University of Nevada, Reno

Diet related adaptations across a small mammal hybrid zone

A dissertation submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in
Ecology, Evolution, and Conservation Biology

by

Danny Nielsen

Dr. Marjorie D. Matocq/Dissertation Advisor

December 2022

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

Danny P. Nielsen

entitled

Diet related adaptations across a small mammal hybrid zone

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Marjorie D. Matocq
Advisor

Chris R. Feldman
Committee Member

Matthew L. Forister
Committee Member

Lora A. Richards
Committee Member

Bradley S. Ferguson
Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean
Graduate School

December, 2022

Abstract

Hybrid zones have long been used to investigate divergence and speciation. Many hybrid zones occur across sharp ecotones—areas characterized by transition in biological community composition. Such hybrid zones are often the result of secondary contact—when populations that descended from a common ancestral population come into contact after a period of allopatry. These populations may have accumulated differences via neutral process (*i.e.,* genetic drift) or adaptation to differing environments. In the absence of complete reproductive isolation, genes may then flow between these differentially adapted populations. Vegetation turnover is common across ecotones, and plant availability is important to mammalian herbivores that consume plants that often produce toxic plant secondary compounds (PSCs). Availability of diet plants for which mammalian herbivores are adapted may limit movement and underlie pre- and post-zygotic isolating mechanisms across sharp ecotones.

I studied diet and diet-related adaptions in a mammalian hybrid zone between two species of woodrat (*Neotoma*) that occurs across a sharp ecotone characterized by differences in plant community composition. Using live-trapping and field-based experiments, coupled with amplicon sequencing of DNA extracted from woodrat feces, I quantified variation in diet, diet preference, and gut microbiome composition between *N. lepida* (desert woodrat) and *N. bryanti* (Bryant's woodrat), and F1 and backcross hybrids. I found that each parental species maintains distinct diets that contain plants that produce toxic PSCs but these plants were also among the most nutritional across the site. Furthermore, these dietary differences were maintained across seasons and years that

spanned more wet to more dry periods. These diets were also associated with differences in microbiome composition, and while diet was primarily predicted by habitat, microbiome composition was constrained by genotype. I then used laboratory-based feeding experiments to determine how each species responds—physiologically, genetically, and behaviorally—to their native and non-native diet. Diet experiments were followed with 16S rRNA sequencing of contents from woodrat caecum, as well as RNA sequencing of tissue from the liver and caecum of woodrats. Response to diet treatments was asymmetrical, with *N. lepida* exhibiting greater response behaviorally, genetically, and in gut microbiome composition. Gene expression in liver was strongly influenced by species and exhibited little effect of diet treatment, but differential expression of genes in the caecum exhibited strong species by diet interaction effects. *Neotoma lepida* exhibited a strong diet effect in genes expressed in the caecum, as well as in differences in microbiota of the caecum. These results suggest that interactions between host genes and microbes contained in the caecum may play a role in the metabolism of plant PSCs. These field-based observations and laboratory-based experiments add to our understanding of how diet and diet-related adaptations may influence gene flow across this small mammal hybrid zone.

Dedication

To my mom and dad: Marty and Brent

Acknowledgements

Finishing a dissertation is a herculean task, and many people have been critical in helping me complete this work. Foremost, my advisor, Dr. Marjorie Matocq. I would have never had the opportunity to study these woodrats, and in such a beautiful desert landscape without her support. I would not have finished without her mentorship, guidance, and unyielding support in this endeavor. I also thank my committee: Drs. Chris Feldman, Matt Forister, Lora Richards, and Brad Ferguson. Whenever I had questions about design, analysis, or theoretical details, I could count on each of these inspiring scientists.

I also owe a big thank you to the many great field and lab technicians over the years. None of this would have been finished without their help. Dan Moore and Alex Mitchell spent countless hours in the desert trapping woodrats with me. I spent many days and nights with Jason Wurtz in the field. His encyclopedic knowledge ranges from geomorphology to satellites, and my time in the desert was enriched in his company. Other former field and lab techs helped this come to reality: Michael Jee, Cameron Demaranville, Mara MacKinnon, Sydney Stephens, and Karli Graski. I also thank Mary Whitfield of the Southern Sierra Research Station, and Reed Tollefson of the Kern River Audubon Preserve for support in the field.

I am also grateful for the EECB community. I have made friends with many current and former students of this program, and have found not just community, but a family of great scientific minds. I thank my lab mates Casey Adkins, Ally Coconis, and Miranda Crowell, for their support and encouragement over the years. I enjoyed many

odd hours with office mates over the years: Annie Espeset, Katie Uckele, Cynthia Scholl, Santiago Villamarín-Cortez. Many guitar-noodling sessions with Jake Francis and Thomas Riecke. Great times in the mountains with Valentina Alaasam, Anson Call, Chris Halsch, Trevor Faske, and Anna Tatarko. And a special thanks to Josh Jahner. I could show up in his office with questions, or just to chat, and he always made time.

Finally, my family. To Sunny, Skylar, and Harper. Thank you for giving me meaning and for your support through this long journey. To my mom, Marty, and my dad, Brent—thank you for encouraging my imagination and curiosity of the natural world. Thanks to my brothers, Peter, Ben and Greg. I count myself lucky to have brothers that I can call any time. Even through our rowdy younger years, and through many fights, I consider them my best friends today.

Table of Contents

Abstract …………………………………………………………………………………....i

Dedication……………………………………………………………………………...iii

Acknowledgements………………………………………………………………….....iv

Table of Contents……………………………………………………………………....vi

List of tables…………………………………………………………………………..vii

List of figures…………………………………………………………………...….viii

Introduction…………………………………………………………………………...1

Chapter 1: Differences in dietary composition and preference maintained despite gene flow across a woodrat hybrid zone……………………………………………………..8

Chapter 2: The gut microbiome reflects ancestry despite dietary shifts across a hybrid zone…….……………………………………………………………………...…....61

Chapter 3: Differential response of specialist and generalist herbivores when switching diets: host gene expression, the gut microbiome, and their potential interaction….........122

Concluding remarks…………………………………………………………….....172

List of tables

Table 1-1……………………………………………………………..…………..39

Table 1-2……………………………………………………………..…………..40

Table 1-3……………………………………………………………..…………..41

Table 2-1……………………………………………………………..…………..91

Table 3-1……………………………………………………………..………....160

Table 3-2……………………………………………………………..………....161

List of figures

Figure 1-1…………………………………………………………………..…….42

Figure 1-2…………………………………………………………………..…….43

Figure 1-3…………………………………………………………………..…….44

Figure 2-1…………………………………………………………………..…….92

Figure 2-2…………………………………………………………………..…….93

Figure 2-3…………………………………………………………………..…….94

Figure 2-4…………………………………………………………………..…….95

Figure 2-5…………………………………………………………………..…….96

Figure 3-1………………………………………………………………..……...155

Figure 3-2………………………………………………………………..……...156

Figure 3-3………………………………………………………………..……...157

Figure 3-4………………………………………………………………..……...158

Figure 3-5………………………………………………………………..……...159

Introduction

Hybrid zones provide "natural laboratories" in which to study the mechanisms underlying the generation and maintenance of biodiversity (Hewitt 1988). Secondary contact occurs when populations, descended from a common ancestral population, come into contact after a period of allopatry (Moore 1977, Barton and Hewitt 1985). If reproductive isolation is incomplete, genes may be exchanged between these independently evolved populations. Many vertebrate hybrid zones are narrow and occur across sharp ecotones (Moore 1977), an area characterized by a transition in biological communities (Kark and Rensburg 2006). Sharp turnover in vegetation composition is one common characteristic of ecotones (Walker et al. 2003). For mammalian herbivores, spatial turnover in available food plants may be an important source of selection due to the need to balance nutrient acquisition and exposure to potentially toxic plant secondary compounds in unfamiliar food plants (PSCs; Freeland and Janzen 1975). Food plant availability, then, may influence gene flow across ecotones by limiting dispersal and reducing fitness of migrants (*i.e.,* prezygotic isolating mechanism), or by reducing fitness of hybrids (*i.e.*, postzygotic isolating mechanism), or both (Via 1999, Via et al. 2000, Nosil et al. 2005).

We have studied a naturally occurring mammalian hybrid zone that occurs in southern California at an ecotone between the Mojave Desert and the Sierra Nevada Mountain range. At this site (hereafter referred to as Whitney Well), *Neotoma lepida* (desert woodrat) and *N. bryanti* (Bryant's woodrat) hybridize and produce a spectrum of F1 and backcross hybrids (Shurtliff et al. 2014, Jahner et al. 2021). The two species are

estimated to have diverged ~1.6 mya before coming into secondary contact at this location (Patton et al. 2007). *Neotoma lepida* primarily occupy the xeric Mojave Desert (hereafter referred to as "flats") and *N. bryanti* primarily inhabit the relatively mesic Sierra Nevada foothill (hereafter referred to as "hill"). Backcross hybrids are primarily located in their 'parental-specific' habitat (i.e., BC-*lepida* in flats, BC-*bryanti* on hill), and F1 hybrids are distributed evenly across this hybrid zone. Even though parental and backcross individuals are largely ecologically segregated, approximately 14% of individuals occupy the alternative habitat (hereafter referred to as 'mismatched' individuals). Although relatively rare, these mismatched individuals are an important source of interspecific contact and hybridization across this ecotone and provide an opportunity to explore the role of dietary plasticity and selection across this sharp environmental gradient.

In chapter 1 (Nielsen and Matocq 2020), we investigate patterns of dietary differentiation in the field across relatively wet (*i.e.,* spring) and dry years (*i.e.*, summer). Using cafeteria style choice trials, we measured dietary preference for the most common plants in the field diets of parental *N. lepida* living in the flats and *N. bryanti* living on the hill. This chapter confirms that diet composition is distinct between the species in their native habitats, and that these dietary differences are maintained across seasons despite seasonal changes in plant availability. We also found that each species exhibited preference for the plant most common in their native habitat (*N. lepida*- flats, *N. bryanti*- hill). Furthermore, the primary plants consumed are both potentially toxic, but also may provide greater nutritional benefit than other widely available plants. In combination, our field measures of seasonal and annual diet plasticity and preference trials support the

classification of *N. bryanti* as a facultative generalist and *N. lepida* as a facultative specialist.

There is still a gap in our understanding of the relative influences of host and environmental factors in shaping gut microbiome composition. For mammalian herbivores, the gut microbiome plays an important role in the acquisition of nutrients (Dearing and Kohl 2017) and has been found to facilitate the metabolism of toxic plant PSCs (Kohl et al. 2014). In chapter 2 (Nielsen et al. 2022), we investigate the relationship between diet and gut microbiome composition across this hybrid zone. We collected fresh fecal pellets across multiple season and years and from individuals that span genomic spectrum that characterizes the hybrid zone. We found that both diet and microbiome varied among genotypic classes, as well as across habitats. However, diet was most constrained by habitat while microbiome was best predicted by genotype. In addition, F1 hybrids exhibited intermediate microbiome composition, while a species signal was evident among backcross hybrids. While these findings are largely observational, we identified several bacterial lineages that may contribute otherwise unavailable metabolic pathways for plants PSCs. While the gut microbiome is a promising avenue of research into its role in metabolism of PSCs, woodrats in the genus *Neotoma* have also evolved a variety of genetic mechanisms for the detoxification of toxic plant compounds—a topic we explore in Chapter 3.

Woodrats in the genus *Neotoma* are known to consume a variety of chemically defended plants through their range including creosote bush (*Larrea tridentata;* Mangione et al. 2000) and juniper (*Juniperus monosperma*; Skopec et al. 2007). Detoxification enzymes in the family cytochrome P450 have been found to play an

important role in metabolism of plant toxins within woodrats (Malenke et al. 2012). This gene family is highly diverse and may be central to dietary adaptations in the genus *Neotoma* throughout its range. More recent work has identified substantial gene copy number expansion and variation in xenobiotic metabolizing genes among species of *Neotoma* (Greenhalgh et al. 2022; Holding et al. in prep). In chapter 3, we conduct lab-based feeding experiments to better understand the physiological and genetic response to the natural and toxic diets observed at Whitney Well. By using a reciprocal cross design, we investigated how parental *N. lepida* and *N. bryanti* respond to both their native and non-native diet types. Through daily observations, we quantified response to diet treatments such as water consumption, food intake, locomotion, and change in body mass. At the conclusion of each trial, we measured metabolic rate, and performed dissections to harvest liver and GI tract tissues for both RNA sequencing, as well as gut microbiome composition in the caecum. We found that during-trial response (*i.e.*, maximum tolerable dose, water consumption, wheel running), as well as microbial and genetic responses to diet treatment were asymmetrical, with *N. lepida*, the relative dietary specialist, exhibiting stronger diet effects. Our results also revealed differential expression of genes that may play a role in the metabolism of plant PSCs, but also in modifying feeding behavior and interactions with the microbiome.

Overall, this work suggests that *N. lepida* and *N. bryanti* maintain distinct dietary adaptations that span from behavior to physiology, that in turn influences gene flow across this species boundary.

Literature Cited

Barton, N. H., & Hewitt, G. M. (1985). Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, *16*, 113–148.

Dearing, M. D., & Kohl, K. D. (2017). Beyond fermentation: Other important services provided to endothermic herbivores by their gut microbiota. *Integrative and Comparative Biology*, *57*(4), 723–731.

Dearing, M. D., Mangione, A. M., & Karasov, W. H. (2000). Diet breadth of mammalian herbivores: Nutrient versus detoxification constraints. *Oecologia*, *123*(3), 397–405. https://doi.org/10.1007/s004420051027

Freeland, W. J., & Janzen, D. H. (1974). Strategies in Herbivory by Mammals: The Role of Plant Secondary Compounds. *The American Naturalist*, *108*(961), 269–289.

Greenhalgh, R., Holding, M. L., Orr, T. J., Henderson, J. B., Parchman, T. L., Matocq, M. D., Shapiro, M. D., & Dearing, M. D. (n.d.). Trio-binned genomes of the woodrats Neotoma bryanti and Neotoma lepida reveal novel gene islands and rapid copy number evolution of xenobiotic metabolizing genes. *Molecular Ecology Resources*, *n/a*(n/a). https://doi.org/10.1111/1755-0998.13650

Hewitt, G. M. (1988). Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, *3*(7), 158–167. https://doi.org/10.1016/0169-5347(88)90033-X

Jahner, J. P., Parchman, T. L., & Matocq, M. D. (n.d.). Multigenerational backcrossing and introgression between two woodrat species at an abrupt ecological transition. *Molecular Ecology*, *n/a*(n/a). https://doi.org/10.1111/mec.16056

Kark, S., & van Rensburg, B. J. (2006). Ecotones: Marginal or Central Areas of Transition? *Israel Journal of Ecology & Evolution*, *52*(1), 29–53. https://doi.org/10.1560/IJEE.52.1.29

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Dearing, M. D. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology Letters*, *17*(10), 1238–1246. https://doi.org/10.1111/ele.12329

Malenke, J. R., Magnanou, E., Thomas, K., & Dearing, M. D. (2012). Cytochrome P450 2B Diversity and Dietary Novelty in the Herbivorous, Desert Woodrat (Neotoma lepida). *PLOS ONE*, *7*(8), e41510. https://doi.org/10.1371/journal.pone.0041510

Mangione, A. M., Dearing, M. D., & Karasov, W. H. (2000). Interpopulation Differences in Tolerance to Creosote Bush Resin in Desert Woodrats (neotoma Lepida). *Ecology*, *81*(8), 2067–2076. https://doi.org/10.1890/0012-9658(2000)081[2067:IDITTC]2.0.CO;2

Moore, W. S. (1977). An Evaluation of Narrow Hybrid Zones in Vertebrates. *The Quarterly Review of Biology*, *52*(3), 263–277.

Nosil, P., Vines, T. H., & Funk, D. J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, *59*(4), 705–719.

Patton, J. L., Huckaby, D. G., & Álvarez-Castañeda, S. T. (2007). *The evolutionary history and a systematic revision of woodrats of the Neotoma lepida group* (Vol. 135). Univ of California Press.

Shurtliff, Q. R., Murphy, P. J., & Matocq, M. D. (2014). Ecological segregation in a small mammal hybrid zone: Habitat-specific mating opportunities and selection

against hybrids restrict gene flow on a fine spatial scale. *Evolution*, *68*(3), 729–742.

Skopec, M. M., Haley, S., & Dearing, M. D. (2007). Differential hepatic gene expression of a dietary specialist (Neotoma stephensi) and generalist (Neotoma albigula) in response to juniper (Juniperus monosperma) ingestion. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *2*(1), 34–43. https://doi.org/10.1016/j.cbd.2006.11.001

Via, S. (1999). Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, *53*(5), 1446–1457.

Via, S., Bouck, A. C., & Skillman, S. (2000). Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution*, *54*(5), 1626–1637.

Walker, S., Wilson, J. B., Steel, J. B., Rapson, G. L., Smith, B., King, W. M., & Cottam, Y. H. (2003). Properties of ecotones: evidence from five ecotones objectively determined from a coastal vegetation gradient. *Journal of Vegetation Science*, *14*(4), 579-590.

Chapter 1: Differences in dietary composition and preference maintained despite gene flow across a woodrat hybrid zone

Danny P. Nielsen[1,2] | Marjorie D. Matocq[1,2]

[1]Department of Natural Resources and Environmental Science; University of Nevada, Reno, Nevada

[2]Graduate Program in EECB, University of Nevada, Reno, Nevada

Citation: Nielsen, D. P., & Matocq, M. D. (2021). Differences in dietary composition and preference maintained despite gene flow across a woodrat hybrid zone. *Ecology and evolution*, 11(9), 4909-4919.

Abstract

1. Ecotones, characterized by adjacent yet distinct biotic communities, provide natural laboratories in which to investigate how environmental selection influences the ecology and evolution of organisms. For wild herbivores, differential plant availability across sharp ecotones may be an important source of dietary-based selection.
2. We studied small herbivore diet composition across a sharp ecotone where two species of woodrat, *Neotoma bryanti* and *N. lepida,* come into secondary contact with one another and hybridize. We quantified woodrat dietary preference through trn*L* metabarcoding of field-collected fecal pellets and experimental choice trials. Despite gene flow, parental *N. bryanti* and *N. lepida* maintain distinct diets across this fine spatial scale, and across temporal scales that span both wet and dry conditions.
3. *Neotoma bryanti* maintained a more diverse diet, with *Frangula californica* (California coffeeberry) making up a large portion of its diet. *Neotoma lepida* maintains a less diverse diet, with *Prunus fasciculata* (desert almond) comprising more than half of its diet. Both *F. californica* and *P. fasciculata* are known to produce potentially toxic plant secondary compounds (PSCs), which should deter herbivory, yet these plants have relatively high nutritional value as measured by crude protein content.
4. *Neotoma bryanti* and *N. lepida* consumed *F. californica* and *P. fasciculata*, respectively, in greater abundance than these plants are available on the landscape – indicating dietary selection. Finally, experimental preference trials revealed that

N. bryanti exhibited a preference for *F. californica*, while *N. lepida* exhibited a relatively stronger preference for *P. fasciculata*. We find that *N. bryanti* exhibit a generalist herbivore strategy relative to *N. lepida*, which exhibit a more specialized feeding strategy in this study system.

5. Our results suggest that woodrats respond to fine-scale environmental differences in plant availability that may require different metabolic strategies in order to balance nutrient acquisition while minimizing exposure to potentially toxic PSCs.

Keywords: adaptation, detoxification, herbivore, hybridization, *Neotoma,* toxin tolerance, woodrat

Introduction

Ecotones are characterized by spatial transition in environmental variables that can create selective gradients that generate or maintain diversity (Smith et al. 1997). When sharp abiotic gradients support the establishment of spatially proximate but distinct vegetation communities (Walker et al. 2003), animals must respond to abrupt spatial transitions in abiotic and biotic resources. Such spatially proximate, yet dissimilar selective environments have the potential to generate or reveal the ecological adaptations or forms of phenotypic plasticity that permit species to exist in disparate environments (Ghalambor et al. 2007, West-Eberhard 1989).

At sharp environmental transitions, one of the primary challenges facing herbivores is the abrupt transition in food plant availability. For herbivores, space use and movement across ecotones is largely determined by the distribution of plants that allow acquisition of adequate nutrition, while minimizing exposure to toxic plant secondary compounds (PSCs; Dearing et al. 2000, Dearing et al. 2005, Westoby 1978, Freeland and Janzen 1974). Mammalian herbivores have evolved numerous behavioral and physiological adaptations to maximize nutrition while minimizing toxin exposure including regulation of liver detoxification enzymes (Malenke et al. 2012), decrease in metabolic rate and physical activity when exposed to dietary PSCs (Sorensen et al. 2005b), and maintenance of a microbiome that facilitates nutrient acquisition and detoxification (Kohl et al. 2014). Mammalian herbivores may also diversify their diets to minimize overexposure to, or neutralize, toxins present (Iason and Villalba 2006). Based on the degree to which mammalian herbivores modify their diets either spatially or

temporally, they can be classified along a continuum of foraging strategies from generalist to specialist consumers (Shipley et al. 2009).

When mammals consume toxic plants they are not adapted to, they suffer energetic consequences that can lead to rapid weight loss and lowered body condition (Sorensen et al. 2005a, Sorensen et al. 2005b, Mangione 2004). Given these consequences, we would expect mammalian herbivores to develop dietary preferences for plants with which they are familiar and which they can efficiently digest (Partridge 1981). Hence, for herbivorous mammals, distinct vegetation communities across sharp ecotones may produce spatial variation in selection that leads to or reinforces distinct dietary preferences, which may in turn determine fine-scale space use and a range of intra- and interspecific interactions (Via 1999, Via et al. 2000, Nosil et al. 2005).

One such ecotone exists on the western edge of the Kelso Valley, California where the southeastern slopes of the Sierra Nevada meet the valleys of the western Mojave Desert. Two closely related species of woodrat meet at this sharp ecotone: *Neotoma bryanti* (Bryant's woodrat) that primarily occur in the relatively mesic woodland and chaparral habitat of a rocky hill (hereafter called the "hill"), and *N. lepida* (Desert woodrat) that occur primarily in the adjacent Mojave Desert scrub habitat (hereafter called the "flats"; Shurtliff et al. 2014, Fig. 2). The two species are estimated to have diverged ~1.6 mya based on mtDNA (Fig. 1; Patton et al. 2007), and while they are largely spatially segregated between the two adjacent habitats, they do occasionally hybridize. These hybridization events lead to approximately 14% of individuals across the study site having hybrid ancestry, with backcrossing and introgression in both parental directions (Shurtliff et al. 2014; Jahner et al. unpublished data). Previous diet

analyses (Shurtliff et al. 2014; Matocq et al. 2020) suggest that *N. bryanti* and *N. lepida* consume distinct diets in the hill and flats, respectively. As such, this system offers an opportunity to investigate dietary choices across a sharp ecotone, as well as the potential role of dietary differences in limiting interspecific contact and hybridization.

Here, we sought to further characterize the degree to which dietary composition and preference differ between pure *N. bryanti* and *N. lepida* in their respective native habitats, and to uncover the potential ecological correlates maintaining species differences in diet across this ecotone. We integrate both field and laboratory studies to ask the following questions: 1) Do *N. bryanti* and *N. lepida* maintain distinct diets across this sharp ecotone in both wet and dry seasons, and in wet and dry years? 2) Do these species consume plants in the wild in proportion to their availability in the habitat, or do they exhibit selection/preference for particular plants? 3) When given a choice in experimental trials, do woodrats exhibit the same dietary preferences as exhibited in field-collected samples? 4) To what degree are plant nutritional content and plant secondary compounds correlated with dietary preferences? To address these questions, we quantify diet preferences in the wild using high-throughput sequencing of the chloroplast *trn*L intron from woodrat fecal samples collected across the ecotone. We further examine these apparent patterns of preference by conducting an experimental choice trial. To understand the underlying drivers of fine-scale diet differentiation in this system, we place these dietary preferences within the context of availability of these plants on the landscape, the plant secondary compound composition of these plants, and their nutritional quality. Our study provides a well-developed example of fine-scale diet

differentiation in mammalian herbivores: differences across an ecotone that are maintained between the species in both wet and dry conditions.

Methods

Study system

The study site is located in Kelso Valley, Kern Co., California, where *N. bryanti* and *N. lepida* meet and hybridize at the southern end of the Sierra Nevada mountain range (35°25'45 N, 118°15' 2 W). The mesic "hill" habitat sharply transitions to the xeric "flats" habitat (Fig. 1), and both parental species and hybrids can be found across a span of as little as tens of meters. The total area of the study site is approximately 50 hectares, approximately centered at the base of the hill (Fig. 1). We conducted vegetation surveys in 27 plots (hill = 16, flats = 11) to estimate the abundance of the most common shrubs and trees (details in supporting information).

Woodrat species identity

We identified individuals as *N. bryanti* or *N. lepida* using microsatellite loci previously developed for these species (Sousa et al. 2007). For animals included in the preference trials (see below), we obtained ear biopsies from each individual and conducted DNA extraction, amplification and scoring of microsatellite loci as described in Coyner et al. (2015). We established species identity by conducting a Bayesian assignment test as implemented in STRUCTURE (Pritchard et al. 2000 and Falush et al. 2003) at K = 2 as in Shurtliff et al. (2014) and used q_{lepida} values > 90% to assign

individuals to *N. lepida* and q_{lepida} values < 10% to assign individuals to *N. bryanti*. To confirm the species identity of individuals included in the fecal metabarcoding, we used the same genotyping approach, but started with the gDNA extractions used for *trn*L sequencing (see below) and performed three replicate PCRs per sample.

Fecal metabarcoding

To determine the dietary composition of *N. bryanti* and *N. lepida* at our site, fecal samples were collected from 35 unique woodrat nests during March-August of 2016. Species identity for fecal pellets was confirmed with microsatellite markers as described above. Clusters of approximately 10 - 20 fresh pellets were collected from *N. bryanti* nests in the hill habitat (n = 19) and *N. lepida* nests in the flats habitat (n = 16). These samples provided insight into diet in the spring months of March - May (*N. bryanti* = 11, *N. lepida* = 11) and summer months of July and August (*N. bryanti* = 8, *N. lepida* = 5). To ensure fecal pellets were characteristic of the sampling period, we located active latrines at woodrat nests and swept away all fecal material; after 3-4 nights, we collected fresh, adult-sized fecal pellets. It is important to note that woodrat houses are solely occupied by one adult woodrat, and these animals are highly territorial, so there is limited chance that more than one woodrat contributed to the fresh fecal pellets we collected. We placed pellets into coin envelopes to dry, and stored them long-term at -20°C. We submitted samples to Jonah Ventures LLC (Boulder, CO) for sequencing of a portion of the chloroplast *trn*L intron to reconstruct relative summer diet composition (methods including extraction, PCR amplification, sequencing, and raw data processing in supporting information). We removed operational taxonomic units (OTUs) that did not

occur in at least one sample with more than 1% abundance. We confirmed identity of remaining OTUs by conducting a BLASTn search (https://blast.ncbi.nlm.nih.gov). The potential presence of a plant at the study site was determined based on our own field collections and the CalFlora database (https://www.calflora.org/). If a resulting search returned more than one possible species, genus, or family that might occur at the site, we report the highest level of taxonomy (*i.e.* genus, family). Finally, to confirm the *trn*L primers used would detect the most common plants at the site, and to generate known sequences (*i.e.* vouchers) for these plants, we sequenced the following collected at the study site: *Ericameria nauseosa*, *Artemisia tridentata*, *Eriogonum fasciculatum*, *Prunus fasciculata*, *Frangula californica*, and *Phacelia tanacetefolia*.

Diet composition

We used read counts of all identified plants to calculate Shannon diversity for diets of *N. bryanti* and *N. lepida* and performed a two-sample t-test in R to compare diversity in diet composition (R Core Team 2016). We used read counts to determine if diets between the two species were distinct by performing a PERMANOVA using Bray-Curtis distances with the adonis function in the *vegan* package (Oksanen et al. 2013, R Core Team 2016).

To estimate individual and population-level (*i.e.* species) consumption of particular plants, we used both frequency of occurrence (FOO) and relative read abundance (RRA) of plant taxa identified in fecal samples. We considered a plant taxon present if it made up 1% or more of the total reads in a sample (Deagle et al. 2019). We calculated RRA for each plant within individual samples, and then averaged RRA values

for *N. bryanti* and *N. lepida*. We used the signassoc function in the R package *indicspecies* (De Caceres and Legendre 2009) on the resulting presence/absence matrix and RRA datasets to determine which plants were significantly associated with either *N. bryanti* or *N. lepida.* Average RRA values have traditionally been viewed with caution as they are prone to recovery bias and other artifacts, but the information contained within read counts can still provide important insights into the relative importance of certain plants at the population level (Deagle et al. 2019). Previous authors have reported correlation between relative abundance of plants consumed and raw number of reads obtained ($r^2 = 0.75$, $P < 10^{-15}$; Willerslev et al. 2014), and while FOO is less affected by recovery bias, RRA can provide a more accurate characterization of population-level diet (Deagle et al. 2019). We sought to incorporate measures of presence/absence (FOO) and relative abundance (RRA) to characterize dietary differences at the population level in this study.

In order to take individual variation into account in estimates of population-level consumption we used a hierarchical Bayesian approach implemented in R using the *bayespref* package (Fordyce et al. 2011) to estimate population-level consumption of the 5 most common plants identified in woodrat diets, which comprised ~80-90% of total reads (Tables 2, S1-S3). We pooled the remaining read counts from all other plant taxa into an “other” group. Rather than relying simply on RRA (as described above) to infer relative degree to which plants are consumed, this hierarchical Bayesian approach incorporates individual variation in our population-level consumption estimates (Fordyce et al. 2011, Forister et al. 2013). We used raw read count data to run models. Raw read counts were not normally distributed, therefore we square-root transformed read counts

prior to analysis. We ran models for 50,000 iterations, with a burn-in of 5,000 iterations and visually confirmed adequate chain-mixing. Hereafter, we will refer to these estimates simply as consumption.

Lastly, we considered diet composition of *N. bryanti* and *N. lepida* in this study (2016, a wet year) relative to that found previously (2013, a dry year; Matocq et al. 2020). We compare RRA values as consumption of plant food was not modeled for 2013 data.

Crude protein content of common shrubs

We characterized the nutritional value of common shrubs in each habitat and/or those that were most common in woodrat diets (see below) by measuring relative crude protein content. Crude protein content is considered the best single factor for determining nutritional value of forage plants (Sampson and Jespersen 1963, pg. 20). We collected leaves and fresh green growth of *F. californica*, *P. fasciculata, E. nauseosa*, *A. tridentata*, and *E. fasciculatum* in summer and dried at ambient temperature. We estimated crude protein on the dry matter basis using the Kjeldahl method (Association of Official Analytical Chemists 2002). In short, one gram of dried plant material was ground and digested in boric acid prior to titration to measure nitrogen content, which was multiplied by a factor of 6.25 (Association of Official Analytical Chemists 2002).

Preference trials

We conducted preference trials in the field from Jun-Aug of 2016 and 2017 to quantify dietary preference in *N. lepida* (n = 12; 3 F, 9 M) and *N. bryanti* (n = 15; 8 F, 7

M) for the two most common plants recovered from field diets (see below): *F. californica* and *P. fasciculata*. We provide all trapping and feeding trial details in the supporting information. All animal procedures were reviewed and approved by the University of Nevada Reno Institutional Animal Care and Use Committee, the California Department of Fish and Wildlife, and were consistent with the guidelines developed by the American Society of Mammalogists (Sikes et al. 2016).

We calculated a preference index with the following formula: $\text{Preference} = (p - f)/T$; where *p* is the total amount of *P. fasciculata* consumed during a trial, *f* is the total amount of *F. californica* consumed, and *T* is the total amount (grams) consumed. The resulting single response variable for preference during a given trial is bounded by $^{-}1$ and $^{+}1$; with positive values indicating preference for *P. fasciculata* and negative values indicating preference for *F. californica*. Results of a Shapiro-Wilk normality test conducted in R found these data to be normal ($W = 0.95$, $P = 0.21$). To test for confounding covariates, we used a linear model created in R with preference index as the response variable and species ID, and potentially confounding covariates (*i.e.* total time in trial, year, mixed vs. foliage food type, sex), as independent variables. This enabled us to rule out the possibility of confounding effects of these covariates on our independent variable of primary interest, species identity.

Results

Vegetation community

The most common shrubs and trees on the hill were *E. nauseosa (*33%), *E. fasciculatum* (16%), *F. californica* (13%), *Ephedra* sp. (11%), *Hesperoyucca whipplei* (7%) and multiple species of *Pinus* (5%). The most common shrubs and trees in the flats were *E. nauseosa* (60%), *E. fasciculatum* (11%), *Yucca brevifolia* (11%), *P. fasciculata* (10%), and *A. tridentata* (6%). Relative proportions of all subshrubs, shrubs and trees are provided in the supporting materials (Fig. S1, Table S4). Vegetation diversity was greater on the hill (H = 1.50) than the flats (H = 0.93; t = -4.40, df = 16.93, $P < 0.001$), and vegetation community composition differed between the hill and flats (MS = 1.83, r^2 = 0.33, P = 0.001). Of 91 woodrat nests in the flats, 59% were either directly at the base of *P. fasciculata* or were located in rocks with *P. fasciculata* adjacent, while the remaining were in *Y. brevifolia*, *E. nauseosa*, and *R. amarum*. Woodrat nests on the hill were primarily within large boulders with little if any immediately surrounding vegetation.

Diet composition, relative frequency of occurrence and relative read abundance

After filtering and verifying OTU representative sequences, we retained 847,690 reads from 35 woodrat fecal samples that represented 33 plant taxa (Tables 2, S1-S3). During spring, diet diversity was greater in *N. bryanti* (H = 1.32) than in *N. lepida* (H = 0.71; t = 4.30, df = 19.87, $p < 0.001$), and diet composition was also distinct between *N. bryanti* and *N. lepida* (MS = 2.81, r^2 = 0.46, P = 0.001). During summer, diet diversity was also greater in *N. bryanti* (H = 1.16) than in *N. lepida* (H = 0.41; t = 5.51, df = 9.69, $P < 0.001$), and diet composition was also distinct between *N. bryanti* and *N. lepida* (MS

= 1.99, r^2 = 0.60, *P* = 0.003). When data from both seasons were combined, diet diversity in *N. bryanti (H* = 1.25), was twice that of *N. lepida (H* = 0.62; t = 5.77, 32.92, *P* < 0.001), and diet composition was also distinct between the species (MS = 4.82, r^2 = 0.50, *P* = 0.001). In addition to the plants recovered from fecal samples, we confirmed that our primer set was able to recover the five common shrubs on which we tested them. Of note is that our known sequences for *Ericameria* and *Artemisia* are not different from many other species in the Asteraceae, thus all these similar sequences are collapsed into the Asteraceae (Tables 2, S1-S3).

Overall, *N. bryanti* and *N. lepida* exhibit distinctly different diets, but do consume some of the same plants. The frequency of occurrence (FOO) and relative abundance (RRA) of all 33 plant taxa identified may be found in Tables 2 and S1-S3. Notably, *N. bryanti* exhibited a more diverse diet with *F. californica* as the most abundant food item in spring and summer combined (FOO = 0.89, RRA = 0.41; Table 2). *Pinus* spp. and *Phacelia tanacetefolia* also occurred in the diet of *N. bryanti* with greater than 80% FOO and over 10% RRA in spring and summer combined (Table 2). *Neotoma bryanti* increased consumption of *F. californica* in summer relative to spring evidenced by increases in both FOO and RRA (Fig. 2, Tables S1 & S2). *Neotoma lepida* consumed a less diverse diet, with *P. fasciculata* being the most abundant in spring and summer (FOO = 1.00, RRA = 0.79; Table 2). *Neotoma lepida* increased consumption of *P. fasciculata* from spring to summer (RRA_{spring} = 0.74, RRA_{summer} = 0.91; Fig. 2, Tables 2 and S1 & S2). Overall, RRA for the Asteraceae family did not exceed 2% for either *N. bryanti* or *N. lepida* and the overall frequency of occurrence was also low ($FOO_{bryanti}$ = 0.37, FOO_{lepida} = 0.12). Thus we are confident that, even with our inability to discriminate within the

Asteraceae family, woodrats consume very little if any *E. nauseosa* or *A. tridentata* at our site.

Results of our hierarchical Bayesian modeling were consistent with diet composition based on FOO and RRA estimates. Notably, estimates of consumption using *bayespref* were less extreme than those from average RRA values (Tables 2, S1-S3 & S5, Fig. 2). *Frangula californica* was still the most common single plant in the diet of *Neotoma bryanti* and increased from spring to summer ($consumption_{spring}$ = 0.22 [95% CI 0.14-0.30], $consumption_{summer}$ = 0.36 [95% CI 0.26-0.45]; Table S5, Fig. 2). More than half the diet of *N. lepida* was composed of *P. fasciculata* also increased from spring to summer ($consumption_{spring}$ = 0.54 [95% CI 0.45-0.61], $consumption_{summer}$ = 0.65 [95% CI 0.56-0.71]; Table S5, Fig. 2). While diets of *N. bryanti* and *N. lepida* were vastly different, *Phacelia tanacetefolia,* an annual forb, was found to make up ~13-19% of the diet of both species (Table S5, Fig. 2).

Our measures of diet composition in this study were largely consistent with those previously described in Matocq et al. (2020). Their measure of diet occurred during the summer of 2013, an extreme drought year, wherein *Neotoma bryanti* consumed a high level of *F. californica* (RRA = 0.52) and *N. lepida* consumed large amounts of *P. fasciculata* (RRA = 0.59; Matocq et al. 2020). During spring of a wet year (2016; this study), when more vegetation diversity was available, *N. bryanti* reduced consumption of the 'difficult' *F. californica* relative to summer (RRA_{spring} = 0.35, RRA_{summer} = 0.51). In contrast, *N. lepida* maintained high levels of *P. fasciculata* in its diet whether an extreme drought year summer (see above) or a wet-year summer (*i.e.* 2016, RRA_{summer} = 0.91.

Even during a 'superbloom' spring, arguably the highest diversity this site experiences, *N. lepida* still consumed high quantities of *P. fasciculata* ($RRA_{spring} = 0.74$).

Crude protein content

Prunus fasciculata and *F. californica* had among the highest levels of summer crude protein content, 15.1% and 12.4% respectively (Table 1). Our measurements of crude protein for *E. nauseosa, A. tridentata,* and *E. fasciculatum* were 8.0, 8.4, and 5.1, respectively. Sampson and Jespersen (1963) reported average crude protein content of *F. californica* leaves as high as 19% from April to August. Summer crude protein content of *A. tridentata* was reported at 9.9% during August, with values as high as 15 % during spring (Sampson and Jespersen 1963, Welch 1989). Crude protein content of *E. nauseosa* can range from a minimum of 9% to a high of 11.8% when new growth forms (Sampson and Jespersen, 1963). Crude protein content of *E. fasciculatum* varied from 5.4% in summer to 8.6% for new growth (Genin and Badan-Dangon 1991). We were unable to find reported crude protein content of *P. fasciculata* in the literature.

Preference trials

A total of 27 individuals were included in diet trials: *N. bryanti* (n = 15), *N. lepida* (n = 12). We found that preference was significantly different between species ($P <$ 0.001, Table 3). *N. bryanti* exhibited a preference for *F. californica* (preference = $^{-}0.47$ [95% CI $^{-}0.66$ - $^{-}0.28$], while *Neotoma lepida* preferred *P. fasciculata* (preference = 0.61 [95% CI 0.41- 0.81]; Fig. 3). There was variation in preference index among individuals. However, all *N. lepida* individuals showed preference for *P. fasciculata* with two

individuals consuming only that plant, and all *N. bryanti* individuals showed preference for *F. californica* with two individuals consuming only that plant.

Discussion

Despite ongoing hybridization between *N. bryanti* and *N. lepida* (Shurtliff et al. 2014), we found differences in dietary preference and dietary composition between these two species; differences that were maintained in both wet and dry years, and across seasons. The primary plants differentially preferred by each species are nutritious relative to other available plants, but also potentially toxic in unique ways, suggesting these species may have evolved or developed distinct metabolic strategies to reduce toxin exposure. Given the degree of dietary plasticity we observed across seasons in natural diets and in preference trials, we find that *N. bryanti* is more of a dietary generalist than *N. lepida*. Dietary differences between the species likely contribute to their spatial segregation across the ecotone, which ultimately determines their opportunities for interspecific interactions, including hybridization.

At this ecotone, the vegetation of the hill community is more diverse and largely distinct from that of the flats, and this diversity and differentiation is partly reflected in the diets of the woodrats that occupy these habitats (Figs. 2 & S1, Tables 2 & S1-S3). Overall, dietary diversity of *N. bryanti* on the hill was twice that of *N. lepida* individuals living in the flats. Despite the diversity of plants consumed by *N. bryanti*, *F. californica* appears to predominate their diet. In contrast, *N. lepida* in the flats have a diet dominated by *P. fasciculata.* For both *N. bryanti* and *N. lepida*, these food plants (*i.e., F. californica*

and *P. fasciculata*) were consumed at higher rates than their availability on the landscape, suggesting dietary selection (Hodgson 1979).

During spring and summer of 2016, we show that the diet of *N. bryanti* on the hill is dominated by *F. californica* while the diet of *N. lepida* on the flats is dominated by *P. fasciculata*. However, we did find that some *N. bryanti* on the hill consumed a small amount of *P. fasciculata,* while *N. lepida* on the flats infrequently consumed *F. californica*. This result from our sample of wild diets is at least partly due to the relative rarity of these two plants in the 'alternate' habitat. However, results of our 2-choice trial show that even when given a choice of both plants, on average, *N. bryanti* primarily consumed *F. californica* and *N. lepida* primarily consumed *P. fasciculata*. As such, on average, individuals in our experimental trial showed a preference for the plant they most commonly consume in the natural environment. Overall, our field and experimental results demonstrate that *N. bryanti* show a preference for *F. californica* and *N. lepida* show a preference for *P. fasciculata*, which may reflect differences in behavioral acclimation to different resources and/or underlying species differences in their ability to metabolize these particular plants.

Despite the overall preference *N. bryanti* and *N. lepida* exhibit for these plants, there was a great deal of individual variation in our experimental trials. Specifically, most individuals consumed at least some of the presumably novel plant. This is a foraging behavior animals may employ to identify new food resources (Partridge 1981), and one we might expect when individuals are exposed to novel food items. This short-term consumption of a potentially novel, chemically distinct plant did not appear to have negative consequences for experimental animals as none lost excessive weight over this

short period (*i.e.* > 10% body mass) and animals remained alert and responsive. Overall, *N. bryanti* showed less extreme preference than *N. lepida* and these results are consistent with several *N. bryanti* on the hill consuming *P. fasciculata,* albeit in very low amounts, while *N. lepida* on the flats rarely consume *F. californica*. Both of these lines of evidence suggest that *N. bryanti* may be further towards the generalist end of the spectrum of specialization, while *N. lepida* may be further towards the specialist end of the spectrum (Shipley et al. 2009).

Herbivores may employ a range of dietary strategies, from specialist to generalist, to balance nutrient acquisition and exposure to plant PSCs. Specifically, facultative specialists exhibit diets largely restricted to a single 'difficult' (*i.e.* potentially toxic) food item, but are capable of expanding their diet when resource availability allows. In contrast, facultative generalists typically maintain more diverse diets, but are capable of restricting their diets to a 'difficult' plant when environmental conditions limit food resources (Shipley et al 2009). *Frangula californica* and *P. fasciculata* are known to contain PSC's that deter herbivory. *Frangula californica* contains anthraquinones, that can cause severe damage to the intestinal lining of mammals, and have hepatotoxic effects (Qin et al. 2016, Jung et al. 2011). In contrast, *P. fasciculata* contains cyanogenic glycosides, also highly toxic once cyanide is released from the parent compound (Vetter 2000). Chemical analysis of plants from the Kelso Valley have shown that *F. californica* and *P. fasciculata* contain chemical peaks consistent with the anthraquinone, emodin, in *F. californica* and the cyanogenic glycoside, prunasin, in *P. fasciculata* (Matocq et al. 2020). Given the potential toxicity of these plants, why do woodrats eat so much of them? On one hand, from a chemical perspective, these plant species may be among the

best of a bad lot. The other common shrubs present - *E. nauseosa, A. tridentata,* and *Ephedra* are also known to be chemically well-defended and/or energetically costly to consume (Johnson et al. 1976, Halls et al. 1994, Dial 1988). In addition to this, though, our nutritional analyses coupled with information available in the literature suggest that *F. californica* and *P. fasciculata* are among the most nutritious plants at this site in terms of crude protein. The composition of woodrat diets is likely a result of how *N. bryanti* and *N. lepida* have come to balance access to nutrition while minimizing their over-exposure to plant secondary compounds, as seen in other small mammals (Ulappa et al. 2014).

The degree to which woodrats and other herbivores can minimize their exposure to toxins by diversifying their diets (Freeland and Jansen 1974) depends on environmental conditions and associated plant availability. For this study site, we can begin to assess dietary plasticity under different seasonal and annual conditions by combining current results with data collected in previous years (Matocq et al. 2020). At one extreme is the 2013 snapshot of diet composition, which was taken in summer of an extreme drought wherein California received less precipitation than in any previous year in the 119-year observational record (Swain et al. 2014), and few annual forbs were observed at the site (Matocq pers. obs). This is in contrast to conditions at the site during 2016 - a wet and warm year facilitated by El Niño conditions that led to a spectacular 2016 spring ‘superbloom’ event (Treonis et al. 2019) characterized by high annual forb diversity across the Mojave desert. These snapshots of diet composition (*i.e.* 2013 and 2016) capture aridity extremes from centennial-scale drought, to wet year-summer, to wet year-spring, and thus, a plant diversity/availability gradient from low to high for this site.

As expected, if *N. bryanti* is a facultative generalist, high plant diversity in spring 2016 led to a decrease in consumption of *Frangula* whereas, *N. lepida* maintained high consumption of *Prunus*, regardless of availability of spring forbs. Others have classified *N. lepida* as a facultative specialist (Dial 1988, Shipley et al. 2009, Skopec et al. 2015), and our data support this classification. Indeed, *N. lepida* can consume large quantities of what is considered to be a potentially toxic plant. Although *N. lepida* is capable of consuming other plants at this site, individuals appear to prefer *P. fasciculata* even when other options are available, suggesting local specialization on this plant. *Neotoma bryanti* is also capable of consuming large quantities of a potentially toxic plant, *F. californica*. Yet, when given the option, we observed that *N. bryanti* will diversify its diet while still maintaining a high proportion of the 'difficult' plant in its diet - further supporting *N. bryanti* at this site as a facultative generalist. It should be noted that any study of diets in wild rodents that cache or hoard, as woodrats do, cannot discriminate between items that were eaten fresh versus those eaten after storage. Caching may reduce toxin content in plants, especially those with volatile compounds (*i.e.*, *Juniperus*, Torregrossa and Dearing 2009). While the primary compounds in *Frangula* and *Prunus* are not volatile, we do not know how these compounds would degrade over time if stored. Likewise, we do not know the extent to which woodrats at this site cache these plants.

Another critical ecological driver of diet composition and breadth at this study site is simply the presence of a closely related congener. Specifically, the narrower dietary niche of *N. lepida* at this site could in part be driven by competition with *N. bryanti*. Shurtliff et al. (2013) showed in laboratory trials that the relatively large-bodied *N. bryanti* is more aggressive than the relatively small-bodied *N. lepida*. Interspecific

competition is thought to be an important driver of dietary differentiation and fine-scale space use in interspecific contact zones between woodrats, with the large-bodied species typically monopolizing optimal nest sites (Cameron 1971, Dial 1988). *Neotoma bryanti* at this site monopolize what is likely the more optimal, relatively thermally stable boulder nesting area of the hill (Brown 1968). We suspect inherent differences in behavioral, physical, and metabolic capabilities have allowed *N. bryanti* to monopolize the hill habitat with its diversity of dietary plants, while *N. lepida* have persisted at the site in part because of its ability to locally specialize on *P. fasciculata*.

Woodrats are well-known for their capacity to consume large quantities of potentially toxic plants (*Larrea tridentata* - Mangione et al. 2000; *Juniperu*s sp. - Dial 1988, Skopec et al. 2007). In particular, *N. lepida* is known to locally specialize on chemically distinct plants across its range *(Larrea tridentata* - Mangione et al. 2000; *Juniperus* - Stones and Hayward, and here, desert almond). The mechanisms that underlie a woodrat's capacity to detoxify these diets likely include expression of their own detoxifying enzymes (Malenke et al. 2012, Kitanovic et al. 2018) and the activity of their gut microbiota (Kohl et al. 2014). Studies are needed to identify loci that are responsible for detoxification of different compounds, the degree to which specific alleles or pathways effectively metabolize particular PSC's, and the interaction between mammalian and microbial genomes in creating toxin resistant phenotypes (Forbey et al. 2018). If unique metabolic adaptations or microbial combinations allow *N. bryanti* and *N. lepida* to metabolize different plants in their respective habitats, then migrant and hybrid individuals that do not possess habitat-specific genomic or microbial combinations may suffer reduced fitness (Via 1999, Via et al. 2000, Nosil et al. 2005). Selection against

migrants would minimize opportunities for interspecific contact and mating (pre-zygotic isolation), while selection against hybrids with suboptimal allelic or microbial combinations would further limit introgression between the species (post-zygotic isolation). Continued integration of field and laboratory studies will be needed to identify the mechanisms that underlie metabolic processing of these diets, and how diet-related selection is influencing the evolutionary trajectory of these species.

Acknowledgements

We thank Michael Jee, Jason Wurtz, Dan Moore, and Angela Mitchell for assistance in field work. We thank Mozart Fonseca for use of his lab and equipment and Teshome Shenkoru for assistance in measuring plant crude protein content. We thank the members of the UNR Evoldoers lab group for comments on earlier drafts of this manuscript. We thank Mary V. Price and one anonymous reviewer for helpful comments on this manuscript. We also thank M. Forister, J. Forbey, and D. Dearing for helpful discussions. This research was supported in part by the National Science Foundation (IOS-1457209 and OIA-1826801).

Literature Cited

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J., Bierne, N., ... & Butlin, R. K. (2013). Hybridization and speciation. *Journal of evolutionary biology*, *26*(2), 229-246.

AOAC. (2002). Official method 2001.11 protein (crude) in animal feed, forage (plant tissue), grain and oilseeds. *Official Methods of Analysis of the Association of Official Analytical Chemists*.

Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*, *341*(6242), 497-503.

Brown, J. H. (1968). Adaptation to environmental temperature in two species of woodrats, *Neotoma cinerea* and *N. albigula*. (No. 135) Miscellaneous publications, Museum of Zoology. University of Michigan.

Cameron, G. N. (1971). Niche overlap and competition in woodrats. *Journal of Mammalogy*, *52*(2), 288-296.

Coyner, B. S., Murphy, P. J., & Matocq, M. D. (2015). Hybridization and asymmetric introgression across a narrow zone of contact between *Neotoma fuscipes and N. macrotis* (Rodentia: *Cricetidae*). *Biological Journal of the Linnean Society*, *115*(1), 162-172.

Cáceres, M. D., & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology*, *90*(12), 3566-3574.

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., ... & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How

should we convert sequence reads to dietary data? *Molecular Ecology*, *28*(2), 391-406.

Dearing, M. D., Mangione, A. M., & Karasov, W. H. (2000). Diet breadth of mammalian herbivores: nutrient versus detoxification constraints. *Oecologia*, *123*(3), 397-405.

Dearing, M. D., Foley, W. J., & McLean, S. (2005). The influence of plant secondary metabolites on the nutritional ecology of herbivorous terrestrial vertebrates. *Annu. Rev. Ecol. Evol. Syst.*, *36*, 169-189.

Dial, K. P. (1988). Three sympatric species of *Neotoma*: dietary specialization and coexistence. *Oecologia*, *76*(4), 531-537.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, *164*(4), 1567-1587.

Forbey, J. S., Liu, R., Caughlin, T. T., Matocq, M. D., Vucetich, J. A., Kohl, K. D., Dearing, M. D. & Felton, A. M. (2018). Using physiologically based models to predict population responses to phytochemicals by wild vertebrate herbivores. *animal*, *12*(s2), s383-s398.

Fordyce, J. A., Gompert, Z., Forister, M. L., & Nice, C. C. (2011). A hierarchical Bayesian approach to ecological count data: a flexible tool for ecologists. *PloS one*, *6*(11), e26785.

Forister, M. L., Scholl, C. F., Jahner, J. P., Wilson, J. S., Fordyce, J. A., Gompert, Z., ... & Nice, C. C. (2013). Specificity, rank preference, and the colonization of a non-native host plant by the Melissa blue butterfly. *Oecologia*, *172*(1), 177-188.

Freeland, W. J., & Janzen, D. H. (1974). Strategies in herbivory by mammals: the role of plant secondary compounds. *The American Naturalist*, *108*(961), 269-289.

Genin, D., & Badan-Dangon, A. (1991). Goat herbivory and plant phenology in a Mediterranean shrubland of northern Baja California. *Journal of arid environments*, *21*(1), 113-121.

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional ecology*, *21*(3), 394-407.

Halls, S. C., Gang, D. R., & Weber, D. J. (1994). Seasonal variation in volatile secondary compounds of *Chrysothamnus nauseosus* (Pallas) britt.; asteraceae ssp. hololeucus (Gray) hall. & clem. Influences herbivory. *Journal of chemical ecology*, *20*(8), 2055-2063.

Hodgson, J. (1979). Nomenclature and definitions in grazing studies. *Grass and forage science*, *34*(1), 11-17.

Iason, G. R., & Villalba, J. J. (2006). Behavioral strategies of mammal herbivores against plant secondary metabolites: the avoidance–tolerance continuum. *Journal of chemical ecology*, *32*(6), 1115-1132.

Johnson, A. E., James, L. F., & Spillett, J. (1976). The abortifacient and toxic effects of big sagebrush (*Artemisia tridentata*) and juniper (*Juniperus osteosperma*) on domestic sheep. *Journal of Range Management*, 278-280.

Jung, K. A., Min, H. J., Yoo, S. S., Kim, H. J., Choi, S. N., Ha, C. Y., ... & Lee, J. S. (2011). Drug-induced liver injury: twenty five cases of acute hepatitis following ingestion of *Polygonum multiflorum* Thunb. *Gut and liver*, *5*(4), 493.

Kelsey, R. G., Stephens, J. R., & Shafizadeh, F. (1982). The chemical constituents of sagebrush foliage and their isolation. *Journal of Range Management*, 617-622.

Kitanovic, S., Orr, T. J., Spalink, D., Cocke, G. B., Schramm, K., Wilderman, P. R., ... & Dearing, M. D. (2018). Role of cytochrome P450 2B sequence variation and gene copy number in facilitating dietary specialization in mammalian herbivores. *Molecular ecology*, *27*(3), 723-736.

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Denise Dearing, M. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology letters*, *17*(10), 1238-1246. doi: 10.1111/ele.12329

Kohl, K. D., & Dearing, M. D. (2012). Experience matters: prior exposure to plant toxins enhances diversity of gut microbes in herbivores. *Ecology letters*, *15*(9), 1008-1015.

Malenke, J. R., Magnanou, E., Thomas, K., & Dearing, M. D. (2012). Cytochrome P450 2B diversity and dietary novelty in the herbivorous, desert woodrat (*Neotoma lepida*). *PloS one*, *7*(8), e41510.

Mangione, A. M., Dearing, M. D., & Karasov, W. H. (2004). Creosote bush (*Larrea tridentata*) resin increases water demands and reduces energy availability in desert woodrats (*Neotoma lepida*). *Journal of chemical ecology*, *30*(7), 1409-1429.

Mangione, A. M., Dearing, M. D., & Karasov, W. H. (2000). Interpopulation differences in tolerance to creosote bush resin in desert woodrats (*Neotoma lepida*). *Ecology*, *81*(8), 2067-2076.

Matocq, M. D., Ochsenrider, K., Jeffrey, C. S., Nielsen, D., & Richards, L. A. (2020). Fine-scale differentiation in diet and metabolomics of small mammals across a sharp ecological transition. *Frontiers in Ecology and Evolution*, *8*, 282.

Nosil, P., Vines, T. H., & Funk, D. J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, *59*(4), 705-719.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Oksanen, M. J. (2013). Package 'vegan'. *Community ecology package, version*, *2*(9), 1-295.

Partridge, L. (1981). Increased preferences for familiar foods in small mammals. *Animal Behaviour*, *29*(1), 211-216.

Patton, J. L., Huckaby, D. G., & Álvarez-Castañeda, S. T. (2007). The evolutionary history and a systematic revision of woodrats of the *Neotoma lepida* group (Vol. 135). Univ of California Press.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Qin, B., Xu, Y., Chen, J., Huang, W., Peng, Y., & Zheng, J. (2016). Chemical reactivity of emodin and its oxidative metabolites to thiols. *Chemical Research in Toxicology*, *29*(12), 2114-2124.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sampson, A. W., & Jesperson, B. S. 1963. California range brushlands and browse plants. *Calif. Agric. Exp. Stn. Ext. Serv. Man*, *33*, 162.

Shipley, L. A., Forbey, J. S., & Moore, B. D. (2009). Revisiting the dietary niche: when is a mammalian herbivore a specialist? *Integrative and Comparative Biology*, *49*(3), 274-290.

Shurtliff, Q. R., Murphy, P. J., & Matocq, M. D. (2014). Ecological segregation in a small mammal hybrid zone: habitat-specific mating opportunities and selection against hybrids restrict gene flow on a fine spatial scale. *Evolution*, *68*(3), 729-742.

Shurtliff, Q. R., Murphy, P. J., Yeiter, J. D., & Matocq, M. D. (2013). Experimental evidence for asymmetric mate preference and aggression: behavioral interactions in a woodrat (*Neotoma*) hybrid zone. *BMC evolutionary biology*, *13*(1), 220.

Sikes, R. S., & Animal Care and Use Committee of the American Society of Mammalogists. (2016). 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *Journal of Mammalogy*, *97*(3), 663-688.

Skopec, M. M., Haley, S., & Dearing, M. D. (2007). Differential hepatic gene expression of a dietary specialist (*Neotoma stephensi*) and generalist (*Neotoma albigula*) in response to juniper (*Juniperus monosperma*) ingestion. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *2*(1), 34-43.

Skopec, M. M., Kohl, K. D., Schramm, K., Halpert, J. R., & Dearing, M. D. (2015). Using the specialization framework to determine degree of dietary specialization in a herbivorous woodrat. *Journal of chemical ecology*, *41*(12), 1059-1068.

Smith, T. B., Wayne, R. K., Girman, D. J., & Bruford, M. W. (1997). A role for ecotones in generating rainforest biodiversity. *Science*, *276*(5320), 1855-1857.

Sorensen, J. S., McLister, J. D., & Dearing, M. D. (2005). Novel plant secondary metabolites impact dietary specialists more than generalists (*Neotoma* spp.). *Ecology*, *86*(1), 140-154.[a]

Sorensen, J. S., McLister, J. D., & Dearing, M. D. (2005). Plant secondary metabolites compromise the energy budgets of specialist and generalist mammalian herbivores. *Ecology*, *86*(1), 125-139.[b]

Sousa, B., Svensson, L. M. E., & Patton, J. L. (2007). Characterization of 18 microsatellite loci for woodrats of the *Neotoma lepida* group (Rodentia, *Cricetidae*, Neotominae). *Molecular ecology notes*, *7*(5), 868-870.

Stones, R. C., & Hayward, C. L. (1968). Natural history of the desert woodrat, *Neotoma lepida*. *American Midland Naturalist*, 458-476.

Swain, D. L., Tsiang, M., Haugen, M., Singh, D., Charland, A., Rajaratnam, B., & Diffenbaugh, N. S. (2014). The extraordinary California drought of 2013/2014: Character, context, and the role of climate change. *Bull. Am. Meteorol. Soc*, *95*(9), S3-S7.

Torregrossa, A. M., & Dearing, M. D. (2009). Caching as a behavioral mechanism to reduce toxin intake. *Journal of Mammalogy*, *90*(4), 803-810.

Treonis, A. M., Sutton, K. A., Unangst, S. K., Wren, J. E., Dragan, E. S., & McQueen, J. P. (2019). Soil organic matter determines the distribution and abundance of nematodes on alluvial fans in Death Valley, California. *Ecosphere*, *10*(4), e02659.

Ulappa, A. C., Kelsey, R. G., Frye, G. G., Rachlow, J. L., Shipley, L. A., Bond, L., ... & Forbey, J. S. (2014). Plant protein and secondary metabolites influence diet selection in a mammalian specialist herbivore. *Journal of mammalogy*, *95*(4), 834-842.

Via, S. (1999). Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, *53*(5), 1446-1457.

Via, S., Bouck, A. C., & Skillman, S. (2000). Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution*, *54*(5), 1626-1637.

Vetter, J. (2000). Plant cyanogenic glycosides. *Toxicon*, *38*(1), 11-36.

Walker, S., Wilson, J. B., Steel, J. B., Rapson, G. L., Smith, B., King, W. M., & Cottam, Y. H. (2003). Properties of ecotones: evidence from five ecotones objectively determined from a coastal vegetation gradient. *Journal of Vegetation Science*, *14*(4), 579-590.

Welch, B. L. (1989). Nutritive value of shrubs. In C.M. McKell (Ed.), *The biology and utilization of shrubs* (pp. 405-424). Elsevier.

West-Eberhard, M. J. (1989). Phenotypic plasticity and the origins of diversity. *Annual review of Ecology and Systematics*, *20*(1), 249-278.

Westoby, M. (1978). What are the biological bases of varied diets? *The American Naturalist*, *112*(985), 627-631.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., ... & Craine, J. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, *506*(7486), 47.

Table 1: Crude protein content (percent dry matter basis) of 5 common perennial shrubs found at the study site.

Species	**Crude Protein This Study**	**Crude Protein Literature**	**Reference**
Artemisia tridentata	8.4 ± 0.9	8.5 - 15	Welch 1989; Sampson and Jespersen 1963; Cook and Harris 1950; Kelsey et al. 1982
Ericameria nauseosa	8.0 ± 1.4	7.8 - 11.8	Welch 1989; Sampson and Jespersen 1963
Eriogonum fasciculatum	5.1 ± 1.1	5.1-5.7	Genin and Badan-Dangon 1991
Prunus fasciculata	15.1 ± 0.1	N/A	No published record
Frangula californica	12.4 ± 0.2	7.5 - 19	Sampson and Jespersen 1963

Table 2: Frequency of occurrence (FOO), relative read abundance (RRA), and where applicable, the percent abundance of woody plants in each habitat of plants identified in the diets of *N. bryanti* and *N. lepida*. Here we include only those plants that occurred with FOO > 15% in spring and summer 2016 combined (full dietary plant list in Tables S1-S3). *P*-values are corrected for multiple comparisons. We confirmed the presence of *E. nauseosa* voucher sequences in some samples and therefore list FOO and RRA for those within the larger Asteraceae family. Where applicable, the percent abundance of woody plants in each habitat is reported (full plant list in Table S4).

Taxa Identified	***N. bryanti* (n = 19)**			***N. lepida* (n = 16)**			***P-value***	
	FOO	RRA	%hill	FOO	RRA	%flats	FOO	RRA
Prunus fasciculata	0.21	0.04	0.04	1.00	0.79	0.10	**0.01**	**0.01**
Frangula californica	0.89	0.41	0.13	0.06	<0.01	0.01	**0.01**	**0.01**
Phacelia tanacetefolia	0.89	0.11		0.82	0.14		0.65	0.34
Pinus spp.	0.84	0.19	0.05	0.12	<0.01	<0.01	**0.01**	**0.01**
Eriogonum umbellatum	0.68	0.08		0.24	0.03		**0.01**	0.16
Ribes amarum	0.32	0.04	0.03	0.00	0.00	0	0.06	**0.01**
Acmispon americanus	0.32	0.03		0.00	0.00		0.08	**0.03**
Asteraceae	0.37	0.02		0.12	<0.01		0.14	0.17
Ericameria nauseosa$_{voucher}$	0.21	<0.01	0.33	0.00	0.00	0.60	NA	NA
Euphorbia maculata	0.00	0.00		0.24	0.02		0.10	0.06
Cercocarpus betuloides	0.16	0.01	<0.01	0.00	0.00	0	0.23	0.29
Salvia columbariae	0.21	<0.01		0.00	0.00		0.20	0.10

Table 3: Effects of variables included in linear model of preference trials. The dependent variable was the preference index for either *F. californica* or *P. fasciculata* - measured as the amount of *P. fasciculata* minus the amount of *F. californica* consumed divided by the total amount of food consumed during the trial.

Variable	**Estimate**	**Std. Error**	**t value**	***P*-value**
(Intercept)	-0.19333	0.17713	-1.091	0.2880
N. lepida	1.17135	0.14086	8.316	**<0.001**
Sex	-0.13788	0.13989	-0.986	0.3361
Mass change in trial	-0.16515	0.08127	-2.032	0.0556
Year	-0.50832	0.31591	-1.609	0.1233
Duration of Trial	-0.17816	0.16896	-1.054	0.3042
Food Type	0.04444	0.20654	0.215	0.8318

Results of overall model: Residual standard error: 0.3225 on 20 degrees of freedom; multiple R-squared: 0.8023; adjusted R-squared: 0.743; F-statistic: 13.53 on 6 and 20 degrees of freedom; *P*-value: **< 0.001**

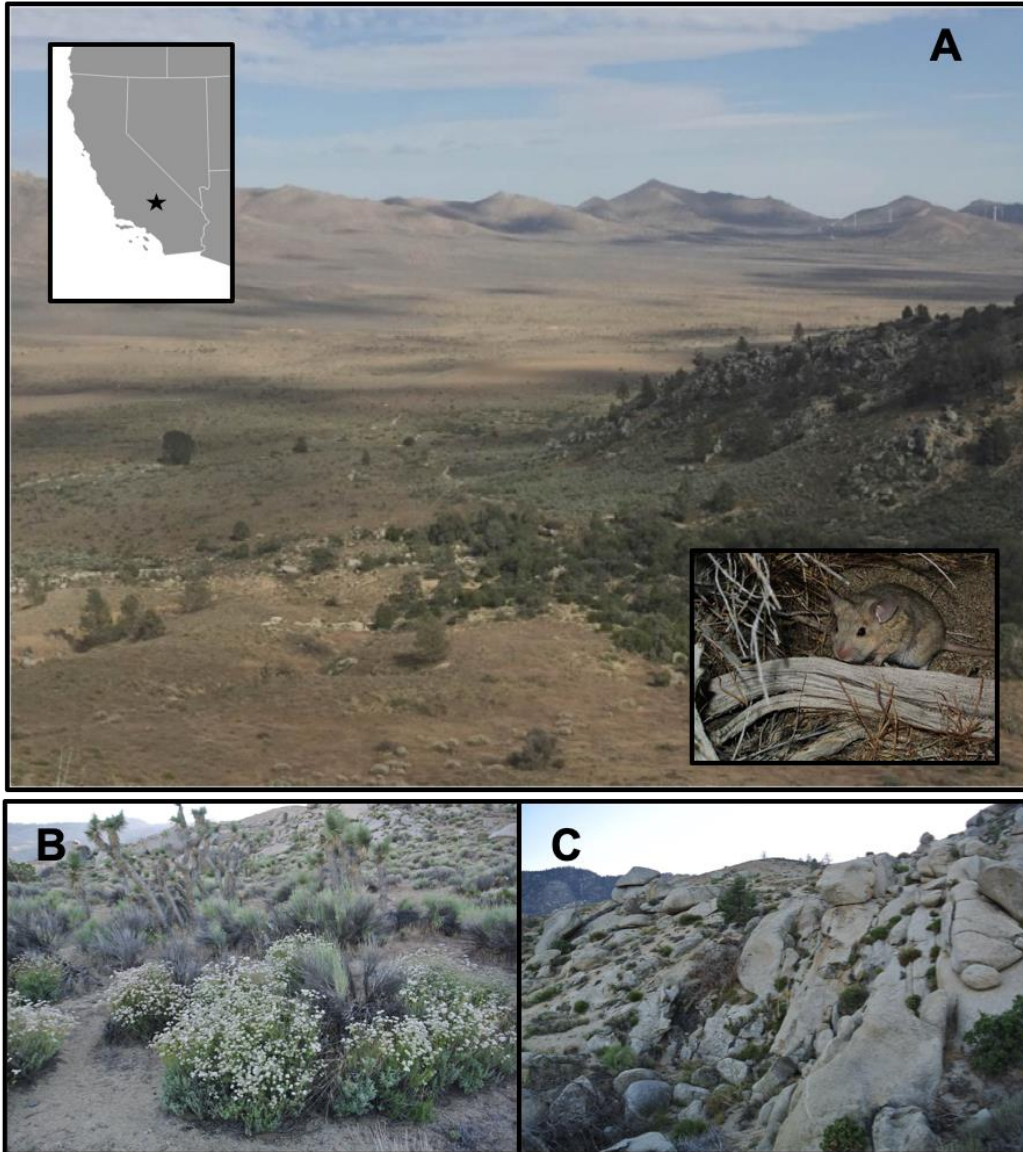

Figure 1: Panel (A) depicts the study site where the mesic hill transitions to the xeric flats. Photo taken from the north looking south. Black star in inset map represents approximate location of the study in Kelso Valley, California. Panels (B) and (C) depict habitat of the flats and hill habitats, respectively. Inset photo of woodrat is *Neotoma lepida*.

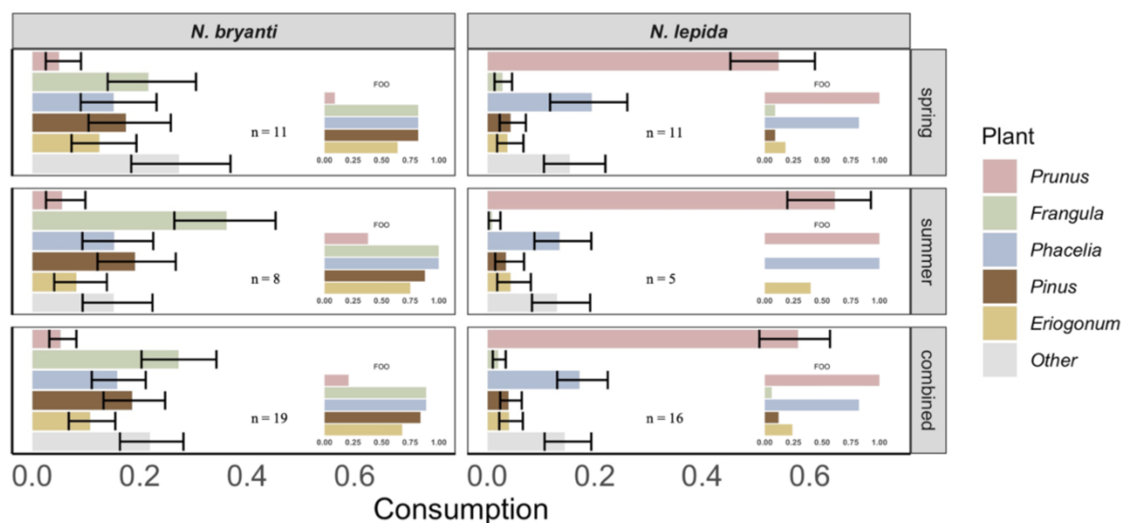

Figure 2: Consumption of the five most abundant plant taxa identified in woodrat diets in 2016 estimated for spring and summer individually, and both seasons combined. Consumption was estimated with *bayespref* using square root transformed read counts. Large bars are medians with 95% credible intervals from Bayesian posterior distributions. Insets represent frequency of occurrence (FOO) of these same plants.

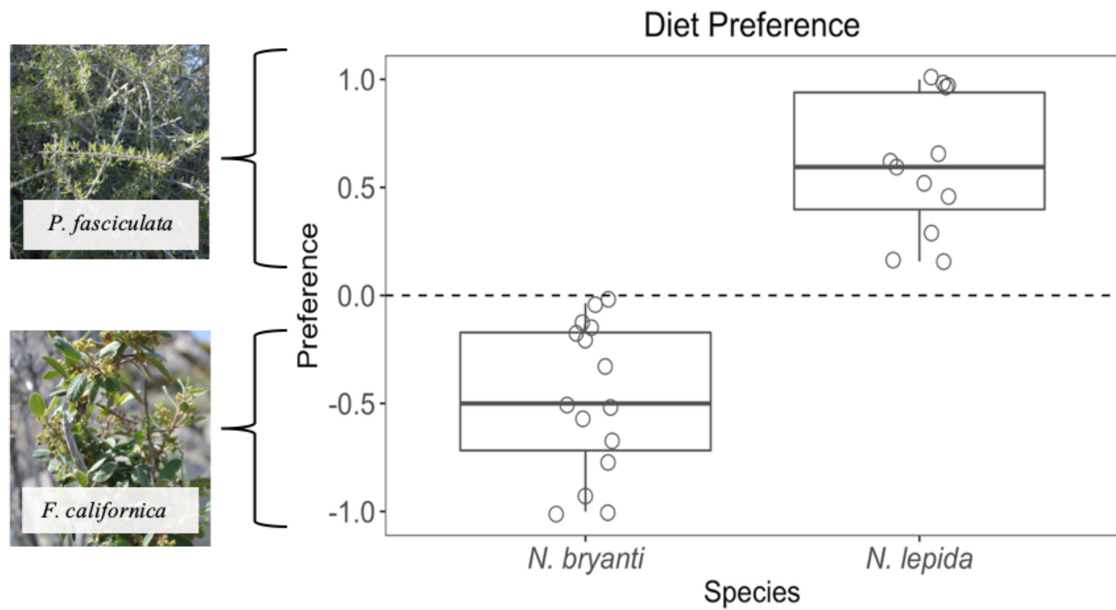

Figure 3: Group level average preference index for each species; *N. bryanti* and *N. lepida* along the x-axis. Preference as measured here is an index of the amount of *P. fasciculata* minus the amount of *F. californica* consumed divided by the total amount of food consumed during the trial. The y-axis represents this index: positive values indicate preference for *Prunus fasciculata* and negative values indicate preference for *Frangula californica*.

Chapter 1—Supporting Information

Text S1: Relative proportion of woody shrubs and trees

To quantify the perennial dietary landscape available to woodrats, we measured the relative proportion of woody shrubs and trees within a 25 m radius plot centered on active woodrat nests in both the hill (n = 16) and flats (n = 11) habitat. Nests were located and 25 meter transect lines were run from the center in each cardinal direction to delineate the plot. We identified and counted all shrubs and trees in each plot and estimated the relative number of individuals. We calculated Shannon diversity for each plot and performed a two-sample t test to compare diversity in woody shrubs and trees between the hill and flats (R Core Team 2016). To determine if the two habitats were characterized by distinct vegetation communities, we performed a PERMANOVA using the Bray-Curtis distances and with the adonis function in the *vegan* package (Oksanen et al. 2013). Because most *N. lepida* nests in the flats habitat occurred at the base of shrubs, we noted the species of plant in which nests were constructed. Average relative proportions were calculated for 15 common shrubs and trees occurring across the study site (Fig. S1, Table S4).

Text S2: Methods for Fecal Metabarcoding

Fecal samples were submitted in barcode-labeled vials to Jonah Ventures LLC (Boulder, CO) for *trn*L sequencing. Samples were assigned to a well in a 96 well plate. Then, a sterile cotton swab was dipped in nuclease free water before swabbing the sample. Approximately 0.25 grams of each sample was used for DNA extractions.

Sample swabs were placed in corresponding well with sterile tweezers. Samples were then processed or stored in -20°C until the extraction process could be performed.

Genomic DNA was extracted using the DNeasy PowerSoil HTP 96 Kit (Cat # 12955-4) according to the manufacturer's instructions. For each sample, a portion of the chloroplast *trn*L intron was PCR amplified using the c and h *trn*L primers (Taberlet et al. 2007):, *c* – CGAAATCGGTAGACGCTACG and *h* – CCATTGAGTCTCTGCACCTATC. Primers also had a 5' adaptor sequence for indexing and Illumina sequencing. Each 25µL PCR included 0.4 µM of each primer and 1µl of gDNA (Promega catalog # M5133, Madison, WI). DNA was PCR amplified with the following: initial denaturation at 94°C for 3 minutes, followed by 40 cycles of 30 seconds at 94°C, 30 seconds at 55°C, and 1 minute at 72°C, and a final elongation at 72°C for 10 minutes.

Amplicons were cleaned with UltraClean 96 PCR Cleanup Kit (384) (cat#12596-4) (MoBio) according to manufacturer's instructions. Another round of PCR was used to give samples unique index sequences. Successful barcode attachment was confirmed on 2% agarose gel. Final amplicons from each sample were cleaned and normalized using SequalPrep Normalization Plates (Life Technologies, Carlsbad, CA) according to manufacturer's instructions; samples were pooled and sequenced on an Illumina MiSeq (San Diego, CA) in the CU Boulder BioFrontiers Sequencing Center.

Sequences were demultiplexed using Golay barcodes via QIIME v1.9.1 (Caporaso et al. 2010). Forward and reverse read sequences were trimmed to 235 nucleotides using the option -fastx_truncate, and then merged with the -fastq_mergepairs option in usearch8 (Edgar 2010). Primers were removed with cutadapt (Martin 2011) and *trn*L amplicons

were processed using the UPARSE pipeline (Edgar 2013). Plant taxonomy was assigned with the SINTAX protocol (http://www.drive5.com/usearch/manual/utax_user_train.html) available in usearch (v8.1.1861; Edgar 2010). Sequences were quality trimmed and OTUs were clustered at 99% similarity with de novo chimera checking. A custom SINTAX *trn*L reference database was constructed by downloading any annotated GenBank (Benson et al. 2005) records that contain the *trn*L gene. This was then used to assign potential taxonomic groups. All extracted amplicon regions were dereplicated to 100% sequence identity and any identical sequence across lineages were collapsed to the lowest-common-ancestor. Closed-reference OTUs were generated by searching against the *trn*L reference database at 99% sequence similarity. Additional OTUs were generated from plant voucher specimens that were collected by D. Nielsen from the study site. Finally, representative sequences that were present with at least 1% abundance were manually blasted in GenBank to confirm taxonomic identification of plants known to be present at the study site (resulting plant IDs in Tables S1-S3).

Text S3: Methods for Experimental Preference Trials

To capture live animals for preference trials, we placed two tomahawk live traps wherever a nest or sign of woodrat activity was present. Trap treadles were scented with a very small amount of peanut butter and oats in order to lure woodrats, but not enough to be consumed and potentially alter fecal composition or microbiome. Each new animal trapped was given a unique ear tag for identification, sexed, weighed, measured, and a portion of the ear pinnae was removed with sterilized surgical scissors and stored in 95%

ethanol at ambient temperature for genotyping. Animals used in preference trials were transported in tomahawk traps back to the temporary field lab, were provided with cotton bedding, water, and were placed under dark cover for the remainder of the day. Woodrats were fasted for at least 8 hours prior to trials. We included only adult individuals, and females showed no sign of reproductive activity.

Enclosures for preference trials were designed from plastic containers with lids that allowed air flow (60.3cm L x 40.6 W x 34.3 H). Fresh cotton was provided during each trial for bedding material. Because fresh vegetation was provided from known diets at this site, water was not provided after we confirmed that woodrats did not readily consume freely available water. Clippings of *F. californica* and *P. fasciculata* were collected haphazardly from the study site and provided in equal amount by weight to each individual woodrat during each trial. For *P. fasciculata*, clippings included both leaves and freshly grown stems as bark of *P. fasciculata* is consumed by woodrats as well (personal obs. D. Nielsen). Fruits of *F. californica* were used for some of the 2016 trials because they were available at the site. All other trials consisted of only foliage from each plant species. We recorded plant type as mixed (fruits and foliage), and foliage (foliage only) for each trial. We conducted preference trials at night between the hours of 2000-0500; the active hours of these nocturnal herbivores. Researchers conducting cafeteria-style preference tests with other species of *Neotoma* have reported detecting dietary preference within only 30 minutes (McEachern et al. 2006). As woodrats are known to be novelty seekers, we allowed trials to run between 4 and 8 hours, or as long as they exhibited foraging behavior, to allow sufficient time for woodrats to exhibit feeding on ‘novel’ plants as well as preferred plants. We monitored woodrats and removed them

from the trial as soon as activity and feeding had ceased. No woodrat completely consumed all of either plant material provided.

We accounted for evaporative water loss of plant material during trials using control enclosures maintained during each trial. After each trial (experimental or control), the remaining material was removed and weighed to the nearest 0.1 gram using a digital scale to calculate the amount of plant mass lost during the trial. For each trial night, we corrected the amount consumed by woodrats with the amount of change in mass of each plant in control enclosures. This allowed us to account for evaporative change in plant mass and assign a conservative consumption value to each plant type during each trial. We did not include trials in which woodrats consumed less than 1 gram of food material total. At the end of each trial, animals were re-weighed and released back at the point of capture. Animals were kept no longer than ~ 24 hours before being released.

Literature Cited

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2005). GenBank. *Nucleic acids research*, *33*(suppl_1), D34-D38.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... & Huttley, G. A. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), 335-336.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460-2461.

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods*, *10*(10), 996.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*(1), 10-12.

McEachern, M.B., A Eagles-Smith, C., M Efferson, C., & H Van Vuren, D. (2006). Evidence for local specialization in a generalist mammalian herbivore, *Neotoma fuscipes*. *Oikos*, *113*(3), 440-448.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Oksanen, M. J. (2013). Package 'vegan'. *Community ecology package, version*, *2*(9), 1-295.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Taberlet P., Coissac E, Pompanon F, et al. (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, 35(3), e14-e14 doi:10.1093/nar/gkl938.

Table S1: Spring 2016 frequency of occurrence (FOO) and relative read abundance (RRA) of 33 unique plant taxa identified in the diets of *N. bryanti* and *N. lepida*. *P*-values are corrected for multiple comparisons. Where applicable, percent abundance of woody plants in each habitat is reported.

Taxa Identified	***N. bryanti* (n = 11)**			***N. lepida* (n = 11)**			***P-value***	
	FOO	RRA	%hill	FOO	RRA	%flats	FOO	RRA
Prunus fasciculata	0.09	0.04	0.04	1.00	0.74	0.10	0.01	0.01
Frangula californica	0.82	0.35	0.13	0.09	<0.01	0.01	0.01	0.01
Phacelia tanacetefolia	0.82	0.12		0.82	0.18		0.88	0.33
Pinus	0.82	0.18	0.05	0.09	<0.01	<0.01	0.01	0.01
Eriogonum umbellatum	0.64	0.10		0.18	0.03		0.05	0.34
Ribes amarum	0.45	0.05	0.03	0.00	0.00	0	0.04	0.03
Acmispon aericanus	0.55	0.05		0.00	0.00		0.02	0.03
Asteraceae	0.45	0.01		0.18	<0.01		0.38	0.23
Ericameria nauseosa voucher	0.27	<0.01	0.33	0.00	0.00	0.60	NA	NA
Euphorbia maculata	0.00	0.00		0.18	0.02		0.43	0.23
Cercocarpus betuloides	0.18	<0.01	<0.01	0.00	0.00	0	0.44	0.42
Mentzelia nitens	0.18	0.01		0.09	<0.01		0.71	0.63
Salvia columbariae	0.36	0.02		0.00	0.00		0.13	0.05
Eriastrum densifolium	0.09	0.01		0.00	0.00		0.76	0.14
Populus	0.09	0.01		0.00	0.00		0.72	0.68
Yucca brevifolia	0.00	0.00	0.01	0.00	0.00	0.11	1.00	0.02
Leptosyne	0.18	<0.01		0.00	0.00		0.37	0.20
Ephedra	0.09	<0.01	0.11	0.00	0.00	<0.01	0.74	0.01
Chenopodium	0.00	0.00		0.09	<0.01		0.73	0.37
Camissonia campestris	0.00	0.00		0.09	<0.01		0.70	0.26
Stephanomeria	0.09	<0.01		0.00	0.00		0.72	0.13
Descurainia pinnata	0.00	0.00		0.09	<0.01		0.68	0.08
Cuppressaceae	0.09	<0.01		0.00	0.00		0.73	0.77
Cirsium arvense	0.00	0.00		0.00	0.00		1.00	0.44
Erodium cicutarium	0.00	0.00		0.09	<0.01		0.68	0.05
Claytonia sp	0.18	<0.01		0.00	0.00		0.40	0.02

Poaceae	0.00	0.00	0.09	<0.01	0.70	0.28
Lupinus	0.09	<0.01	0.00	0.00	0.75	0.06
Ceanothus cordulatus	0.00	0.00	0.00	0.00	1.00	1.00
Phragmites australis	0.00	0.00	0.09	<0.01	0.70	0.12
Camissonia kernensis	0.00	0.00	0.09	<0.01	0.74	0.15
Thysanocarpus	0.09	<0.01	0.00	0.00	0.72	0.68
Scrophularia desertorum	0.00	0.00	0.00	0.00	1.00	0.25
Bromus	0.00	0.00	0.09	<0.01	0.70	0.25

Table S2: Summer 2016 frequency of occurrence (FOO) and relative read abundance (RRA) of 33 unique plant taxa identified in the diets of *N. bryanti* and *N. lepida*. *P*-values are corrected for multiple comparisons. Where applicable, the percent abundance of woody plants in each habitat is reported.

Taxa Identified	***N. bryanti* (n = 8)**			***N. lepida* (n = 5)**			***P-value***	
	FOO	RRA	%hill	FOO	RRA	%flats	FOO	RRA
Prunus fasciculata	0.38	0.05	0.04	1.00	0.91	0.10	0.10	0.01
Frangula californica	1.00	0.51	0.13	0.00	0.00	0.01	0.01	0.01
Phacelia tanacetefolia	1.00	0.08		1.00	0.03		1.00	0.39
Pinus	0.88	0.20	0.05	0.00	0.00	<0.01	0.01	0.01
Eriogonum umbellatum	0.75	0.06		0.40	0.01		0.38	0.09
Ribes amarum	0.13	0.03	0.03	0.00	0.00	0	0.82	0.06
Acmispon americanus	0.00	0.00		0.00	0.00		1.00	0.34
Asteraceae	0.25	0.03		0.00	0.00		0.62	0.60
Ericameria nauseosa$_{voucher}$	0.13	<0.01	0.33	0.00	0.00	0.60	NA	NA
Euphorbia maculata	0.00	0.00		0.40	0.02		0.27	0.20
Cercocarpus betuloides	0.13	0.02	<0.01	0.00	0.00	0	0.89	0.74
Mentzelia nitens	0.00	0.00		0.20	<0.01		0.63	0.66
Salvia columbariae	0.00	0.00		0.00	0.00		1.00	0.60
Eriastrum densifolium	0.13	<0.01		0.00	0.00		0.88	0.61
Populus	0.00	0.00		0.00	0.00		1.00	1.00
Yucca brevifolia	0.00	0.00	0.01	0.20	<0.01	0.11	0.70	0.01
Leptosyne	0.00	0.00		0.00	0.00		1.00	1.00
Ephedra	0.13	<0.01	0.11	0.00	0.00	<0.01	0.86	0.16
Chenopodium	0.00	0.00		0.00	0.00		1.00	1.00
Camissonia campestris	0.00	0.00		0.20	<0.01		0.65	0.60
Stephanomeria	0.13	<0.01		0.00	0.00		0.85	0.80
Descurainia pinnata	0.00	0.00		0.00	0.00		1.00	0.67
Cuppressaceae	0.00	0.00		0.00	0.00		1.00	1.00
Cirsium arvense	0.13	<0.01		0.00	0.00		0.86	0.58
Erodium cicutarium	0.00	0.00		0.00	0.00		1.00	0.25
Claytonia sp	0.00	0.00		0.00	0.00		1.00	1.00
Poaceae	0.00	0.00		0.00	0.00		1.00	0.26
Lupinus	0.00	0.00		0.00	0.00		1.00	0.34
Ceanothus cordulatus	0.13	<0.01		0.00	0.00		0.82	0.89
Phragmites australis	0.00	0.00		0.00	0.00		1.00	0.73

Camissonia kernensis	0.00	0.00	0.00	0.00	1.00	0.60
Thysanocarpus	0.00	0.00	0.00	0.00	1.00	0.89
Scrophularia desertorum	0.13	<0.01	0.00	0.00	0.86	0.85
Bromus	0.00	0.00	0.00	0.00	1.00	0.60

Table S3: Frequency of occurrence (FOO) and relative read abundance (RRA) of 33 unique plant taxa identified in the diets of *N. bryanti* and *N. lepida* in spring and summer 2016 combined. *P*-values are corrected for multiple comparisons. We confirmed the presence of *E. nauseosa* voucher sequences in some samples and therefore include these within the Asteraceae family. Where applicable, the percent abundance of woody plants in each habitat is reported.

Taxa Identified	***N. bryanti* (n = 19)**			***N. lepida* (n = 16)**			***P-value***	
	FOO	RRA	%hill	FOO	RRA	%flats	FOO	RRA
Prunus fasciculata	0.21	0.04	0.04	1.00	0.79	0.10	**0.01**	**0.01**
Frangula californica	0.89	0.41	0.13	0.06	<0.01	0.01	**0.01**	**0.01**
Phacelia tanacetefolia	0.89	0.11		0.82	0.14		0.65	0.34
Pinus	0.84	0.19	0.05	0.12	<0.01	<0.01	**0.01**	**0.01**
Eriogonum umbellatum	0.68	0.08		0.24	0.03		**0.01**	0.16
Ribes amarum	0.32	0.04	0.03	0.00	0.00	0	0.06	**0.01**
Acmispon americanus	0.32	0.03		0.00	0.00		0.08	**0.03**
Asteraceae	0.37	0.02		0.12	<0.01		0.14	0.17
*Ericameria nauseosa*voucher	0.21	<0.01	0.33	0.00	0.00	0.60	NA	NA
Euphorbia maculata	0.00	0.00		0.24	0.02		0.10	0.06
Cercocarpus betuloides	0.16	0.01	<0.01	0.00	0.00	0	0.23	0.29
Mentzelia nitens	0.11	<0.01		0.12	<0.01		0.88	0.71
Salvia columbariae	0.21	<0.01		0.00	0.00		0.20	0.10
Eriastrum densifolium	0.11	<0.01		0.00	0.00		0.44	0.15
Populus	0.05	<0.01		0.00	0.00		0.75	0.74
Yucca brevifolia	0.00	0.00	0.01	0.12	<0.01	0.11	0.34	**0.01**
Leptosyne	0.11	<0.01		0.00	0.00		0.47	0.25
Ephedra	0.11	<0.01	0.11	0.00	0.00	<0.01	0.50	**0.01**
Chenopodium	0.00	0.00		0.06	<0.01		0.72	0.41
Camissonia campestris	0.00	0.00		0.12	<0.01		0.34	0.09
Stephanomeria	0.11	<0.01		0.00	0.00		0.45	0.07
Descurainia pinnata	0.00	0.00		0.06	<0.01		0.75	**0.04**
Cuppressaceae	0.05	<0.01		0.00	0.00		0.76	0.80
Cirsium arvense	0.05	<0.01		0.00	0.00		0.72	0.17

Erodium cicutarium	0.00	0.00	0.06	<0.01	0.75	**0.02**
Claytonia sp	0.11	<0.01	0.00	0.00	0.45	**0.01**
Poaceae	0.00	0.00	0.06	<0.01	0.71	0.27
Lupinus	0.05	<0.01	0.00	0.00	0.80	**0.02**
Ceanothus cordulatus	0.05	<0.01	0.00	0.00	0.80	0.83
Phragmites australis	0.00	0.00	0.06	<0.01	0.71	0.25
Camissonia kernensis	0.00	0.00	0.06	<0.01	0.72	**0.02**
Thysanocarpus	0.05	<0.01	0.00	0.00	0.75	0.45
Scrophularia desertorum	0.05	<0.01	0.00	0.00	0.72	0.16
Bromus	0.00	0.00	0.06	<0.01	0.71	0.29

Table S4: Average values of relative proportion of the 15 common shrubs and trees in the hill and flats habitat shown in Figure S1.

Plant	hill	flats
YUBR	0.011	0.108
RIAM	0.033	0
PUTR	0.001	0
PRFA	0.035	0.1
PI*sp.*	0.052	0.001
HEWH	0.071	0.006
FRCA	0.133	0.013
ERNA	0.329	0.599
ERFA	0.157	0.105
ERCU	0.034	0.005
EPHE	0.108	0.003
CHOL	0	0.001
CEBE	0.002	0
ATCA	0	0.004
ARTR	0.034	0.055

Table S5: Bayesian posterior probabilities (medians and 95% credible intervals) of relative consumption of plants identified in woodrat diets from 2016. Raw relative read counts from trn*L* sequences were used to estimate population level consumption of plants using *bayespref*. The 5 listed plants here comprise > 90% of the total reads of 33 total plants identified. The "Other" category contains the sum of all reads of the remaining 28 plants in dataset.

	Spring		**Summer**		**Combined**	
Plant	***N. bryanti*** **(n = 11)**	***N. lepida*** **(n = 11)**	***N. bryanti*** **(n = 8)**	***N. lepida*** **(n = 5)**	***N. bryanti*** **(n = 19)**	***N. lepida*** **(n = 16)**
Prunus fasciculata	0.05 (0.03, 0.09)	0.54 (0.45, 0.61)	0.06 (0.03, 0.10)	0.65 (0.56, 0.71)	0.05 (0.03, 0.08)	0.58 (0.51, 0.64)
Frangula californica	0.22 (0.14, 0.30)	0.03 (0.01, 0.05)	0.36 (0.26, 0.45)	0.01 (0.00, 0.02)	0.27 (0.20, 0.34)	0.02 (0.01, 0.03)
Phacelia tanacetefolia	0.15 (0.09, 0.23)	0.19 (0.11, 0.26)	0.15 (0.09, 0.23)	0.13 (0.09, 0.19)	0.16 (0.11, 0.21)	0.17 (0.13, 0.22)
Pinus sp.	0.17 (0.10, 0.26)	0.04 (0.02, 0.07)	0.19 (0.12, 0.27)	0.03 (0.01, 0.07)	0.19 (0.13, 0.25)	0.04 (0.02, 0.06)
Eriogonum umbellatum	0.12 (0.07, 0.19)	0.04 (0.02, 0.07)	0.08 (0.04, 0.14)	0.04 (0.02, 0.08)	0.11 (0.07, 0.15)	0.04 (0.02, 0.07)
Other	0.27 (0.18, 0.37)	0.15 (0.10, 0.22)	0.15 (0.09, 0.22)	0.13 (0.08, 0.19)	0.22 (0.16, 0.28)	0.14 (0.11, 0.19)

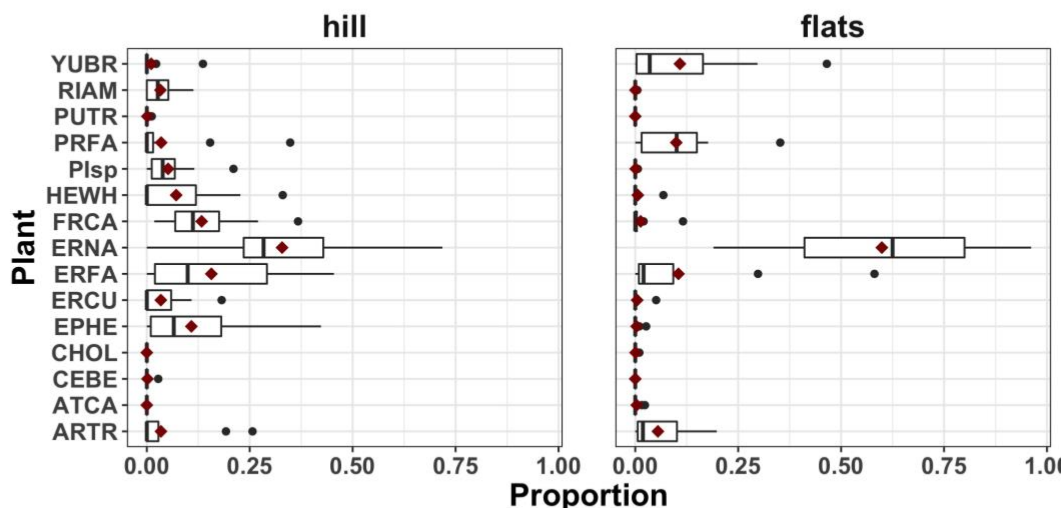

Figure S1: Relative proportion of shrubs and trees present at the study site estimated from counts within plots in both hill and flats. Black bars are median values; dark red diamonds are the mean of each plant in that habitat. Four-letter codes: YUBR – *Yucca brevifolia*; RIAM – *Ribes amarum*; PUTR – *Purshia tridentata*; PRFA – *Prunus fasciculata*; PIsp – *Pinus* sp.; HEWH – *Hesperoyucca whipplei*; FRCA – *Frangula californica*; ERNA – *Ericameria nauseosa*; ERFA – *Eriogonum fasciculatum*; ERCU – *Ericameria cuneata*; EPHE – *Ephedra* sp.; CHOL – *Cholla* sp.; CEBE – *Cercocarpus betuloides;* ATCA – *Atriplex canescens* ; ARTR – *Artemisia tridentata*.

Chapter 2: The gut microbiome reflects ancestry despite dietary shifts across a hybrid zone

Danny P. Nielsen[1,3*] | Joshua G. Harrison[4] | Nathan W. Byer[1] | Trevor M. Faske[2,3] | Thomas L. Parchman[2,3] | W. Brian Simison[5] | Marjorie D. Matocq[1,3]

[1]Department of Natural Resources and Environmental Science; University of Nevada, Reno, Nevada

[2]Department of Biology, University of Nevada, Reno, Nevada

[3]Graduate Program in EECB, University of Nevada, Reno, Nevada

[4]Botany Department, University of Wyoming, Laramie, Wyoming

[5]Center for Comparative Genomics, California Academy of Sciences, San Francisco, California

Accepted: *Ecology Letters* 2022

Abstract

The microbiome is critical to an organism's phenotype, and its composition is shaped by, and a driver of, eco-evolutionary interactions. We investigated how host ancestry, habitat, and diet shape gut microbial composition in a mammalian hybrid zone that occurs across an ecotone between distinct vegetation communities. We found that habitat is the primary determinant of diet, while host genotype is the primary determinant of the gut microbiome—a finding further supported by intermediate microbiome composition in first generation hybrids. Despite these distinct primary drivers, microbial richness was correlated with diet richness, and individuals that maintained higher dietary richness had greater gut microbial community stability. Both relationships were stronger in the relative dietary generalist of the two parental species. Our findings show that host ancestry interacts with dietary habits to shape the microbiome, ultimately resulting in the organismal phenotypic plasticity that host-microbial interactions allow.

Introduction

The gut microbiome contains a diverse community of microorganisms central to host health (Arnolds & Lozupone, 2016; Sender et al., 2016). For mammalian herbivores, the gut microbiome is critical to nutrient acquisition by performing metabolic functions otherwise unavailable to the host (Dearing & Kohl, 2017). Additionally, the functional diversity contained in the gut microbiome provides a source of phenotypic plasticity that is important for host survival and that can drive host evolution (Kolodny & Schulenburg 2020, Moeller & Sanders 2020). Despite general recognition of the importance of the gut microbiome in the ecology and evolution of hosts, little is known of how genetic and environmental factors interact to influence gut microbiome communities, especially in wild animal populations.

Individual-level traits such as diet, seasonal change in diet, host sex, and disease state are known to influence microbiome composition (Amato et al. 2015; Gilbert et al. 2016; Kartzinel et al. 2019). However, host genotype is an important driver of microbial composition that may supersede environmental effects (Knowles et al. 2019, Spor et al. 2011). For example, host phylogeny and microbiome composition often mirror one another (Brucker & Bordenstein 2013), suggesting that animals and their gut microbiomes remain associated over macroevolutionary timescales (Weinstein et al., 2021). However, other studies report prevailing environmental effects in shaping microbial composition (Grieneisen et al., 2019; Grond et al., 2020). Elucidating the relative influences of the environment and host genotype in shaping gut microbiome variation is central to understanding the role of microbial plasticity in individual fitness and dietary adaptation.

Hybrid zones that occur across ecotones provide an ideal arena in which to study the relative influences of environment and host genotype on microbial community composition. Such hybrid zones offer natural laboratories in which to investigate how mismatches between habitat, diet, host genomes, and gut microbial composition may influence individual fitness and rates of hybridization. For example, adaptation to divergent habitats can reinforce reproductive isolation by way of selection against migrants (Nosil et al., 2005; Via, 1999; Via et al., 2000). It is possible that gut microbial mismatches with novel habitats and diets may underlie selection against migrants, yet, to our knowledge, few *in situ* studies have examined microbiome variation in natural hybrid zones (Grieneisen et al., 2019; Lin et al., 2020; Wang et al., 2015).

Over a four-year period, we studied a woodrat hybrid zone that occurs at a sharp habitat transition in southern California between the southern Sierra Nevada and western Mojave Desert (Figure 1; Shurtliff et al., 2014). Here, *Neotoma bryanti* (Bryant's woodrat) occurs primarily within the rocky, and relatively mesic, hill habitat (hereafter referred to as "hill") and *Neotoma lepida* (desert woodrat) occurs primarily within the relatively xeric Mojave Desert scrub (hereafter referred to as "flats"; Figure 1). Here, the two species hybridize and generate a spectrum of F1 and backcross hybrid genotypes that are distributed within approximately ½ km^2 (Jahner et al., 2021; Patton et al., 2007; Shurtliff et al., 2014). The parental species maintain distinct, habitat-specific, and toxic, plant-based diets across the ecotone during both wet and dry seasons (Matocq et al., 2020; Nielsen & Matocq, 2021). *Neotoma bryanti* on the hill maintain a more diverse diet with *Frangula californica* (California coffeeberry) comprising 25% or more of the "hill diet"; while *N. lepida* in the flats consume a less diverse diet with *Prunus fasciculata*

(desert almond) comprising over 50% of the "flats diet" (Matocq et al., 2020, Nielsen & Matocq, 2021). Individuals of pure parental ancestry and parental backcross hybrids are largely spatially segregated (*N. lepida*-like genomes in the flats; *N. bryanti*-like genomes on the hill), and F1 hybrids are distributed throughout the site. However, some *N. lepida*-like and *N. bryanti*-like individuals occupy the alternative habitat (hereafter 'mismatched'). The distribution of genotypes across the genomic spectrum and across habitat types provides an opportunity to quantify the effects of host genotype and host environment on the gut microbiome.

We characterize diet and gut microbiome composition of *N. bryanti, N. lepida*, and hybrids to ask: 1) how do diet and the gut microbiome vary spatially and temporally across the genotypic spectrum between *N. lepida* and *N. bryanti*? 2) what are the relative influences of environment and host genotype on diet and gut microbiome composition? 3) are certain microbial lineages associated with genotypic classes or diet types, and what might that suggest from a metabolic perspective? To address these questions, we use high-throughput sequencing of field collected fecal samples to characterize covariation between diet and microbiome across the genomic spectrum of woodrats at this site. Our study identifies the primary drivers of variation in diet and the gut microbiome and provides insight into the functional significance of differing gut microbial communities.

Methods

Trapping and fecal sample collection

We collected fecal samples from live-trapped woodrats from 2016-2019 in Kelso Valley, Kern Co., California, as part of long-term sampling (Supporting Information Text S1). Here, hybridization occurs between *N. bryanti* and *N. lepida* across a Sierra Nevada – Mojave Desert ecotone within an area approximately ½ km^2 in size (35°25'45 N, 118°15' 2 W; Fig. 1). We grouped samples in two seasonal categories: March-June, spring; and July or later, summer/fall. Animal handling was approved by the University of Nevada Reno Institutional Animal Care and Use Committee, the California Department of Fish and Wildlife, and were consistent with the guidelines developed by the American Society of Mammalogists (Sikes et al., 2016).

Establishing genotypic classes

We extracted DNA from woodrat ear tissue using the Qiagen DNeasy blood and tissue kit as previously described in Nielsen and Matocq (2021). For fecal samples collected from active woodrats nests, we identified genotypes using amplification of microsatellites previously described and used for *N. bryanti* and *N. bryanti* at our study site (Coyner et al., 2015; Nielsen and Matocq, 2021). For woodrats from which ear biopsies were collected, we generated a Single-Nucleotide Polymorphism dataset (SNPs) using a double-digest restriction site associated sequencing (ddRADseq) protocol previously used in this hybrid zone (Jahner et al. 2021; Parchman et al. 2012; Peterson et al., 2012). Reads were aligned to the *N. lepida* genome (Greenhalgh et al., 2022). We used Stacks v. 2.53 to identify SNPs and call genotypes for each locus (Catchen et al.,

2013), and genotype and population structure was inferred using FastSTRUCTURE (Raj et al., 2014; see S.I. Text S2).

Rather than treat genetic variation as a continuous variable, we categorized individuals into genotypic classes based on the proportion of their genome assigned to *N. bryanti* ($q_{bryanti}$) using K = 2 as follows: > 0.90 = *N. bryanti*, 0.90-0.60 = BC-*bryanti*, 0.60-0.40 = F1, 0.40-0.10 = BC-*N. lepida*, and < 0.10 = *N. lepida*. We do not classify advanced hybrids as previous analysis found no evidence of F2 or advanced generation hybrids (Jahner et al. 2021). Four individuals were removed from the STACKS analysis due to low coverage and were assigned genotypes using microsatellites as described above.

Metagenomic data for characterizing diet and microbiome

We submitted trap collected fecal samples to Jonah Ventures LLC (Boulder, CO) for DNA extraction, PCR amplification, and sequencing. DNA for both *trnL* and 16S was sequenced on the Illumina MiSeq platform (San Diego, CA) using the v2 500-cycle kit. Methods of DNA extraction, amplification, and sequencing for diet followed Nielsen and Matocq (2021). We removed plants from the dataset that fell below a 5% relative read abundance threshold. We filtered out reads originating from bait (genus *Arachis*). After filtering and taxonomic identification, we retained 4,015,325 reads across 46 plant taxa. We restricted our analyses to the five most common diet plants identified previously: *Prunus fasciculata*, *Frangula californica*, *Eriogonum*, *Pinus*, and *Phacelia*. The remaining plant reads were grouped together as an ‘other’ category (S.I. Table S5 for full plant list).

Microbiome composition was characterized using the 515f and 806R 16S rRNA primers (Caporaso et al., 2011). 16S sequences were processed using the standard MOTHUR SOP pipeline (Kozich et al., 2013, assessed Oct. 25, 2020). Sequences from mitochondria, chloroplast, archaea, and eukaryota were excluded using the *remove.lineage* command. 16S sequences were clustered into operational taxonomic units (OTUs) based on 97% similarity and taxonomy was assigned using the SILVA reference database (accessed Oct. 25, 2020). This initial dataset, including singletons, contained 3,552,497 reads across 242 unique samples and was used to calculate microbial richness (see below). Use of trap collected feces for studying the gut microbiome has been validated (Kohl et al., 2015); but to confirm that environmental contamination did not contribute to overall patterns in microbial composition of fecal samples, we conducted a validation study (see S.I. Text S1). A negative lab control was included, and we removed any OTU for which 5% or more of its reads was contained within the blank.

We used a hierarchical Bayesian approach implemented with CNVRG to model the proportion of both plants and microbes within samples (Harrison et al., 2020; details in S.I. Text S3). The resulting proportional dataset included 3,655 microbial OTUs across 242 unique individual woodrats. We calculated Bray-Curtis distances from the resulting diet and microbiome proportional data. Finally, we also used CNVRG multinomial estimates to infer differentially abundant microbial taxa (S.I. Text S3).

Variation in diet and microbiome diversity

We estimated Shannon diversity of diet and microbiome across genotypes, habitats, and spring and summer seasons using the phyloseq package in R (McMurdie &

Holmes, 2013). To assess differences in diversity in diet and microbiome across genotypes, habitats, and spring and summer, we performed ANOVA using the *aov* function in the stats library and using Shannon's diversity index. We calculated richness for diet and microbiome as the total number of plants and microbial (including singletons) OTUs observed in a sample, respectively. We tested whether diet and microbial richness and distance were correlated within woodrat genotypes as well as within individuals sampled multiple times (S.I. Text S4).

Quantifying host and environmental effects on diet and microbiome

We estimated seasonal diet and microbiome turnover among genotypic classes and across habitat as the average Bray-Curtis distance between individual samples within a group from spring to summer. We conducted a principal components analysis (PCA) using square-root transformed read counts from diet data with the *prcomp* function. We visualized microbiome and diet composition using principal coordinates analysis (PCoA) and PCA, respectively. We assessed whether F1 hybrid microbiome composition was intermediate by comparing pairwise Bray-Curtis distances among conspecific (e.g., *N. bryanti* or *N. lepida* comparisons), heterospecific (i.e. *N. bryanti – N. lepida* comparisons), or either parental species versus F1 hybrids (see S.I. Text S8).

To guide dimensionality reduction, we calculated Pearson's correlation coefficients (Pearson's *r*) between host and environmental variables (i.e., genotype, habitat, season, year) and the first four axes of both diet (PCA) and microbiome (PCoA; S.I. Text S5). We estimated the amount of variance explained in diet and microbiome composition by individual host and environmental variables using partial distance-based

redundancy analysis (dbRDA) using the *dbrda* function in vegan (Oksanen, 2020; see S.I. Text S6). We further explored the individual and combined contributions of habitat, genotype, and either diet PC1 or microbiome PCoA2 on overall variation in diet and microbiome composition using variance partitioning (*varpart* in vegan).

To determine if microbial community composition was significantly associated with the most common diet plants, we implemented a constrained analysis of principal coordinates (CAP) with the *ordinate* function in phyloseq using the model: *OTU ~ Prunus fasciculata + Frangula californica + Eriogonum + Condition (Year + Age + Sex)*. We used an ANOVA with 999 permutations to test for model significance. We conducted analyses in R (R Core Team 2020).

Results

Genotypic variation among sampled individuals

After filtering and removal of loci, our genomic dataset contained 154,022 SNPs. Of 242 unique individuals sampled between 2016-2019, we identified 83 *N. bryanti*, 14 BC-*bryanti*, 22 F1 hybrids, 28 BC-*lepida*, and 95 *N. lepida*. Sample sizes within genotype, habitat, and seasonal groups are reported in Figure 2 (panels A & D).

Diet varies across habitat and among genotypes

We collected a total of 334 fecal samples from 242 individuals from 2016-2019. Overall, diet and microbiome were distinguishable among genotypic classes and across habitat types (Fig. 1 B&C). The three most abundant plants in spring and summer diets were: *P. fasciculata* (desert almond) predominantly in the flats diet, *Frangula californica* (Californica coffeeberry) in the hill diet, and *Eriogonum* (buckwheat) was consumed in

both habitats (Figs. 1B & 2). *Phacelia*, a spring forb, was consumed in both habitats when available (Fig. 2a). Diet diversity differed among genotypes (df = 4, F = 6.5, $P < 0.001$) and across habitats (df = 1, F = 4.7, $P = 0.03$; Fig. 2B). Seasonal turnover of *N. lepida* diet in the flats was significantly lower than that of *N. bryanti* on the hill ($P < 0.05$; Fig. 2C).

Microbiome varies among genotypes and across habitats

Microbiome composition varied among genotypic classes and habitat types, with F1 hybrids exhibiting a microbial community composition intermediate to pure parental individuals (Figs. 1C & Fig. S5). Firmicutes and Bacteroidetes were the most abundant microbial phyla across samples. Common microbial families included Lactobacillaceae, Porphyromonadaceae, Lachnospiraceae, Erysipelotrichaceae, Ruminococcaceae, and Bifidobacteriaceae (Fig. 2d). Shannon diversity of the microbial community did not differ among genotypic classes, but did differ across seasons (df = 1, F = 11.3, $P < 0.001$; Fig. 2E). Seasonal turnover of the gut microbiome community was greater in *N. bryanti* than *N. lepida* ($P < 0.05$; Fig. 2F).

Across the entire dataset microbiome richness was positively correlated with dietary richness in ($P = 0.002$; Fig. 3A). When evaluated individually, only *N. bryanti* and BC-*bryanti* on the hill exhibited a positive relationship between diet and microbiome richness (R = 0.3, 0.64; $P < 0.01$, 0.05; Fig. 3B). There was a significant linear relationship between diet and microbiome distance across the entire dataset (Mantel, $r = 0.13$, $P = 0.001$; Fig. S3), but no linear relationship within individual genotypes. Among individuals sampled multiple times, *N. lepida* exhibited a trend toward a positive

relationship between diet and microbiome distance (Spearman; $r = 0.4$, $P = 0.08$) while *N. bryanti* exhibited a negative trend (Spearman; $r = -0.4$, $P = 0.07$; Fig. 3C). Diet richness was negatively correlated with microbiome distance in *N. bryanti* individuals sampled multiple times (Spearman; $r = -0.65$, $P = 0.003$; Fig. 3D).

Drivers of diet and microbiome composition

Habitat and genotype were correlated with diet PC1 which explained 43% of variation; whereas habitat and genotype were correlated with axis 2 of the microbiome PCoA which explained 9.4% of variation (Tables S2&3). Based on these results, we used diet PC1 and microbiome PCoA2 as variables in further analyses (see below).

Partial distance-based redundancy analysis (dbRDA) revealed that habitat explained the greatest amount of variation in diet, while host genotype explained the greatest amount of variation in the gut microbiome (Table S4). See supporting information for more detail.

Variance partitioning provided further insight into the relative contributions of genotype, habitat, diet, and microbiome on variation in diet and microbiome across this ecotone (Table 1). For this analysis and based on results summarized above (Tables S2&3), we simplified diet and microbiome to values along PC1 and PCoA2, respectively. Habitat (*PVE* = 68.6%) explained more variation in diet than genotype (*PVE* = 49.7%) or microbiome (PCoA2; *PVE* = 54.5%). When the partial contributions of these variables were modeled alone, habitat (adj. $r^2 = 15.5\%$, Table 1) explained the most variation in diet while microbiome (PCoA2) only explained 1.4% and genotype did not explain any further variation. In contrast, microbiome appears to be primarily constrained by

ancestry. Genotype (*PVE* = 70.1%) explained more variation in the microbiome than habitat (*PVE* = 54.5%) and diet (PC 1; *PVE* = 54.5%). When removing the effects of the other variables genotype (adj. r^2 = 15%) still explained the most variation in microbiome while diet PC1 only explained 1.3%, and habitat did not explain any further variation. To verify that these results were not biased due to dimension reduction, we confirmed the order of variable importance by performing variance partitioning using all diet and microbiome dimensions.

Microbiome community composition was significantly associated with the most common diet plants, with *F. californica* and *Eriogonum* exhibiting strong associations with microbiome composition of *N. bryanti* on the hill, and *P. fasciculata* significantly associated with *N. lepida* microbiome composition in the flats (Fig. 4). BC-*lepida* individuals exhibited a *P. fasciculata*-associated microbiome like their parental counterparts in the flats. F1 hybrids exhibited a range of diet by microbiome associations, but an overall intermediate microbiome composition in comparison to pure-bred individuals (Figs. 4, 5, S5).

Mismatched individuals exhibit more variable diets and microbiomes

Individuals occupying the "mismatched" habitat (i.e., *N. bryanti* in flats, *N. lepida* on hill) exhibited reduced preference for the plant most consumed in that habitat. For instance, the relatively rare *N. bryanti* in the flats consumed some *P. fasciculata* (FOO = 42.8%, RRA = 36.8%), but much less in comparison to *N. lepida* in the flats (FOO = 94%, RRA=70.7%). Likewise, the rare *N. lepida* on the hill consumed some *F. californica* (FOO = 60%, RRA = 17.1%), but less than *N. bryanti* on the hill (FOO =

88.7%, RRA=40.3%). *Neotoma bryanti* in the flats that consumed more *P. fasciculata* exhibited a more intermediate microbiome than those that consumed a more hill-like diet (Fig. 5). *Neotoma lepida* on the hill consumed more diverse diets, including increased consumption of *Eriogonum*, and exhibited variable microbiome composition (Fig. 5). Backcross and F1 hybrids primarily consumed habitat-specific plants and exhibited more intermediate microbiome composition (Figs. 5 & S5).

Differential abundance of microbial taxa

More than 80% of differentially abundant taxa between *N. bryanti* and *N lepida* belonged to Bacteroidetes and Firmicutes (CNVRG analysis; S.I. Fig. S4.). *Lactobacillus* were more abundant in woodrats consuming a *P. fasciculata*-dominated diet, and in *N. lepida* and F1 hybrids in the flats (Fig. S6). We also detected microbial lineages that are expected to modify hydrogen cyanide, including members of the Pseudomonadaceae (Zhu et al., 2018). For complete results and a list of genera that differed with greater than 95% probability is provided in the supporting information (Tables S5-S8).

Discussion

The distribution of *N. bryanti, N. lepida*, and their hybrids across an ecotone allowed us to investigate the individual and joint effects of environment, host genotype, and diet on gut microbial composition. We found that habitat-specific diets are accompanied by distinct microbial communities and that, at the individual level, microbial diversity is correlated with diet diversity. Nonetheless, we found that diet is

most influenced by habitat, while microbial composition is primarily determined by host genotype. The latter of the two findings was further supported by our observation that admixed genomes were more likely to harbor microbial communities that were intermediate to those typically associated with pure parental genomes, regardless of diet (Figs. 1, 4 & 6). Our findings suggest that gut microbiome composition in woodrats is primarily driven by host genotype, yet within that overall constraint, individual variation in diet is accompanied, and could be facilitated (Kohl et al., 2014), by gut microbial community diversity.

Relationship between diet and gut microbial community composition

We found a strong signature of habitat-specific diets within parental species, their respective backcross hybrids, and F1 hybrids. *Frangula californica* and *P. fasciculata* were the most abundant diet plants in the hill and flats, respectively, and consumption of these plants was maintained across seasons. Given that both plants are known to contain compounds that can be toxic to mammals—anthraquinones in *F. californica* and cyanogenic glycosides in *P. fasciculata* (Matocq et al., 2020; Qin et al., 2016; Vetter, 2000)—we would expect these animals to diversify their diets when given the seasonal opportunity to shift away from these toxins (Nielsen and Matocq 2021). Diet turnover across seasons was significantly higher for *N. bryanti* in the hill habitat than *N. lepida* in the flats, suggesting *N. bryanti* is more of a dietary generalist than *N. lepida*—a result consistent with previous studies (Nielsen and Matocq 2021). Further, *N. lepida* maintains a high proportion of *Prunus* in their diet even in spring when a higher diversity of plants

become available, consistent with findings that this species is a facultative specialist (Nielsen and Matocq 2021, Shipley et al. 2009).

We found that diet and microbial richness were positively correlated, a pattern evident at broad spatial scales across populations and species of woodrats (Weinstein et al., 2021), but also evident at the fine spatial scale of this study. The positive relationship between diet and microbial richness could be the result of multiple factors. It is possible that consuming a more nutritionally and chemically diverse diet requires or results in a more functionally diverse microbiome (Heiman & Greenway, 2016), consistent with the expectation of close ecological interactions between the gut microbiome and specific dietary components (Kartzinel et al., 2019; Knowles et al., 2019; Ren et al., 2017). Alternatively, microbial lineages detected in the feces of woodrats may have been associated with the plants consumed and not persistent members of the gut microbiome (Kohl et al. 2014). Nonetheless, even transient plant-associated microbial lineages can contribute to metabolic functioning in the intestinal microflora (Zeibich et al., 2019), and may augment the functional capacity of host gut microbes through horizontal gene transfer (Hehemann et al., 2010; Wybouw et al., 2014). However, when broken down by genotype alone, diet and microbiome richness were only significantly positively correlated in *N. bryanti* and BC-*bryanti*. As these genotypic classes are those that exhibit a more generalist dietary strategy, this result supports the supposition that microbial diversity is at least partly driven by feeding strategy (Reese & Dunn, 2018).

Given the broad concordance between diet and microbiome richness and the close ecological association this relationship suggests, we anticipated that changes in diet would be correlated with changes in microbiome composition, resulting in a positive

relationship between individual pairwise diet distances and their gut microbial distance. We found this relationship to be significantly positive for the entire dataset, but not when evaluated by individual genotypic class (Fig. S3). For individuals resampled across time, allowing further examination of whether changes in diet are correlated with changes in the microbiome, we found a positive trend within *N. lepida*, and a negative trend in *N. bryanti*. The relative dietary specialist, *N. lepida*, may have a gut microbial community more tuned to a low diversity, albeit toxic, diet and when these diets shift, more microbial turnover occurs. On the other hand, the relative dietary generalist *N. bryanti* maintains a more diverse diet and microbiome, the latter of which may have the capacity to metabolize new dietary components without a compositional shift. This would lead to the expectation that *N. bryanti* individuals with the most diverse diets may exhibit greater stability (less turnover) in their microbial composition from one sampling point to the next, and that is indeed what we observe (Fig 3D). As seen in humans (Johnson et al. 2019), *N. bryanti* individuals with the most diverse diets appear to have the greatest stability in their microbial community composition. The potential relationships between specialists and generalists and their respective microbial communities warrants further investigation. Overall, though, despite the associations we detected between diet and microbiome, both appear to be primarily driven by different factors.

Influence of environment and host genotype on diet and gut microbial composition

Diet and microbiome composition were differentially influenced by each other, host genotype and habitat (Tables 1 & S4). In this system, diet is influenced most by habitat, then moderately by an individual's microbiome, with no additional variation

explained by host genotype. Conversely, the microbiome is influenced most by genotype, then moderately by diet, with no additional variation explained by habitat.

The importance of genotype, or individual ancestry, in shaping microbiome composition is further supported by the microbial communities that characterize F1 hybrids. With half the *N. lepida* genome and half the *N. bryanti* genome, we might expect the microbiomes of F1 hybrids to be intermediate between parental types regardless of habitat diet. Indeed, F1 hybrids exhibited intermediate microbiome composition, particularly in the flats (Figs. S5 & 5). This pattern was evident even though most F1 individuals inhabiting the flats eat a *Prunus*-rich diet, and most on the hill eat a characteristic hill diet. Further, although we cannot confirm maternal genotype of individuals, hybridization at this site is thought to primarily occur via female *N. bryanti* mating with male *N. lepida* (Shurtliff et al., 2013). Given the importance of maternally inherited microbes in mammalian microbiomes (Funkhouser & Bordenstein, 2013), it is possible that the pronounced intermediacy of microbiome composition of F1 hybrids in the flats is due to having *N. bryanti* mothers from which they inherit a more “hill-like” microbiome. The primacy of host genotype as a driver of microbiome composition has recently been shown at broad spatial scales in woodrats (Weinstein et al. 2021), but ours is the first investigation to support these findings with hybrid individuals and at a fine spatial scale. If intermediate microbial communities allow these individuals to have greater flexibility in habitat association, this could be an important mechanism determining rates of hybridization in this system.

In our study, genotype and habitat were highly correlated due to the strong spatial segregation of parentals and their respective backcrosses. However, although rare,

parental and backcross individuals that occupy the mismatched habitat provide a decoupling of the dominant habitat and genotype association at this site. For *N. bryanti*, the relative dietary generalist, some individuals (N = 3; Fig. 5) that occupy the flats maintain a characteristic *Prunus* diet and their microbiome is intermediate between hill- and flats-like microbial communities, while other *N. bryanti* (N = 4) maintain a more diverse or hill-like diet and microbiome all while living in the flats. On the other hand, mismatched *N. lepida* on the hill consumed little *Prunus* and shifted their diet to a more hill-like diet, albeit dominated by *Eriogonum* rather than *Frangula*. However, these *N. lepida* genomes, appear to sustain a wide range of microbial communities including those with intermediate composition, and even one example of a hill-like community. It appears that at least some *N. lepida* and *N. bryanti* can diversify their diets and sustain themselves in the alternate habitat. *Neotoma lepida* primarily rely on *Eriogonum* and other plants on the hill, while some *N. bryanti* can consume large amounts of *Prunus*, as might be expected for this facultative generalist. Many of these mismatched individuals have microbial communities that are intermediate in composition in comparison to their counterparts in their native habitat. This result suggests that the microbiome is not absolutely constrained by host genotype, and that plastic response to environmental conditions can occur.

Ecological and evolutionary implications

Our study adds to a small number of studies that investigate the relative role of host genotype and host environment in shaping the gut microbiome across mammalian hybrid zones (Grieneisen et al., 2019; Lin et al., 2020; Wang et al., 2015). There is

complexity in the diet and microbiome datasets stemming from temporal variation (Tables S2-S4), as well as individual-level variation evident in high baseline Bray-Curtis distances within individuals sampled multiple times (Figure 3C-D). However, we still detected signals from habitat and genotype which allowed us to explore the relative contributions of these variables in shaping diet and microbiome composition. Diet was most influenced by habitat; and, consistent with other studies, we find that host genotype is the primary driver of microbial composition (Knowles et al., 2019; Weinstein et al., 2021). Additionally, dietary richness and composition was associated with microbial richness, composition, and stability, but the nature of these relationships may differ between dietary specialists and generalists. Among differentially abundant microbial taxa detected, some may be functionally important in this system. *Lactobacillus* was strongly associated with *N. lepida* and F1 hybrids in the flats, as well as diets composed of greater than 50% *P. fasciculata*. Some *Lactobacillus* species are known to metabolize cyanogenic glycosides (Lei et al., 1999), which have been found in *P. fasciculata* at this site (Matocq et al., 2020). Presence of specific microbial taxa, even lineages in low abundance, may be critical to individual fitness across this hybrid zone.

This hybrid zone is characterized by strong ecological segregation, with across-habitat dispersal occurring in only ~4% of captured individuals, despite spatial distances between these habitats being within individual dispersal capabilities (Shurtliff et al. 2014). Many mismatched individuals also took on a mismatched diet which appeared to be accompanied by a shift in their microbiome. Individual variation in dietary and gut microbial plasticity can influence the ability to acclimate to new habitats (Alberdi et al., 2016), which may have important ecological and evolutionary consequences (Vander

Zanden et al., 2010), including the facilitation or encumbrance of gene flow. Individuals may persist in the mismatched habitat if they host lineages that aid in the metabolism of toxic plant compounds. Further genomic and demographic study is needed to identify the host metabolic and microbial traits that determine survival and reproductive success of these mismatched individuals and the potential for interspecific gene flow they create.

Acknowledgments

We wish to thank D. Moore, A. Mitchell, J. Wurtz, and M. Jee for assistance in the field. We also thank M. Whitfield at the Southern Sierra Research Station and R. Tollefson at the Audubon Kern River Preserve for hospitality and use of resources while in the field. We thank L. Galland and K. Graski for assistance in the lab; J. Jahner and J. Craine for helpful bioinformatics discussions; and J. Patton, J. Forbey, and D. Dearing for helpful woodrat discussions. We thank M. Forister, C. Feldman and members of the UNR Evoldoers lab group for comments on earlier drafts of this paper. This research was funded in part by National Science Foundation grants to MDM (IOS-1457209 and OIA-1826801) and a Society of Integrative and Comparative Biology Grant-in-Aid of Research grant to DPN.

References

Alberdi, A., Aizpurua, O., Bohmann, K., Zepeda-Mendoza, M. L., & Gilbert, M. T. P. (2016). Do Vertebrate Gut Metagenomes Confer Rapid Ecological Adaptation? *Trends in Ecology & Evolution*, *31*(9), 689–699. https://doi.org/10.1016/j.tree.2016.06.008

Amato, K. R., Leigh, S. R., Kent, A., Mackie, R. I., Yeoman, C. J., Stumpf, R. M., Wilson, B. A., Nelson, K. E., White, B. A., & Garber, P. A. (2015). The gut microbiota appears to compensate for seasonal diet variation in the wild black howler monkey (Alouatta pigra). *Microbial Ecology*, *69*(2), 434–443.

Arnolds, K. L., & Lozupone, C. A. (2016). Striking a Balance with Help from our Little Friends – How the Gut Microbiota Contributes to Immune Homeostasis. *The Yale Journal of Biology and Medicine*, *89*(3), 389–395.

Brucker, R. M., & Bordenstein, S. R. (2013). The hologenomic basis of speciation: Gut bacteria cause hybrid lethality in the genus Nasonia. *Science*, *341*(6146), 667–669.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, *108*(Supplement 1), 4516–4522.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Coyner, B. S., Murphy, P. J., & Matocq, M. D. (2015). Hybridization and asymmetric introgression across a narrow zone of contact between Neotoma fuscipes and N. macrotis (Rodentia: Cricetidae). *Biological Journal of the Linnean Society*, *115*(1), 162–172. https://doi.org/10.1111/bij.12487

Dearing, M. D., & Kohl, K. D. (2017). Beyond fermentation: Other important services provided to endothermic herbivores by their gut microbiota. *Integrative and Comparative Biology*, *57*(4), 723–731.

Funkhouser, L. J., & Bordenstein, S. R. (2013). Mom knows best: The universality of maternal microbial transmission. *PLoS Biol*, *11*(8), e1001631.

Gilbert, J. A., Quinn, R. A., Debelius, J., Xu, Z. Z., Morton, J., Garg, N., Jansson, J. K., Dorrestein, P. C., & Knight, R. (2016). Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature*, *535*(7610), 94–103.

Greenhalgh, Robert, Matthew L. Holding, Teri J. Orr, James B. Henderson, Thomas L. Parchman, Marjorie D. Matocq, Michael D. Shapiro, and M. Denise Dearing. “Trio-Binned Genomes of the Woodrats *Neotoma bryanti* and *Neotoma lepida* Reveal Novel Gene Islands and Rapid Copy Number Evolution of Xenobiotic Metabolizing Genes.” *Molecular Ecology Resources* n/a, no. n/a. Accessed May 24, 2022. https://doi.org/10.1111/1755-0998.13650.

Grieneisen, L. E., Charpentier, M. J., Alberts, S. C., Blekhman, R., Bradburd, G., Tung, J., & Archie, E. A. (2019). Genes, geology and germs: Gut microbiota across a primate hybrid zone are explained by site soil properties, not host species. *Proceedings of the Royal Society B*, *286*(1901), 20190431.

Grond, K., Bell, K. C., Demboski, J. R., Santos, M., Sullivan, J. M., & Hird, S. M. (2020). No evidence for phylosymbiosis in western chipmunk species. *FEMS Microbiology Ecology*, *96*(1), fiz182. https://doi.org/10.1093/femsec/fiz182

Harrison, J. G., Calder, W. J., Shastry, V., & Buerkle, C. A. (2020). Dirichlet-multinomial modelling outperforms alternatives for analysis of microbiome and other ecological count data. *Molecular Ecology Resources*, *20*(2), 481–497. https://doi.org/10.1111/1755-0998.13128

Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., & Michel, G. (2010). Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, *464*(7290), 908–912. https://doi.org/10.1038/nature08937

Heiman, M. L., & Greenway, F. L. (2016). A healthy gastrointestinal microbiome is dependent on dietary diversity. *Molecular Metabolism*, *5*(5), 317–320. https://doi.org/10.1016/j.molmet.2016.02.005

Jahner, J. P., Parchman, T. L., & Matocq, M. D. (2021). Multigenerational backcrossing and introgression between two woodrat species at an abrupt ecological transition. *Molecular Ecology*. https://doi.org/10.1111/mec.16056

Johnson, A. J., Vangay, P., Al-Ghalith, G. A., Hillmann, B. M., Ward, T. L., Shields-Cutler, R. R., Kim, A. D., Shmagel, A. K., Syed, A. N., Walter, J., Menon, R., Koecher, K., & Knights, D. (2019). Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. *Cell Host & Microbe*, *25*(6), 789-802.e5. https://doi.org/10.1016/j.chom.2019.05.005

Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R., & Pringle, R. M. (2019). Covariation of diet and gut microbiome in African megafauna. *Proceedings of the National Academy of Sciences*, *116*(47), 23588–23593.

Knowles, S. C. L., Eccles, R. M., & Baltrūnaitė, L. (2019). Species identity dominates over environment in shaping the microbiota of small mammals. *Ecology Letters*, *22*(5), 826–837.

Kohl, K. D., Luong, K., & Dearing, M. D. (2015). Validating the use of trap-collected feces for studying the gut microbiota of a small mammal (Neotoma lepida). *Journal of Mammalogy*, *96*(1), 90–93.

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Dearing, M. D. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology Letters*, *17*(10), 1238–1246. https://doi.org/10.1111/ele.12329

Kolodny, O., & Schulenburg, H. (2020). Microbiome-mediated plasticity directs host evolution along several distinct time scales. *Philosophical Transactions of the Royal Society B*, *375*(1808), 20190589.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, *79*(17), 5112–5120.

Lei, V., Amoa-Awua, W. K. A., & Brimer, L. (1999). Degradation of cyanogenic glycosides by Lactobacillus plantarum strains from spontaneous cassava fermentation and other microorganisms. *International Journal of Food Microbiology*, *53*(2), 169–184. https://doi.org/10.1016/S0168-1605(99)00156-7

Lin, D., Lacey, E. A., Bach, B. H., Bi, K., Conroy, C. J., Suvorov, A., & Bowie, R. C. (2020). Gut microbial diversity across a contact zone for California voles: Implications for lineage divergence of hosts and mitonuclear mismatch in the assembly of the mammalian gut microbiome. *Molecular Ecology*, *29*(10), 1873–1889.

Matocq, M. D., Ochsenrider, K., Jeffrey, C. S., Nielsen, D., & Richards, L. A. (2020). Fine-scale differentiation in diet and metabolomics of small mammals across a sharp ecological transition. *Frontiers in Ecology and Evolution*, *8*, 282.

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.

Moeller, A. H., & Sanders, J. G. (2020). Roles of the gut microbiota in the adaptive evolution of mammalian species. *Philosophical Transactions of the Royal Society B*, *375*(1808), 20190597.

Nielsen, D. P., & Matocq, M. D. (2021). Differences in dietary composition and preference maintained despite gene flow across a woodrat hybrid zone. *Ecology and Evolution*, *11*(9), 4909–4919. https://doi.org/10.1002/ece3.7399

Nosil, P., Vines, T. H., & Funk, D. J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, *59*(4), 705–719.

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., Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package* (2.5-7) [Computer software]. https://CRAN.R-project.org/package=vegan

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Patton, J. L., Huckaby, D. G., & Álvarez-Castañeda, S. T. (2007). *The evolutionary history and a systematic revision of woodrats of the Neotoma lepida group* (Vol. 135). Univ of California Press.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLOS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Qin, B., Xu, Y., Chen, J., Huang, W., Peng, Y., & Zheng, J. (2016). Chemical Reactivity of Emodin and Its Oxidative Metabolites to Thiols. *Chemical Research in Toxicology*, *29*(12), 2114–2124. https://doi.org/10.1021/acs.chemrestox.6b00191

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Reese, A. T., & Dunn, R. R. (2018). Drivers of Microbiome Biodiversity: A Review of General Rules, Feces, and Ignorance. *MBio*, *9*(4). https://doi.org/10.1128/mBio.01294-18

Ren, T., Boutin, S., Humphries, M. M., Dantzer, B., Gorrell, J. C., Coltman, D. W., McAdam, A. G., & Wu, M. (2017). Seasonal, spatial, and maternal effects on gut microbiome in wild red squirrels. *Microbiome*, *5*(1), 1–14.

Sender, R., Fuchs, S., & Milo, R. (2016). Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell*, *164*(3), 337–340.

Shipley, L. A., Forbey, J. S., & Moore, B. D. (2009). Revisiting the dietary niche: When is a mammalian herbivore a specialist? *Integrative and Comparative Biology*, *49*(3), 274–290. https://doi.org/10.1093/icb/icp051

Shurtliff, Q. R., Murphy, P. J., & Matocq, M. D. (2014). Ecological segregation in a small mammal hybrid zone: Habitat-specific mating opportunities and selection against hybrids restrict gene flow on a fine spatial scale. *Evolution*, *68*(3), 729–742.

Shurtliff, Q. R., Murphy, P. J., Yeiter, J. D., & Matocq, M. D. (2013). Experimental evidence for asymmetric mate preference and aggression: Behavioral interactions in a woodrat (Neotoma) hybrid zone. *BMC Evolutionary Biology*, *13*(1), 220. https://doi.org/10.1186/1471-2148-13-220

Sikes, R. S., Care, A., & Mammalogists, U. C. of the A. S. of. (2016). 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *Journal of Mammalogy*, *97*(3), 663–688.

Spor, A., Koren, O., & Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology*, *9*(4), 279–290.

Vander Zanden, H. B., Bjorndal, K. A., Reich, K. J., & Bolten, A. B. (2010). Individual specialists in a generalist population: Results from a long-term stable isotope series. *Biology Letters*, *6*(5), 711–714. https://doi.org/10.1098/rsbl.2010.0124

Vetter, J. (2000). Plant cyanogenic glycosides. *Toxicon*, *38*(1), 11–36. https://doi.org/10.1016/S0041-0101(99)00128-2

Via, S. (1999). Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, *53*(5), 1446–1457.

Via, S., Bouck, A. C., & Skillman, S. (2000). Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution*, *54*(5), 1626–1637.

Wang, J., Kalyan, S., Steck, N., Turner, L. M., Harr, B., Künzel, S., Vallier, M., Häsler, R., Franke, A., & Oberg, H.-H. (2015). Analysis of intestinal microbiota in hybrid house mice reveals evolutionary divergence in a vertebrate hologenome. *Nature Communications*, *6*(1), 1–10.

Weinstein, S. B., Martínez-Mota, R., Stapleton, T. E., Klure, D. M., Greenhalgh, R., Orr, T. J., Dale, C., Kohl, K. D., & Dearing, M. D. (2021). Microbiome stability and structure is governed by host phylogeny over diet and geography in woodrats (*Neotoma* spp.). *Proceedings of the National Academy of Sciences*, *118*(47), e2108787118. https://doi.org/10.1073/pnas.2108787118

Wybouw, N., Dermauw, W., Tirry, L., Stevens, C., Grbić, M., Feyereisen, R., & Van Leeuwen, T. (2014). A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. *ELife*, *3*, e02365. https://doi.org/10.7554/eLife.02365

Zeibich, L., Schmidt, O., & Drake, H. L. (2019). Fermenters in the earthworm gut: Do transients matter? *FEMS Microbiology Ecology*, *95*(2). https://doi.org/10.1093/femsec/fiy221

Table 1: Variance partitioning of the individual and combined influences of habitat and genotype on variation in diet PCA axis 1 (explained 43.0% of total variation; Fig. 1 panel B) and microbiome PCoA axis 2 (explained 9.4% of total variation; Fig. 1 panel C). Respective axes of variation for diet and microbiome were also included in models for each other. Together, these three variables explained 72.2% of diet variation (PC1) and 74.5% of microbiome variation (PCoA 2). Adjusted r^2 values represent the amount of variation each variable explained after removing the effect of the other variable.

Model	Adjusted r^2 (%)	Proportion of Variance Explained (%)
Diet ~ Habitat + Genotype + Microbiome PCoA 2	46.0—	72.2*
Diet ~ Habitat	15.5*	68.6*
Diet ~ Genotype	0.2	49.7*
Diet ~ Microbiome PCoA 2	1.4*	54.5*
Residuals	28.0	
16S ~ Habitat + Genotype + Diet PC 1	47.0—	74.5*
16S ~ Habitat	0.2	54.5*
16S ~ Genotype	15.0*	70.1*
16S ~ Diet PC 1	1.3*	54.5*
Residuals	25.5—	

* model significance $P \leq 0.001$)
—cannot be tested

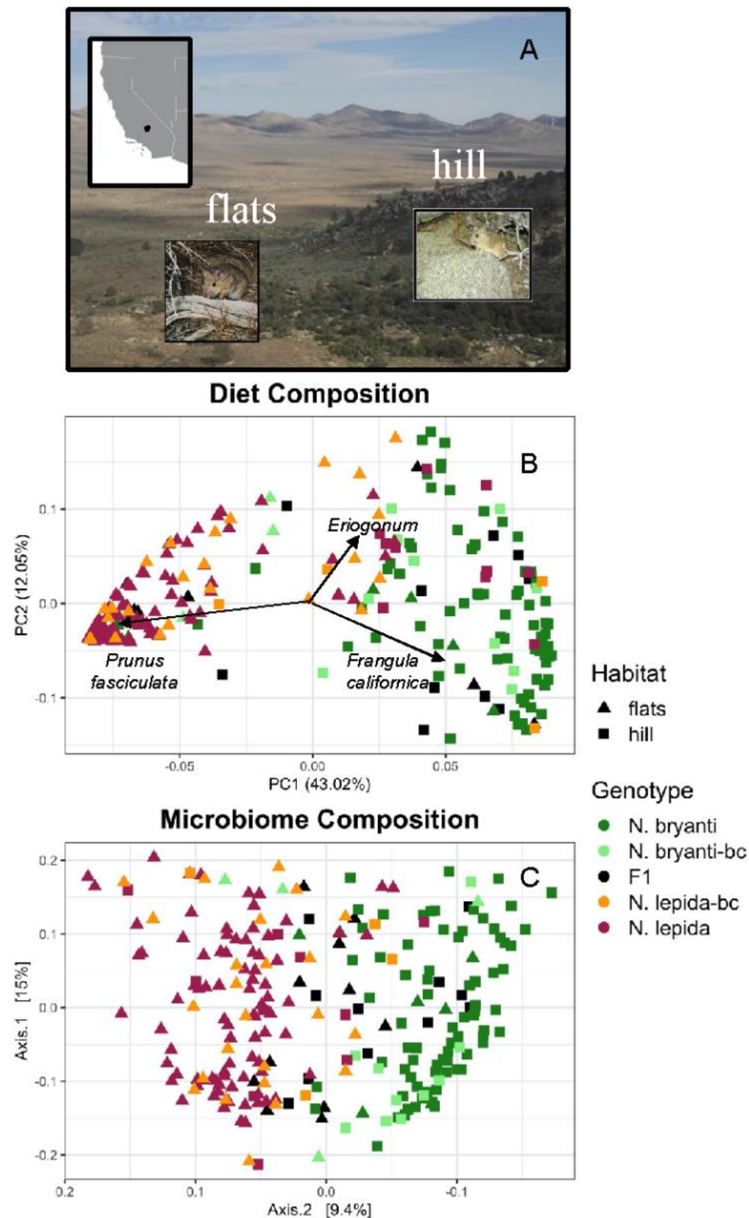

Figure 1: The study site in southern California, where the mesic hill habitat transitions to the xeric flats. Photo taken from north looking southeast. Inset photos of *N. lepida* in flats, and *N. bryanti* on hill (Panel A). Diet composition was largely distinct among genotypic classes and across habitats (Panel B). Loadings in Panel B are from principal components analysis (PCA). Microbiome composition also differed among genotypic classes and across both habitats (Panel C).

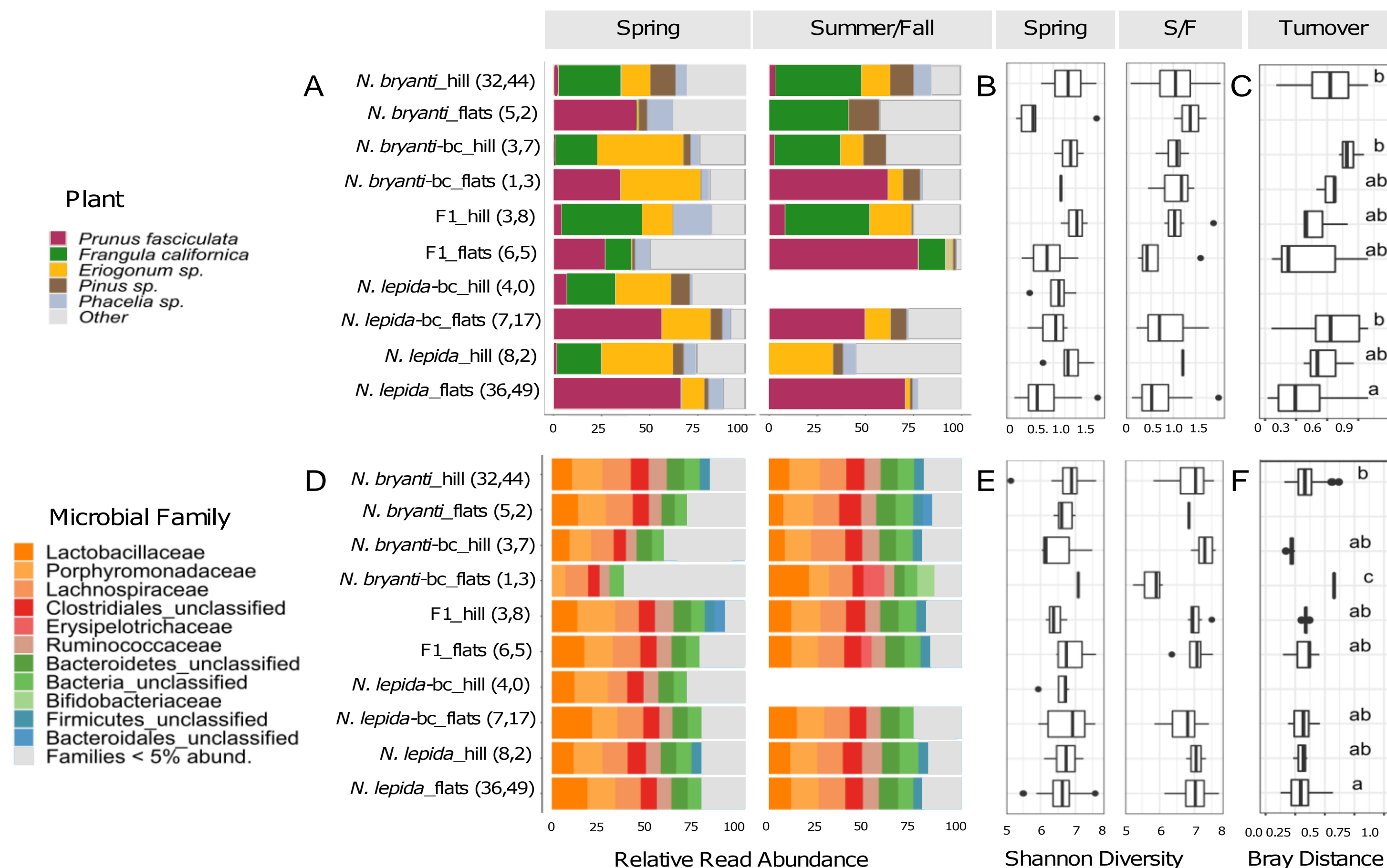

Figure 2: Composition, diversity, and turnover varied among genotype, habitat, and season for diet (panels A-C) and microbiome (panels D-F). Numbers in parentheses on the y axis are sample sizes of each category of woodrat samples for spring and summer/fall, respectively. Mean and sd are plotted in B & E. Turnover for each group was calculated as the average Bray-Curtis distance between spring and summer/fall samples from each category. Letters indicate significant differences (Tukey; $P < 0.05$).

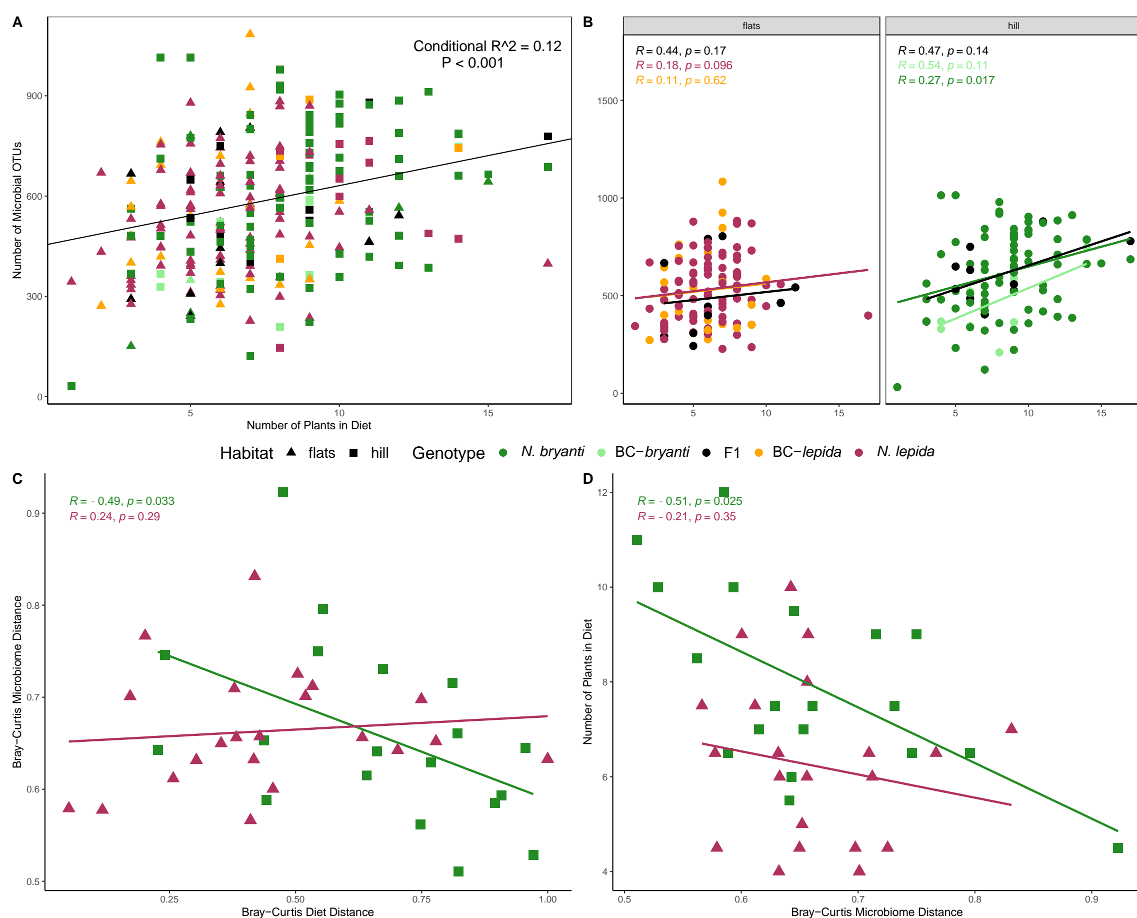

Figure 3: Correlation between diet and microbiome diversity. (A) Overall, diet and microbial richness were significantly correlated ($P = 0.002$). The regression line in panel a is from the *glmmTMB* model using diet richness, season, habitat, and genotype as fixed effects. (B) The relationship between diet and microbiome richness varied when evaluated individually among genotype classes and habitats. (C) Diet and microbiome distance were marginally positively correlated in *N. lepida* individuals sampled multiple times, and marginally negatively correlated in *N. bryanti* sampled multiple times. (D) Diet richness was negatively correlated with microbiome distance in *N. bryanti* individuals sampled multiple times. The open triangles and dashed lines in c & d correspond with *N. lepida*, and the black squares and solid lines correspond with *N. bryanti*.

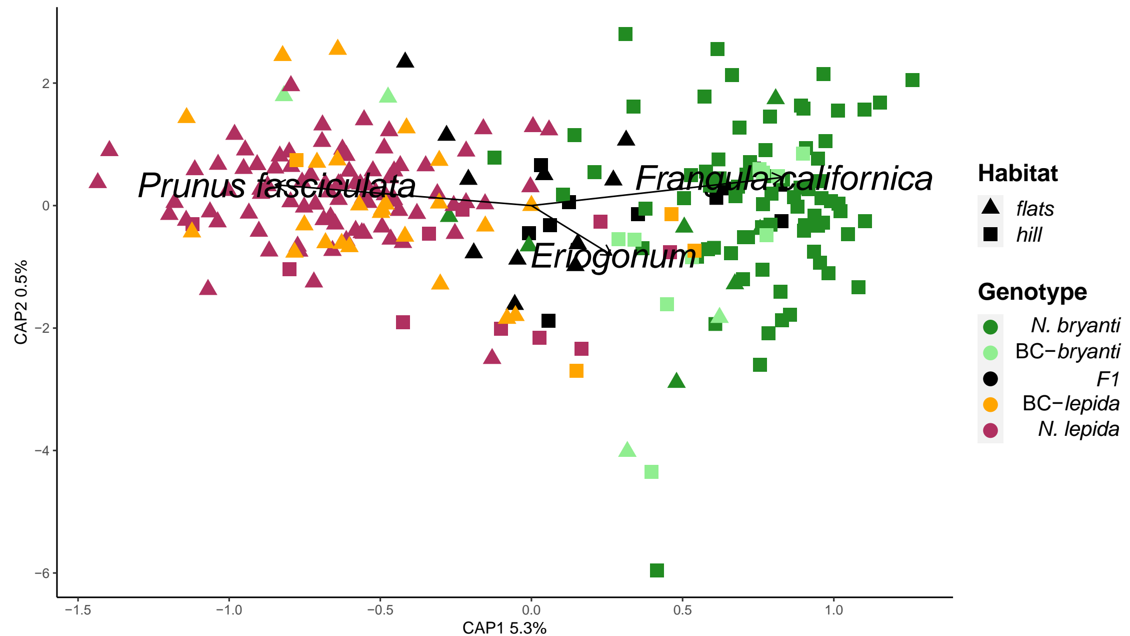

Figure 4: Constrained analysis of principal coordinates (CAP). Each point represents an individual's microbiome community composition using Bray-Curtis dissimilarity. We tested for significant associations between diet plants and microbiome composition, and removed the effects of year, age, and sex with this following formula: OTU ~ *Prunus fasciculata* + *Frangula californica* + *Eriogonum* + *Condition (year + sex + age)*. The association between microbiome and diet plants included was significant (ANOVA with 999 permutations; df = 3, F = 4.95, $P = 0.001$, adj. $r^2 = 4.6$).

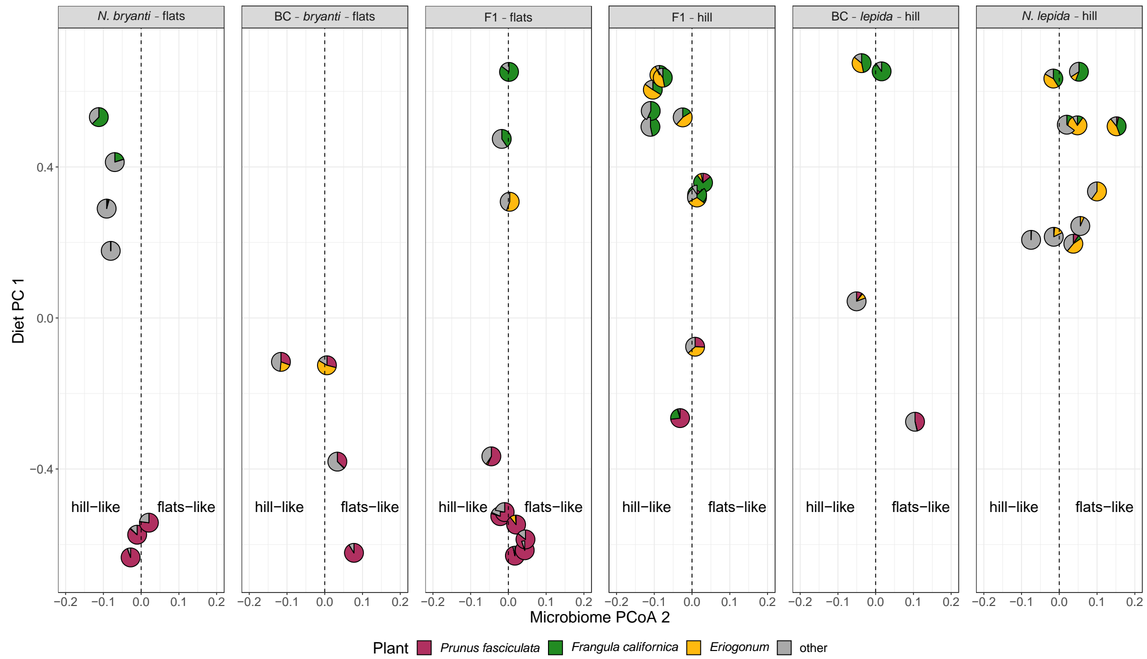

Figure 5: Microbiome PCoA 2 plotted against diet PC 1 for mismatched parental and backcross hybrids, and F1 hybrids. Pie charts represent relative abundance of *Prunus fasciculata*, *Frangula californica*, *Eriogonum*, or the remaining 'other' plants identified in diets using *trn*L metabarcoding. Among parental *N. bryanti* and *N. lepida* individuals, the imprint of ancestry is evident by the maintenance of genotype-specific microbiome, with some exceptions in individuals that consumed habitat-specific diets. The strong influence of genotype on microbiome is also evident in F1 hybrids which exhibit intermediacy (particularly in the flats).

Chapter 2 Supporting Information

The gut microbiome reflects ancestry despite dietary shifts across a hybrid zone

Text S1: Woodrat trapping, sampling, and validation

We set two Tomahawk traps near active nests during sampling periods of 3-4 nights; traps were checked during early morning hours. Fecal samples were either collected in coin envelopes and transported in coolers with dry ice or stored in cryovials and transported in liquid nitrogen to the lab where they were stored at -80 ° C. Woodrats were uniquely marked, and a portion of ear tissue was removed with scissors and stored in 95% EtOH for later DNA sequencing. We recorded body length, ear length, hind foot length, body mass, age class, reproductive condition, and sex before releasing individuals at the point of capture.

As a basis for characterizing microbiome variation, we sampled feces in 2016, prior to trapping, by locating active nests and sampling fresh fecal pellets from latrines as described previously (Nielsen and Matocq 2021). Fecal pellets were stored in coin envelopes and dried until storage in -20 ° C. From 2017-2019, we sampled trap-collected feces while monitoring the population with established traplines and baited minimally with peanut butter to create a scent lure and to minimize impact on the natural gut microbiome.

Validation study

The use of trap-collected fecal samples for microbiome studies of woodrats has previously been validated (Kohl et al., 2015). However, to confirm our sampling was not

biased by environmental contamination, we conducted a validation study specific to our study site.

We collected fecal and soil samples from 6 trapped woodrats. Fresh fecal samples were collected from beneath tomahawk traps, and a paired soil sample was collected from adjacent substrate at the same time for a total of 6 fecal and 6 soil samples. Three pairs each were collected from the hill and flats habitats.

DNA from fecal and soil samples was extracted, amplified, and sequenced by Jonah Ventures LLC (Boulder, CO) using the DNeasy PowerSoil HTP 96 kit. 16S rRNA was PCR amplified using the 515f and 806R 16S rRNA primers (Caporaso et al., 2011): forward - GTGYCAGCMGCCGCGGTAA; reverse - GGACTACNVGGGTWTCTAAT. 16S amplicons were bioinformatically processed using QIIME and UPARSE jointly (Caporaso et al., 2010; Edgar, 2013) following steps similar to Andrei et al. (2015). Operational taxonomic units (OTUs) were generated from clustered reads at 97% similarity.

Read counts for OTUs for each fecal and soil sample were converted to relative read abundance (RRA). PERMANOVA (vegan; Oksanen et al., 2020) and principal coordinates analysis (PCoA) were conducted with Bray-Curtis distances using the *ordinate* function in phyloseq (McMurdie & Holmes, 2013).

We next used the *plot_heatmap* function in phyloseq to visualize RRA of the top 100 microbial OTUs present in fecal and soil samples. The heatmap was structured using principal coordinates analysis with Bray-Curtis distance, and microbial taxa were labeled at the family level.

Source was the only factor that had a significant effect on microbial composition of samples and explained 58% of variation (*P*=0.001; Table S1). Soil and fecal samples clustered in ordination space (Fig. S1). Notably, soil samples formed two clusters that were mostly divided along habitat. However, all fecal sample formed one uniform cluster (Fig. 1). Composition of the top 100 microbial OTUs in the dataset were clearly distinct between soil and fecal samples; and only 10 OTUs that were common across all soil samples were detected in fecal samples, and in low abundance (~0.0002; Fig. S2). These results provide further evidence that lab contamination was not appreciable, as this would have had a homogenizing effect on these samples.

Text S2: Woodrat genotyping and SNP analysis

Reads were aligned to an unpublished *N. lepida* genome using BWA-MEM aligner v. 0.7.17 (r1188) (Li 2013) followed by Samtools v. 1.11 (Li 2011) to sort the reads. The resulting bam files were processed with Stacks v. 2.53 to identify SNPs and call genotypes for each locus (Catchen et al., 2013). We filtered SNPS using the populations script of Stacks. We set the minimum number of populations a locus must be present in to process to 1 (-p flag) and set the minimum percentage of samples in the population required to process a locus for the population to 60% (-r flag). We required a minor allele frequency of at least 0.05 to process a nucleotide site (*–min_maf flag*). To avoid including SNPs in elevated linkage disequilibrium, we included only the first SNP in each fragment (*–write_single_snp* flag). To infer population structure, we used Stacks structure file as input for FastSTRUCTURE (Raj et al., 2014). FastSTRUCTURE implements the Bayesian model used by STRUCTURE (Falush et al., 2003; Pritchard et

al., 2000), but uses computationally efficient solutions for estimating individual ancestries in very large data sets.

Text S3: Modeling the proportion of plants and microbes in samples using CNVRG

The CNVRG approach models sequence counts for plant or microbe OTUs as proportions using the Dirichlet and multinomial distributions. The hierarchical nature of this model allows for sharing of information among individuals within a sampling group. Samples were placed in one of five populations according to the genotype of the individual woodrat sampled: *N. bryanti*, BC-*bryanti*, F1, BC-*lepida*, and *N. lepida*. To remove effects of pseudoreplication, we randomly selected one sample from individuals we sampled multiple times.

Text S4: Correlation between diet and microbial richness and turnover

We used a linear model implemented in *glmmTMB* to determine if diet and microbial richness were correlated using diet richness, season, genotype, and habitat as fixed effects (glmmTMB; Brooks et al. 2017). We used Spearman correlation tests to evaluate the linear relationship between diet and microbiome richness for each genotypic class within their matched habitat. Leveraging the large number of parental *N. bryanti* and *N. lepida* individuals sampled multiple times, we further explored this relationship using Spearman correlation tests to evaluate the following relationships: diet and microbiome Bray-Curtis distances; and microbiome Bray-Curtis distance and diet richness. Duplicate pairwise comparisons due to more than two samples from the same individual were averaged. Due to the issue of pseudoreplication in using these data for

the CNVRG model (see above), we used raw *trn*L and 16S read counts for this analysis. To reduce bias from differences in sequencing depth across these replicate samples, we rarefied diet and microbiome read counts to 2000 reads.

Text S5: Correlation of host and environmental variables that contribute to diet and microbiome composition

We calculated Pearson's correlation coefficients (Pearson's r) between host and environmental variables and the first four axes of both diet (PCA) and microbiome (PCoA). We do not differentiate between positive or negative correlations, and thus report the absolute value of these estimates (Tables S2-S3). Of the first four diet PCs, habitat (Pearson's $r = 0.83$, $p < 0.001$) and genotype (Pearson's $r = 0.71$, $p < 0.001$) were only correlated with PC1, which explained 43% of variation in diet (Fig. 1b). Season (Pearson's $r = 0.15$, $p = 0.02$) and year (Pearson's $r = 0.\ 16$, $p = 0.01$) were correlated most strongly with diet PC2, which explained 12% of diet variation. Among the first four microbiome PCoA axes, habitat (Pearson's $r = 0.74$, $p < 0.001$) and genotype (Pearson's = 0.84, $p < 0.001$) were only correlated with microbiome axis 2, which explained 9.4% of variation in microbiome (Fig. 1c). Microbiome composition along PCoA axis 1 explained 15% of variation and was strongly correlated with season (Pearson's $r = 0.17$, $p = 0.009$) and year (Pearson's $r = 0.30$ $p < 0.001$).

Text S6: Partial distance-based redundancy analysis

For partial distance-based redundancy analyses, we included genotype, habitat (i.e., hill or flats), season, sex, and age. We also used loadings from diet PC1 and

microbiome PCoA 2 as variables. We removed the effect of year using the Condition () statement.

Genotype explained 23.3% and 8.2% of variation in diet and microbiome composition, respectively (Table 1). Habitat explained 31.6% and 5.6% of variation in diet and microbiome, respectively (Table 1). Season explained only a small amount of variation in diet (< 1%) but explained more variation in microbiome (2.1%). Diet PC1 explained 5.3% of the variation in microbiome composition, while microbiome PCoA2 explained 24.9% variation in diet composition.

Text S7: Linear Relationships between diet and microbiome distance

We tested for linear relationships between diet and microbiome Bray-Curtis distance using mantel tests within each genotypic class, and collectively with the entire dataset. We calculated Pearson correlation coefficient between pairwise distances using *stat_cor* function of the ggpubr package (Kassambara, 2020).

Overall, diet and microbiome distance were correlated, but not when compared within genotypic classes: *N. bryanti*-like, *N. lepida*-like, or F1 hybrids (Fig. S3 a-f).

Text S8: Testing F1 hybrid microbiome intermediacy

To test whether F1 hybrids harbor intermediate microbiome composition, we calculated pairwise Bray-Curtis microbiome distances across the following categories: conspecific (distance between either two *N. bryanti* or *N. lepida* individuals); F1

(distance between either parental species and an F1 individual); and heterospecific (distance between *N. bryanti* and *N. lepida)* individuals. We then performed analysis of variance using the *aov* function followed by Tukey test using *TukeyHSD*, which provides adjusted p values for each comparison. Average pairwise distances within all three categories were significantly different ($P < 0.001$), with conspecifics exhibiting the most similar compositions, then F1 comparisons, followed by heterospecific comparisons (Fig. S5).

Text S9: Differentially abundant microbial taxa and potential function

We used multinomial estimates of microbial OTU proportions to identify microbial taxa that were differentially abundant between various categorical groupings. We collapsed the microbial dataset to the genus level and estimated the probability of differing and effect size for all genera. We investigated differential abundances of genera in five comparisons: Pure *N. bryanti* versus *N. lepida* in 1) native habitat and 2) mismatched habitat; 3) F1s on the hill versus the flats; and 4) *Prunus* consumers (> 50% of diet) versus *Frangula* consumers (> 25%) regardless of genotype or habitat. Because woodrats consume *P. fasciculata*, which is known to produce cyanogenic glycosides (Matocq et al., 2020), we also examine the among-sample distribution of microbial lineages that produce rhodanese-like enzymes (Zhu et al., 2018).

Three hundred and seventeen taxa within 7 microbial phyla and 29 genera differed significantly between *N. bryanti* and *N. lepida*. Eight genera differed between F1 hybrids on the hill versus flats (Fig. S6a). Diet preference for either *F. californica* or *P. fasciculata* produced the most pronounced differences among microbiomes that we

observed. We found ten genera differed between woodrats, regardless of genotype, that consumed greater than 50% *P. fasciculata* or greater than 25% *F. californica* (Fig. S6b). Seven genera differed between *N. lepida* on the hill versus the flats (Fig. S6c). Six genera differed between *N. bryanti* on the hill versus flats (Fig. S6d). Unclassified genera within the family Porphyromonadaceae, were significantly more abundant in *N. bryanti* in the hill habitat compared to the flats.

The gut microbiome can provide an array of diet-related services otherwise unavailable to host animals and, within the woodrat genus *Neotoma,* has been shown to enable metabolism of certain toxic plant compounds (Dearing & Kohl, 2017; Kohl et al., 2014). While microbes can contribute to nutrient acquisition and detoxification (Wybouw et al., 2014, Dearing and Kohl, 2017), it should be noted that microbial activity can also transform plant secondary compounds into their potentially toxic and bioavailable form. For example, both anthraquinones (Surh et al. 2013) and cyanogenic glycosides (Carter et al., 1980) would otherwise be able to pass through an animal gut in their intact form and be excreted, but once bacterial beta-glucosidase activity in the gut cleaves the sugar molecule from the parent compound, these metabolites become potentially toxic (*i.e.,* emodin and hydrogen cyanide, respectively).

Lactobacillus was strongly associated with *N. lepida* and F1 hybrids in the flats, as well as diets composed of greater than 50% *P. fasciculata*. While we do not have adequate resolution to identify individual species within the genus *Lactobacillus*, some species are known to metabolize cyanogenic glycosides (Lei et al., 1999). Specifically, some Lactobacilli have beta-glucosidase activity that cleaves the sugar molecule from cyanogenic glycosides, causing the release of hydrogen cyanide during digestion. As

such, it appears that gut microbial activity in *N. lepida* and animals consuming high *Prunus* diets may be releasing hydrogen cyanide. This potential release of hydrogen cyanide in the gut is in addition to that which is likely released early in the consumption process from mastication (Suchard et al., 1998). From either or both sources, animals consuming *Prunus* are likely being exposed to hydrogen cyanide. In mammals, hydrogen cyanide is typically modified in the liver by the host enzyme rhodanese (thiocyanate sulfurtransferase), resulting in thiocyanate is eliminated in urine. Alternatively, or in addition, many bacterial lineages have rhodanese-like activity, and thus, could detoxify hydrogen cyanide within the woodrat gut. Pandas consume diets high in cyanogenic glycosides (bamboo) and have gut microbial communities enriched in bacteria that have rhodanese-like activity (Zhu et al. 2018). These rhodanese-like lineages include members of the Pseudomonadaceae, which we found to be among the bacterial lineages significantly differentially abundant in certain genotype-habitat combinations. Surprisingly, these lineages were not found in heavy *Prunus*-consumers whose guts might contain hydrogen cyanide, but rather in *N. lepida* and BC-*lepida* living on the hill habitat-- animals that are primarily consuming *Eriogonum*. We hypothesize that because these animals may only rarely consume *Prunus* in the hill habitat, they may not be effectively inducing their own rhodanese pathways. Therefore, gut bacteria may be exposed to hydrogen cyanide providing a substrate for rhodanese-like lineages. Such exposure may not occur in an animal that routinely consumes *Prunus*, has consistent or more rapid induction of their own rhodanese pathway, thus clearing this toxin before the lower gut bacteria is exposed.

The microbiomes of *N. bryanti* in their native hill habitat, or individuals that consuCmed large amounts of *F. californica*, were enriched with various genera in the family Porphyromonadaceae. Lineages within this family are important for fiber degradation (Ivarsson et al., 2014) and may promote more efficient digestion of the particular plants and plant parts that characterize the diets of *N. bryanti* and animals occupying the hill habitat. It is also possible that *N. lepida* and animals on the flats consume less fiber overall because they may include more non-fibrous items in their diets (*e.g.,* insects, fungi). We do not have reason to expect a difference in consumption of non-fibrous diet items across this study site or species, but insects have been found to compose nearly 7% of *N. floridana*'s diet (Kanine et al. 2015), while *N. magister* (Castleberry et al. 2002) are known to consume fungi. It should be noted that all the microbial differences we uncovered were of small effect size, yet they provide the groundwork for future studies to study gut microbes that facilitate digestion and detoxification of habitat-specific diets. Further, while the taxonomic shifts in frequency of microbial lineages we uncovered were subtle, differences in rare taxa may be functionally significant in terms of gene expression and enzymatic activity (Wei et al., 2019).

Chapter 2 Supplemental Tables and Figures

Table S1: PERMANOVA results from paired soil and fecal samples. Source (either feces or soil) was the only significant factor.

Variable	**F statistic**	**% Variation explained (R^2)**	***P***
Source	14.7	58.0	0.001
Habitat	1.4	5.4	0.233
Nest	1.0	15.8	0.458
Source x Habitat	1.4	5.4	0.245

PERMANOVA with 999 permutations.

Table S2: Pearson's r coefficients between host and environmental variables and the first four principal components of diet composition cumulatively explain 69.4% of variation. The percent variation explained by each PC are in parentheses.

Variable	PC1 (43%)	PC2 (12.1%)	PC3 (8.9%)	PC4 (5.4%)
Habitat	0.83*	0.04	0.01	0.12
Genotype	0.71*	0.1	0.07	0.08
Season	0.04	0.15*	0.13*	0.36*
Year	0.004	0.16*	0.1	0.16*

* $P \leq 0.05$

Table S3: Pearson's r coefficients between host and environmental variables and the first four principal coordinate axes of microbiome composition which cumulatively explain 36.6% of variation. The percent variation explained by each axis are in parentheses.

Variable	Axis 1 (15%)	Axis 2 (9.4%)	Axis 3 (6.7%)	Axis 4 (5.5%)
Habitat	0.05	0.74*	0.09	0.06
Genotype	0.05	0.84*	0.01	0.01
Season	0.17*	0.01	0.22*	0.41*
Year	0.3*	0.03	0.04	0.19*

* $P \leq 0.05$

Table S4: Results of partial distance-based redundancy analysis (dbRDA) of environmental and host variables on diet and microbiome composition. Bray-Curtis distances were used for both diet and microbiome. Each variable was evaluated individually for its contribution to explained variance (adjusted r^2 %) and with the effect of year removed.

Variable	adjusted r^2 (%)	
	Diet	Microbiome
Genotype	23.3**	8.2**
Habitat	31.6**	5.6**
Season	0.6*	2.1**
Diet PC 1	NA	5.3**
Microbiome PCoA 2	24.9**	NA
Sex	0	0.1
Age	0.1	0

Significance evaluated using ANOVA
with 999 permutations: * $P < 0.05$; ** $P < 0.001$

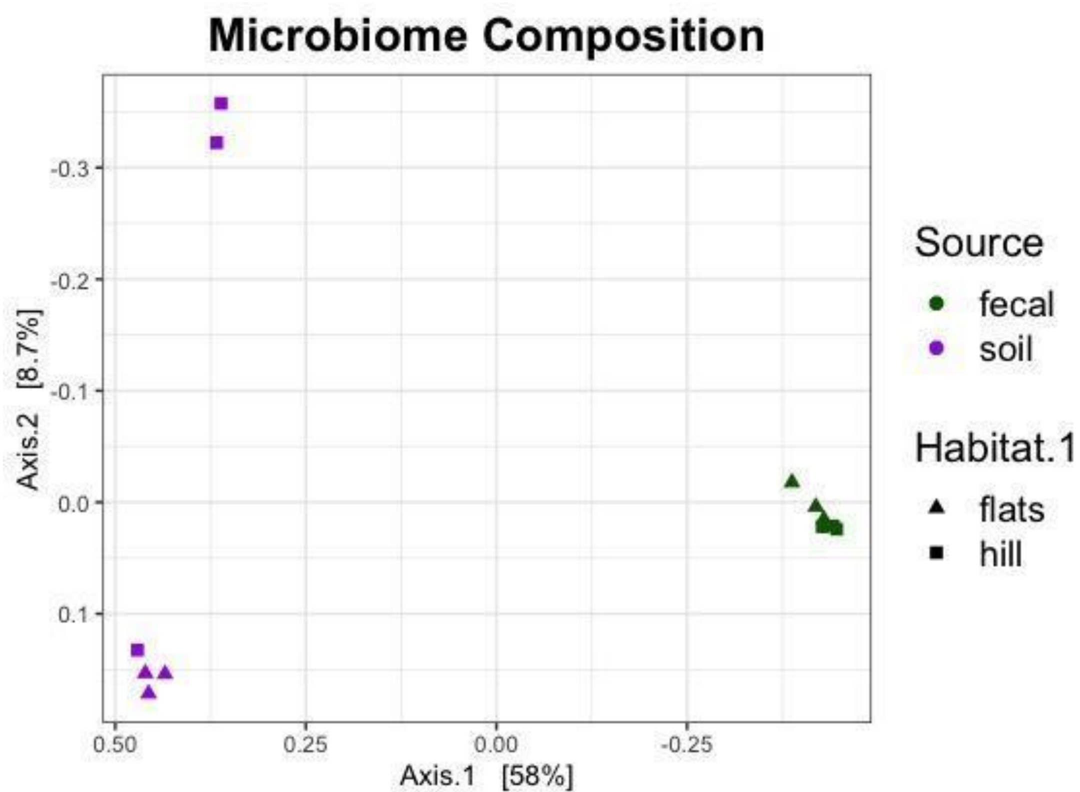

Figure S1: Principal coordinates analysis using Bray-Curtis distance of microbial composition of paired fecal and soil samples.

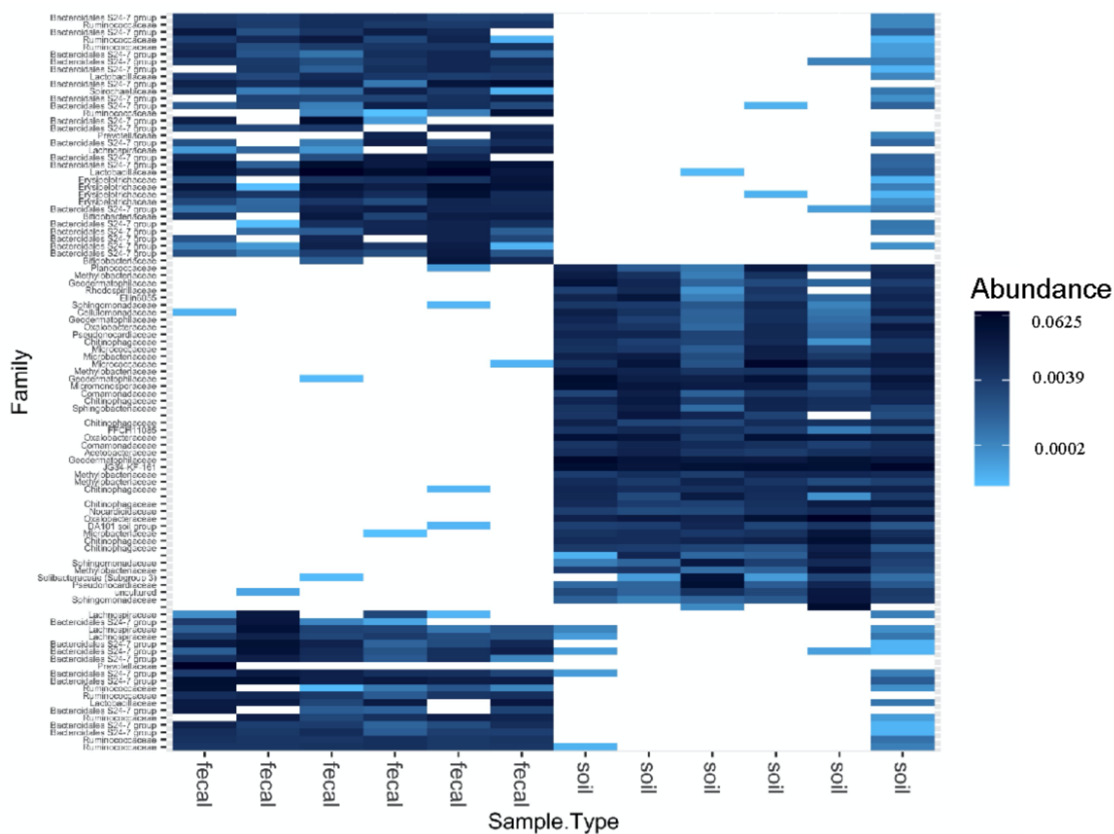

Figure S2: Heat map of relative read abundance of the 100 most prevalent microbial families identified in fecal and soil samples. The heat map was structured using principal coordinates analysis with Bray-Curtis distance.

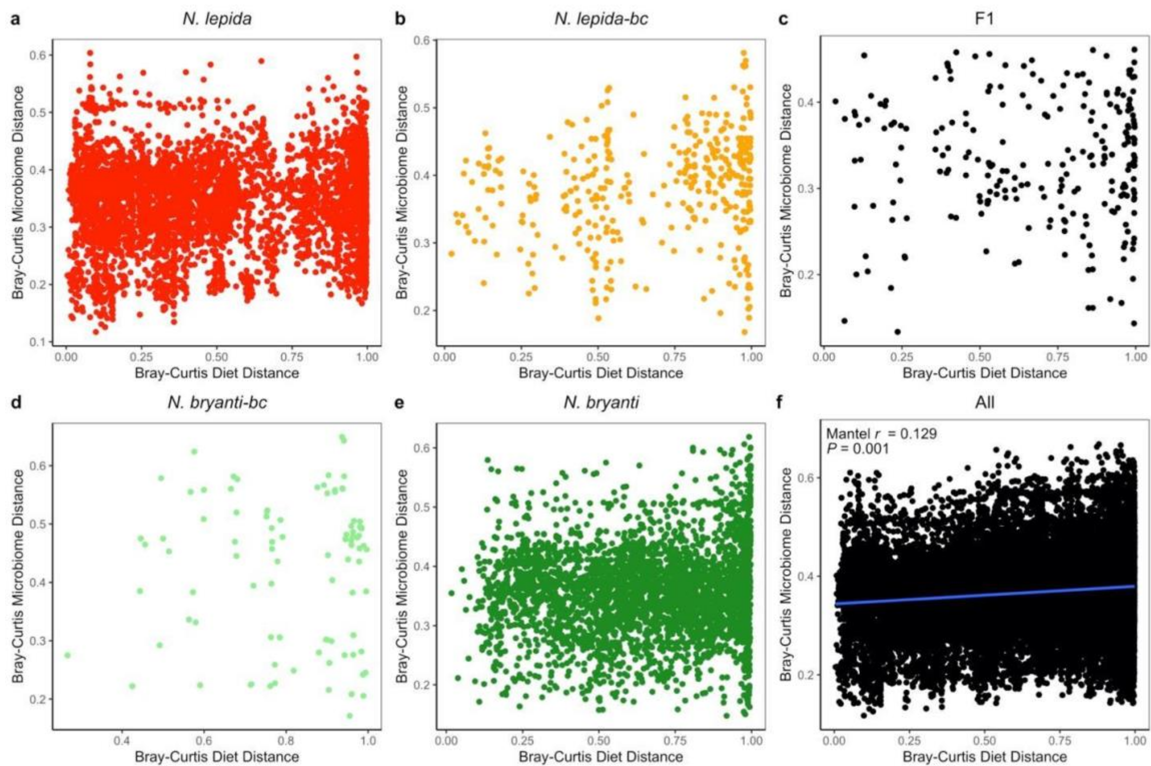

Figure S3: Diet and microbiome distances were not significantly correlated within individual genotypic groups (a-e). Diet and microbiome distance were significantly correlated when evaluated across the entire dataset (f).

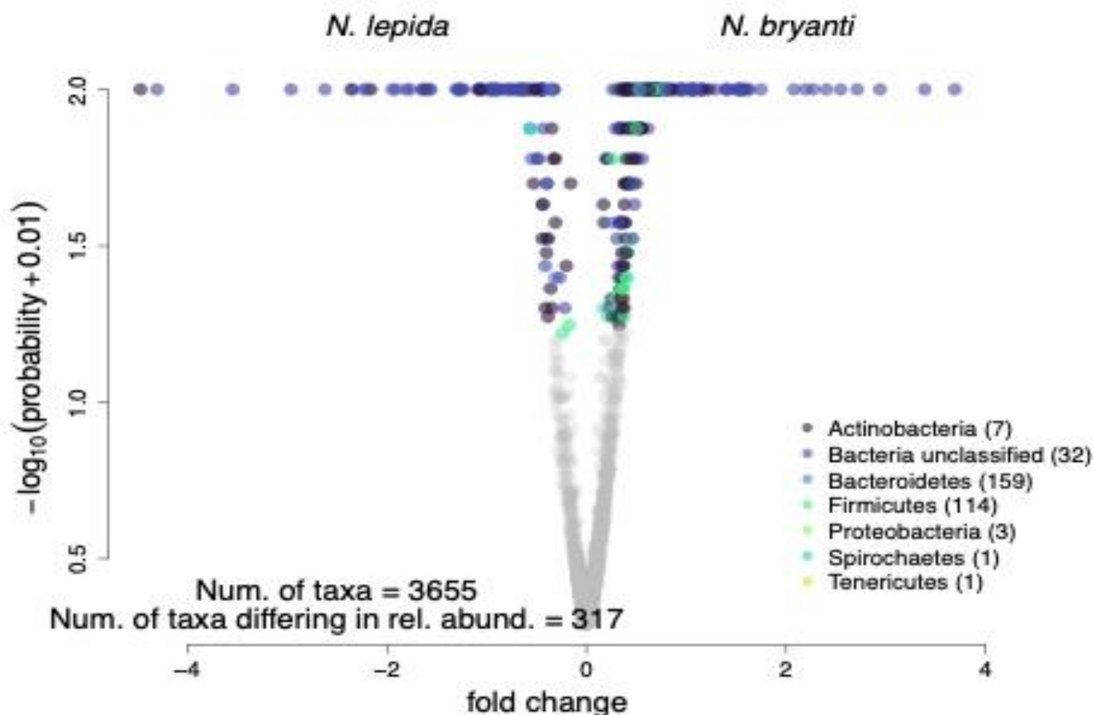

Figure S4: A total of 317 microbial OTUs were differentially abundant between pure *N. lepida* and *N. bryanti*. CNVRG estimates were used to identify features that differed. Colors represent microbial phyla with the number of differentially abundant OTUs in parentheses.

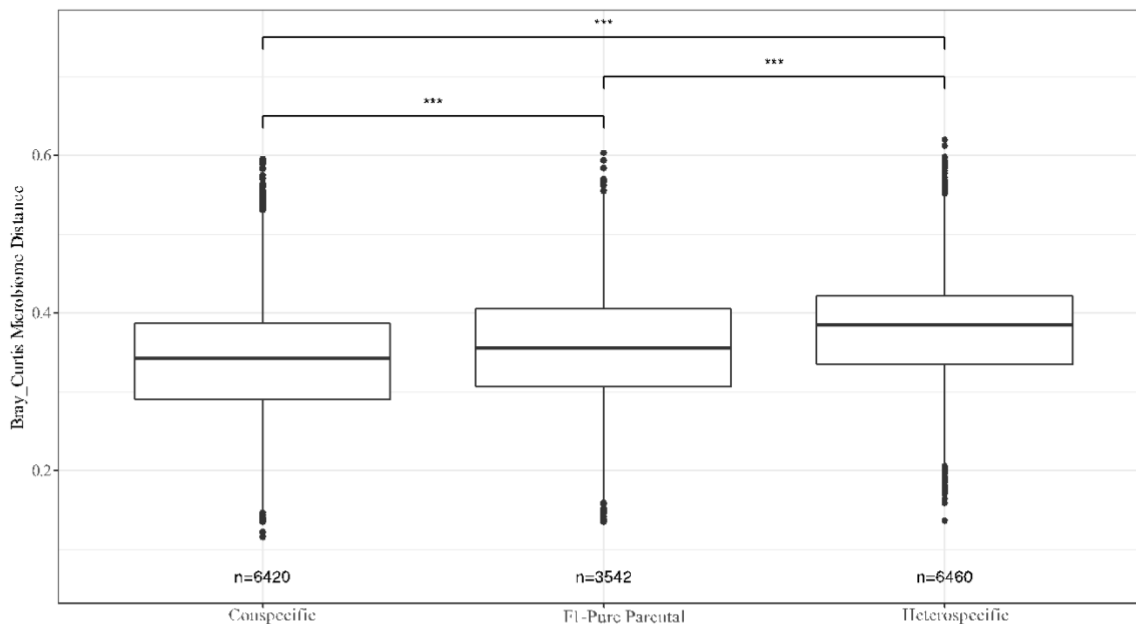

Figure S5: Average Bray-Curtis microbiome distance among pairwise comparisons of conspecifics, heterospecifics, and F1 vs. pure parentals. Average pairwise microbiome distances were greatest in heterospecific comparisons (i.e., *N. lepida* and *N. bryanti* pairwise distances), lowest in conspecific comparisons (i.e., between *N. lepida* in the flats and between *N. bryanti* on the hill), and intermediate between F1 hybrids and pure parentals.

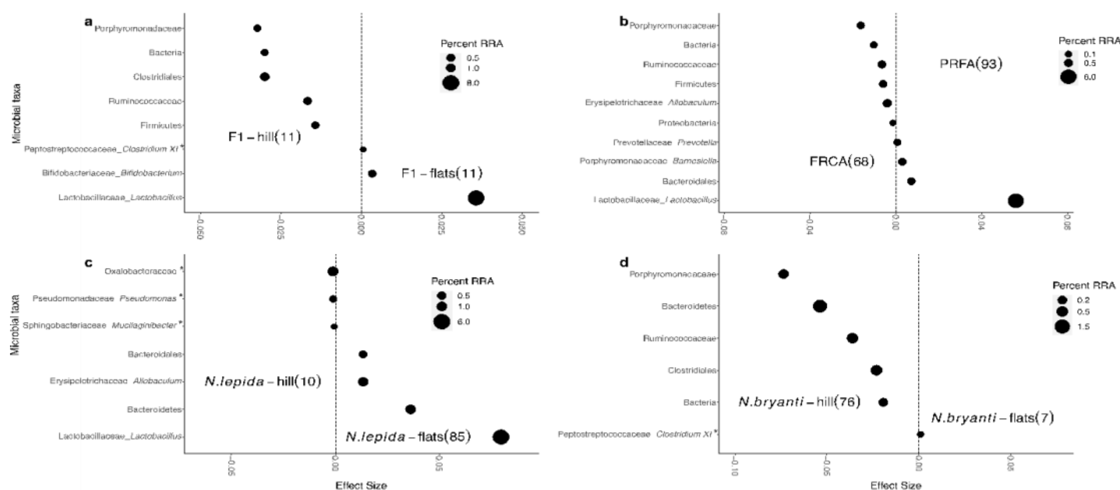

Figure S6: Results of CNVRG modeling of differentially abundant microbial genera. Individuals had to consume either 50% or more PRFA or 25% FRCA or more to be classified as either diet type (Panel b). Taxa displayed here are those that differed with 100% probability, or lineages that are known to have rhodanese activity (denoted with asterisk). A complete list of taxa that differed with > 95% probability is in Tables S3-S6. The size of the points represents the percent relative read abundance (RRA) of that microbial taxa within the group for which the feature was greater in abundance.

References

Andrei, A.-Ş., Robeson, M. S., Baricz, A., Coman, C., Muntean, V., Ionescu, A., Etiope, G., Alexe, M., Sicora, C. I., & Podar, M. (2015). Contrasting taxonomic stratification of microbial communities in two hypersaline meromictic lakes. *The ISME Journal*, *9*(12), 2642–2656.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). GlmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/10.3929/ethz-b-000240890

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, *108*(Supplement 1), 4516–4522.

Carter, J. H., McLafferty, M. A., & Goldman, P. (1980). Role of the gastrointestinal microflora in amygdalin (laetrile)-induced cyanide toxicity. *Biochemical Pharmacology*, *29*(3), 301–304. https://doi.org/10.1016/0006-2952(80)90504-3

Castleberry, N. L., Castleberry, S. B., Ford, W. M., Wood, P. B., & Mengak, M. T. (2002). Allegheny Woodrat (Neotoma magister) Food Habits in the Central Appalachians. *The American Midland Naturalist*, *147*(1), 80–92. https://doi.org/10.1674/0003-0031(2002)147[0080:AWNMFH]2.0.CO;2

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Dearing, M. D., & Kohl, K. D. (2017). Beyond fermentation: Other important services provided to endothermic herbivores by their gut microbiota. *Integrative and Comparative Biology*, *57*(4), 723–731.

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998. https://doi.org/10.1038/nmeth.2604

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*(4), 1567–1587.

Ivarsson, E., Roos, S., Liu, H. Y., & Lindberg, J. E. (2014). Fermentable non-starch polysaccharides increases the abundance of Bacteroides–Prevotella–Porphyromonas in ileal microbial community of growing pigs. *Animal*, *8*(11), 1777–1787. https://doi.org/10.1017/S1751731114001827

Kanine, J. M., Castleberry, S. B., Mengak, M. T., & Winchester, C. (2015). Seasonal Variation in Composition of Key Largo Woodrat (Neotoma floridana smalli) Diets. *Southeastern Naturalist*, *14*(2), 405–414. https://doi.org/10.1656/058.014.0220

Kassambara A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version
0.4.0. https://CRAN.R-project.org/package=ggpubr

Kohl, K. D., Luong, K., & Dearing, M. D. (2015). Validating the use of trap-collected feces for studying the gut microbiota of a small mammal (Neotoma lepida). *Journal of Mammalogy*, *96*(1), 90–93.

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Dearing, M. D. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology Letters*, *17*(10), 1238–1246. https://doi.org/10.1111/ele.12329

Lei, V., Amoa-Awua, W. K. A., & Brimer, L. (1999). Degradation of cyanogenic glycosides by Lactobacillus plantarum strains from spontaneous cassava fermentation and other microorganisms. *International Journal of Food Microbiology*, *53*(2), 169–184. https://doi.org/10.1016/S0168-1605(99)00156-7

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv:1303.3997 [q-Bio]*. http://arxiv.org/abs/1303.3997

Matocq, M. D., Ochsenrider, K., Jeffrey, C. S., Nielsen, D., & Richards, L. A. (2020). Fine-scale differentiation in diet and metabolomics of small mammals across a sharp ecological transition. *Frontiers in Ecology and Evolution*, *8*, 282.

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.

Nielsen, D. P., & Matocq, M. D. (2021). Differences in dietary composition and preference maintained despite gene flow across a woodrat hybrid zone. *Ecology and Evolution*, *11*(9), 4909–4919. https://doi.org/10.1002/ece3.7399

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., Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package* (2.5-7) [Computer software]. https://CRAN.R-project.org/package=vegan

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Suchard, J. R., Wallace, K. L., & Gerkin, R. D. (1998). Acute Cyanide Toxicity Caused by Apricot Kernel Ingestion. *Annals of Emergency Medicine*, *32*(6), 742–744. https://doi.org/10.1016/S0196-0644(98)70077-0

Surh, I., Brix, A., French, J. E., Collins, B. J., Sanders, J. M., Vallant, M., & Dunnick, J. K. (2013). Toxicology and Carcinogenesis Study of Senna in C3B6.129F1-Trp53tm1Brd N12 Haploinsufficient Mice. *Toxicologic Pathology*, *41*(5), 770–778. https://doi.org/10.1177/0192623312464304

Wei, Z., Gu, Y., Friman, V.-P., Kowalchuk, G. A., Xu, Y., Shen, Q., & Jousset, A. (2019). Initial soil microbiome composition and functioning predetermine future plant health. *Science Advances*, *5*(9), eaaw0759. https://doi.org/10.1126/sciadv.aaw0759

Wybouw, N., Dermauw, W., Tirry, L., Stevens, C., Grbić, M., Feyereisen, R., & Van Leeuwen, T. (2014). A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. *ELife*, *3*, e02365. https://doi.org/10.7554/eLife.02365

Zhu, L., Yang, Z., Yao, R., Xu, L., Chen, H., Gu, X., Wu, T., & Yang, X. (2018). Potential Mechanism of Detoxification of Cyanide Compounds by Gut Microbiomes of Bamboo-Eating Pandas. *MSphere*, *3*(3). https://doi.org/10.1128/mSphere.00229-18

Chapter 3:

Differential response of specialist and generalist herbivores when switching diets: host gene expression, the gut microbiome, and their potential interaction

Abstract

Closely related species that have diverged and independently evolved during periods of geographic isolation may interbreed upon secondary contact if reproductive isolation is incomplete. Differential adaptations accumulated during isolation can mediate the frequency with which heterospecifics interact and the fitness of hybrids that may result from secondary contact. For mammalian herbivores, diet may be an agent of selection as the plants they eat often produce toxic compounds, but little is known concerning: 1) the underlying mechanisms by which mammalian herbivores metabolize these toxins; nor 2) the degree to which these herbivores may shift to novel diet plants. We have studied diet and diet-related adaptations of two species of woodrats (*Neotoma*) that hybridize across a sharp ecotone characterized by differences in plant community composition. At this site, *N. lepida* consumes primarily *Prunus fasciculata* and is considered a dietary specialist, while *N. bryanti*—considered a dietary generalist—consumes a more variable diet, with *Frangula californica* being the single most common item in its diet. We conducted laboratory-based diet experiments to understand how each species metabolizes its native diet and the degree to which they can switch to the primary plant consumed by the other species. On native and non-native diets, we quantified change in daily body mass, voluntary wheel running, daily food and water intake, metabolic rate, and gene expression in the liver and caecum. We also sequenced a portion of the 16S rRNA gene from DNA extracted from caecum contents to characterize changes in gut microbial composition across species and diet treatments. We found that *N. lepida*, the dietary specialist, showed greater response to diet changes by exhibiting

lower tolerance for the non-native diet and dramatic shifts in gene expression and gut microbial composition. Our results suggest that diet-related adaptations may indeed mediate hybrid zone dynamics, but asymmetrically, whereby the dietary generalist may be less constrained by diet changes in new habitats than the dietary specialist.

Introduction

When closely related, and independently evolved lineages come into secondary contact after a period of spatial separation, hybridization may occur if reproductive isolation is incomplete (Barton and Hewitt, 1985). How biological diversity is then maintained despite such gene flow is a central question in ecology and evolution. At secondary contact zones, barriers to gene flow may be weakened, or strengthened, due to hybridization (Abbot et al. 2013). Secondary contact zones often occur across ecotones where sharp transitions in both biotic and abiotic factors between adjacent habitats promote local adaptation and habitat segregation of sister taxa (Via 1999). Across such ecotones, selective gradients may reduce the fitness of migrants (*i.e.,* pre-zygotic reproductive isolation) and/or reduce fitness of the hybrids themselves (*i.e.,* post-zygotic reproductive isolation; Nosil et al. 2005).

Mammalian herbivores are thought to experience strong selection from the plants they consume (Freeland and Janzen 1974). Mammalian herbivores that consume plants for which they are not adequately adapted may suffer energetic costs (Sorenson et al. 2005), such that dispersal into new habitats with novel diet plants may be costly. Woodrats of the genus *Neotoma* can consume toxic diets (Dial 1988, Mangione et al. 2000, Skopec et al. 2007) and likely minimize exposure to toxins through both behavioral modification of the plants they consume (Iason and Villalba 2006, Nielsen et al. 2021), expression of liver detoxification enzymes (Kitanovic et al. 2018), and through enzymatic activity produced by their gut microbiome (Kohl et al. 2014, Nielsen et al. 2022). Moreover, interactions between host gene expression and gut microbial composition can

have positive, or negative, impacts on host health (Nichols and Davenport 2021). However, the degree to which dietary adaptations mediate the potential for hybridization and interspecific gene flow has not been explored in mammals.

Woodrats (*Neotoma*) provide a unique opportunity to examine the role of dietary adaptations in the maintenance of species boundaries. We have been studying the ecology of a woodrat hybrid zone between *Neotoma lepida* (desert woodrat) and *N. bryanti* (Bryant's woodrat) in southern California. The two species hybridize across a sharp ecotone between the xeric Mojave Desert (hereafter "flats"), and a relatively mesic rocky slope of the Sierra Nevada foothills (hereafter "hill"). Here, *N. lepida* primarily occupies the flats where they consume *Prunus fasciculata* (desert almond), while *N. bryanti* primarily occupies the hill where *Frangula californica* (California coffeeberry) is the most abundant plant in their diet (Nielsen and Matocq 2021). These principal diet plants contain potentially toxic plant secondary compounds (PSCs): prunasin, a cyanogenic glycoside, in *P. fasciculata* and emodin, an anthraquinone, in *F. californica* (Matocq et al. 2020). These dietary differences are maintained across seasons and years of high and low precipitation, although both diet choice trials and field sampling suggest that *N. bryanti* is more of a dietary generalist while *N. lepida* is more of a dietary specialist (Nielsen and Matocq 2021, Nielsen et al. 2022). Diet composition in this system appears to be strongly driven by the habitat in which an individual lives, while gut microbiome composition appears to be primarily driven by host genotype (Nielsen et al. 2022).

We conducted laboratory diet manipulations to gain insight into the physiological mechanisms that woodrats use to minimize exposure to plant-based toxins and to understand the degree to which woodrats can shift their consumption of potentially toxic

plants. Specifically, we build on our previous work to better understand the differential responses of pure *N. bryanti* and *N. lepida* to their native and non-native diets by testing the following predictions: 1) woodrat species will exhibit differential response to diet treatments, with the dietary specialist showing reduced tolerance of their non-native diet; 2) gene expression will differ between woodrat species and will shift with diet changes, although the dietary specialist will exhibit greater change in gene expression with dietary change; and 3) microbiome composition will differ between woodrat species and will shift with diet changes, with the expectation that the dietary specialist will show greater change in microbiome composition in response to dietary change. We tested these predictions using lab-based experimental feeding trials, 16S rRNA sequencing of gut contents, and RNA sequencing of both liver and caecum tissue. Understanding how *N. lepida* and *N. bryanti* respond physiologically and genetically to both their native and non-native diets will improve our understanding of the role of diet and diet-related adaptations in maintaining this species boundary.

Methods

Field collection and transport

To prevent possible confounding effects of admixture at the hybrid zone, we collected *N. bryanti* and *N. lepida* from populations on either side of the hybrid zone, but where either *Frangula californica* (California coffeeberry) or *Prunus fasciculata* (desert almond) was present (S.I. Table S1). Woodrats were trapped using tomahawk live traps baited with peanut butter and oats. We only collected adult individuals for use in

laboratory feeding trials and females that were not pregnant or lactating. Upon capture, individuals were given a unique identifying ear tag, a piece of ear pinnae was removed and stored in ethanol for DNA sequencing, and body measurements were recorded as well as sex. Captured animals were temporarily housed and transported in shoebox cages and given food and water until arrival at the University of Nevada, Reno.

Urine samples were collected to screen for hantavirus before animals were placed in individual cages (32 x 35 x 26 cm) prepared with woodchip bedding (Teklad Sani-Chips 7090), and a tube for hiding. Animals were maintained with water ad libitum and a high fiber rabbit chow (Teklad formula 2031). Temperature in the housing room was kept at 23° C and lighting was maintained on a 15:9 hour light:dark cycle. During experiments, woodrats were fed the same diet in powder form to prevent caching. All animal procedures were reviewed and approved by the University of Nevada Reno Institutional Animal Care and Use Committee, the California Department of Fish and Wildlife, and were consistent with the guidelines developed by the American Society of Mammalogists (Sikes et al., 2016).

Feeding Trials

During feeding experiments, woodrats were placed into shoebox cages outfitted with running wheel, food hopper, water bottle, and a wire floor to allow feces and urine to be collected. Beneath the wire floor of each cage was a mesh screen to collect fecal pellets, and below that, a tray to collect urine (as in Sorenson et a. 2005). Feces and urine were collected daily during the feeding trials and stored at -80 C. Each day of the trial, running wheel activity was recorded, food and water consumption was measured,

woodrat body mass recorded, and food and water was reapportioned as needed and recorded (these measurements will be hereafter referred to as during-trial responses).

Experimental diet was prepared via methanol extraction of dry leaves of *P. fasciculata* and *F. californica*. After soaking overnight, extracts were dried using a Genevac EZ-2 (SP Scientific), followed by overnight high-vacuum to remove residual solvent. For each plant we divided the mass (in grams) of dried extract by the mass of starting material to determine the amount of extract to incorporate into rabbit chow for each desired dosage. When we refer to 100% dose this is the maximum amount of extract an animal would be exposed to in the wild if they only consumed that plant. Extract was thoroughly mixed with powdered rabbit chow so that woodrats could not parse chow from experimental compounds. Woodrats were maintained at each dose for 2 days, and then for 3 days on the maximum dose determined for each treatment group to ensure liver enzyme induction. For the first trial, diets were adjusted as follows: 0%, 1%, 5%, 10%, then increased in 10% intervals until reaching 100%. We conducted this first trial as such to determine the maximum tolerable dose of each treatment group. Maximum tolerable dose (MTD) was identified by the percentage just prior to a significant reduction in food intake, a 10% loss in weight, or a significant increase in water intake indicative of exposure to toxins (ref).

Measurement of resting metabolic rate

On the last day of the diet trial, and between 10 a.m. and 2 p.m., woodrats' resting metabolic rate was measured at 23° C. Each metabolic chamber received dry, filtered air at 800 ml min^{-1}. Water and CO_2 were removed from the excurrent air with a column

containing Drierite and Ascarite II. LabVIEW 7.1 (National Instruments, Austin, TX, USA) was used to control incurrent air-flow rate for all chambers for the duration of the trial and to switch solenoid valves to allow for sampling of excurrent air. Control chambers without woodrats were treated identically except they did not contain animals. Initial and final baselines samples (~ 2 min each) of ambient air from the control chambers were obtained immediately before and after each animal was monitored. The data from the control chambers was used to correct for linear drift in the baseline concentration of O_2 during the measurements. Two dual-channel oxygen analyzers enabled us to monitor up to 4 animals simultaneously. Each animal was sampled 3 times per hour for 4 hours with each sampling interval lasting 15 minutes after allowing for washout and collection of baseline data. Animal data for the last period of the 4th hour was not collected for every animal so we use only the first 11 sampling periods to estimate RMR (that is 3 samples per hour for the first three hours and the first two samples for the fourth hour). Excurrent oxygen concentration was averaged and recorded every 5 s. RMR was estimated as the lowest 5-minute steady state rate of O_2 consumption from the eleven 15-min measurement periods for each woodrat using equation 4 from Hill (1972, p. 261). Air flow was regulated using upstream CMOSens mass flow controllers (Sensirion, Zurich, Switzerland), and oxygen content was analyzed using Oxilla II dual channel/differential oxygen analyzers (Sable Systems, Las Vegas, NV, USA).

Dissection

Immediately following metabolic measurements, woodrats were euthanized with isoflurane and cervical dislocation as secondary euthanasia. Tissue from the liver and caecum were rapidly harvested, rinsed with sterile saline, minced on an ice-cooled tray, and immediately placed in a cryovial in liquid nitrogen. All tissues were handled with sterile tools and changed in between different tissue types and across individuals to minimize risk of cross contamination. To reduce batch effect on gene expression, we conducted all dissection in the afternoon between the hours of 1400 – 1600. Following dissections, tissues were moved from liquid nitrogen to -80° C storage.

RNA isolation, sequencing, and analysis

RNA was isolated from liver and caecum tissues using a QIAzol (Qiagen) following Evans et al. (2020), and RNA quality was measured using an Agilent RNA Bioanalyzer for which all samples had RIN scores > 7.

For liver RNA, library preparation and sequencing were performed by the QB3-Berkeley Genomics core labs (Berkeley, CA). rRNA was depleted from samples with a Ribo-Zero Plus rRNA Depletion kit (Illumina 20037135). Both total RNA and depleted RNA quality was assessed on an Agilent 2100 Bioanalyzer. Libraries were prepared using the KAPA RNA Hyper Prep kit (Roche KK8540). Truncated universal stub adapters were ligated to cDNA fragments, which were then extended via PCR using unique dual indexing primers into full length Illumina adapters. Library quality was checked on an AATI Fragment Analyzer. Library molarity was measured via quantitative PCR with the KAPA Library Quantification Kit (Roche KK4824) on a BioRad CFX

Connect thermal cycler. Libraries were then pooled by molarity and sequenced on an Illumina NovaSeq 6000 S4 flowcell for 2 x 150 cycles, targeting at least 25M reads per sample. Fastq files were generated and demultiplexed using Illumina bcl2fastq2 v2.20 and default settings.

One ug of RNA from the caecum was submitted to Novogene (Sacramento, CA) for cDNA library preparation and sequencing. Libraries were made using NEBNext Ultra Directional RNA library prep kit for Illumina (cat# E7420S, New England Biolabs, Ipswich, MA) according to the manufacturer’s protocol. Briefly, mRNA was enriched using oligo(dT) beads followed by two rounds of purification, and fragmented randomly by adding fragmentation buffer. The first strand cDNA was synthesized using random hexamers primer, after which a custom second-strand synthesis buffer (Illumina), dNTPs, RNase H and DNA polymerase I were added to generate the second-strand (ds cDNA). After a series of terminal repair, poly-adenylation, and sequencing adaptor ligation, the double-stranded cDNA library was completed following size selection and PCR enrichment. The resulting 250-350 bp insert libraries were quantified using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA) and quantitative PCR. Size distribution was analyzed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Qualified libraries were sequenced on an Illumina NovaSeq Platform (Illumina, San Diego, CA) using a paired-end 150 run (2×150 bases). 30 M raw reads were generated from each library.

We used TrimGalore (Babraham Bioinformatics) to trim raw RNAseq reads and remove Illumina adapters. We also removed bases with phred scores below 20 and removed any read pairs with less than 20 bp length. Trimmed RNAseq reads were then

aligned to the *N. bryanti* genome (Greenhalgh et al. 2022) using HISAT2 (Kim et al. 2015) and alignments were then converted to sorted and indexed bam files using samtools (Danecek et al. 2021). The *N. bryanti* genome was annotated to contig level using Geneious software. The annotated *N. bryanti* contig-level assembly was then used to count reads using HTSeqv v.2.0.1 (Putri et al. 2022).

16S sequencing and analysis

The caecum of each woodrat was removed, and contents were sampled with either a sterile spatula or cotton swab, placed in a cryovial, and flash frozen in liquid nitrogen. Samples were stored at -80° C until being shipped on ice to Jonah Ventures (Boulder, CO). DNA was extracted from ~0.25 grams per sample using the DNeasy Powersoil HTP 96 kit. Part of the mitochondrial 16S rRNA gene was amplified from each DNA sample using the following primers: 515F GTGYCAGCMGCCGCGGTAA, 806R GGACTACNVGGGTWTCTAAT (Caporaso et al., 2011). Sequences were demultiplexed using Golay barcodes in Qiime v1.9.1 (Caporaso et al., 2010). Pooled libraries were sequenced on Illumina MiSeq (San Diego, CA).

16S bioinformatic processing

Demultiplexed 16S reads (fastq) were processing using QIIME2 version 2021.8.0 (Bolyen et al., 2019). We removed the forward and reverse primers using cutadapt (Martin, 2011). Sequences were denoised using DADA2 (Callahan et al., 2016), with forward and reverse reads truncated at 231 and 230, respectively. We performed de novo clustering using *vsearch* (Rognes et al., 2016), and generated operational taxonomic units

(OTUs) at 99% identity. We used the Naive Bayes classifier to assign taxonomy using the Greengenes database (DeSantis et al. 2006) version 13_8 for 99% OTUs and specific to 515F/806R primers used. We constructed a phylogeny by first aligning sequences with the *alignment mafft* command, then *alignment mask*, and finally *phylogeny fasttree*. The phylogeny was rooted using the *midpoint-root* command.

Statistical analysis

All analyses were conducted in R version 4.2.0 (R Core Team 2022).

During-trial response

We first estimated maximum tolerable dose (MTD), which was determined to be the percent dose of plant extract in chow at which an individual either lost 10% of its initial body mass, or the inflection point at which daily water consumption sharply increased. When challenged by toxic PSCs, mammalian herbivores may increase water consumption to maintain water balance (Dearing et al. 2002). No individuals lost greater than 10% body mass and so water consumption was used to determine MTD. We then averaged MTD for each species by treatment group and performed analysis of variance using the *aov* function in R to test whether MTD differed among treatment groups.

We next examined mass-adjusted during-trial response in water and food intake, wheel running, and body mass to test whether individuals showed significant increases or decreases in these measures between the initial 0% dose and the maximum dose attained. For each species by diet treatment group (4 groups), we used either t tests (*t.test* function) or Kruskal-Wallis (*kruskal.test* function) tests from the *stats* package to compare average

baseline and maximum dose measurements of food consumption, water intake, maximum daily speed (meters per second), daily rotations on wheel (total distance traveled), total time in minutes active on wheel. We visualized these results with boxplots using ggplot2 (Wickham 2016). We also tested for differences in metabolic rate using analysis of covariance on log transformed metabolic rate and adjusted for body mass.

Differential gene expression in liver and caecum

With RNAseq data from the liver and the caecum, we used DESeq2 (Love et al. 2014) to identify genes that were differentially expressed between species, diet treatments, and any that showed a species x diet interaction effect. We then used the *makeContrasts* of the limma (Ritchie et al. 2015) package to explore individual pairwise comparisons (i.e., *N. lepida* versus *N. bryanti*, PRFA versus FRCA). We used a p-value < 0.05 and log2 fold change of > 2 as a threshold for detecting differentially expressed genes. We used *vst* in DESeq2 to estimate dispersion and apply variance stabilizing transformation, then performed a principal components analysis on the resulting data using the *prcomp* function with scale and center set to true. We then plotted PCA using the *fviz_pca_biplot* function in the factoextra (Kassambra and Mundt 2020) package.

Microbial alpha diversity

We imported QIIME2 artifacts including the feature count table, phylogenetic tree, taxonomy table, and sample metadata into R as a phyloseq object (McMurdie and Holmes 2013) using the *qza_to_phyloseq* function of the qiime2R package (Bisanz 2018). Any OTUs representing mitochondria or chloroplast were filtered and we removed

singleton OTUs. We then estimated Faith's phylogenetic diversity and calculated microbial richness using the *estimate_pd* function of the btools package (Battaglia 2022). We plotted diversity using ggplot2 (Wickham 2016) and performed Wilcox tests for differences between diet treatments within *N. lepida* and *N. bryanti* using the *geom_signif* function in the ggpubr package (Kassambra 2020).

Microbial composition between species and diet treatments

We visualized relative read abundances (RRA) of microbial lineages at the level of phylum, family and genus across woodrat species and diet treatments. We first agglomerated OTU reads to the desired taxonomic level using the *tax_glom* function of phyloseq before merging sample types using the *merge_samples* function. We then transformed the merged read counts to RRA using the *transform_sample_counts* function of phyloseq, and taxa that were below 5% RRA were combined into an 'other' category. Results were plotted using ggplot2.

We used principal components analysis (PCA) to characterize differences in microbial composition in the caecum between *N. lepida* and *N. bryanti* and between *P. fasciculata* and *F. californica* diet treatments. We removed reads that originated from either mitochondria or chloroplast and agglomerated the remaining reads at the family level using the *tax_glom* function of phyloseq. We then performed PCA using *vst* to transform read counts and plotted the results using the *fviz_pca_biplot* function as with gene counts (as above).

To quantify the influence of woodrat species, sex, diet, and any species by diet interaction on microbial composition, we performed permutational analysis of variance

(PERMANOVA) using the *adonis2* function of the vegan package (Oksanen et al. 2022). We converted microbial read counts to relative read abundance then used Bray-Curtis and weighted and unweighted distance matrices generated with the *distance* function of phyloseq to perform PERMANOVAs with 999 permutations.

Identifying differentially abundant microbial taxa

We used DESeq2 (Love et al. 2014) to identify microbial OTUs that significantly differed between diet treatments within *N. lepida* and *N. bryanti* individually. For *DESeq* models, we used the *Wald* test, *parametric* fitType and *poscounts* sfType parameters. An adjusted P value of < 0.05 was used to identify OTUs that were significantly differentially abundant. We plotted results using ggplot2.

Co-expression of genes and associated microbiota

Of the caecum genes that exhibited a significant species x diet interaction effect, we performed weighted gene co-expression network analysis using the WGCNA package (Langfelder and Horvath 2008, Langfelder and Horvath 2012) to identify modules of genes that were co-expressed and to associate those modules with experimental treatment groups. Before construction of the network, we determined the power threshold that would result in a scale-free topology, this resulted in a power of 12 for network construction. We used the *signed* TOMtype parameter which results in modules of co-expressed genes that exhibit positive correlations. We also set the minimum module size to 5. To determine the relative expression of gene modules, we used the *lmFit* function of the limma package to perform linear models for each gene module across the four

treatment groups. After the model was run, we used the *eBayes* function (`limma`) to smooth standard errors.

To identify potentially interacting gene networks and individual microbes in the caecum, we then used the *cor* function of WGCNA to estimate Pearson's correlation between microbiota that exhibited significant species x diet interaction effects and gene modules identified. As we are most interested in gene-microbe relationships within each species x diet treatment group, we repeated this analysis for each of the 4 experimental treatment groups. To understand the functional significance of identified gene modules, we input the list of genes in each module to ShinyGO (http://bioinformatics.sdstate.edu/go/) using the Geno Ontology (GO) biological process database.

Results

A total of 29 woodrats from 5 populations were used in laboratory feeding trials (*N. lepida* = 14; *N. bryanti* = 15; S.I. Table S1). After RNA sequencing and processing, we retained 619,253,852 reads in the liver and 881,845,520 reads in the caecum across 22,515 genes. After filtering potential contaminants and removing singletons, we retained 202,388 high-quality reads across 1,444 microbial 16S rRNA OTUs.

Neotoma lepida shows greater during-trial response to diet treatment

Neotoma lepida consuming their non-native *F. californica* was the only treatment group that exhibited a significant during-trial response, had an average maximum

tolerable dose (MTD) of ~60%, and was the only group not to reach 100% MTD (Fig. 1A). This group also exhibited a significant increase in water intake between 0% added plant extract and MTD (Fig. 1b; $P < 0.05$), and a marginal decrease in total minutes per day running between 0% and MTD (Fig. 1c; $P = 0.08$). We did not detect significant differences in other during-trial responses or in resting metabolic rate at the end of the trial (Figs. S1-4).

Strong species by diet interaction effect on caecum gene expression, but not in liver

In the liver, we detected 942 genes that were differentially expressed between species, 17 of which belonged to the cytochrome P450 family (Fig. 2A). We detected 6 genes that were differentially expressed between diets, and 4 genes that exhibited a species by diet interaction. In the caecum, 722 genes were differentially expressed between species, 304 differed between diet, and 367 genes exhibited a species x diet interaction (Fig. 2B). The composition of genes expressed in the caecum when *N. lepida* consumed *F. californica* differed markedly from the composition of the other treatment groups and was strongly driven by expression of sulfotransferase (SULT2B) and the gene ApoA-IV (Fig. 2B). We also identified a different version of a sulfotransferase (SULT2A) expressed in the caecum of *N. bryanti*.

Species and diet influence microbiome composition

We did not detect significant differences in either of two measures of microbial alpha diversity (*i.e.*, Faith's phylogenetic diversity and microbial richness). Despite the similarities in alpha diversity, principal components analysis identified variation in

microbial composition across species and diet treatments (Fig. 3). Furthermore, PCA revealed associations of several microbial families with species and diets—including *Erysipelotrichaceae* with *N. bryanti,* and perhaps more generally with the *F. californica* diet treatment, while the family *Lactobacillaceae* was associated with *N. lepida* (Fig. 3).

We detected a signal from both species and diet in microbiome composition. Using Bray-Curtis distances, species (PERMANOVA, $R^2 = 10\%$, $P < .001$, Table 1) had a stronger effect than diet (PERMANOVA, $R^2 = 6.0\%$, $P < .01$) in shaping microbiome composition. We also estimated beta diversity using both weighted and unweighted UniFrac distances, which incorporate phylogenetic information into distance estimates. Weighted UniFrac distance, which measures community composition (*i.e.*, relative abundances), was similarly influenced by both diet and species ($R^2 = 11.0\%$, $P < .02$; S.I. Table S2). Unweighted UniFrac distance, which measures community membership (*i.e.*, presence or absence), was slightly more influenced by species (PERMANOVA, $R^2 = 7.0\%$, $P < .001$) than diet ($R^2 = 6.0\%$, $P < .001$; S.I. Table S3).

Greater number of differentially abundant microbes between diet treatments in Neotoma lepida than *Neotoma bryanti*

We detected 12 microbial OTUs within two phyla and six families that were differentially abundant between diet treatments in *N. lepida* (S.I. Fig. S5). Of these detected in *N. lepida*, three belonged to *Ruminococcaceae*, three belonged to *Lachnospiraceae*, and one each of *Desulfovibrionaceae*, *Erysipelotrichaceae*, and *Clostridiaceae*. Conversely, we only detected 1 differentially abundant lineage in *N.*

bryanti that increased on the *F. californica* diet—*Ruminococcus flavefaciens*—which also increased in *N. lepida* on its native diet (S.I. Figs. 5&6).

Co-expressed gene networks and their microbial correlates

WGCNA network analysis identified 10 co-expressed sets of genes (*i.e.,* modules) ranging from 5 to 79 genes. These modules were primarily related to immune and metabolism function (Fig. 4A, Table 2). *Neotoma lepida* on the *F. californica* diet exhibited contrasting patterns of module expression when compared to other treatment groups (Fig. 4B&C). The pink, green, and red modules were relatively upregulated in *N lepida* consuming *F. californica*, while these modules exhibited minimal variation among the other species x diet treatment groups (Fig. 4B). These three modules are related to, among other pathways, amino acid, lipid, and carbohydrate metabolism and folate biosynthesis. Among the remaining 7 gene modules, variation existed in expression levels among species x diet treatment groups but were markedly downregulated in *N. lepida* consuming *F. californica* (Fig. 4C).

Within *N. lepida* consuming *F. californica,* bacteria in the order Clostridiales were negatively associated with the pink module, which contains genes related to metabolism (Fig. 5B, Table 2). Also, within *N. lepida* consuming *F. californica*, one species in the genus *Oscillospira* and one in the genus *Clostridium* were also positively associated with the red module, which contained genes involved in metabolism, amino acid synthesis, and folate biosynthesis.

Discussion

The maintenance of unique species—*N. bryanti* and *N. lepida*—despite ongoing hybridization indicates that selection is involved in maintaining this species boundary. Hybridization occurs between these sister species across an ecotone characterized by a sharp transition in vegetation communities, in which the two species maintain distinct and toxic diets (Matocq et al. 2020, Nielsen and Matocq 2021). This suggests that differential diet-related adaptations may influence the distribution of parental individuals, and therefore the degree to which heterospecifics interact and potentially breed. We found an asymmetry in the ability of parental *N. bryanti* and *N. lepida* to consume the natural diet of the other species, and thus tolerate a habitat and diet shift. Furthermore, genes expressed in the caeca of *N. lepida* consuming *F. californica* exhibited markedly different composition than that of other treatment groups, and we identified more differentially abundant microbiota in the caeca of this same group than others. Our results suggest that diet-related adaptations may contribute to an environmental selective gradient operating across this ecotonal hybrid zone, but that this selection may be asymmetrical.

Asymmetry in during-trial response

Neotoma lepida exhibited decreased maximum tolerable dose (MTD) to their non-native *F. californica* diet, while *N. bryanti* did not differ in MTD between diet treatments. *Neotoma lepida* consuming *F. californica* also exhibited increased water intake and reduced time spent wheel running—both ecologically relevant responses that may reduce

fitness of individuals in arid environments. Water balance is important to the physiology of mammalian herbivores, and increased water consumption has been found in woodrats consuming toxic plant PSCs (Dearing et al, 2002). Based on field-measured diets and cafeteria choice trials, *N. bryanti* is considered a facultative dietary generalist and *N. lepida* a facultative dietary specialist (Nielsen and Matocq 2021, Shipley et al. 2009). Results of the experimental feeding trials reported here further support these classifications and indicate that *N. lepida* may be more constrained by diet in their ability to move between habitats than *N. bryanti*.

Variation in gene expression in liver and caecum

Differences in gene expression in the liver of woodrats was strongly driven by species identity, with relatively few genes exhibiting diet or species x diet interaction effects. Among liver genes that differed between species, we identified several within the cytochrome P450 family, a diverse gene family linked to dietary novelty and specialization (Kitanovic et al. 2018, Malenke et al. 2012). This suggests that both *N. bryanti* and *N. lepida* express liver genes in response to dietary toxins in a species-specific manner, regardless of diet.

Whereas gene expression in the liver was strongly influenced by species, gene expression in the caecum, a chambered organ of the gastrointestinal (GI) tract, exhibited remarkable variation among diet treatments, particularly in the dietary specialist—*N. lepida*. Among genes associated with *N. lepida* consuming its non-native diet *F. californica*, we identified a sulfotransferase (SULT2B1), which belongs to a class of enzymes that participate in conjugation and metabolism of a variety of plant toxins

(Dearing et al. 2005). We also detected a different version (SULT2A1) of this gene expressed in the caecum of *N. bryanti*, suggesting that both species utilize these genes, but have come to use differentially adapted families of sulfotransferases. Another locus highly expressed by *N. lepida* consuming *F. californica* was an apolipoprotein (ApoA-IV), known to be involved in metabolism and the stimulation of acute satiation (Wang et al. 2015). This suggests that the response of *N. lepida* to dietary change includes expression of detoxification loci (*i.e.*, SULT), but also may include the induction of a satiety effect (ApoA-IV) to reduce ingestion of a potentially toxic, and unfamiliar, diet item.

Variation in the microbiome of the caecum

We detected a signal of both species and diet in the composition of the caecum microbiome of woodrats (Fig. 3, Table 1). Species signal in shaping microbiome composition was almost twice that of diet when using Bray-Curtis distances, which considers relative abundances of microbial taxa (Table 1). These results are consistent with our previous study in that genotype was the strongest predictor of composition across the hybrid zone (Nielsen et al. 2022), and more broadly across the genus *Neotoma* (Weinstein et al. 2021). Species signal was also slightly greater than diet when using unweighted UniFrac distance, which considers only the presence or absence (*i.e.*, unweighted) and the phylogenetic relationships of microbial taxa (*i.e.*, UniFrac; Table S.I. S3). However, both species and diet explained ~11% of gut microbiome composition when using weighted UniFrac distance, which considers the relative abundances (*i.e.*, weighted) of microbial taxa for which phylogenetic information is considered. These

results suggest that both species and diet influence microbial composition of the caecum, but that woodrat species (*i.e.,* evolutionary history) may be a stronger driver of composition. However, the difference in signal between species and diet decreased when considering the phylogenetic relationships among microbial taxa (*i.e.*, UniFrac), suggesting close evolutionary relationships between the microbiomes of these closely related woodrat species.

We identified several microbial lineages that were differentially abundant between diet treatments within *N. lepida* and *N. bryanti*. We recovered 12 differentially abundant lineages in *N. lepida* and only 1 in *N. bryanti. Neotoma lepida*, the dietary specialist, exhibited greater microbial response to diet treatments, suggesting they harbor a gut microbiome more sensitive to dietary changes. This result fits the pattern of heightened diet response by *N. lepida*, relative to the facultative dietary generalist, *N. bryanti* (Nielsen et al. 2021; Shipley et al. 2009). In non-human primates, dietary specialization has been linked to heightened sensitivity of the microbiome to dietary changes (Frankel et al. 2019). The heightened sensitivity of the microbiomes of specialists could be the result of having a more specialized microbiome community tuned to metabolism of specific plant compounds. Alternatively, but not exclusive of the latter, specialist herbivores may experience direct physiological impacts of altered diets leading to an increase in pathology-related microbial lineages.

When eating their non-native *F. californica* diet, *N. lepida* displayed an increased abundance of several microbial lineages that provide insight into the complexity of host and microbial response to altered diets. These individuals had greater abundance of *Allobaculum*, a microbial genus linked to lipid metabolism and reduction of inflammation

(Zheng et al. 2021). The genus *Desulfovibrio* also increased in abundance in the caecum, which may be linked to ulcerative colitis in the colon (Rowan et al. 2010). As such, the response of some microbes may provide a benefit to the woodrat host (*i.e*., metabolism, reduction of inflammation), while that of other lineages may signal a pathological response. Interestingly, both *N. lepida* and *N. bryanti,* when feeding on their respective native diet, had increased abundance of *Ruminococcus flavefaciens,* a bacterial species that degrades cellulose (Varel et al. 1984). This implies that both these species harbor microbial lineages important to digestion of plants and that the relative abundance of these core functional lineages may be heightened when consuming a familiar diet.

Associations between gene co-expression and microbial taxa

The GI tract maintains a symbiosis with gut microbiota, and its composition is important to host health (Arnolds and Lozupone 2016). In the woodrat caecum, microbes may contribute to metabolism of plant toxins (Kohl et al. 2018, Miller et al. 2014). Within *N. lepida* consuming their non-native diet, several microbial lineages exhibited significant associations with gene co-expression modules. For example, genes underlying one-carbon metabolism and related biosynthesis of lipids and amino acids were upregulated and negatively correlated with the bacterial order Clostridiales (Figs. 4B & 5B). In contrast, there was a positive association between genes related to metabolism and amino acid synthesis, and the genus *Oscillospira*, which produce short-chain fatty acids and are generally considered to be beneficial (Yang et al. 2021; Fig. 5B).

There is growing appreciation of the importance of interactions between host gene expression and the gut microbiome (Nichols and Davenport 2021). Expression of some

CYP P450 enzymes in the intestines of germ-free mice was lower compared to conventional mice, suggesting an important role of gut microbes in regulating xenobiotic-metabolizing genes of the host (Fu et al. 2017). Another study characterized microbe-mediated regulation of host genes along the GI tract, including anti-microbial genes (Larsson et al. 2012). Our gene expression results are limited to the genes identified from RNAseq, not the translated proteins themselves. Likewise, our sampling of the microbiome only reflects the presence of lineages and not an inventory of the genes and proteins being expressed by the microbiome. However, our findings provide a foundation for generating and testing hypotheses related to interactions between the genomes of woodrats and the microbes they host.

Eco-evolutionary implications for hybridization between N. lepida and *N. bryanti*

For mammalian herbivores, toxic plant compounds play a critical role in nutritional ecology (Dearing et al. 2005). Based on mtDNA evidence, *N. lepida* and *N. bryanti* are estimated to have diverged ~1.6 million years ago (Patton et al. 2007). During this period of allopatric divergence, both species have accumulated different ecological adaptations, including those related to diet. We found that the species had distinct responses to diet shifts that spanned physiology/behavior (*i.e.,* maximum tolerable dose, water intake, wheel running), gene expression, and compositional shifts in the gut microbiome. However, the magnitude of these responses appears to be asymmetrical, with the dietary specialist, *N. lepida,* exhibiting greater dietary response than the generalist, *N. bryanti*.

Of particular interest is our finding that gene expression in the caecum showed much stronger species by diet interaction effects than the liver. Liver gene expression is a major source of metabolic and detoxification enzymes that has been well studied in the genus *Neotoma* (Kitanovic et al. 2018, Malenke et al. 2012), and that differed greatly between the species studied here, regardless of diet treatment. While less studied than the liver, metabolic and detoxification enzymes—including cytochrome P450s—are also expressed by the host in the intestinal lining (Hall et al. 1999); and in the caecum, there is great opportunity for host-microbial interaction. Indeed, we found a significant species by diet interaction effect in caecum gene expression and microbial community changes, especially in the dietary specialist, *N. lepida*, on its non-native diet. Microbial changes in the caecum are likely a combination of direct response of the microbial community to changes in dietary plant compounds (*i.e.*, changes in nutrients and PSCs), but also indirect response to dietary change through interaction with host gene expression and physiological condition.

Both species are largely spatially segregated across the ecotone where secondary contact occurs (Shurtliff et al. 2014), and plant community composition also sharply differs across this site (Nielsen and Matocq 2021). Availability of suitable diet plants likely produces a selective gradient, across which migrants may be selected against, reducing the frequency with which heterospecific individuals interact and produce hybrid offspring (Via 1999, Via et al. 2000, Nosil et al. 2005). However, some individuals of each species can occupy the alternative habitat, and both species will consume some of the alternative diet plant (Nielsen et al. 2022). The asymmetry in diet response we identified suggests that *N. lepida* may be more limited by diet than *N. bryanti* in its ability

to occupy the alternative habitat. Diet-related selection is only one of many factors determining the degree to which these species remain largely spatially segregated across the ecotone of this site, but our study provides mechanistic insight into the centrality of ecology in maintaining species boundaries.

Acknowledgements

This chapter was a huge collaborative effort with current and former lab members, as well as others of the UNR EECB community, and of which will be coauthors on the future publication. Matt Holding was critical in analysis of RNAseq data, which would have taken me years. Jack Hayes provided lab space and expertise for measuring metabolic rate and conducting dissections. Josh Hayes programmed the ’Ratalyzer’ Raspberry PI units for collecting running wheel data. Denise Dearing loaned us experimental caging for use during these trials. I also thank Jennifer Forbey and Denise Dearing for helpful discussions about experimental design.

I also thank lab technicians who made these trials possible: Dan Moore, Angela Mitchell, and Sydney Stephens. Also Karli Graski for managing and organizing lab samples.

Literature Cited

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Arnolds, K. L., & Lozupone, C. A. (2016). Striking a Balance with Help from our Little Friends – How the Gut Microbiota Contributes to Immune Homeostasis. *The Yale Journal of Biology and Medicine*, *89*(3), 389–395.

Barton, N. H., & Hewitt, G. M. (1985). Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, *16*, 113–148.

Battaglia T (2022). btools: A suite of R function for all types of microbial diversity analyses_. R package version 0.0.1.

Bisanz J. E. (2018). qiime2R: Importing QIIME2 artifacts and associated data into R sessions. https://github.com/jbisanz/qiime2R.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, *108*(Supplement 1), 4516–4522.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Dearing, D. M., Mangione, A. M., & Karasov, W. H. (2002). Ingestion of plant secondary compounds causes diuresis in desert herbivores. *Oecologia*, *130*(4), 576–584. https://doi.org/10.1007/s00442-001-0834-z

Dearing, M. D., Foley, W. J., & McLean, S. (2005). The Influence of Plant Secondary Metabolites on the Nutritional Ecology of Herbivorous Terrestrial Vertebrates. *Annual Review of Ecology, Evolution, and Systematics*, *36*, 169–189.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., & Andersen, G. L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology*, *72*(7), 5069–5072. https://doi.org/10.1128/AEM.03006-05

Evans, L. W., Bender, A., Burnett, L., Godoy, L., Shen, Y., Staten, D., Zhou, T., Angermann, J. E., & Ferguson, B. S. (2020). Emodin and emodin-rich rhubarb inhibits histone deacetylase (HDAC) activity and cardiac myocyte hypertrophy. *The Journal of Nutritional Biochemistry*, *79*, 108339. https://doi.org/10.1016/j.jnutbio.2019.108339

Frankel, J. S., Mallott, E. K., Hopper, L. M., Ross, S. R., & Amato, K. R. (2019). The effect of captivity on the primate gut microbiome varies with host dietary niche. *American Journal of Primatology*, *81*(12), e23061. https://doi.org/10.1002/ajp.23061

Freeland, W. J., & Janzen, D. H. (1974). Strategies in Herbivory by Mammals: The Role of Plant Secondary Compounds. *The American Naturalist*, *108*(961), 269–289.

Fu, Z. D., Selwyn, F. P., Cui, J. Y., & Klaassen, C. D. (2017). RNA-Seq Profiling of Intestinal Expression of Xenobiotic Processing Genes in Germ-Free Mice. *Drug Metabolism and Disposition*, *45*(12), 1225–1238. https://doi.org/10.1124/dmd.117.077313

Greenhalgh, R., Holding, M. L., Orr, T. J., Henderson, J. B., Parchman, T. L., Matocq, M. D., Shapiro, M. D., & Dearing, M. D. (2021). *Trio-binned genomes of the woodrats* Neotoma bryanti *and* N. lepida *reveal novel gene islands and rapid copy number evolution of xenobiotic metabolizing cytochrome p450 genes* [Preprint]. Genomics. https://doi.org/10.1101/2021.03.08.434435

Hall, S. D., Thummel, K. E., Watkins, P. B., Lown, K. S., Benet, L. Z., Paine, M. F., ... & Wrighton, S. A. (1999). Molecular and physical mechanisms of first-pass extraction. *Drug Metabolism and Disposition*, *27*(2), 161-166.

Hill RW (1972). Determination of oxygen consumption by use of the paramagnetic oxygen analyzer. J. Appl. Physiol. 33: 261-263.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

Kitanovic, S., Orr, T. J., Spalink, D., Cocke, G. B., Schramm, K., Wilderman, P. R., Halpert, J. R., & Dearing, M. D. (2018). Role of cytochrome P450 2B sequence variation and gene copy number in facilitating dietary specialization in mammalian herbivores. *Molecular Ecology*, *27*(3), 723–736. https://doi.org/10.1111/mec.14480

Kohl, K. D., Oakeson, K. F., Orr, T. J., Miller, A. W., Forbey, J. S., Phillips, C. D., Dale, C., Weiss, R. B., & Dearing, M. D. (2018). Metagenomic sequencing provides insights into microbial detoxification in the guts of small mammalian herbivores

(Neotoma spp.). *FEMS Microbiology Ecology*, *94*(12), fiy184. https://doi.org/10.1093/femsec/fiy184

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Dearing, M. D. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology Letters*, *17*(10), 1238–1246. https://doi.org/10.1111/ele.12329

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Langfelder, P., & Horvath, S. (2012). Fast R Functions for Robust Correlations and Hierarchical Clustering. *Journal of Statistical Software*, *46*, 1–17. https://doi.org/10.18637/jss.v046.i11

Larsson, E., Tremaroli, V., Lee, Y. S., Koren, O., Nookaew, I., Fricker, A., Nielsen, J., Ley, R. E., & Bäckhed, F. (2012). Analysis of gut microbial regulation of host gene expression along the length of the gut and regulation of gut microbial ecology through MyD88. *Gut*, *61*(8), 1124–1131. https://doi.org/10.1136/gutjnl-2011-301104

Love, M.I., Huber, W., Anders, S. Moderated estimation of fold change and dispersion for

RNA-seq data with DESeq2 Genome Biology 15(12):550 (2014)

Malenke, J. R., Magnanou, E., Thomas, K., & Dearing, M. D. (2012). Cytochrome P450 2B Diversity and Dietary Novelty in the Herbivorous, Desert Woodrat (Neotoma lepida). *PLOS ONE*, *7*(8), e41510. https://doi.org/10.1371/journal.pone.0041510

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200

Matocq, M. D., Ochsenrider, K., Jeffrey, C. S., Nielsen, D., & Richards, L. A. (2020). Fine-scale differentiation in diet and metabolomics of small mammals across a sharp ecological transition. *Frontiers in Ecology and Evolution*, *8*, 282.

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.

Miller, A. W., Kohl, K. D., & Dearing, M. D. (2014). The Gastrointestinal Tract of the White-Throated Woodrat (Neotoma albigula) Harbors Distinct Consortia of Oxalate-Degrading Bacteria. *Applied and Environmental Microbiology*, *80*(5), 1595–1601. https://doi.org/10.1128/AEM.03742-13

Nichols, R. G., & Davenport, E. R. (2021). The relationship between the gut microbiome and host gene expression: A review. *Human Genetics*, *140*(5), 747–760. https://doi.org/10.1007/s00439-020-02237-0

Nielsen, D. P., & Matocq, M. D. (2021). Differences in dietary composition and preference maintained despite gene flow across a woodrat hybrid zone. *Ecology and Evolution*, *11*(9), 4909–4919. https://doi.org/10.1002/ece3.7399

Nosil, P., Vines, T. H., & Funk, D. J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, *59*(4), 705–719.

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.,

Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package* (2.5-7) [Computer software]. https://CRAN.R-project.org/package=vegan

Patton, J. L., Huckaby, D. G., & Álvarez-Castañeda, S. T. (2007). *The evolutionary history and a systematic revision of woodrats of the Neotoma lepida group* (Vol. 135). Univ of California Press.

Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E., & Zanini, F. (2022). Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics*, *38*(10), 2943–2945. https://doi.org/10.1093/bioinformatics/btac166

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma
powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 43(7), e47.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Rowan, F., Docherty, N. G., Murphy, M., Murphy, B., Coffey, J. C., & O'Connell, P. R. (2010). Desulfovibrio Bacterial Species Are Increased in Ulcerative Colitis. *Diseases of the Colon & Rectum*, *53*(11), 1530–1536. https://doi.org/10.1007/DCR.0b013e3181f1e620

Shipley, L. A., Forbey, J. S., & Moore, B. D. (2009). Revisiting the dietary niche: When is a mammalian herbivore a specialist? *Integrative and Comparative Biology*, *49*(3), 274–290. https://doi.org/10.1093/icb/icp051

Sikes, R. S., Care, A., & Mammalogists, U. C. of the A. S. of. (2016). 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *Journal of Mammalogy*, *97*(3), 663–688.

Sorensen, J. S., McLister, J. D., & Dearing, M. D. (2005). Plant Secondary Metabolites Compromise the Energy Budgets of Specialist and Generalist Mammalian Herbivores. *Ecology*, *86*(1), 125–139. https://doi.org/10.1890/03-0627

Varel, V. H., Fryda, S. J., & Robinson, I. M. (1984). Cellulolytic bacteria from pig large intestine. *Applied and environmental microbiology*, *47*(1), 219-221.

Via, S. (1999). Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, *53*(5), 1446–1457.

Via, S., Bouck, A. C., & Skillman, S. (2000). Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution*, *54*(5), 1626–1637.

Wang, F., Kohan, A. B., Lo, C.-M., Liu, M., Howles, P., & Tso, P. (2015). Apolipoprotein A-IV: A protein intimately involved in metabolism. *Journal of Lipid Research*, *56*(8), 1403–1418. https://doi.org/10.1194/jlr.R052753

Weinstein, S. B., Martínez-Mota, R., Stapleton, T. E., Klure, D. M., Greenhalgh, R., Orr, T. J., Dale, C., Kohl, K. D., & Dearing, M. D. (2021). Microbiome stability and structure is governed by host phylogeny over diet and geography in woodrats (*Neotoma* spp.). *Proceedings of the National Academy of Sciences*, *118*(47), e2108787118. https://doi.org/10.1073/pnas.2108787118

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Yang, J., Li, Y., Wen, Z., Liu, W., Meng, L., & Huang, H. (n.d.). Oscillospira—A candidate for the next-generation probiotics. *Gut Microbes*, *13*(1), 1987783. https://doi.org/10.1080/19490976.2021.1987783

Zheng, Z., Lyu, W., Ren, Y., Li, X., Zhao, S., Yang, H., & Xiao, Y. (2021). Allobaculum Involves in the Modulation of Intestinal ANGPTLT4 Expression in Mice Treated by High-Fat Diet. *Frontiers in Nutrition*, *8*. https://www.frontiersin.org/articles/10.3389/fnut.2021.690138

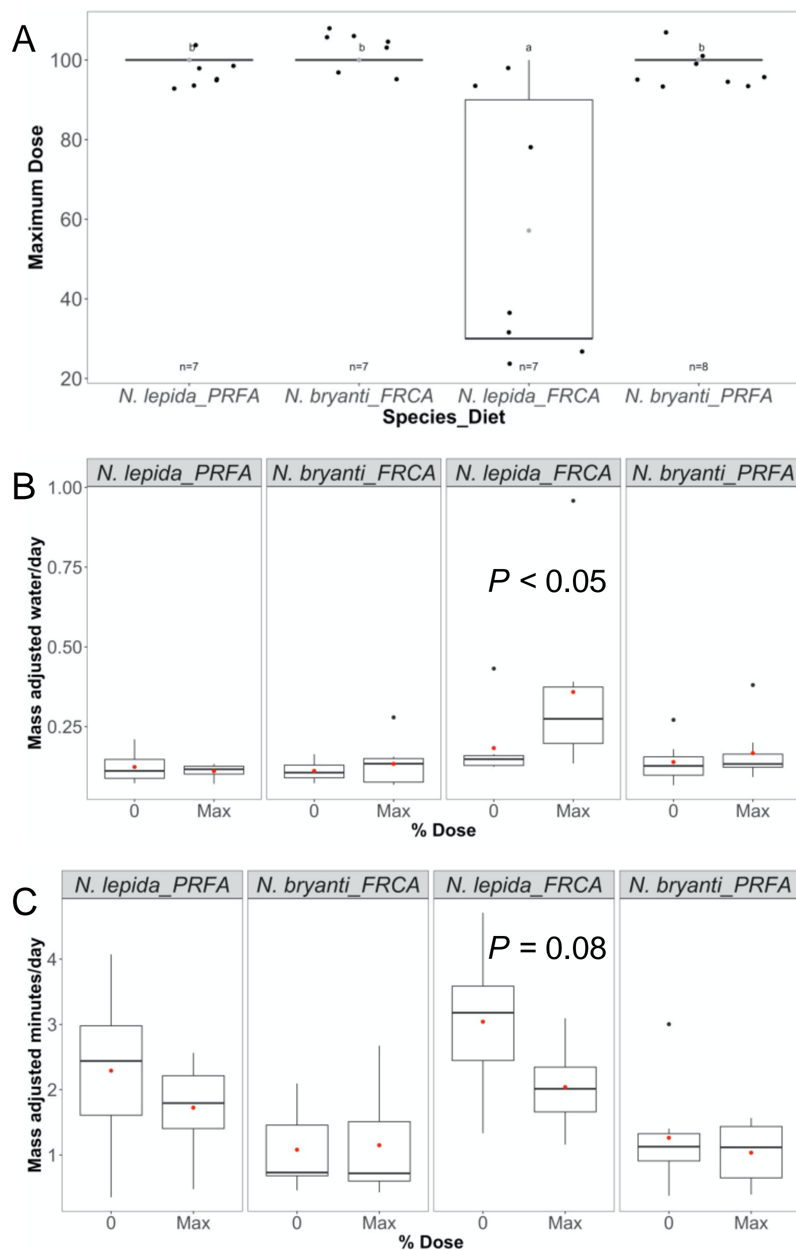

Figure 1: Maximum dose (A), mass adjusted daily water consumption (B), and mass adjusted daily total time spent running on wheel (C). *Neotoma lepida* consuming its non-native *Frangula californica* diet had a significantly lower maximum sustainable dose than the other treatment groups. *N. lepida* consuming *F. californica* also significantly increased their water consumption at maximum tolerable dos (MTD), and showed a marginal decrease in minutes of wheel running per day at the maximum sustainable dose. *P*-values calculated using Kruskal-Wallis tests.

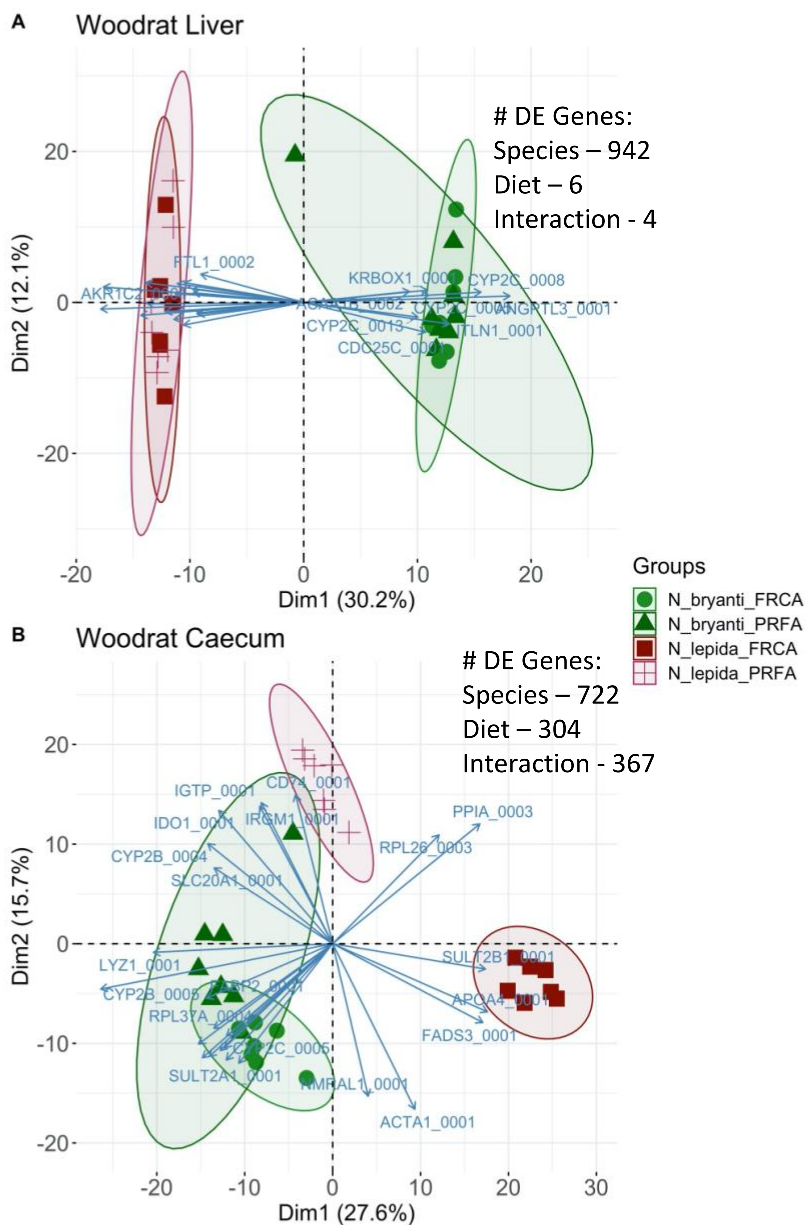

Figure 2: Principal components analysis of the composition of genes expressed in the liver of experimental treatment groups with 25 genes with the highest loadings. Also, are the number of genes that exhibited species, diet, and interaction effects. Points clustering closer together indicate similar composition of expressed genes. Genes expressed in the liver showed strong species signal; only 6 genes showed differential expression between diet treatments, and 4 genes exhibited a species x diet interaction effect in the liver (A). Genes expressed in the caecum showed strong diet and interaction effects (B).

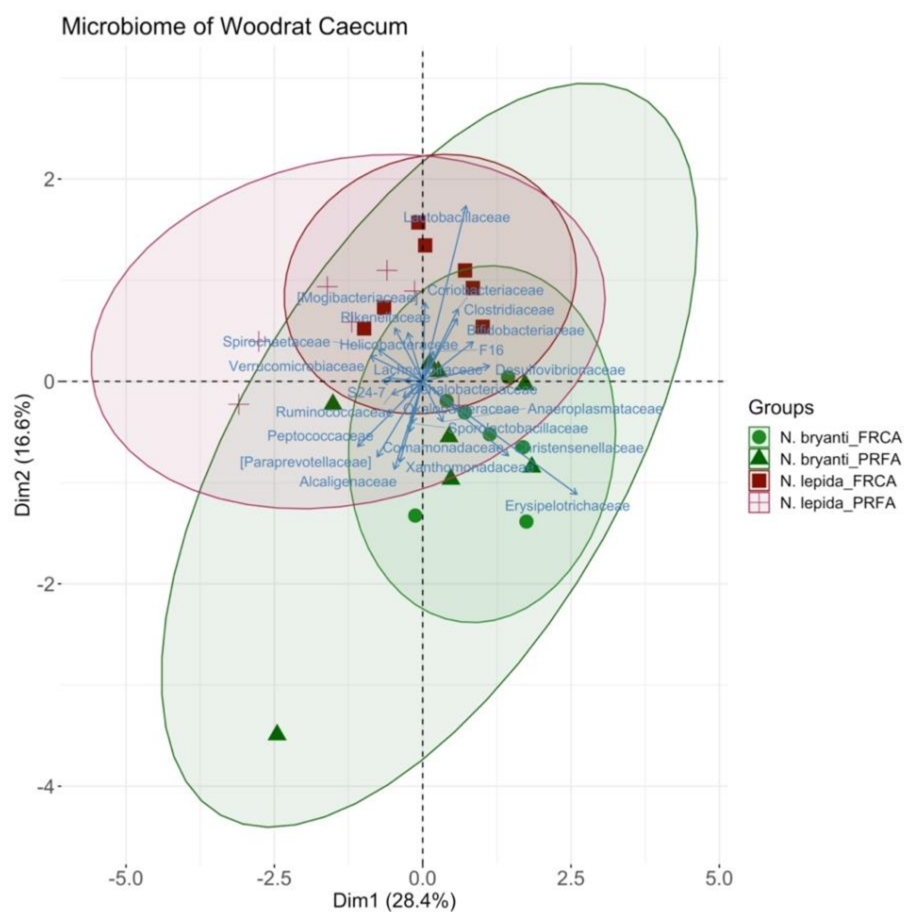

Figure 3: Principal components analysis of the microbiome composition of the caecum contents of woodrats with 25 microbial families with the highest loadings.

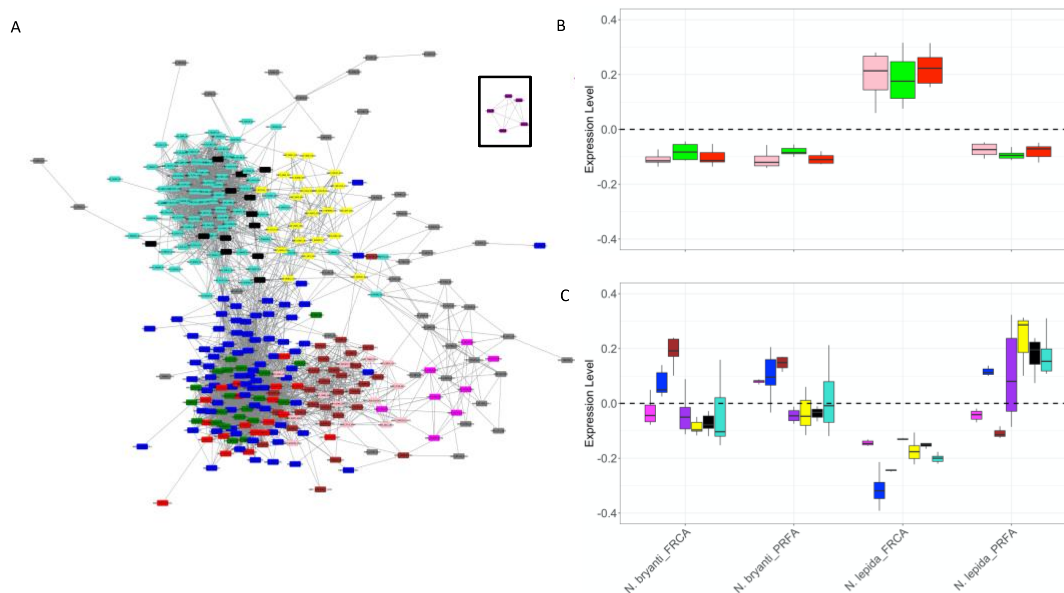

Figure 4: WGCNA gene network (A), and relative expression level of each module across species x diet treatment groups. The y-axis represents coefficients of linear models estimated using the *lmFit* function of the limma R package. The pink, green and red modules were similar across treatment groups, except in *N. lepida* consuming *F. californica* where relative expression of these modules was greater (B). Conversely, the remaining modules showed moderated levels of variability across treatment groups, but a clear decrease in relative expression in *N. lepida* consuming *F. californica* (C).

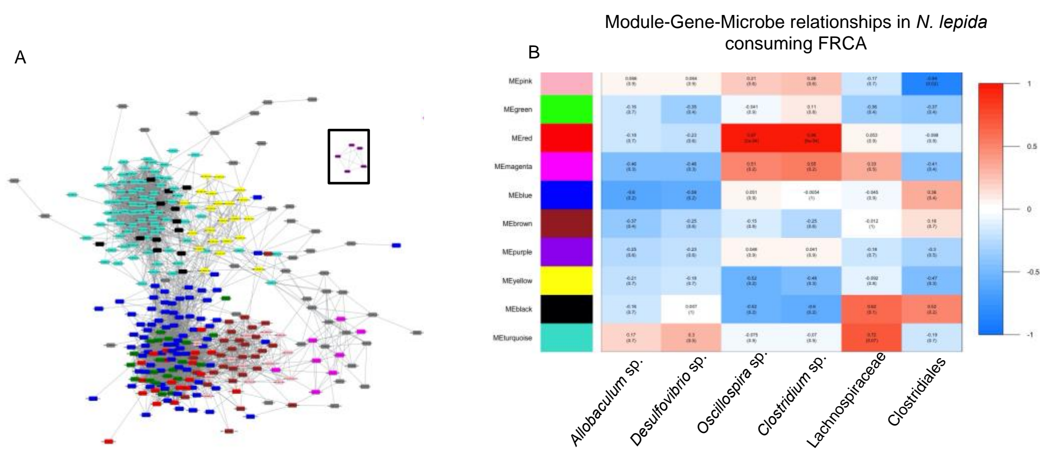

Figure 5: Network diagram of ten gene modules identified from WGCNA network analysis (A). Correlation heatmap of microbial taxa in caecum contents of *N. lepida* consuming *F. californica* diet (B). Values in the heatmap are the Pearson's correlation coefficient (top row) and the p-value (bottom row in parentheses).

Table 1: PERMANOVA with Bray-Curtis distances and using 999 permutations. Species had the greatest effect on caecum microbiome composition, with diet treatment also influencing microbiome.

Bray-Curtis	Df	SumOfSqs	R2	F	Pr(>F)
Species	1.00	0.93	0.10	3.17	0.001
Diet_treatment	1.00	0.55	0.06	1.87	0.01
Sex	1.00	0.39	0.04	1.33	0.08
Species:Diet_treatment	1.00	0.34	0.04	1.14	0.22
Residual	24.00	7.03	0.76	NA	NA
Total	28.00	9.23	1.00	NA	NA

Table 2: WGCNA network module colors and the number of genes within each eigengene. The grey module captures remaining genes that did not fit into any cluster of genes.

Module color	# Genes	General Function
Turquoise-1	79	Immune; response to other organisms
Blue-2	76	Metabolism
Brown-3	29	Metabolism; response to, and metabolism of, xenobiotic stimulus
Yellow-4	22	Reproduction; tissue development
Green-5	21	Development
Red-6	20	Metabolism
Black-7	13	Metabolism
Pink-8	13	Metabolism
Magenta-9	8	Immune
Purple-10	5	Metabolism
Grey	81	—

Supporting Information to Chapter 3:

Differential response of specialist and generalist herbivores when switching diets: host gene expression, the gut microbiome, and their potential interaction

Danny Nielsen, Matthew Holding, Bradley Ferguson, Lora Richards, Jack Hayes, Denise Dearing, Jennifer Forbey, Marjorie Matocq

Table S1: Number of individual woodrats collected from 5 populations and used in laboratory feeding trials.

Population	Species	Lat	Long	# Samples
Pine Tree Canyon	N. lepida	35.232113	-118.09236	10
Cameron Rd.	N. lepida	35.090861	-118.30958	4
Erskine Creek	N. bryanti	35.584644	-118.42871	12
Kernville canyon; Mountain Rd.	N. bryanti	35.981128	-118.43782	1
Kernville canyon; Chico Flat	N. bryanti	35.827301	-118.45883	2

Table S2: PERMANOVA with Weighted UniFrac distances and using 999 permutations. Species and diet had similar effect on caecum microbiome composition.

Weighted UniFrac	**Df**	**SumOfSqs**	**R2**	**F**	**Pr(>F)**
Species	1	0.07	0.11	3.83	0.01
Diet_treatment	1	0.07	0.11	3.86	0.02
Sex	1	0.02	0.03	1.18	0.26
Species:Diet_treatment	1	0.02	0.03	0.99	0.37
Residual	24	0.46	0.71	NA	NA
Total	28	0.65	1.00	NA	NA

Table S3: PERMANOVA with Unweighted UniFrac distances and using 999 permutations. Species had slightly more influence on caecum microbiome composition than did diet.

Unweighted UniFrac	**Df**	**SumOfSqs**	**R2**	**F**	**Pr(>F)**
Species	1	0.45	0.07	2.04	0.00
Diet_treatment	1	0.42	0.06	1.91	0.00
Sex	1	0.27	0.04	1.25	0.06
Species:Diet_treatment	1	0.26	0.04	1.20	0.11
Residual	24	5.24	0.79	NA	NA
Total	28	6.64	1.00	NA	NA

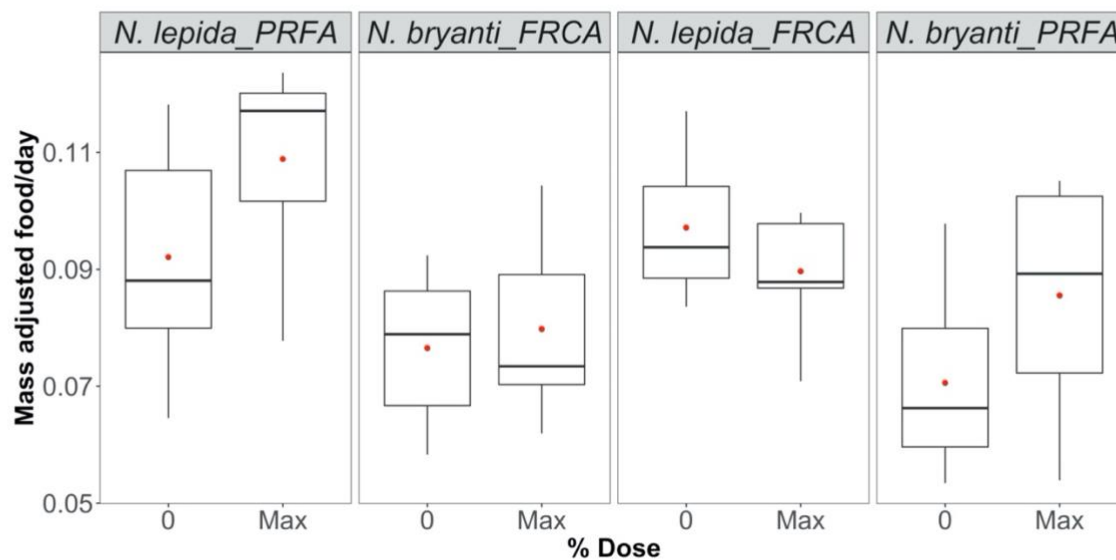

Figure S1: Mass adjusted food consumed per day (grams) did not differ significantly between 0% and maximum dose for any treatment group. Plant four letter codes are for *Prunus fasciculata* (PRFA) and *Frangula californica* (FRCA).

‘

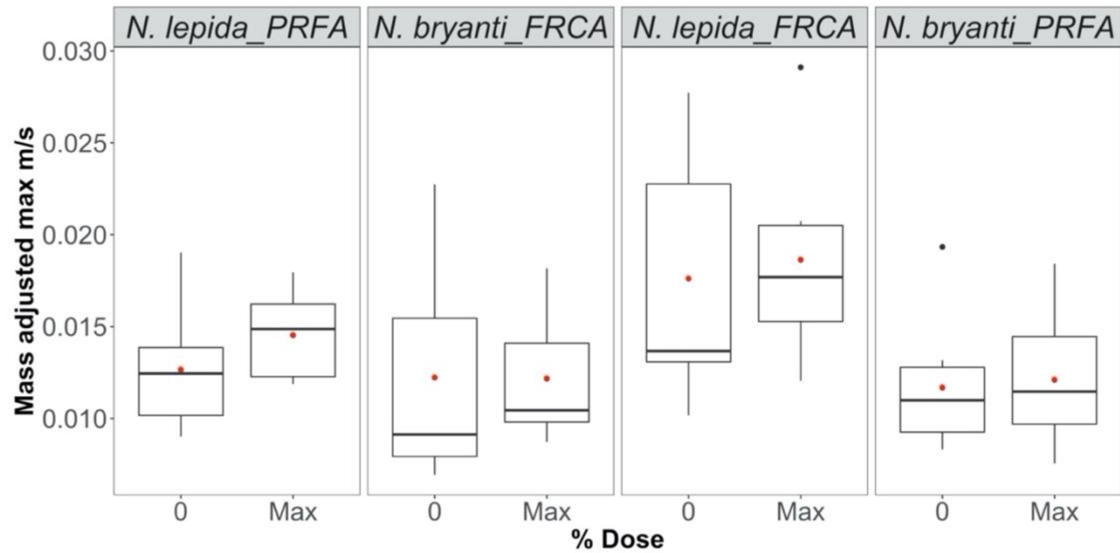

Figure S2: Mass adjusted maximum speed (meters per second) did not differ significantly between 0% and maximum dose for any treatment group. Plant four letter codes are for *Prunus fasciculata* (PRFA) and *Frangula californica* (FRCA).

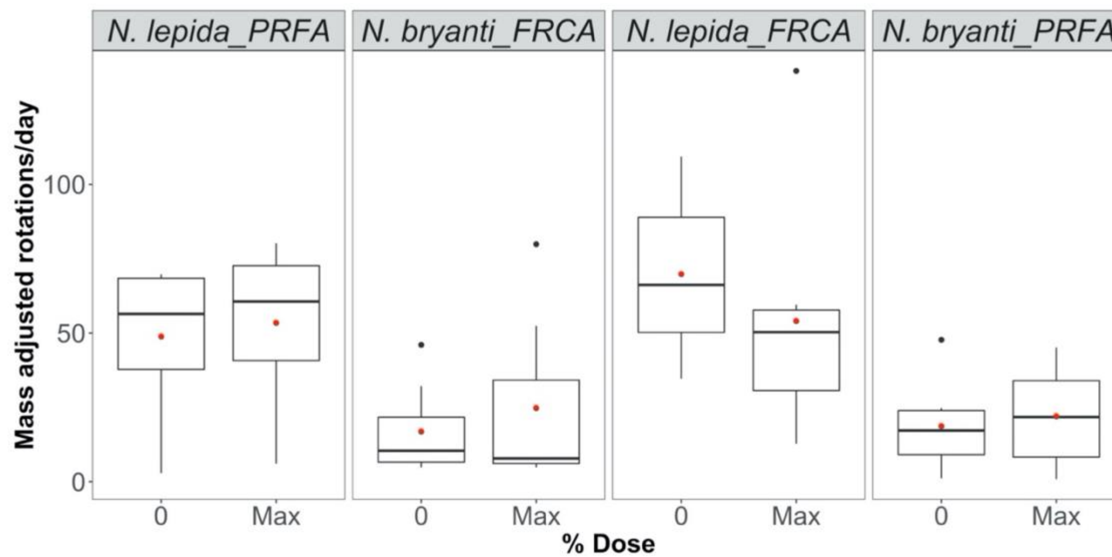

Figure S3: Mass adjusted total wheel rotations per day did not differ significantly between 0% and maximum dose for any treatment group. Plant four letter codes are for *Prunus fasciculata* (PRFA) and *Frangula californica* (FRCA).

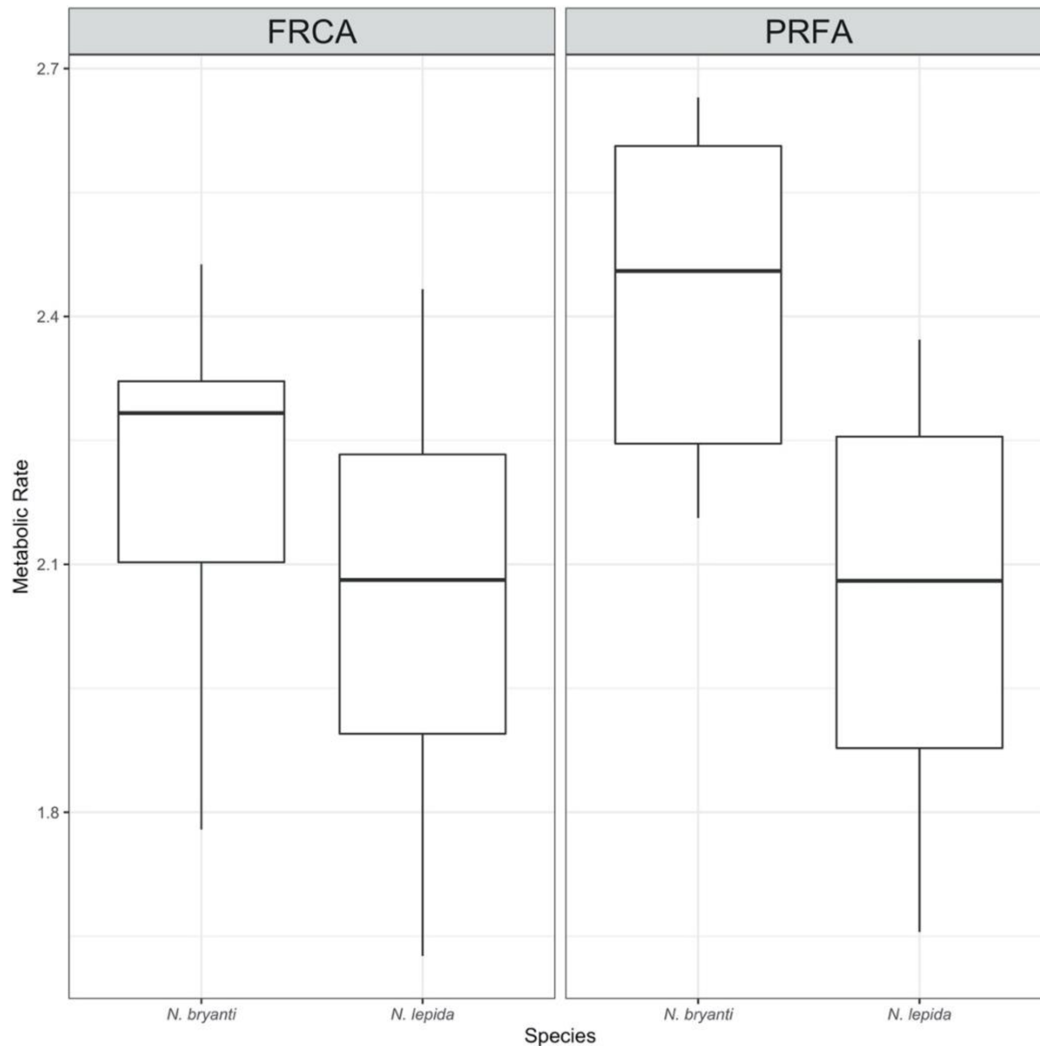

Figure S4: Resting metabolic rate (ml O_2/min) at 23 C did not differ significantly between species, or between diet types. Although *N. bryanti* appears to have greater metabolic rate, particularly on the *P. fasciculata* diet, there is no significant difference when adjusting for body mass. Plant four letter codes are for *Prunus fasciculata* (PRFA) and *Frangula californica* (FRCA).

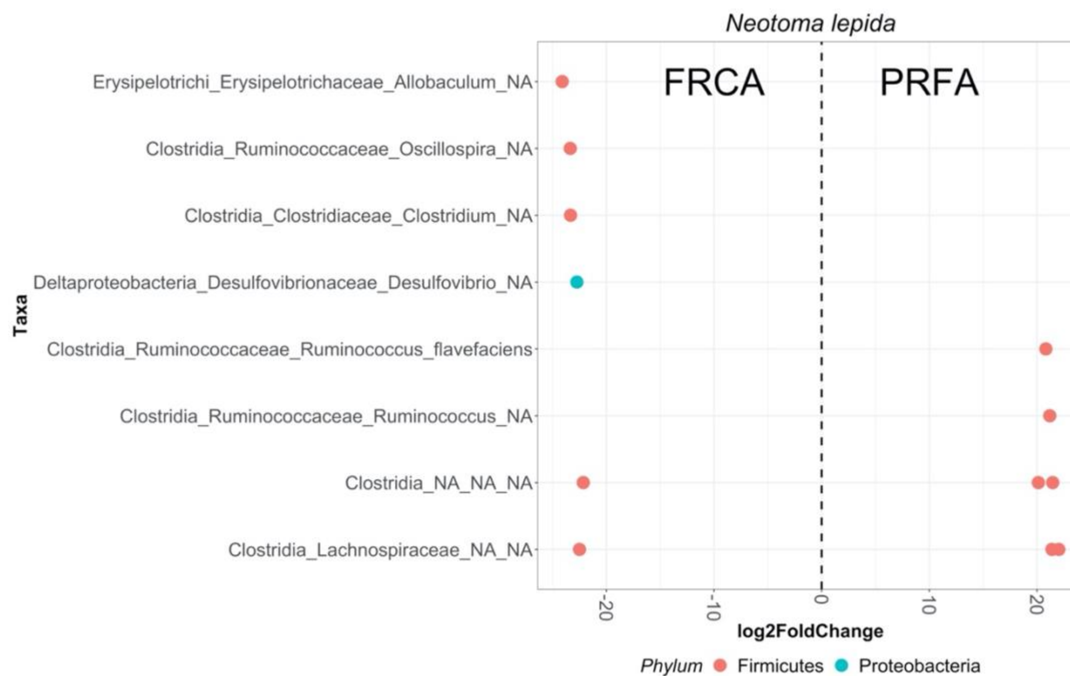

Figure S5: Microbial taxa in the caecum of *Neotoma lepida* that were differentially abundant between diet treatments. Log 2 fold change values less than 0 were more abundant in *N. lepida* individuals on the *F. californica* diet treatment, and values greater than 0 were more abundant on the *P. fasciculata* diet treatment. Where points are on either side of the 0 line indicate multiple microbial OTUs detected within that taxonomic level. Taxa were identified with P-value < 0.05 using `DESeq2`.

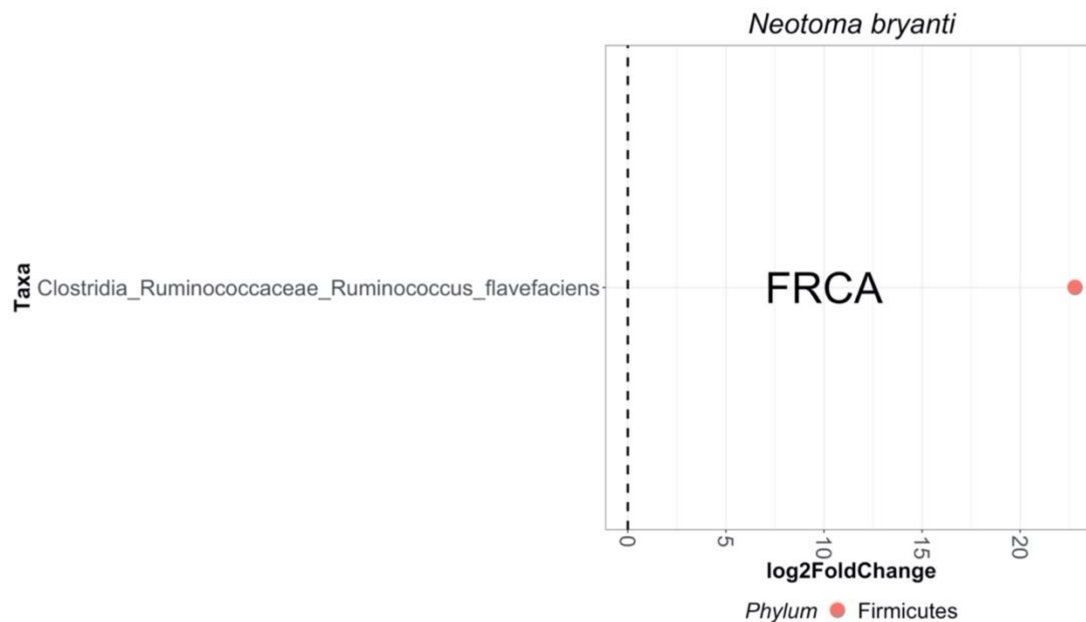

Figure S6: A single microbial taxa in the caecum of *Neotoma bryanti* that was differentially abundant between diet treatments. Log 2 fold change value greater than 0 indicates that *Ruminococcus flavefaciens* was more abundant in individuals consuming the *F. californica* diet. Taxa were identified with P-value < 0.05 using `DESeq2`.

Concluding remarks

This dissertation investigated the role of diet and diet related differential adaptation in determining patterns of gene flow across a small mammal ecotonal hybrid zone. The dietary adaptations my research suggests are among other ecological mechanisms that likely contribute to patterns of hybridization in this system. For example, while *N. lepida* may have more difficulty switching to a hill-type diet, they are also (perhaps more proximally) excluded from occupying the more environmentally buffered rocky hill habitat hill by the larger and more aggressive *N. bryanti*. Conversely, while the dietary generalist *N. bryanti* may be able to shift its diet to that available on the flats, their larger body size and presumed mesic adaptations may limit their ability to occupy the more xeric flats habitat. As such, diet is likely only one axis of variation in a complex web of ecological interactions that operate across this species boundary. Nonetheless, how these woodrats acquire nutrients while limiting toxin exposure plays a central role in energy balance, which ultimately constrains or buffers the degree to which individual fitness changes with environmental conditions.