Micrococcus	MSL	6	3	0	1	0.02	2.00	0
Sporulation	Ralstonia	MSL	5	1	0	3	0.06	5.00	0
Sporulation	Weizmannia	MSL	100	1	0	4	1.18	100.00	1
Sporulation	Alkalihalophilus	Odyssey	131	1	0	8	1.54	131.00	1
Sporulation	Georgenia	Odyssey	11	1	0	5	0.13	11.00	0
Sporulation	Rothia	Odyssey	1	1	0	1	0.01	1.00	0
Sporulation	Rothia	Phoenix	1	1	0	1	0.01	1.00	0
Sporulation	Sutcliffiella	Odyssey	138	1	0	9	1.62	138.00	1
Sporulation	Hydrogenophaga	Phoenix	153	1	0	14	1.80	153.00	1
Sporulation	NA	Mars Pathfinder	110	1	0	8	1.29	110.00	1
Sporulation	NA	MER	431	2	0	36	2.54	215.50	2
Sporulation	NA	MSL	641	3	0	22	2.51	213.67	2
Sporulation	NA	Phoenix	633	5	0	10	1.49	126.60	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
WTF Processes	Bacillus	Mars Pathfinder	113	6	0	5	0.99	18.83	1
WTF Processes	Bacillus	MER	372	20	0	7	0.98	18.60	0
WTF Processes	Bacillus	MSL	157	8	0	6	1.03	19.63	1
WTF Processes	Bacillus	Phoenix	139	7	0	6	1.05	19.86	1
WTF Processes	Solibacillus	Mars Pathfinder	17	1	0	4	0.89	17.00	0
WTF Processes	Staphylococcus	Mars Pathfinder	10	1	0	3	0.53	10.00	0
WTF Processes	Staphylococcus	MER	100	8	0	6	0.66	12.50	0
WTF Processes	Staphylococcus	MSL	40	4	0	3	0.53	10.00	0
WTF Processes	Staphylococcus	Odyssey	30	3	0	3	0.53	10.00	0
WTF Processes	Staphylococcus	Phoenix	10	1	0	3	0.53	10.00	0
WTF Processes	Agromyces	MER	2	1	0	2	0.11	2.00	0
WTF Processes	Alkalihalobacillus	MER	31	2	0	5	0.82	15.50	0
WTF Processes	Alkalihalobacillus	Odyssey	48	2	0	6	1.26	24.00	0
WTF Processes	Bhargavaea	MER	16	1	0	3	0.84	16.00	1
WTF Processes	Brevibacillus	MER	20	1	0	6	1.05	20.00	0

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
WTF Processes	Brevibacillus	MSL	42	3	0	4	0.74	14.00	0
WTF Processes	Brevibacillus	Viking	28	2	0	4	0.74	14.00	0
WTF Processes	Carnobacterium	MER	8	1	0	3	0.42	8.00	0
WTF Processes	Curtobacterium	MER	4	1	0	3	0.21	4.00	0
WTF Processes	Curtobacterium	Odyssey	4	1	0	3	0.21	4.00	0
WTF Processes	Cytobacillus	MER	67	3	0	8	1.18	22.33	1
WTF Processes	Cytobacillus	MSL	22	1	0	6	1.16	22.00	1
WTF Processes	Cytobacillus	Odyssey	17	1	0	4	0.89	17.00	1
WTF Processes	Cytobacillus	Phoenix	70	4	0	6	0.92	17.50	0.5
WTF Processes	Exiguobacterium	MER	9	1	0	3	0.47	9.00	0
WTF Processes	Fictibacillus	MER	15	1	0	4	0.79	15.00	0
WTF Processes	Fictibacillus	Phoenix	32	2	0	5	0.84	16.00	0
WTF Processes	Kocuria	MER	19	1	0	5	1.00	19.00	0
WTF Processes	Kocuria	MSL	3	1	0	1	0.16	3.00	0
WTF Processes	Kocuria	Odyssey	4	1	0	1	0.21	4.00	0

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
WTF Processes	Mesobacillus	MER	118	5	0	13	1.24	23.60	1
WTF Processes	Mesobacillus	MSL	27	1	0	7	1.42	27.00	1
WTF Processes	Mesobacillus	Odyssey	23	1	0	7	1.21	23.00	1
WTF Processes	Metabacillus	MER	89	3	0	12	1.56	29.67	1
WTF Processes	Metabacillus	MSL	15	1	0	4	0.79	15.00	0
WTF Processes	Microbacterium	MER	2	1	0	1	0.11	2.00	0
WTF Processes	Microbacterium	Odyssey	6	3	0	1	0.11	2.00	0
WTF Processes	Neobacillus	MER	86	5	0	8	0.91	17.20	0
WTF Processes	Neobacillus	MSL	19	1	0	5	1.00	19.00	0
WTF Processes	Niallia	MER	39	3	0	9	0.68	13.00	0
WTF Processes	Niallia	MSL	21	1	0	7	1.11	21.00	1
WTF Processes	Nocardioides	MER	5	1	0	2	0.26	5.00	0
WTF Processes	Oceanobacillus	MER	17	1	0	4	0.89	17.00	1
WTF Processes	Paenibacillus	MER	202	11	0	9	0.89	18.36	0
WTF Processes	Paenibacillus	MSL	71	4	0	5	0.93	17.75	0

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
WTF Processes	Paenibacillus	Phoenix	41	2	0	6	1.08	20.50	0
WTF Processes	Peribacillus	MER	24	1	0	8	1.26	24.00	1
WTF Processes	Peribacillus	MSL	24	1	0	6	1.26	24.00	1
WTF Processes	Planococcus	MER	25	1	0	6	1.32	25.00	1
WTF Processes	Priestia	MER	202	8	0	12	1.33	25.25	1
WTF Processes	Priestia	MSL	72	3	0	9	1.26	24.00	1
WTF Processes	Priestia	Odyssey	50	2	0	10	1.32	25.00	1
WTF Processes	Priestia	Phoenix	48	2	0	10	1.26	24.00	1
WTF Processes	Psychrobacillus	MER	24	1	0	8	1.26	24.00	1
WTF Processes	Psychrobacillus	MSL	26	1	0	7	1.37	26.00	1
WTF Processes	Rummeliibacillus	MER	21	1	0	6	1.11	21.00	1
WTF Processes	Sphingomonas	MER	12	1	0	5	0.63	12.00	0
WTF Processes	Sphingopyxis	MER	12	1	0	5	0.63	12.00	0
WTF Processes	Sporosarcina	MER	48	3	0	5	0.84	16.00	0
WTF Processes	Stenotrophomonas	MER	10	1	0	4	0.53	10.00	0

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
WTF Processes	Streptococcus	MER	22	1	0	6	1.16	22.00	1
WTF Processes	Streptomyces	MER	11	2	0	5	0.29	5.50	0
WTF Processes	Terribacillus	MER	20	1	0	5	1.05	20.00	1
WTF Processes	Ureibacillus	MER	18	1	0	5	0.95	18.00	0
WTF Processes	Caldibacillus	MSL	32	2	0	5	0.84	16.00	0
WTF Processes	Caldibacillus	Viking	19	1	0	6	1.00	19.00	0
WTF Processes	Cellulosimicrobium	MSL	3	1	0	2	0.16	3.00	0
WTF Processes	Cupriavidus	MSL	15	1	0	7	0.79	15.00	0
WTF Processes	Domibacillus	MSL	24	1	0	5	1.26	24.00	1
WTF Processes	Heyndrickxia	MSL	27	1	0	11	1.42	27.00	1
WTF Processes	Heyndrickxia	Phoenix	27	1	0	11	1.42	27.00	1
WTF Processes	Lederbergia	MSL	17	1	0	5	0.89	17.00	0
WTF Processes	Micrococcus	MSL	12	3	0	1	0.21	4.00	0
WTF Processes	Ralstonia	MSL	8	1	0	3	0.42	8.00	0
WTF Processes	Weizmannia	MSL	16	1	0	4	0.84	16.00	1

Table S3.5: Summary Statistics of HMM Hit Frequencies for Resistance Gene Categories by Isolate Genus From Each Mission

CATEGORY	genus	mission	total_hmm_hits	genus_per_mission	min_hmm_hits	max_hmm_hits	mean_hmm_hits_per_HMM	mean_hmm_hits_per_genomeofGenus_per_category_per_mission	median_hmm_hits
WTF Processes	Alkalihalophilus	Odyssey	15	1	0	6	0.79	15.00	0
WTF Processes	Georgenia	Odyssey	4	1	0	2	0.21	4.00	0
WTF Processes	Rothia	Odyssey	3	1	0	1	0.16	3.00	0
WTF Processes	Rothia	Phoenix	3	1	0	1	0.16	3.00	0
WTF Processes	Sutcliffiella	Odyssey	19	1	0	5	1.00	19.00	0
WTF Processes	Hydrogenophaga	Phoenix	14	1	0	4	0.74	14.00	1
WTF Processes	NA	Mars Pathfinder	14	1	0	5	0.74	14.00	0
WTF Processes	NA	MER	61	2	0	10	1.61	30.50	1
WTF Processes	NA	MSL	119	3	0	13	2.09	39.67	1
WTF Processes	NA	Phoenix	101	5	0	7	1.06	20.20	0