Microbiome community composition and zoonotic bacterial pathogen prevalence in synanthropic *Peromyscus* mice

Running title: Microbiome composition of synanthropic mice

Janine Mistrick [1], Evan J. Kipp [2], Sarah I. Weinberg [3], Collin C. Adams [4], Peter A. Larsen [2†], Meggan E. Craft [1†]

[†] co-senior authors

[1] Department of Ecology, Evolution, and Behavior, College of Biological Sciences, University of Minnesota, St. Paul, MN, USA

[2] Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, St. Paul, MN, USA

[3] School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA

[4] Itasca Biological Station and Laboratories, University of Minnesota, Lake Itasca, MN, USA

Corresponding author: Janine Mistrick - mistr033@umn.edu

ABSTRACT

Rodents are key reservoirs of zoonotic pathogens and play an important role in disease transmission to humans. Importantly, anthropogenic land-use change has been found to increase the abundance of synanthropic rodents, particularly rodent reservoirs of zoonotic disease. Anthropogenic environments also affect the microbiome of synanthropic wildlife, influencing wildlife health and potentially introducing novel pathogens. Our objective was to characterize the microbiome and investigate the prevalence of zoonotic bacterial pathogens in synanthropic rodents in native and anthropogenic environments to better understand their role in pathogen maintenance and transmission. We sampled wild *Peromyscus* mice in agricultural and undeveloped landscapes and forest and synanthropic habitat in Minnesota, USA and conducted 16S amplicon sequencing using long-read Nanopore sequencing technology on fecal samples to characterize the rodent microbiome. We compared community composition and diversity between habitats and screened for the presence of putative pathogenic bacteria species. Microbiome community composition differed significantly between agricultural and undeveloped landscapes and forest and synanthropic habitat while microbiome richness, diversity, and evenness were lower in undeveloped-forest habitat compared to all other habitats. We detected overall low abundance and diversity of putative pathogenic bacteria, though the greatest number of pathogenic bacteria were detected in the agricultural-forest habitat. Our findings show that rodent microbiome community composition differs across landscapes and habitat types but suggest that landscape-level anthropogenic factors may be most important to predict zoonotic pathogen abundance. Ultimately, understanding how anthropogenic land-use change and synanthropy affect rodent microbiomes and pathogen prevalence is important to managing transmission of rodent-borne zoonotic diseases to humans.

KEYWORDS

microbiome, Nanopore sequencing, *Peromyscus*, synanthropy, zoonoses, 16S amplicon sequencing

1 INTRODUCTION

Rodents are an important source of zoonotic disease spillover, accounting for a greater diversity of zoonotic pathogens than any other mammalian order (Han et al., 2016). While many factors have been proposed to contribute to this (e.g. fast-paced life history, Han et al., 2015; cyclic population fluctuations, Kallio et al., 2009) recent studies have suggested that the tendency of particular rodent species to occasionally or exclusively live in human-built environments (synanthropy) is likely a key factor (Ecke et al., 2022).

Anthropogenic land-use change, leading to habitat fragmentation and the intensification of agricultural development and urbanization, is the major driver of zoonotic pathogen spillover (Gottdenker et al., 2014). Indeed, urbanized habitat has been found to have a significant, positive effect on the abundance of rodent hosts of zoonotic pathogens compared to areas of native vegetation (Mendoza et al., 2019). Shifts in rodent biodiversity in anthropogenic landscapes could further increase zoonotic risk, as rodent hosts and non-host rodents show opposite responses to agricultural and urban habitat, with the abundance of host species increasing and non-host species decreasing compared to areas of minimally disturbed primary vegetation (Gibb et al., 2020).

However, spillover of zoonotic pathogens at the human-wildlife interface does not solely flow from wildlife into humans. Synanthropic wildlife (including rodents) also show increased prevalence of human pathogens: *Escherichia coli*, *Clostridioides difficile*, *Salmonella enterica* in Norway rats in New York City, New York (Firth et al., 2014); antimicrobial-resistant *E.coli* in racoons in Chicago, Illinois (Worsley-Tonks et al., 2021); *Salmonella* in urbanized white ibis in

southern Florida (Hernandez et al., 2016), representing both a concern for wildlife health and a potential source for spillback into human populations. As such, while the relationships between land-use change, rodents, and zoonotic pathogen prevalence are still being explored, synanthropic wildlife represent both important reservoirs for zoonotic pathogens and likely drivers of pathogen maintenance and spillover in anthropogenic landscapes (Hassell et al., 2017).

Synanthropy has also been shown to impact the gut microbiome of wildlife. The gut microbiome plays a role in host health (Marchesi et al., 2016) and immune function (Schluter et al., 2020) and disruption of the normal microbiome has been linked to various health conditions in wildlife, livestock, and domestic animals (Funosas et al., 2021; Monteiro & Faciola, 2020; Suchodolski, 2022). Wildlife living in close proximity to humans often experience changes to the composition of their microbiome compared to counterparts in native habitat (e.g. rodents, Anders et al., 2022; sparrows, Berlow et al., 2021) though whether anthropogenic habitats decrease or increase microbiome diversity may vary by species (Diaz et al., 2023; Dillard et al., 2022). It is likely that changes in microbiome diversity associated with synanthropy could increase the prevalence of pathogenic bacteria in wildlife, but studies linking microbiome shifts with pathogen prevalence are limited (but see Murray et al., 2020).

Here, we characterize the microbiome and compare the abundance of zoonotic bacterial pathogens in *Peromyscus* mice in agricultural developed and undeveloped landscapes and forest and synanthropic habitat in Minnesota, USA. Our research questions were two-fold: 1) How does the microbiome community of *Peromyscus* mice differ between forest and synanthropic habitat? and 2) Are zoonotic bacterial pathogens more abundant in agricultural developed landscapes?

We expected the microbiome community of *Peromyscus* to be shaped by the surrounding landscape and specific habitat as they influence the availability of food resources and exposure to humans and their pathogens. We predicted that microbiome richness and diversity would be lower in the agricultural landscape and synanthropic habitat compared to the undeveloped landscape and forest habitat due lower diversity of food resources. We predicted that the agricultural landscape would have a higher abundance and diversity of pathogenic bacteria since the area is dominated by crop fields and human habitation and thus increased exposure to manure as fertilizer, wastewater and runoff, and trash; whereas we predicted that the undeveloped landscape would have lower pathogen abundance because the surrounding area is largely forested with little anthropogenic development. Characterizing rodent microbiomes across development gradients is important for quantifying the risk of rodent-borne zoonotic pathogen spillover and understanding how microbiome shifts associated with synanthropy may influence pathogen abundance.

2 MATERIALS & METHODS

2.1 Study Sites

Three major North American biomes intersect in Minnesota: the eastern deciduous forest, northern coniferous forest, and western prairie, providing diverse habitats and biological communities of hosts and pathogens. With respect to land-use, the state is dominated by agricultural cropland and forest with interspersed developed areas ranging from dense metropolitan areas to small, rural communities. Together, the biological and anthropogenic factors create a heterogeneous landscape of natural areas mixed with agricultural and urban developed landscapes where synanthropic rodents have many opportunities to overlap with humans. We focus our study on mice of the genus *Peromyscus* (i.e. *Peromyscus leucopus* and *Peromyscus maniculatis*) which are highly adaptable generalists that are common throughout Minnesota and can thrive in agricultural and urban settings as well as forests and grasslands.

Importantly, *Peromysucus* mice are known reservoirs of zoonotic and foodborne pathogens (e.g. *Borrelia*, *Campylobacter* spp., *E. coli*, *Giardia* spp., hantavirus; Jahan et al., 2021).

For our study, we focused on two landscape types: native, contiguous forest with little permanent human habitation or agriculture (hereafter "undeveloped landscape") and a mosaic of fragmented forest interspersed with crop fields and low-density housing (hereafter "agricultural landscape"). Within each landscape, four study sites were chosen to represent two habitat types (two sites per habitat): forest habitat and synanthropic habitat around human-frequented structures (e.g. cabins, tent platforms, field station buildings, maintenance sheds and garages). Rodent sampling was conducted at two locations: the Itasca Biological Station and Laboratories at Itasca State Park ("Itasca", undeveloped landscape) and Cedar Creek Ecosystem Science Reserve ("Cedar Creek", agricultural landscape). Itasca is located in northern Minnesota in the northern coniferous boreal forest biome. Though the state park is frequented by hikers and visitors, the surrounding landscape is contiguous forest with no agricultural development and very sparse permanent human habitation (Figure 1-A). Cedar Creek is located in central Minnesota in the eastern deciduous forest and oak savanna biome. The landscape surrounding the reserve is dominated by agricultural development (e.g. pasture/hay, cultivated crops), woody and herbaceous wetlands, and low-medium intensity housing communities (Figure 1-B).

2.2 Rodent Trapping

Two consecutive nights of rodent trapping were conducted at each study site (a "trapping session") using 100 Sherman live-capture traps baited with oats. Traps were opened at dusk and checked at dawn the following morning. Traps were closed during the day between trap nights at a single site and were reopened at dusk for the second night. Captured *Peromyscus* mice were sampled and then released at the point of capture. Due to the difficulty in

distinguishing *P. leucopus* and *P. maniculatus* – two species found across our study landscapes in Minnesota – based solely on morphologic features, we did not attempt to identify captured *Peromyscus* mice to the species level. Captured non-target (i.e. non-*Peromyscus*) species were released immediately and were not sampled. Longitudinal trapping was conducted at the agricultural landscape sites. Each site was sampled three times throughout the summer (June, July, and August 2019) with 3-4 weeks between trapping sessions. Captured *Peromyscus* mice were marked with a metal ear tag to identify individuals at subsequent recaptures. Only one trapping session (July 2019) was conducted at the undeveloped landscape sites. For each captured *Peromyscus*, a fecal sample was collected and body mass, sex, and reproductive status were recorded (reproductive individuals identified by the presence of scrotal testes for males or any of the following traits for females: perforate vagina, enlarged nipples, palpable embryos). Individuals captured a second time within a trapping session were not resampled and were promptly released at the point of capture.

All rodent trapping and handling methods were reviewed and approved by the University of Minnesota Institutional Animal Care and Use Committee (protocol no. 1903-36892A). The objective of this study was live-capture and release but trap fatalities (3.4% [16/477] of capture events of target and non-target species) were collected with approval by the Minnesota Department of Natural Resources (MN-DNR) under Special Permit No. 28440 and were accessioned with the Bell Museum of Natural History collections.

2.3 DNA Extraction

We collected 176 fecal samples representing 153 unique individuals. Fecal samples of up to 250 mg were stored without buffer or ethanol and frozen at -80°C immediately after sampling until DNA extraction. DNA was extracted using a QIAamp PowerFecal Pro kit (Qiagen, Hilden, Germany) following manufacturer instructions both manually and using a QIAcube robotic

workstation (Qiagen, Hilden, Germany). DNA extracts were quantified using a Qubit 4 fluorometer (Thermofisher Scientific, Waltham, MA, USA) using the Qubit dsDNA BR Assay Kit (Thermofisher Scientific, Waltham, MA, USA) following the manufacturer's instructions. Samples with low DNA yield (<24 ng/µL, n=16) were excluded from downstream analysis.

2.4 Library Prep & Nanopore Sequencing

For the remaining 160 samples, the Rapid 16S Barcoding Kit (SQK-16S024 [utilizing 'Kit 9' chemistry]; Oxford Nanopore Technologies [ONT], Oxford, UK) was used to prepare barcoded amplicon libraries for sequencing, largely following the manufacturer's protocol (methods described in detail in Jahan et al., 2021). First, all fecal DNA extracts were diluted in nuclease-free water to a concentration of 10-30 ng/µl. The full-length bacterial 16S rRNA gene region (1.6kb) was amplified via PCR using specific primers and between 20-40 ng of DNA template, a long-range master mix (LongAmp Hot Start Taq, 2X; New England Biolabs, Ipswich, MA, USA), and sample-specific barcode identifier. PCR products were purified and prepared for sequencing through a series of magnetic bead wash steps (AMPure XP beads; Beckman Coulter Life Sciences, Indianapolis, IN, USA). Barcoded samples were pooled with ONT rapid sequencing adapter mixture into a final library for sequencing. Seven sequencing runs were performed with a total of 160 samples, including 24 (run 1, 2, 4), 23 (run 6), 22 (run 3, 5), and 21 (run 7) barcoded samples from individual mice. Libraries were sequenced on a FLO-MIN106 MinION flow cell utilizing R9 sequencing chemistry (Oxford Nanopore Technologies, Oxford, UK), run for 24 hours using the ONT MinKNOW GUI (v4.3.20; Oxford Nanopore Technologies, Oxford, UK).

2.5 Bioinformatic Pipeline

Raw Fast5 data from the sequencing runs were base-called using the ONT Guppy basecaller using the 'super accuracy' basecalling model (ONT configuration file:

dna_r9.4.1_450bps_sup.cfg). The barcoded samples were further de-multiplexed using the Guppy barcoder to identify reads as belonging to one of the 24 unique barcodes. Reads were quality filtered (Nanopore Q score ≥ 8, corresponds to 84.15% base call accuracy) and filtered for target length (full-length bacterial 16S region approx. 1600 bp in length) using NanoFilt (v2.8.0; De Coster et al., 2018). Only reads 1200-1800 bp in length were retained for onward analysis. Summary reports were generated using Nanoq (v0.9.0; Steinig & Coin, 2022).

Taxonomic abundance profiles were generated using Emu, an expectation-maximization algorithm designed specifically to account for the increased read length and error rate often associated with long-read data such as ONT-generated sequences (v3.4.4; Curry et al., 2022). Compared to conventional taxonomic identification algorithms, Emu is able to reduce the false positive rate of identification and accurately identify long reads to species level (Curry et al., 2022). Reads were mapped using the Emu default database settings: a combination of rrnDB v5.6 (Stoddard et al., 2015) and NCBI (National Center for Biotechnology Information) 16S RefSeq downloaded on 17 September 2020 (O'Leary et al., 2016). The output of Emu is an estimated abundance (read count) of each identified species in a given sample. Because read counts are estimated based on likelihood probabilities, outputted values are not necessarily integer counts. Values were rounded to the nearest integer for analysis.

2.6 Data Analysis

The full fecal microbiome was characterized at the sample level using measures of alpha and beta diversity (to quantify within-sample and between sample bacterial diversity, respectively). Alpha diversity indices included species richness, Shannon-Weiner diversity, Simpson diversity, and species evenness. Shannon diversity and Simpson diversity make different assumptions about species evenness and how it contributes to diversity: Shannon diversity assumes all species are present and are randomly sampled while Simpson diversity gives more weight to

common species. Calculating both indices can suggest how common or rare species may affect diversity estimates for different populations. Beta diversity was quantified using the Bray-Curtis dissimilarity index to compare bacterial microbiome community composition at the species level between all pairs of samples. As a subset of the full fecal microbiome, the presence of pathogenic bacteria (foodborne and zoonotic pathogens of concern for human, livestock, and domestic animal health) was quantified at the sample level, then grouped by landscape-habitat pairing.

Rodent sampling was conducted across three months (June, July, and August) in the agricultural landscape and 18 individuals were captured and sampled in multiple months. To control for non-independence between repeat samples of the same individuals, only one sample per mouse (n=140) was included in the alpha and beta diversity analyses. We chose to include only the July sample for all recaptured mice to avoid introducing variation based on sampling month (all recaptured animals were sampled in July, but not in June or August) and to better align with the undeveloped landscape sampling (which was only conducted in July). For the analysis of pathogenic bacteria species, all samples (n=160) were used.

For the analyses of alpha and beta diversity, all samples were rarefied to the number of reads of the least abundant sample using the ‘phyloseq’ R package (v1.38.0; McMurdie & Holmes, 2013). Alpha diversity indices (richness, Shannon, Simpson, and evenness) were estimated from the rarefied data using the ‘vegan’ R package (v2.6.4; Oksanen et al., 2022). We investigated whether alpha diversity was affected by landscape or habitat type by developing a linear regression model for species richness and Shannon diversity and a beta regression model for Simpson diversity and species evenness. In all models, the response variable was the alpha diversity index and the explanatory variables were landscape type (i.e. anthropogenic or undeveloped), habitat type (forest or synanthropic), the interaction of landscape and habitat

type, mouse sex, reproductive status (reproductive or non-reproductive), body mass, and sampling month (June, July, or August). Beta diversity was visualized using non-metric multidimensional scaling (NMDS) ordination performed on the rarefied data using the Bray-Curtis dissimilarity index in the 'vegan' package. NMDS was first performed with 2-dimensions (k) and the k value was iteratively increased until the stress value was below 0.1. Non-parametric statistical analyses were performed on the rarefied distance matrices using the 'adonis2', 'anosim', 'betadisper', and 'permutest' functions also in the 'vegan' package.

For the analysis of pathogenic bacteria, species-level abundances were not rarefied and the raw estimated read counts output by the Emu pipeline were used. A list of 209 putative pathogenic bacteria species was generated using the PHI-base pathogen database (Urban et al., 2020; accessed on 13 Feb. 2023, plant-specific pathogens removed); resources from the U.S. Centers for Disease Control and Prevention on 'foodborne germs and illnesses' (CDC, 2022); and foodborne and mastitis-causing pathogens screened for by Jahan *et al.* 2021 (Jahan et al., 2021; For full list of pathogens, see Table S1). The species-level read count abundance data from the sequenced samples was filtered for reads assigned to the pathogen species on this list. We thresholded read count per pathogen species to at least 50 reads and visualized the patterns of pathogen read count per mouse, grouped by landscape-habitat pairing.

All analyses were performed in R Statistical Software (v4.1.2; R Core Team, 2021).

3 RESULTS

3.1 Rodent Samples

160 fecal samples were sequenced, representing 140 unique *Peromyscus* mice. In the agricultural landscape, 50 and 39 total fecal samples were collected from forest and synanthropic habitats, respectively, across three months of rodent trapping (Figure 1-C).

Excluding recaptures, 40 and 29 unique mouse fecal samples were collected in forest and synanthropic habitats, respectively. In the undeveloped landscape, 31 and 40 unique mouse fecal samples were collected from forest and synanthropic habitats, respectively (Figure 1-D).

3.2 Nanopore Sequencing Summary

After quality filtering, over 32.7 million high quality reads were retained (mean Q score 12.8 ± 0.31 s.d; Q score of 12.8 corresponds to base call accuracy of 94.75%). The mean number of reads per sample was 204,772.4, though the number of reads per sample was highly variable (standard deviation: 82,970.5; range: 74,517-517,058 reads; Table 1).

The Emu algorithm identified 1212 unique bacterial species across the 160 fecal samples. The mean number of species per sample was 211 (standard deviation: ± 55.8; range: 82-367).

Rarefaction curves were plotted for all sequenced samples (n=160). The asymptotic nature of these curves suggest reasonable sequencing depth was achieved for all samples (Figure S1). To enable direct comparisons between samples for the alpha and beta diversity analyses, the samples used in the diversity analysis (n=140) were rarefied to the minimum number of reads per sample (74,517 reads) and species were selected without replacement to reach the desired number of reads. After rarefaction, 36 species were removed because they were no longer present in any sample after random subsampling.

3.3 Alpha Diversity

The interaction of landscape and habitat type had a moderate effect on observed species richness, Shannon diversity and Simpson diversity indices (all $p<0.05$; Table S2). The effect of the interaction of landscape and habitat type on species evenness was weaker and only marginally significant ($p=0.087$; Table S2), though there was a significant effect of landscape

alone (p=0.016; Table S2). Mean observed species richness, Shannon diversity, Simpson diversity, and species evenness were lower in the undeveloped-forest habitat compared to all other landscape-habitat pairings (Figure 2; Table S3). However, contrary to our hypotheses, there was no difference in species richness, diversity, or evenness between forest and synanthropic habitats in the agricultural landscape or between agricultural-synanthropic and undeveloped-synanthropic habitat. Reproductive status (reproductive or non-reproductive individual, as noted by external morphology) had a moderate effect on Shannon and Simpson diversity and species evenness (all $p<0.01$; Table S2). None of the other parameters tested (sex, body mass, sampling month) had an effect on any alpha diversity index.

3.4 Beta Diversity

Across the four landscape-habitat pairings, the microbiome communities of sampled mice were dominated by three phyla: Firmicutes, Proteobacteria, and Bacteroidetes (relative abundance ≥5%) though Melainabacteria (a candidate phylum related to Cyanobacteria, Di Rienzi et al., 2013) and Deferribacteres were observed at relative abundances ≥1% in some samples (Figure 3). Firmicutes was the dominant phyla in most samples (relative abundance 90.1% mean ± 11.1 s.d.) followed by Proteobacteria (16.8% ± 20.0) and Bacteriodetes (8.92% ± 3.07).

Bacterial microbiome community composition at the species level was compared between all pairs of samples using the Bray-Curtis dissimilarity index based on rarefied count data. A nonparametric analysis of similarities test (‘anosim’ function, ‘vegan’ R package) comparing dissimilarity indices between samples from the four landscape-habitat pairings suggested that the between-group dissimilarity in microbiome community composition was significantly greater than the within-group dissimilarity (p=0.001).

An NMDS ordination plot calculated based on Bray-Curtis dissimilarity indices showed a high degree of overlap between samples from the four landscape-habitat pairings (Figure 4). Samples from agricultural-synanthropic and undeveloped-forest habitat showed the greatest dissimilarity while samples from agricultural-forest and undeveloped-synanthropic habitat were more similar. The variability among samples was high, but an analysis of multivariate homogeneity of group dispersion ('betadisper' and 'permutest' functions, 'vegan' R package) by landscape-habitat pairing showed no significant difference in variance between the groups (permutation test, $p=0.96$), indicating that the differences in community composition were not only due to differences in sample variance.

A nonparametric PERMANOVA analysis was used to test the effects of landscape, habitat type, mouse sex, reproductive status, body mass, and sampling month on differences in microbiome community composition using the 'adonis2' function in the 'vegan' R package with the by="margin" option to determine the marginal effect of each parameter. There was a small but significant effect of landscape and habitat, suggesting that the microbiome of sampled mice was different between agricultural and undeveloped landscapes and between forest and synanthropic habitats (PERMANOVA, $R^2_{Landscape}=0.06$, $R^2_{Habitat}=0.04$, both $p=0.001$; Table S4). Mouse reproductive status and body mass also had small, but significant effects (both $p<0.05$). However, much of the variance in microbiome community composition was not explained by the modeled parameters (residual $R^2=0.85$).

3.5 Putative Pathogen Detection

The presence of putative pathogenic bacteria was investigated using raw read counts of all sequenced samples (n=160). Read counts from mice captured in more than one month in the agricultural landscape were pooled by bacterial species across fecal samples from a single mouse. Of the 209 putative pathogenic bacteria species screened for, 18 were identified in

sampled mice (read count ≥ 50). At the population level, putative pathogen species richness was higher in agricultural-forest and undeveloped-synanthropic habitat (13 species identified; Figure 5) compared to agricultural-synanthropic and undeveloped-forest habitat (7 species identified). However, at the individual level, putative pathogen species richness was higher in mice in the agricultural landscape (agricultural-forest: mean putative pathogen species/mouse 1.42 ± 1.17 s.d.; agricultural-synanthropic: 1.24 ± 1.06) compared to mice in the undeveloped landscape (undeveloped-forest: 0.42 ± 0.77; undeveloped-synanthropic: 0.83 ± 0.93).

Read counts of detected putative pathogens were similar across landscape-habitat pairings with many mice having low read counts (<200 reads), though the number of mice with high read counts (>500 reads) was greatest in the agricultural-forest habitat (Figure 5). Across all sampled mice, *Clostridioides difficile*, *Streptococcus sanguinis*, *Enterococcus gallinarum*, *Citrobacter freundii*, and *Morganella morganii* were the most frequently detected putative pathogens (Figure 5).

4 DISCUSSION

Our objective was to characterize and compare the microbiome of synanthropic rodents and the abundance of zoonotic bacterial pathogens in agricultural landscapes and synanthropic habitat in Minnesota. We found that landscape-habitat pairing affected microbiome richness and diversity but species evenness was only affected by landscape. Overall, undeveloped-forest habitat had lower mean alpha diversity (richness, Shannon and Simpson diversity, evenness) than the other three landscape-habitat pairings. Microbiome community composition at the species level was also significantly different between landscapes (agricultural versus undeveloped) and habitat types (forest vs. synanthropic). We detected reads for a number of putative pathogenic bacteria across the four habitats, mostly at low read counts. The mean

number of putative pathogenic bacteria detected per mouse was higher in the agricultural landscape than the undeveloped.

Across landscape-habitat pairings, the microbiome of sampled mice was dominated by three phyla (Firmicutes, Bacteroidetes, Proteobacteria). These phyla are typical of the gut microbiome of wild *Peromyscus*, though we observed higher levels of Firmicutes and lower levels of Bacteroidetes compared to previous studies (e.g. Diaz et al., 2023; Schmidt et al., 2019). This suggests that the core fecal microbiome of the mice in our study is similar to *Peromyscus maniculatus* in other regions of North America. Only one other study has compared microbiome communities of free living *Peromyscus* in developed and undeveloped habitats (Diaz et al., 2023). We found lower richness and alpha diversity in the undeveloped-forest habitat compared to all other habitats, conversely, Diaz *et al.* found lower mean richness and Shannon diversity in urban habitats compared to undeveloped habitats. However, the directionality of alpha diversity shifts between undeveloped and developed populations is likely affected by multiple species- and system-specific factors; research in other wildlife systems has documented an increase in alpha diversity between undeveloped and developed populations (Dillard et al., 2022). Despite the differences in the direction of alpha diversity shifts, our finding that the microbiome community composition (beta diversity) between mice from undeveloped and agricultural developed landscapes was significantly different aligned with the findings of Diaz *et al.* These shifts in microbiome composition could be attributed to dietary shifts based on habitat type and food availability, particularly in synanthropic environments (Anders et al., 2022). In future studies, stable isotope analysis similar to those conducted by Anders *et al*. could provide additional insights into the diet of synanthropic and forest mice. Such information would likely inform the microbiome composition observed in our data, as the PERMANOVA modeling approach utilized herein indicated a high degree of unexplained microbiome composition variability that was not accounted for by landscape or habitat type.

We detected 16S sequences of a number of putative pathogenic bacteria in samples from all four landscape-habitat pairings. The greatest number of mice carrying putative pathogenic bacteria and the highest mean diversity of putative pathogen species per mouse was found in the anthropogenic-forest habitat while the lowest was found in the undeveloped-forest habitat. These differences are likely explained by the landscape surrounding our sampling locations which could represent a source of infection for many of these pathogens. The forest sampling sites in the agricultural landscape were located on the periphery of a research reserve which is surrounded by crop fields, pastures, and low-density housing. By contrast, the forest sites in the undeveloped landscape were contained in a state park and the forest continues uninterrupted beyond the park boundary with little agricultural development, limiting sources of pathogen exposure. *Peromyscus* are known to forage in crop fields as well as forest habitat, so it is likely that the abundance of putative pathogens in mice in the anthropogenic-forest habitat are representative of exposure to the surrounding agricultural landscape. Indeed, *Clostridioides difficile* was the most frequently detected putative pathogenic bacteria in the agricultural landscape, aligning with literature documenting this bacteria in many species of livestock and wildlife, including antimicrobial resistant strains in urban rodents and those living on or near farms (reviewed in Weese, 2020). In agricultural settings, manure used as fertilizer may serve as a source of environmental contamination for *C. difficile* (Frentrup et al., 2021) which could provide a transmission route to rodents and other wildlife. Contrary to our predictions, the mean number of putative pathogenic bacteria per mouse was similar between forest and synanthropic habitat within a landscape, suggesting similar levels of pathogen exposure for mice between these two habitats. The synanthropic habitats sampled were all at the interface of forest and human-habitated areas. It is possible the synanthropic mice only occasionally visit the human structures where they were trapped (maintenance garages and storage areas, cabins and tent platforms, etc.) and predominantly reside in the nearby forest. Frequent movement of mice

between native vegetation and synanthropic habitat could account for similar putative pathogen exposure within a landscape type.

Accurate detection and taxonomic assignment of reads is a key assumption for community diversity and metagenomic analyses. Species richness and diversity estimates can be sensitive to the presence of rare species. The Emu algorithm has a built-in abundance threshold of 10 reads for large samples (over 1,000 reads) to control against long tails of low-abundance species which are an artifact of the probabilistic expectation-maximization model (Curry et al., 2022). As a result, Emu has a limited ability to detect rare species and thus our estimates of species richness and diversity are likely underestimations of the true community composition. However, Emu's strength is that it was specifically designed for taxonomic identification of long-read sequence data. The Emu pipeline helps to correct errors and improve the accuracy of Nanopore 16S amplicon sequencing through the expectation-maximization algorithm and has been shown to outperform algorithms designed for short-read (i.e. Illumina) data when classifying 16S Nanopore sequences (Curry et al., 2022). Because we were most interested in the species-level identification of reads for the detection of putative pathogenic bacteria, we chose to prioritize accurate taxonomic assignment over the ability to detect rare species and more accurately estimate species richness and diversity. Furthermore, Nanopore sequencing provides a key advancement over short-read microbiome sequencing in that species-level identification is possible and accurate. In future research, we see great utility for taxonomic assignment algorithms like Emu designed specifically for long-read Nanopore sequences and expect these novel methods to continue to improve the ability to accurately characterize and study species-level microbiome composition. Indeed, already the Nanopore 'Kit 12' chemistry and R10 flow cells (released in late 2021) are able to outperform Illumina sequencing with less noise and higher accuracy, specifically for species-level classification of 16S amplicon sequencing of gut microbiota (Szoboszlay et al., 2023).

It is important to clarify that, while we can be confident in accurate taxonomic assignment of the bacterial species detected in the sampled mice, their presence does not guarantee zoonotic potential. Many of these bacteria are commensal in the human and mammalian gut and may only be opportunistic pathogens or only certain serotypes possess virulence factors capable of infecting humans. Determining pathogenicity requires more in-depth genotyping or lab cultures that were outside the scope of this research. Nonetheless, our detection of these bacteria species serves to inform the potential of *Peromyscus* mice to be reservoirs for zoonotic pathogens and can inform future studies that characterize the pathogenicity of these bacteria.

Our research supports and expands upon previous work done in Minnesota using Nanopore sequencing to identify pathogenic bacteria in synanthropic rodents. Jahan *et al.* pointed to the role that farms play in the increased abundance of putative pathogenic bacteria in synanthropic rodents (Jahan et al., 2021). However, farms are a unique anthropogenic environment with many routes of pathogen introduction, and rodents at this interface may not be representative of synanthropic rodents more broadly. Our work expands upon the foundation set by Jahan *et al.* by investigating less disturbed environments to understand the abundance and diversity of zoonotic bacterial pathogens in undeveloped and agricultural (cropland) landscapes. The diversity of putative pathogenic genera found in *Peromyscus* mice generally align between our studies: Jahan *et al.* similarly identified putative pathogenic genera including *Bacillus, Clostridium*, *Enterococcus*, and *Streptococcus* circulating in synanthropic rodents on Minnesota farms. However, we identified a higher abundance of *Clostridioides* and no pathogenic species of *Helicobacter* in our study. It is possible that these differences can be attributed to differences in how pathogen abundance was quantified: Jahan *et al*. reported abundance of reads identified at the genus level (summed across all sampled *Peromyscus*) whereas we focused on read abundance of specific pathogenic species per individual mouse. Interestingly, Jahan *et al.* found

lower abundance of putative pathogenic genera in *Peromyscus* mice compared to other rodent species trapped on farms including *Mus musculus*, *Microtus pennsylvanicus*, and *Rattus norvegicus*. While our study did not include other rodent species, the limited abundance of putative pathogenic bacteria found in *Peromyscus* herein corroborates the findings of Jahan *et al.* and could indicate lower exposure for these mice compared to other synanthropic rodents.

Overall, we found that *Peromyscus* in undeveloped and agricultural landscapes in Minnesota carried low abundance and diversity of putative pathogenic bacteria (we detected, on average, 1-2 putative pathogens per mouse and zero putative pathogens in many mice). Further, many of these were opportunistic pathogens which may pose a limited risk to zoonotic transmission in the human population. Our findings suggest that agricultural landscapes play a role in increasing the abundance of zoonotic pathogens in wild rodents; however, synanthropic habitat may be less informative of the abundance of zoonotic bacterial pathogens, particularly in environments where mice are expected to be highly mobile across interfaces between native vegetation and synanthropic areas. Taken together, our research suggests that *Peromyscus* are occasional hosts of zoonotic bacterial pathogens when sources of exposure are high (i.e. agricultural settings like crop fields and farms) but their flexibility to thrive in natural vegetation as well synanthropic habitat may act as a buffer to higher levels of zoonotic pathogen abundance.

5 CONCLUSIONS

The data presented herein provide a glimpse into the gut microbiome of *Peromyscus* mice in diverse landscapes of Minnesota. By sampling from populations in agricultural and undeveloped landscapes and in forest and synanthropic habitat, we find that landscape and habitat are important factors influencing microbiome community composition in wild rodents. We also identify low abundance of putative pathogenic bacteria species in these populations and

suggest the role of agricultural landscapes in increasing rodent exposure to putative pathogens. Even where transmission risk seems low, infection in wildlife populations could represent sources of novel pathogenic strains, bridge hosts linking environmental contamination back to human or livestock infection, or vectors to translocate pathogens across the landscape. As such, this research underscores the importance of investigating zoonotic pathogen prevalence in synanthropic rodents and other wildlife to better characterize their potential as reservoirs and vectors for pathogen spillover at the human-wildlife interface.

ACKNOWLEDGEMENTS

We thank Dr. Forest Isbell, Dr. Caitlin Barale Potter, and the staff of the Cedar Creek Ecosystem Science Reserve; Dr. Jonathan Schilling and the staff at Itasca Biological Station and Laboratories and Itasca State Park for their coordination and support of this research. Rodent trapping at Cedar Creek was conducted under Research Project Proposal RP922, Experiment No. E331. Rodent trapping at Itasca State Park was conducted under approval of the MN-DNR Division of Parks and Trails Special Permit No. 201938. We thank Dr. Sharon Jansa for providing rodent sampling permits and field supplies and Dr. Laramie Lindsey, Lauren Agnew, and Abby Guthmann for assisting with field sampling. Suzanne Stone provided logistical assistance with the molecular lab work. Dr. Pat Schloss, Dr. Noelle Noyes, and Jasmine Veitch provided guidance on data analysis. The Minnesota Supercomputing Institute provided computational and data storage resources. Research funding was provided by: the Itasca Biological Station American Indian Fund student research internship; Alexander and Lydia Anderson Fellowship through the Graduate School at the University of Minnesota (UMN); the Zoological Society Fund & the Dayton Bell Museum Fund of the Bell Museum of Natural History; the Cedar Creek Graduate Fellows program; the Florence Rothman Research Fellowship through the Department of Ecology, Evolution, and Behavior at UMN; and the Graduate Student Grant-in-Aid of Research from the American Society of Mammalogists. JM was supported by

the National Science Foundation Graduate Research Fellowship Program under Award No. 2237827. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. EJK was supported by UMN startup funds awarded to PAL and SIW was supported by the Van Sloun Foundation.

REFERENCES

Anders, J. L., Mychajliw, A. M., Moustafa, M. A. M., Mohamed, W. M. A., Hayakawa, T., Nakao, R., & Koizumi, I. (2022). Dietary niche breadth influences the effects of urbanization on the gut microbiota of sympatric rodents. *Ecology and Evolution*, *12*(9), e9216. https://doi.org/10.1002/ece3.9216

Berlow, M., Phillips, J. N., & Derryberry, E. P. (2021). Effects of urbanization and landscape on gut microbiomes in white-crowned sparrows. *Microbial Ecology*, *81*(1), 253–266. https://doi.org/10.1007/s00248-020-01569-8

CDC. (2022, December 19). *Foodborne Illnesses and Germs*. Centers for Disease Control and Prevention. https://www.cdc.gov/foodsafety/foodborne-germs.html

Curry, K. D., Wang, Q., Nute, M. G., Tyshaieva, A., Reeves, E., Soriano, S., Wu, Q., Graeber, E., Finzer, P., Mendling, W., Savidge, T., Villapol, S., Dilthey, A., & Treangen, T. J. (2022). Emu: Species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nature Methods*, *19*(7), Article 7. https://doi.org/10.1038/s41592-022-01520-4

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Dewitz, J. (2021). *National Land Cover Database (NLCD) 2019 Products* [Data set]. U.S. Geological Survey. https://doi.org/10.5066/P9KZCM54

Di Rienzi, S. C., Sharon, I., Wrighton, K. C., Koren, O., Hug, L. A., Thomas, B. C., Goodrich, J. K., Bell, J. T., Spector, T. D., Banfield, J. F., & Ley, R. E. (2013). The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *ELife*, *2*, e01102. https://doi.org/10.7554/eLife.01102

Diaz, J., Redford, K. H., & Reese, A. T. (2023). Captive and urban environments are associated with distinct gut microbiota in deer mice (Peromyscus maniculatus). *Biology Letters*, *19*(3), 20220547. https://doi.org/10.1098/rsbl.2022.0547

Dillard, B. A., Chung, A. K., Gunderson, A. R., Campbell-Staton, S. C., & Moeller, A. H. (2022). Humanization of wildlife gut microbiota in urban environments. *ELife*, *11*, e76381. https://doi.org/10.7554/eLife.76381

Ecke, F., Han, B. A., Hörnfeldt, B., Khalil, H., Magnusson, M., Singh, N. J., & Ostfeld, R. S. (2022). Population fluctuations and synanthropy explain transmission risk in rodent-borne zoonoses. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-35273-7

Firth, C., Bhat, M., Firth, M. A., Williams, S. H., Frye, M. J., Simmonds, P., Conte, J. M., Ng, J., Garcia, J., Bhuva, N. P., Lee, B., Che, X., Quan, P.-L., & Lipkin, W. I. (2014). Detection of zoonotic pathogens and characterization of novel viruses carried by commensal

Rattus norvegicus in New York City. *MBio*, *5*(5), e01933-14. https://doi.org/10.1128/mBio.01933-14

Frentrup, M., Thiel, N., Junker, V., Behrens, W., Münch, S., Siller, P., Kabelitz, T., Faust, M., Indra, A., Baumgartner, S., Schepanski, K., Amon, T., Roesler, U., Funk, R., & Nübel, U. (2021). Agricultural fertilization with poultry manure results in persistent environmental contamination with the pathogen Clostridioides difficile. *Environmental Microbiology*, *23*(12), 7591–7602. https://doi.org/10.1111/1462-2920.15601

Funosas, G., Triadó-Margarit, X., Castro, F., Villafuerte, R., Delibes-Mateos, M., Rouco, C., & Casamayor, E. O. (2021). Individual fate and gut microbiome composition in the European wild rabbit (Oryctolagus cuniculus). *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-020-80782-4

Gibb, R., Redding, D. W., Chin, K. Q., Donnelly, C. A., Blackburn, T. M., Newbold, T., & Jones, K. E. (2020). Zoonotic host diversity increases in human-dominated ecosystems. *Nature*, *584*(7821), Article 7821. https://doi.org/10.1038/s41586-020-2562-8

Gottdenker, N. L., Streicker, D. G., Faust, C. L., & Carroll, C. R. (2014). Anthropogenic land use change and infectious diseases: A review of the evidence. *EcoHealth*, *11*(4), 619–632. https://doi.org/10.1007/s10393-014-0941-z

Han, B. A., Kramer, A. M., & Drake, J. M. (2016). Global patterns of zoonotic disease in mammals. *Trends in Parasitology*, *32*(7), 565–577. https://doi.org/10.1016/j.pt.2016.04.007

Han, B. A., Schmidt, J. P., Bowden, S. E., & Drake, J. M. (2015). Rodent reservoirs of future zoonotic diseases. *Proceedings of the National Academy of Sciences*, *112*(22), 7039–7044. https://doi.org/10.1073/pnas.1501598112

Hassell, J. M., Begon, M., Ward, M. J., & Fèvre, E. M. (2017). Urbanization and disease emergence: Dynamics at the wildlife–livestock–human interface. *Trends in Ecology & Evolution*, *32*(1), 55–67. https://doi.org/10.1016/j.tree.2016.09.012

Hernandez, S. M., Welch, C. N., Peters, V. E., Lipp, E. K., Curry, S., Yabsley, M. J., Sanchez, S., Presotto, A., Gerner-Smidt, P., Hise, K. B., Hammond, E., Kistler, W. M., Madden, M., Conway, A. L., Kwan, T., & Maurer, J. J. (2016). Urbanized white ibises (Eudocimus albus) as carriers of Salmonella enterica of significance to public health and wildlife. *PLOS ONE*, *11*(10), e0164402. https://doi.org/10.1371/journal.pone.0164402

Jahan, N. A., Lindsey, L. L., Kipp, E. J., Reinschmidt, A., Heins, B. J., Runck, A. M., & Larsen, P. A. (2021). Nanopore-based surveillance of zoonotic bacterial pathogens in farm-dwelling peridomestic rodents. *Pathogens*, *10*(9), Article 9. https://doi.org/10.3390/pathogens10091183

Kallio, E. R., Begon, M., Henttonen, H., Koskela, E., Mappes, T., Vaheri, A., & Vapalahti, O. (2009). Cyclic hantavirus epidemics in humans—Predicted by rodent host dynamics. *Epidemics*, *1*(2), 101–107. https://doi.org/10.1016/j.epidem.2009.03.002

Marchesi, J. R., Adams, D. H., Fava, F., Hermes, G. D. A., Hirschfield, G. M., Hold, G., Quraishi, M. N., Kinross, J., Smidt, H., Tuohy, K. M., Thomas, L. V., Zoetendal, E. G., & Hart, A. (2016). The gut microbiota and host health: A new clinical frontier. *Gut*, *65*(2), 330–339. https://doi.org/10.1136/gutjnl-2015-309990

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Mendoza, H., Rubio, A. V., García-Peña, G. E., Suzán, G., & Simonetti, J. A. (2019). Does land-use change increase the abundance of zoonotic reservoirs? Rodents say yes. *European Journal of Wildlife Research*, *66*(1), 6. https://doi.org/10.1007/s10344-019-1344-9

Monteiro, H. F., & Faciola, A. P. (2020). Ruminal acidosis, bacterial changes, and lipopolysaccharides. *Journal of Animal Science*, *98*(8), skaa248. https://doi.org/10.1093/jas/skaa248

Murray, M. H., Lankau, E. W., Kidd, A. D., Welch, C. N., Ellison, T., Adams, H. C., Lipp, E. K., & Hernandez, S. M. (2020). Gut microbiome shifts with urbanization and potentially facilitates a zoonotic pathogen in a wading bird. *PLOS ONE*, *15*(3), e0220926. https://doi.org/10.1371/journal.pone.0220926

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., & Stevens, M. H. H. (2022). *Vegan: Community Ecology Package. R package version 2.5-2.7. 2020.*

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/

Schluter, J., Peled, J. U., Taylor, B. P., Markey, K. A., Smith, M., Taur, Y., Niehus, R., Staffas, A., Dai, A., Fontana, E., Amoretti, L. A., Wright, R. J., Morjaria, S., Fenelus, M., Pessin, M. S., Chao, N. J., Lew, M., Bohannon, L., Bush, A., … Xavier, J. B. (2020). The gut microbiota is associated with immune cell dynamics in humans. *Nature*, *588*(7837), 303–307. https://doi.org/10.1038/s41586-020-2971-8

Schmidt, E., Mykytczuk, N., & Schulte-Hostedde, A. I. (2019). Effects of the captive and wild environment on diversity of the gut microbiome of deer mice (Peromyscus maniculatus). *The ISME Journal*, *13*(5), Article 5. https://doi.org/10.1038/s41396-019-0345-8

Steinig, E., & Coin, L. (2022). Nanoq: Ultra-fast quality control for nanopore reads. *Journal of Open Source Software*, *7*(69), 2991. https://doi.org/10.21105/joss.02991

Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K., & Schmidt, T. M. (2015). rrnDB: Improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Research*, *43*(D1), D593–D598. https://doi.org/10.1093/nar/gku1201

Suchodolski, J. S. (2022). Analysis of the gut microbiome in dogs and cats. *Veterinary Clinical Pathology*, *50*(S1), 6–17. https://doi.org/10.1111/vcp.13031

Szoboszlay, M., Schramm, L., Pinzauti, D., Scerri, J., Sandionigi, A., & Biazzo, M. (2023). Nanopore is preferable over Illumina for 16S amplicon sequencing of the gut microbiota when species-level taxonomic classification, accurate estimation of richness, or focus on rare taxa is required. *Microorganisms*, *11*(3), Article 3. https://doi.org/10.3390/microorganisms11030804

Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S. Y., De Silva, N., Martinez, M. C., Pedro, H., Yates, A. D., Hassani-Pak, K., & Hammond-Kosack, K. E. (2020). PHI-base: The pathogen–host interactions database. *Nucleic Acids Research*, *48*(D1), D613–D620. https://doi.org/10.1093/nar/gkz904

Weese, J. S. (2020). Clostridium (Clostridioides) difficile in animals. *Journal of Veterinary Diagnostic Investigation : Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *32*(2), 213–221. https://doi.org/10.1177/1040638719899081

Worsley-Tonks, K. E. L., Miller, E. A., Anchor, C. L., Bender, J. B., Gehrt, S. D., McKenzie, S. C., Singer, R. S., Johnson, T. J., & Craft, M. E. (2021). Importance of anthropogenic sources at shaping the antimicrobial resistance profile of a peri-urban mesocarnivore. *Science of The Total Environment*, *764*, 144166. https://doi.org/10.1016/j.scitotenv.2020.144166

DATA ACCESSIBILITY STATEMENT

Upon manuscript submission, all Nanopore sequence data will be uploaded to the National Center for Biotechnology Information Sequence Read Archive. Metadata associated with all samples will be made available on Dryad. The R code used for analysis will be archived on Zenodo.

BENEFIT-SHARING STATEMENT

This research was supported by the Itasca Biological Station American Indian Fund whose goal is to foster scientific growth and collaboration between local students from area high schools and the researchers at the Itasca Biological Station and Laboratories. A recent high school graduate (CCA) collaborated on the field research and is included as a co-author. As described above, all data resulting from this research have been publicly shared via appropriate research databases.

AUTHOR CONTRIBUTIONS

JM, MEC, and PAL conceived of and designed the research and acquired funding. JM, SIW, and CCA performed the field sampling. JM and EJK conducted sequencing. JM conducted the bioinformatic and statistical analyses with support from EJK and PAL. JM led the writing of the manuscript under the mentorship of MEC and PAL. All authors contributed to revisions and editing. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FIGURES & TABLES

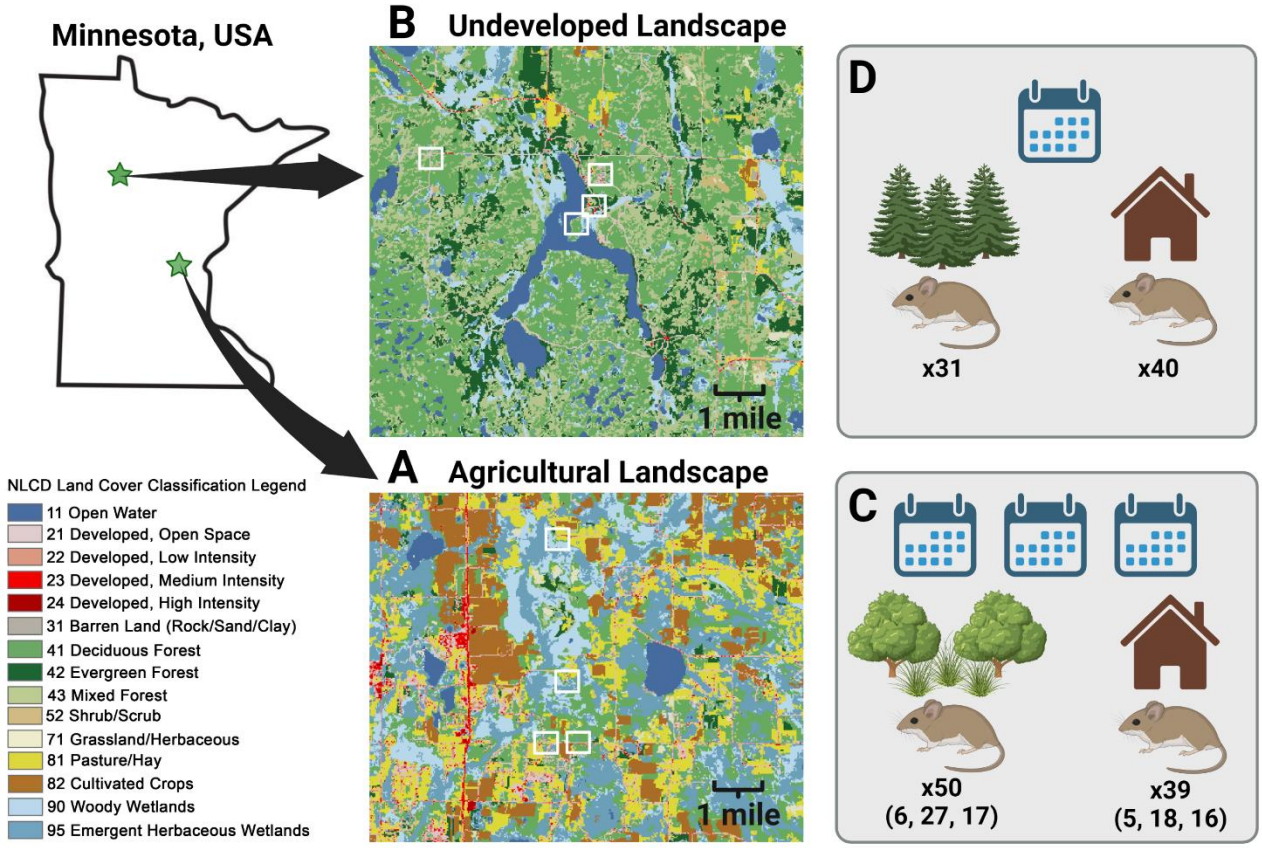


Figure 1. Rodent sampling locations and sample size summary. Sampling was conducted at two locations in Minnesota, USA representing undeveloped and agricultural landscapes. Study sites are outlined with white boxes (A, B). Sample size (total number of fecal samples) in forest and synanthropic habitat is shown for each landscape (C, D). Sampling was conducted once in the undeveloped landscape and three times in the agricultural landscape. Total number of samples per landscape-habitat pairing is noted first with samples per month in parentheses below (includes multiple samples from individual mice). Maps and land cover classification legend from National Land Class Database (NLCD) 2019 (Dewitz, 2021). Figure created with BioRender.com.

Table 1. Summary statistics of 16S Nanopore sequencing data of mouse fecal sample DNA (after quality filtering) by landscape, habitat type, and sampling month. Mean and standard deviation are reported for number (N) of reads per sample (reported in units of thousands of reads), number of basepairs per sample (reported in units of millions of basepairs [Mb]), and read quality (Q) score. Individual sampling months in the agricultural landscape shown in italics, rows shaded in gray. Mean values across all three months shown in bold. Number of samples represents total number of fecal samples sequenced (includes repeat sampling of unique mice).

Landscape	Habitat Type	Month	N samples	N reads/sample *(thousands of reads)*	N basepairs/sample *(Mb)*	Q Score
Agricultural	*Forest*	*June*	*6*	*272.18 ± 39.12*	*433.27 ± 61.68*	*13.08 ± 0.04*
Agricultural	*Forest*	*July*	*27*	*262.35 ± 62.91*	*418.69 ± 100.39*	*13.14 ± 0.06*
Agricultural	*Forest*	*August*	*17*	*248.61 ± 116.91*	*395.76 ± 186.39*	*12.84 ± 0.31*
Agricultural	**Forest**	**Summer**	**50**	**258.86 ± 82.36**	**412.64 ± 131.34**	**13.03 ± 0.23**
Agricultural	*Synanthropic*	*June*	*5*	*326.35 ± 73.84*	*517.32 ± 115.67*	*13.08 ± 0.08*
Agricultural	*Synanthropic*	*July*	*18*	*215.45 ± 21.66*	*342.53 ± 33.57*	*12.82 ± 0.04*
Agricultural	*Synanthropic*	*August*	*16*	*88.02 ± 10.85*	*139.68 ± 17.33*	*12.4 ± 0*
Agricultural	**Synanthropic**	**Summer**	**39**	**177.39 ± 88.31**	**281.72 ± 139.93**	**12.68 ± 0.25**
Undeveloped	**Forest**	**July**	**31**	**139.49 ± 22.44**	**223.04 ± 35.6**	**12.43 ± 0.22**
Undeveloped	**Synanthropic**	**July**	**40**	**214.45 ± 59.77**	**342.01 ± 95.9**	**12.88 ± 0.11**

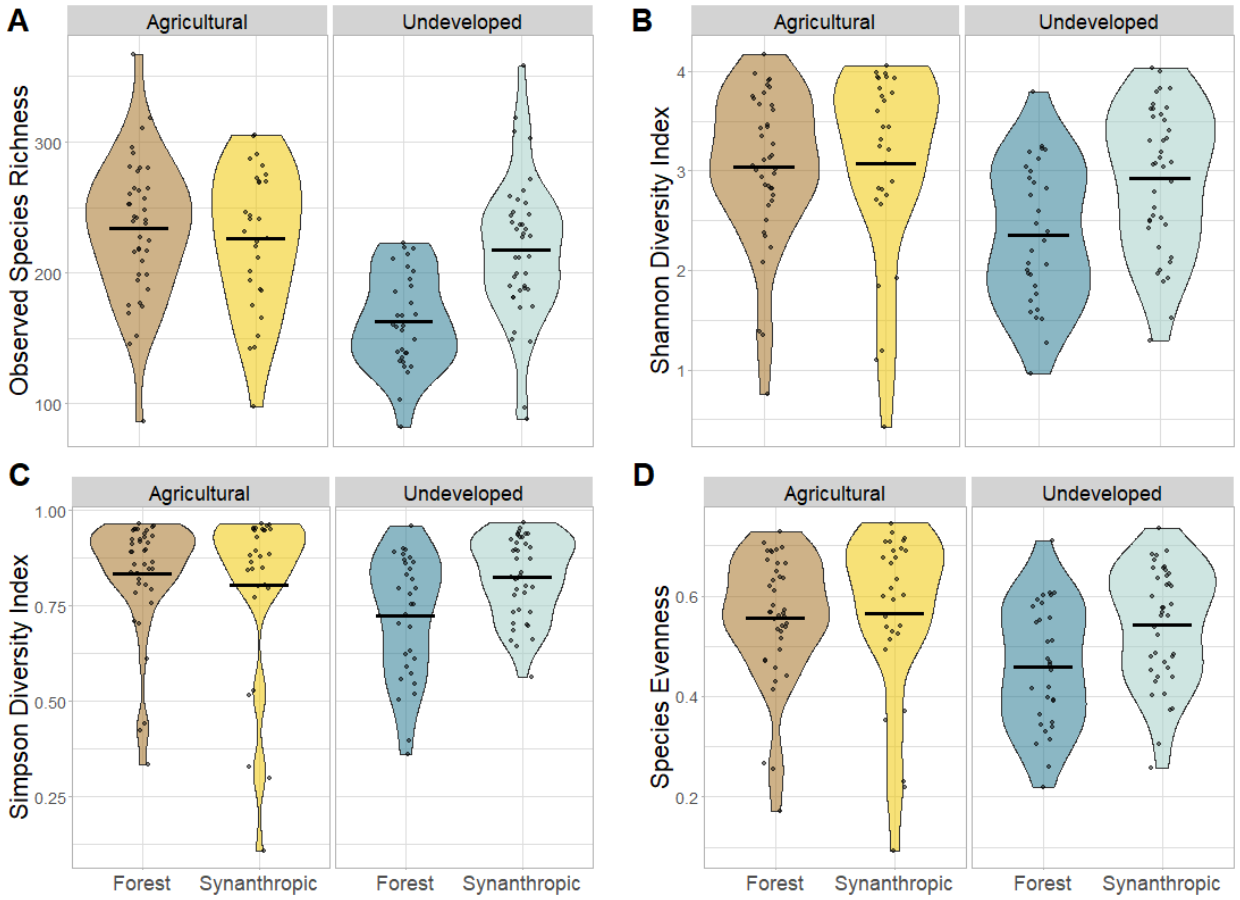
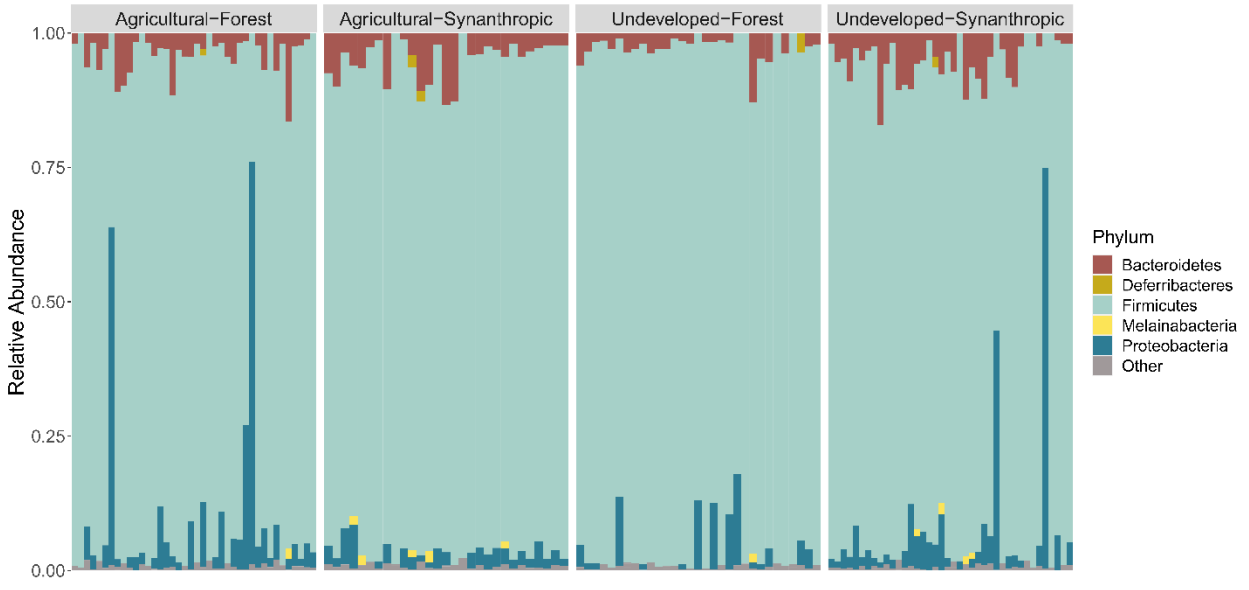
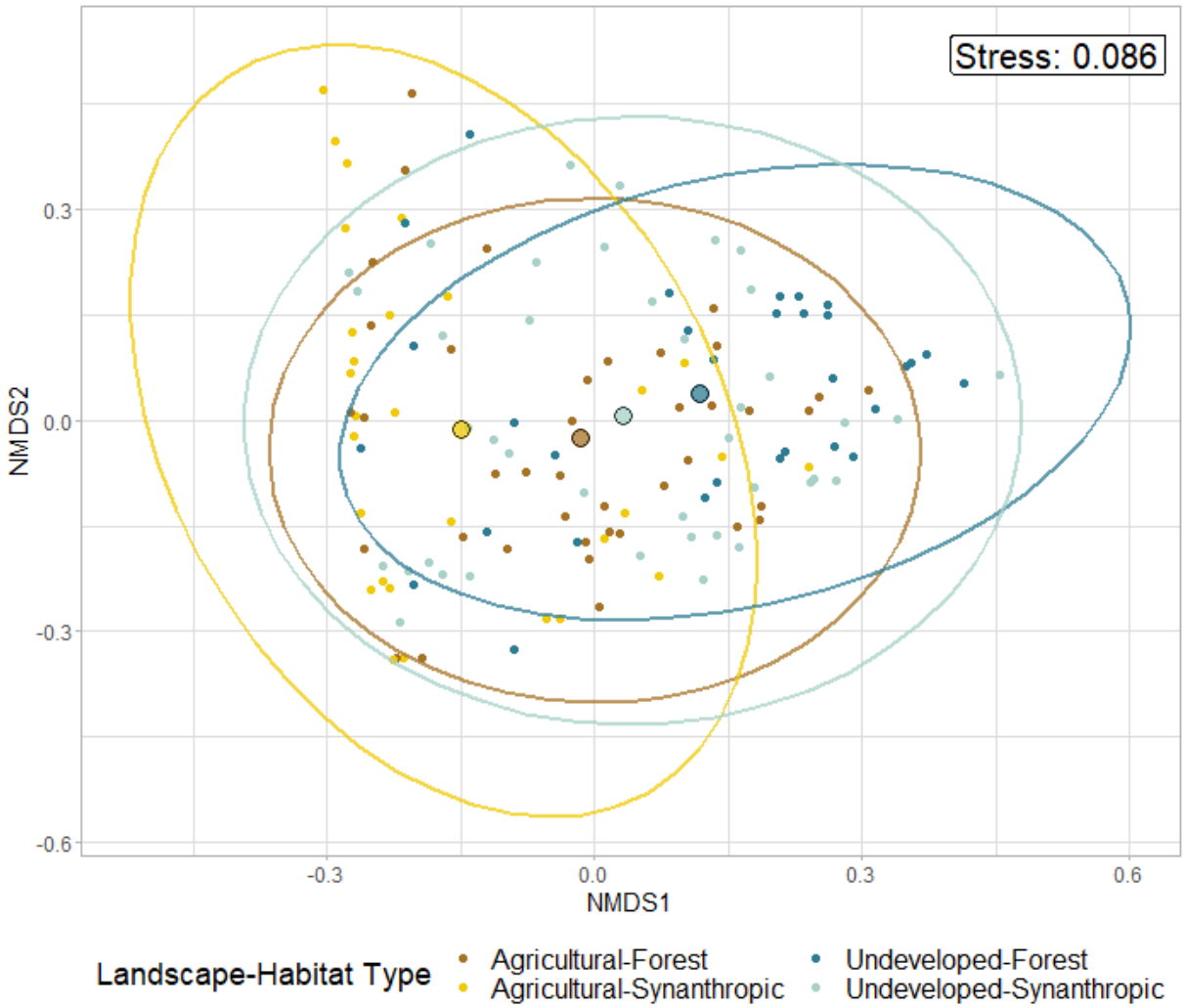


Figure 2. Alpha diversity for all unique mouse fecal samples (n=140) in anthropogenic and undeveloped landscapes and in forest and synanthropic habitat according to A) observed species richness B) Shannon diversity index C) Simpson diversity index and D) species evenness.



729

730 Figure 3. Relative abundance of bacteria phyla per sample (n=140) by landscape-habitat pairing
731 showing phyla present at ≥1% relative abundance. Phyla observed at <1% relative abundance
732 were grouped in a single category “Other”. The microbiome of sampled mice was dominated by
733 three phyla: Bacteroidetes, Firmicutes, and Proteobacteria.
734



735

736 Figure 4. Non-metric multidimensional scaling ordination on microbiome community composition
737 by Bray-Curtis dissimilarity index. Points represent individual samples, colored by landscape-
738 habitat pairing. Ellipses denote the 95% confidence level for a multivariate t-distribution of the
739 data points per group (centroids marked with larger points). Stress value: 0.086 (k=4).
740

741

742 Figure 5. Heatmap of read counts of putative pathogenic bacteria species per mouse in each
743 landscape-habitat pairing (count threshold >50 reads). The vertical axis represents samples
744 from an individual mouse. Warmer colors indicate higher read abundance (natural log scale).