A peer-reviewed version of this preprint was published in PeerJ on 27 October 2017.

View the peer-reviewed version (peerj.com/articles/3944), which is the preferred citable publication unless you specifically need to cite this preprint.

Winter AS, Hathaway JJM, Kimble JC, Buecher DC, Valdez EW, Porras-Alfaro A, Young JM, Read KJH, Northup DE. 2017. Skin and fur bacterial diversity and community structure on American southwestern bats: effects of habitat, geography and bat traits. PeerJ 5:e3944 https://doi.org/10.7717/peerj.3944

External bacterial diversity on bats in the southwestern United States: Changes in bacterial community structure above and below ground

Ara S. Winter[1], Jason C. Kimble[1], Jesse M. Young[1], Debbie C. Buecher[2], Ernest W. Valdez[3], Jennifer J. M. Hathaway[1], Andrea Porras-Alfaro[4], Kaitlyn J. H. Read, and Diana E. Northup[1]*

[1]Department of Biology, University of New Mexico, Albuquerque, New Mexico, United States of America

[2]Buecher Biological Consulting, Tucson, Arizona, United States of America

[3]United States Geological Survey, Fort Collins Science Center, Colorado; Biology Department, MSC03 2020, University of New Mexico, Albuquerque, New Mexico, United States of America

[4]Department of Biological Sciences, Western Illinois University, Macomb, Illinois, United States of America

*Corresponding author
E-mail: dnorthup@unm.edu

Abstract

Microorganisms that reside on and in mammals, such as bats, have the potential to influence their host's health and to provide potential defenses against invading pathogens. However, we have little to no understanding of the external bacterial microbiome on bats, or factors that influence the structure of these communities. The southwestern United States offers excellent sites for the study of external bat bacterial microbiomes due to the diversity of bat species, the variety of abiotic and biotic factors that may govern bat bacterial microbiome communities, and the lack of white-nose syndrome (a newly emergent fungal disease of bats) presence in the Southwest. We studied the extent to which changes in distributions of bacteria on external bat surfaces are a function of geographic location and ecoregion, and whether the sampled bats were caught in caves or surface-netted. To test these variables we used 16S rRNA gene 454 pyrosequencing from swabs of external skin and fur surfaces from 186 bats from 14 species sampled across southeastern New Mexico to northwestern Arizona. Community similarity patterns and random forest models, and generalized linear mixed-effects models show that factors such as location (e.g. cave-caught vs. surface-netted) and ecoregion are major contributors to the structure of bacterial communities on bats. Bats caught in caves had a distinct microbial community compared to those that were netted on the surface. Our results provide a first insight into the distribution of external bat bacteria in a WNS-free environment and provide a baseline of bat external microbiomes that can be explored for potential natural defenses against pathogens.

Keywords: bat microbiome, biogeography, microbiome, Chiroptera, white-nose syndrome

Introduction

Recent studies of microbiota associated with humans and some other species have shown that external and internal microbiota play critical roles in maintaining the health and well-being of these organisms (Apprill et al, 2014, Human Microbiome Project Consortium, 2012). In contrast to humans, the nature of the microbiomes associated with bats, in particular with their external surfaces, is poorly studied. Furthermore, we know very little about what role bacteria play in defense against invading pathogenic microorganisms in bats, a diverse group of mammals that plays key roles in our agriculture and natural ecosystems.

Second to rodents at 2,277 species, bats are the most numerous mammal in the world (Wilson and Reeder, 2005). They are represented by approximately 1,116 different species that occupy habitats ranging from the wet tropics of the equator to dry lowland deserts in temperate latitudes (Wilson and Reeder, 2005). There are approximately 45 species of bats that occur throughout the continental United States and in the Southwest there are approximately 28 different species belonging to Vespertilionidae, Molossidae, Phyllostomidae, and Mormoopidae. Many of these bat species are sympatric and syntopic, especially in New Mexico and Arizona (Findley et al. 1975, Humphrey, 1975; Hall, 1981; Hoffmeister, 1986; Frey, 2004; Harvey et al., 2011) (Figure 1).

The high diversity of bats in the Southwest is attributed to the presence of some species occurring at the northern limits of their range from Mexico (Findley et al., 1975; Hoffmeister, 1986; Frey, 2004). Species diversity is also attributed to the diverse topography of the Southwest (e.g., Colorado Plateau and Sky Islands) that contributes to suitable habitat for roosts that range from rock outcrops, crevices, caves, and lava caves to tree cavities and under exfoliating bark during different periods of the year (Bogan et al. 2003). Within the Southwest, the ecology, life histories, and morphology of each bat species during different times of the year are also diverse. For example, within Vespertilionidae, there are 21 species of bats that represent different body, wing, and ear sizes and shapes, that can affect flight speed and maneuverability, thus resulting in specialized feeding strategies within each species group (Findley et al., 1975; Frey, 2004; Hoffmeister, 1986; Harvey et al., 2011). Vespertilionidae includes many species that use hibernation (a unique trait used by many bat species in temperate regions) during winter months when they face greater thermoregulatory demands (cold temperatures) and reduced food resources. Hibernating bats also suppress their immune system, which would require more calories to sustain in deep hibernation. This evolutionary mechanism has served bats over time, but has now become a risk to their survival because of wildlife disease known as white-nose syndrome (WNS).

WNS, which was introduced into the eastern region of the United States 10 years ago (Frick et al., 2010), is caused by a psychrophilic, keratinophilic fungus (*Pseudogymnoascus destructans*) that attacks, during hibernation, the bats' wings and uropatagium (tail membrane), thus degrading the physiological function of a large surface area on the bats as the disease, as well as causing disruption to fat storage and water regulation. Currently, WNS has killed millions of hibernating bats in the East and is spreading westward. Given the high diversity of bat species in the western and southwestern United States, the potential threat to bat diversity at a regional-scale is very high. Arizona and New Mexico have nine species of *Myotis*, some of which are western analogs to eastern species currently impacted by WNS. It is therefore critical that we determine which western species will be negatively impacted by WNS prior to its predicted arrival (Maher et al.,

2012) in order to target our monitoring for WNS. It is possible that certain bacteria present on some bat species can influence the progression and outcome of WNS (Hoyt et al. 2015).

Because *P. destructans* is a novel species for cave ecosystems in North America, it is likely affecting the natural external microbiome of bats and caves. Current microbiome studies on bats focus on the gut or fecal microbiome (Carrillo-Araujo et al., 2015; Borda et al., 2014), and knowledge on a regional-scale of the external bat microbiome in a WNS-free area is lacking. The influence of local factors including abiotic and biotic variables in geographic patterns of the bat external microbiome at the local and regional-scale is needed in order to understand the potential natural defenses of the natural external bat microbiota.

In this study we sampled 186 bats collected from southeastern New Mexico to northwestern Arizona to gain insights into regional-scale patterns of external bat bacteria and the factors that drive these patterns. Specifically, we address two questions: First, to what extent are the changes in distributions of bat bacteria a function of geographic location, ecoregion (Omernik and Griffith, 2008), and climatic variables? Second, does being in a cave for 6-8 hours before sampling cause changes in the external bat microbiome? This is of importance given that bats are susceptible to WNS while hibernating in caves and differential exposure to microbes might explain differing levels of susceptibility.

Methods

Sampling. We sampled 186 bats belonging to 14 species (*Myotis ciliolabrum*, *M. californicus, M. evotis*, *M. occultus, M. thysanodes M. velifer, M. volans*, *Corynorhinus townsendii, Eptesicus fuscus, Tadarida brasiliensis, Antrozous pallidus, Parastrellus hesperus*, *Lasionycteris noctivagans* and *Lasiurus cinereus*, S2) using 16S rRNA gene analysis for external microbiome identification. These samples came from five study locations in the Southwest: Grand Canyon-Parashant National Monument (PARA), in Arizona, and Carlsbad Caverns National Park (CCNP), Fort Stanton-Snowy River Cave National Conservation Area (FS), El Malpais National Monument (ELMA), and Bureau of Land Management high grasslands (HGL) caves near Roswell, in New Mexico (Figure 2). Bat sample collection was allowed under the following permits: 2014 Arizona and New Mexico Game and Fish Department Scientific Collecting Permit (SP670210, SCI#3423, SCI#3350), National Park Service Scientific Collecting Permit (CAVE-2014-SCI-0012, ELMA-2013-SCI-0005, ELMA-2014-SCI-0001, PARA-2012-SCI-0003), Fort Collins Science Center Standard Operating Procedure (SOP) SOP#: 2013-01, and an Institutional Animal Care and Use Committee (IACUC) Permit from the University of New Mexico (Protocol #15-101307-MC) and from the National Park Service (Protocol #IMR_ELMA.PARA.CAVE.SEAZ_Northup_Bats_2015.A2).

Samples were collected from the spring to early autumn from 2011 through 2014. Cave-caught bats were either plucked from the walls of the caves in ELMA, FS, and HGL or netted in sterilized nets in Carlsbad Cavern in CCNP in a location along their flight path out of the cave. Cave-caught bats were typically sampled 6-8 hours after returning to the cave in the early morning. Surface-netted bats were netted after sundown using sterilized nets near water sources in CCNP, ELMA, FS, and PARA. All bats were handled with clean gloves and swabbed for DNA before other measurements were taken to limit contamination by human-associated microbiota. Using a sterile swab moistened with Ringer's Solution (Hille, 1984), the entire skin (i.e., ears, wings and uropatagia) and furred surfaces of each bat were thoroughly swabbed. While the bat biologist held the bat in appropriate positions to give access to the area to be swabbed, the

microbiologist rubbed each area with the sterile swab approximately 3-5 times, rotating the swab as the action was performed.

Each swab was placed in a sterile 1.7 ml snap-cap microcentrifuge tube containing 100 ul of RNAlater, and immediately frozen in a liquid nitrogen dry shipper or placed on dry ice. Samples were transported to the University of New Mexico and stored in a -80°C freezer. Samples were sent to MR DNA Molecular Research LP, Shallowater, Texas (http://www.mrdnalab.com/) for genomic DNA extraction and 454 sequencing diversity assays of bacterial 16S rRNA genes. The 186 samples were sequenced in nine runs. Barcoded amplicon sequencing processes were performed by MR DNA® under the trademark service (bTEFAP®). The 16S rRNA gene universal PCR primers 27F (5′– AGRGTTTGATCMTGGCTCAG -3′) and 519 R (5′– GWATTACCGCGGCKGCTG-3′) (Englebrektson et al. 2010), were used in a single-step 30 cycle PCR using the HotStarTaq Plus Master Mix Kit (Qiagen, USA) under the following conditions: 94°C for 3 minutes, followed by 28 cycles (5 cycle used on PCR products) of 94°C for 30 seconds, 53°C for 40 seconds and 72°C for 1 minute, after which a final elongation step at 72°C for 5 minutes was performed. Sequencing with the 27F primer was performed at MR DNA on a Roche 454 FLX titanium following the manufacturer's guidelines.

Sequence Processing. All 454 reads were processed in QIIME (Caporaso, et al., 2010). Primer and linker sequences were removed before analysis. Bacterial sequences shorter than 200 bp or longer than 500 bp, or containing bases with a quality score lower than 30, were excluded. The quality control and trimming was computed using the split_libraries command. Bacterial samples were denoised and clustered into operational taxanomic units (OTU) (at the 97% level) with pick_denovo_otus.py pipeline using the sumaclust option (Mercier et al., 2013). Chimera checking was done using usearch (Edgar, 2010) to detect artifacts created during sequencing. Taxonomy was assigned using SILVA123 database with uclust. Full QIIME workflow with all parameters used is available at: https://zenodo.org/record/17577#.

Alpha diversity analysis and normalization. Alpha diversity indices were carried out in QIIME using alpha_diversity.py command. Rarefaction curves plotted against observed species, chao1, chao1 standard error, and Shannon are available in the supplemental data (S1). Transformation of the count and richness data was carried out using the normalize_table.py in QIIME with both the DESeq2 and cumulative sum scaling (CSS) (Paulson, et al., 2013) options. The data was also rarified to a depth of 1500 for comparison purposes only.

Distribution of major phyla on bats. Bar plots of major phylum of interest were run in ggplot. Proteobacteria were broken out into the following classes: Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria. Each bar is the relative abundance of the phylum within a sample. Three plots (S4 a,b,c) show relative abundance by cave-caught or surface-netted, by ecoregion (split by cave or surface), and by bat species (split by cave or surface).

Testing grouping of categorical data. Random forest models were run in QIIME (supervised_learning.py) using 10-fold cross-validation with 1,000 trees. The random forest models were run to test if our classes of samples were predictive of the bacterial community composition. Random forest model, a type of supervised classification, was used to test the predictive power of the ecological classes. The goal of random forest model is to classify unlabeled communities based on a set of labeled training communities. This will generate a ratio of estimated generalization error and baseline error. A reasonable ratio of the estimated

generalization error compared to the baseline error should be two or greater, i.e. the random forests classifier does at least twice as well as random guessing for an unlabeled community.

Differential abundance between categories. DESeq2 (Love et al., 2014) was used to identify taxa that were differentially proportional between cave-caught and surface-netted bats and across ecoregions. DESeq2 was picked due to it's ability to correct for large differences in sample library size without loss of statistical power or increase false positive rates. A custom script for running DESeq2 in R was used and last accessed March, 4th, 2016 is here: (http://userweb.eng.gla.ac.uk/umer.ijaz) by Umer Zeeshan Ijaz.

Geographic distance and community similarity. Mantel tests were carried out on geographic distance using the vegan (Oksanen et al., 2007) package in R with 999 permutations. Multiple regression on distance matrices (MRM) was done in the ecodist (Goslee and Urban, 2007) package in R with 1,000 permutations. The paired geographic distance matrix for these analyses was calculated from the latitude and longitude using an R function written by Peter Rosenmai, last accessed at: http://eurekastatistics.com/calculating-a-distance-matrix-for-geographic-points-using-r March, 4th, 2016. Sorting of the distance matrix was done using the dendextend (Galili, 2015) package. Retrieval of the paired scores (distance and similarity) was done using an R function from http://stackoverflow.com/questions/21180464/distance-matrix-to-data-frame-pairs-in-r, last accessed March, 4th, 2016.

Ordination of bacteria OTUs. NMDS analysis was carried out using the phyloseq package (McMurdie and Holmes, 2013) and ggplot2 (Wickham, 2009) in R (R development core team, 2012). The main analysis was focused on drivers of beta diversity across different categories. The large differences in bacterial counts between samples was dealt with by running DESeq2 on the whole dataset before running a non-metric dimensional scaling (NMDS) with the Bray-Curtis distance. DESeq2 does a variance stabilized transformation of the data. However, NMDS is robust to large differences in counts so the DESeq2 transformation minimally changes the NMDS. The Bray-Curtis distance was used because it is invariant to changes in units and unaffected by additions of new communities, and NMDS was chosen because it uses rank orders and does not assume linear relationships. NMDS can make use of a variety of distance measures.

Modeling bacteria similarity and richness. Elevation data were taken from USGS NED1 courtesy of the U.S. Geological Survey, mean annual precipitation and temperature from World Climate database 1.4 (http://www.worldclim.org), net primary productivity (NPP) was sourced from MOD17A3 (http://images.ntsg.umt.edu/alg_desc.php?caid=6139), soil organic carbon from ORCDRC, and soil pH from PHIHOX (https://soilgrids.org/). Modeling of environmental parameters and grouping data were done in R using the rstanarm (Gabry and Goodrich, 2016) package using a generalized linear mixed effects model (glmer). We choose a gaussian family; a normal, weakly informed prior (normal(location = 0, scale = 8)); and 10,000 iterations. Grouping data were treated as random effects in a partial pooling model. Bacterial richness (response variable) was calculated from the variance stabilized data (VDS) from DESeq2 (which accounts for differences in library size, retains statistical power, and doesn't remove samples or taxa) in phyloseq using the sample_sums on the DESeq2 VDS phyloseq object. The full model call in R was:

```
SEED <- 101
fit_partialpool_richness <- stan_glmer(vsd_richness ~ (1 | species) + (1|month_cat) + (1|
ecoregion_iv) + elevation + mean_annual_prep_mm + mean_temp + log_npp + soil_org_c +
```

soil_ph, data = metadata, family = gaussian, adapt_delta = 0.95, prior = normal(location = 0, scale = 8), seed = SEED, iter=10000, cores=6)
The Rhat statistic was used to measure if the MCMC chains converged. Rhat measures the ratio of the average variance of the draws within each chain to the variance of the pooled draws across chains.

Associated taxa. Environmentally associated taxa were taken from Barberán et al. (2015), with the exception of the freshwater taxa (Newton et al., 2011) and cave taxa (from this study: Nitrospiraceae, Nitrospira, Acidimicrobiia, Rubrobacteria, Thermoleophilia , Acidobacteria-6, Nitriliruptoria, Sphingobacteriia, Gemm-1) which occurred in cave-caught bats. Differences between associated taxa groupings were tested using Bayesian First Aid (Bååth, 2013) with the Bayes t-test.

Normalizing the data. Microbiome studies deal with differences in library sizes (number of sequences per sample) in a variety of ways. Once standard practice of rarefying data (subsampling to an even depth), developed for plants communities is statistically inadmissible for microbial abundance data. While many important discoveries were made with rarefied data, doing this: removes real data (removal of OTUs); removes samples that can be clustered meaningfully by other methods (NMDS, DESeq2); results in loss of statistical power; and increases false positive rates when comparing abundance data across categories (see McMurdie and Holmes, 2014 for further details). McMurdie and Holmes carried out these tests on simulated data with known richness and abundances and real data sets. The authors also suggested better alternatives to rarefied data.

In our data there are 111,199 total taxa present in the data. When the data are rarefied the number of taxa drops to 40,163 and fifteen samples are removed. The DESeq2 transformation retains 80,329 OTUs and the CSS retains 102,653 OTUs. Both previous methods retain all samples. In this data set presented here the DESeq2 and CSS transformed data both correlated highly (similar correlation reported in Barberán, A, et al. 2015) with the rarefied (depth of 1500) data (Pearson's correlation 0.65 and 0.75, respectively). We used the DESeq2 transformed data for differential abundance across categories and ordination using NMDS. All other richness calculations used the CSS transformed data.

Data and workflow availability. Biome files, QIIME mapping files, workflow, and R scripts are available at https://github.com/bioinfonm/microBat/tree/batmicrobiom and are archived at https://zenodo.org/record/17577#. All raw sequence data with the quality files and mapping files are available at: https://zenodo.org/record/50976 . The full metadata table is available in the supplemental data (S2 table). A Binder (http://mybinder.org/) ipython notebook with the full dataset is available at: https://github.com/bioinfonm/bat_microbiome_plots . Cave names and location are encoded to protect park and BLM resources. The full cave names and sampling locations are protected by law by federal law and their respective agencies.

Results and Discussion

Microbial diversity on bats. Our study stands apart from culture-based studies and other next generation sequencing studies by focusing on the diversity of the external bacteria from 186 bats from 14 bat species across a broad range of environments. The number of reads after quality control range from 843 to 20,515 per sample. Sample coverage was measured by calculating the

Good's coverage, whose values (S4) ranged from 81% to 99%, with an average of 95.3%. Bacterial sequences that were unable to be assigned to the SILVA_123 database ranged from 0.7% to 50% across all samples. Very small portions (0% - 0.25%) of the sequences in the data could not be assigned to a phylum, but were identified as bacterial. At the class level between 0% to 0.55% could not be assigned to a class. Cave-caught bats were dominated by the phylum Actinobacteria (S1 a), whereas surface-netted bats were dominated by Cyanobacteria, Actinobacteria, and Alphaproteobacteria (S1 b).

The data were tested using a random forest model to see if the data could be classified by our metadata categories. Random forest models were minimally successful for determining sampling site (2.76), bat species (2.28) (see also S1 c for phylum distribution by bat species), and seasonality of sampling (2.61) associated with each sample. The models were successful for determining cave-caught or surface-netted with a ratio of 8.43 and ecoregion with a ratio of 3.20. Since the random forest model takes the OTU counts as predictors and the metadata (i.e ecoregion) as classes, we can classify a given bat as cave-caught or surface-netted and from what ecoregion it came from. The results from the random forest model confirmed visual differences that were seen in the phylum bar plots, along with the proportional changes in OTU abundance across ecoregions.

Proportional changes in OTUs between cave-caught and surface-netted bats were quantified using DESeq2. Cave-caught bats had proportionally greater amounts of Actinobacteria and Nitrospirae compared to the surface-netted bats (Figure 3). Surface-caught bats had proportionally higher Synergistetes, Cyanobacteria, FBP, Armatimonadetes, Thermi, Firmicutes, Alphaproteobacteria, TM7, Betaproteobacteria, GAL15, Fusobacteria, SBR1093, and Tenericutes. Across all ecoregions there were many phyla that varied in their proportions (S5), including : Acidobacteria, Cyanobacteria, Firmicutes, Deltaproteobacteria, Chloroflexi, Gemmatimonadetes, Planctomycetes, Synergistes, Thermi, Armatimonadetes, FBP, Alphaproteobacteria, Tenericutes, Actinobacteria, Fusobacteria, Verrucomicrobia, Betaproteobacteria, Chlorobi, Nitrospirae, and Epsilonproteobacteria. Some of the variation among ecoregions is due to bats only being caught in the cave (Chihuahuan Basins and Playas) or netted on the surface (i.e. Mojave ecoregions), with the remaining ecoregions having mixed cave-caught and surface-netted. The variation across ecoregions is due to local bacteria being picked up on the bats as opposed to more cosmopolitan bacteria seen on all bats.

Many of the OTUs were restricted to relatively few samples with very few shared taxa (S6), while a few OTUs were found more widely. Eighty percent of surface-netted bats shared 15 OTUs at the genus level. In cave-caught bats, 80% of the samples shared only eight OTUs. Across 80% of all bats sampled only six OTUs were shared and they belonged to the classes *Actinobacteria*, *Flavobacteria*, and *Gammaproteobacteria*. These bacterial classes are widely distributed across a range of environments. Bats are exposed to bacteria common in air and soil, and chloroplasts. Given the ability of bacteria to disperse over long ranges, one might expect surface-netted bat bacterial communities to be more homogenous than caves, but this is not the case. As shown in the DESeq2 result (Figure 3) cave-caught bats' external bacteria shift towards being dominated by *Actinobacteria* and *Nitrospirae*.

In a culture-dependent study by Borda et al. (2014), common bacteria found in the air above guano piles included: *Chryseomonas*, *Klebsiella*, *Micrococcus*, *Salmonella*, *Staphylococcus*, and *Streptococcus*. *Bacillus*, *Enterobacter*, *Enterococcus*, *Escherichia*, *Klebsiella*, *Pantoea*,

Pseudomonas and *Serratia* were found in the gut of the short-nosed fruit bat (Daniel et al., 2013). On the ocular surfaces (Leigue dos Santos et al., 2014) of 36 bats, the most common isolated bacteria were *Staphylococcus*, *Bacillus*, *Corynebacterium*, *Shigella*, *Hafnia*, *Morganella*, and *Flavobacterium*. In our study many of the same bacterial genera were present on the external surfaces of bats. The exception was the lack of *Chryseomonas*, *Klebsiella*, *Salmonella*, *Pantoea*, *Serratia* (found in six samples in low amounts), *Shigella*, and *Hafnia* in our samples. Habitat and host species were primary drivers of bacteria diversity and taxa composition on captive neotropical bats (Lemieux-Labonté, et al., 2016).

Structuring of Community Similarity Patterns and Richness. The bacterial community composition was highly variable within sampling sites and across geographic regions (S1 b). Community similarity patterns (NMDS) and the random forest model show that factors such as location (e.g. cave-caught vs. surface-netted) and ecoregion help to structure the bacterial communities on bats (Figure 4 and 5). In addition, bacterial community similarity was related to geographic distance (Figure 6). Communities that were geographically close were more similar, as indicated by a Mantel test [R_m=0.09, P=0.003], but the regression coefficient was weak.

Other microbiome projects from continental scale soil microbiomes (Ma et al., 2016), to whales and shrimp, noted that factors such as net primary productivity (NPP), rainfall, temperature, soil properties, and seasonality were correlated with patterns of bacterial richness and similarity (Apprill et al., 2014; Larsen et al., 2015). The Bayesian model for community similarity (Figure 7a and 7b) showed that regional variables (i.e. NPP, soil pH) were predictive for the MDS1 axis, but slightly less for the MDS2 axis. Three bats species contributed to the MDS1 (*Myotis velifer*, *Myotis volans*, *Tadarida brasiliensis*) and MDS2 (*Myotis velifer*, *Myotis occultus*, *Myotis evotis*) axes. The ecoregions Chihuahuan Basins and Playas, Chihuahuan Desert Slopes, and Conifer Woodlands and Savannas contributed to the MDS1 axis, while for the MDS2 axis the ecoregion contributed minimally. Cave-caught or surface-netted contributed to the MDS2 axis the most, while it minimally impacted the MDS1 axis. The Bayesian model for VDS corrected richness (Figure 7c) showed that regional variables (i.e NPP, soil pH), ecoregion, and months were predictive of richness. Bat species was also predictive of richness with values from ~ -60 to 90 but not to the extent of regional variables. Variables that contributed to bacterial richness were: NPP, soil pH, Conifer Woodlands and Savannas, and Chihuahuan Basins and Playas.

Climate, NPP, and soil makeup are highly linked, so it is difficult to tease apart which factors are directly responsible for structuring the bacterial community similarity and richness of bats. The month data (seasonality) is cofounded with whether during a given month bats were just cave-caught or surface-netted. In addition, some species of bats dominated the samples during a given month. Therefore, it is unlikely that a seasonal signal for the external bat bacteria communities can be determined from this dataset. Full model results are available in S7. Future studies should be designed to target the specific effects of soil properties, climate variables, seasonality, and bat species richness on the external microbiome.

Effects of cave and surface habitats on bacteria. Basic information on how roosting in a cave or flying on the surface affect a bat's external microbiome is lacking. This is particularly important to understand when addressing novel wildlife diseases, such as WNS, that may alter naturally occurring microbiomes. Because bats contract WNS while hibernating in caves, it is possible that the external microbiome offer natural defenses against WNS for some bat species (Hoyt et al., 2015). Thus, the overall distribution of bacteria among phyla changing between

cave-caught or surface-netted bats (Figure 3) after a period of 6-8 hours is important for bats vulnerable to WNS. Cave-caught bats have proportionally more Actinobacteria and Nitrospirae, while surface-netted bats had proportionally more Cyanobacteria, Firmicutes, and Synergistetes. Earlier studies in caves showed differences in community structure between surface soil and cave samples. This was seen in a carbonate cave speleothems in Arizona (Ortiz, et al., 2014) and the photic and aphotic zone in samples from two caves in the Antarctic (Tebo, et al., 2015).

In addition to seeing the effects of roosting in the cave on the external microbiome, we expected the source of microbes to vary between cave-caught and surface-netted bats. To test this hypothesis, we identified specific bacterial taxa typically associated with environmental sources. We visualized source associated taxa using violin plots (Figure 8). Violin plots are similar to box plots, but also show the probability density at a given value. Environmentally associated taxa included sources from: plants (Chloroplasts), soil, insect, freshwater, and caves. We would expect that for bats netted on the surface there would be proportionally more surface-associated taxa than for cave-caught bats. For example, we would expect more plant or freshwater associated bacteria in bats netted on the surface. We did detect differences, using a Bayesian t-test, in the mean proportions for plant-associated taxa (BEST mean difference for cave -0.11, 95% CI -0.15 - -0.083) and weak evidence for cave-associated taxa (BEST mean difference for cave 0.058 , 95% CI 0.017-0.061). There was no evidence for differences between insect-, soil-, and freshwater-associated taxa proportions (BEST mean difference for cave -0.00047, -0.0077, and -0.0026, respectively). We hypothesize that the few samples with high freshwater associated taxa are likely bats who were netted shortly after dipping into local water sources. In addition, there were several samples (~20) with high numbers of insect-associated taxa; likely these belong to bats that recently fed on insects before being netted, or had a high parasite load. Future bat microbiome studies should test these hypotheses.

Conclusions

Overall, our results show that the external microbial communities on bats follow similar local and regional-scale bacteria patterns as noted in the eastern China soils study (Ma, et al., 2016), bat fecal and internal microbiome studies (Daniel et al., 2013; Borda et al., 2014; Leigue dos Santos et al., 2014), and share predictors with bat species richness in Arizona and New Mexico (S8). We might expect some of the microbial patterns to be driven by differences in rates of bacterial dispersion. Unmeasured variables can contribute to both local and regional patterns. For example, average plant height and composition at a sampling site, and local bat foraging and roosting behavior might influence the bacterial communities on a smaller scale. From our study, we can show the importance of sourcing associated taxa with bats, such as foraging habits. Our data show that surface-caught bats carry proportionally more plant taxa (i.e. Chloroplasts), Cyanobacteria, and surface soil associated bacteria (Synergistetes, Armatimonadetes, and Firmicutes). ave roosting bats, on the other hand, have a greater proportion of Nitrospirae and Actinobacteria.

The bacteria found on bats caught in the cave tend to be more homogeneous, i.e. collapsing into two dominant phyla. We think that bats in caves are exposed to cave bacteria and a reduced number of surface taxa. Therefore, bats caught in caves after spending 6-8 hours in the cave trended towards having more Actinobacteria, while other taxa (except Nitrospirae) were reduced in proportions when compared with surface-netted bats. In general, we conjecture that the number of species, life histories, morphology, and ecology of bats occurring in the Southwest is

diverse and at times can be complicated when trying to categorize species into defined groups. Therefore, there are many factors that can contribute to the presence of certain bacteria found on the external surfaces of the bats, both among and between species. However, it does appear that some of the aforementioned aspects of bats found in the Southwest give insight to unique trends observed in the bacteria found in this study. For example, bacterial richness of bats can be predicted by ecoregion, regional variables such as NPP and soil pH. While bat species does contribute to variation in bacterial richness, only a small number of species contribute to the differences in community similarity.

Our results shed new light on the external microbiome of southwestern bat species and the extent to which geographic, biotic, and abiotic factors influence the bacterial diversity patterns observed on different bat species. These results provide an important baseline characterization of bat bacterial microbiomes in non-WNS affected area, and provide the basis for exploration of potential bacterial defenses possessed by different bat species.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Acknowledgments

We thank the staff at El Malpais and Grand Canyon-Parashant National Monuments, Carlsbad Caverns National Park, Bureau of Land Management, and the Fort Stanton Cave Study Project. Funding was provided by the National Park Service (CPCESU) and Western National Park Association for work in El Malpais National Monument, Carlsbad Caverns National Monument, and Grand Canyon-Parashant National Monument. The Bureau of Land Management and Fort Stanton Cave Study Project funded work in Fort Stanton and BLM Caves 45 and 55. Additional funding was provided by the New Mexico Game and Fish Department Share with Wildlife Program, Cave Conservancy Foundation, Eppley Foundation, National Speleological Society Rapid Response Fund, Western National Park Association, and T&E, Inc. We thank Graham Walmsley for writing suggestions; Brennen Reece for graphic design and typographic help; and Ken of Kenneth Ingham Photography for the bat photo.

Author Contributions A.S.W. contributed to the bacterial data analysis, methods, results, discussion and fieldwork. J.C.K. contributed to the writing, results, and discussion. J.M.Y. contributed to the data analysis, methods, results, discussion, and fieldwork. E.W.V. contributed to funding acquisition, data collection, writing, and interpretation of data related to bat ecology. A. P-A. contributed to writing, editing, and workflow. J.J.H. contributed to the writing, editing, and data analysis. D.E.N. contributed to study design, funding acquisition, data collection, editing, and interpretation of habitat characteristics and bacterial sequencing results. D.C.B. contributed to study design, funding acquisition, data collection and discussions regarding bat ecology. K. J. H. R. contributed to study design and fieldwork.

References

Apprill, A., Robbins, J., Eren, A.M., Pack, A.A., Reveillaud, J., Mattila, D., Moore, M., Niemeyer, M., Moore, K.M., Mincer, T.J. (2014). Humpback whale populations share a core skin bacterial community: towards a health index for marine mammals?. PloS one. Mar 26;9(3):e90785.

Bååth, R. (2013). Bayesian First Aid, https://github.com/rasmusab/bayesian_first_aid

Barberán, A., Ladau, J., Leff, J.W., Pollard, K.S., Menninger, H.L., Dunn, R.R. and Fierer, N., (2015). Continental-scale distributions of dust-associated bacteria and fungi. Proceedings of the National Academy of Sciences, 112(18), pp. 5756-5761.

Bogan, M.A., Cryan, P.M., Valdez, E.W., Ellison, L.E., and O'Shea, T.J. (2003). Western crevice and cavity-roosting bats. In Monitoring trends in bat populations of the United States and territories: problems and prospects: U.S. Geological Survey, Biological Resources Discipline, Information and Technology Report, USGS/BRD/ITR--2003--0003, 274p.

Borda, D.R., Nastase-Bucur, R.M., Spînu, M., Uricariu, R., Mulec, J. (2014). Aerosolized microbes from organic rich materials: case study of bat guano from caves in Romania. Journal of Cave and Karst Studies. Aug 1;76(2):114.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 7(5), 335-336.

Carrillo-Araujo, M., Taş, N., Alcántara-Hernández, R.J., Gaona, O., Schondube, J.E., Medellín, R.A., Jansson, J.K., Falcón, L.I. (2015). Phyllostomid bat microbiome composition is associated to host phylogeny and feeding strategies. Frontiers in microbiology.

Daniel, D.S., Ng, Y.K., Chua, E.L., Arumugam, Y., Wong, W.L., Kumaran, J.V. (2013). Isolation and identification of gastrointestinal microbiota from the short-nosed fruit bat Cynopterus brachyotis brachyotis. Microbiological research. Oct 1;168(8):485-96.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 26, 2460-2461.

Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H. and Hugenholtz, P., (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *The ISME journal*, *4*(5), pp.642-647.

Findley, J.S., Harris, A.H. , Wilson, D.E., and Jones, C. (1975). Mammals of New Mexico. University of New Mexico Press, 360 p.

Frey, J.K. (2004). Taxonomy and distribution of the mammals of New Mexico: an annotated checklist. Museum of Texas Tech University.

Frick, W.F., Pollock, J.F., Hicks, A.C., Langwig, K.E., Reynolds, D.S., Turner, G.G., Butchkoski, C.M., Kunz, T.H. (2010). An emerging disease causes regional population collapse of a common North American bat species. Science. Aug 6;329(5992):679-82.

Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics. Jul 23:btv428.

Gabry, J. and Goodrich, B. (2016). rstanarm: Bayesian Applied Regression Modeling via Stan. R package version 2.11.1. https://CRAN.R-project.org/package=rstanarm

Goslee, S.C., Urban, D.L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software. Sep 30;22(7):1-9.

Hall, E.R., (1981). The mammals of North America. 2nd Edition. John Wiley and Sons, New York, 1181 pp.

Harvey, M.J., Altenbach, J.S., Best, T.L. (2011). Bats of the United States and Canada. The John Hopkins University Press, 224p.

Hille, B. (1984). Ionic Channels of Excitable Membranes. Sunderland, Mass. 01375, Sinauer Associates, Inc.

Hoffmeister, D.F. (1986). Mammals of Arizona. Tucson: University of Arizona Press.

Hoyt, J.R., Cheng, T.L., Langwig, K.E., Hee, M.M., Frick, W.F. and Kilpatrick, A.M. (2015). Bacteria isolated from bats inhibit the growth of *Pseudogymnoascus destructans*, the causative agent of white-nose syndrome. PLoS One, 10(4), p.e0121329.

Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. Nature. 486(7402), pp.207-214.

Humphrey, S.R. (1975). Nursery roosts and community diversity of Nearctic bats. Journal of mammalogy. May 30;56(2):321-46.

Larsen, A.M., Bullard, S.A., Womble, M., Arias, C.R. (2015). Community structure of skin microbiome of gulf killifish, Fundulus grandis, is driven by seasonality and not exposure to oiled sediments in a Louisiana salt marsh. Microbial Ecology. Aug 1;70(2):534-44.

Leigue dos Santos, L., Montiani−Ferreira, F., Lima, L., Lange, R. (2014). Bacterial microbiota of the ocular surface of captive and free−ranging microbats: *Desmodus rotundus*, *Diameus youngi* and *Artibeus lituratus*. Veterinary ophthalmology. May 1;17(3):157-61.

Love, M.I., Huber, W., Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biology.

Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., Brookes, P.C., Xu, J. and Gilbert, J.A., (2016). Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. The ISME journal. 10, 1891–1901.

Maher, S.P., Kramer, A.M., Pulliam, J.T, Zokan, M.A., Bowden, S.E., Barton, H.D., Magori, K. and Drake, J.M. (2012). Spread of white-nose syndrome on a network regulated by geography and climate. Nature Communication 3:1306.

McMurdie PJ, Holmes S. (2014) Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol. Apr 3;10(4).

Mercier, C., Boyer, F., Bonin, A., Coissac, E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. SeqBio Invited Talks. Available at: http://www.gdr-bim.cnrs.fr/seqbio2013/wp-content/uploads/2013/12/seqbio2013-actes.pdf#page=28. Accessed through the QIIME 1.9.1 software package.

Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. Microbiology and Molecular Biology Reviews. Mar 1;75(1):14-49.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H., Oksanen, M.J., Suggests, M.A. (2007). The vegan package. Community ecology package. Oct 3;10.

Omernik, J., and Griffith, G. (2008). Ecoregions of the United States-Level IV (EPA). Retrieved from http://www.eoearth.org/view/article/152243

Ortiz, M., Legatzki, A., Neilson, J.W., Fryslie, B., Nelson, W.M., Wing, R.A., Soderlund, C.A., Pryor, B.M., Maier, R.M. (2014). Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. The ISME journal. Feb 1;8(2):478-91.

Paulson, J.N., Stine, O.C., Bravo, H.C., Pop, M. (2013) Differential abundance analysis for microbial marker-gene surveys. Nature methods. Dec 1;10(12):1200-2.

R development core team. (2012). R: A language and environment for statistical computing. Vienna: R foundation for statistical computing. Available: http://www.R-project.org.

Tebo, B.M., Davis, R.E., Anitori, R.P., Connell, L.B., Schiffman, P., Staudigel, H. (2015). Microbial communities in dark oligotrophic volcanic ice cave ecosystems of Mt. Erebus, Antarctica. Frontiers in Microbiology. 6.

Wickham H. (2009). ggplot2: elegant graphics for data analysis. (New York: Springer)

Wilson, D.E., and Reeder, D.M. (2005). Mammal Species of the World: A Taxonomic and Geographic Reference. 3rd edition (D.E. Wilson and D.M. Reeder, eds.). The John Hopkins University Press. Vol. 1, 743p.

Disclaimer:
"This draft manuscript is distributed solely for purposes of scientific peer review. Its content is deliberative and predecisional, so it must not be disclosed or released by reviewers. Because the manuscript has not yet been approved for publication by the U.S. Geological Survey (USGS), it does not represent any official USGS finding or policy."

Figure 1.

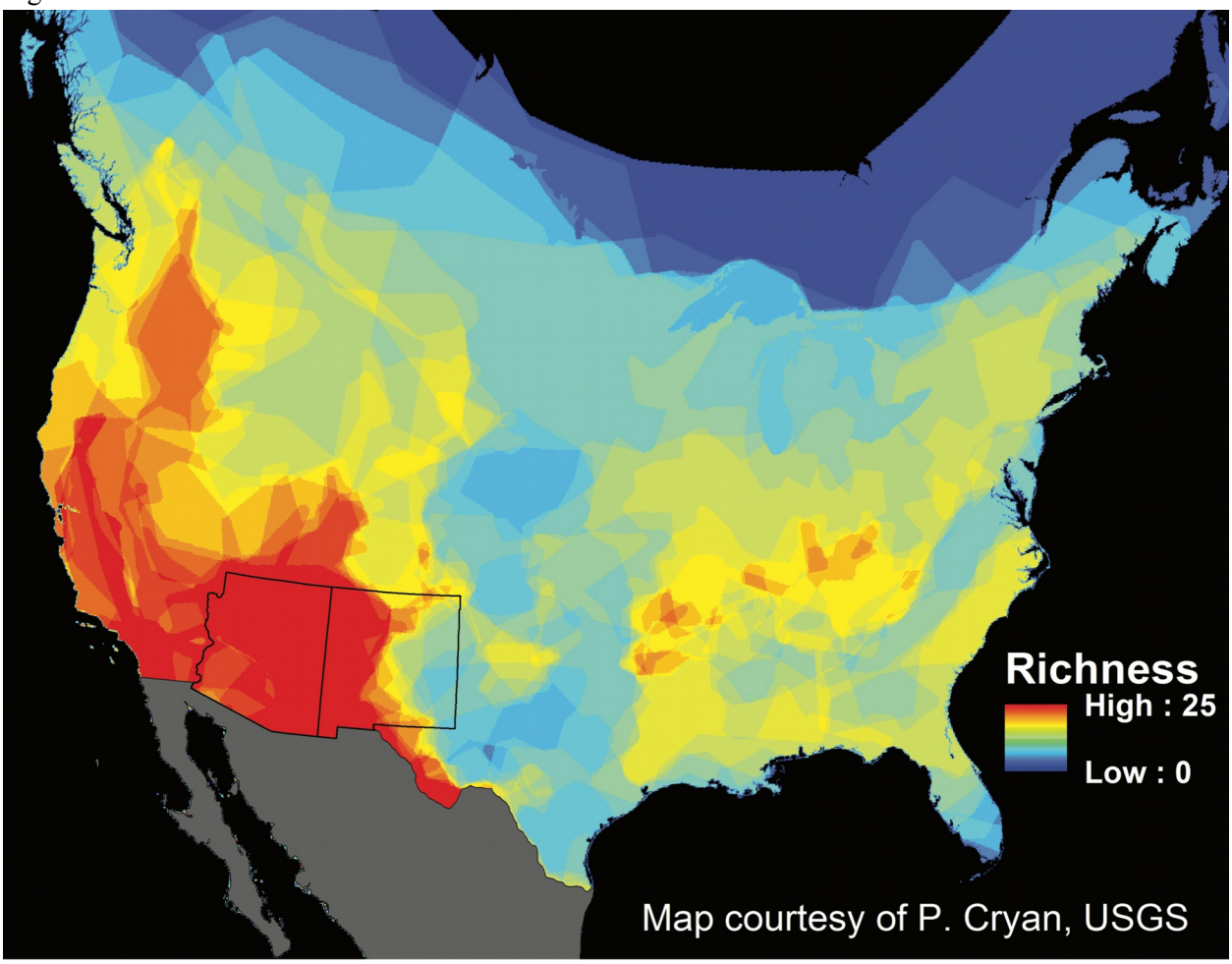


Figure 1. Map of bat species richness in the United States (US) and Canada. Total number of bat species occurring in each area calculated by counting the number of overlapping species distributions, as represented by the US National Atlas Bat Ranges geospatial data set (available at https://catalog.data.gov/dataset/north-american-bat-ranges-direct-download). Warmer colors represent areas with higher species richness and cooler colors represent areas with lower species richness. Map courtesy of P. Cryan, US Geological Survey.

Figure 2.

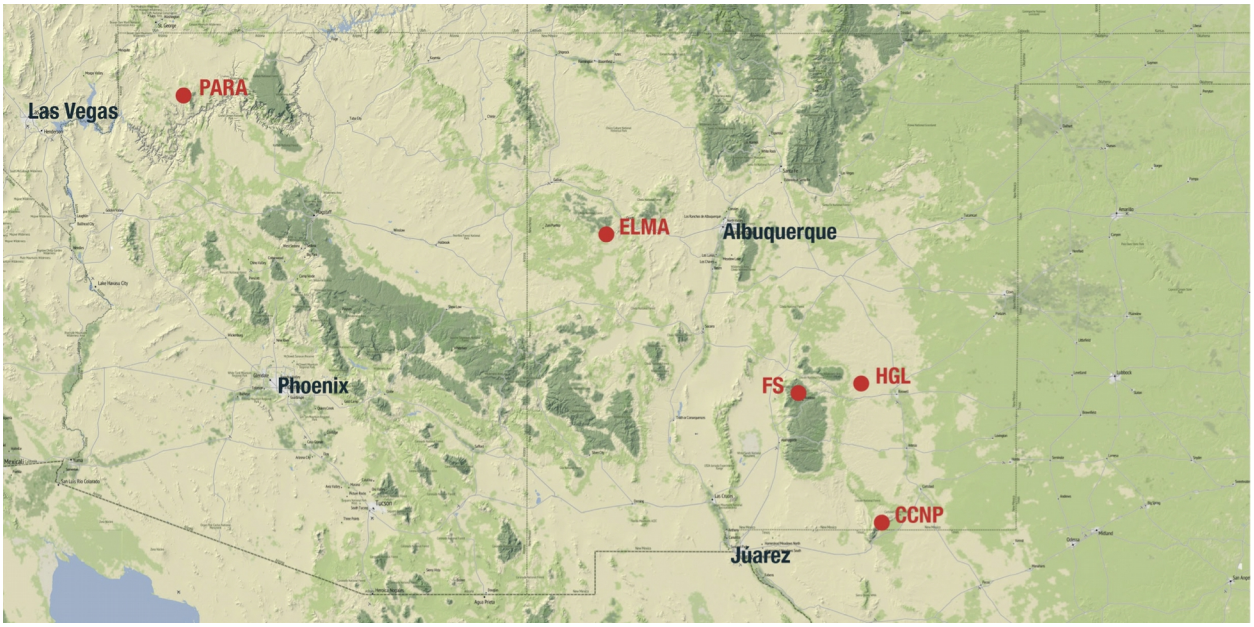


Figure 2. Map showing the general locations of the sampling sites in the southwestern United States. PARA (Grand Canyon Parashant National Monument), ELMA (El Malpais National Monument), FS (Fort Stanton-Snowy River Cave National Conservation Area), HGL (High Grasslands), CCNP (Carlsbad Caverns National Park). Elevation base map by Stamen, CC-BY OpenStreetMap Terrain.

665 Figure 3.

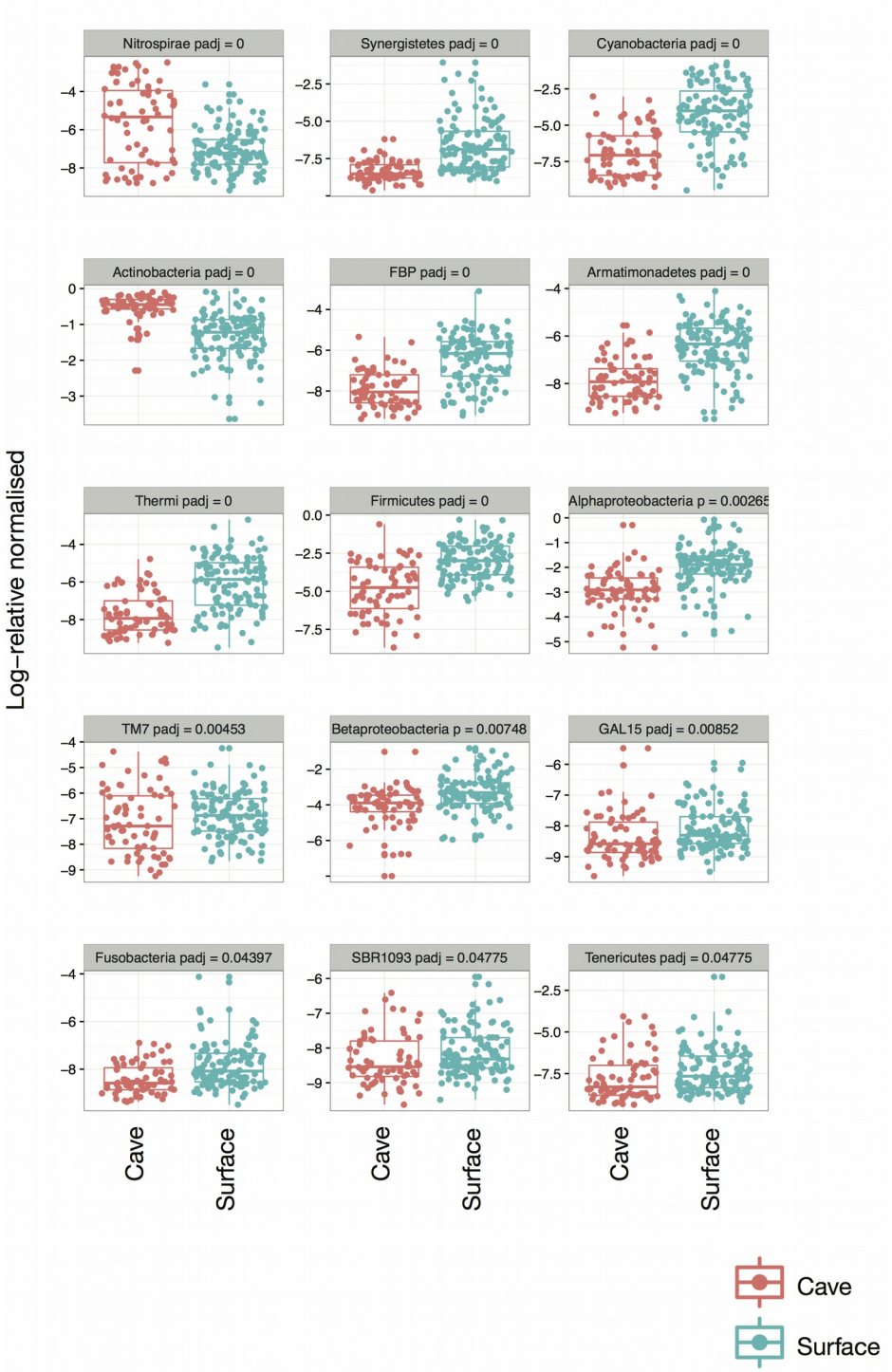


Figure 3. Differential abundance of OTUs between cave-caught and surface-netted bats using DESeq2. Two OTUs at the phylum level were higher in cave-caught bats: Actinobacteria and Nitrospira.

670

Figure 4.

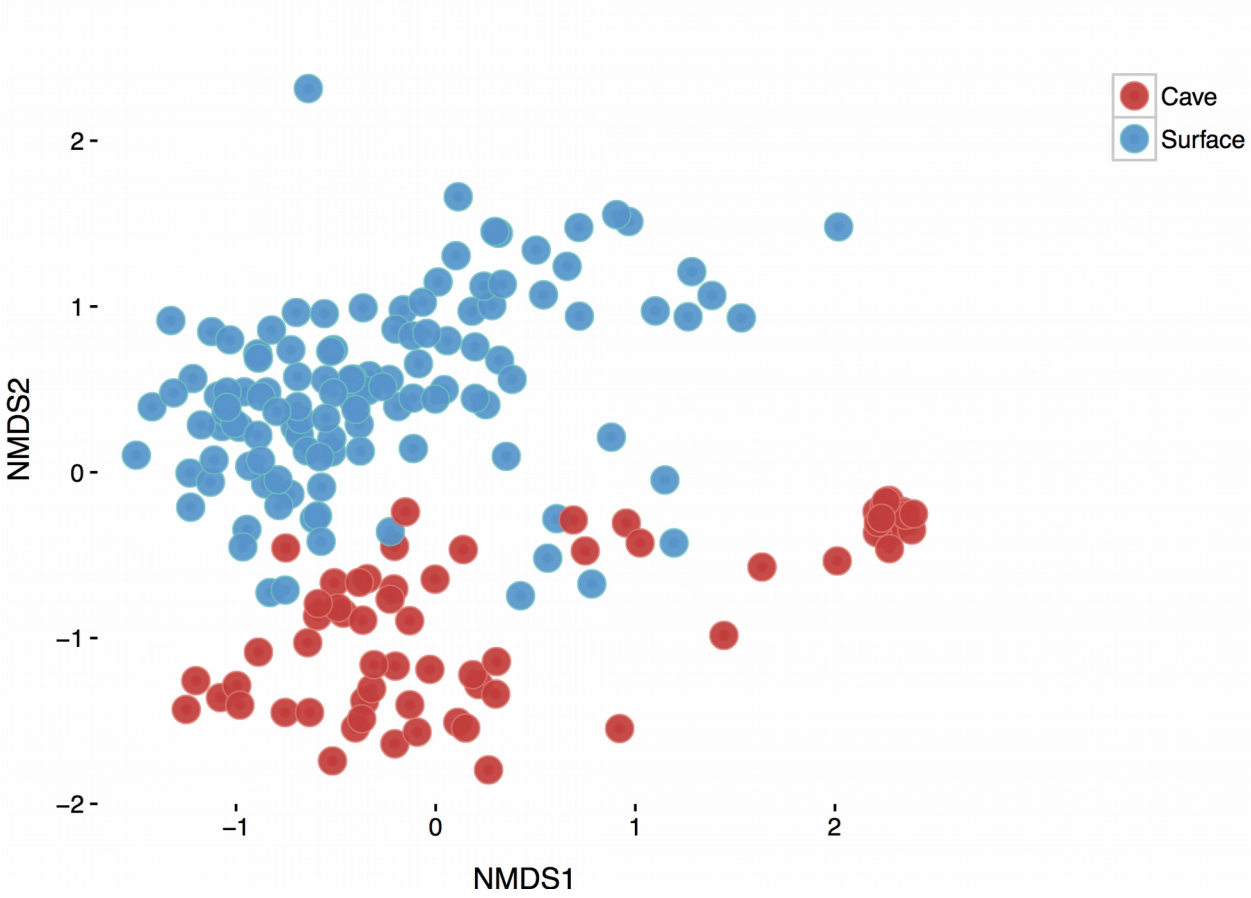
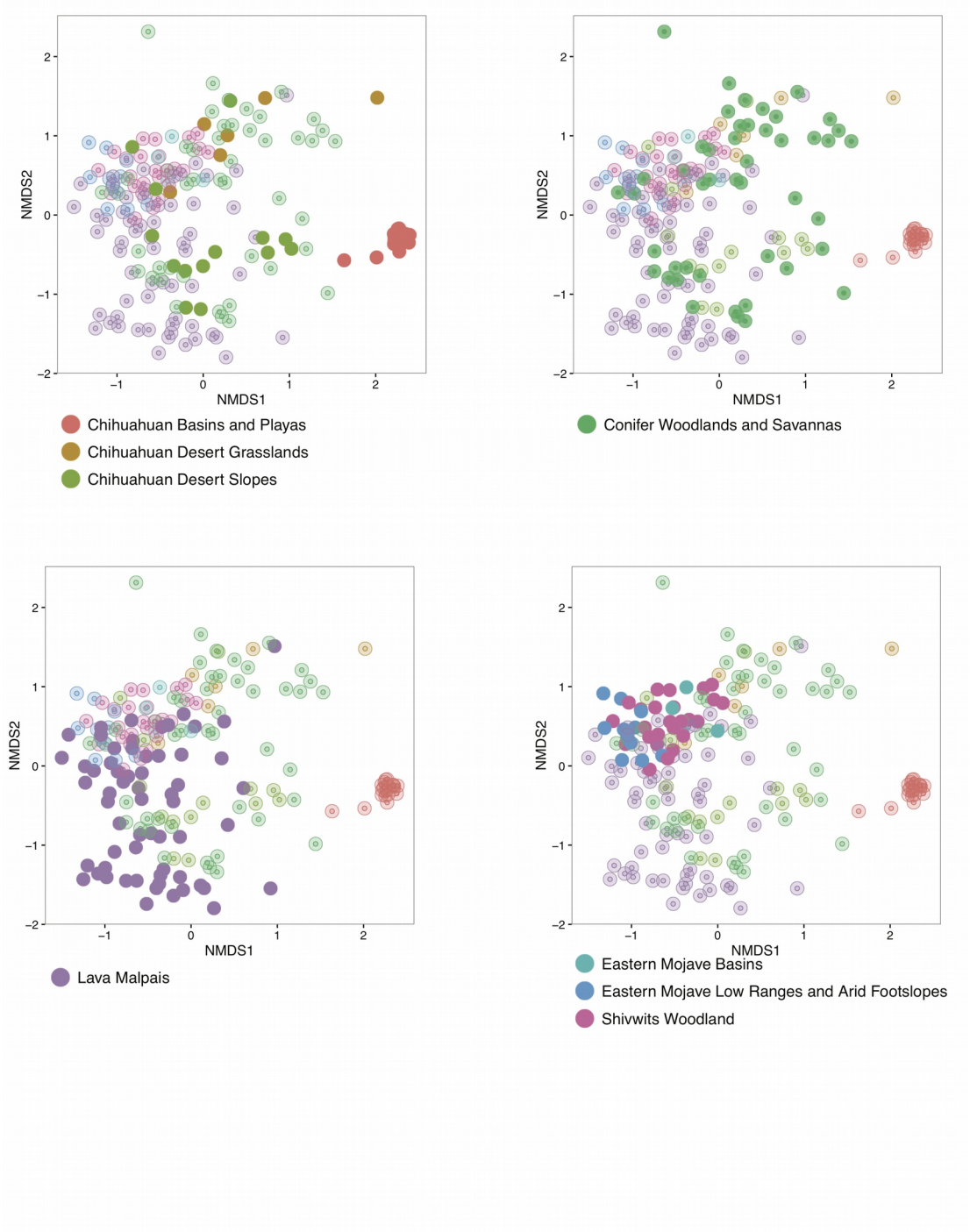


Figure 4. Similarity in the composition of the bacterial communities was quantified using NMDS (stress = 0.084) with the Bray-Curtis distance metric. Symbols are colored by location of capture. Samples closer together represent samples with more similar bacterial communities. The samples tend to cluster by cave-caught or surface-netted.

Figure 5.



695 **Figure 5.** Similarity in the composition of the bacterial communities was quantified using NMDS (stress = 0.084) with the Bray-Curtis distance metric. Symbols are colored by EPA Ecoregion IV. Samples closer together represent samples with more similar bacterial communities. The grey line represents the split between surface-netted and cave-caught bats.

700

Figure 6.

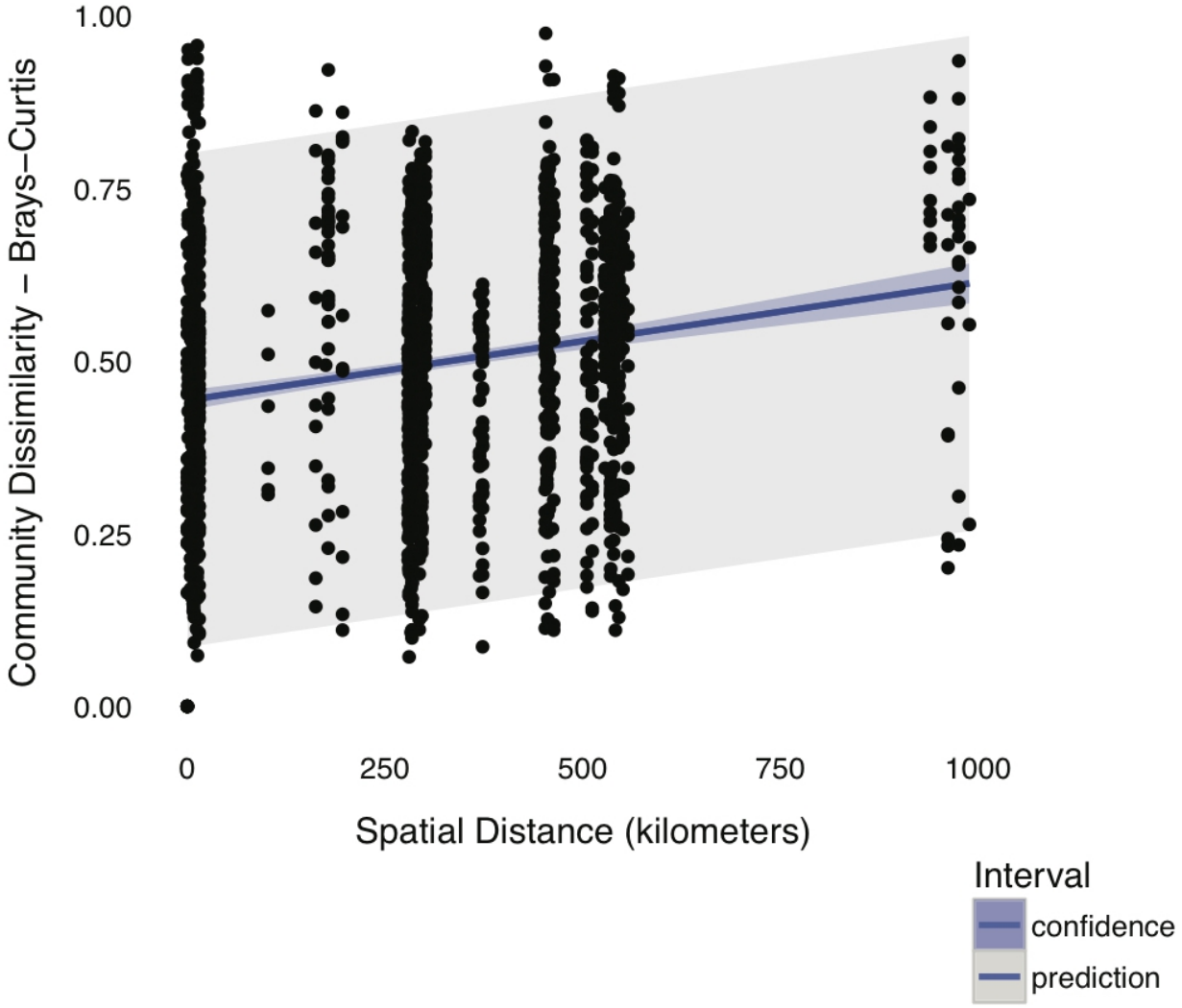
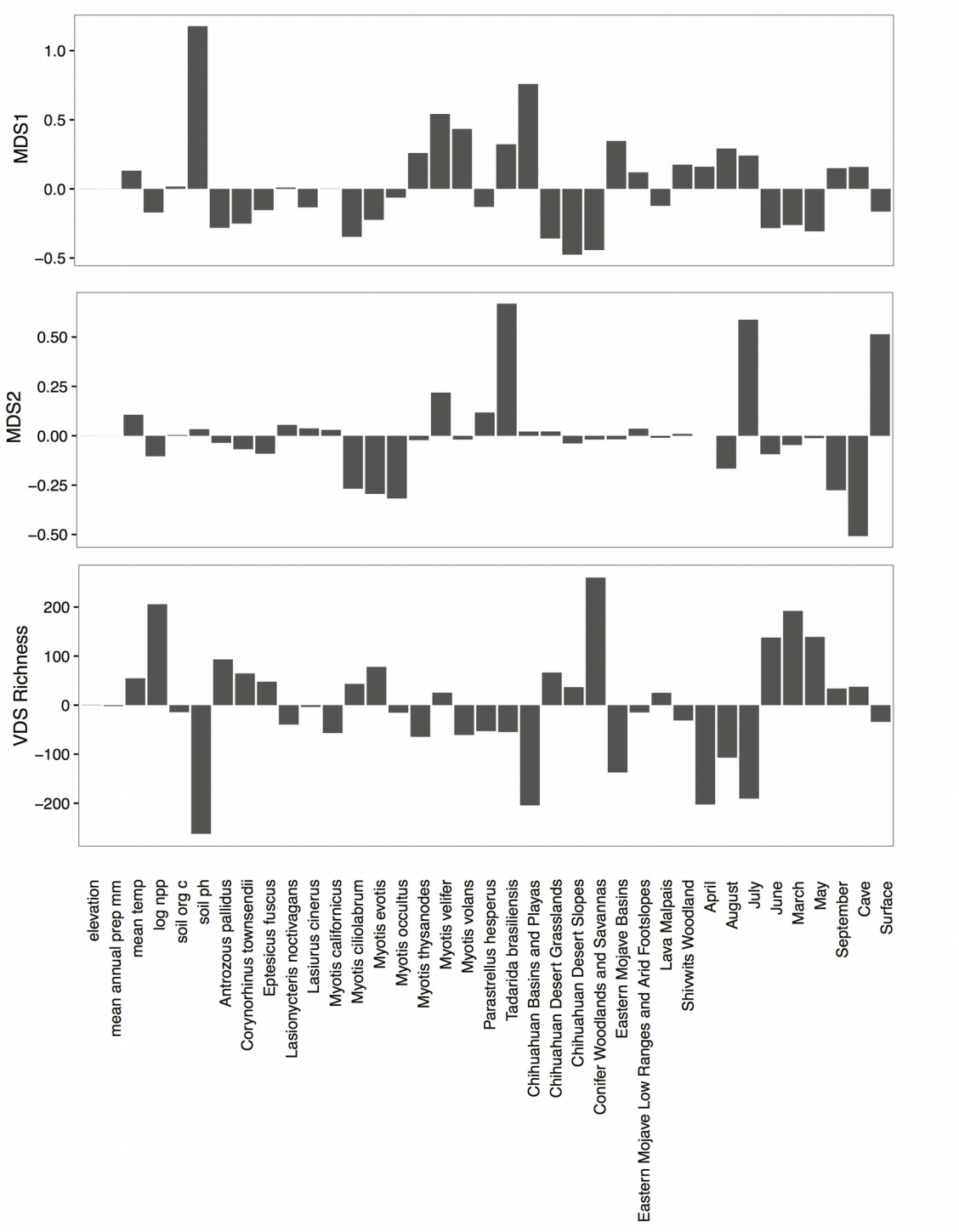


Figure 6. Relationship between paired community similarity and distance scores. The slope is 1.710e-07 with a corrected R2 of 0.04. Blue is the 95% confidence interval and grey is the predicted from the linear model.

Figure 7.



720

Figure 7. Bayesian model (glmer) values of predictors for a) MDS1 and MDS2 (community similarity), and b) bacterial richness (total OTUs).

Figure 8.

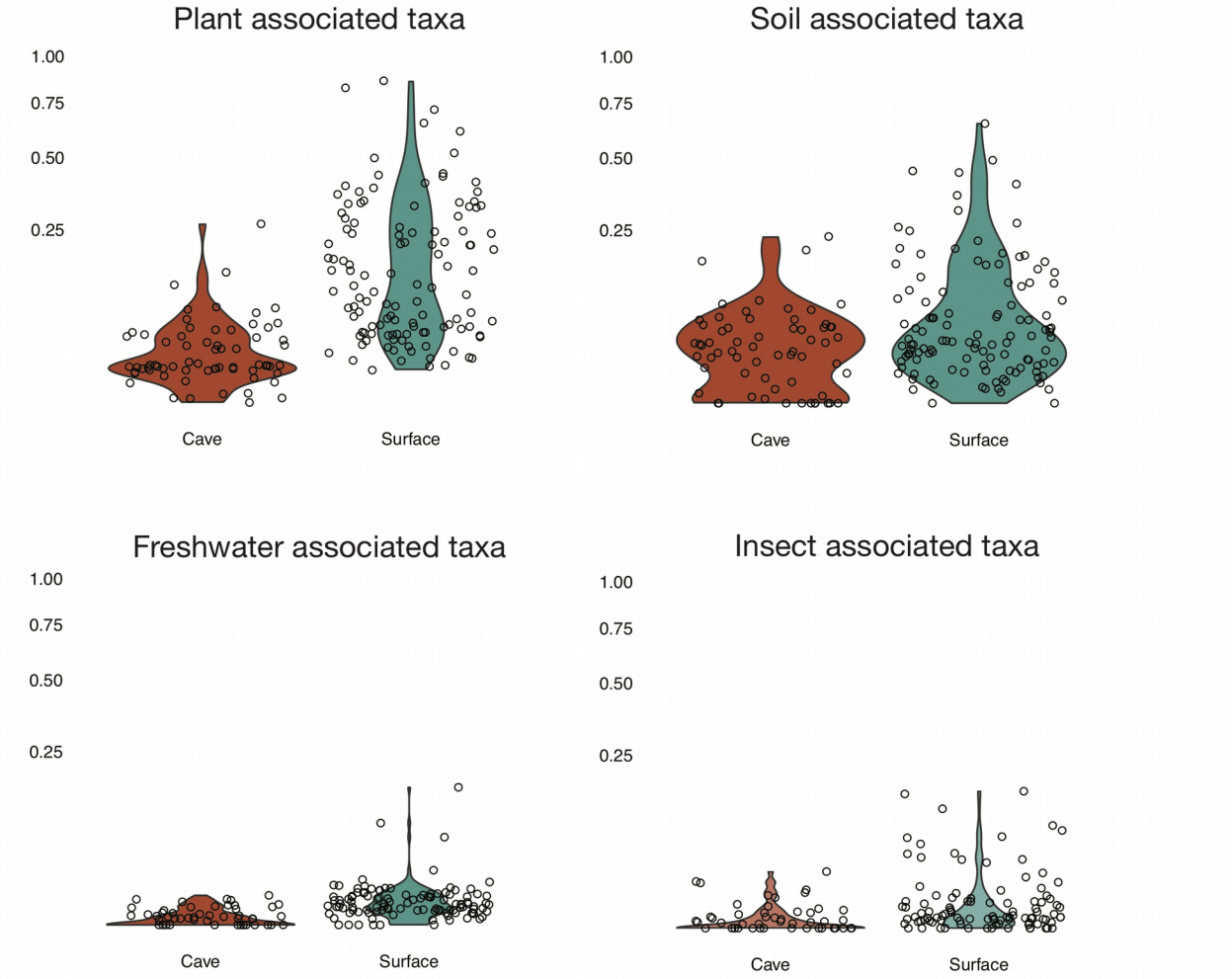


Figure 8. Square root proportion of bacterial sequences identified as indicator taxa of cave-caught or surface-netted bats. Scale is the proportion of the total number of OTUs in a sample.