2023

WHITE-TAILED DEER IN A WINTER-WONDERLAND: LONG-TERM DEER YARD USE AND METHODOLOGICAL CONSIDERATIONS FOR UNGULATE FECAL DNA METABARCODING

Melanie A. Ottino
Michigan Technological University, maottino@mtu.edu

WHITE-TAILED DEER IN A WINTER-WONDERLAND: LONG-TERM DEER YARD USE AND METHODOLOGICAL CONSIDERATIONS FOR UNGULATE FECAL DNA METABARCODING

By

Melanie A. Ottino

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Forest Molecular Genetics and Biotechnology

MICHIGAN TECHNOLOGICAL UNIVERSITY

2023

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Forest Molecular Genetics and Biotechnology.

College of Forest Resources and Environmental Science

Thesis Co-Advisor: *Dr. Kristin E. Brzeski*

Thesis Co-Advisor: *Dr. Carsten Külheim*

Committee Member: *Dr. Christopher R. Webster*

College Dean: *David J. Flaspohler*

Table of Contents

List of Figures

List of Tables

Acknowledgements

To begin, I offer a most heartfelt thank you to each of my incredible advisors, Dr. Kristin Brzeski, Dr. Carsten Külheim, and Dr. Christopher Webster. The value of your mentorship, collaboration, and support throughout this project is immeasurable, and I am grateful to have had the opportunity to work with each of you.

On the funding front, a special thanks to Dr. Webster and Dr. Brzeski for supporting me as a Graduate Teaching and Research Assistant respectively—the luckiest of us get paid to spend time outside talking about plants, and I have been lucky indeed. Additionally, I would like to recognize Michigan Technological College of Forest Resources and Environmental Science, the Ecosystem Science Center, the Graduate Student Government, and the National Service Fellowship for their financial support.

To my peers within the Conservation Genetics Laboratory—your support, assistance, and friendship lightened the burden of every failed PCR and tortuous bead-clean. To my broader UP community, you made this cold town feel vibrantly warm. To Jazz and Katie in particular, I couldn't have done this without you. Nathan, thanks for catching frogs with (*for*) me.

Another huge round of thank yous to my friends and family back home and everywhere, you make my world go round. Lastly, to the beautiful planet we live on; may my work and life continue to respect, protect, conserve, and restore these spaces and the diverse life within them for many generations to come.

List of Abbreviations

ASV	amplicon sequence variant
BA	Basal Area (m2/ha)
CL	Crown Length (m)
CR	Crown Ratio
CSR	Crown Spread Ratio
CT	Crown Thickness
CW	Crown Width
DNA	deoxyribose nucleic acid
eDNA	environmental DNA
FI	Form Index
HWA	hemlock wooly adelgid
ITS2	nuclear Internal Transcribed Spacer DNA
Msat	microsatellites
mtDNA	mitochondrial DNA
rbcL	chloroplast rbcL gene
TPH	Tree Per Hectare (tree/ha)
trnL	chloroplast trnL (UAA) gene
UP	Upper Peninsula

Abstract

Northern white-tailed deer (*Odocoileus virginianus*) mitigate the increased energetic costs of severe winter conditions through obligate migration to densely congregated areas with abundant conifer cover, a behavior referred to as yarding. “Deer yards” in the Western Upper Peninsula (UP) of Michigan are principally located within the reduced snowpack and increased ambient temperature microclimates of densely canopied eastern hemlock (*Tsuga canadensis*) stands, but specific drivers of long-term site fidelity and utilization are largely uncharacterized. As an important game species of high economic and cultural value and a keystone herbivore with critical impact of plant community composition and structure, identifying winter yarding site selection requirements across the landscape is necessary for effective management strategies. The first chapter of this MS thesis research leverages 18 years of deer fecal pellet-group counts along with tree inventory measurements from 39 relict eastern hemlock stands across the Western UP to investigate the spatiotemporal predictive power of structural overstory traits for use estimation of individual deer yards. We found that high deer use is associated with overstory traits of larger hemlock including increased crown width, basal area, and height, and that deer yarding complexes with larger hemlock are especially critical during severe winter events. The second chapter uses feces collected from the same 39 winter yarding sites from 2006-2022 to assess impacts of methodological choices for fecal DNA metabarcoding and fecal DNA host analyses. We compared three commercial DNA extraction kits targeting different sample types with deer fecal samples stored for varying lengths of time (0-16 years) and utilized fecal metabarcoding to target both plant and microbial species within the fecal extract. We assessed metabarcoding results from 3 universal plant primer pairs targeting diverse genomic regions (trnL, rbcL, and ITS2) and a single 16S rRNA v3-v4 microbial amplicon, and quantified host DNA quality with deer mitochondrial Sanger Sequencing. Target locus selection was the most significant factor for winter diet item detection and taxonomic resolution, with rbcL exhibiting the best overall performance. DNA extraction kit selection was most significant for host DNA sequencing, and sample storage time had little impact aside from slight differences in microbial community composition. Overall, this research has important implications for white-tailed deer and eastern hemlock management across the Upper Midwest and provides valuable recommendations for methodological development of fecal metabarcoding for wildlife ecology.

1 Introduction

White-tailed deer (*Odocoileus virginianus*) are broadly distributed large mammalian herbivores in North America with immense value as a wildlife species. The economic and cultural value related to the highly profitable game industry layers upon the ecological importance of deer as ecosystem engineers, wherein the selective generalists have disproportionate impact on ecosystems by altering plant community structure and composition.

Winter yarding in northern deer populations, an adaptive behavior to mitigate increased energetic costs of winter conditions by densely congregating in areas of high conifer cover, represents an important spatiotemporal movement pattern important for management. In the western Upper Peninsula (UP) of Michigan, the majority of deer yards are located within remnant patches of eastern hemlock (*Tsuga canadensis*), a declining foundational forest species in the Midwest region whose dense canopy intercepts snowfall and traps thermal energy to provide high-quality winter cover for deer.

In this Thesis research, we investigate different aspects of UP white-tailed deer yarding, including habitat suitability and long-term use, along with methods of estimating diet, host population structure, and microbial communities using noninvasive environmental DNA samples. In Chapter 1 we leverage an 18-year spatiotemporal dataset of deer yarding to assess patch-level overstory structure of eastern hemlock as a predictor of long-term deer use intensity, frequency, and variability. In Chapter 2 we assess key methodological considerations of noninvasive molecular techniques such as fecal metabarcoding and host DNA analyses, and the impact the choices have on our ability to answer ecological questions related to herbivore diet, gut microbial communities, and host population dynamics.

Overall, the findings of this thesis research are broadly applicable to both wildlife and forestry management and the importance of active management for larger mature hemlock stands across the Upper Midwest for high quality white-tailed deer winter habitat. Additionally, we demonstrate the importance of methodological design for molecular ecology research and different factors impacting our ability to answer ecological questions from noninvasive wildlife fecal samples.

2 Influence of Stand Structure on Long-Term Use of Relict Eastern Hemlock Stands by Overwintering White-Tailed Deer

2.1 Introduction

Winter yarding is a critical behavioral adaptation of northern white-tailed deer (*Odocoileus virginianus*) populations for survival in the harsh winter conditions of the Upper Peninsula (UP) of Michigan (Doepker et al., 2017.) With annual snowfall surpassing 300 inches annually in the most severe areas (Brugam et al., 2004; Eichenlaub, 1970), the onset of dropping temperatures and snowfall initiate fall migration to traditional overwintering yarding sites (Nelson, 1995) characterized by abundant coniferous cover (Ozoga & Gysel, 1972). UP deeryards are predominantly associated with dense northern white-cedar (*Thuja occidentalis*) swamps in the eastern portion of the region and eastern hemlock (*Tsuga canadensis*) in the western-half, the latter also often on wind-protected slopes and river bottoms that increase ambient temperature (Suggitt et al., 2011; Verme, 1973). In addition to trapping warmth, reduced within-stand snowpack due to snowfall interception by dense tree canopies facilitates critical energy conservation for deer during winter periods where browse is less available but movement is energetically costlier (LaGory et al., 1985; Moen, 1976). Predator safety in high-density congregations at deeryards increases in association with social organization (Nelson & Mech, 1981) more abundant runaway tracks (Messier & Barrette, 1985), and there is evidence of sociobiological matrilineal genetic structure related to winter yarding site fidelity (Chepko-Sade & Halpin, 1987; Cronin et al., 1991).

Despite the well-established importance of eastern hemlock stands for deer winter survivorship, the foundational forest species has seen vast declines in its historic range across the United States due factors such as logging, invasive species, and shifting climate patterns (Ellison et al., 2018; Hart, 2008). Excessive seed source removal and slash fires have eliminated large-diameter decomposing woody debris from conifer forest floors that serve as important microsites for hemlock regeneration (Mladenoff & Stearns, 1993). Although geographic distance and colder temperatures have historically buffered the Upper Midwest from the infestations of hemlock wooly adelgid (HWA) that are decimating eastern populations (Orwig et al., 2002), regional climatic shifts towards warmer and drier conditions are accelerating northward expansion of this emergent threat (Hart, 2008; Paradis et al., 2008). Additionally, these climate trends are altering forest composition and tree recruitment patterns to the detriment of remnant hemlock patches, which now occupy only about 0.5% of the northern Great Lakes landscape (Eckstein, 1980, Foster et al., 2006). As a result, the region is experiencing shifts towards deciduous species in stands historically dominated by the extremely shade-tolerant and slow-growing hemlock, particularly with sugar maple (*Acer saccharum*) (Bradshaw & Waller, 2016).

Given the continued reduction of hemlock cover across the landscape, the concentrated use of remnant patches for winter yarding combined with increased deer populations levels in a system historically limited by severe winter conditions is an additional driver of hemlock failed regeneration at the site-level (Michigan Department of Natural Resources, Mladenoff & Stearns, 1993). However, despite reduced availability, the remaining patches of high conifer cover are not uniformly utilized by northern yarding deer (Eichenlaub, 1970; Witt et al., 2012). While environmental and landscape features such as snow depth and topography (Eichenlaub, 1970; Witt et al., 2012) along with availability of hardwood browse (Morrison et al., 2003) and increased ground-layer community heterogeneity (Jensen et al., 2011) have been shown to influence winter habitat selection at varying degrees of winter severity (DelGiudice et al., 2013; Morrison et al., 2002), the specific importance of hemlock overstory structure for deer use is not well-evidenced. Short-term assessments of deer yarding over the course of two to three seasons have found that overnight wintering bed-sites are strongly associated with hemlock cover (Armstrong et al., 1983) and spatial distribution of softwood shelter is the best predictive factor for deer use variation during moderate to severe winters (Morrison et al., 2003), but long-term monitoring data is needed to evaluate the significance of hemlock cover for deer yarding amidst complex interactions with highly-variable environmental conditions and fluctuating population levels.

To help bridge this gap and provide useful information for management of both eastern hemlock and white-tailed deer populations in the Upper Midwest, our primary objective was to leverage 18-years of winter pellet survey data in 39 western Upper Peninsula deer yards to investigate stand-level hemlock overstory structure as a predictor of long-term deer use. We hypothesized that yarding sites with hemlock overstory metrics associated with increased cover and larger trees would have higher deer use intensity and consistency over time, and that complexes with optimal shelter would be particularly important during highly severe winter seasons.

2.2 Methods

2.2.1 Field Site Description

Our study system is comprised of 39 relict eastern hemlock stands across the western Upper Peninsula of Michigan (Figure 1.1) selected as potential winter yarding sites for white-tailed deer based on patch size and relative snow depth data from the US Forest Service and Michigan Department of Natural Resources (Witt & Webster, 2010). Patch sizes of the selected stands range from 0.4 to 60 ha and were located across four different land ownerships: Michigan Technological University, Gratiot Lake Conservatory, US Forest Service, and Michigan Department of Natural Resources. Within each stand 3-9 vegetation plots (400 m^2) and 3-30 pellet-group plots (9.27 m^2) were randomly distributed using ArcView 3.3 Random Point Generator based on patch size calculated using US Geological Survey Digital Orthophoto Quadrangles (Witt & Webster, 2010).

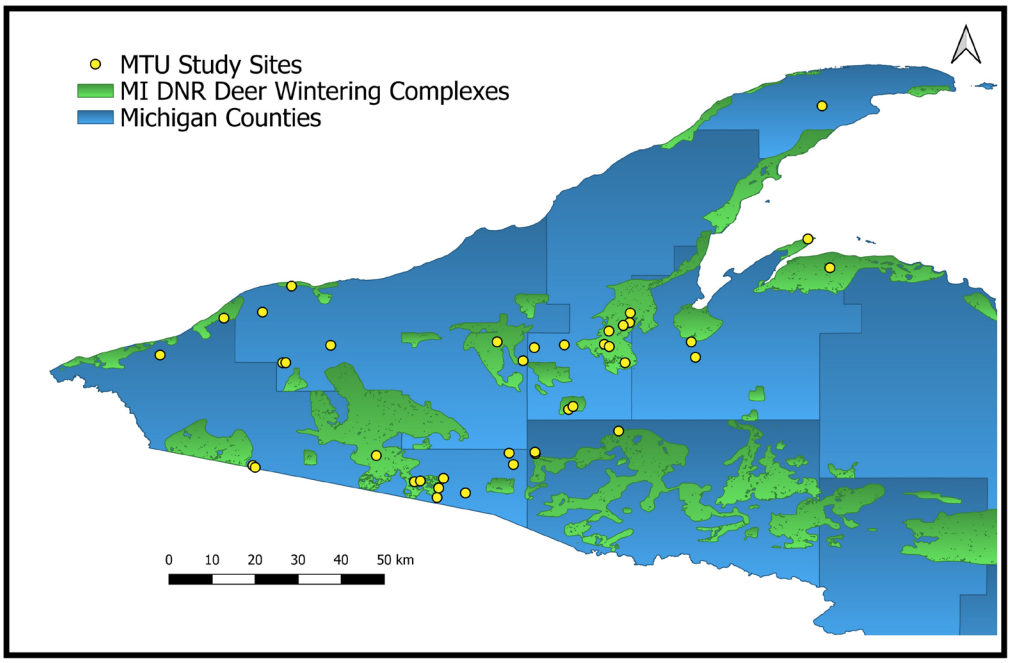


Figure 2.1. Map of 39 study sites and Michigan DNR defined overwintering complexes of white-tailed deer. The study sites are located within relict eastern hemlock stands in the western Upper Peninsula of Michigan and were monitored annually for deer presence from 2006-2023 using pellet counts. The GIS layer of historic overwintering deer complexes is sourced from open access Michigan Department of Natural Resources data layers.

2.2.2 Tree Measurements and Deer Fecal Pellet Sampling

To quantify overstory composition and structure, species, diameter at breast height (dbh) (cm), height (m), mid-canopy (m), and height at base of crown (m) was recorded for all overstory and midstory trees (dbh > 4.0 cm) within each vegetation plot (400 m^2) across the 39 sites during the summer of 2006 (n=8653) (Witt & Webster, 2010). Utilization by white-tailed deer of the hemlock stands as winter yarding sites was measured with annual spring pellet-group counts within the 3-30 smaller circular sampling plots (9.29 m^2). Pellet plots were cleared of feces prior to winter during site establishment in 2005 and then rechecked for pellet-groups immediately following spring snowmelt and cleared for the subsequent year ("reset") (Witt et al., 2012; Witt & Webster, 2010). Field crews employed a concurrent recheck system for accurate counts, and as suggested in previous studies (Forsyth et al., 2007), a pellet-group was only counted if it included ≥10 intact pellet in single defecation. Pellet surveys were conducted from 2006-2023 resulting in a total of 3196 pellet-groups counted across 39 sites over the 18-year period.

2.2.3 Calculating Tree Variables and Deer Usage

The 2006 stand inventory measurements were used to calculate the following individual tree metrics: basal area (BA), tree per hectare (TPH), crown length (CL), crown width (CW), crown ratio (CR), form index (FI), crown thickness (CT), and crown spread ratio (CSR) (Table 1.1). Crown characteristics are difficult to measure in the field as well as accurately model due to complex species-specific interactions with stand density (Hu et al., 2023; Miraki et al., 2021), so we employed a local basal area adjustment with individual species equations (Table B.1) to calculate crown width (Bragg, 2001). These metrics are useful for modeling stand attributes such as canopy cover and tree density previously suggested to be important for ungulate winter habitat selection (McIntosh et al., 2012; Telfer, 1970). Individual tree measures were summed (TPH only) or averaged to determine stand-level values. We calculated stand-level overstory structural values for eastern hemlock in addition to the percentage of basal area per stand represented by the following three groupings: northern white-cedar (*Thuja canadensis*), all conifers, and all deciduous species.

Table 2.1. Individual tree metrics and equations used for determining stand structure. Two letter codes for each variable are listed along with final units. Variables designated with '*' code were modified from species-specific equations from Bragg et al. 2001, and those designated with '**' were calculated from the estimated variable.

Tree Variable	Code	Equation	Final Unit
Basal Area/Tree	BA	$= \frac{(\text{dbh}_{\text{cm}})^2 \text{ x } 0.00007854}{400} * (1000)$	m^2/ha
Tree Per Hectare	TPH	$= \frac{1 \text{ tree}}{400\text{m}^2} \text{ x } \frac{10{,}000\text{m}^2}{1 \text{ ha}} = 25$	tree/ha
Crown Length	CL	$= \text{Tree Height (m)} - \text{Low Canopy (m)}$	m
Crown Ratio	CR	$= \frac{\text{Crown Length (m)}}{\text{Tree Height (m)}}$	-
Crown Width*	CW	Calculated with species-specific formulas	m
Form Index**	FI	$= \frac{\text{Crown Length (m)}}{\text{Crown Width (m)}}$	-
Crown Thickness**	CT	$= \frac{\text{Crown Width (m)}}{\text{Crown Length (m)}}$	-
Crown Spread Ratio**	CSR	$= \frac{\text{Crown Width (m)}}{\text{Tree Height (m)}}$	-

*Crown Width (CW) was calculated using species-specific formulas adapted from Bragg et al. 2001
**Variables were calculated using the estimated crown width measurement

Three measures were used to quantify winter deer use across the study period. As a measure of intensity, we calculated Mean Annual Use (MA), or the average total pellet-groups per hectare per year at each site. As a measure of frequency, we calculated Stand Vacancy Rate (VR), or the percentage of years with no pellet-groups found. Finally, as a measure of variability, we calculated Use Variability (UV), of the standard error of annual use at each site. Variation in stand area and number of plots across the stands was minimized by using pellet-groups per hectare. Qualitative classes for each usage measure were defined using sample means and standard deviations of normalized distributions.

2.2.4 Environmental Data

Winter severity measures of localized snow depth and ambient temperature are important predictors of winter deer use of yarding sites (Morrison et al., 2003; Parikh & Webster, 2019), so we calculated the Winter Severity Index (WSI) used by the Wisconsin Department of Natural Resources for white-tailed deer management (Verme, 1968). WSI is the number of days with a snowpack depth exceeding 46cm summed with the number of days with a minimum temperature below -18 °C (Parikh & Webster, 2019; Witt et al., 2012). Snow depth data were obtained from the Snow Data Assimilation System (SNODAS) contiguous United States model (National Operational Hydrologic Remote Sensing Center, 2004) and temperature data were sourced from local weather stations (Michigan Technological University Keweenaw Research Center, 2023; National Buoy Data Center 2023). We generated daily snow depth estimates at each stand by georeferencing the daily SNODAS interpolated depth model in RStudio (2023.09.1) with the R-package SNODASR (R version 4.3.1) and averaging QGIS sampled raster values within sites. We compared annual averages of WSI and mean annual deer use (pellets ha-1) to qualitatively assess temporal trends of winter yarding site use with climate variation. Additionally, we evaluated differences in deer use intensity for years with decreased ($WSI_{Diff} < -20$), similar ($-20 < WSI_{Diff} < 20$), increased ($WSI_{Diff} < 20$) winter severity compared to the previous season using chi-squared goodness of fit tests.

2.2.5 Nonmetric Multidimensional Scaling (NMDS)

To evaluate predictive power of stand overstory structure for winter deer use, we employed non-metric multidimensional scaling (NMDS) to characterize variation in key overstory variables across our 39 hemlock stands. The species-matrix included stand-level variables averaged for all species (TPH, BA), eastern hemlock (dbh, BA, CL, CR, FI, CT, CSR, CW, height), northern white-cedar (BA, TPH), and all conifers (BA). All structure metrics were relativized by maximum (Legendre & Gallagher, 2001) but required no further transformation. Ordination with Sorensen-Bray distances and standard parameters was performed in PCord (Wild Blueberry Media LLC, Corvallis, OR). A secondary matrix with continuous and qualitative deer-use measures was correlated with

the final NMDS ordinations using PCord and the 'vegan' R-package envfit function (permutation=999). Multiple Response Permutation Procedures (MRPPs) for significantly correlated deer-use measures (R^2>0.1) were run in PCord to generate pairwise group comparisons (Cai, 2006). Indicator Species Analysis (ISA) was also conducted to determine if particular stand structural variables were significantly correlated to specific deer use categories (Bakker, 2008; Dufrêne & Legendre, 1997).

2.3 Results

2.3.1 Deer Usage Measures

We detected notable spatiotemporal deer use variability across the 39 winter yarding sites over the 18-year study period. Total study mean annual use (MA) was 407 ± 82 pellet-groups ha-1 with stand average values ranging from 0 to 4042.6 (Table 2.2). Stand vacancy rate (VR) ranged from 0% (5 sites) to 100% vacancy (4 sites) with an average of 37.3% vacancy, or 7 out of 18 years with no detected winter deer use. We observed high interannual within-stand use variability (UV), with an average standard error of 91.34 ± 20.63 pellets-groups ha-1 and a range of 5.98 to 498.26 excluding the four hemlock stands that had no pellet groups detected over the entire study period. For categorical assessments we defined qualitative usage classes for each of the three deer use metrics based on normalized data distributions. The mean annual usage (MA) classes (units= pellet-groups ha-1) were never (0), below-average(1-169), above-average (170-310), and high (>311). Stand vacancy (SV) classes were always vacant (100%), frequently vacant (50-99%), infrequently vacant (1-49%), and never vacant (0%). Use variability (UV) classes (units= pellet-groups ha-1) were no-use (0), low variability (0.1-33.1), average variability (33.2-88.7), and high variability (>87.2). For simplified qualitative comparisons of vacancy rates, we also assigned each stand as either vacant (VR>50%) or occupied (VR <50%).

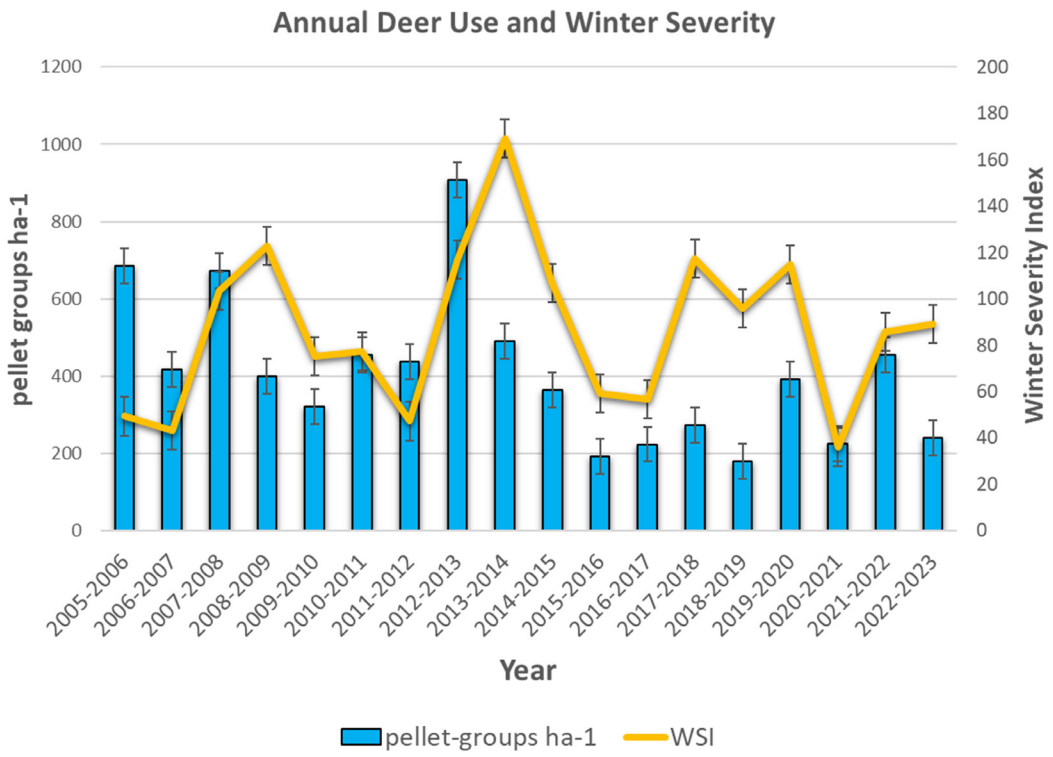


Figure 2.2. Average annual deer use and winter severity index from 2006-2023. Deer use was measured as the pellet-groups ha-1 across all 39 study sites across the western Upper Peninsula for each year. Winter severity index (WSI) is an index used by the Wisconsin Department of Natural Resources summing days with snow depth > 46cm and a minimum temperature < 18 °C from December 1st – April 30th of each winter season. Snow depth data were obtained from the SNODAS database. Daily minimum temperatures were sourced from local weather stations.

Table 2.2. Stand attribute summary for select eastern hemlock traits, snow depth, and deer usage measures for 39 study sites. Stands are ordered by descending total stand area (ha) and study averages with standard error are included for each measure.

Stand Area	Basal Area (BA)	Trees ha -1	% *Tsuga* BA	dbh *Tsuga*	Height *Tsuga*	CW *Tsuga*	Days Snow Depth >46cm	Mean Annual Deer Use		
(ha)	(m^2/ha)			(cm)	(m)	(m)		Pellets ha-1	Min	Max
59.6	37.6	658.3	55.5	35.4	19.5	6.6	59.1	324.9	35.9	1112.3
46.7	40.7	730.6	55.1	32.4	18.0	6.2	55.8	1146.2	322.9	2296.4
38.8	47.8	691.7	48.5	33.5	18.3	5.9	48.3	169.4	0.0	358.8
34.4	46.7	1588.9	45.5	24.8	13.5	4.8	54.3	418.6	0.0	2583.4
29.4	47.8	613.9	63.5	41.3	21.4	6.7	72.8	4042.6	1722.3	8611.4
29.1	35.3	813.9	37.9	32.8	16.3	6.5	80.2	3115.7	861.1	6781.5
29.1	74.9	390.0	77.0	43.1	21.5	5.7	50.3	484.4	0.0	1829.9
23.9	57.2	475.0	50.2	37.1	17.0	5.8	50.0	233.2	0.0	753.5
21.8	37.1	730.6	38.7	39.1	19.8	6.8	75.4	753.5	0.0	8288.5
21.6	36.9	1035.7	31.9	29.7	15.0	5.9	75.8	336.9	35.9	968.8
21.2	42.3	1244.4	44.5	20.9	12.2	4.4	42.1	296.0	0.0	1614.6
19.2	47.1	1125.0	28.9	24.2	11.9	4.8	79.9	287.0	0.0	1076.4
19.1	38.2	1253.6	43.9	21.5	13.8	4.8	70.7	123.6	0.0	430.6
18.3	53.8	503.6	47.9	39.0	19.0	6.2	51.3	165.5	0.0	753.5
15.7	49.9	607.1	62.6	30.6	15.3	5.3	49.2	131.6	0.0	484.4
14.8	34.9	790.0	55.4	38.2	20.3	7.0	61.8	233.2	0.0	968.8
14.7	51.0	1239.3	38.6	29.7	16.7	5.4	34.2	68.8	0.0	215.3
14.6	33.7	1357.1	63.9	38.1	17.5	6.9	50.7	140.5	0.0	484.4
14.1	35.8	1496.4	45.3	15.8	10.0	4.0	55.9	221.6	0.0	633.2
12.2	52.2	660.0	45.9	42.4	19.3	6.7	60.7	158.5	0.0	538.2
10.5	41.6	650.0	36.5	55.1	19.8	8.4	76.3	263.1	0.0	1507.0
7.7	42.7	1305.0	50.1	17.2	11.7	4.1	47.3	41.9	0.0	107.6
7.4	49.0	1190.0	47.9	34.4	17.6	5.9	27.1	412.6	0.0	968.8
7.1	51.0	1260.0	65.1	25.1	13.9	4.7	81.8	6.0	0.0	107.6
6.5	37.8	1660.0	35.1	15.6	10.0	4.0	46.2	263.1	0.0	1076.4
4.7	60.9	500.0	80.8	38.5	22.1	6.0	62.3	71.8	0.0	287.0
3.5	46.8	1330.0	40.8	17.7	11.4	4.0	55.4	1016.6	0.0	2583.4
3.1	38.9	1040.0	28.1	49.0	20.9	7.9	46.9	191.4	0.0	502.3
3.0	42.4	891.7	80.9	43.1	18.2	7.2	42.9	51.8	0.0	215.3
2.4	77.2	785.0	71.7	29.9	15.3	4.3	50.1	191.4	0.0	753.5
1.6	72.8	625.0	46.8	32.9	15.1	4.8	63.4	0.0	0.0	0.0
1.4	43.5	2033.3	19.2	18.7	11.1	4.2	43.1	107.6	0.0	645.9
0.9	59.7	583.3	69.1	37.1	19.0	5.8	45.3	12.0	0.0	107.6
0.7	36.7	1450.0	31.4	13.0	8.0	3.6	83.0	0.0	0.0	0.0
0.5	60.5	900.0	62.4	34.7	18.2	5.7	77.0	155.5	0.0	861.1
0.4	46.3	1608.3	25.0	16.3	9.5	3.8	68.1	0.0	0.0	0.0
0.4	63.7	1100.0	53.5	32.0	19.8	5.2	59.3	137.5	0.0	322.9
0.2	55.2	716.7	46.1	49.0	20.0	7.2	53.0	95.7	0.0	645.9
0.0	49.5	1350.0	25.5	31.8	15.5	5.8	76.8	0.0	0.0	0.0
Study Averages										
	41.1	999.6	48.6	31.8	16.3	5.62	58	407	76.4	1319
	± 0.04	±160	±7.8	±1.6	±0.6	±0.19	±7.4	±82	±49	±328

2.3.2 Temporal Environmental Trends and Deer Usage

Our comparison of average annual WSI and mean annual deer usage (pellet-groups ha-1) across the study period demonstrated that use intensity increases during high-severity winters with dramatic declines in proceeding years (Figure 2.2). This pattern was observed in all three severe winter cycles (WSI >100) documented throughout our study period, namely the 2007-2009, 2012-2015, and 2017-2019 seasons, with a maximum WSI of 169 in the 2013-2014 winter (WSI_{avg} = 87) . While the average difference in pellet-groups ha-1 between subsequent seasons was -25 ± 55.2, winters with WSI decreases ≥ 20 saw only 51% of stands with increased use (mean difference = -109.39 ± 24.7) compared to 64.5% in those with WSI increases ≥ 20 (mean difference = 77.4 ± 116.9). Chi-squared goodness of fit testing produced non-significant results (p=0.057) for differences in deer usage based on interannual WSI trends.

2.3.3 Overstory Structure Variation and NMDS

Composition and structural overstory traits of eastern hemlock varied at the stand level (Table 1.2). Average dbh (31.8 ± 1.6 cm) and height (16.3 ± 0.6 m) exhibited wide ranges from 15.6 - 49 cm and 8 - 22.1 m respectively, and 27 stands (69.2%) contained at least one mature hemlock of dbh > 61 cm (USDA Forest Service) within the sampled vegetation plots. Crown width, an important component for cover in winter yarding sites, averaged 5.62 ± 0.19 m, ranging from 3.6 to 7 m.

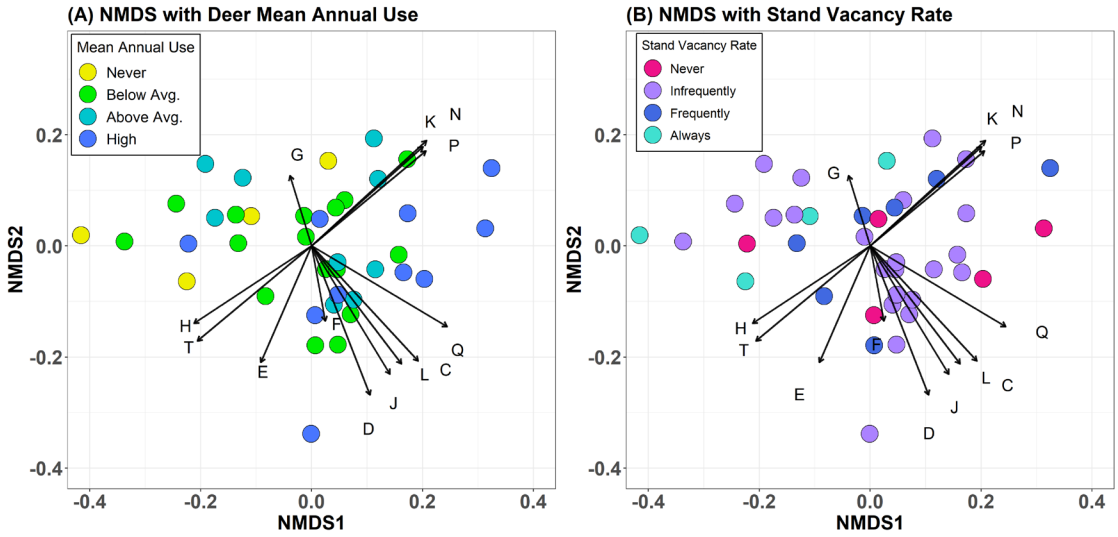


Figure 2.3. NMDS ordinations of 39 relict eastern hemlock stands using stand structure variables coded by (A) mean annual deer use (pellets ha-1) and (B) stand vacancy rate (% years unoccupied). Fitted vector codes are diameter at breast height (dbh) hemlock (C), crown length hemlock (D), crown ratio hemlock (E), form index hemlock (F), crown thickness hemlock (G), crown spread ratio hemlock (H), crown width hemlock (J), total basal area hemlock (K), basal area hemlock (L), basal area conifers (M), basal area per hectare conifers (N), % basal area conifers (P), height hemlock (Q) , total stand tree per hectare (R), total stand basal area (S).

NMDS Bray-Curtis ordination with stand structural traits and winter deer use measures as a secondary matrix evidenced significant hemlock overstory differences between the lowest and highest mean annual usage (MA) classes (Figure 2.3 and Figure 2.4). All input traits were significantly correlated (R^2 >0.5, p <0.001) with the final solution (k=3, stress=9.3) except for basal area and density of northern white-cedar (p=0.849). Form index (FI) and crown thickness (CT) were less strongly correlated than the other traits. Although the ordination showed no clear stand hemlock structure differences for vacancy rate (R^2 = 0.118, p =0.17), mean annual use had significant goodness of fit (R^2 =0.19, p=0.022) with low vs. high use sites separating along the first axis. Ellipse centroids for the "High" use intensity class correlate with fitted vectors for hemlock crown length, crown width, basal area, and height (Figure 2.3B, vectors D, J, L, and Q), evidenced by the first axis separation of sites with and without mature hemlock classes (Figure 2.4D). The use variability (UV) measure was also significant in the ordination (R^2 =0.17, p=0.045), with higher variability associated with larger average hemlock traits (Figure 2.4C). Multi-response permutation procedures for all deer measures were insignificant (p>0.05), but indicator species analysis identified candidate predictors of specific usage levels. Hemlock crown spread ratio and deciduous basal area were significant predictors of MA usage level of "never" used (p=0.04, p=0.025), with deciduous basal area also a predictor of sites that are "always" vacant (p=0.30). Hemlock crown width was a significant indictor of the "never" vacant use frequency class (p=0.05), along with crown length (p=0.022) and hemlock height (p=0.016) (Table 2.3). Additionally, higher coniferous basal area was a significant indicator of sites with low variability in annual use (p=0.028).

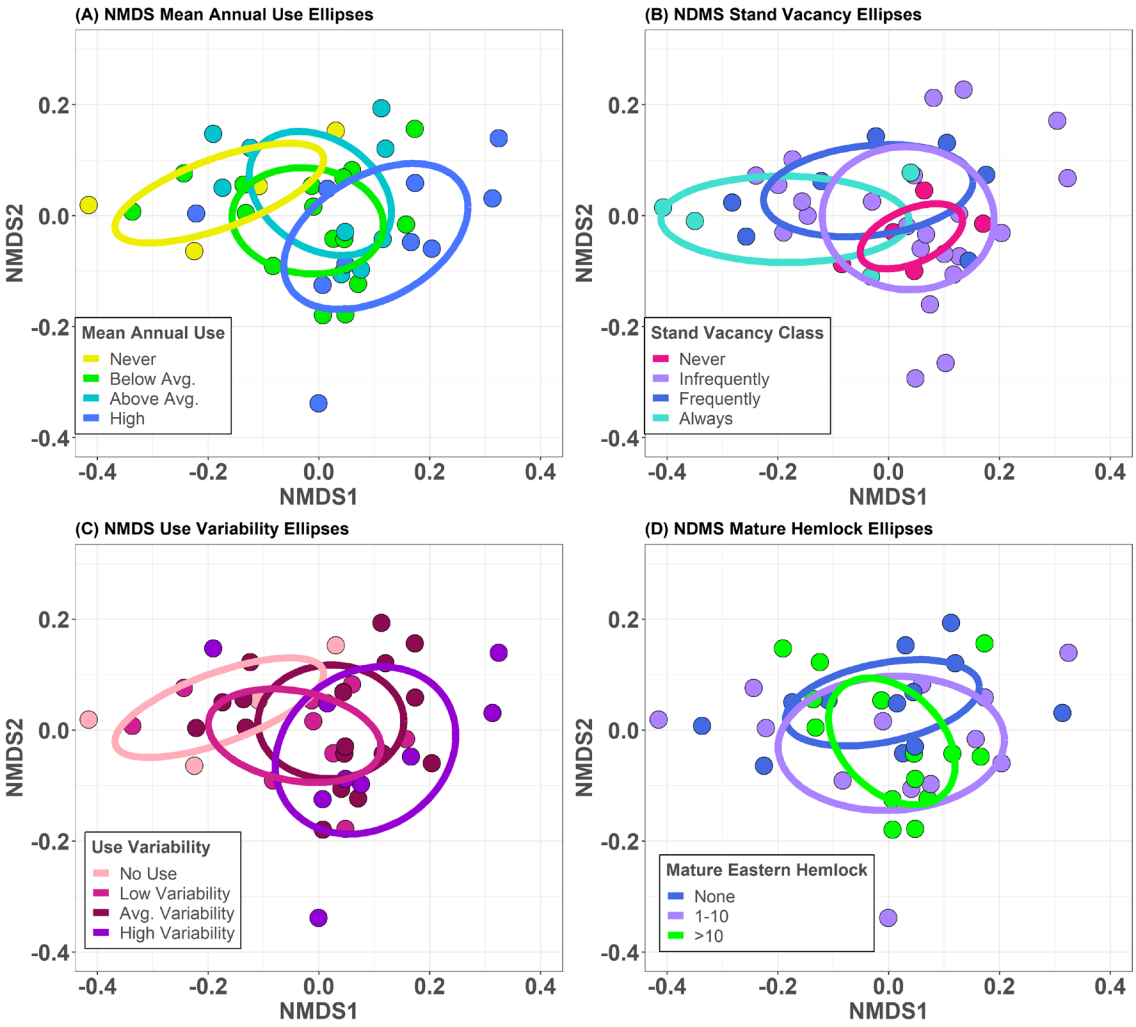


Figure 2.4. NMDS ordinations of 39 relict eastern hemlock stands with ellipse centroids of deer use measures. The ellipses are coded by class for (A) mean annual deer usage levels and (B) stand vacancy rates, (C) standard error in mean annual use, and (D) number of mature hemlock trees samples.

Table 2.3. Indicator species analysis correlating stand structural variables with levels of deer use and occupancy in 39 winter yarding sites. Indicators were only included if p <0.10, and Group denotes class with the Deer Use Estimate.

Deer Use Estimate	Structural Variable	Group	p-value
Mean Annual Use	Hemlock Crown Spread Ratio	Never	0.040*
	Deciduous Basal Area	Never	0.025*
Vacancy Rate	Hemlock Crown Spread Ratio	Always	0.083
	Hemlock Height	Never	0.076
	Hemlock Crown Width	Never	0.050*
	Deciduous Basal Area	Always	0.030*
Simplified Vacancy Rate	Hemlock dbh	>50% Occupied	0.066
	Hemlock Basal Area	>50% Occupied	0.061
	Hemlock Crown Length	>50% Occupied	0.022*
	Hemlock Crown Width	>50% Occupied	0.021*
	Stand Tree Per Hectare	>50% Vacant	0.021*
	Hemlock Height	>50% Occupied	0.016*
Use Variability	Coniferous Basal Area	Low Variability	0.028*

*Denotes significant p-values (<0.05)

2.4 Discussion

In support of previous findings within our same system and beyond that winter severity alters deer yarding behavior (Morrison et al., 2003; Ozoga & Gysel, 1972; Parikh & Webster, 2019), we observed a consistent trend over the 18-year study period of deer use intensity increasing during severe winter seasons coupled with dramatic declines in the proceeding years. These usage spikes occurring in high cover areas are consistent with energy conservation and a constricted energetic budget during seasons where snow depth and cold temperatures dramatically limit movement (Dumont et al., 2005; Schmitz, 1991). The sharp decreases in deer abundance following heavy winters has been linked to increases in overwintering mortality with decreased deer harvest levels due to starvation and prolonged exposure (Michigan Department of Natural Resources, 2016; Parikh & Webster, 2019). Higher mortality is expected given the heavier winter conditions characteristic of the Upper Midwest, although deer population is still recovering from a notably severe winter over the 2013-2014 season, colloquially termed the “polar-vortex”. Within our study system we observed four newly abandoned stands following the “vortex”, indicating the importance of high-quality winter habitat for deer survivorship. The sudden decline in usage at these sites may be a product of predation, a risk positively

correlated with winter severity (Norton et al., 2021), but more likely it reflects inadequate coniferous cover with the yarding complex (DeLgiudice et al., 2002).

Importantly, our results provide evidence that the overstory structure of eastern hemlock is predictive of long-term deer use in Upper Midwest overwintering complexes. We found significant correlations between traits associated with larger hemlock trees including crown width, tree height, and dbh. This finding is harmonious with snow depth as a primary driver of yarding behavior, as larger hemlock promote denser canopies capable of intercepting up to 60% of annual snowfall (Ozoga & Gysel, 1972; Varhola et al., 2010). The data suggests that stands with large, older hemlock trees with spreading crowns are the most valuable yarding locations for deer. The association of lower mean annual deer usage at sites with higher percent deciduous basal area highlights the concern of well-documented declines of eastern hemlock in the Upper Peninsula of Michigan negatively impacting deer survivorship (Davis et al., 1995; Frelich & Lorimer, 1985; Hix & Barnes, 1984). Although northern deer population levels are inherently limited by winter-severity compared to southern US regions (Michigan Department of Natural Resources), accelerated successional shifts to hardwood species in historically hemlock dominant sites may increase annual mortality and destabilize population growth cycles (Bradshaw & Waller, 2016). While winters in the Upper Midwest are trending warmer with decreased late season snow-cover (April-May), future climate projections for the region include increased total winter precipitation levels in conjunction with "lake-effect" snow (Demaria et al., 2016; Wuebbles & Hayhoe, 2004), suggesting that high-quality deeryards may be important buffers for sporadic extreme intra-seasonal weather events (Morrison et al., 2003; Ozoga & Gysel, 1972). However, concurrent degradation of habitat and environmental conditions best-suited for eastern hemlock regeneration including cooler soil temperatures, high moisture levels, and low leaf litter induced by climate shifts may necessitate active forest management to guarantee successful seedling recruitment in important yarding complexes.

However, we found evidence supporting the spatial arrangement of mature hemlock patches across the landscape within overwintering complexes in combination with proximity to available browse as key predictors of temporal variation in yarding site use across all variations environmental conditions (Morrison et al., 2003). Stands with higher conifer basal area and density demonstrated the lowest variability of year-to-year usage throughout our study system. While tradeoffs between browse availability and quality cover are strongly correlated with use intensity, higher spatial connectedness can dilute concentrated use and increase access to broader areas for browsing opportunities with lowered energetic costs (Van Deelen et al., 1998; Witt et al., 2012). Unfortunately, given the long legacy of land use in the region and the resultant patchiness of hemlock on forest landscape-scales, connected corridors indicative of high-quality overwintering deer complexes are not highly the norm and restoration would also require active and integrated forest management practices (Morrison et al., 2003).

In addition to overstory structure, composition, and spatial distribution of hemlock, there are a number of additional features potentially driving deer yarding behavior. Supplemental feeding of deer, a difficult to enforce practice regulated by state agencies

due to increased disease risk (Brown & Cooper, 2006), alters population distribution across the landscape by concentrating deer to sites proximal to feeders (Osborn & Jenks, 1998) such that potential winter yarding sites closer to supplemental food resources are likely to have increased temporal utilization. Vicinity to anthropogenically altered areas alters seasonal ranges of white-tailed deer, with previous research demonstrating 78% of female deer range during severe winters was associated with residential neighborhoods (Grund et al., 2002). These areas often feature snow removal, increased temperatures, and landscaping, providing less energetically-taxing browse. Additionally, the dynamics of matrilineally-linked selection of yarding location is little understood. Social structure and kinship are suggested to affect overwintering site selection and fidelity in northern deer populations (Ozoga & Verme, 1984; Van Deelen et al., 1998), but no research on how this social tradition varies with environmental factors, anthropogenically altered landscapes, and shifting browse composition has been conducted. A deeper understanding of the full breadth of factors influencing white-tailed deer yarding is critical for developing localized and effective management practices.

In conclusion, the findings here highlight the critical role of larger eastern hemlock overstory structures for winter yarding site quality and long-term deer use. The increased capacity of older hemlock stands to intercept a substantial portion of annual snowfall provides heightened protection and energetic benefits for deer and is essential for deer survival. The potential of larger-hemlock dominated stands to also serve as a buffer against increased variability in winter weather patterns may grow in significance as if climate patterns continue in the Upper Midwest region. The relationship between overstory structure and deer yarding demonstrated here underscores the importance of protecting and managing mature hemlock patches in northern forested landscapes. This is especially important given the incredible length of time required to establish mature hemlock on the landscape. Furthermore, a comprehensive approach to wildlife and forestry management that synthesizes the ecological value of mature hemlock stands for flora and fauna species adapted to its microclimates is essential for protecting persistence of biodiversity across the landscape. As forest composition and environmental conditions continue to shift, active management practices that promote eastern hemlock regeneration will be crucial for sustained deer populations across the Upper Peninsula of Michigan.

2.5 References

Armstrong, E., Euler, D., & Racey, G. (1983). Winter Bed-Site Selection by White-Tailed Deer in Central Ontario. *The Journal of Wildlife Management*, *47*(3), 880–884. https://doi.org/10.2307/3808632

Bakker, J. D. (2008). Increasing the utility of Indicator Species Analysis. *Journal of Applied Ecology*, *45*(6), 1829–1835. https://doi.org/10.1111/j.1365-2664.2008.01571.x

Bradshaw, L., & Waller, D. M. (2016). Impacts of white-tailed deer on regional patterns of forest tree recruitment. *Forest Ecology and Management*, *375*, 1–11. https://doi.org/10.1016/j.foreco.2016.05.019

Bragg, D. C. (2001). A Local Basal Area Adjustment for Crown Width Prediction. *Northern Journal of Applied Forestry*, *18*(1), 22–28. https://doi.org/10.1093/njaf/18.1.22

Brown, R. D., & Cooper, S. M. (2006). In My Opinion: The Nutritional, Ecological, and Ethical Arguments Against Baiting and Feeding White-Tailed Deer. *Wildlife Society Bulletin*, *34*(2), 519–524. https://doi.org/10.2193/0091-7648(2006)34[519:TNEAEA]2.0.CO;2

Brugam, R. B., Owen, B., & Kolesa, L. (2004). Continental-scale climate forcing factors and environmental change at Glimmerglass Lake in the upper Peninsula of Michigan. *The Holocene*, *14*(6), 807–817. https://doi.org/10.1191/0959683604hl761rp

Chepko-Sade, B. D., & Halpin, Z. T. (1987). *Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics*. University of Chicago Press.

Cronin, M. A., Nelson, M. E., & Pac, D. F. (1991). Spatial Heterogeneity of Mitochondrial DNA and Allozymes among Populations of White-Tailed Deer and Mule Deer. *Journal of Heredity*, *82*(2), 118–127. https://doi.org/10.1093/oxfordjournals.jhered.a111045

Davis, M. B., Sugita, S., Calcote, R. R., Parshall, T. E., & Ferrari, J. B. (1995). *3000 years of abundant hemlock in upper Michigan*. http://conservancy.umn.edu/handle/11299/178229

DelGiudice, G. D., Sampson, B. A., & Giudice, J. H. (2013). A long-term assessment of the effect of winter severity on the food habits of white-tailed deer. *The Journal of Wildlife Management*, *77*(8), 1664–1675. https://doi.org/10.1002/jwmg.616

Demaria, E. M. C., Roundy, J. K., Wi, S., & Palmer, R. N. (2016). The Effects of Climate Change on Seasonal Snowpack and the Hydrology of the Northeastern and Upper Midwest United States. *Journal of Climate*, *29*(18), 6527–6541. https://doi.org/10.1175/JCLI-D-15-0632.1

Dufrêne, M., & Legendre, P. (1997). Species Assemblages and Indicator Species:the Need for a Flexible Asymmetrical Approach. *Ecological Monographs*, *67*(3), 345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2

Dumont, A., Ouellet, J.-P., Crête, M., & Huot, J. (2005). Winter foraging strategy of white-tailed deer at the northern limit of its range. *Écoscience*, *12*(4), 476–484. https://doi.org/10.2980/i1195-6860-12-4-476.1

Eichenlaub, V. L. (1970). lake effect snowfall to the lee of the Great Lakes: Its role in Michigan. *Bulletin of the American Meteorological Society*, *51*(5), 403–413. https://doi.org/10.1175/1520-0477(1970)051<0403:LESTTL>2.0.CO;2

Ellison, A. M., Orwig, D. A., Fitzpatrick, M. C., & Preisser, E. L. (2018). The Past, Present, and Future of the Hemlock Woolly Adelgid (Adelges tsugae) and Its Ecological Interactions with Eastern Hemlock (Tsuga canadensis) Forests. *Insects*, *9*(4), 172. https://doi.org/10.3390/insects9040172

Forsyth, D. M., Barker, R. J., Morriss, G., & Scroggie, M. P. (2007). Modeling the Relationship between Fecal Pellet Indices and Deer Density. *The Journal of Wildlife Management*, *71*(3), 964–970.

Foster, D. R., Oswald, W. W., Faison, E. K., Doughty, E. D., & Hansen, B. C. S. (2006). A Climatic Driver for Abrupt Mid-Holocene Vegetation Dynamics and the Hemlock Decline in New England. *Ecology*, *87*(12), 2959–2966. https://doi.org/10.1890/0012-9658(2006)87[2959:ACDFAM]2.0.CO;2

Frelich, L. E., & Lorimer, C. G. (1985). Current and predicted long-term effects of deer browsing in hemlock forests in Michigan, USA. *Biological Conservation*, *34*(2), 99–120. https://doi.org/10.1016/0006-3207(85)90103-X

Gold, Z., Curd, E. E., Goodwin, K. D., Choi, E. S., Frable, B. W., Thompson, A. R., Walker Jr., H. J., Burton, R. S., Kacev, D., Martz, L. D., & Barber, P. H. (2021). Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. *Molecular Ecology Resources*, *21*(7), 2546–2564. https://doi.org/10.1111/1755-0998.13450

Grund, M. D., McAninch, J. B., & Wiggers, E. P. (2002). Seasonal Movements and Habitat Use of Female White-Tailed Deer Associated with an Urban Park. *The Journal of Wildlife Management*, *66*(1), 123–130. https://doi.org/10.2307/3802878

Hart, J. L. (2008). *EASTERN HEMLOCK DECLINE AND PERSISTENCE OF DISJUNCT POPULATIONS NEAR ITS SOUTHERN BOUNDARY*. *79*(3).

Hart, M. L., Meyer, A., Johnson, P. J., & Ericsson, A. C. (2015). Comparative Evaluation of DNA Extraction Methods from Feces of Multiple Host Species for Downstream Next-Generation Sequencing. *PLOS ONE*, *10*(11), e0143334. https://doi.org/10.1371/journal.pone.0143334

Hix, D. M., & Barnes, B. V. (1984). Effects of clear-cutting on the vegetation and soil of an eastern hemlock dominated ecosystem, western Upper Michigan. *Canadian Journal of Forest Research*, *14*(6), 914–923. https://doi.org/10.1139/x84-163

Hu, L., Xu, X., Wang, J., & Xu, H. (2023). Individual tree crown width detection from unmanned aerial vehicle images using a revised local transect method. *Ecological Informatics*, *75*, 102086. https://doi.org/10.1016/j.ecoinf.2023.102086

Jensen, N. R., Webster, C. R., Witt, J. C., & Grant, J. B. (2011). Ungulate winter habitat selection as a driver of herbaceous-layer heterogeneity in northern temperate forests. *Ecosphere*, *2*(6), art67. https://doi.org/10.1890/ES10-00189.1

LaGory, M. K., LaGory, K. E., & Taylor, D. H. (1985). Winter Browse Availability and Use by White-Tailed Deer in Southeastern Indiana. *The Journal of Wildlife Management*, *49*(1), 120–124. https://doi.org/10.2307/3801855

Mason, R. (n.d.). *Michigan Deer Management Plan*.

McIntosh, A., Gray, A., & Garman, S. (2012). Estimating canopy cover from standard forest inventory measurements in western Oregon. *Forest Science. 58(2): 154-167*, *58*, 154–167.

Miraki, M., Sohrabi, H., Fatehi, P., & Kneubuehler, M. (2021). Individual tree crown delineation from high-resolution UAV images in broadleaf forest. *Ecological Informatics*, *61*, 101207. https://doi.org/10.1016/j.ecoinf.2020.101207

Mladenoff, D. J., & Stearns, F. (1993). Eastern Hemlock Regeneration and Deer Browsing in the Northern Great Lakes Region: A Re-examination and Model Simulation. *Conservation Biology*, *7*(4), 889–900.

Moen, A. N. (1976). Energy Conservation by White-Tailed Deer in the Winter. *Ecology*, *57*(1), 192–198. https://doi.org/10.2307/1936411

Morrison, S. F., Forbes, G. J., & Young, S. J. (2002). Browse occurrence, biomass, and use by white-tailed deer in a northern New Brunswick deer yard. *Canadian Journal of Forest Research*, *32*(9), 1518–1524. https://doi.org/10.1139/x02-081

Morrison, S. F., Forbes, G. J., Young, S. J., & Lusk, S. (2003a). Within-yard habitat use by white-tailed deer at varying winter severity. *Forest Ecology and Management*, *172*(2), 173–182. https://doi.org/10.1016/S0378-1127(01)00809-X

Nelson, M. E. (1993). Natal Dispersal and Gene Flow in White-Tailed Deer in Northeastern Minnesota. *Journal of Mammalogy*, *74*(2), 316–322. https://doi.org/10.2307/1382386

Nelson, M. E. (1995). Winter range arrival and departure of white-tailed deer in northeastern Minnesota. *Canadian Journal of Zoology*, *73*(6), 1069–1076. https://doi.org/10.1139/z95-127

Nelson, M. E., & Mech, L. D. (1981). Deer Social Organization and Wolf Predation in Northeastern Minnesota. *Wildlife Monographs*, *77*, 3–53.

Norton, A. S., Storm, D. J., & Van Deelen, T. R. (2021). White-Tailed Deer, Weather and Predation: A New Understanding of Winter Severity for Predicting Deer Mortality. *The Journal of Wildlife Management*, *85*(6), 1232–1242. https://doi.org/10.1002/jwmg.22083

Orwig, D. A., Foster, D. R., & Mausel, D. L. (2002). Landscape patterns of hemlock decline in New England due to the introduced hemlock woolly adelgid. *Journal of Biogeography*, *29*(10–11), 1475–1487. https://doi.org/10.1046/j.1365-2699.2002.00765.x

Osborn, R. G., & Jenks, J. A. (1998). Assessing Dietary Quality of White-Tailed Deer Using Fecal Indices: Effects of Supplemental Feeding and Area. *Journal of Mammalogy*, *79*(2), 437–447. https://doi.org/10.2307/1382974

Ozoga, J. J., & Gysel, L. W. (1972a). Response of White-Tailed Deer to Winter Weather. *The Journal of Wildlife Management*, *36*(3), 892–896. https://doi.org/10.2307/3799445

Ozoga, J. J., & Verme, L. J. (1984). Effect of Family-Bond Deprivation on Reproductive Performance in Female White-Tailed Deer. *The Journal of Wildlife Management*, *48*(4), 1326–1334. https://doi.org/10.2307/3801793

Paradis, A., Elkinton, J., Hayhoe, K., & Buonaccorsi, J. (2008). Role of winter temperature and climate change on the survival and future range expansion of the hemlock woolly adelgid (Adelges tsugae) in eastern North America. *Mitigation and Adaptation Strategies for Global Change*, *13*(5), 541–554. https://doi.org/10.1007/s11027-007-9127-0

Parikh, G. L., & Webster, C. R. (2019). Press-Pulse Odocoileus Virginianus Herbivory in Relict Tsuga Canadensis Stands in the Western Upper Peninsula of Michigan, USA. *Forests*, *10*(6), Article 6. https://doi.org/10.3390/f10060496

Schmitz, O. J. (1991). Thermal constraints and optimization of winter feeding and habitat choice in white-tailed deer. *Ecography*, *14*(2), 104–111. https://doi.org/10.1111/j.1600-0587.1991.tb00640.x

Suggitt, A. J., Gillingham, P. K., Hill, J. K., Huntley, B., Kunin, W. E., Roy, D. B., & Thomas, C. D. (2011). Habitat microclimates drive fine-scale variation in extreme temperatures. *Oikos*, *120*(1), 1–8. https://doi.org/10.1111/j.1600-0706.2010.18270.x

Telfer, E. S. (1970). Winter Habitat Selection by Moose and White-Tailed Deer. *The Journal of Wildlife Management*, *34*(3), 553–559. https://doi.org/10.2307/3798862

Van Deelen, T. R., Campa, H., Hamady, M., & Haufler, J. B. (1998a). Migration and Seasonal Range Dynamics of Deer Using Adjacent Deeryards in Northern Michigan. *The Journal of Wildlife Management*, *62*(1), 205–213. https://doi.org/10.2307/3802280

Varhola, A., Coops, N. C., Bater, C. W., Teti, P., Boon, S., & Weiler, M. (2010). The influence of ground- and lidar-derived forest structure metrics on snow accumulation and ablation in disturbed forests. *Canadian Journal of Forest Research*, *40*(4), 812–821. https://doi.org/10.1139/X10-008

Verme, L. J. (1968). An Index of Winter Weather Severity for Northern Deer. *The Journal of Wildlife Management*, *32*(3), 566–574. https://doi.org/10.2307/3798937

Witt, J. C., & Webster, C. R. (2010a). Regeneration dynamics in remnant Tsuga canadensis stands in the northern Lake States: Potential direct and indirect effects of herbivory. *Forest Ecology and Management*, *260*(4), 519–525. https://doi.org/10.1016/j.foreco.2010.05.007

Witt, J. C., Webster, C. R., Froese, R. E., Drummer, T. D., & Vucetich, J. A. (2012). Scale-dependent drivers of ungulate patch use along a temporal and spatial gradient of snow depth. *Canadian Journal of Zoology*, *90*(8), 972–983. https://doi.org/10.1139/z2012-065

Wuebbles, D. J., & Hayhoe, K. (2004). Climate Change Projections for the United States Midwest. *Mitigation and Adaptation Strategies for Global Change*, *9*(4), 335–363. https://doi.org/10.1023/B:MITI.0000038843.73424.de

3 Impact of Storage Time, DNA Extraction Kit, and Target Locus on Fecal Metabarcoding for Diet, Host, and Microbial Analysis in Upper Peninsula of Michigan White-Tailed Deer

3.1 Introduction

3.1.1 Fecal DNA Metabarcoding as a Molecular Tool

Fecal DNA Metabarcoding is an increasingly used amplicon sequencing technique within molecular ecology in which universal primers targeting genomic regions with high interspecies variation amplify templates within fecal samples belonging to the target group. As a type of environmental DNA (eDNA), feces contain genetic material from not only the host species but also prey items, gut microbes, and environmental contaminants, making it an extremely useful sample type for diet and biodiversity studies. For example, previous studies have leveraged fecal metabarcoding to quantify seasonal diet variation (Goldberg et al., 2020), monitor aquatic and terrestrial invasive or introduced species (Guillerault et al., 2017; Westfall et al., 2020), characterize remote community biodiversity (Nørgaard et al., 2021), and identify disease-associated parasites on local and population level scales (Davey et al., 2021; Garwood et al., 2023). Additionally, fecal sample collection is generally lower-effort, cost-effective, and noninvasive, increasing the benefits for research efforts focused on sensitive, endangered, or cryptic species (Andriollo et al., 2021; Young et al., 2020), especially in more remote global areas (Kartzinel et al., 2015; Schuette et al., 2022). However, despite the power of this new molecular tool, there remains a lack of methodological standardization and consensus even within studies targeting similar taxonomic groups or habitat types (Ando et al., 2020), with a limited number of studies examining how methodology impacts eDNA sequencing results (De Barba et al., 2014; Divoll et al., 2018; Ruppert et al., 2019).

3.1.2 Major Sources of Error and Bias

Amidst the murky methodology, fecal metabarcoding is susceptible to various sources of bias and error that can substantially reduce the accuracy and reliability of results (Figure 2.1). Sample collection and storage introduce potential issues such as environmental contamination, DNA degradation, and low DNA-quantity due to improper handling or prolonged storage. While Krehenwinkel et al., 2018 found biased taxon recovery attributed to differential DNA degradation in mock arthropod community with different preservation regimes, a comparison of fresh and degraded wolf scat found decreased sequence abundance did not impact prey detection (Massey et al., 2021), indicating the contribution of variation in sample quality to error is unclear. DNA extraction introduces further bias through tube-to-tube contamination, PCR inhibitors, and specific extraction chemistries, all of which can affect the quality and quantity of extracted fecal DNA. Fecal matter contains abundant PCR inhibitors such as bile salts and complex polysaccharides that interfere in enzymatic activity and DNA binding during PCR amplification, with

feces from herbivores harboring even higher inhibitor levels due to plant secondary metabolites common in leaf tissue (Schrader et al., 2012; Sidstedt et al., 2020). Commercially available extraction kits optimize chemistries of buffers, washing solutions, and filtration methods for specific tissue or sample types to remove PCR inhibitors most efficiently without reducing purified DNA yield (Vishnivetskaya et al., 2014), but there is no consensus on optimal kit selection. Although a majority of animal-targeted studies utilize a stool specific kit like QIAamp (Fast) DNA Stool (Ando et al., 2020), comparative studies like Ingala et al., 2021 have found kits optimized for plant tissue perform better in herbivores ((Galan et al., 2018).

PCR amplification specific biases may arise from factors such as universal primer choice, which might preferentially amplify certain taxa, and target amplicon length, which reduces amplification efficiency and decreases coverage of diverse taxa. The highly degraded and low-quantities of fecal DNA causes technical PCR issues often circumvented through increased cycle numbers and sample-optimized annealing temperature (van der Loos & Nijland, 2021), but researchers also tend to select organellular genomic loci with shorter amplicon lengths to increase amplification success (Ando et al., 2020; Mallott et al., 2018; Taberlet et al., 2007). Lab contamination poses another risk, particularly in low-biomass fecal samples, and has been shown to produce false positives through introduction of exogenous DNA or tube-to-tube contamination. Efforts to mitigate this include establishment of designated low-copy spaces and strict adherence to sterilization and workflow protocols, but reaction plate designs incorporating biological replicates and multiple negative, positive, and tag-jumping controls facilitate the estimation of error rates and enhance confidence in dietary profiling outcomes (Alberdi et al., 2018; Ragot et al., 2023; van der Loos & Nijland, 2021). Perhaps the largest obstacle, taxonomic resolution can be compromised by the availability of reference databases and inherent diversity of prey across different loci, limiting the precise classification of certain taxa at high resolution. Public databases like NCBI or EMBL are incomplete with uneven representation of species or taxonomic groups increasing misclassification, particularly when taxonomic ambiguity is high at the selected metabarcoding marker. Therefore, careful consideration, optimization, and standardization of protocols is essential to a metabarcoding design that mitigates issues surrounding eDNA studies.

3.1.3 The Power of Poop in Molecular Ecology

Nevertheless, fecal metabarcoding is an extremely powerful high-throughput molecular method enabling high resolution of target taxonomic groups within a host species diet along with their microbiome (Ando et al., 2020; de Sousa et al., 2019; Thuo et al., 2019). Diet shapes trophic interaction and food web dynamics within ecosystems that control the flow of energy and nutrients supporting biodiversity, population health, ecosystem services (Carreon-Martinez & Heath, 2010; Ferguson et al., 2018; Polis & Strong, 1996). In forested systems, keystone large mammalian herbivores disproportionately influence plant community structure and diversity– Suzuki et al., 2013 identified deer herbivory as the primary determinant of floristic composition in a productive mesic forest, whereas Forrester et al., 2014 evidenced how ungulate browsing reduced tree regeneration and

shifted species composition in temperate forest gaps. These alterations radiate to higher trophic levels, and the impacts may be heightened in sensitive ecosystems or those with acute temporal herbivory. For example, northern white-tailed deer populations practice winter yarding that concentrate herbivory impact to specific forest habitat dominated by slow-growing conifers eastern hemlock (*Tsuga canadensis*) and northern white-cedar (*Thuja canadensis*) (Bradshaw & Waller, 2016). Complex interactions of regeneration dynamics with acute-herbivory associated severe winter conditions and subsequent population declines may be accelerating forest composition shifts toward deciduous species in hemlock stands (Parikh & Webster, 2019; Salk et al., 2011). A more in-depth understanding of temporal variation in winter diet of white-tailed deer and potential links to landscape, anthropogenic, and climatic factors is critical for effective management of both deer populations and forest ecosystems.

Beyond diet, feces can also be used to investigate gut microbiome communities with potential implications for individual health and disease-risk. For example, assessment of gut microbiome communities from yarding deer could elucidate the impact of the spatiotemporal impact of supplemental feeding for deer health and survivorship, with potential links to changes in diet composition over time (Cooper et al., 2006; Miller et al., 2003). Additionally, host DNA within feces allows for higher level population structure analyses via mitochondrial sequencing, or individual identification via microsatellite genotyping (Brinkman et al., 2010). Given the high socioeconomic impact of white-tailed deer as a prominent game species (Grado et al., 2007), the accuracy and cost-effectiveness of noninvasive fecal sampling as means to estimate landscape-scale population dynamics makes it an extremely valuable tool for management agencies.

However, the cost of molecular analyses can be extremely limiting for research. Although advances in NGS sequencing technologies and high-throughput lab procedures over the past two decades have made molecular work more accessible for ecological studies (Ekblom & Galindo, 2011), the supply and labor cost of processing the same sample with different methods and extraction kits remains a prominent barrier, often reducing the number of samples included in a study or types of analyses (Ando et al., 2020; Beauchamp et al., 2011). Development of streamlined protocols for noninvasive fecal samples that mitigate the numerous potential sources of error and bias would be extremely beneficial for increasing the accessibility and applicability of robust and powerful techniques for answering complex ecological questions.

3.1.4 Study Objectives

To that end, in this study we leveraged a robust temporal dataset of white-tailed deer fecal samples from an Upper Peninsula of Michigan deer yard to assess how methodological factors impact DNA analyses from noninvasive feces quantifying winter diet, host populations genetics, and microbial communities. Our first objective was to validate fecal DNA metabarcoding and host DNA analyses on ungulate fecal samples stored for variable lengths and determine the impact of storage time. Our second objective was to assess how choice of DNA extraction kit and target locus would impact diet, host, and microbial sequencing results of fecal metabarcoding and host

mitochondrial and microsatellite analyses. Finally, our third objective to was develop an optimized single-extraction protocol for downstream analyses of all DNA components found within fecal samples for cost and time efficiency without sacrificing data quality. Our findings can be utilized to further investigate interactions between ungulates, winter dietary preferences, population structure, and associated microbiomes specifically, and more broadly provide insights into the methodological considerations crucial for accurate and comprehensive eDNA analysis.

3.2 Materials and Methods

3.2.1 Study Area and Sample Collection

As part of an ongoing collaborative monitoring project of white-tailed deer winter habitat structure and usage in the western Upper Peninsula of Michigan, 39 remnant eastern hemlock patches varying from 0.4 to 0.6 ha in size were established in 2005 (Witt and Webster 2010). Fecal samples were collected from these stands during annual springtime pellet-group surveys of 3-30 (plot number dependent on stand area) randomly-placed 9.29 m^2 circular plots (radius = 1.72m) conducted directly following spring snowmelt from 2006 to 2023. A minimum threshold of 10 intact pellets was required per pellet-group and groups deposited on wet soil were excluded from collection. Selected fecal samples were kept on ice post field-sampling and immediately transferred to Michigan Technological University (Houghton, MI, USA) for long-term storage at -20°C . We used a subset of these fecal samples for our study testing fecal metabarcoding methodology selected from a single high-usage stand near Point Abbaye, MI.

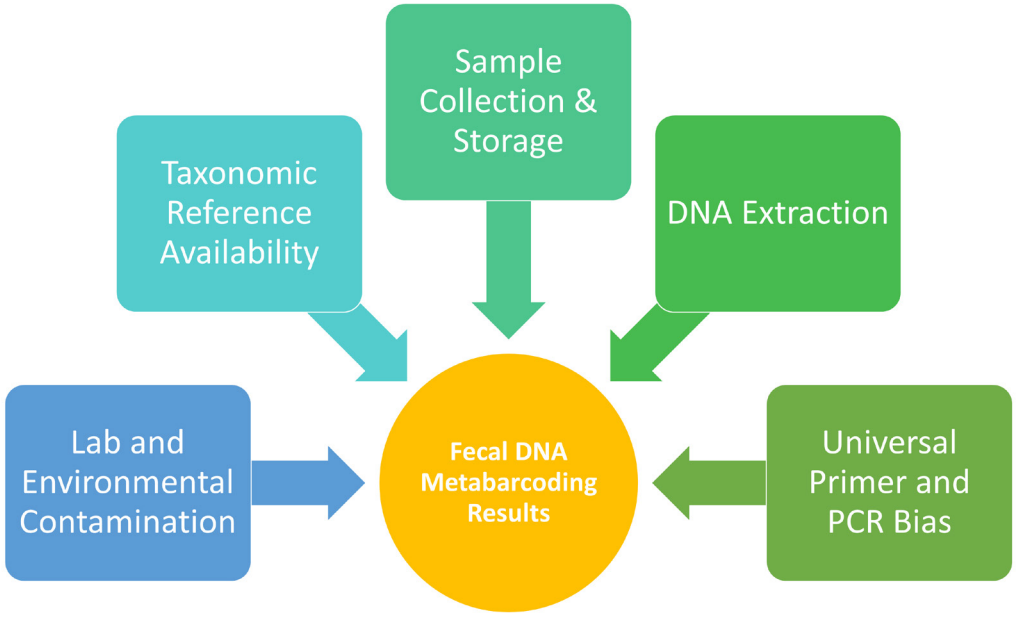


Figure 3.1. Flowchart of major potential biases impacting fecal DNA metabarcoding.

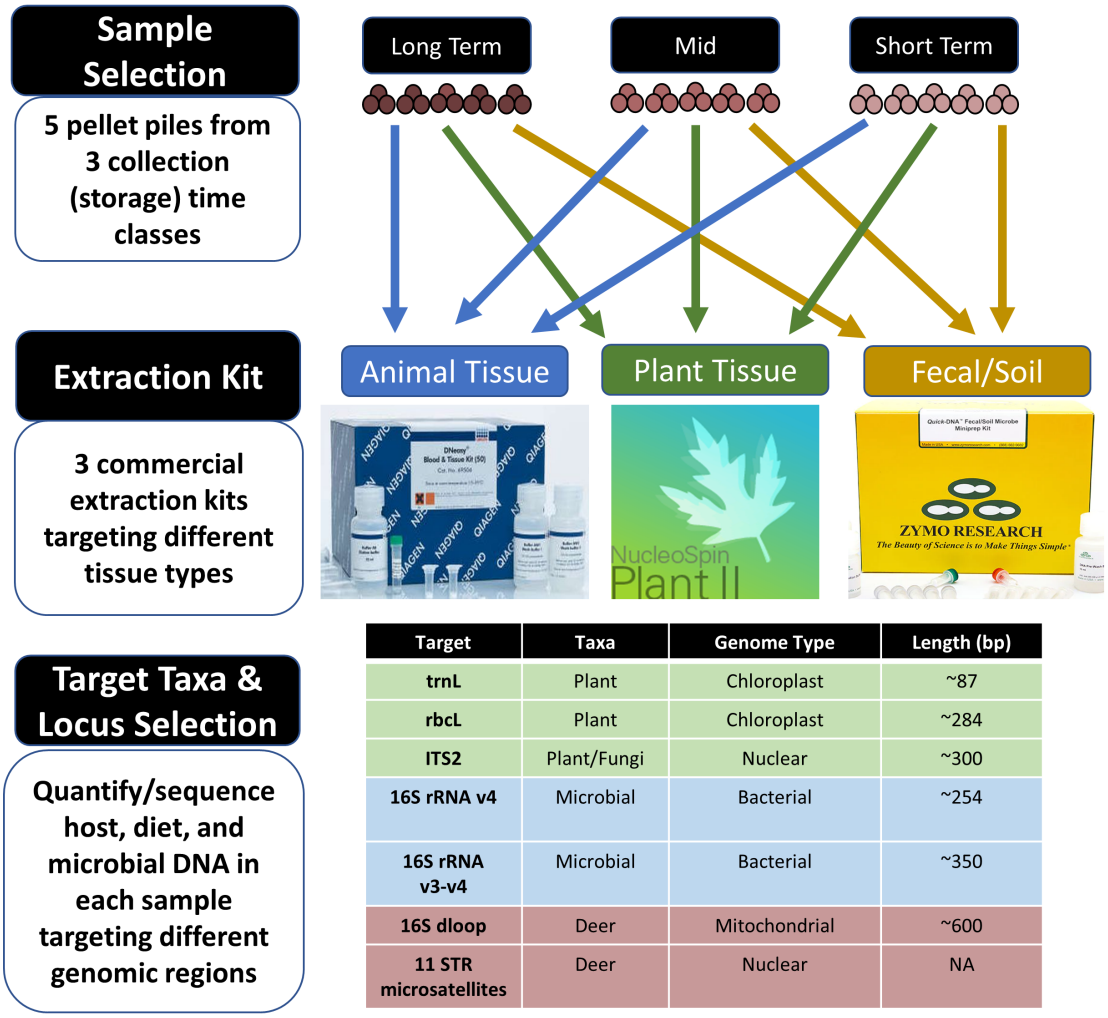
3.2.2 Sample Selection and Processing

For a total of 15 fecal samples with variable storage length, we selected five fecal-group samples from 3 different storage time classes: long-term (2006), mid-term (2012-2015), and short-term (2020-2021). Due to sample availability limitations, all pellet-groups for the long-term storage class were from 2006 only, where the other time classes span multiple years within the designated range.

For individual pellet-group processing for DNA extraction, 5 fully-intact frozen pellets were randomly selected to minimize environmental contamination (Ando et al., 2018) and then manually homogenized into a fine powder within a UV-sterilized processing hood. Homogenized samples were immediately returned to the -20°C freezer to minimize any genomic DNA shearing due to thawing and refreezing the samples (Röder et al., 2010; Wu et al., 2019).

3.2.3 DNA Extraction

DNA extraction for each homogenized fecal sample (n=15) was conducted using three different commercially-available extraction kits: Qiagen DNeasy Blood & Tissue (QIAGEN, Valencia, CA), TakaraBio NucleoSpin Plant II (Clontech Laboratories, Inc., Mountain View, CA), and Zymo Research Fecal/Soil Miniprep (Zymo Research, Irvine, CA). All extractions were carried out in a laboratory space designated for low-quantity DNA samples with strict sterilization and workflow procedures, and extraction for each commercial kit was carried out on different days. A negative control was included in every extraction to quantify initial cross-sample contamination. All three extractions followed recommended kit protocols provided by manufacturers with slight modifications of fecal eDNA samples. In general, all protocols were modified to include a bead-beating procedure and separation of fecal matter from supernatant following lysis. Detailed descriptions of protocol modifications per kit are included in Chapter 4 of this thesis.



Target	Taxa	Genome Type	Length (bp)
trnL	Plant	Chloroplast	~87
rbcL	Plant	Chloroplast	~284
ITS2	Plant/Fungi	Nuclear	~300
16S rRNA v4	Microbial	Bacterial	~254
16S rRNA v3-v4	Microbial	Bacterial	~350
16S dloop	Deer	Mitochondrial	~600
11 STR microsatellites	Deer	Nuclear	NA

Figure 3.2. Overview of pilot study design testing impacts of storage time, DNA extraction kit, and target locus on fecal DNA metabarcoding for diet, host, and microbial components within ungulate feces. Five fecal samples within 3 storage time classes were each extracted using 3 commercially available DNA extraction kits optimized for different sample types: animal tissue (Qiagen DNeasy Blood & Tissue), plant tissue (TakaraBio NucleoSpin Plant II), and environmental samples (Zymo Research Fecal/Soil Miniprep). A summary of universal plant and microbial metabarcoding primers and deer host specific mitochondrial and microsatellite primers include target taxa, genomic type, and expected amplicon length is included.

3.2.4 Metabarcoding Primers and Library Preparation

We assessed the diet component of the extracted fecal DNA with a dual-PCR metabarcoding methodology for Illumina MiSeq Sequencing using workflows modified from previous diet metabarcoding studies (Shi et al. 2021, Erickson et al. 2017). Given the typical low-quality DNA extracted from environmental samples, we tested three commonly used universal plant metabarcoding primers of varying amplicon lengths and genomes for ability to quantify and resolve herbivorous diets from fecal samples. The shortest amplicon (~86 bp) targeted the P6 loop of the chloroplast trnL (UAA) locus, with the locus specific forward primer (g) sequence: 5'- GGG CAA TCC TGA GCC AA-3', and the reverse primer (h) sequence: 5'- CCA TTG AGT CTC TGC ACC TAT C -3' (Taberlet et al. 2007). We also amplified at a second chloroplast marker, rbcL, targeting a mini-barcoded region with a larger amplicon size (~379 bp), with locus-specific forward primer: 5'- CTT ACC AGY CTT GAT CGT TAC AAA GG-3', and reverse primer 5'- GTA AAA TCA AGT CCA CCR CG-3′ (Erickson et al. 2017). The third universal plant primer pair targeted a nuclear locus, the second internal transcribed spacer (ITS2), with an expected amplicon size of 187-387bp with locus-specific forward primer UniPlantF (5′-TGT GAA TTG CAR RAT YCM G-3′) and reverse primer UniPlantR (5′-CCC GHY TGA YYT GRG GTC DC-3′) (Moorhouse-Gann et al. 2018). Primers were synthesized as fusion primers with 5' adapters for custom Illumina barcoded sequencing primers (see 16S Metagenomic Protocol) and a trimming adaptor for bioinformatic processing.

We implemented an identical metabarcoding experimental design for the fecal microbial DNA assessment. We conducted two library preparations with 16S rRNA universal primers targeting bacteria and archaea, with one shorter amplicon with just the V4 region (amplicon size ~390 bp) and another longer amplicon with the V3 and V4 regions (amplicon size: 400-600 bp). The short amplicon used a modified universal primer set developed by Parada et al. 2016 (515Fmod: 5'- GTG YCA GCM GCC GCG GTA A -3' and 806Rmod: 5'- GGA CTA CNV GGG TWT CTA AT -3'). The locus-specific primers for the V3-V4 region amplicon were Pro341F (5′-C-CTA CGG GNB GCA SCA G-3′) and Pro805R (5′-GAC TAC NVG GGT ATC TAA TC-C-3′), the reverse primer including an additional degenerate base pair (Schmidt et al. 2019).

Plant and bacterial DNA from fecal samples was amplified in locus-specific PCR reactions with optimized reagent concentrations and thermocycling conditions for each primer set (Table 6.2, Table 6.3). All 45 fecal extractions were amplified in triplicate 20 µL reactions subjected to identical PCR conditions for all 5 universal primers, where reaction plates included blank wells for tracking contamination and tag-jumping. The reaction cocktail for the two chloroplast plant primers (trnL & rbcL) consisted of 2 µL (1X) of 10X Buffer, 1.6 µL (2mM) of MgCl2, 0.4 µL (0.2 µM) each of 10 uM forward and reverse primer, 0.14 µL (1.75X) of AmpliTaq Gold 360 polymerase, 0.1 µL (0.1mM) BSA, 2 µL of DNA, and 12.56 µL of nanopure water. The reaction cocktail for the nuclear plant primer (ITS2) and the two 16S rRNA primers consisted of 10 µL (1X) of 2X AmpliTaq Gold 360 PCR Master Mix, 0.4 µL (0.2 µM) each of 10 uM forward and reverse primer, 0.1 µL (0.1mM) BSA, 2 µL of DNA, and 7.1 µL of nanopure water. Thermocycling was performed on an Applied Biosystems Verti or 2720 Thermal Cycler

depending on availability in the lab. DNA from each locus-specific reaction was purified with Serapure beads (1.5x to PCR reaction volume) and washed with 100% ethanol twice before elution in nanopure water to recover 20 µL of clean DNA solution. Amplicons of expected size range were verified before and after bead cleaning with gel electrophoresis imaging for a subset of samples.

Cleaned PCR1 product from the locus-specific amplifications was used in a second PCR with custom primers containing a unique combination of 8bp barcodes and Illumina MiSeq flow cell adapters. For all indexing PCR reactions (20 µL total volume), we used: 2 µL (1X) of 10X Buffer, 1.6 µL (2mM) of MgCl2, 0.8 µL (2mM each) dNTPs, 1 µL (0.5 µM) each of 10 uM forward and reverse primer, 0.1 µL (1.25X) of AmpliTaq Gold 360 polymerase, 6 µL of DNA, and 7.5 µL of nanopure water. Each reaction was performed on an Applied Biosystems thermocycler with the program: 1 cycle 95 °C 10 min; 8 cycles of: 94 °C for 30 s, 55 °C for 60 s, 72 °C for 60 s, 1 final extension of 72 °C for 10 min; hold at 4 °C. Cleanup and normalization was performed with Charm Biotech Just-a-Plate 96 PCR Purification Kit (Charm Biotech, San Diego, CA, USA) using manufacturers recommended protocol, and 5 µL of all equal molar indexed samples were pooled into a single solution. The final concentrations of all five sequencing libraries were measured with a dsDNA Qubit assay and a high-sensitivity Agilent TapeStation (Santa Clara, CA, USA) verified library size. All libraries consisted of 142 uniquely barcoded samples, including 135 fecal DNA products and 7 extraction and PCR controls.

Four of the indexed libraries with amplicon sizes ranging from 410-589 bp, rbcL (513 bp), ITS2 (482 bp), 16S rRNA V4 (406bp), and 16S rRNA V3-V4 (584 bp), were diluted to 1 nM with nanopure water and 10 µL of each library was pooled into a single 40 µL solution. The pooled libraries were diluted and denatured to 20 pM following Illumina's recommended protocol for low concentration libraries (< 2 nM) and sequenced using an Illumina MiSeq with a 600 bp V3 kit (301 cycles). Due to concerns about sequencing bias with shorter amplicon sizes, the indexed trnL library was sequenced separately with a 300 bp V2 Nano kit (151 cycles) using the same low concentration denaturation and dilution protocol. Both sequencing runs included a 20% PhiX spike to increase base-calling diversity.

3.2.5 Mitochondrial Sequencing and Microsatellite Fragment Analysis

We employed two methodologies for assessing host DNA extracted from the fecal samples. First, the mitochondrial genome was targeted using deer-specific 16S rRNA primers (Forward: 5' – CGC CTG TTT ATC AAA AAC AT -3' and Reverse: 5'- CTC CGG TTT GAA CTC AGA TC- 3') (Hoffman et al. 2015). The same 45 purified DNA samples representing the 15 fecal samples extracted with 3 different commercial kits were amplified in duplicate 15 µL reactions with a PCR reagent cocktail containing: 7.5 µL (1X) of 2X AmpliTaq Gold 360 PCR Master Mix, 0.3 µL (0.2 µM) each of 10 uM forward and reverse primer, 2 µL of DNA, and 4.9 µL of nanopure water. All sample reactions were run on an Applied Biosystems Verti system with the program: 1 cycle 95

°C 10 min; 30 cycles of: 94 °C for 30 s, 49 °C for 30 s, 72 °C for 45 s, 1 final extension of 72 °C for 10 min; hold at 4 °C. After verifying successful amplification of a ~550 bp fragment with gel electrophoresis, all samples were purified with an Exo-Sap-IT protocol and sent to GeneWiz (South Plainfield, New Jersey, USA) for Sanger Sequencing.

Second, we assessed nuclear host DNA with an 11-locus short-tandem repeat (STR) panel protocol modified from Miller, W.L., et.al. 2019 that targets 11 informative deer microsatellites with 5 multiplexed PCR reactions. Due to material constraints and depleted fecal DNA, each multiplex consisted of 35 of the original 45 fecal samples run in duplicate or triplicate to generate samples for comparison sufficient for our purposes versus the 4 replicates per samples standard for heterozygous allele verification commonly used in population studies. The five multiplexed PCR reaction cocktails consisted of 5 µL of 2x Qiagen Multiplex PCR Master Mix, 3 µL of forward and reverse primer master mix, 1 µL of 5X Q-solution, and 1 µL of purified fecal DNA. PCR reactions were performed on an Applied Biosystems Verti system with the program: 1 cycle 95 °C for 15 min; 35 cycles of: 94 °C for 30 s, variable annealing temperature for 30 s, 72 °C for 60s, 1 final extension of 72 °C for 10 min; hold at 4 °C. The multiplexed PCR products were combined with LiZ500 size standard and HiDi and sent to the Cornell Institute of Biotechnology Genomics Facility (Ithaca, New York, USA) for Microsatellite Fragment Analysis on an Applied Biosystems 3730xl instrument.

3.2.6 Metabarcoding Bioinformatics

All five plant (trnL, rbcL, ITS2) and microbial (16S rRNA V3, V3-V4) sequencing libraries were subjected to identical bioinformatic processing pipelines adjusted for amplicon specific information. The sequences produced by both Illumina MiSeq runs were trimmed and demultiplexed on the instrument by bcl2fastq Conversion Software (v2.20, https://support.illumina.com/) using the provided trimming adaptor sequences and barcode combinations, with each of the 710 sequenced samples (675 fecal samples, 35 negative PCR and extraction controls) resulting in separate reverse and forward read FASTQ sequence files. Raw sequence quality reports for all sequence files were generated using MultiQC v.1.14 (Ewels et al. 2016), and all reads were filtered for a minimum quality score of Q=30 and minimum length of 100 bp (50 bp for trnL) with bbduk v.35.74 (Bushnell 2017). Locus-specific primer sequences for each amplicon were identified and removed from the 5’ and 3’ ends with cutadapt v.4.2c (Martin 2011) from all forward and reverse reads, with a minimum length of 1 bp to exclude any sequences with length of zero post-filtering and primer removal. Primer sequences were not removed from trnL reads due to short amplicon target length and downstream issues with minimum length requirements for taxonomic classification (>50bp). Additional quality filtering, error correction, read merging, and chimera detection was performed using the DADA2 v.3.14 package for R (Callahan et al. 2016) using a modified workflow outlined in detail on the DADA2 webpage (https://benjjneb.github.io/dada2/tutorial.html) . Additional quality filtering removed reads with ambiguous bases, >2 expected errors in either of the forward or reverse reads, PhiX matching sequence, and length < threshold

set based on expected size for the given amplicon after filtering. Error rate was estimated with filtered sequences and amplicon sequence variants (ASVs) were inferred using DADA2's core denoising algorithm (Callahan et al. 2016) for each sample. Forward and reverse reads were merged into a single contig requiring a minimum overlap of 12 bp and allowing a 4bp mismatch, and then utilized to construct a sequence table of read counts per ASV per sample. Chimeric ASVs were flagged on a sample-by-sample basis and removed if consensus across all samples was found (bimeric in >90% of samples it occurs in).

For all of the target amplicons besides trnL ASV taxonomy was assigned using default parameters of the naïve Bayesian classifier implemented in DADA2 with amplicon-specific reference databases. These databases were the rbcL-specific reference database adapted from Bell et al. 2017 (Bell 2021, FigShare Download), the ITS region UNITE v.9 all eukaryotic reference database, and the SSU Silva Ref nr 99 v138.1 16S/18S rRNA reference database (Quast et al. 2013). For trnL, a custom-formatted NCBI BLAST returning the top 10 hits per ASV was used as input for the Assign-Taxonomy-with-BLAST python script (2020, https://github.com/Joseph7e/Assign-Taxonomy-with-BLAST) available on GitHub with default parameters to determine best hits for taxonomic resolution. The CRAN taxonomizer v.0.10.2 package for R was then implemented using NCBI accession numbers from the top hits to resolve taxonomy for each trnL ASV (https://CRAN.R-project.org/package=taxonomizr) to the highest taxonomic level. For the amplicons targeting plant genetic markers, ASVs unable to be assigned with >80% to phylum "Viridiplantae" or "Streptophyta" were considered off-target sequencing and excluded from further analysis (Davey et al. 2023). However, the ITS2 marker is also able to identify fungi and other eukaryotes, and ASVs associated with these taxonomic groups were considered separately. The vast majority (>98%) of 16S rRNA V4 ASVs were unable to be taxonomically assigned beyond the kingdom level and we consequently excluded all sequencing results for the V4 region from further analysis. The successful taxonomic assignment of the longer V3-V4 region 16S rRNA amplicon suggests that issues with the shorter fragment can be attributed to off-target amplification or sequencing issues. For the three plant metabarcoding primers we compared both the number of unbiased ASVs generated across all the samples, the percentages that resolve to each taxonomic level, and the taxonomic overlap in diet detection for plant families and genera.

3.2.7 Host DNA Quantification

For mitochondrial sequencing, we used trace files to manually remove low-quality or ambiguous base calls from the FASTA sequences for all samples and then submitted the sequences to NCBI BLAST to verify identification of white-tailed deer. To detect differences in host DNA amplification related to DNA extraction kit or storage time an average quality score for the two replicates produced by Sanger Sequencing was calculated for each sample, with n>40 high-quality, 25>n>40 acceptable quality, and n<25 low-quality or failure as defined by GENEWIZ.

For microsatellites, we used Geneious Prime 2023 (https://www.geneious.com) to compare the electropherograms with a known DNA size standard LiZ500 (ThermoFisher Scientific, Waltham, MA, USA) and call alleles for each of the 11 loci in the STR panel (Miller et al. 2019). In general, we used the Microsatellites plugin in Geneious Prime to automate binning and proposed bin ranges for each allele, adjusting bins based on histograms instead of expected repeat motifs to account for potential mutations (Miller et al. 2019). Allele calls were manually edited and verified for all samples (Flores-Rentería & Krohn, 2013; Pálsson et al., 1999). As a qualitative estimate of how DNA extraction method and storage time impact ability to obtain microsatellite data from fecal DNA, the number of assigned loci was averaged across replicates for an average assigned loci per sample. Any samples that produced no peaks were assigned “NA " and excluded from statistical analyses to distinguish from samples whose peaks could not be called as alleles.

3.2.8 Statistical Analysis

Our first objective was to evaluate the impact of storage time, DNA extraction kit, and target locus on diet fecal metabarcoding for plant DNA. All statistical analyses for metabarcoding data were conducted in the R statistical environment (R version 4.1.3). Due to differences in the inherent resolution of genetic markers and gaps in reference databases, we used ASV count average across all three replicates per sample in combination with taxonomic resolution information to minimize comparison bias (Ando et al., 2020; Gold et al., 2021). For every sample the number of raw ASVs and the number of ASVs with an average > 10 reads, ASV10, were summarized as a measure of detection breadth and efficiency. As it is common for multiple ASVs to represent the same biological species (Joos et al., 2020; Schloss, 2021), we also calculated a response variable representing the number of unique genera with > 10 average reads, UniqGen10, for each sample to assess taxonomic detection diversity. Each potential explanatory variable impacting plant ASV detection had three qualitative factor levels: Short (1-2 yrs), Mid (7-10 yrs), and Long (16 yrs) for storage time, Qiagen (Q), Nucleospin (N), and Zymo (Z) for extraction kit, and trnL (tr), rbcL (rb) and ITS2 (it) for target locus. Initial exploration of differences between each factor level was done with repeated measures ANOVA tests in R with package *rstatix* with significance testing employing Bonferroni multiple-testing corrections. We used generalized mixed effect modeling to evaluate the importance of each explanatory variable for predicting plant ASV10 detection and UniqGen10. To deal with over-dispersed positively skewed count data, the glmer.nb() function from the MASS package in R was implemented instead of the standard lmer4. All tested models had ASV10 or UniqGen10 as the response variable and fecal sample ID as a random effect. The three explanatory variables of DNA extraction kit, storage time class, and target locus were differentially combined in a total of nine models, including null and full interacted models. Each model was assessed for residual normality, collinearity, homoscedasticity, and overdispersion using R-packages “performance” and “AICcmodavg”. Model selection was conducted by comparing AICc values for the most parsimonious model, and the final model was compared to a null model with an ANOVA to determine significance ($p<0.05$).

Our second objective was to assess how the same methodological differences of DNA extraction and storage time impact microbial community detection with fecal metabarcoding from the same extraction protocol. Using the same approach described for the plant-specific amplicons, generalized mixed effect modeling of microbial ASV10 (ASVs with > 10 average reads) was implemented with fecal sample ID as a random effect. We used the "phyloseq" r-package as described in the Dada2 tutorial documentation (License: CC-BY 4.0) to examine the microbial metabarcoding data with a variety of methods. We first filtered for the 20 most abundant sequences and constructed relative abundance histograms of bacterial families grouped by kit and storage time to assess general community differences, and then calculated alpha diversity measures and compared both factors with an ANOVA test. We used non-metric multidimensional scaling (NMDS) with a Bray-Curtis algorithm to assess community differences separately for extraction kit and storage time and determined statistical significance of variances using multiple response permutation procedure (MRPP) implemented with *vegan* in RStudio (R package version 2.6-4). Finally, we generated and compared generalized mixed effect models predicting microbial ASV detection in R as detailed for the plant metabarcoding data to determine significance of kit and storage as predictor variables.

Our third and final objective was to quantity the host component of the fecal samples. For the mitochondrial sequencing, univariate and multivariate mixed effect models with DNA extraction kit and storage time as predictors of sequencing quality were produced in the R statistical environment with lme4 package, with final model selection and comparison of model fit conducted with the same workflow described for the plant and microbial analyses. We then conducted one-way repeated measures ANOVAs with Bonferroni correction testing for all factor combinations of the predictor variables with a 0.05 significance threshold and used the multcomp and glht R-packages (R version 4.3.2) to generate coefficient estimates for each comparison. For microsatellites, generalized mixed effect model selection and one-way repeated measures ANOVA significance testing for the same predictor variables were also performed. Grouped bar plots comparing both mitochondrial and microsatellite host assessment metrics for DNA extraction kit and storage time were graphed using "ggplot2" in R.

3.3 Results

3.3.1 Illumina MiSeq Sequencing

Illumina MiSeq sequencing was successful for all libraries at sufficient sequencing depth for downstream analyses from noninvasive samples. The 5 metabarcoding libraries yielded a total of 22,432,216 reads for the v3 kit (4 libraries) and 6,277,546 reads for the v2 (trnL only) with 79% and 85% passing instrument quality filtration respectively. After additional quality (Q>30) and length (locus-specific) filtering, the trnL amplified samples had the lowest sequencing success with 11 samples (24%) with less than 10,000 high-quality reads across three replicates, a conservative cutoff for diet resolution (Dully et al., 2021). The average read depth per sample was 13,917, 45,003, 36,964, and 21,751 for trnL, rbcL, ITS2, and 16S rRNA v3-v4 respectively. The reduction in trnL read depth is

likely due to only 11% of third replicates surpassing 1000 reads per sample, compared to 94% and 74% for the first and second replicates. This pattern was not observed in any of the other target loci, suggesting a PCR thermocycling or library prep error for that reaction plate, and had no impact on downstream analysis methods for ASV counts.

Table 3.1. Summary of plant ASVs identified for each target plant locus, including broad taxonomic group assignment. Percentages of ASVs resolved to taxonomic levels of order, family, genus, and species as an estimate of resolution power.

	trnL	rbcL	ITS2
	ASV Counts		
Total	75	184	1421
Plant	74	184	286
Fungi	0	0	72
Other	0	0	506
Unidentified	1	0	557
	Taxonomic Resolution		
to Order	69.3% | 52 ASVs	96.1% | 177 ASVs	100% | 286 ASVs
to Family	69.3% | 52 ASVs	84.8% | 156 ASVs	100% | 286 ASVs
to Genus	58.7% | 44 ASVs	70.7% | 130 ASVs	94.8% | 271 ASVs
to Species	50.6% | 38 ASVs	40.8% | 75 ASVs	59.7% | 171 ASVs

3.3.2 Plant and Microbial Metabarcoding

We observed high variability in both ASV, or unique sequence, counts and capacity to assign fine-scale taxonomy between our three diet barcodes. The DADA2 pipeline produced the following number of ASVs per plant locus: 75 (trnL), 184 (rbcL), and 1421 (ITS2) (Table 3.1). After assigning taxonomy and filtering for Kingdom Viridiplantae or Streptophyta to exclude any non-plant sequences, 74 (99%) of the trnL ASVs compared to 184 (100%) for rbcL and 286 (20%) for ITS2 were classified as plants. The number of ASVs unable to be taxonomically resolved and considered off-target or contamination varied per locus, with 0.01% for trnL, 0% for rbcL, and 39.2% for ITS2. In addition to those assigned as plants, 72 (5.1%) and 506 (35.6%) of the ITS2 ASVs were assigned as fungi and general eukaryotes respectively (Table 3.1). We also assessed the resolution power of each target locus by quantifying resolution as each taxonomic level beyond Order. The nuclear ITS2 locus has the largest percentage of ASVs resolved to the species level (59.7%), followed by trnL (50.6%) and then trnL (40.8%), although given

differences in total ASVs resolved nearly double the total number of ASVs to species (Table 3.1). The vast majority of ITS2 ASVs were resolved to at least genus (94.8%), with rbcL (70.7% to genus) and ITS2 having similar resolution at the family level. Overall, resolution power was lowest for the trnL locus, with lowest total ASVs (69.3%) with taxonomy assigned to order.

We also compared the plant families and genera represented across all samples between the three loci to determine what degree of taxonomic overlap exists for diet item detection. We found 16 plant families identified by all loci, but each locus also represented 6-10 unique families not detected by either of the other two loci (Figure 3.3). This pattern continued at the genus level, with 16 genera concurrent across all loci but 22 genera identified only by the ITS2 locus compared to 18 each for rbcL and trnL. Although our objectives did not include a complete assessment of winter diet in Upper Peninsula deer, a subset of potential diet items such as lycophytes and eastern hemlock that were identified only by specific plant primers are highlighted in Figure 3.3. We did not assess overlap at the species level given the vast variability of resolution power and inherent biases of reference databases, but information on specific families and genera identified by each locus is provided in the Chapter 6 of this thesis (Table 6.4).

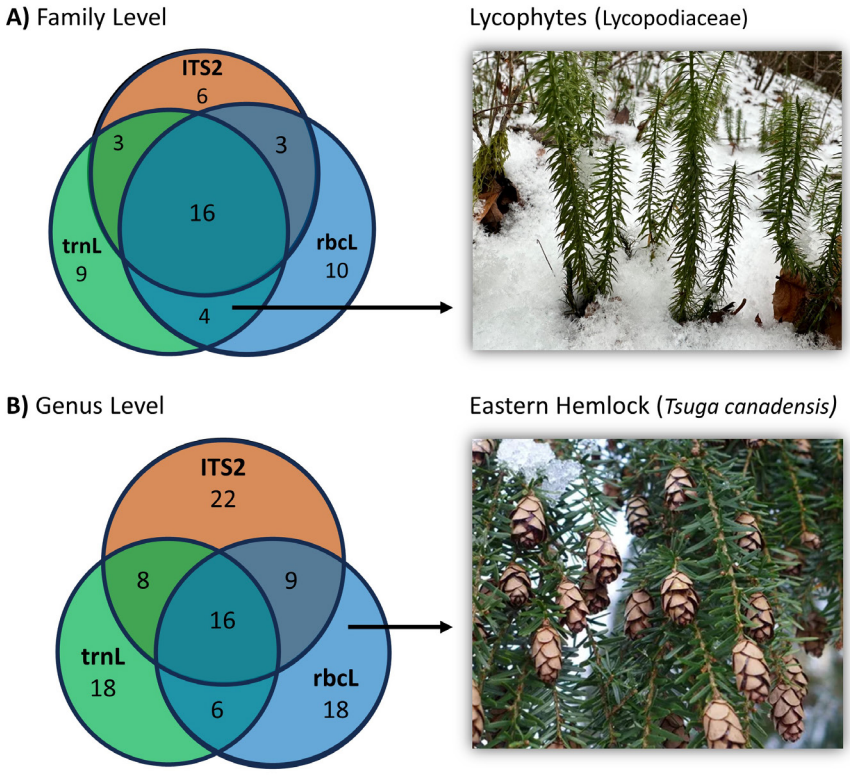


Figure 3.3. Venn diagrams overlap between the number of distinct (a) families and (b) genera identified by the three universal plant primers. The two pictures represent two important winter diet items for white-tailed deer and the primers that detected them within the fecal samples. Photographs taken by Melanie Ottino.

We detected evidence of strong PCR bias for trnL and ITS2 for conifers and dicot flowering plants respectively. When assessing the percentage of sequencing reads and ASVs assigned to different plant classes and orders, we found the majority of ASVs assigned to Magnoliopsida, followed by Pinopsida and Liliopsida across all three plant loci (Figure 3.4). The rbcL chloroplast marker detected the greatest amount of taxonomic diversity including ASVs assigned to Bryopsida, Lycopodiopsida, and Sphagnopsida, classes also resolved by trnL but not ITS2 (Figure 3.4). Regarding potential amplification or sequencing bias, we found significant differences in read percentages assigned to major plant classes. Notably, 72.1% of trnL reads were assigned to Pinopsida compared to 34.4% and 1% for rbcL and ITS2 respectively (Figure 3.4). Conversely, nearly all processed plant sequencing reads from ITS2 were designated as Magnoliopsida (Figure 3.4).

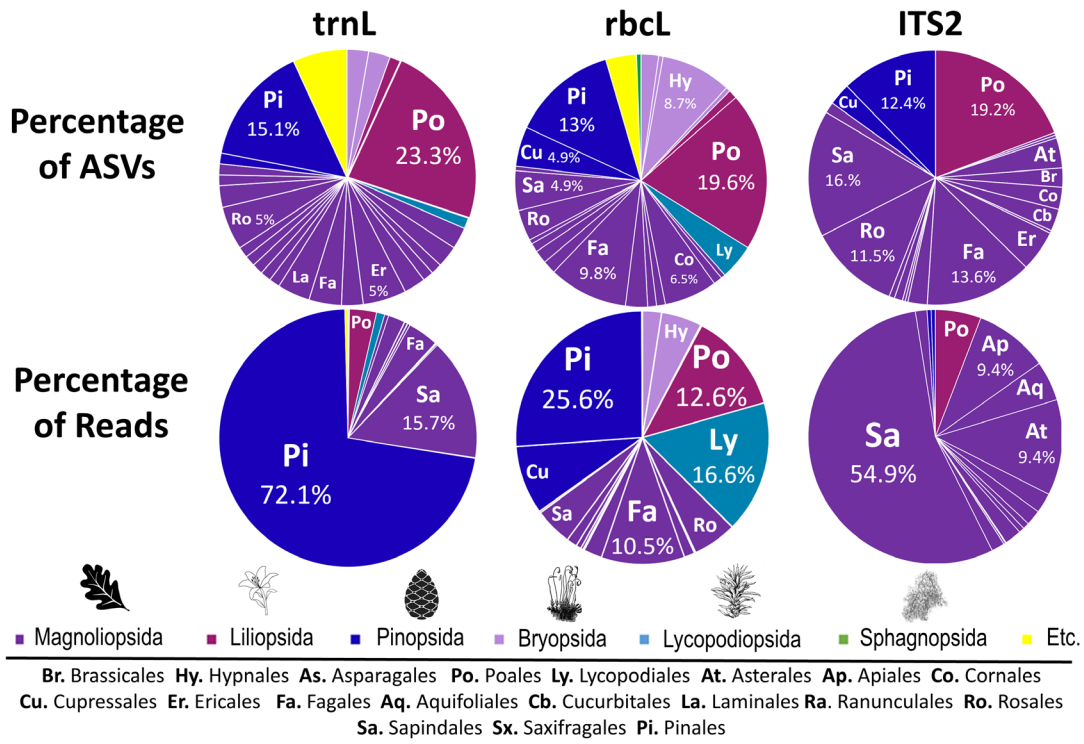


Figure 3.4. Percentage of Amplicon Sequence Variants (ASVs) and sequencing reads taxonomically assigned to major plant classes and orders. Plant classes are designated by color with orders representing more than 5% of total labelled with a two-letter code. Codes for each order are listed in the legend.

For our microbial sequencing library targeting the 16S rRNA V3-V4 region, 461 ASVs were identified post-DADA2 processing and alignment to the curated UNITE reference database. All ASVs were classified as belonging to phylum Bacteria, and 453 (98%) taxonomically resolved to at least order. A total of 61 different bacterial families were represented across the 15 fecal samples, with 351 (76%) of ASVs belonging to the following four families: Sphingobacteriaceae, Pseudomonadaceae, Nitrosomonadaceae, and Flavobacteriaceae. We generated relative abundance histograms with the top 20 most

abundant ASVs comparing proportions of each family grouped by both DNA extraction kit and storage class (Figure 3.7). Qualitative assessment indicated that short-term storage samples had higher relative abundance of Rhizobiaceae bacteria compared to those stored longer and a more balanced representation of the top 5 most abundant families (Figure 3.7). Additionally, no major variation in bacterial family representation was detected between the different extraction kits (Figure 3.7).

3.3.3 Metabarcoding Statistical Results

For our first objective, qualitative comparisons of ASV10 grouped by DNA extraction kit and target locus indicated that kit has a relatively small effect on ASV10 compared to locus, with the Zymo kit performing slightly better for all loci and trnL detecting noticeably fewer ASV10s on average than the longer targets (Figure 3.5). Repeated measures ANOVA comparing all factor level combinations for extraction kit type, time class, and target locus found statistically significant differences between all target plant loci ($p_{adj} <<< 0.05$) but only Zymo vs. TakaraBio ($p_{adj} = 0.026$) for extraction kits and long-term vs. mid-term ($p_{adj} =0.005$) for storage time class. Qualitative comparisons of ASV10 grouped by DNA extraction kit and target locus indicated that kit has a relatively small effect on ASV10 compared to locus, with the Zymo kit performing slightly better for all loci and trnL detecting noticeably fewer ASV10s on average than the longer targets (Figure 3.5). The most parsimonious model predicting plant ASV10 for our dataset included extraction kit and target locus but excluded storage (AICc = 909.54, K=7) and had a relatively high marginal and conditional R^2 (R^2_m= 0.55, R^2_c=0.66) (Table 3.2). Incidence rate ratios of independent variables for our selected mixed effect model were statistically significant for the ITS2 and rbcL loci compared to trnL ($p < 0.001$), but only nearly significantly for the plant-specific NucleoSpin II extraction kit compared to the Zymo Fecal/Soil kit (p=0.055).

Table 3.2. Ranked AIC table of negative binomial mixed effect models predicting plant ASVs (minimum reads >10). The top model is bolded.

Model	k	AICc	ΔAICc	χ^2	p-value
Locus + Kit + Storage	9	908.3	0	107.1***	< 0.001
Locus + Kit	**7**	**909.5**	**1.19**	**101.3*** **	**< 0.001**
Locus + Storage	7	915.0	6.65	95.9***	< 0.001
Locus	5	916.1	7.71	90.4***	< 0.001
Locus + Kit + Storage + Interacted (All)	13	917.0	8.64	108.0***	< 0.001
Null	3	1002.2	93.8	-	-

Our second objective was to evaluate the impact of kit type and storage time on microbial metabarcoding outcomes. We found that the most complex model including an interaction term between kit and storage time had the lowest AICc (AICc=402.98, K=11)

and a delta AICc > 2 for the next best model (Table 3.3). However, an ANOVA between our best model and the null was statistically insignificant ($\chi 2$ =13.635, p=0.125). We also evaluated differences of alpha diversity between kit types and storage classes using ANOVA and found no statistically significant results (p >0.05), suggesting that kit and storage do not impact the total number of ASVs detected via microbial fecal metabarcoding. Our assessment of beta diversity using nonmetric multidimensional scaling (NMDS) resulted in a 2-dimension solution with 0.22 final stress value. Ordination plots coded by factor type showed little separation in ordination space related to DNA extraction kit type but clear community differences for storage time (Figure 5.1). We confirmed this with Multiple Response Permutation Procedure (MRPP) tests resulting statistical significance for storage class (p=0.001) but non-significance for kit type (p=0.99), indicating that there is detectable bacterial community dissimilarity between fecal samples of variable storage age.

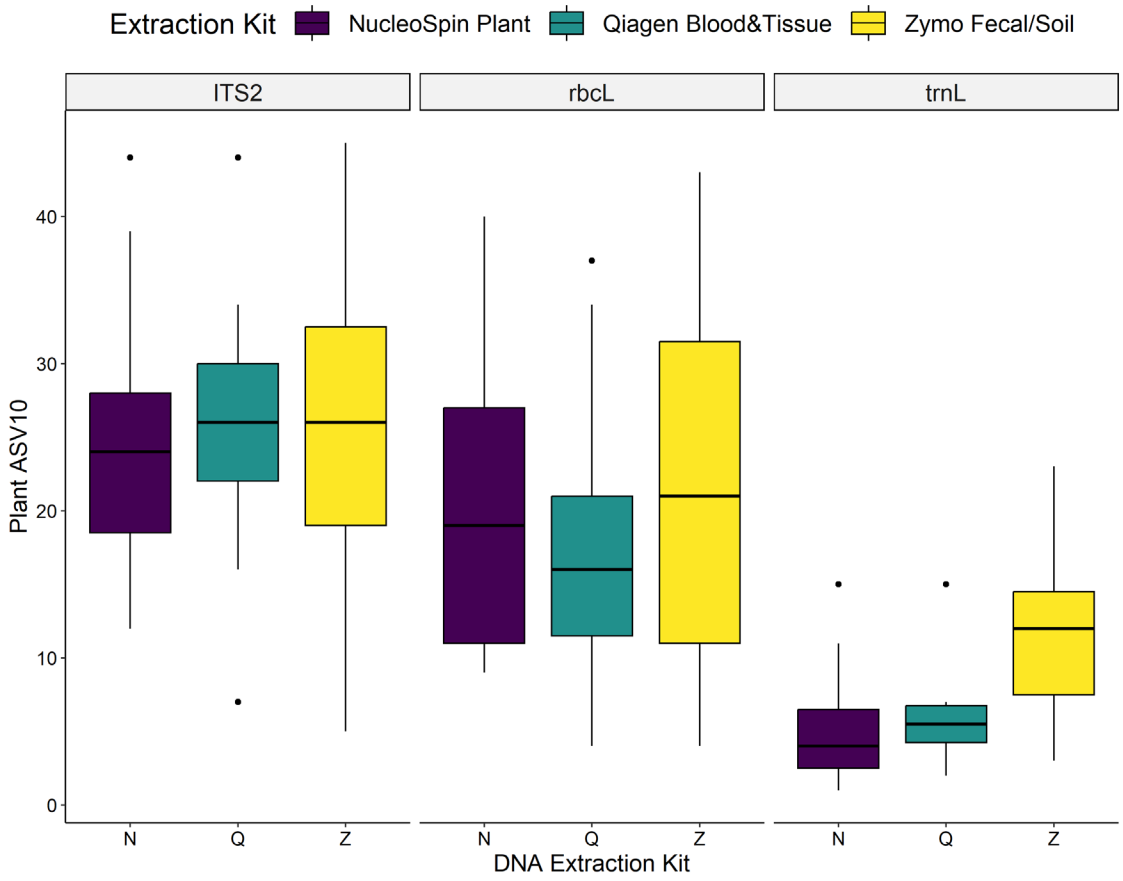


Figure 3.5. Grouped boxplot demonstrating impact of target loci and DNA extraction kit on number of plant ASVs identified with fecal metabarcoding. Plots are separated by target locus (ITS2, rbcL, and trnL) and colored by extraction kit (N= TakaraBio NucleoSpin Plant II, Q= Qiagen DNeasy Blood & Tissue, Z= Zymo Research Fecal/Soil Miniprep). Outliers per group are represented as dots and error bars to the first and third quantiles are included.

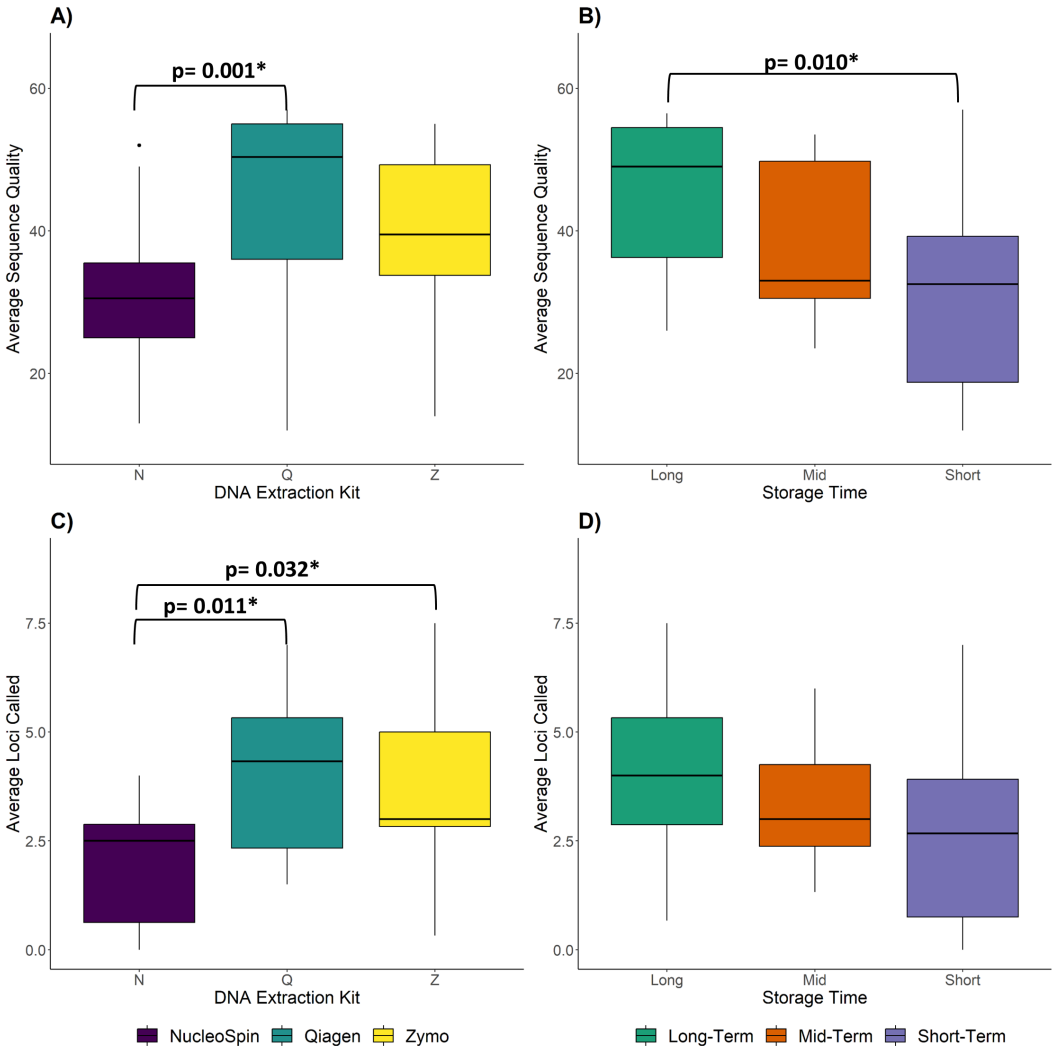


Figure 3.6. Boxplot summary of host DNA sequencing for test extraction kits and storage classes. The two different measures of host quality are (A-B) average quality score of fecal host DNA 16S dloop mtDNA Sanger Sequencing and (C- D) average loci called for the 11 loci STR microsatellite panel. Each measure is grouped by (A,C) DNA extraction kit and (B,D) storage time. Significant values ($p<0.05$) from pairwise t-tests with Bonferroni multiple-testing corrections are bolded and designated with a '*'.

3.3.4 Host Mitochondrial and Microsatellite Results

For the third objective, Sanger sequencing of the 45 purified fecal DNA samples in duplicate (n=90) resulted in 23 samples classified as either "poor quality" or "no priming" due to lower quality scores (QS <24). In total, 6 of the 45 samples had both replicates fail this minimum quality threshold. However, even low-quality sequences can be used to investigate sample taxonomy if informative amplicon regions are present (Azenta Life Sciences 2021). In total 15 samples with QS<15 found no significant similarity matches, but the remaining 76 sequences with QS ranging from 12-52 had 3-10 top hits matching genus Odocoileus ($P_{ID}>0.85$). One-way repeated measures ANOVA

and post-hoc Tukey testing of univariate mixed effect models predicting average sequence quality with extraction kit and storage time determined statistically significant differences for both factor types (Figure 3.6, A-B). The significant relationships identified were that the Qiagen kit increased average quality score compared to the TakaraBio kit (coefficient= 13.72, p=0.001) while samples stored for shorter amounts of time had decreased quality scores (coefficient= -13.12, p=0.01). The best-fit model for mtDNA Sanger Sequencing average quality included both kit and storage time as predictor variables (AICc=335.75) and significantly improved prediction power compared to the null model ($\chi 2$ =19.24, p<<<0.001).

Table 3.3. Ranked AIC table of negative binomial mixed effect models for microbial community detection and host DNA quality (mtDNA, msat) components of ungulate fecal samples. Number of parameters (k), adjusted AIC, and ΔAIC are listed for each model along with chi-squared and p-value for comparison to null model. The best model for each type is bolded if applicable.

Type	Model	k	AICc	ΔAICc	χ^2	p-value
Microbiome	Full + Interaction	11	402.98	0	12.64	0.125
	Storage + Kit	7	429.35	26.37	4.52	0.340
	Storage	5	439.05	36.07	2.24	0.326
	Kit	5	441.10	38.12	2.28	0.320
	Null	3	451.33	48.35	-	-
Mitochondrial	**Storage + Kit**	7	335.75	0	19.24	0.001
	Kit	5	347.18	11.43	10.98	0.004
	Storage	5	350.01	14.26	7.94	0.019
	Null	3	361.97	26.22	-	-
Microsatellite	**Kit**	5	157.3	0	8.17	0.017
	Storage + Kit	7	157.7	0.4	9.24	0.056
	Storage	5	162.4	5.14	0.89	0.640
	Null	3	162.6	5.31	-	-

For the microsatellite data, only 1 of the 38 included fecal samples had no electropherogram peaks in any of the replicates, with an additional 15 having fewer than 3 loci called on average. All three positive controls could only be confirmed at 10 of the 11 loci included on the STR panel, so any peak information from the dropout locus was disregarded. Although means differed slightly between variably aged fecal samples, the repeated measures ANOVA found no significant comparative difference between storage classes (Figure 3.6, D). However, we found statistically significant differences between NucleoSpin and both of the other kits with lower numbers of average called loci

compared to Zymo (p=0.032) and Qiagen (p=0.011) (Figure 3.6, C). By comparing AICc values and parameter numbers, we selected the mixed effect model with only kit as the best model for the fecal microsatellite data and excluded storage time (Table 3.3). We compared our selected model to the null model with an ANOVA and found a decreased but still statistically significant improvement for average called loci prediction compared to the mitochondrial models for host DNA ($\chi 2$ =8.85, p=0.012).

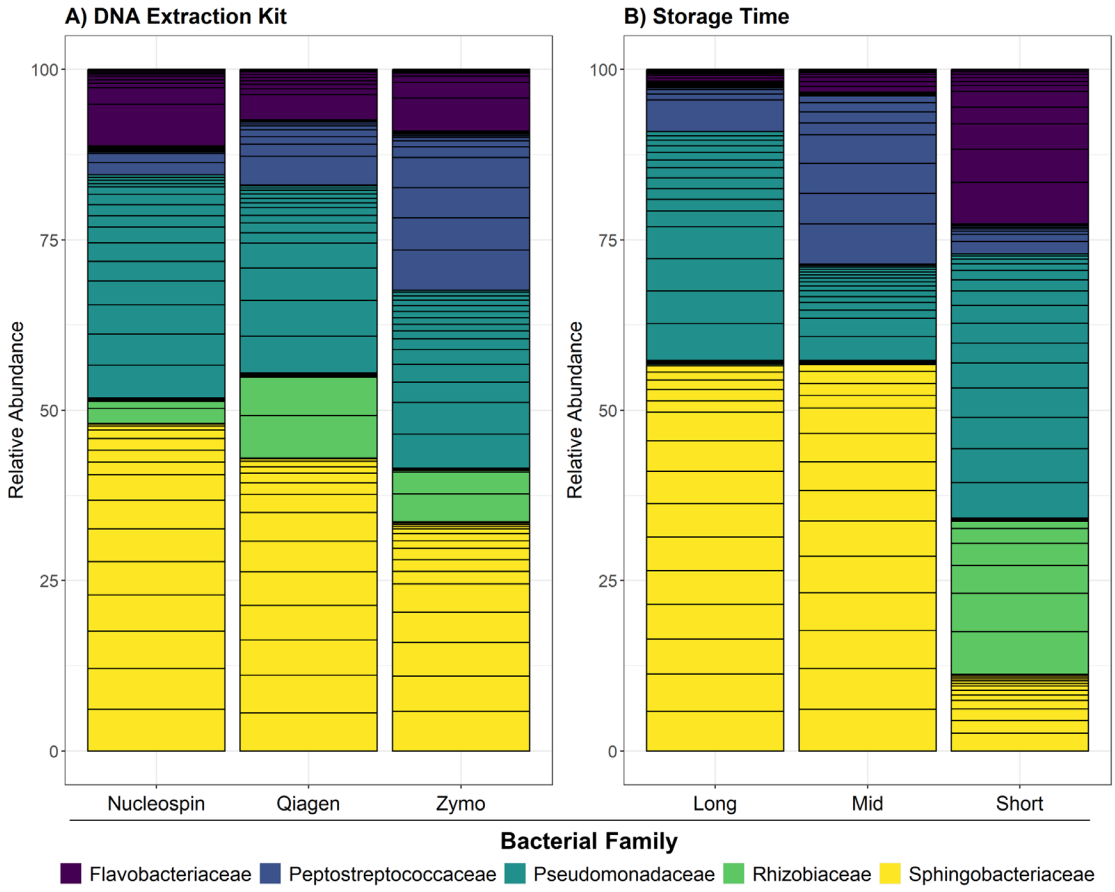


Figure 3.7. Relative abundance histogram with top 20 most-common bacterial families identified by 16S rRNA v3-v4 sequencing from ungulate feces. Data is grouped by (A) DNA extraction kit and (B) storage time.

3.3.5 Methodological Comparisons

A rank table generated to compare different methodological factor performance for diet, host, and microbial sequencing found the highest general success with the Zymo Research Fecal/Soil Miniprep kit, the mini-barcoded rbcL target plant locus, and shorter storage time (Table 2.4). Associated costs were calculated assuming 100 samples and found that combination is also the most expensive option, with the Qiagen DNeasy Blood & Tissue kit and trnL amplification and sequencing being the most economical choices.

Additionally, the highest performing DNA extraction kit for host-specific analyses was Qiagen DNeasy Blood and Tissue.

Table 3.4. Rank table for different DNA extraction kits, storage lengths, plant barcodes, and target taxa. Cost per sample (USD) is calculated for n=100 samples with available pricing from each company as of August 2023.

	Cost/Sample*	*Plant ASV*	*Plant Genera*	Plant (Total)	Microbial	Host	Summed Rank
DNA Extraction Kit							
Qiagen Blood & Tissue	$3.50	*1*	*1*	2	1	3	**6**
NucleoSpin Plant II	$3.42	*1*	*1*	2	1	1	**4**
Zymo Fecal/Soil	$4.99	*2*	*2*	4	2	2	**8**
Target Locus (Plant)							
trnL	$2.17	*1*	*2*	3	-	-	**3**
rbcL	$7.55	*2*	*3*	5	-	-	**5**
ITS2	$7.55	*3*	*1*	4	-	-	**4**
Storage Time							
Long	-	*1*	*1*	2	1	1	**4**
Mid	-	*1*	*1*	2	1	1	**4**
Short	-	*1*	*1*	2	2	1	**5**

3.4 Discussion

3.4.1 Storage Method, Storage Time, and Sample Quality

These results highlight how environmental DNA sample collection and storage can introduce error into molecular analyses, especially given the elevated contamination risks (Ando et al., 2020). The deer feces used in this study were collected in the field in early-spring conditions directly following snowmelt with average daily temperatures around 50.2 °F (data from 1991-2020) and immediately transferred to a -20°C freezer for long-term storage (National Weather Service). Freezing at -20 to -80 °C is considered best-practice for eDNA sample storage, but the potential for freezer temperature fluctuations, accidental thawing, and machine malfunctions with samples stored for 1-16 years could result in increased DNA degradation or fungal growth on older samples (Seeber et al., 2022; Wietz et al., 2022). Our results largely showed the opposite trend of what we anticipated, in that older samples with longer storage times performed better in ASV detection for both total number of plant and microbial ASVs and host DNA sequencing. One possible explanation is related to the freeze-thawing impacts on PCR inhibitors and cell lysis efficiency. Although shearing of large fragments of genomic DNA due to freeze-thaw cycles is well-characterized (Ross et al., 1990; Shao et al., 2012), multiple studies have found that repeated freeze-thaw cycles increased bacterial eDNA yield due to more complete lysis and potential inhibition of PCR inhibitors immediately following freezing (Sluter et al., 1997; Xin et al., 2021). Additionally, there is evidence that freeze-thaw cycles alter the physical properties of polysaccharides, one of the main PCR

inhibitors found in herbivore fecal samples, which may decrease interference with DNA amplification (Acharya et al., 2017; Monteiro et al., 1997; Wang et al., 2023).

However, a more likely explanation is that storage time acts as a proxy for sample quality and the trends observed are reflecting inherent fecal sample quality produced by year-specific variables such as winter temperature stability, deposition to collection time, and deposition location (Ruppert et al., 2019). The noninvasive collection methodology utilized in this study is such that fecal sample deposition time is both unknown and difficult to accurately estimate based of visual characteristics and includes a critical assumption that samples remain consistently frozen throughout the winter season, potentially compromising best-practice recommendations for field collection like limited environmental exposure and time from deposition (Ando et al., 2020; Jones et al., 2021). The Upper Midwest region, which historically has winters characterized by a large snowpack and cold temperatures, has experienced increasingly erratic weather patterns over the past decade with unseasonably warm periods through the winter season (Demaria et al., 2016). Although our sample sizes were not selected for a large-scale temporal comparison(Pryor, 2014), these analyses suggest that warmer winters negatively impact metabarcoding results. It is critical for molecular ecology studies to mitigate these potential biases in field design as much as possible, although with appropriate models and climate information, mathematically controlling for random annual climate and collection variability for long-term metabarcoding studies may be sufficient for limiting error introduced by sample quality (Ando et al., 2020; Ruppert et al., 2019). However, the only GLMM models that included storage time as a statistically significant predictor were for microbial and host mitochondrial DNA components (Table 2.3), and the effect size of these were very small.

In contrast to the trend that older samples performed better, we evidenced that samples stored for shorter periods had higher amounts of environmental and soil-associated bacteria, namely from Flavobacteriaceae and Rhizobiaceae (Alves et al., 2014). Fecal sampling protocols minimizing time from deposition and contact with wet soil have been shown to decrease environmental contamination ((Ando et al., 2018), but our results support previous findings short term storage time and temperature can also noticeably alter community composition and structure (Lauber et al., 2010; Rubin et al., 2013). Overall, our results validate fecal metabarcoding for variably-aged frozen fecal samples and suggest that storage time has little to no impact on ASV detection for herbivore diet studies but may remain relevant for microbial metabarcoding studies. However, measures should be taken to minimize variability in sample quality resulting from field conditions, whether by methodological design or robust data collection enabling more accurate modeling.

3.4.2 DNA Extraction Kit

We also found that DNA extraction kit was a significant factor for both diet metabarcoding and host DNA analyses from our ungulate fecal samples. In contrast with a study comparing 8 extraction kits that found NucleoSpin Plant II performed the best for both host and herbivorous diet detection (Galan et al., 2018), our results showed little to

no statistically significant difference in the number of plant ASVs or genera identified by any of the three commercial kits we compared (Figure 3.4, Figure 3.6, Table 3.1). There was a slight indication that the eDNA optimized Zymo Fecal/Soil kit performed marginally better for all plant loci, with particularly pronounced improvements for the shorter length trnL amplicon. This was unexpected since longer fragments are usually more difficult to amplify in the presence of PCR inhibitors (Sidstedt et al., 2020). A possible explanation for this is that although shorter fragments amplify more easily, PCR inhibition may increase the likelihood that a small subset of shorter fragments will swamp out broad amplification (Krehenwinkel et al., 2017), as evidenced by equivalent raw reads but reduced ASV counts between the three kits. Additionally, PCR inhibitors may decrease the universal taxa-specific primer binding efficiency and exacerbate inherent preferential annealing biases due to sequence divergence that will differentially impact different genomic regions (Stadhouders et al., 2010), and this effect may be more pronounced with smaller fragments. Moreover, despite our study quantifying an herbivorous diet, the plant-tissue optimized kit performed least optimally for both ITS2 and trnL loci. Cell digestion in the rumen reduces plant cell wall particle size and lignin content (Flint & Bayer, 2008; Smith et al., 1983), which may reduce any benefit from additional lysis steps like RNase A treatment in plant kits for ungulates compared to other mammals.

For the deer host DNA component, the Qiagen DNeasy Blood & Tissue kit performed best with improved average sequence quality of mitochondrial amplicons and increased number of microsatellite loci genotyped compared to the other two kits, particularly the NucleoSpin Plant II kit. Host DNA within feces is generally <1% of total fecal DNA (He et al., 2019), and the additional enzymes present in plant, fecal, and soil specific kits for digesting cell walls and removing additional contaminants likely reduces successful isolation and capture of the degraded and low-quantity genetic fragments (Manen et al., 2005). Moreover, additional filtration steps to remove PCR inhibitors or contaminants common in plant and fecal/soil optimized kits will reduce overall DNA yield, and this will especially impact low-quantity host DNA yield (Katevatis et al., 2017). Similar to the plant component, our results evidence that the extraction kit did not significantly impact bacterial sequencing, although the Zymo kit produced slightly more balanced detection of diverse bacterial taxonomic families (Figure 2.7). Although differences between extraction kits had only minor differences in sequencing output and taxonomic resolution, our findings suggest that choice of DNA extraction method for feces should be considered within the framework of study objective.

3.4.3 Impact of Target Locus Selection

Target locus was the most significant factor for ASV detection and taxonomic resolution power for the diet component of our fecal metabarcoding study. The most commonly utilized plant metabarcoding universal primer set for herbivorous diet studies targets the shorter trnL P6 loop locus due to its shorter amplicon length increasing amplification success for degraded eDNA samples like feces or soil (Abdullah-Fauzi et al., 2022; Kartzinel et al., 2015; Robeson II et al., 2018). In contrast to studies such as Mallott et al. 2018 that found trnL outperforming longer amplicons like rbcL and ITS2 in both

accuracy and taxonomic resolution, our results demonstrate that despite the longer amplicon length, the minibarcoded rbcL and ITS2 primers did not have decreased amplification or sequencing success but actually detected significantly more ASVs and resolved a greater diversity of plant families and genera. Discrepancies across the literature may be due to a number of contributing factors, including host species, study environment, and available reference databases. Interspecies physical, chemical, and compositional differences of fecal matter, particularly from separate taxonomic groups, has been suggested to affect rates of decomposition (Agetsuma-Yanagihara et al., 2017; Jung & Kukka, 2016), ease of DNA extraction (Hart et al., 2015), and quality differences of nuclear and organellar genetic material (Ernest et al., 2000; Kovach et al., 2003). Although our study found that longer target amplicons from both chloroplast and nuclear DNA provided superior diet information, this result may not be directly applicable to other studies that do not focus on ungulate ruminant host species, winter-preserved samples, and boreal forest habitat.

Inherent genetic diversity is also critical for locus selection, as interspecies variation at different conserved sites can make unique diet item detection difficult. When information about potential diet items is known beforehand, initial screening sequence variation at the site to predict taxonomic resolution power given available genetic resources can aid in locus selection that minimizes primer bias (Piñol et al., 2019). Primer bias is particularly important for fecal metabarcoding studies where detection of specific or rare species such as endangered or invasive diet items (Walker et al., 2022; Westfall et al., 2020) or disease-associated gut parasites (Davey et al., 2023) is a primary objective. Although the ITS2 primer pair generated the most dietary ASVs, we found significant potential amplification bias in both sequencing read and ASV percentages. Less than 5% of reads assigned as conifers compared to 55% to Sapindales, a class representative of maples in Upper Midwest region. This indicates potential sequence divergence at the ITS2 genomic binding site between Gymnosperms (Pinopsida) and Angiosperms (Magnoliopsida) with the UniF/R primer pair preferentially binding and amplifying Angiosperm sequences, possibly due to the development and testing of this primer pair being heavily weighted towards Magnoliopsida species (Moorhouse-Gann et al, 2018). However, this apparent primer bias may also reflect differences in nuclear plant DNA quality in the deer feces included in this study between coniferous and deciduous species which are known to differ in leaf structure and cell wall thickness along with other physicochemical properties (Côté, 1968; Donaldson et al., 2018). Furthermore, the plant fecal DNA in our samples represents the winter diet meaning conifer species are sourced from leaf tissue whereas deciduous are from twig or buds. Supporting suggestions that selecting plant metabarcoding primers targeting chloroplast loci due to higher abundance and reduced degradation, both rbcL and trnL detected a more taxonomically diverse distribution of ASVs in the fecal samples, although only rbcL showed little primer bias with evenly-distributed sequencing reads. Importantly, rbcL also identified key winter diet items hypothesized to be present, including eastern hemlock (*Tsuga canadensis*), which dominated the study sites, and Lycophytes, which deer could forage by digging, a behavior documented in other studies in up to 12 inches of snowpack (Dorn, 1992.; Rogers, 1981). The notable differences in taxonomic resolution power of the three tested plant universal primer sets suggest careful consideration of target locus needs to be

tailored to specific research goals. Moreover, as suggested by previous studies, our findings provide further evidence that economical and robust diet detection requires a multi-locus approach given that each primer set identified unique plant families and genera (Gillet et al., 2015; Zeale et al., 2011).

The benefit of multiple barcoding markers is even more apparent given gaps in taxonomic coverage in publicly available reference databases for even commonly used metabarcoding markers like those utilized here. Studies like Erickson et al. 2017 have attempted to bridge this gap by sampling and sequencing potential prey items for a customized local database, but these approaches demand substantial effort, time, and financial resources and are therefore unfeasible as a universally viable solution. Increasing availability of user-friendly bioinformatic tools such as CRABS (Creating Reference Databases for Amplicon-Based Sequencing) to curate, train, and filter custom reference databases using public sequence repositories like GenBank, EMBL, and BOLD are increasing diversity incorporated into reference databases, but lack of geographic or ecological information on published sequences thwarts improved localized taxonomic assignment (Jeunen et al., 2023; Mugnai et al., 2023). While our study design does not include a positive control for diet detection and taxonomic assignment accuracy comparisons due to the noninvasive nature of eDNA sampling, manual examination of assignments using regional flora surveys, field notes, and citizen-science reports can provide qualitative measurements of accuracy. However, this highlights the need for quantifiable and structured validation of taxonomy assignments across fecal metabarcoding studies to minimize false positives in diet descriptions that may have implications for forest and wildlife management.

3.4.4 Optimal Fecal DNA Metabarcoding Methodology

A key objective of this study was to evaluate not just how different methodological choices impact molecular analyses, but if a single-extraction protocol from feces could be used for downstream analyses for diet, host, and microbial communities. Although differences related to focal species, diet type, ecosystem, and lab resources will factor into the optimal methodological design, our results support use of a single-extraction method, with both the Qiagen DNeasy Blood & Tissue and the Zymo Fecal/Soil Miniprep kits performing equivalently for host and diet analyses in particular. However, the higher cost per sample (Table 3.4) combined with higher number of filtration steps in the eDNA optimized kit may be critical considerations for researchers with substantial time and economic limitations.

To conclude, fecal DNA metabarcoding on variably-aged winter ungulate fecal samples is a valid molecular ecology tool for diet detection on a finer taxonomic resolution scale. We tested three widely available commercial DNA extraction kits optimized for different tissue types and found that the Zymo Research Fecal/Soil Miniprep kit detected more ASVs and unique plant genera, although Qiagen DNeasy Blood & Tissue more successfully captured host DNA without compromising significantly on diet detection. Most importantly, our results demonstrate the importance of locus selection and provide evidence that universal primers with longer expected amplicon lengths can be

successfully leveraged for more genus and species level identification of diet items. Our assay suggests that for single extraction analyses of diet, host, and microbial DNA, the Zymo Fecal/Soil kit and the mini-barcoded rbcL primer set were most effective, but variable funding limitations, sample type, and research objectives must be considered when selecting and finalizing a fecal metabarcoding protocol.

3.5 References

16S-metagenomics-faq-1270-2014-003.pdf. (n.d.). Retrieved November 1, 2023, from https://support.illumina.com/content/dam/illumina-marketing/documents/products/other/16s-metagenomics-faq-1270-2014-003.pdf

Abdullah-Fauzi, N. A. F., Karuppannan, K. V., Mohd-Radzi, N. H. S., Gani, M., Mohd-Ridwan, A. R., Othman, N., Haris, H., Sariyati, N. H., Aifat, N. R., Abdul-Latiff, M. A. B., Abdul-Razak, M. F. A., & Md-Zain, B. M. (2022). Determining the Dietary Preferences of Wild Asian Elephants (Elephas maximus) in Taman Negara National Park, Malaysia Based on Sex and Age using trnL DNA Metabarcoding Analysis. *Zoological Studies*, *61*, e60. https://doi.org/10.6620/ZS.2022.61-60

Acharya, K. R., Dhand, N. K., Whittington, R. J., & Plain, K. M. (2017). PCR Inhibition of a Quantitative PCR for Detection of Mycobacterium avium Subspecies Paratuberculosis DNA in Feces: Diagnostic Implications and Potential Solutions. *Frontiers in Microbiology*, *8*. https://www.frontiersin.org/articles/10.3389/fmicb.2017.00115

Agetsuma-Yanagihara, Y., Inoue, E., & Agetsuma, N. (2017). Effects of time and environmental conditions on the quality of DNA extracted from fecal samples for genotyping of wild deer in a warm temperate broad-leaved forest. *Mammal Research*, *62*(2), 201–207. https://doi.org/10.1007/s13364-016-0305-x

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134–147. https://doi.org/10.1111/2041-210X.12849

Alves, L., Alves, C., Jackson, A., De Souza, J. A., Varani, A., Gertrudes, E., & Lemos, E. (2014). 9 The Family Rhizobiaceae. *The Prokaryotes*. https://doi.org/10.1007/978-3-642-30197-1_297

Ando, H., Fujii, C., Kawanabe, M., Ao, Y., Inoue, T., & Takenaka, A. (2018). Evaluation of plant contamination in metabarcoding diet analysis of a herbivore. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-32845-w

Ando, H., Mukai, H., Komura, T., Dewi, T., Ando, M., & Isagi, Y. (2020a). Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. *Environmental DNA*, *2*(4), 391–406. https://doi.org/10.1002/edn3.117

Andriollo, T., Michaux, J. R., & Ruedi, M. (2021). Food for everyone: Differential feeding habits of cryptic bat species inferred from DNA metabarcoding. *Molecular Ecology*, *30*(18), 4584–4600. https://doi.org/10.1111/mec.16073

Bagchi, R., Brown, L. M., Elphick, C. S., Wagner, D. L., & Singer, M. S. (2018). Anthropogenic fragmentation of landscapes: Mechanisms for eroding the specificity of plant–herbivore interactions. *Oecologia*, *187*(2), 521–533. https://doi.org/10.1007/s00442-018-4115-5

Beauchamp, J. P., Cesarini, D., Johannesson, M., van der Loos, M. J. H. M., Koellinger, P. D., Groenen, P. J. F., Fowler, J. H., Rosenquist, J. N., Thurik, A. R., & Christakis, N. A. (2011). Molecular Genetics and Economics. *Journal of Economic Perspectives*, *25*(4), 57–82. https://doi.org/10.1257/jep.25.4.57

Bell, K. L., Loeffler, V. M., & Brosi, B. J. (2017). An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Applications in Plant Sciences*, *5*(3), apps.1600110. https://doi.org/10.3732/apps.1600110

Bessetti, J. (n.d.). *An Introduction to PCR Inhibitors*.

Bradshaw, L., & Waller, D. M. (2016). Impacts of white-tailed deer on regional patterns of forest tree recruitment. *Forest Ecology and Management*, *375*, 1–11. https://doi.org/10.1016/j.foreco.2016.05.019

Brinkman, T. J., Person, D. K., Schwartz, M. K., Pilgrim, K. L., Colson, K. E., & Hundertmark, K. J. (2010). Individual identification of Sitka black-tailed deer (Odocoileus hemionus sitkensis) using DNA from fecal pellets. *Conservation Genetics Resources*, *2*(1), 115–118. https://doi.org/10.1007/s12686-010-9176-7

Brown, R. D., & Cooper, S. M. (2006). In My Opinion: The Nutritional, Ecological, and Ethical Arguments Against Baiting and Feeding White-Tailed Deer. *Wildlife Society Bulletin*, *34*(2), 519–524. https://doi.org/10.2193/0091-7648(2006)34[519:TNEAEA]2.0.CO;2

Bushnell, B., Rood, J., & Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE*, *12*(10), e0185056. https://doi.org/10.1371/journal.pone.0185056

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), Article 7. https://doi.org/10.1038/nmeth.3869

Carreon-Martinez, L., & Heath, D. D. (2010). Revolution in food web analysis and trophic ecology: Diet analysis by DNA and stable isotope analysis. *Molecular Ecology*, *19*(1), 25–27. https://doi.org/10.1111/j.1365-294X.2009.04412.x

Cooper, S. M., Owens, M. K., Cooper, R. M., & Ginnett, T. F. (2006). Effect of supplemental feeding on spatial distribution and browse utilization by white-tailed deer in semi-arid rangeland. *Journal of Arid Environments*, *66*(4), 716–726. https://doi.org/10.1016/j.jaridenv.2005.11.015

Côté, W. A. (1968). The Structure of Wood and the Wood Cell Wall. In F. F. P. Kollmann & W. A. Côté (Eds.), *Principles of Wood Science and Technology: I Solid Wood* (pp. 1–54). Springer. https://doi.org/10.1007/978-3-642-87928-9_1

Davey, M. L., Kamenova, S., Fossøy, F., Solberg, E. J., Davidson, R., Mysterud, A., & Rolandsen, C. M. (2023a). Faecal metabarcoding provides improved detection and taxonomic resolution for non-invasive monitoring of gastrointestinal nematode parasites in wild moose populations. *Parasites & Vectors*, *16*(1), 19. https://doi.org/10.1186/s13071-022-05644-6

Davey, M. L., Utaaker, K. S., & Fossøy, F. (2021). Characterizing parasitic nematode faunas in faeces and soil using DNA metabarcoding. *Parasites & Vectors*, *14*(1), 422. https://doi.org/10.1186/s13071-021-04935-8

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, *14*(2), 306–323. https://doi.org/10.1111/1755-0998.12188

de Sousa, L. L., Silva, S. M., & Xavier, R. (2019). DNA metabarcoding in diet studies: Unveiling ecological aspects in aquatic and terrestrial ecosystems. *Environmental DNA*, *1*(3), 199–214. https://doi.org/10.1002/edn3.27

Demaria, E. M. C., Roundy, J. K., Wi, S., & Palmer, R. N. (2016). The Effects of Climate Change on Seasonal Snowpack and the Hydrology of the Northeastern and Upper Midwest United States. *Journal of Climate*, *29*(18), 6527–6541. https://doi.org/10.1175/JCLI-D-15-0632.1

Divoll, T. J., Brown, V. A., Kinne, J., McCracken, G. F., & O'Keefe, J. M. (2018). Disparities in second-generation DNA metabarcoding results exposed with accessible and repeatable workflows. *Molecular Ecology Resources*, *18*(3), 590–601. https://doi.org/10.1111/1755-0998.12770

Donaldson, L. A., Cairns, M., & Hill, S. J. (2018). Comparison of Micropore Distribution in Cell Walls of Softwood and Hardwood Xylem. *Plant Physiology*, *178*(3), 1142–1153. https://doi.org/10.1104/pp.18.00883

Dorn, R. D. (n.d.). *White-tailed Deer in Southeastern Minnesota: Winter Observations*. *Effect of supplemental feeding on spatial distribution and browse utilization by white-tailed deer in semi-arid rangeland—ScienceDirect*. (n.d.). Retrieved December 4, 2023, from https://www-sciencedirect-com.services.lib.mtu.edu/science/article/pii/S0140196305003010?casa_token=e_PEuI113v4AAAAA:7QgUaky5tFAPVf2UPzms_jn-FYEA9o0sOJXAYMyeXclr3FbdWuXE0Q0--5ksdZPsQ0Oyy5o

Ekblom, R., & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, *107*(1), Article 1. https://doi.org/10.1038/hdy.2010.152

Erickson, D. L., Reed, E., Ramachandran, P., Bourg, N. A., McShea, W. J., & Ottesen, A. (2017). Reconstructing a herbivore's diet using a novel rbcL DNA mini-barcode for plants. *AoB PLANTS*, *9*(3), plx015. https://doi.org/10.1093/aobpla/plx015

Ernest, H. B., Penedo, M. C., May, B. P., Syvanen, M., & Boyce, W. M. (2000). Molecular tracking of mountain lions in the Yosemite valley region in California: Genetic analysis using microsatellites and faecal DNA. *Molecular Ecology*, *9*(4), 433–441. https://doi.org/10.1046/j.1365-294x.2000.00890.x

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Faecal metabarcoding provides improved detection and taxonomic resolution for non-invasive monitoring of gastrointestinal nematode parasites in wild moose populations | Parasites & Vectors. (n.d.). Retrieved November 1, 2023, from https://link.springer.com/article/10.1186/s13071-022-05644-6

Ferguson, J. M., Hopkins III, J. B., & Witteveen, B. H. (2018). Integrating abundance and diet data to improve inferences of food web dynamics. *Methods in Ecology and Evolution*, *9*(6), 1581–1591. https://doi.org/10.1111/2041-210X.13001

Flint, H. J., & Bayer, E. A. (2008). Plant Cell Wall Breakdown by Anaerobic Microorganisms from the Mammalian Digestive Tract. *Annals of the New York Academy of Sciences*, *1125*(1), 280–288. https://doi.org/10.1196/annals.1419.022

Flores-Rentería and Krohn—2013—Scoring Microsatellite Loci.pdf. (n.d.). Retrieved October 12, 2023, from http://potpathodiv.org/static/papers/Flores-Renteria%20-%20Scoring%20microsatellite%20loci.pdf

Forrester, J. A., Lorimer, C. G., Dyer, J. H., Gower, S. T., & Mladenoff, D. J. (2014). Response of tree regeneration to experimental gap creation and deer herbivory in north temperate forests. *Forest Ecology and Management*, *329*, 137–147. https://doi.org/10.1016/j.foreco.2014.06.025

Galan, M., Pons, J.-B., Tournayre, O., Pierre, É., Leuchtmann, M., Pontier, D., & Charbonnel, N. (2018). Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis. *Molecular Ecology Resources*, *18*(3), 474–489. https://doi.org/10.1111/1755-0998.12749

Garwood, T. J., Moore, S. A., Fountain-Jones, N. M., Larsen, P. A., & Wolf, T. M. (2023). SPECIES IN THE FECES: DNA METABARCODING TO DETECT POTENTIAL GASTROPOD HOSTS OF PARELAPHOSTRONGYLUS TENUIS Consumed By Moose (Alces Alces). *Journal of Wildlife Diseases*, *59*(4), 640–650. https://doi.org/10.7589/JWD-D-22-00120

Gillet, F., Tiouchichine, M.-L., Galan, M., Blanc, F., Némoz, M., Aulagnier, S., & Michaux, J. R. (2015). A new method to identify the endangered Pyrenean desman (Galemys pyrenaicus) and to study its diet, using next generation sequencing from faeces. *Mammalian Biology*, *80*(6), 505–509. https://doi.org/10.1016/j.mambio.2015.08.002

Gold, Z., Curd, E. E., Goodwin, K. D., Choi, E. S., Frable, B. W., Thompson, A. R., Walker Jr., H. J., Burton, R. S., Kacev, D., Martz, L. D., & Barber, P. H. (2021). Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. *Molecular Ecology Resources*, *21*(7), 2546–2564. https://doi.org/10.1111/1755-0998.13450

Grado, S. C., Hunt, K. M., & Whiteside, M. W. (2007). *Economic Impacts of White-tailed Deer Hunting in Mississippi*.

Guillerault, N., Bouletreau, S., Iribar, A., Valentini, A., & Santoul, F. (2017). Application of DNA metabarcoding on faeces to identify European catfish Silurus glanis diet. *Journal of Fish Biology*, *90*(5), 2214–2219. https://doi.org/10.1111/jfb.13294

Hart, M. L., Meyer, A., Johnson, P. J., & Ericsson, A. C. (2015). Comparative Evaluation of DNA Extraction Methods from Feces of Multiple Host Species for Downstream Next-Generation Sequencing. *PLOS ONE*, *10*(11), e0143334. https://doi.org/10.1371/journal.pone.0143334

He, K., Fujiwara, H., Zajac, C., Sandford, E., Reddy, P., Choi, S. W., & Tewari, M. (2019). A Pipeline for Faecal Host DNA Analysis by Absolute Quantification of LINE-1 and Mitochondrial Genomic Elements Using ddPCR. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-41753-6

Hoffmann, G. S., Johannesen, J., & Griebeler, E. M. (2015). Species cross-amplification, identification and genetic variation of 17 species of deer (Cervidae) with microsatellite and mitochondrial DNA from antlers. *Molecular Biology Reports*, *42*(6), 1059–1067. https://doi.org/10.1007/s11033-014-3845-7

Hotvedt, J. E., & Luzar, E. J. (1989). The economic value of white-tailed deer hunting on the Sherburne Wildlife Management Area and the Atchafalaya National Wildlife Refuge. *The Economic Value of White-Tailed Deer Hunting on the Sherburne Wildlife Management Area and the Atchafalaya National Wildlife Refuge.*, *No. 820*. https://www.cabdirect.org/cabdirect/abstract/19901876781

Jeunen, G.-J., Dowle, E., Edgecombe, J., von Ammon, U., Gemmell, N. J., & Cross, H. (2023). crabs—A software program to generate curated reference databases for metabarcoding sequencing data. *Molecular Ecology Resources*, *23*(3), 725–738. https://doi.org/10.1111/1755-0998.13741

Jones, J., Reinke, S. N., Ali, A., Palmer, D. J., & Christophersen, C. T. (2021). Fecal sample collection methods and time of day impact microbiome composition and short chain fatty acid concentrations. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-93031-z

Joos, L., Beirinckx, S., Haegeman, A., Debode, J., Vandecasteele, B., Baeyen, S., Goormachtig, S., Clement, L., & De Tender, C. (2020). Daring to be differential: Metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics*, *21*(1), 733. https://doi.org/10.1186/s12864-020-07126-4

Jung, T. S., & Kukka, P. M. (2016). Influence of habitat type on the decay and disappearance of elk Cervus canadensis pellets in boreal forest of northwestern Canada. *Wildlife Biology*, *22*(4), 160–166. https://doi.org/10.2981/wlb.00186

Kartzinel, T. R., Chen, P. A., Coverdale, T. C., Erickson, D. L., Kress, W. J., Kuzmina, M. L., Rubenstein, D. I., Wang, W., & Pringle, R. M. (2015). DNA

metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proceedings of the National Academy of Sciences*, *112*(26), 8019–8024. https://doi.org/10.1073/pnas.1503283112

Kovach, A., Litvaitis, M., & Litvaitis, J. (2003). Evaluation of Fecal mtDNA Analysis as a Method to Determine the Geographic Distribution of a Rare Lagomorph. *Wildlife Society Bulletin*, *31*, 1061–1065. https://doi.org/10.2307/3784452

Krehenwinkel, H., Fong, M., Kennedy, S., Huang, E. G., Noriyuki, S., Cayetano, L., & Gillespie, R. (2018). The effect of DNA degradation bias in passive sampling devices on metabarcoding studies of arthropod communities and their associated microbiota. *PLOS ONE*, *13*(1), e0189188. https://doi.org/10.1371/journal.pone.0189188

Lauber, C. L., Zhou, N., Gordon, J. I., Knight, R., & Fierer, N. (2010). Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiology Letters*, *307*(1), 80–86. https://doi.org/10.1111/j.1574-6968.2010.01965.x

Low concentration DNA extraction and recovery using a silica solid phase | PLOS ONE. (n.d.). Retrieved November 1, 2023, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0176848

Mallott, E. K., Garber, P. A., & Malhi, R. S. (2018). trnL outperforms rbcL as a DNA metabarcoding marker when compared with the observed plant component of the diet of wild white-faced capuchins (Cebus capucinus, Primates). *PLOS ONE*, *13*(6), e0199556. https://doi.org/10.1371/journal.pone.0199556

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

Massey, A. L., Roffler, G. H., Vermeul, T., Allen, J. M., & Levi, T. (2021). Comparison of mechanical sorting and DNA metabarcoding for diet analysis with fresh and degraded wolf scats. *Ecosphere*, *12*(6), e03557. https://doi.org/10.1002/ecs2.3557

McShea, W. J. (2012). Ecology and management of white-tailed deer in a changing world. *Annals of the New York Academy of Sciences*, *1249*(1), 45–56. https://doi.org/10.1111/j.1749-6632.2011.06376.x

McShea, W. J., & Rappole, J. H. (2000). Managing the Abundance and Diversity of Breeding Bird Populations through Manipulation of Deer Populations. *Conservation Biology*, *14*(4), 1161–1170. https://doi.org/10.1046/j.1523-1739.2000.99210.x

Miller, W.L., Edson, J., Pietrandrea, P. *et al.* Identification and evaluation of a core microsatellite panel for use in white-tailed deer (*Odocoileus virginianus*). *BMC Genet* **20**, 49 (2019). https://doi.org/10.1186/s12863-019-0750-z

Miller, R., Kaneene, J. B., Fitzgerald, S. D., & Schmitt, S. M. (2003a). EVALUATION OF THE INFLUENCE OF SUPPLEMENTAL FEEDING OF WHITE-TAILED DEER (ODOCOILEUS VIRGINIANUS) ON THE PREVALENCE OF BOVINE TUBERCULOSIS IN THE MICHIGAN WILD DEER POPULATION. *Journal of Wildlife Diseases*, *39*(1), 84–95. https://doi.org/10.7589/0090-3558-39.1.84

Monteiro, L., Bonnemaison, D., Vekris, A., Petry, K. G., Bonnet, J., Vidal, R., Cabrita, J., & Mégraud, F. (1997). Complex polysaccharides as PCR inhibitors in feces: Helicobacter pylori model. *Journal of Clinical Microbiology*, *35*(4), 995–998.

Moorhouse-Gann, R. J., Dunn, J. C., de Vere, N., Goder, M., Cole, N., Hipperson, H., & Symondson, W. O. C. (2018). New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-26648-2

Mugnai, F., Costantini, F., Chenuil, A., Leduc, M., Ortega, J. M. G., & Meglécz, E. (2023). Be positive: Customized reference databases and new, local barcodes balance false taxonomic assignments in metabarcoding studies. *PeerJ*, *11*, e14616. https://doi.org/10.7717/peerj.14616

Nørgaard, L., Olesen, C. R., Trøjelsgaard, K., Pertoldi, C., Nielsen, J. L., Taberlet, P., Ruiz-González, A., De Barba, M., & Iacolina, L. (2021). eDNA metabarcoding for biodiversity assessment, generalist predators as sampling assistants. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-85488-9

Pálsson, B., Pálsson, F., Perlin, M., Gudbjartsson, H., Stefánsson, K., & Gulcher, J. (1999). Using Quality Measures to Facilitate Allele Calling in High-Throughput Genotyping. *Genome Research*, *9*(10), 1002–1012.

Parikh, G. L., & Webster, C. R. (2019). Press-Pulse Odocoileus Virginianus Herbivory in Relict Tsuga Canadensis Stands in the Western Upper Peninsula of Michigan, USA. *Forests*, *10*(6), Article 6. https://doi.org/10.3390/f10060496

Parikh, G. L., Webster, C. R., & Vucetich, J. A. (2021). A Microhistological Investigation of Winter Diets of White-Tailed Deer in Relict Eastern Hemlock Stands, Upper Peninsula of Michigan. *Northeastern Naturalist*, *28*(3), 296–310. https://doi.org/10.1656/045.028.0305

Polis, G. A., & Strong, D. R. (1996). Food Web Complexity and Community Dynamics. *The American Naturalist*, *147*(5), 813–846. https://doi.org/10.1086/285880

Ragot, R., Lessard, F., Bélanger, A., & Villemur, R. (2023). Assessment of multiple fecal contamination sources in surface waters using environmental mitochondrial DNA metabarcoding. *Science of The Total Environment*, *898*, 165237. https://doi.org/10.1016/j.scitotenv.2023.165237

Robeson II, M. S., Khanipov, K., Golovko, G., Wisely, S. M., White, M. D., Bodenchuck, M., Smyser, T. J., Fofanov, Y., Fierer, N., & Piaggio, A. J. (2018). Assessing the utility of metabarcoding for diet analyses of the omnivorous wild pig (Sus scrofa). *Ecology and Evolution*, *8*(1), 185–196. https://doi.org/10.1002/ece3.3638

Röder, B., Frühwirth, K., Vogl, C., Wagner, M., & Rossmanith, P. (2010). Impact of Long-Term Storage on Stability of Standard DNA for Nucleic Acid-Based Methods. *Journal of Clinical Microbiology*, *48*(11), 4260–4262. https://doi.org/10.1128/jcm.01230-10

Rogers, L. L. (1981). *Foods of White-tailed Deer in the Upper Great Lakes Region: A Review*. North Central Forest Experiment Station, Forest Service, U.S. Department of Agriculture.

Ross, K. S., Haites, N. E., & Kelly, K. F. (1990). Repeated freezing and thawing of peripheral blood and DNA in suspension: Effects on DNA yield and integrity. *Journal of Medical Genetics*, *27*(9), 569–570.

Rubin, B. E. R., Gibbons, S. M., Kennedy, S., Hampton-Marcell, J., Owens, S., & Gilbert, J. A. (2013). Investigating the Impact of Storage Conditions on Microbial Community Composition in Soil Samples. *PLOS ONE*, *8*(7), e70460. https://doi.org/10.1371/journal.pone.0070460

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, *17*, e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Salk, T. T., Frelich, L. E., Sugita, S., Calcote, R., Ferrari, J. B., & Montgomery, R. A. (2011). Poor recruitment is changing the structure and species composition of an old-growth hemlock-hardwood forest. *Forest Ecology and Management*, *261*(11), 1998–2006. https://doi.org/10.1016/j.foreco.2011.02.026

Schloss, P. D. (n.d.). Amplicon Sequence Variants Artificially Split Bacterial Genomes into Separate Clusters. *mSphere*, *6*(4), e00191-21. https://doi.org/10.1128/mSphere.00191-21

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

Schmitz, O. J. (1990). Management Implications of Foraging Theory: Evaluating Deer Supplemental Feeding. *The Journal of Wildlife Management*, *54*(4), 522–532. https://doi.org/10.2307/3809345

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors – occurrence, properties and removal. *Journal of Applied Microbiology*, *113*(5), 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Schuette, P., Ebbert, S., Droghini, A., & Nawrocki, T. (2022). Small mammal diet indicates plant diversity, vegetation structure, and ecological integrity in a remote ecosystem. *Biodiversity and Conservation*, *31*(3), 909–924. https://doi.org/10.1007/s10531-022-02370-4

Seeber, P. A., von Hippel, B., Kauserud, H., Löber, U., Stoof-Leichsenring, K. R., Herzschuh, U., & Epp, L. S. (2022). Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems. *Environmental DNA*, *4*(5), 1150–1163. https://doi.org/10.1002/edn3.315

Shao, W., Khin, S., & Kopp, W. C. (2012). Characterization of effect of repeated freeze and thaw cycles on stability of genomic DNA using pulsed field gel electrophoresis. *Biopreservation and Biobanking*, *10*(1), 4–11. https://doi.org/10.1089/bio.2011.0016

Sidstedt, M., Rådström, P., & Hedman, J. (2020). PCR inhibition in qPCR, dPCR and MPS—mechanisms and solutions. *Analytical and Bioanalytical Chemistry*, *412*(9), 2009–2023. https://doi.org/10.1007/s00216-020-02490-2

Sluter, S. D., Tzipori, S., & Widmer, G. (1997). Parameters affecting polymerase chain reaction detection of waterborne Cryptosporidium parvum oocysts. *Applied*

Microbiology and Biotechnology, *48*(3), 325–330. https://doi.org/10.1007/s002530051057

Smith, L. W., Weinland, B. T., Waldo, D. R., & Leffel, E. C. (1983). Rate of Plant Cell Wall Particle Size Reduction in the Rumen. *Journal of Dairy Science*, *66*(10), 2124–2136. https://doi.org/10.3168/jds.S0022-0302(83)82059-1

Stadhouders, R., Pas, S. D., Anber, J., Voermans, J., Mes, T. H. M., & Schutten, M. (2010). The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay. *The Journal of Molecular Diagnostics: JMD*, *12*(1), 109–117. https://doi.org/10.2353/jmoldx.2010.090035

Suzuki, M., Miyashita, T., Kabaya, H., Ochiai, K., Asada, M., & Kikvidze, Z. (2013). Deer herbivory as an important driver of divergence of ground vegetation communities in temperate forests. *Oikos*, *122*(1), 104–110. https://doi.org/10.1111/j.1600-0706.2012.20431.x

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trn L (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14. https://doi.org/10.1093/nar/gkl938

The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools—PMC. (n.d.). Retrieved November 1, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531112/

Thuo, D., Furlan, E., Broekhuis, F., Kamau, J., Macdonald, K., & Gleeson, D. M. (2019). Food from faeces: Evaluating the efficacy of scat DNA metabarcoding in dietary

analyses. *PLOS ONE*, *14*(12), e0225805. https://doi.org/10.1371/journal.pone.0225805

van der Loos, L. M., & Nijland, R. (2021). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology*, *30*(13), 3270–3288. https://doi.org/10.1111/mec.15592

Verkuil, Y. I., Nicolaus, M., Ubels, R., Dietz, M. W., Samplonius, J. M., Galema, A., Kiekebos, K., de Knijff, P., & Both, C. (2022). DNA metabarcoding quantifies the relative biomass of arthropod taxa in songbird diets: Validation with camera-recorded diets. *Ecology and Evolution*, *12*(5), e8881. https://doi.org/10.1002/ece3.8881

Vishnivetskaya, T. A., Layton, A. C., Lau, M. C. Y., Chauhan, A., Cheng, K. R., Meyers, A. J., Murphy, J. R., Rogers, A. W., Saarunya, G. S., Williams, D. E., Pfiffner, S. M., Biggerstaff, J. P., Stackhouse, B. T., Phelps, T. J., Whyte, L., Sayler, G. S., & Onstott, T. C. (2014). Commercial DNA extraction kits impact observed microbial community composition in permafrost samples. *FEMS Microbiology Ecology*, *87*(1), 217–230. https://doi.org/10.1111/1574-6941.12219

Walker, F. M., Sanchez, D. E., Froehlich, E. M., Federman, E. L., Lyman, J. A., Owens, M., & Lear, K. (2022). Endangered Nectar-Feeding Bat Detected by Environmental DNA on Flowers. *Animals*, *12*(22), Article 22. https://doi.org/10.3390/ani12223075

Waller, D. M., & Alverson, W. S. (1997). The White-Tailed Deer: A Keystone Herbivore. *Wildlife Society Bulletin (1973-2006)*, *25*(2), 217–226.

Wang, J., Chapman, S. J., & Yao, H. (2015). The effect of storage on microbial activity and bacterial community structure of drained and flooded paddy soil. *Journal of Soils and Sediments*, *15*(4), 880–889. https://doi.org/10.1007/s11368-014-1053-7

Wang, Y.-C., Liang, Y.-C., Huang, F.-L., & Chang, W.-C. (2023). Effect of Freeze–Thaw Cycles on Physicochemical and Functional Properties of Ginger Starch. *Processes*, *11*(6), Article 6. https://doi.org/10.3390/pr11061828

Westfall, K. M., Therriault, T. W., & Abbott, C. L. (2020). A new approach to molecular biosurveillance of invasive species using DNA metabarcoding. *Global Change Biology*, *26*(2), 1012–1022. https://doi.org/10.1111/gcb.14886

Wietz, M., Metfies, K., Bienhold, C., Wolf, C., Janssen, F., Salter, I., & Boetius, A. (2022). Impact of preservation method and storage period on ribosomal metabarcoding of marine microbes: Implications for remote automated samplings. *Frontiers in Microbiology*, *13*. https://www.frontiersin.org/articles/10.3389/fmicb.2022.999925

Witt, J. C., & Webster, C. R. (2010). Regeneration dynamics in remnant Tsuga canadensis stands in the northern Lake States: Potential direct and indirect effects of herbivory. *Forest Ecology and Management*, *260*(4), 519–525. https://doi.org/10.1016/j.foreco.2010.05.007

Witt, J. C., Webster, C. R., Froese, R. E., Drummer, T. D., & Vucetich, J. A. (2012). Scale-dependent drivers of ungulate patch use along a temporal and spatial gradient of snow depth. *Canadian Journal of Zoology*, *90*(8), 972–983. https://doi.org/10.1139/z2012-065

Wu, W.-K., Chen, C.-C., Panyod, S., Chen, R.-A., Wu, M.-S., Sheen, L.-Y., & Chang, S.-C. (2019). Optimization of fecal sample processing for microbiome study—The journey from bathroom to bench. *Journal of the Formosan Medical Association*, *118*(2), 545–555. https://doi.org/10.1016/j.jfma.2018.02.005

Xin, Y., Xie, J., Nan, B., Tang, C., Xiao, Y., Wu, Q., Lin, Y., Zhang, X., & Shen, H. (2021). Freeze-Thaw Pretreatment Can Improve Efficiency of Bacterial DNA Extraction From Meconium. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.753688

Young, M. J., Dutoit, L., Robertson, F., Heezik, Y. van, Seddon, P. J., Robertson, B. C., Young, M. J., Dutoit, L., Robertson, F., Heezik, Y. van, Seddon, P. J., & Robertson, B. C. (2020). Species in the faeces: DNA metabarcoding as a method to determine the diet of the endangered yellow-eyed penguin. *Wildlife Research*, *47*(6), 509–522. https://doi.org/10.1071/WR19246

Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C., & Jones, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, *11*(2), 236–244. https://doi.org/10.1111/j.1755-0998.2010.02920.x

4 Metabarcoding Methodology

4.1 Modifications to Commercial DNA Extraction Kit Protocols

4.1.1 Overview of Alterations to DNA Extraction Protocol

The extraction kits optimized for non-eDNA sample types were modified from stand protocols provided by manufacturers to increase suitability of kits for feces. No modifications were made to the Zymo Research Fecal/Soil Miniprep kit standard practices.

4.1.2 Qiagen DNeasy Blood & Tissue

For Qiagen DNeasy Blood & Tissue, steps 1-3 of the Qiagen DNeasy standard protocol for animal tissue were slightly modified to the following: 200 mg of frozen homogenized fecal matter was added to a 2mL microcentrifuge tube containing 10-20 1.0mm Zirconia/Silica beads, 20 µL proteinase K, and 180 µL Buffer ATL along with a negative extraction control. These tubes were vortexed on high-speed for 10 minutes, then incubated overnight at 56°C on a thermomixer (500rpms). After incubation, samples were centrifuged for 5 minutes at maximum speed to pellet fecal matter and beads, and supernatant was transferred to a new 2mL microcentrifuge tube (~200 µL) before addition of 200 µL of Buffer AL. After thorough mixing by vortex at a 10-minute incubation at 56°C, 200 µL of 96% ethanol was added and the resulting solution vortexed and quickly spun down before transferring to a spin column. Steps 4-8 of the standard protocol were followed as written in the Qiagen handbook, resulting in 100 µL of final purified DNA eluted.

4.1.3 TakaraBio NucleoSpin Plant II

For the TakaraBio NucleoSpin Plant II kit, steps 1-2 were modified from the NucleoSpin Plant II protocol documentation, from both the standard plant tissue and soil/fecal recommendations. Into labeled 2mL microcentrifuge tubes, 10-20 1.0mm Zirconia/Silica beads and 200 mg of homogenized fecal matter were added along with 10 µL of RNase A solution and 500 µL of Lysis Buffer PL1. The samples were vortexed at maximum speed for 10 minutes, then incubated for a minimum of 1 hour at 65 in a thermomixer at 400 rpm. Lysed samples were centrifuged for 10 minutes at 5000 x g to pellet fecal matter and beads, and 300 µL of clear supernatant was transferred to the first spin column (violet ring) with a new 2mL collection tube. Steps 3-7 of NucleoSpin Plant II recommended protocol were then followed as written, resulting in 100 µL of purified DNA eluted.

5 Supplementary Figures

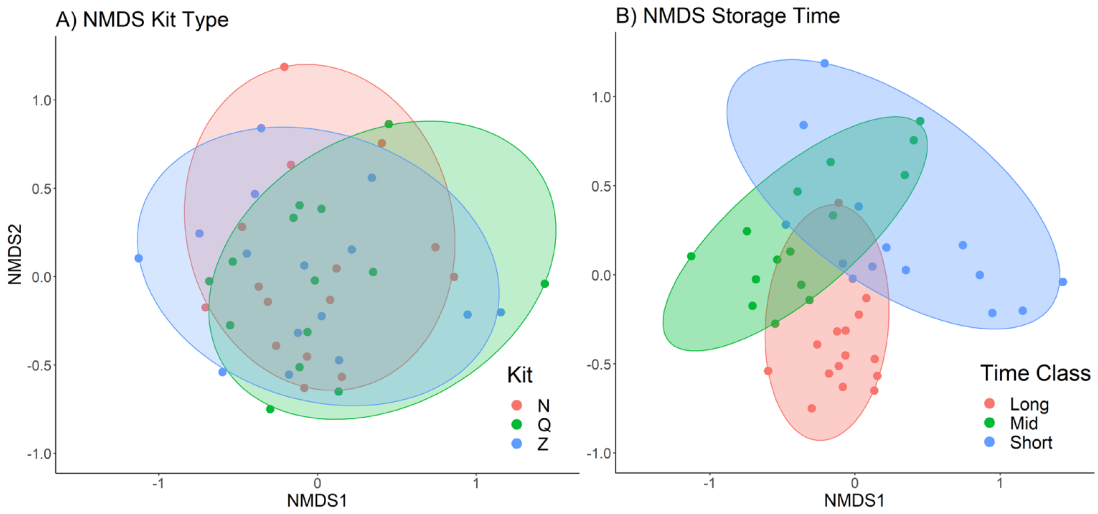


Figure 5.1. NMDS Ordination of microbial communities detected in fecal samples based on storage time and DNA extraction kit type. Ordinations were generated in R using vegan and similarity measures (A) and p-values (p) from multiple response permutation procedures between factor groups provided for each ordination.

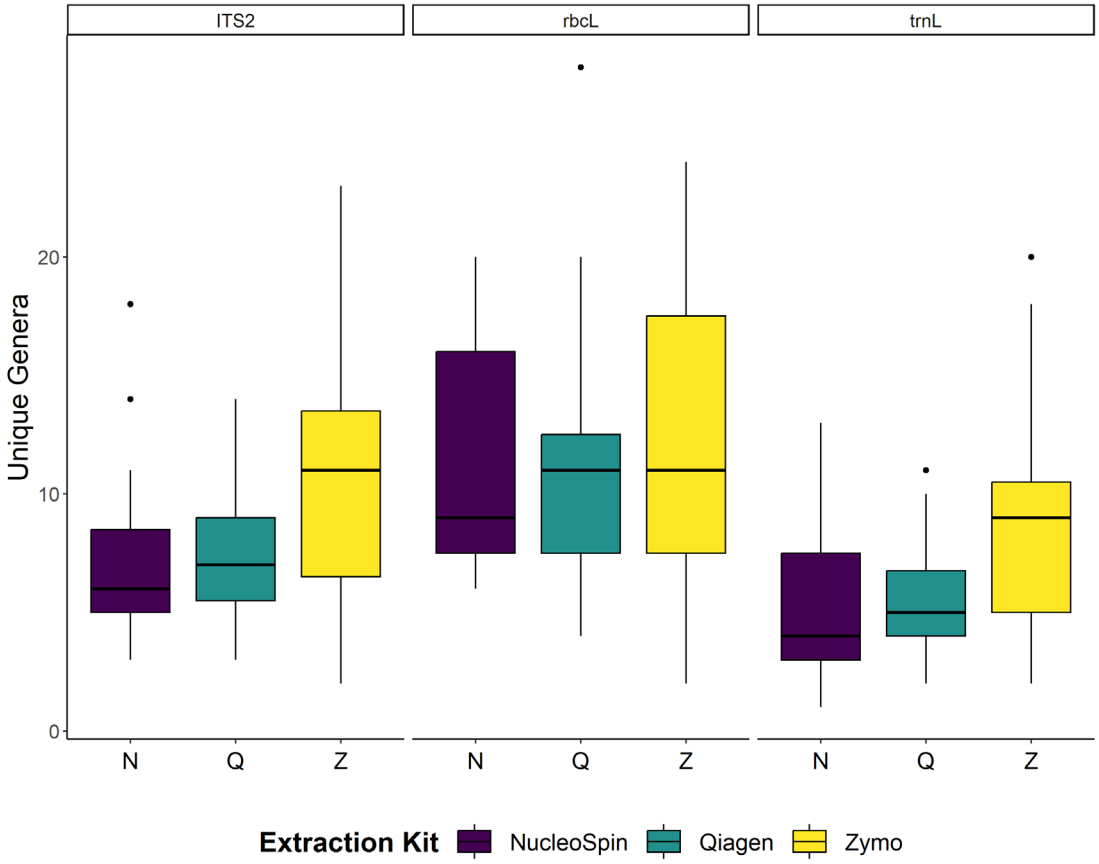


Figure 5.2. Boxplots of unique plant genera identified by three universal plant primers across three different commercial DNA extraction kits.

6 Supplementary Tables

Table 6.1. Species-specific equations for estimating crown width adapted from Bragg et al. 2001

Species	Equation
American basswood	2.563005 +0.069720(DBH)^1.133592 - 0.016178(LBA)
American elm	4.097250 +0.025662(DBH)^ 1.343831 -0.026273(LBA)
Balsam fir	-1.148370+ 1.484774(DBH)^ 0.398922 +0.000989(LBA)
Bigtooth aspen	1.899554+ 0.127094(DBH)^ 1.045309 -0.033460(LBA)
Black ash	1.044675 +0.429465(DBH)^ 0.719439 -0.009582(LBA)
Black cherry	0.396623 +1.047267 (DBH)^0.502013 -0.026799(LBA)
Black spruce	1.500497 +0.013990(DBH)^ 1.662184 -0.017577(LBA)
Eastern hemlock	1.442937 +0.991358 (DBH)^0.540750-0.040759(LBA)
Eastern white pine	1.419708 +0.367860(DBH)^ 0.762768-0.048905(LBA)
Ironwood	-11.033100 +10.786200 *(DBH)^0.146459 +0.005400 (LBA)
Northern red oak	1.796712 +0.546874*(DBH)^ 0.758820 -0.077570(LBA)
Northern white-cedar	1.933244 +0.146711*(DBH)^ 0.898806-0.018520(LBA)
Paper birch	2.223017 +0.067197*(DBH)^ 1.243736-0.022022(LBA)
Pin cherry	1.260218 +0.130841*(DBH)^ 1.130661 +0.004834(LBA)
Quaking aspen	0.917085+ 0.426571*(DBH)^ 0.772969 -0.034042(LBA)
Red maple	1.946356 +0.277289*(DBH)^ 0.852833 -0.014818(LBA)
Red Pine	1.454644 + 0.131228*(DBH)^ 1.004795 -0.023031 (LBA)
Sugar maple	2.119782 + 0.346366(DBH)^ 0.813395 -0.017616(LBA)
Tamarack	1.325384 +0.127903(DBH)^ 1.064072 -0.055023(LBA)
White ash	4.067896 + 0.126510(DBH)^ 1.055638 -0.087031(LBA)
White spruce	1.653061 +0.290016(DBH)^ 0.730953 -0.026987(LBA)
Yellow birch	1.297470 +0.697196(DBH)^ 0.670675 -0.023695(LBA)

Table 6.2. PCR reagent concentrations for each metabarcoding primer pair for plant and microbial taxa. All reactions were run at 20 µL volumes with 2 µL of purified fecal DNA.

Reagent	trnL	rbcL	ITS2	16S (V4)	16S (V3-V4)
AmpliTaq Gold 360 Polymerase	1.75x	1.75x			
AmpliTaq Gold 360 Master Mix			1X	1X	1X
Forward Primer	0.2 µM	0.2 µM	0.4 µM	0.4 µM	0.4 µM
Reverse Primer	0.2 µM	0.2 µM	0.4 µM	0.4 µM	0.4 µM
MgCl	2 mM	2 mM			
dNTPs	200 µM	200 µM			
BSA	0.1 mM	0.1 mM	0.1 mM	0.1 mM	0.1 mM

Table 6.3. Summary of PCR thermocycling conditions for all plant and microbial metabarcoding primers.

Step	trnL	rbcL	ITS2	16S (v4)	16S (v3-v4)
Activation	1 Cycle	1 Cycle	1 Cycle	1 Cycle	1 Cycle
	95°C for 10 minutes	95°C for 10 min	95°C for 10 min	95°C for 10 min	95°C for 10 min
Amplification	35 Cycles	35 Cycles	40 Cycles	35 Cycles	30 Cycles
Denature	95°C for 30 sec	94°C for 30 sec	94°C for 30 sec	94°C for 45 sec	94°C for 30 sec
Anneal	55°C for 30 sec	55°C for 30 sec	55°C for 30 sec	52°C for 45 sec	53°C for 45 sec
Extend	72°C for 30 sec	72°C for 60 sec	72°C for 60 sec	72°C for 60 sec	72°C for 60 sec
Final Extension	1 Cycle	1 Cycle	1 Cycle	1 Cycle	1 Cycle
	72°C for 4 min	72°C for 7 min	72°C for 7 min	72°C for 7 min	72°C for 10 min
Hold	4°C until stopped	4°C until stopped	4°C until stopped	4°C until stopped	4°C until stopped

Table 6.4. Summary table of plant order, family, and genus levels identified by the three tested universal plant metabarcoding primers. Overlapping taxonomic identifications are listed with combinations of initial letter of locus, including IRT (all loci), IR (ITS2 and rbcL), IT (ITS2 and trnL), and RT (rbcL and trnL).

	Genus Level						
	IRT	**IR**	**IT**	**RT**	**ITS2 Only**	**rbcL Only**	**trnL Only**
	Abies, Acer, Alnus, Betula, Carex, Cornus, Glyercia, Juncus, Oxalis, Picea, Pinus, Prunus, Quercus, Ribes, Thuja	Chamaedaphne, Coptis, Fraxinus, Lolium, Milium, Poa, Rhododendron, Rubus, Vaccinium	Agrostis, Berteroa, Epilobium, Festuca, Gaultheria, Juniperus, Lactuca, Pilosella	Avenella, Huperzioideae, Larix, Linnaea, Luzula, Rumex	Anthoxanthum, Artemisia, Boehmeria, Brassica, Cinna, Cucumis, Cucurbita, Dactylis, Deschampsia, Fragaria, Humulus, Ilex, Impatiens, Juglans, Leucanthemum, Ostrya, Populus, Ranunculus, Rostraria, Sambucus, Tanacetum, Urtica	Amphineurion, Arrhenatherum, Castanea, Castanopsis, Cerastium, Cupressus, Datisca, Dendrolycopodium, Fagus, Imperata, Lepidozia, Pohlia, Polytrichastrum, Prunella, Sorbus, Sphagnum, Tsuga, Viburnum	Arachis, Athyrium, Chimonocalamus, Cosmarium, Dicranella, Dicranum, Digitalis, Dryopteris, Gaylussacia, Glycine, Keteleeria, Lonicera, Maianthemum, Melocalamus, Musa, Oryza, Salix, Sisymbrium
TOTALS	**16**	**9**	**8**	**6**	**22**	**18**	**18**
	Family Level						
	IRT	**IR**	**IT**	**RT**	**ITS2 Only**	**rbcL Only**	**trnL Only**
	Asteraceae, Betulaceae, Cornaceae, Cupressaceae, Cyperaceae, Ericaceae, Fagaceae, Grossulariaceae, Juncaceae, Oleaceae, Oxalidaceae, Pinaceae, Plantaginaceae, Poaceae, Ranunculaceae, Rosaceae, Sapindaceae	Adoxaceae, Cucurbitaceae	Brassicaceae, Onagraceae, Salicaceae	Asparagaceae, Caprifoliaceae, Lycopodiaceae, Polygonaceae	Aceraceae, Aquifoliaceae, Araliaceae, Balsaminaceae, Cannabaceae, Juglandaceae, Urticaceae	Apocynaceae, Caryophyllaceae, Datiscaceae, Lamiaceae, Lepidoziaceae, Mniaceae, Orthotrichaceae, Polytrichaceae, Pylaisiadelphaceae, Sphagnaceae	Athyriaceae, Desmidiaceae, Dicranaceae, Dryopteridaceae, Drypoteridaceae, Fabaceae, Musaceae, Neckeraceae
TOTALS	**17**	**2**	**3**	**4**	**7**	**10**	**7**
	Order Level						
	IRT	**IR**	**IT**	**RT**	**ITS2 Only**	**rbcL Only**	**trnL Only**
	Asterales, Cornales, Cupressales, Dipsacales, Ericales, Fagales, Lamiales, Oxalidales, Pinales, Poales, Ranunculales, Rosales, Sapindales, Saxifragales	Cucurbitales	Brassicales, Malpighiales, Myrtales	Asparagales, Caryophyllales, Dicranales, Hypnales, Lycopodiales	Apiales, Aquifoliales	Bryales, Gentianales, Jungermanniales, Orthotrichales, Polytrichales, Sphagnales	Desmidiales, Fabales, Polypodiales, Zingiberales
TOTALS	**14**	**1**	**3**	**5**	**2**	**6**	**4**