

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Determinants and Implications of Variation in the Avian Gut Microbiota

Permalink

<https://escholarship.org/uc/item/9xn036k9>

Author

Vo, Anh-Thu

Publication Date

2014

Peer reviewed|Thesis/dissertation

Determinants and Implications of Variation in the Avian Gut Microbiota

By

Anh-Thu Elaine Vo

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Rauri C. K. Bowie

Dr. Michael Shapira

Professor Rodrigo P. P. Almeida

Spring 2014

© Copyright by

Anh-Thu Elaine Vo

2014

Abstract

Determinants and Implications of Variation in the Avian Gut Microbiota

by

Anh-Thu Elaine Vo

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Rauri C. K. Bowie, Chair

Microbes occur in extraordinary abundance and diversity in vertebrates, and a growing body of work, particularly on humans and mice, has highlighted the pervasive role that microbiota may play in host physiology and health. Yet factors underlying the assembly of microbiota and the potential for microbiota to mediate host fitness in free-living vertebrate populations remain understudied. Investigating host-microbial interactions in natural systems is necessary to understand the extent to which ecologically relevant levels of environmental heterogeneity shape host-microbial relationships and to gain insight into the role that microbiota may play in the ecology and evolution of non-model host taxa. The neonatal vertebrate life stage is of particular interest during which substantial physiological development occurs and through which the initial microbial colonization can have longlasting effects on metabolic programming. To better understand mechanisms underlying microbial variation and associations with host condition during early periods of vertebrate development, this study characterized the gut microbiota of nestlings and their mothers in a natural population of Western Bluebirds *Sialia mexicana*. A novel method based on solid phase reversible immobilization beads was developed to extract metagenomic DNA from minimally invasive avian fecal and swab specimens. Illumina sequence analysis of bacterial 16S rRNA amplicons was performed on oral, cloacal, and fecal samples collected longitudinally from nestlings throughout the nestling period across multiple nests and over three consecutive breeding seasons (2011-2013). Environmental selection appeared to drive variation in microbiota across gut sites within individuals as well as between nests among individuals. Ambient temperature and sampling year correlated substantially with variation in the gut microbiota. In contrast, surprisingly no association was recovered between host age and alpha diversity, and increased similarity of microbiota across gut sites with host age was observed. Also unexpectedly, shared nest environment and geographical distance were weak predictors of microbial similarity among individuals. Variation in the gut microbiota significantly associated with morphological metrics of nestling growth, even after accounting for the influence of several temporal and environmental variables known to mediate nestling development. However, gut microbiota comprised a less robust predictor of nestling morphology than host age, sampling year, and ambient temperature. Furthermore, perturbation in the form of a viral immune challenge caused measurable shifts in the bacterial

gut microbiota, which has previously been undocumented, but no measured effect of the immune challenge was observed on morphological metrics. These data inform intriguing hypotheses for future research on the drivers of avian host-microbial associations and the extent to which microbiota may impact host growth in natural systems. Discerning how microbiota may generate phenotypic or fitness differences among vertebrate host organisms is essential to understanding the evolution and maintenance of biodiversity.

Dedication

To Bunnie, my mossy rock

Table of Contents

Dedicationi

Table of Contentsii

Introductionv

 Referencesvii

Acknowledgements..... ix

Chapter 1: A primer on avian microbiology..... 1

 Abstract 1

 Introduction..... 1

 Background..... 2

 (i) Microbial infection of eggs 3

 (ii) Bacterial colonization of young birds..... 3

 (iii) Variation in the gut microbiota of adults..... 4

 (iv) Host genetic versus environmental determinants of biogeographical variation 4

 (v) Sexual transmission of bacteria 5

 (vi) Feather-degrading bacteria..... 5

 (vii) Microbial effects on avian physiology..... 6

 (viii) The role of birds in pathogen and disease ecology..... 6

 Conclusion 7

 References..... 8

 Figures 14

Chapter 2: Protocols for metagenomic DNA extraction and Illumina amplicon library
preparation for fecal and swab samples 15

 Abstract 15

 Introduction..... 15

 Materials and methods 17

 Results 21

 Discussion 22

 References..... 26

 Tables 31

 Figures 32

 Supporting Information..... 35

Chapter 3: Host-microbial relationships in a changing world: determinants of natural variation in the avian gut microbiota.....	65
Abstract	65
Introduction.....	65
Materials and methods	67
Results	70
Discussion	72
Conclusion	77
References.....	77
Tables	82
Figures	83
Supporting Information.....	90
Chapter 4: Relationships between the avian gut microbiota and nestling development.....	97
Abstract	97
Introduction.....	97
Materials and methods	98
Results	101
Discussion	102
Conclusion	104
References.....	104
Tables	108
Figures	110
Supporting Information.....	111
Chapter 5: Impacts of an immune challenge on the gut microbiota and growth performance of neonatal hosts: insights from an avian model system	141
Abstract	141
Introduction.....	141
Materials and methods	142
Results	144
Discussion	145
Conclusions.....	148
References.....	148
Tables	152

Figures	154
Supporting Information.....	158
Synthesis	168
Introduction.....	168
Empirical findings	168
Implications	169
Limitations.....	171
Exciting avenues for future research	171
Conclusion	174
References.....	175

Introduction

Microorganisms pervade all living systems on earth, from common habitats such as seawater and soil to unique environments that include hot springs, hydrothermal vents, polar ice caps, hypersaline settings, and extreme pH conditions (Keller & Zengler 2004). The estimated global number of archaeal and bacterial microorganisms as well as their total amount of cellular carbon is particularly staggering: $4\text{--}6 \times 10^{30}$ cells and 350–550 Pg (1 Pg = 10^{15} g), respectively (Whitman *et al.* 1998). In relative terms, archaea and bacteria contain 60–100% of the estimated total carbon and as much as ten-fold more nitrogen and phosphorus than all plants. Microorganisms thereby comprise a substantial component of biota on earth and play fundamental roles in its biogeochemical processes.

Microbes have also long been known to occur in extraordinary abundance within animal hosts (Savage 1977). The microbial assemblage living in and on a host is known as the microbiota, and their collective genetic material, the microbiome (Table 1). Although archaea, bacteria, viruses and eukaryotes comprise the microbiota, most studies on host-associated microbiota have focused on bacteria (Ley *et al.* 2008). In recent years, massively parallel sequencing (Schuster 2008) has remarkably enhanced the sequencing depth achievable for host-associated samples, and metagenomic analyses (Wooley *et al.* 2010) have revealed microbes to be fascinatingly more diverse as well as integral to the biology of host animals than previously appreciated. For example, obesity is inducible through transplant of fecal microbiota from obese mice to both lean mice and germ-free mice (Ridaura *et al.* 2013). Immunological development and function are severely defective in germ-free mice, whereby the mice are highly susceptible to infectious agents (Lee & Mazmanian 2010). Furthermore, gut microbiota have been shown to modulate murine brain development and behavior (Heijtz *et al.* 2011). In humans, populations at higher latitudes exhibit greater body masses and, perhaps partially explaining this finding, have relative abundances of bacterial phyla (increased Firmicutes and decreased Bacteroidetes) that are associated with obesity (Suzuki & Worobey 2014). Thus, microbiota harbor a profound potential to impact host phenotype across diverse host taxa and spatiotemporal scales.

Nevertheless, a striking paucity of studies in natural populations of non-human and non-rodent hosts exists (McFall-Ngai *et al.* 2013), despite the increasing relevance of knowledge on host-microbial interactions for understanding the implications of global change (e.g., Sutherst 2004). Similar to the challenges faced at the macroecological scale (Tylianakis *et al.* 2008), in order to interpret the extent to which environmental heterogeneity can impact present day microecological or metaorganismal (Bosch & McFall-Ngai 2011) function, baseline variation within these systems as well as the processes underlying their form and function must be characterized. Doing so will empower predictive models to clarify ecological and evolutionary dynamics in a constantly changing world, which is anticipated to be widely relevant to fields such as disease ecology, epidemiology, and wildlife conservation.

To contribute to the current dearth of studies on host-microbial relationships in free-living vertebrates, this dissertation focuses on the gut microbiota in a natural population of avian hosts. More specifically, this work addresses the hypothesis that variation in the avian gut microbiota incurs biologically relevant impacts on host physiological performance, and consequently, avian-associated microbiota comprise valuable model systems for generating

ecological and evolutionary insights. The neonatal vertebrate life stage is studied in particular as a developmentally critical period during which assembly of the gut microbiota can impart lifelong effects on host metabolic programming (Bäckhed 2011). Additionally, the Western Bluebird *Sialia mexicana* is developed as a free-living yet tractable study system because it is an abundant, year-round, resident passerine found in open woodlands throughout California. Western Bluebirds are cavity nesters that readily breed in manmade nest boxes, thereby supporting repeated sample collection (i.e., longitudinal sampling), robust sample sizes, and feasible experimental manipulations. The annual grasslands, foothill oak-pine woodlands, and historical nest box system of the San Joaquin Experimental Range (SJER, Madera County, CA) comprise favorable habitat for Western Bluebirds. SJER is thereby chosen as the primary field site to support the work herein.

In this dissertation, **Chapter One** provides background on host-associated microbial ecology and historical advancements in methodology that have led to the current state of the field. Likewise, it introduces the field of avian microbiology and offers a review of the existing relevant literature to date. Given that avian microbiology comprises a rapidly emerging field, comprehensive methodological guidelines do not yet exist to support field collection and laboratory analyses of avian-associated microbial samples using the latest technological advancements available (i.e., massively parallel sequencing). Consequently, **Chapter Two** develops cost-effective protocols for sample preservation, metagenomic DNA extraction, and Illumina amplicon library preparation for studies in avian microbiology using minimally invasive samples. **Chapter Three** then applies the methods developed from **Chapter Two** to characterize natural variation in the avian gut microbiota over broad spatiotemporal scales and considers the mechanistic drivers of the observed patterns. **Chapter Four** contextualizes the microbial data from **Chapter Three** with host morphological data in order to investigate associations between variation in the avian gut microbiota and host growth performance. To study resilience as well as to demonstrate the utility of avian-associated microbiota as a model system for broader questions in the life sciences, **Chapter Five** investigates the heretofore undocumented relationship between vaccine use and gut microbial ecology, with potentially widespread implications for both human and veterinary medicine. Finally, the **Conclusion** evaluates the findings of the dissertation as a whole, discusses their synergistic derivations, and outlines exciting avenues for future research.

References

- Bäckhed F (2011) Programming of host metabolism by the gut microbiota. *Annals of nutrition & metabolism*, **58 Suppl 2**, 44–52.
- Bosch TCG, McFall-Ngai MJ (2011) Metaorganisms as the new frontier. *Zoology (Jena, Germany)*, **114**, 185–190.
- Heijtz RD, Wang S, Anuar F *et al.* (2011) Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 3047–3052.
- Keller M, Zengler K (2004) Tapping into microbial diversity. *Nature Reviews Microbiology*, **2**, 141–150.
- Lee YK, Mazmanian SK (2010) Has the Microbiota Played a Critical Role in the Evolution of the Adaptive Immune System? *Science*, **330**, 1768–1773.
- Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of Mammals and Their Gut Microbes. *Science*, **320**, 1647–1651.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3229–3236.
- Ridaura VK, Faith JJ, Rey FE *et al.* (2013) Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. *Science*, **341**, 1241214.
- Savage DC (1977) Microbial ecology of the gastrointestinal tract. *Annual Reviews in Microbiology*, **31**, 107–133.
- Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nature Methods*, **5**, 16–18.
- Sutherst RW (2004) Global Change and Human Vulnerability to Vector-Borne Diseases. *Clinical Microbiology Reviews*, **17**, 136–173.
- Suzuki TA, Worobey M (2014) Geographical variation of human gut microbial composition. *Biology Letters*, **10**, 20131037.
- Tylianakis JM, Didham RK, Bascompte J, Wardle DA (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351–1363.
- Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 6578–6583.
- Wooley JC, Godzik A, Friedberg I (2010) A Primer on Metagenomics. *PLoS Computational Biology*, **6**, e1000667.

Table 1. Glossary of terms used in this dissertation.

Term	Definition
Diversity	A measure of variety present in a community, consists of richness and evenness.
Evenness	The distribution of individuals across types.
Metagenomics	A culture-independent, sequence-based analysis of total environmental, community DNA.
Metaorganism/metaorganismal	Any multicellular organism comprised of a macroscopic host and its associated microorganisms, which may include bacteria, archaea, fungi, algae, and other microscopic eukaryotes.
Microbiome	The collective genomic content of a community.
Microbiota/microbial assemblage/microbial community	The collection of microorganisms existing in the same place at the same time.
Microorganism	A microscopic organism. Examples include bacteria, fungi, archaea, and protists.
Richness	The number of different types.

Acknowledgements

I am deeply grateful for financial support from the UC Berkeley Chancellor's Fellowship, NSF Graduate Research Fellowship Program, Berkeley Natural History Museum's GK-12 Fellowship Program, American Vineyard Foundation, UC Berkeley Museum of Vertebrate Zoology, UC Berkeley Department of Integrative Biology, UC Berkeley Graduate Division, North American Bluebird Society, and Berkeley Chapter of Sigma Xi. I am thankful for the training opportunities provided by the NSF Research Coordination Network in Ecoimmunology and the Joint Genome Institute's Microbial Genomics and Metagenomics Workshop. I would also like to acknowledge Wiley-Blackwell for publishing part of my dissertation work in *Molecular Ecology Resources*, and I am grateful to anonymous reviewers for thoughtful comments on earlier versions of various parts of the dissertation.

I wish to express my utmost gratitude to my advisor, Dr. Rauri Bowie, for supporting my broad and varied research interests throughout the years. I would also like to thank my committee members, Drs. Rodrigo Almeida and Michael Shapira, for their time and critical guidance in every step of my research. Furthermore, I am thankful for the logistical support provided by so many without whom my dissertation would not have been possible: Dr. Kathryn Purcell for instrumental field resources and mentorship, including the "sisterhood of the highway dresses"; Dr. Jeanne Fair for fundamental training in field sampling and ecoimmunology; Dr. Julie Jedlicka for problem-solving laboratory challenges and moral support; Bob and Marilyn Evans for providing a heartwarming home and family through three brutally hot field seasons; Viet Cai for being the panacea to any and all research woes; Rachael Joakim, Kelsey Bain, Cheongwai Martin Wong, Nathaniel Simon, and Taylor Miller for being the best research assistants and friends that I could have hoped to find for my project; Lydia Smith for countless brainstorming sessions and testing of novel techniques in the lab; Renee Denton for kindly supporting two summers of enthusiastic field crews; the San Joaquin Experimental Range; the USFS Pacific Southwest Research Station; students and teachers of the Center for Advanced Research and Technology; and the UC Davis Genome Center. I would also like to thank my fellow doctoral students, past and present, for their invaluable friendship, inspiration, and feedback. Lastly, I thank my friends and family for their endless encouragement throughout my academic journey.

Chapter 1: A primer on avian microbiology

Abstract

Over the past decade, examination of host-associated microbiota has greatly raised awareness of the fundamental role that microbes play in the physiology, phenotypic diversity, ecology, and evolution of host organisms. However, host-microbial relationships in avian systems remain exceedingly understudied relative to those in mammalian systems. This chapter applies the term 'avian microbiology' to this emerging yet exciting interdisciplinary field. The existing relevant body of work is reviewed with the goal of encouraging avian studies that address microbiota to develop a more comprehensive understanding of the processes underlying variation in life-history traits, including but not limited to: immune response, digestion, olfaction, diet, and overall health.

Introduction

Recently, the "metaorganismal" biological unit, encompassing a host and all of its associated microbiota, has increasingly taken precedence in the concept of the "individual" and has brought to light fascinating complexity in thinking about species interactions (Bosch & McFall-Ngai 2011). To date, studies of host-associated microbiota have primarily focused on the gut microbiota of humans and laboratory mice (Nicholson *et al.* 2012; Faith *et al.* 2014; but see Nyholm and McFall-Ngai 2004), as the number of microbes found in the human gastrointestinal tract far surpasses those associated with all other body surfaces (Savage 1977). The most prominent theme among these studies has been the role of microbiota in host health and disease (Sekirov *et al.* 2010). However, technological advances in analytical methods (e.g., Caporaso *et al.* 2010; Kuczynski *et al.* 2012) now pave the way for researchers to use a comparative framework to pursue studies of host-associated microbiota in non-model systems as well as to tackle broader questions in the life sciences.

Knowledge on avian-microbial relationships remains far behind current understanding of the structure, function, and evolutionary implications of microbiota in mammalian hosts. For example, a search of Web of Science using the terms 'variation' and 'microbiota' (conducted on February 4, 2014) returned 529 matches. Of the matches, 177 studies investigated factors underlying patterns in host-associated microbial diversity, and only six of the 177 studies (less than four percent) focused on avian hosts (Fig. 1.1). Yet, birds exhibit remarkably diverse life-history traits that span various dietary (e.g. Snow 1981; Robinson & Holmes 1982), developmental (Starck & Ricklefs 1998), physiological (e.g. Olson & Owens 2005), reproductive (Jetz *et al.* 2008), migratory (Greenberg & Marra 2005) and behavioral (Revis *et al.* 2004) strategies. Accordingly, they comprise a fabulous study system to begin understanding, for example, the role that microbiota and various levels of selection play in life-history evolution.

Research opportunities abound at the exciting intersection of community ecology, microbiology, disease ecology, and ornithology. The term 'avian microbiology' is proposed here to best describe this emerging, interdisciplinary field. In this review, relevant research to date in avian microbiology is covered, and some of the motivating insights that the existing body of work has provided are highlighted. Ultimately, now is the opportune time for ornithologists to leverage their natural history knowledge and existent research efforts to incorporate synergistic questions in avian microbiology. Doing so will expand the conceptual framework underlying

current avian research and address a more complex yet comprehensive understanding of organismal biology (McFall-Ngai *et al.* 2013).

Background

Host-associated microorganisms were first observed in the 17th century when Anton van Leeuwenhoek discovered 'animalcules' in dental plaque under a microscope. Microscopic analysis and culture of microbes remained the primary tools for the study of microbiology until the late 20th century. During this time, discordance in the diversity of microbes observed microscopically versus in culture was frequently encountered, wherein the former far exceeded the latter. In 1980, Carl Woese and colleagues transformed microbial methodology through sequencing 16S ribosomal RNA (rRNA) to study bacterial phylogenetics (Fox *et al.* 1980). The 16S rRNA gene was favored as a molecular marker for a number of reasons, including its nearly universal distribution among bacteria and its informative overall length (1,500 bp). More specifically, the function of the 16S rRNA gene is highly conserved, which affords effective target sites for universal primers, yet the gene contains nine hypervariable regions across its full length, which allows species-specific differentiation. Thus, the 16S rRNA approach offered the detection sensitivity necessary to address prior discrepancies between microscopy and culturing. Thereafter, analysis of 16S rRNA genotypes became routine in microbial studies.

However, until recently, 16S rRNA sequence analyses of complex microbial communities required isolation via cloning and Sanger sequencing (Sanger *et al.* 1977) of all unique genotypes, a time-intensive process that often could not yield satisfactory saturation of rarefaction curves for most samples. In 2006, the advent of massively parallel sequencing (MPS) revolutionized microbial analysis (Gill *et al.* 2006). Contemporary Sanger sequencing involves incorporation of chain-terminating, labelled dideoxynucleotides and capillary electrophoresis to produce sequence chromatograms (Prober *et al.* 1987). It is mainly limited by scalability, as a maximum of only 96 samples may be analyzed per run. In contrast, MPS drastically reduces the reaction volume of sequencing reactions and enables multiplexing of several million templates per run (Schuster 2008). Consequently, the high throughput capacity of MPS makes it feasible to generate orders of magnitude more sequence data per sample than is ever possible with Sanger sequencing. It has supported interrogation of diverse host-microbial associations and interactions with fine-scale resolution (Lamendella *et al.* 2012), and it has given rise to an explosion of published studies, currently numbering hundreds per year.

There is a growing body of research on the assembly of host-associated microbiota and the effects of microbiota on host biology. Deterministic processes including diet, environmental exposure, host genotype, and interspecific microbial interactions as well as stochastic processes are known to underlie host-microbial assembly (Spor *et al.* 2011; Costello *et al.* 2012). Likewise, microbiota dynamically impact host physiology, from gut anatomy to nutrient metabolism, immune function, overall growth and development, as well as susceptibility to pathogens and non-infectious diseases (Maslowski & Mackay 2011; Sommer & Bäckhed 2013). Here, avian studies are reviewed, which have primarily covered eight areas: i) microbial infection of eggs, ii) bacterial colonization of nestlings, iii) variation in the gut microbiota of adult birds, iv) genetic versus environmental determinants of biogeographical variation, v) sexual transmission of bacteria, vi) feather-degrading bacteria, vii) microbial effects on host physiology, and viii) the role of birds in pathogen and disease ecology.

The majority of avian studies (Lombardo *et al.* 1996; Mills *et al.* 1999; Stewart & Rambo 2000; Hupton *et al.* 2003; Blanco *et al.* 2006, 2009; Shawkey *et al.* 2007; Skov *et al.* 2008; Radhouani *et al.* 2012; Oravcova *et al.* 2013) have relied on culture-dependent analyses, which afford limited microbial characterization as less than one percent of all bacteria can be cultured (Amann *et al.* 1995). Thus, there has been strong motivation to employ culture-independent methods, including: G+C profiling (Apajalahti *et al.* 2001), EcoPlates (Maul *et al.* 2005), analysis of fragment length polymorphisms of rRNA genes or their intergenic spacer regions (Knarreborg *et al.* 2002; Lucas & Heeb 2005; Klomp *et al.* 2008; van Dongen *et al.* 2013), sequencing clone libraries of 16S rRNA gene amplifications (Lu *et al.* 2003; Shawkey *et al.* 2005, 2006; Jeter *et al.* 2009; Grizard *et al.* 2014), and microarray analysis (Shawkey *et al.* 2009). Few studies have incorporated MPS (Singh *et al.* 2012; Stanley *et al.* 2012; Vasaï *et al.* 2014). Nevertheless, these prior studies have offered valuable insights into avian-associated microbial ecology.

(i) Microbial infection of eggs

Microorganisms enter eggs through eggshell pores that are uncovered by the shell cuticle, and if allowed to persist, microbial overgrowth can render eggs unviable (Board 1966). In the presence of water, microbes are capable of digesting the cuticle, unplugging substantial numbers of eggshell pores, and increasing the susceptibility of eggs to microbial penetration (Board *et al.* 1979). Exposed eggs are colonized by microbiota within one to five days, which comprises a proposed constraint on the onset of incubation and may contribute to the habit of partial incubation among birds (Cook *et al.* 2003). Incubation of eggs protects them from humidity, establishes suboptimal egg temperatures for microbial growth (Board & Fuller 1994), and thereby inhibits the growth of bacteria on eggshells (Cook *et al.* 2005; Shawkey *et al.* 2009; D'Alba *et al.* 2010a). Fungi also comprise a substantial component of the microbiota on eggshells, and incubation reduces the diversity of both bacterial and fungal microbiota (Grizard *et al.* 2014). Interestingly, within brood parasite systems, parasitized nests and host eggs harbor greater bacterial loads than unparasitized nests and parasitic eggs, respectively (Soler *et al.* 2011). Deposition of antimicrobial proteins in eggs also serves to inhibit microbial infection, and incubation generates egg temperatures that are conducive to optimal antimicrobial protein function (Board & Fuller 1994). Variation in the deposition of antimicrobial proteins in eggs occurs, whereby females mated to more attractive or monogamous males deposit more antimicrobial proteins than those mated to less attractive or polygynous males (D'Alba *et al.* 2010b). Taken altogether, control of microbial infections of eggs appears highly relevant to the evolution of a range of life history traits from incubation behavior to brood parasitism and immune function.

(ii) Bacterial colonization of young birds

The gastrointestinal tract of embryos is generally presumed to be internally sterile such that birds are immediately colonized from the surrounding environment upon hatching (Wielen *et al.* 2002). Microbial sources present in the environment include parental saliva, food, nesting material, and ectoparasites. Cloacal microbiota establish in Tree Swallow *Tachycineta bicolor* nestlings soon after hatching (two days), and subsequently, both microbial species richness and total abundance increase with age (Mills *et al.* 1999). Similarly, significant changes in ileal and cecal microbial diversity occur in maturing broiler chickens from age three to 49 days, wherein

ileal and cecal composition begin to diverge after 14 days of age (Lu *et al.* 2003). The cloacal microbiota of older Black-legged Kittiwake *Rissa tridactyla* chicks are also more diverse than assemblages of younger chicks, but interestingly, when compared to the microbiota of adult kittiwakes, those of the chicks are significantly more diverse though less abundant (van Dongen *et al.* 2013). Thus overall, a general increase in microbial diversity and divergence between body sites (e.g., ileal vs. cecal, oral vs. cloacal) occurs within chicks post-hatching (Mills *et al.* 1999; Lu *et al.* 2003), but more comparisons between microbiota of young birds and their parents are needed to determine the degree of transience versus overlap and stability that occurs in microbial assemblages over developmental timeframes.

(iii) Variation in the gut microbiota of adults

Microbiota in Chicken *Gallus gallus* ceca (Zhu *et al.* 2002), Turkey *Meleagris gallopavo* ceca (Scupham *et al.* 2008), Ostrich *Strutio camelus* ceca (Matsui *et al.* 2010), Mule Duck (hybrid Muscovy *Cairina moschata* X Mallard *Anas platyrhynchos*) ilea and ceca (Vasai *et al.* 2014), Hoatzin *Opisthocomus hoazin* crops (Godoy-Vitorino *et al.* 2008), gull feces (Lu *et al.* 2008), and Canada Goose *Branta canadensis* feces (Lu *et al.* 2009) are dominated by Firmicutes and Bacteroidetes. Cecal microbiota of Emus *Dromaius novaehollandiae* (Bennett *et al.* 2013) as well as choanal and cloacal microbiota of parrots (Xenoulis *et al.* 2010; Waite *et al.* 2012) are dominated by Bacteroidetes and Proteobacteria whereas fecal microbiota of Adelie Penguins *Pygoscelis adeliae* are dominated by Firmicutes and Actinobacteria (Banks *et al.* 2009).

Although most studies principally analyzed samples from one or two gut sites (e.g., cecal, cloacal, or fecal), analysis of microbiota along the entire gastrointestinal tract of chickens revealed high variation in microbial diversity between sites, with a predominance of Firmicutes (Gong *et al.* 2007). The chicken crop, gizzard, duodenum, jejunum, and ileum exhibit low OTU richness and are either predominantly or exclusively dominated by lactobacilli whereas the ceca harbor the highest OTU richness and contain mostly clostridia-related taxa (Gong *et al.* 2007). The fluidity, smooth muscle activity, relative rates of digesta flow, and other microenvironmental differences among the various gastrointestinal regions likely contribute to the variation in microbiota observed along the length of the gut, but more detailed, mechanistic studies to evaluate these factors are currently lacking.

Comparative analysis of mammalian gut microbiota across 80 species has revealed consistent dominance by two main phyla, *Firmicutes* and *Bacteroidetes*, and this pattern is thought to be due to the dependence on carbohydrate energy sources in many mammalian diets (Ley *et al.* 2008). Interestingly, *Firmicutes* and *Bacteroidetes* also figure prominently in the gut microbiota of avian species with known carbohydrate-rich diets (Zhu *et al.* 2002; Lu *et al.* 2008, 2009; Scupham *et al.* 2008; Godoy-Vitorino *et al.* 2008; Matsui *et al.* 2010; Vasai *et al.* 2014). Relative to mammalian microbiota, unique genera and species, but no unique phyla, have been identified in birds using 16S rRNA sequencing.

(iv) Host genetic versus environmental determinants of biogeographical variation

Diet plays a dominant role in shaping ileal (Lu *et al.* 2003) and cecal (Apajalahti *et al.* 2001) bacterial communities in broiler chicks. Likewise, a shared environment dictates similarity in cloacal bacterial communities among partially cross-fostered Great Tit *Parus major* and Eurasian Blue Tit *Cyanistes caeruleus* nestlings (Lucas & Heeb 2005), Tree Swallow nestlings

(Lombardo *et al.* 1996; Mills *et al.* 1999), as well as adult Spotted Towhees *Pipilo maculatus* (Klomp *et al.* 2008). In contrast, the interindividual similarity of microbiota in Adelie Penguins is negatively associated with both host genetic and geographic distances (Banks *et al.* 2009). Furthermore, Great Spotted Cuckoo *Clamator glandarius* and Eurasian Magpie *Pica pica* nestlings sharing the same nest and parents exhibit significantly different microbiota (Ruiz-Rodríguez *et al.* 2009). Therefore, evidence consistent with as well as discordant with the role of environment in driving variation in microbiota has been reported. Intriguingly, although microbial similarity between co-nesting Great Tits and Blue Tits suggest a greater role for environment over host genetics, distinct microbiota between co-nesting Great Spotted Cuckoos and Eurasian Magpies indicate a greater role for host genetics. Accordingly, the relative role of intrinsic versus extrinsic factors in driving variation in avian-associated microbiota remains an exciting topic to be clarified.

(v) Sexual transmission of bacteria

The avian cloaca forms the conduit for both excretion and gamete transfer, which provides the opportunity for intestinal microbes to be incorporated into ejaculate and thereby be sexually transmitted. Cloacal assemblages are significantly correlated among mated pairs of House Sparrows *Passer domesticus* (Stewart & Rambo 2000). Experimental infection of the cloaca of captive Zebra Finches *Taeniopygia guttata* with a single bacterial strain results in predominant transmission from males to females among mated pairs (Kulkarni & Heeb 2007). Likewise, Red-winged Blackbird *Agelaius phoeniceus* ejaculate contains a highly diverse yet low abundance of bacteria (Hupton *et al.* 2003), and mating exposes female Red-winged Blackbirds to bacteria in male semen (Westneat & Birch Rambo 2000). Bacteria are also transmitted during mating among monogamous Black-legged Kittiwakes, with subsequent shifts in female cloacal assemblages (White *et al.* 2010). Interestingly, experimental blocking of insemination during mating results in the divergence of microbiota between mated pairs of kittiwakes, and the diversity of female cloacal assemblages decline over time, exhibiting extinction of mate-shared bacteria and resilience in reestablishment of pre-copulatory microbiota (White *et al.* 2010). Given that the transmitted microbes can be composed of mutualists, commensals, and/or parasites, it is of interest to determine the potential fitness consequences of sexually transmitted microbiota and whether this process may comprise selective pressure underlying the maintenance of monogamy or the evolution of post-copulatory sperm ejection.

(vi) Feather-degrading bacteria

Bacteria are capable of using feathers as habitat substrates, and feather-degrading bacteria (*Bacillus licheniformis*) have been isolated from the plumage of many species of birds (32 of 83 surveyed avian species; Burt & Ichida 1999). Keratinolytic bacteria degrade white, non-melanic feathers more rapidly than black, melanic feathers (Goldstein *et al.* 2004), and bacterial degradation of feathers occurs more rapidly in humid locales than in arid localities (Burt & Ichida 1999). Darkly colored Song Sparrows *Melospiza melodia* are found in humid habitats whereas pale sparrows occur where humidity is low (Goldstein *et al.* 2004). Accordingly, Burt & Ichida (2004) suggested that microbially mediated selection may underlie Gloger's Rule, in which more heavily pigmented phenotypes tend to be found in more humid environments.

In Pied Flycatchers, unmelanized and melanized feather areas serve as honest signals under sexual selection, and Ruiz-De-Castañeda *et al.* (2012) suggested that their differential degradability by microbes indicates a role for feather-degrading bacteria in avian sexual selection. A similar argument was made previously in Eastern Bluebirds *Sialia sialis* in which males with brighter structural plumage color exhibit greater competitive ability and higher reproductive success than duller males (Siefferman & Hill 2003, 2005). Feather-degrading bacteria alter the spectral properties of male Eastern Bluebird feathers through increases in brightness and purity as well as decreases in ultraviolet chroma (Shawkey *et al.* 2007). Since uropygial oil affects bacterial growth (Shawkey *et al.* 2003), Shawkey *et al.* (2007) suggested that plumage color may indicate the ability of individual males to regulate feather microbiota. However, removal of the uropygial gland in House Sparrows (Czirják *et al.* 2013) and Mallard Ducks (Giraudeau *et al.* 2013) effect no change in the abundance of feather degrading bacteria. Thus, the role of uropygial oil in mediating sexual selection through control of feather degrading bacteria remains unsupported.

(vii) Microbial effects on avian physiology

Kohl (2012) provided an extensive review of the physiological effects of the avian gut microbiota. Briefly, fibrolytic and cellulolytic microbes aid in the utilization of plant polysaccharides by ostriches (Matsui *et al.* 2010), hoatzins (Grajal *et al.* 1989), and Rock Doves *Columba livia* (Shetty *et al.* 1990). Saccharolytic microbes are found in the gut of most Galliformes (Vispo & Karasov 1997). Furthermore, microbes that mediate uric acid metabolism have been isolated from chickens, turkeys, guineafowls, ducks, pheasants and Anna's Hummingbirds *Calypte anna* (Barnes 1972; Preest *et al.* 2003). Interestingly, microbial colonization of the chicken gastrointestinal tract corresponds to decreased absorption of glucose and vitamins (Ford & Coates 1971). Although gut microbes have been shown to substantially influence immunological development and function in mammals (Hooper *et al.* 2012), a striking dearth of research exists on microbial interactions with the avian immune system. Regarding microbial effects on host condition, bacterial load does not affect fledging success of Tree Swallows (Mills *et al.* 1999), but gram-negative enteric lactose fermenters positively correlate with a greater degree of wing asymmetry in Tree Swallow nestlings. Similarly, the presence of *Enterococcus faecium* positively associates with nestling mass and size in Pied Flycatchers *Ficedula hypoleuca* (Moreno *et al.* 2003). Thus, microbes are known to significantly associate with avian metabolic function and morphology, but more comprehensive assessments of avian-associated microbiota and their relationships with physiological performance have yet to be pursued to characterize both within-species and between-species variation.

(viii) The role of birds in pathogen and disease ecology

The relationship between wild birds and public health has garnered significant attention in recent years, especially due to emerging and re-emerging zoonoses (diseases transmittable between animals and humans) such as West Nile and avian influenza. West Nile virus has been isolated from over 200 bird species (McLean 2006), and influenza viruses, from more than 100 species across 15 different orders (Wallensten 2007), exemplifying the importance of birds for the study of pathogen and disease ecology. Numerous additional, high priority, pathogenic

microorganisms have been detected in various bird species, including *Salmonella* spp., *Campylobacter* spp., and *Listeria* spp. in Ring-billed Gulls *Larus delawarensis* (Quessy & Messier 1992); *Helicobacter canadensis* in Barnacle Geese *Branta leucopsis* and Canada Geese *B. canadensis* (Waldenström *et al.* 2003); as well as *Salmonella* spp. and *Escherichia coli* in free-living passerines (Morishita *et al.* 1998). Ginsberg *et al.* (2005) identified American Robins *Turdus migratorius* as having the highest reservoir competence for *Borrelia burgdorferi* among Gray Catbirds *Dumetella carolinensis*, Eastern Towhees *Pipilo erythrophthalmus*, Song Sparrows, Northern Cardinals *Cardinalis cardinalis*, and Brown Thrashers *Toxostoma rufum*. Similarly, waterfowl are highly competent reservoirs for influenza viruses (Ito *et al.* 1995). Although migratory waterfowl and shorebirds have been posited to support intercontinental exchange of influenza viruses, phylogenetic analyses revealed the exchange of viruses between Eurasian and American clades to occur infrequently (Krauss *et al.* 2007). Despite the findings of these studies, the role of birds in the propagation, transmission, and spread of human pathogens and disease is largely understudied (Reed *et al.* 2003; Benskin *et al.* 2009).

Conclusion

Taken altogether, it is clear that microbiota are pervasively relevant to avian reproduction, metabolism, phenotypic diversity, condition, and evolution. Likewise, birds play a substantial role in harboring and disseminating microbes of public concern. Nevertheless, the field of avian microbiology remains overflowing with research opportunities to expand upon the implications of previous research findings. For example, what is the relationship between nest attentiveness and deposition of antimicrobial peptides in eggs? Is the relationship between microbial diversity and avian growth performance linear? What factors mediate differential reservoir competence of avian hosts for various pathogens? As methodological, technological, and computational advances progress at unprecedented rates, now is an exciting time to take advantage of the available tools to pursue the plethora of unanswered questions.

References

- Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiology and Molecular Biology Reviews*, **59**, 143.
- Apajalahti JHA, Kettunen A, Bedford MR, Holben WE (2001) Percent G C profiling accurately reveals diet-related differences in the gastrointestinal microbial community of broiler chickens. *Applied and Environmental Microbiology*, **67**, 5656.
- Banks JC, Cary SC, Hogg ID (2009) The phylogeography of Adelie penguin faecal flora. *Environmental Microbiology*, **11**, 577–588.
- Barnes EM (1972) The avian intestinal flora with particular reference to the possible ecological significance of the cecal anaerobic bacteria. *The American journal of clinical nutrition*, **25**, 1475–1479.
- Bennett DC, Tun HM, Kim JE, Leung FC, Cheng KM (2013) Characterization of cecal microbiota of the emu (*Dromaius novaehollandiae*). *Veterinary Microbiology*, **166**, 304–310.
- Benskin CMH, Wilson K, Jones K, Hartley IR (2009) Bacterial pathogens in wild birds: a review of the frequency and effects of infection. *Biological Reviews*, **84**, 349–373.
- Blanco G, Lemus JÚSA, Grande J (2006) Faecal bacteria associated with different diets of wintering red kites: influence of livestock carcass dumps in microflora alteration and pathogen acquisition. *Journal of Applied Ecology*, **43**, 990–998.
- Blanco G, Lemus JA, Grande J (2009) Microbial pollution in wildlife: Linking agricultural manuring and bacterial antibiotic resistance in red-billed choughs. *Environmental research*, **109**, 405–412.
- Board RG (1966) The Course of Microbial Infection of the Hen's Egg. *Journal of Applied Bacteriology*, **29**, 319–341.
- Board RG, Fuller R (1994) *Microbiology of the Avian Egg*. Springer.
- Board RG, Loseby S, Miles VR (1979) A note on microbial growth on hen egg-shells. *British Poultry Science*.
- Bosch TCG, McFall-Ngai MJ (2011) Metaorganisms as the new frontier. *Zoology (Jena, Germany)*, **114**, 185–190.
- Burt EH, Ichida JM (1999) Occurrence of Feather-Degrading Bacilli in the Plumage of Birds. *The Auk*, **116**, 364–372.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, **7**, 335–336.
- Cook MI, Beissinger SR, Toranzos GA, Arendt WJ (2005) Incubation reduces microbial growth on eggshells and the opportunity for trans-shell infection. *Ecology Letters*, **8**, 532–537.
- Cook MI, Beissinger SR, Toranzos GA, Rodriguez RA, Arendt WJ (2003) Trans-shell infection by pathogenic micro-organisms reduces the shelf life of non-incubated bird's eggs: a constraint on the onset of incubation? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 2233–2240.
- Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA (2012) The Application of Ecological Theory Toward an Understanding of the Human Microbiome. *Science*, **336**, 1255–1262.
- Czirják GÁ, Pap PL, Vágási CI *et al.* (2013) Preen gland removal increases plumage bacterial load but not that of feather-degrading bacteria. *Naturwissenschaften*, **100**, 145–151.

- D'Alba L, Oborn A, Shawkey MD (2010a) Experimental evidence that keeping eggs dry is a mechanism for the antimicrobial effects of avian incubation. *Naturwissenschaften*, **97**, 1089–1095.
- D'Alba L, Shawkey MD, Korsten P *et al.* (2010b) Differential deposition of antimicrobial proteins in blue tit (*Cyanistes caeruleus*) clutches by laying order and male attractiveness. *Behavioral Ecology and Sociobiology*, **64**, 1037–1045.
- Van Dongen W, White J, Brandl HB *et al.* (2013) Age-related differences in the cloacal microbiota of a wild bird species. *BMC Ecology*, **13**, 11.
- Faith JJ, Ahern PP, Ridaura VK, Cheng J, Gordon JI (2014) Identifying Gut Microbe–Host Phenotype Relationships Using Combinatorial Communities in Gnotobiotic Mice. *Science Translational Medicine*, **6**, 220ra11–220ra11.
- Ford DJ, Coates ME (1971) Absorption of glucose and vitamins of the B complex by germ-free and conventional chicks. *The Proceedings of the Nutrition Society*, **30**, 10A–11A.
- Fox GE, Stackebrandt E, Hespell RB *et al.* (1980) The phylogeny of prokaryotes. *Science*, **209**, 457–463.
- Gill SR, Pop M, DeBoy RT *et al.* (2006) Metagenomic analysis of the human distal gut microbiome. *Science*, **312**, 1355.
- Ginsberg HS, Buckley PA, Balmforth MG *et al.* (2005) Reservoir competence of native North American birds for the lyme disease spirochete, *Borrelia burgdorferi*. *Journal of Medical Entomology*, **42**, 445–449.
- Giraudeau M, Czirják GÁ, Duval C *et al.* (2013) Effect of preen oil on plumage bacteria: An experimental test with the mallard. *Behavioural Processes*, **92**, 1–5.
- Godoy-Vitorino F, Ley RE, Gao Z *et al.* (2008) Bacterial Community in the Crop of the Hoatzin, a Neotropical Folivorous Flying Bird. *Applied and Environmental Microbiology*, **74**, 5905–5912.
- Goldstein G, Flory KR, Browne BA *et al.* (2004) Bacterial degradation of black and white feathers. *The Auk*, **121**, 656.
- Gong J, Si W, Forster RJ *et al.* (2007) 16S rRNA gene-based analysis of mucosa-associated bacterial community and phylogeny in the chicken gastrointestinal tracts: from crops to ceca. *FEMS Microbiology Ecology*, **59**, 147–157.
- Grajal A, Strahl SD, Parra R, Gloria Dominguez M, Neher A (1989) Foregut fermentation in the hoatzin, a neotropical leaf-eating bird. *Science*, **245**, 1236–1238.
- Greenberg R, Marra PP (2005) *Birds of Two Worlds: The Ecology and Evolution of Migration*. JHU Press.
- Grizard S, Dini-Andreote F, Tieleman BI, Salles JF (2014) Dynamics of bacterial and fungal communities associated with eggshells during incubation. *Ecology and Evolution*, **4**, 1140–1157.
- Hooper LV, Littman DR, Macpherson AJ (2012) Interactions Between the microbiota and the immune system. *Science*, **336**, 1268–1273.
- Hupton G, Portocarrero S, Newman M, Westneat DF (2003) Bacteria in the reproductive tracts of red-winged blackbirds. *The Condor*, **105**, 453–464.
- Ito T, Okazaki K, Kawaoka Y *et al.* (1995) Perpetuation of influenza A viruses in Alaskan waterfowl reservoirs. *Archives of Virology*, **140**, 1163–1172.

- Jeter SN, McDermott CM, Bower PA *et al.* (2009) Bacteroidales diversity in Ring-billed gulls (*Larus delawarensis*) residing at Lake Michigan beaches. *Applied and Environmental Microbiology*, **75**, 1525.
- Jetz W, Sekercioglu CH, Böhning-Gaese K (2008) The worldwide variation in avian clutch size across species and space. *PLoS Biology*, **6**, e303.
- Klomp JE, Murphy MT, Smith SB *et al.* (2008) Cloacal microbial communities of female spotted towhees *Pipilo maculatus*: microgeographic variation and individual sources of variability. *Journal of Avian Biology*, **39**, 530–538.
- Knarreborg A, Simon MA, Engberg RM, Jensen BB, Tannock GW (2002) Effects of dietary fat source and subtherapeutic levels of antibiotic on the bacterial community in the ileum of broiler chickens at various ages. *Applied and Environmental Microbiology*, **68**, 5918.
- Kohl KD (2012) Diversity and function of the avian gut microbiota. *Journal of Comparative Physiology. B, Biochemical, Systemic, and Environmental Physiology*, **182**, 591–602.
- Krauss S, Obert CA, Franks J *et al.* (2007) Influenza in migratory birds and evidence of limited intercontinental virus exchange. *PLoS Pathogens*, **3**, e167.
- Kuczynski J, Lauber CL, Walters WA *et al.* (2012) Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics*, **13**, 47–58.
- Kulkarni S, Heeb P (2007) Social and sexual behaviours aid transmission of bacteria in birds. *Behavioural Processes*, **74**, 88–92.
- Lamendella R, VerBerkmoes N, Jansson JK (2012) “Omics” of the mammalian gut – new insights into function. *Current Opinion in Biotechnology*, **23**, 491–500.
- Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of mammals and their gut microbes. *Science*, **320**, 1647–1651.
- Lombardo MP, Thorpe PA, Cichewicz R *et al.* (1996) Communities of cloacal bacteria in tree swallow families. *The Condor*, **98**, 167–172.
- Lu J, Domingo JWS, Hill S, Edge TA (2009) Microbial diversity and host-specific sequences of Canada Goose feces. *Applied and Environmental Microbiology*, **75**, 5919.
- Lu J, Domingo JWS, Lamendella R, Edge T, Hill S (2008) Phylogenetic diversity and molecular detection of bacteria in gull feces. *Applied and Environmental Microbiology*, **74**, 3969.
- Lu J, Idris U, Harmon B *et al.* (2003) Diversity and succession of the intestinal bacterial community of the maturing broiler chicken. *Applied and Environmental Microbiology*, **69**, 6816.
- Lucas FS, Heeb P (2005) Environmental factors shape cloacal bacterial assemblages in great tit *Parus major* and blue tit *P. caeruleus* nestlings. *Journal of Avian Biology*, **36**, 510–516.
- Maslowski KM, Mackay CR (2011) Diet, gut microbiota and immune responses. *Nature Immunology*, **12**, 5–9.
- Matsui H, Kato Y, Chikaraishi T *et al.* (2010) Microbial diversity in ostrich ceca as revealed by 16S ribosomal RNA gene clone library and detection of novel Fibrobacter species. *Anaerobe*, **16**, 83–93.
- Maul JD, Gandhi JP, Farris JL (2005) Community-level physiological profiles of cloacal microbes in songbirds (Order: Passeriformes): variation due to host species, host diet, and habitat. *Microbial Ecology*, **50**, 19–28.

- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3229–3236.
- McLean RG (2006) West Nile Virus in North American Birds. *Ornithological Monographs*, 44–64.
- Mills TK, Lombardo MP, Thorpe PA (1999) Microbial colonization of the cloacae of nestling tree swallows. *The Auk*, **116**, 947–956.
- Moreno J, Briones V, Merino S *et al.* (2003) Beneficial effects of cloacal bacteria on growth and fledging size in nestling Pied Flycatchers (*Ficedula hypoleuca*) in Spain. *The Auk*, **120**, 784–790.
- Morishita T, Aye P, Ley E, Harr B (1998) Survey of pathogens and blood parasites in free-living passerines. *Avian Diseases*, **43**, 549–552.
- Nicholson JK, Holmes E, Kinross J *et al.* (2012) Host-gut microbiota metabolic interactions. *Science*, **336**, 1262–1267.
- Nyholm SV, McFall-Ngai M (2004) The winnowing: establishing the squid–*Vibrio* symbiosis. *Nature Reviews Microbiology*, **2**, 632–642.
- Olson VA, Owens IPF (2005) Interspecific variation in the use of carotenoid-based coloration in birds: diet, life history and phylogeny. *Journal of Evolutionary Biology*, **18**, 1534–1546.
- Oravcova V, Ghosh A, Zurek L *et al.* (2013) Vancomycin-resistant enterococci in rooks (*Corvus frugilegus*) wintering throughout Europe. *Environmental Microbiology*, **15**, 548–556.
- Preest MR, Folk DG, Beuchat CA, del Rio CM (2003) Decomposition of nitrogenous compounds by intestinal bacteria in hummingbirds. *The Auk*, **120**, 1091–1101.
- Prober JM, Trainor GL, Dam RJ *et al.* (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, **238**, 336–341.
- Quessy S, Messier S (1992) Prevalence of *Salmonella* spp., *Campylobacter* spp. and *Listeria* spp. in Ring-billed Gulls (*Larus delawarensis*). *Journal of Wildlife Diseases*, **28**, 526–531.
- Radhouani H, Poeta P, Goncalves A *et al.* (2012) Wild birds as biological indicators of environmental pollution: antimicrobial resistance patterns of *Escherichia coli* and enterococci isolated from common buzzards (*Buteo buteo*). *Journal of Medical Microbiology*, **61**, 837–843.
- Reed KD, Meece JK, Henkel JS, Shukla SK (2003) Birds, migration and emerging zoonoses: West Nile virus, Lyme disease, influenza A and enteropathogens. *Clinical Medicine & Research*, **1**, 5–12.
- Revis HC, Waller DA, Brittingham M (2004) Bactericidal and fungicidal activity of ant chemicals on feather parasites: an evaluation of anting behavior as a method of self-medication in songbirds. *The Auk*, **121**, 1262–1268.
- Robinson SK, Holmes RT (1982) Foraging behavior of forest birds: the relationships among search tactics, diet, and habitat structure. *Ecology*, **63**, 1918.
- Ruiz-De-Castañeda R, BURTT Jr EH, González-Braojos S, Moreno J (2012) Bacterial degradability of an intrafeather unmelanized ornament: a role for feather-degrading bacteria in sexual selection? *Biological Journal of the Linnean Society*, **105**, 409–419.
- Ruiz-Rodríguez M, Lucas FS, Heeb P, Soler JJ (2009) Differences in intestinal microbiota between avian brood parasites and their hosts. *Biological Journal of the Linnean Society*, **96**, 406–414.

- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, **74**, 5463–5467.
- Savage DC (1977) Microbial ecology of the gastrointestinal tract. *Annual Review of Microbiology*, **31**, 107–133.
- Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nature Methods*, **5**, 16–18.
- Scupham AJ, Patton TG, Bent E, Bayles DO (2008) Comparison of the cecal microbiota of domestic and wild turkeys. *Microbial Ecology*, **56**, 322–331.
- Sekirov I, Russell SL, Antunes LCM, Finlay BB (2010) Gut microbiota in health and disease. *Physiological Reviews*, **90**, 859–904.
- Shawkey MD, Firestone MK, Brodie EL, Beissinger SR (2009) Avian incubation inhibits growth and diversification of bacterial assemblages on eggs. *PLoS ONE*, **4**, e4522.
- Shawkey MD, Hussain MJ, Strong AL *et al.* (2006) Use of culture-independent methods to compare bacterial assemblages on feathers of Crested and Least Auklets (*Aethia cristatella* and *Aethia pusilla*) with those of passerines. *Waterbirds*, **29**, 507–511.
- Shawkey MD, Mills KL, Dale C, Hill GE (2005) Microbial diversity of wild bird feathers revealed through culture-based and culture-independent techniques. *Microbial Ecology*, **50**, 40–47.
- Shawkey MD, Pillai SR, Hill GE (2003) Chemical warfare? Effects of uropygial oil on feather-degrading bacteria. *Journal of Avian Biology*, **34**, 345–349.
- Shawkey MD, Pillai SR, Hill GE, Siefferman LM, Roberts SR (2007) Bacteria as an agent for change in structural plumage color: correlational and experimental evidence. *The American Naturalist*, **169**, S112–S121.
- Shetty S, Sridhar KR, Shenoy KB, Hedge SN (1990) Observations on bacteria associated with pigeon crop. *Folia Microbiologica*, **35**, 240–244.
- Siefferman L, Hill GE (2003) Structural and melanin coloration indicate parental effort and reproductive success in male eastern bluebirds. *Behavioral Ecology*, **14**, 855–861.
- Siefferman L, Hill GE (2005) UV-blue structural coloration and competition for nestboxes in male eastern bluebirds. *Animal Behaviour*, **69**, 67–72.
- Singh KM, Shah T, Deshpande S *et al.* (2012) High through put 16S rRNA gene-based pyrosequencing analysis of the fecal microbiota of high FCR and low FCR broiler growers. *Molecular Biology Reports*, **39**, 10595–10602.
- Skov MN, Madsen JJ, Rahbek C *et al.* (2008) Transmission of *Salmonella* between wildlife and meat-production animals in Denmark. *Journal of Applied Microbiology*, **105**, 1558–1568.
- Snow DW (1981) Tropical frugivorous birds and their food plants: A world survey. *Biotropica*, **13**, 1–14.
- Soler JJ, Peralta-Sánchez JM, Martínez-Bueno M *et al.* (2011) Brood parasitism is associated with increased bacterial contamination of host eggs: bacterial loads of host and parasitic eggs. *Biological Journal of the Linnean Society*, **103**, 836–848.
- Sommer F, Bäckhed F (2013) The gut microbiota — masters of host development and physiology. *Nature Reviews Microbiology*, **11**, 227–238.
- Spor A, Koren O, Ley R (2011) Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology*, **9**, 279–290.

- Stanley D, Denman SE, Hughes RJ *et al.* (2012) Intestinal microbiota associated with differential feed conversion efficiency in chickens. *Applied Microbiology and Biotechnology*, **96**, 1361–1369.
- Starck JM, Ricklefs RE (1998) *Avian Growth and Development: Evolution Within the Altricial-precocial Spectrum*. Oxford University Press.
- Stewart R, Rambo TB (2000) Cloacal microbes in house sparrows. *The Condor*, **102**, 679–684.
- Vasai F, Brugirard Ricaud K, Bernadet MD *et al.* (2014) Overfeeding and genetics affect the composition of intestinal microbiota in *Anas platyrhynchos* (Pekin) and *Cairina moschata* (Muscovy) ducks. *FEMS Microbiology Ecology*, **87**, 204–216.
- Vispo C, Karasov WH (1997) The Interaction of avian gut microbes and their host: an elusive symbiosis. In: *Gastrointestinal Microbiology* Chapman & Hall Microbiology Series. (eds Mackie RI, White BA), pp. 116–155. Springer US.
- Waite DW, Deines P, Taylor MW (2012) Gut microbiome of the critically endangered New Zealand parrot, the Kakapo (*Strigops habroptilus*). *PLoS ONE*, **7**, e35803.
- Waldenström J, On SLW, Ottvall R *et al.* (2003) Avian reservoirs and zoonotic potential of the emerging human pathogen *Helicobacter canadensis*. *Applied and Environmental Microbiology*, **69**, 7523–7526.
- Wallensten A (2007) Influenza virus in wild birds and mammals other than man. *Microbial Ecology in Health and Disease*, **19**, 122–139.
- Westneat DF, Birch Rambo T (2000) Copulation exposes female Red-winged Blackbirds to bacteria in male semen. *Journal of Avian Biology*, **31**, 1–7.
- White J, Mirleau P, Danchin E *et al.* (2010) Sexually transmitted bacteria affect female cloacal assemblages in a wild bird. *Ecology Letters*, **13**, 1515–1524.
- Wielen PWJJ, Keuzenkamp DA, Lipman LJA, Knapen F, Biesterveld S (2002) Spatial and temporal variation of the intestinal bacterial community in commercially raised broiler chickens during growth. *Microbial Ecology*, **44**, 286–293.
- Xenoulis PG, Gray PL, Brightsmith D *et al.* (2010) Molecular characterization of the cloacal microbiota of wild and captive parrots. *Veterinary Microbiology*, **146**, 320–325.
- Zhu XY, Zhong T, Pandya Y, Joerger RD (2002) 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. *Applied and Environmental Microbiology*, **68**, 124–137.

Figures

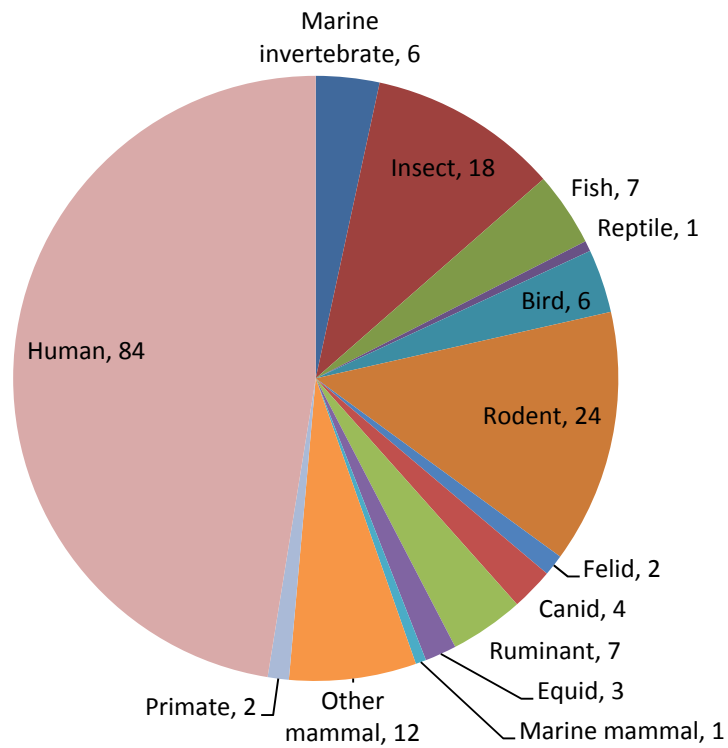


Fig. 1.1. Published studies investigating the factors underlying variation in host-associated microbial diversity, separated by focal host organism. Studies were identified through Web of Science via the search terms 'variation' and 'microbiota' on February 4, 2014.

Chapter 2: Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for fecal and swab samples

Abstract

Massively parallel sequencing (MPS) technology has extraordinarily enhanced the scope of research in the life sciences, and the goal of this study was to broaden the application of MPS to systems that were previously difficult to study. Protocols for processing fecal and swab samples into amplicon libraries amenable to Illumina sequencing are presented. A novel metagenomic DNA extraction approach using solid phase reversible immobilization (SPRI) bead technology was developed and tested on fecal and swab samples collected from Western Bluebirds *Sialia mexicana*. The performance of the SPRI-based extraction protocol was compared with that of the Mo Bio PowerSoil Kit, the current standard for the Human Microbiome Project and the Earth Microbiome Project. The SPRI-based method produced comparable PCR amplification success from fecal extractions but significantly outperformed the PowerSoil Kit in DNA quality, quantity, and PCR success for both cloacal and oral swab samples. Furthermore, published protocols for preparing highly multiplexed Illumina libraries without post-adaptor ligation amplification were modified to minimize sample loss. This modified library preparation protocol was successfully validated on three sets of heterogeneous amplicons (derived from SPRI, PowerSoil, and control extractions of avian fecal and swab samples) that were sequenced across three independent, 250 bp, paired-end runs on Illumina's MiSeq platform. The comprehensive strategies developed in this study focus on maximizing efficiency and minimizing costs. In addition to increasing the feasibility of using minimally invasive sampling and MPS capabilities in avian research, the methods are notably not avian-specific and thus applicable to many research programs that involve DNA extraction and amplicon sequencing.

Introduction

The power of minimally invasive genetic sampling for molecular ecology and conservation genetics has become increasingly realized and applied within the past decade. Genetic analyses of samples such as feces, urine, saliva, feathers, eggshells, hair, and skin have supported research across numerous fields including movement ecology (Ulizio *et al.* 2006), population genetics (Bellemain *et al.* 2005), behavioral ecology (Rudnick *et al.* 2007), disease ecology (Harris *et al.* 2009), and wildlife forensics (Wasser *et al.* 2007). However, although fecal and swab samples are less invasive and more readily attained than blood or tissue samples, they comprise surprisingly underexploited resources in ecological and evolutionary studies within non-mammalian systems (Beja-Pereira *et al.* 2009). Whereas minimally invasive approaches in mammalian studies have utilized a wide range of sample types and availability, comparable approaches in avian studies, for example, have predominantly focused on feather sampling for species identification, gender determination, and population genetics (Segelbacher 2002; Horváth *et al.* 2005).

Molecular analyses of fecal and swab samples afford exciting and underused opportunities to address both long-standing and new research questions in fields such as trophic ecology and microbial ecology. Amplification of prey DNA in fecal metagenomic DNA extracts has enabled characterization of diet and foraging ecology in previously difficult to study taxa such as penguins (Deagle *et al.* 2010) and lizards (Brown *et al.* 2012). Likewise,

amplification of microbial DNA in skin swabs from several species of co-habiting frogs provided insight into processes underlying biogeographic patterns of free-living versus host-associated microbiota (McKenzie *et al.* 2012). Analysis of microbial DNA in fecal extracts has also supported investigation into the relative role of host genetics in determining gut bacterial diversity within an important model species for avian sexual selection and behavioral ecology studies, captive zebra finches (Benskin *et al.* 2010).

Despite the utility and promise of fecal and swab sampling, it remains a challenge to extract high quality and quantities of metagenomic DNA from these sample sources. In the field, limited amenities can interfere with sample preservation and storage conditions, whereby samples containing high concentrations of microbial enzymes can quickly degrade DNA (Ruiz *et al.* 2000). During DNA extraction, non-genetic material such as glycoproteins and phenolics in the samples can also co-extract with the DNA (Monteiro *et al.* 1997). Samples may additionally contain inherently low concentrations of genetic material such that existent extraction methods are too inefficient to yield DNA suitable for PCR amplification. Reviews in the current literature on fecal DNA extraction predominantly focus on samples of mammalian origin (McOrist *et al.* 2002; Tang *et al.* 2008). Although mammals excrete digesta and nitrogenous waste (urea) separately, feces from amphibians, reptiles, and birds are comprised of both nitrogenous waste (uric acid) and gastrointestinal excrement combined (Evans & Heiser 2004). These physiological and sample composition differences pose unique impediments for determining the best DNA extraction method for non-mammalian excreta. The application of standard extraction methods as well as those used successfully on mammalian feces can often be ineffective for samples from other taxa (Idaghdour *et al.* 2003; Jedlicka *et al.* 2013)

The challenges associated with deriving both high quality and high yields of metagenomic DNA from fecal and swab samples also affect the performance of subsequent enzymatic processes using the extracts. For routine analyses involving amplicon sequencing, compensatory strategies are necessary to achieve PCR amplification from poor quality extracts. To address low extraction yields, minimal sample loss must occur during sample handling from PCR through sequencing. The advent of massively parallel sequencing (MPS) has extraordinarily enhanced capabilities to deeply sequence mixed amplicon pools, wherein declining costs continue to render it increasingly feasible to incorporate MPS into ongoing research programs (Neiman *et al.* 2011). Now is the opportune time to capitalize on both the ease of fecal and swab sampling as well as the cost-effectiveness of MPS to pursue studies with amplicon MPS. Unfortunately, although a multitude of criteria must carefully be considered in designing and implementing an amplicon-based MPS study, few resources in the literature are available to inform best practices for the complete workflow from DNA extraction of minimally invasive samples through Illumina sequencing of amplicon libraries.

Therefore, the goals of this study were to: 1) identify effective metagenomic DNA extraction methods for fecal and swab samples, 2) apply cost-efficient library preparation protocols to the metagenomic DNA extracts for amplicon sequencing, using microbial and prey amplicons as case studies, and 3) validate extraction and library preparation approaches for amplicon libraries using Illumina sequencing. In doing so, a novel metagenomic DNA extraction approach was designed based on the repurposing of a reagent commonly used in DNA cleanup, solid phase reversible immobilization (SPRI) beads (Rohland & Reich 2012), and the successful performance of this approach was demonstrated. Field-collected avian fecal and swab samples

were extracted using the SPRI-based method, and the results were compared with a standard commercial kit. Subsequently, a cost-effective amplicon library preparation protocol was synthesized and applied to the field-collected extracts. Finally, the libraries were sequenced on the Illumina platform to report the efficacy of the library preparation protocol and extraction methods for microbial characterization of oral swab, cloacal swab, and fecal samples. With informed approaches for sample collection through extraction and library preparation, PCR-quality DNA can be isolated consistently from minimally invasive samples of non-mammalian systems, and the application of MPS to amplicons comprises an invaluable analytical tool with widespread relevance to diverse questions in ecology and evolution.

Materials and methods

Collection of avian fecal and swab samples

Western Bluebirds *Sialia mexicana*, common nest box occupants in Central California, were monitored among songbird nest boxes from two study sites, a vineyard system with 300 nest boxes erected across 50 vineyard blocks in Napa County, California and a rangeland site with 107 nest boxes erected at the San Joaquin Experimental Range in Madera County, California. Nestlings were sampled between five and 15 days post-hatching. Fecal samples were obtained through sudden lifting of the nestling from the nest, placement of the nestling on a sterilized surface, and encouragement of nestling movement (e.g., flapping and scrambling) to elicit voidance. Sterile nylon flocked mini-tip swabs (Puritan Medical PurFlock 25-3316 1PN) were used to sample cloacal and oral cavities. For cloacal sampling, the cloacal exterior was pre-cleansed with 70% isopropanol before the swab tip was inserted 5–8 mm into the cloaca and gently rotated against the inside circumference of the cloaca two to three times for 5–10 seconds. For oral sampling, a swab was rotated against the tracheal opening, under both sides of the tongue, and along the choanal slit. Samples were initially stored under a variety of buffered and freezer temperature settings for a pilot study to determine the most effective sample preservation conditions before standardization for the rest of the study (described below under *Comparison of SPRI-based versus commercial metagenomic DNA extraction*). All collection tools (fecal collection surface, microspatula, scissors) were sterilized with hydrogen peroxide between samples. Samples were collected from May through July in 2011 and 2012. All protocols were approved by the UC Berkeley Animal Care and Use Committee (MAUPs R217-0512B and R317-0913), California Department of Fish and Game (Permits SC-12125 and SC-11869), and U.S. Geological Survey (Permits 23772 and 21859).

Pilot study for sample preservation and metagenomic DNA extraction

A literature review was performed to derive candidate sample preservation and extraction methods to be compared (Table S2.1, Supporting Information). A pilot study was then conducted on avian fecal and swab samples during which a wide suite of approaches was attempted to gain preliminary insight into their relative performance. For sample preservation, samples were stored dry versus in an ammonium-sulfate saturated nucleic acid stabilization and storage buffer (25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; hereafter DNA buffer). Samples were frozen either at -20°C, temporarily at -20°C before transfer to -80°C, or at -80°C. For metagenomic DNA extraction, eight different

commercial kits from Invitrogen, Zymo, Qiagen, and Mo Bio (Table S2.2, Supporting Information) were tested before developing our novel SPRI-based method (following section).

The quality and quantity of the DNA from the various extraction methods were assessed using three endpoints. The NanoDrop spectrophotometer was used to evaluate DNA purity ($A_{260/280}$, $A_{260/230}$). The Qubit fluorometer with the dsDNA HS (high sensitivity) Assay Kit (Invitrogen Q32854) evaluated DNA yield (ng), and PCR amplification was tested using degenerate primers targeting the bacterial 16S rRNA locus (515F and 806R yields a ~292 bp amplicon; Caporaso *et al.* 2012) and the arthropod mitochondrial cytochrome oxidase *c* subunit I (COI) locus (ZBJ-ArtF1c and ZBJ-ArtR2c produces a ~200 bp amplicon; Zeale *et al.* 2011). These loci were chosen because fecal and swab samples provide unique opportunities to explore host-microbial and dietary relationships, which currently comprise growing fields of study (see Molecular Ecology Special Issue: Nature's Microbiome 2014). All 16S rRNA PCRs were performed in 25 μ L reactions containing 1X HotMasterMix (5 PRIME 2200400), 0.2 μ M of each primer, 0.5 μ g/ μ L BSA (NEB B9000S), 5% DMSO (ACS grade), and 0.1-1 ng metagenomic DNA extract. Thermocycling conditions followed an initial denaturation of 94°C for 2 minutes and 35 cycles of 94°C for 8 seconds, 50°C for 20 seconds, and 72°C for 30 seconds. All COI PCRs were performed in 20 μ L reactions containing 1X Phusion GC Buffer, 200 μ M of each dNTP, 0.5 μ M of each primer, 5% DMSO, 1 U Phusion high fidelity polymerase (NEB M0530S), and 0.1-1 ng of DNA extract. Thermocycling conditions followed an initial denaturation of 98°C for 2 minutes; 35 cycles of 98°C for 8 seconds, 50°C for 20 seconds, and 72°C for 30 seconds; and a final extension of 72°C for 7 minutes. The use of DMSO and BSA as PCR additives prevent the formation of secondary structures in the template and bind PCR inhibitors, respectively (Farell & Alexandre 2012). All PCR reactions were performed in duplicate with appropriate positive and no-template negative controls. PCR set-up was performed in a clean hood, and thermocycling occurred in a separate, isolated room to minimize risk of cross-contamination. PCR reactions were electrophoresed on 1.5% (w/v) agarose gels stained with SYBR-Gold (Invitrogen S-11494) to detect amplification success.

SPRI-based metagenomic DNA extraction from avian fecal and swab samples

Upon obtaining consistently low yields and quality of DNA from fecal and cloacal swab samples across the majority of the tested extraction methods in our pilot study (Table S2.2, Supporting Information), it was noted that nearly all of the methods used silica filter spin columns to bind and wash captured DNA. Improvement upon the DNA binding step during extraction was attempted by employing solid phase reversible immobilization (SPRI) beads (Hawkins *et al.* 1994; DeAngelis *et al.* 1995). SPRI beads, such as Ampure XP beads (Beckman), are commonly used for DNA cleanup of reactions between steps of MPS library preparation (Lundin *et al.* 2010). Briefly, carboxyl-coated magnetic beads reversibly bind DNA based on the concentration of salts and polyethylene glycol in solution with the DNA (Lis & Schleif 1975; Hawkins *et al.* 1994). Due to the specific interaction between the carboxyl groups and DNA, impurities are washed away while the DNA is bound, and the large surface area of the SPRI beads supports greater binding capacity relative to silica purification matrices (Tan & Yiap 2009). Therefore, SPRI beads were substituted for silica spin columns in a unique extraction protocol to try to enhance the resultant DNA yield and quality.

For the SPRI-based method, approximately 220 mg of fecal matter were frozen in 1.5 mL of DNA buffer, and each swab head was stored frozen in 1 mL of DNA buffer. Samples were processed in lysis buffer (166 mM trisodium phosphate, 111 mM guanidine thiocyanate, 11 mM sodium chloride, 0.3% sodium dodecyl sulfate, 37 mM Tris-HCl pH 8.0, 1% beta-mercaptoethanol) with zirconia-silica beads on a Precellys 24 tissue homogenizer (two rounds of 6800 rpm for 30 sec with a 30 sec pause between rounds). Fecal and cloacal lysates were incubated with ammonium acetate to precipitate proteins and aluminum ammonium sulfate to flocculate PCR inhibitors (Braid *et al.* 2003). DNA was purified from the resultant supernatant using SPRI beads, prepared with Sera-Mag (ThermoScientific, 6515-2105-050250) beads following Rohland and Reich (2012) and added at 1.5 times the supernatant volume (see Supporting Information for detailed SPRI-based extraction method protocols).

Comparison of SPRI-based versus commercial metagenomic DNA extraction

The pilot study revealed promising results for fecal and swab samples preserved in DNA buffer at either -20°C or -80°C and extracted using a high-powered homogenizer with either the SPRI-based method or the PowerSoil Kit (Mo Bio 12855; Supporting Information). Therefore, more rigorous testing was pursued for each extraction method on a larger sample size of 13 fecal, five cloacal swab with fecal debris, 18 cloacal swab, and 18 oral swab samples for the SPRI-based method as well as 18 fecal, 18 cloacal swab, and 18 oral samples for the PowerSoil Kit. All were independent samples collected from the rangeland site during the 2012 breeding season, controlled for sample age, stored in DNA buffer on ice immediately after collection, transferred to -20°C storage by the end of the collection day, and maintained at -20°C for a month maximally at the field site before transfer to long-term storage at -80°C in the lab. The same extraction protocols and DNA assessments were followed as in the pilot study except that only bacterial 16S rRNA gene PCRs were conducted, and the PCRs were performed in triplicate per sample to maximize recovery of the targeted diversity in the original samples.

Amplicon library preparation and sequencing

From the literature, several strategies to synthesize a cost-effective approach for high-throughput amplicon library preparation were combined and modified. The synthesized approach was applied to three sets of amplicons to investigate its efficacy. The microbial 16S rRNA amplicons generated from the aforementioned 54 SPRI- and 54 PowerSoil-based extracts comprised the first two sets of amplicons to enable further comparison of the two extraction methods. The third set of amplicons included arthropod COI PCR products, generated as in the pilot study except in triplicate per sample, for 49 fecal samples stored dry at -80°C and extracted using a high powered homogenizer with the Xpedition Soil/Fecal DNA MiniPrep Kit (Zymo D6202). The third set served as a procedural control since this particular combination of sample storage and extraction methods has previously yielded cloning and Sanger sequencing success (Jedlicka *et al.* 2013).

To enable unique identification of each sample after multiplex sequencing (Stiller *et al.* 2009), a combinatorial tagging and sample pooling approach was implemented following Neiman *et al.* (2011) using two different tags (Fig. 2.1). Briefly, the first tag was a 6-nucleotide barcode appended on the 5' end of both the forward and reverse primers and thereby incorporated at both ends of the resultant amplicons during PCR (Table S2.3). The barcodes

were designed to be error-correcting based on the Hamming code so that the true barcode would be recoverable if a single nucleotide sequencing error were to occur therein (Bystrykh 2012). The second tag was a 6-nucleotide index embedded in the Illumina adapter, which was ligated onto pools of uniquely barcoded amplicons during library preparation. To begin amplicon library preparation, 20 μ L of each triplicate PCR reaction was pooled per sample and SPRI-cleaned following Rohland and Reich (2012), except that a 1.2 volume ratio of SPRI bead solution to sample volume was used to discard fragments smaller than 200 bp (e.g., primer dimers; Fig. 2.1). The cleaned samples were quantified using the Qubit dsDNA HS (high sensitivity) Assay Kit and pooled equimolarly into libraries. Each library was then SPRI-cleaned with a 3.0 volume ratio of SPRI bead solution to sample volume in order to standardize the volumes of all libraries yet maximize sample retention.

The libraries were prepared using the “with-bead” method, as the presence of SPRI beads in the reactions does not interfere with reaction performance (Fisher *et al.* 2011). For each library, all reactions and SPRI cleanup steps were performed in the same tube to minimize sample loss from adherence to plastic tubes and pipette tips during sample transfers. The NEBNext End Repair Module (NEB E6050S), NEBNext dA-Tailing Module (NEB E6053S), and Meyer and Kircher (2010) adapter ligation protocols were used with TruSeq Illumina indexed adapters to prepare the libraries. For each SPRI cleanup between library reactions, the same SPRI bead solution from the SPRI-based extraction method was used following Fisher *et al.*'s (2011) SPRI cleanup protocol with the following modifications to maximize DNA recovery. Tubes were always centrifuged for a few seconds before any incubation period to prevent beads from drying onto the tube walls. During the initial DNA binding step, libraries were vortexed at 2000 rpm for 10 minutes and incubated at room temperature for another 10 minutes. Fresh 80% ethanol was used for all washes, and the samples were air-dried for 10 minutes before elution. During elution, the samples were vortexed at 2000 rpm for two minutes and incubated at room temperature for another two minutes. For the final cleanup after adapter ligation, a 0.9 volume ratio of bead suspension buffer to sample was used, to exclude all adapter dimers (see Supporting Information for more detailed Illumina amplicon library preparation protocol). Of note, this library preparation workflow is PCR-free (Kozarewa *et al.* 2009), without a post-ligation PCR enrichment step, to avoid barcode swapping from chimera formation between amplicons of the pooled samples in each library.

The libraries were quantified via qPCR using the KAPA Library Quantification Kit (KK4824), pooled equimolarly into one of three main libraries (PowerSoil-based microbial 16S rRNA, SPRI-based microbial 16S rRNA, or Xpedition-based arthropod COI), and analyzed on the Bioanalyzer (Agilent Technologies G2940CA). The resultant three libraries were submitted to the University of California Davis Genome Center, and each library was analyzed in a separate 250 bp, paired-end run on the Illumina MiSeq platform.

Data analysis

For extraction method comparisons, all Qubit concentration measurements were used to compare total DNA yield in each extract. For fecal extractions, all calculated total DNA yields were normalized by the processed fecal sample size (g wet mass). PCR results were classified into one of five ordinal codings for each sample replicate: (0) no amplification, (1) weak amplification, (2) moderate amplification, (3) strong amplification, and (4) very strong

amplification. The mean value across triplicate PCRs was calculated for each sample as an index of PCR success. Wilcoxon rank sum tests were applied to detect significant differences in DNA quality, quantity, and PCR success between sample types and extraction methods.

The Illumina data were trimmed of adapters and demultiplexed into the original input libraries using CASAVA 1.8. Reads that did not pass filter were removed. Trimmomatic was used to trim the trailing edge of each sequence to a minimum quality score of Q20. Barcodes were error-corrected, and unresolved reads were discarded. PANDAseq was used to merge paired-end reads, using a quality threshold of 0.6 (Masella *et al.* 2012). Merged reads lacking matching barcodes at both ends or correct primer sequences were also discarded. Using QIIME 1.8.0, each library was demultiplexed by barcode, and all barcode and primer sequences were trimmed from the reads (Caporaso *et al.* 2010). Alpha diversity rarefaction analyses were performed to determine minimum sequencing depth to include all samples in beta diversity analyses. Summaries of taxonomic composition of each sample were generated using the Greengenes database (v. 13_8) and Barcode of Life Database (BOLD) as references for taxonomic assignment of bacterial 16S rRNA and arthropod COI amplicons, respectively.

Results

Comparison of SPRI-based versus commercial metagenomic DNA extraction

For fecal samples, the PowerSoil kit produced significantly greater metagenomic DNA yields ($n = 13$ for SPRI and $n = 18$ for Powersoil, $S = 122$, $P = 0.001$), more ideal $A_{260/280}$ ratios ($S = 129.5$, $P = 0.002$), and greater total amplicon yield across triplicate 16S PCRs ($S = 129$, $P = 0.002$; Table S2.4 and Fig. 2.2). However, the SPRI-based method performed equally well in amplification success across triplicate PCRs. For cloacal swab samples, the SPRI-based method yielded significantly more metagenomic DNA ($n = 23$ for SPRI and $n = 18$ for PowerSoil, $S = 201$, $P < 0.001$) and greater performance across triplicate PCRs ($S = 188.5$, $P = <0.0001$; Figs. 2.2 and S2.1). Finally, for oral swab samples, the SPRI-based method produced significantly higher DNA yields ($n = 18$ per method, $S = 495$, $P = <0.0001$), greater DNA purity at both $A_{260/280}$ ($S = 423$, $P = 0.004$) and $A_{260/230}$ ($S = 495$, $P < 0.0001$), as well as greater amplification success across triplicate PCRs ($S = 479$, $P < 0.0001$; Figs. 2.2 and S2.1).

In comparing extraction performance within method, fecal samples generated greater yields, higher quality, and more amplicons than swab samples using the PowerSoil Kit (Table S2.4). Oral swab samples performed the worst of the three sample types with this kit. In contrast, higher quality extracts were derived from the swab samples than the fecal samples with the SPRI-based method. Cloacal swab samples produced the highest yields whereas the oral swab samples provided the purest samples (closest to ideal $A_{260/280}$ and $A_{260/230}$ values). However, despite the observed differences in DNA quality among sample types, all but two fecal extracts amplified across all triplicate 16S PCRs.

Library preparation

With the combinatorial tagging and PCR-free library preparation approach, libraries were generated with 30 to 466 ng of amplicon starting material (Table 2.1). Between 2.5 and 72.7% of starting amplicons were successfully adapter-ligated after library preparation. Across all three MiSeq runs, a mean of 14.4 (± 2.6 SD) million raw reads per run was obtained (Table S2.5). CASAVA 1.8 was used for each run to demultiplex sequences into libraries by adapter

index. Across all libraries, 88.2% of the raw reads passed filter whereas 50.1% passed both filter and quality thresholds for paired-end merging.

Of the merged reads, 6.1% did not have a barcode or contained erroneous barcodes that could not be corrected. These sequences did not produce significant nucleotide BLAST matches. Approximately 6.9% of the merged reads contained barcodes at each end of the read, but the barcodes did not match. These sequences may have resulted from either barcode synthesis errors, barcode sequencing errors, or chimera formation. The 16S rRNA and COI merged reads lacking matching barcodes yielded bacterial and no significant BLAST results, respectively. Additionally, 12.4% of the merged reads possessed correct, matching barcodes but did not contain the primer sequences used in the study. These sequences were typically too short and either matched bacterial sequences or yielded no significant BLAST matches. These sequences may have resulted from either fragmentation of amplicons, primer synthesis errors, or random incorporation of environmental DNA that happened to contain sequences that matched the barcodes used in the study.

Thus overall, 37.4% of the paired-end raw reads or 74.6% of the merged reads contained correct barcodes and primers for retention in downstream analyses. QIIME's "split libraries" function was applied to each library to separate sequences into samples by primer barcode. Quantification via qPCR for equimolar pooling of libraries into a sequencing run generally resulted in an even number of sequences per sample across libraries, although the SPRI libraries exhibited higher variation in number of sequences per library (Table 2.1). Likewise, Qubit quantification for equimolar pooling of amplicons per library produced mainly even sequencing depth across samples, although the SPRI libraries yielded higher variation in the number of sequences per barcode per library.

Taxonomic diversity among PowerSoil- and SPRI-derived microbial amplicons as well as among Xpedition-derived control arthropod amplicons was analyzed (Fig. S2.2). In comparing amplicon sequence diversity from PowerSoil and SPRI-based extracts, samples did not segregate by extraction method for any sample type (Fig. 2.3). More detailed analyses of alpha and beta diversity among sequenced samples are forthcoming (Vo and Jedlicka, in prep).

Discussion

To increase application of MPS technology to broader study systems, there is pressing need for effective protocols on collecting and processing diverse samples compatible with routine molecular analyses such as amplicon sequencing. Although minimally invasive sampling comprises a valuable sampling strategy, there is a striking paucity of non-mammalian genetic studies that capitalize on fecal and swab sampling. Alternative gut content and genetic sampling methods are substantially more invasive, including the use of emetics (Carlisle & Holberton 2006), neck ligatures (Mellott & Woods 1993), gastric lavage (Moody 1970), blood collection (Voss *et al.* 2010), and sacrifice of study organisms (Beal 1915). Of the studies that have involved avian fecal or swab samples, some have pursued genetic sequencing (e.g., Benskin *et al.* 2010; Jedlicka *et al.* 2013), but remarkably, only two have integrated MPS analyses, relying on pyrosequencing (Deagle *et al.* 2010; Singh *et al.* 2012). Therefore, this research addresses the dearth of study design guidelines for embracing minimally invasive sampling and Illumina amplicon sequencing.

Metagenomic DNA extraction from avian fecal and swab samples

Storage of avian fecal and swab samples frozen in ammonium-sulfate saturated DNA stabilization buffer supported metagenomic DNA preservation. In seeking extraction of high quality and quantity metagenomic DNA from these samples, an effective extraction approach that coupled chemical and mechanical steps for cellular lysis as well as solid phase reversible immobilization (SPRI) beads was developed for the isolation of nucleic acids. Although studies have shown superior DNA extraction from human fecal samples using a combination of mechanical and chemical lysis relative to chemical lysis alone (Salonen *et al.* 2010; Ariefdjohan *et al.* 2010; Smith *et al.* 2011), the integration of chemical, mechanical, and SPRI-based approaches to fecal and swab DNA extraction has not been proposed to the best of current knowledge. The SPRI-based method excelled in several measures relative to the PowerSoil Kit employed in the standard operating protocols of both the Human Microbiome Project and the Earth Microbiome Project. Although the SPRI and PowerSoil methods performed comparably in bacterial 16S rRNA amplification success for fecal samples, the SPRI method generated significantly higher DNA quality ($A_{260/280}$, $A_{260/230}$), quantity, and PCR amplification for swab samples in comparison to the PowerSoil Kit.

Given that the composition of the extraction buffers in the SPRI-based and PowerSoil methods were comparable (Brolaski *et al.* 2008), the success of the SPRI method specifically with swab samples is likely due to the higher binding and elution efficiency that SPRI beads exhibit over silica matrices. DNA binds carboxyl groups on the beads with strong affinity in solutions of high ionic strength and dissociates from carboxyl groups in low ionic strength solutions. When SPRI beads were used in extreme excess (three times the sample volume) on clean amplicon samples with the bead cleanup protocol, very high ($87.3 \pm 7.2\%$) recovery of starting material across all samples was observed (Table S2.6). In contrast, studies have documented higher losses of input DNA when using silica filters to clean and/or concentrate DNA samples (e.g., Roossinck *et al.* 2010). The results revealed a stark decline in performance of silica-based extraction for the swab samples relative to the fecal samples which warrants user caution when starting with lower amounts of genetic material with the silica columns.

Concerning sample types, sample complexity plays a pivotal role in determining the difficulty encountered in extracting DNA. Like Idaghdour *et al.* (2003), arthropod exoskeletal remnants were observed among avian fecal matter. Extra measures to remove heavy protein loads and non-nucleic organic compounds were necessary to avoid high levels of co-extracted contaminants in fecal extracts, as seen in Jedlicka *et al.* (2013). It followed that cloacal swab samples performed better when less fecal debris was present and that oral swabs dependably provided high yields as well as relatively pure DNA without the need for special inhibitor removal treatments when extracted using the SPRI-based method. Further, in considering the high recovery efficiency of SPRI beads for nucleic acids, it is likely that host DNA was also co-extracted using the SPRI-based method, but additional application of host-specific primers was beyond the scope of the study. Altogether, the straightforward reagents and supplies needed for the SPRI-based extraction protocol totaled substantially less than the PowerSoil protocol (<\$1/sample for SPRI-based versus \$4.94/sample for PowerSoil), thereby ranking the SPRI-based method as the more cost-effective and reliable approach for processing avian fecal and swab samples.

Illumina library preparation for amplicon sequencing

The modified library preparation protocol resulted in impressively high amplicon template retention rates relative to those reported in other studies. The approach to library preparation synthesized several published, successful, cost-saving strategies from the literature. A combinatorial sample tagging protocol was used (Neiman *et al.* 2011), which permitted the same set of barcoded primer pairs to be used for each indexed library whereby different combinations of barcodes and indices formed unique tags for each sample (Fig. 2.1). The protocol substantially reduced the complexity of sample handling as well as overall costs by eliminating the need for a unique PCR primer pair (~\$60 per pair purified by high performance liquid chromatography) or a unique Illumina adapter (~\$30 per adapter) per sample. Additionally, a with-bead method (Fisher *et al.* 2011) and a PCR-free library preparation (Kozarewa *et al.* 2009) workflow was employed. Modifications to the with-bead SPRI cleanup protocol, including extended DNA binding and elution times as well as vortexing and centrifugation, ensured maximal recovery of template amplicons between library preparation reactions. While Fisher *et al.*'s (2011) with-bead method observed 47% recovery of input library material, the modified library protocol in this study retained 75 (\pm 4)% of the starting template (Table S2.6). This difference is significant when accommodation of low availability of starting material as well as the exclusion of post-ligation PCR enrichment of the libraries are absolutely necessary.

A steady decline in adapter ligation efficiency was observed among sequentially prepared sets of libraries (SPRI to Xpedition to PowerSoil; Table 2.1). Given that each library across all three sets of libraries contained amplicons from successful PCR reactions and that all amplicons were processed with the same SPRI bead cleanup before library preparation, systematic differences in the generation or retention of enzymatic inhibitors between sets of libraries appear unlikely to underlie the pattern in ligation efficiencies. However, increased freeze-thaw cycles of the adapter tubes and resultant adapter degradation may have impacted adapter ligation success in subsequently prepared libraries. Making several sub-aliquots of the adapters in one-time use tubes comprises good general practice for prolonging the use of the purchased adapter stock. Nevertheless, even among libraries that exhibited low adapter ligation efficiency (2.5% of the starting template), sequence analyses suggested that the adapter-ligated products were representative and comprehensive of the diversity of the starting pool of amplicons. No substantial differences in either number of sequences (Table 2.1) or microbial diversity (Fig. 2.3) were observed between samples in libraries with higher versus lower adapter ligation efficiencies. Additionally, there was a decrease in the variability of the number of sequences per library and the number of sequences per barcode per library recovered for subsequent libraries versus initial libraries (again, SPRI to Xpedition to PowerSoil; Table 2.1). Although a possible effect of extraction method cannot be excluded, these differences were likely due to pipet malfunction during equimolar sample pooling in the initial SPRI libraries. Yet, the variation in sequencing depth yielded no detectable differences in microbial diversity between samples (Fig. 2.3).

Sequence analyses of the Xpedition control libraries validated the library preparation approach and demonstrated the strength of MPS application. Using cloning and Sanger sequencing, Jedlicka *et al.* (2013) analyzed the diets of 12 nestlings and discovered a range of one to five and a mean of 2.0 (\pm 1.1 SD) taxa per nestling in their characterization of Western

Bluebird diet from fecal samples. Costs associated with this approach totaled over \$37 per sample as well as an additional \$3.13 per sequenced clone per sample. In contrast, the Illumina approach in this study supported a substantially more extensive characterization of the arthropod prey base for 49 nestlings. A range of one to 56 and a mean of 7.4 (\pm 8.6 SD) taxa per nestling were observed (Fig. S2.2a). Additionally, Illumina costs totaled \$27 per sample and an additional flat rate of \$1,170 for sequencing. Given that a MiSeq run yielded a mean of approximately 14 million paired-end reads, the Illumina approach generated 28,000 times more sequences, and its per sequence cost was four orders of magnitude lower than the combined cloning and Sanger sequencing approach. Thus, the cloning approach was more costly in both relative and absolute terms even though it ultimately acquired less data for fewer samples.

Sequencing analyses of the SPRI- versus PowerSoil-based amplicons revealed little segregation by method. That is, neither extraction method appeared to exhibit inherent bias for sampling a different subset of the total microbial diversity within the samples (Fig. 2.3). This result was interesting for the swab samples as SPRI-based swab extracts were significantly higher quality and amplified more robustly than the PowerSoil swab extracts. Thus, the major advantage with the SPRI-based approach lies in its relative affordability. However regarding swab performance, the SPRI-based method's greater total yield is also highly useful, especially when necessary troubleshooting steps can rapidly deplete extraction stocks and consequently reduce availability of the stocks to support multiple projects. Additionally, the higher quality SPRI-based swab extracts are optimal for preventing PCR inhibition in amplicon-based biodiversity assessments. Overall, the SPRI-based metagenomic DNA extraction and library preparation approach performed exceedingly well in the multiplexing of numerous samples into each library and several libraries into a single Illumina run to support cost-effective, in-depth interrogation of mixed amplicon pools from the avian fecal and swab samples.

Applications

In light of the ever increasing accessibility of MPS technology, investigations of several lines of inquiry in ecology and evolution using avian study systems can benefit tremendously from MPS-mediated genetic analyses. With dietary analyses, morphological examination of prey items from stomach contents, regurgitated pellets, or fecal matter results in over-representation of prey with hard parts that resist digestion (Tollit *et al.* 2003). Methods employing stable isotopes of animal tissues require the additional analysis of all possible prey items to match isotopic signatures and identify potential prey, resulting in low taxonomic resolution for the prey base (Hobson *et al.* 1997). It is possible to use degenerate primers to amplify prey DNA in fecal samples, clone libraries to separate mixed PCR products, and use Sanger sequencing to identify cloned amplicons, but the entire process is tedious, inefficient, and costly for large sample sizes.

Similarly, for the characterization of microbiota or detection of pathogens, microscopy approaches are low-throughput and prone to sensitivity shortcomings as well as inter-observer bias (Pizarro *et al.* 2007). Culture-dependent methods provide equally low sensitivity as less than one percent of bacteria are culturable (Amann *et al.* 1995). G+C profiling (Apajalahti *et al.* 2001), gel electrophoresis and band analysis of rRNA gene amplicons (Knarreborg *et al.* 2002; Lucas & Heeb 2005; Klomp *et al.* 2008), and sequencing of 16S rRNA gene clone libraries (Lu *et al.* 2003; Jeter *et al.* 2009) comprise vast improvements as culture-independent methods of microbial community characterization, but they still afford limited taxonomic resolution and/or

intensive input effort. Microarray analyses can include thousands of probes to provide impressive taxonomic resolution for microbiota, but the probe-dependence restricts the capacity of the method to discover new species (Claesson *et al.* 2009).

MPS can be applied to drastically enhance these research scopes and analyses with greater cost-effectiveness than older, established methods. For microbial and dietary studies, MPS can simultaneously sequence target loci in metagenomic extracts from hundreds to thousands of individuals. Furthermore, MPS analysis of DNA from fecal and swab samples in particular can address a multitude of parallel questions and projects as they contain host, prey, and microbial genetic material.

It is important to note that although the novel SPRI-based extraction approach worked remarkably well with avian fecal and swab samples, the protocol is applicable to many sample types and circumstances in which DNA recovery is the goal. That is, it is relevant for DNA isolation from recalcitrant or degraded samples, such as feathers or museum samples, as well as from typically less problematic samples, such as blood or tissue. Additionally, extracts produced from other methods that exhibit difficulty in PCRs can be cleaned using the SPRI-based cleanup from the library preparation protocol to improve amplification success. The strengths of the SPRI-based extraction lie in the yield and quality of sample DNA derived, which enhance the availability of genetic material for the rest of the workflow from PCR amplification through amplicon library preparation. From the discovery of millions of single nucleotide polymorphisms for phylogenetic studies to the detection of the rare biosphere in biodiversity assessments, MPS has and will continue to empower research in the life sciences, and it is hoped that the strategies and workflows proposed herein will help buttress research with minimally invasive samples to new MPS-mediated heights.

References

- Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, **59**, 143–169.
- Apajalahti JHA, Kettunen A, Bedford MR, Holben WE (2001) Percent G+C profiling accurately reveals diet-related differences in the gastrointestinal microbial community of broiler chickens. *Applied and Environmental Microbiology*, **67**, 5656–5667.
- Ariefdjohan MW, Savaiano DA, Nakatsu CH (2010) Comparison of DNA extraction kits for PCR-DGGE analysis of human intestinal microbial communities from fecal specimens. *Nutrition Journal*, **9**, 23–30.
- Beal FEL (1915) Food of the robins and bluebirds of the United States. *Bulletin of the U.S. Department of Agriculture*, **171**, 1–31.
- Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources*, **9**, 1279–1301.
- Bellemain E, Swenson JE, Tallmon D, Brunberg S, Taberlet P (2005) Estimating population size of elusive animals with DNA from hunter-collected feces: four methods for brown bears. *Conservation Biology*, **19**, 150–161.
- Benskin CMH, Rhodes G, Pickup RW, Wilson K, Hartley IR (2010) Diversity and temporal stability of bacterial communities in a model passerine bird, the zebra finch. *Molecular Ecology*, **19**, 5531–5544.

- Braid MD, Daniels LM, Kitts CL (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. *Journal of Microbiological Methods*, **52**, 389–393.
- Brolaski MN, Venugopal RJ, Stolow D (2008) Kits and processes for removing contaminants from nucleic acids in environmental and biological samples. *US Patent App 2008/US7459548 B2*.
- Brown DS, Jarman SN, Symondson WOC (2012) Pyrosequencing of prey DNA in reptile faeces: analysis of earthworm consumption by slow worms. *Molecular Ecology Resources*, **12**, 259–266.
- Bystrykh LV (2012) Generalized DNA Barcode Design Based on Hamming Codes. *PLoS ONE*, **7**, e36852.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.
- Carlisle JD, Holberton RL (2006) Relative efficiency of fecal versus regurgitated samples for assessing diet and the deleterious effects of a tartar emetic on migratory birds. *Journal of Field Ornithology*, **77**, 126–135.
- Claesson MJ, O’Sullivan O, Wang Q *et al.* (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE*, **4**, e6669–6684.
- Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics*, **11**, 2039–2048.
- DeAngelis MM, Wang DG, Hawkins TL (1995) Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Research*, **23**, 4742–4743.
- Evans H, Heiser J (2004) What’s inside: anatomy and physiology. In: *Cornell Lab of Ornithology Handbook of Bird Biology*. Princeton University Press, Ithaca, New York.
- Farell EM, Alexandre G (2012) Bovine serum albumin further enhances the effects of organic solvents on increased yield of polymerase chain reaction of GC-rich templates. *BMC Research Notes*, **5**, 257–265.
- Fisher S, Barry A, Abreu J *et al.* (2011) A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biology*, **12**, R1–15.
- Harris RN, Brucker RM, Walke JB *et al.* (2009) Skin microbes on frogs prevent morbidity and mortality caused by a lethal skin fungus. *The ISME Journal*, **3**, 818–824.
- Hawkins TL, O’Connor-Morin T, Roy A, Santillan C (1994) DNA purification and isolation using a solid-phase. *Nucleic Acids Research*, **22**, 4543–4544.
- Hobson KA, Sease JL, Merrick RL, Piatt JF (1997) Investigating trophic relationships of pinnipeds in Alaska and Washington using stable isotope ratios of nitrogen and carbon. *Marine Mammal Science*, **13**, 114–132.
- Horváth MB, Martínez-Cruz B, Negro JJ, Kalmár L, Godoy JA (2005) An overlooked DNA source for non-invasive genetic analysis in birds. *Journal of Avian Biology*, **36**, 84–88.
- Idaghdour Y, Broderick D, Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustards. *Conservation Genetics*, **4**, 789–792.

- Jedlicka JA, Sharma AM, Almeida RPP (2013) Molecular tools reveal diets of insectivorous birds from predator fecal matter. *Conservation Genetics Resources*, **5**, 879–885.
- Jeter SN, McDermott CM, Bower PA *et al.* (2009) Bacteroidales diversity in Ring-Billed Gulls (*Larus delawarensis*) residing at Lake Michigan beaches. *Applied and Environmental Microbiology*, **75**, 1525–1533.
- Klomp JE, Murphy MT, Smith SB *et al.* (2008) Cloacal microbial communities of female spotted towhees *Pipilo maculatus*: microgeographic variation and individual sources of variability. *Journal of Avian Biology*, **39**, 530–538.
- Knarreborg A, Simon MA, Engberg RM, Jensen BB, Tannock GW (2002) Effects of dietary fat source and subtherapeutic levels of antibiotic on the bacterial community in the ileum of broiler chickens at various ages. *Applied and Environmental Microbiology*, **68**, 5918–5924.
- Kozarewa I, Ning Z, Quail MA *et al.* (2009) Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nature Methods*, **6**, 291–295.
- Lis JT, Schleif R (1975) Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Research*, **2**, 383–390.
- Lu J, Idris U, Harmon B *et al.* (2003) Diversity and succession of the intestinal bacterial community of the maturing broiler chicken. *Applied and Environmental Microbiology*, **69**, 6816–6824.
- Lucas FS, Heeb P (2005) Environmental factors shape cloacal bacterial assemblages in great tit *Parus major* and blue tit *P. caeruleus* nestlings. *Journal of Avian Biology*, **36**, 510–516.
- Lundin S, Stranneheim H, Pettersson E, Klevebring D, Lundeberg J (2010) Increased throughput by parallelization of library preparation for massive sequencing. *PLoS ONE*, **5**, e10029–10036.
- Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, **13**, 31.
- McKenzie VJ, Bowers RM, Fierer N, Knight R, Lauber CL (2012) Co-habiting amphibian species harbor unique skin bacterial communities in wild populations. *The ISME Journal*, **6**, 588–596.
- McOrist AL, Jackson M, Bird AR (2002) A comparison of five methods for extraction of bacterial DNA from human faecal samples. *Journal of Microbiological Methods*, **50**, 131–139.
- Mellott RS, Woods PE (1993) An improved ligation technique for dietary sampling in nestling birds (Una técnica mejorada para obtener muestras de la dieta ofrecida a pichones). *Journal of Field Ornithology*, **64**, 205–210.
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, **2010**, 1–10.
- Monteiro L, Bonnemaïson D, Vekris A *et al.* (1997) Complex polysaccharides as PCR inhibitors in feces: *Helicobacter pylori* model. *Journal of Clinical Microbiology*, **35**, 995–998.
- Moody DT (1970) A method for obtaining food samples from insectivorous birds. *The Auk*, **87**, 579–579.
- Neiman M, Lundin S, Savolainen P, Ahmadian A (2011) Decoding a substantial set of samples in parallel by massive sequencing. *PLoS ONE*, **6**, e17785–17792.

- Pizarro JC, Lucero DE, Stevens L (2007) PCR reveals significantly higher rates of *Trypanosoma cruzi* infection than microscopy in the Chagas vector, *Triatoma infestans*: High rates found in Chuquisaca, Bolivia. *BMC Infectious Diseases*, **7**, 66–74.
- Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, **22**, 939–946.
- Roossinck MJ, Saha P, Wiley GB *et al.* (2010) Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology*, **19**, 81–88.
- Rudnick JA, Katzner TE, Bragin EA, DeWoody JA (2007) A non-invasive genetic evaluation of population size, natal philopatry, and roosting behavior of non-breeding eastern imperial eagles (*Aquila heliaca*) in central Asia. *Conservation Genetics*, **9**, 667–676.
- Ruiz TR, Andrews S, Smith GB (2000) Identification and characterization of nuclease activities in anaerobic environmental samples. *Canadian Journal of Microbiology*, **46**, 736–740.
- Salonen A, Nikkilä J, Jalanka-Tuovinen J *et al.* (2010) Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: Effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *Journal of Microbiological Methods*, **81**, 127–134.
- Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. *Molecular Ecology Notes*, **2**, 367–369.
- Singh KM, Shah T, Deshpande S *et al.* (2012) High through put 16S rRNA gene-based pyrosequencing analysis of the fecal microbiota of high FCR and low FCR broiler growers. *Molecular Biology Reports*, **39**, 10595–10602.
- Smith B, Li N, Andersen AS, Slotved HC, Krogfelt KA (2011) Optimising Bacterial DNA Extraction from Faecal Samples: Comparison of Three Methods. *The Open Microbiology Journal*, **5**, 14–17.
- Stiller M, Knapp M, Stenzel U, Hofreiter M, Meyer M (2009) Direct multiplex sequencing (DMPS)--a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Research*, **19**, 1843–1848.
- Tan SC, Yiap BC (2009) DNA, RNA, and protein extraction: the past and the present. *BioMed Research International*, **2009**, 1–10.
- Tang J, Zeng Z, Wang H *et al.* (2008) An effective method for isolation of DNA from pig faeces and comparison of five different methods. *Journal of Microbiological Methods*, **75**, 432–436.
- Tollit DJ, Wong M, Winship AJ, Rosen DAS, Trites AW (2003) Quantifying errors associated with using prey skeletal structures from fecal samples to determine the diet of Steller's sea lion (*Eumetopias jubatus*). *Marine Mammal Science*, **19**, 724–744.
- Ulizio TJ, Squires JR, Pletscher DH *et al.* (2006) The Efficacy of Obtaining Genetic-Based Identifications from Putative Wolverine Snow Tracks (Whittaker, Ed.). *Wildlife Society Bulletin*, **34**, 1326–1332.
- Voss M, Shutler D, Werner J (2010) A Hard Look at Blood Sampling of Birds. *The Auk*, **127**, 704–708.
- Wasser SK, Mailand C, Booth R *et al.* (2007) Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 4228–4233.

Zeale MRK, Butlin RK, Barker GLA, Lees DC, Jones G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, **11**, 236–244.

Tables

Table 2.1. Performance of combinatorial tagging approach for Illumina amplicon library preparation protocol used in this study. Means and standard deviations shown.

Preparatory step	SPRI	PowerSoil	Xpedition
Total number of samples	54	54	49
Total number of libraries	9	9	5
Amplicon starting material (ng)	405 ± 40.9	212 ± 63.2	54 ± 17.7
% yield of adapter-ligated amplicons relative to starting amplicon	44 ± 10	2.5 ± 0.7	24 ± 7.2
Number of sequences/library	682,097 ± 698,694	282,823 ± 50,439	399,983 ± 28,605
Number of sequences/barcode/library	113,683 ± 116,449	18,910 ± 5,478	40,843 ± 2,741

Figures

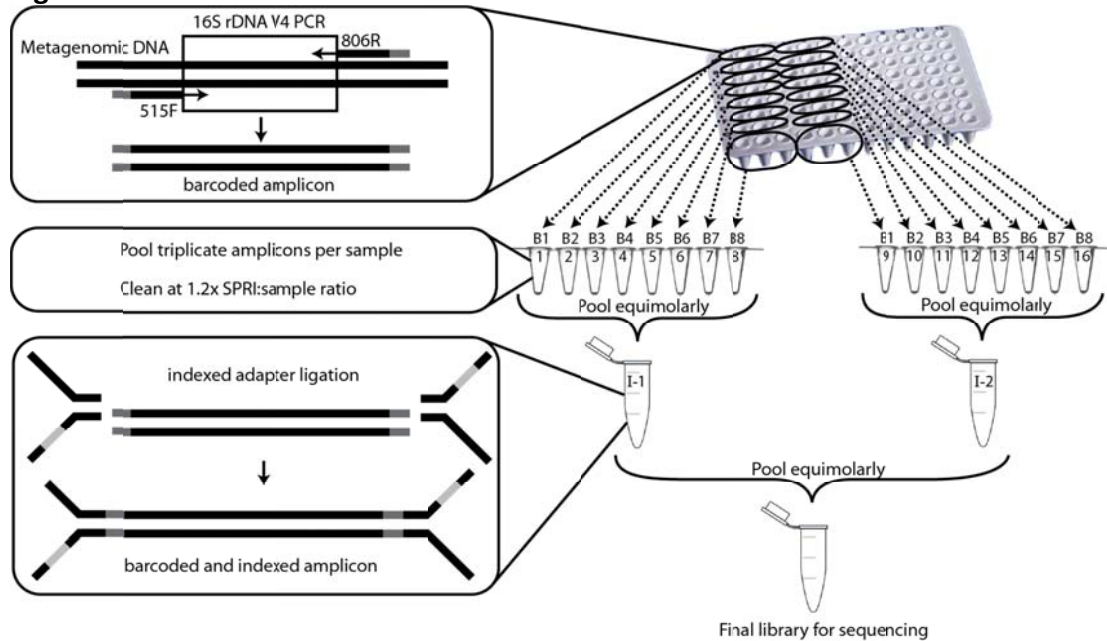


Fig. 2.1. Combinatorial tagging and pooling strategy employed in this study for Illumina amplicon library preparation and sequencing, demonstrated for 16 samples. Dark grey bars indicate barcodes on the 5' end of primers that are incorporated into amplicons during PCR. Light grey bars indicate indexes on the adapter that are attached to amplicons during adapter ligation. B# and I-# denote unique barcode and index sequences, respectively, whereas tube # indicates unique sample ID. The same set of distinct barcodes is repeated with each index to cost-effectively yield unique combinations for each sample ID.

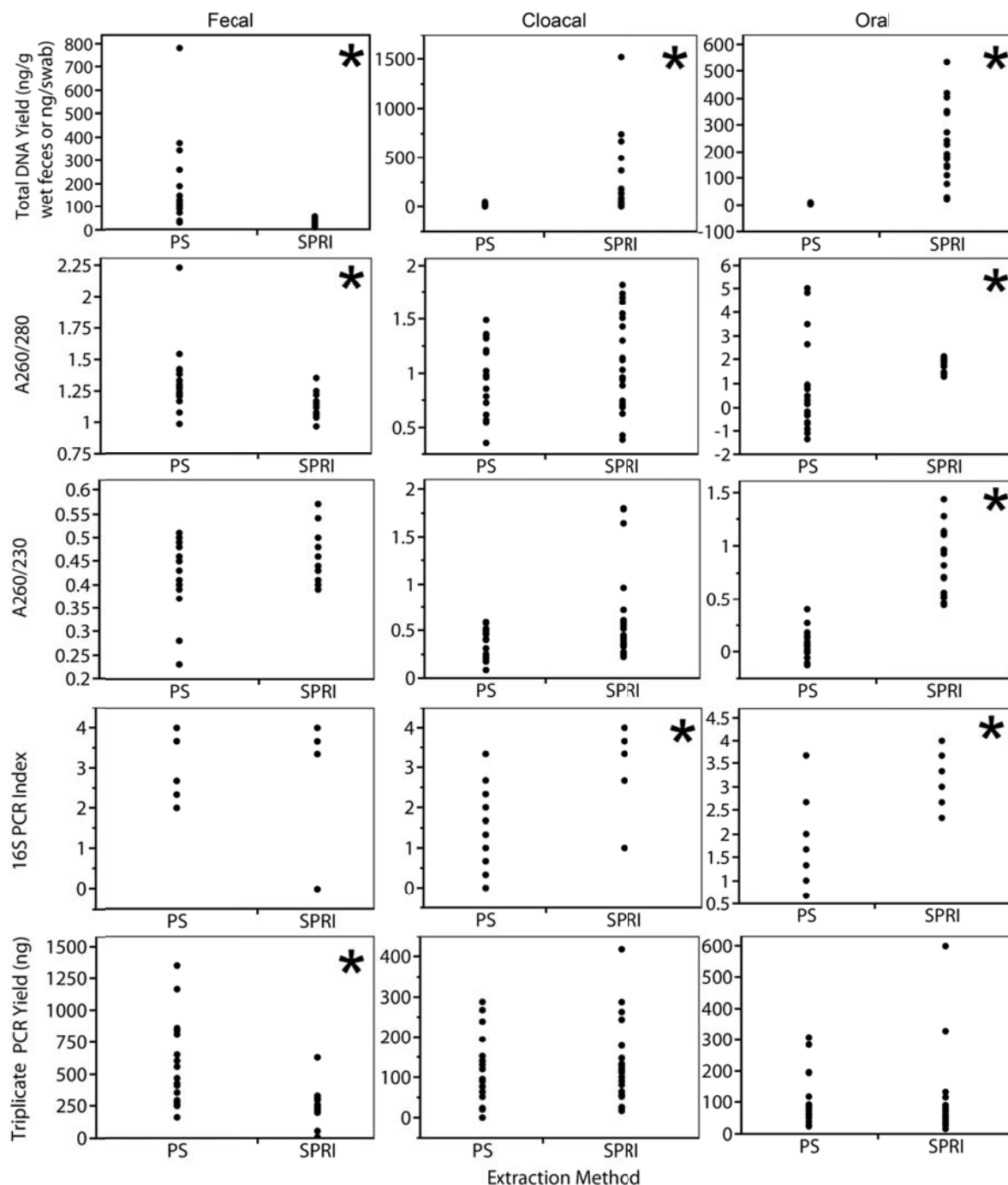


Fig. 2.2. Performance of Mo Bio PowerSoil (PS) Kit versus SPRI-based method (developed in this study) for metagenomic DNA extraction from avian fecal (left column), cloacal swab (middle column), and oral swab (right column) samples assessed via total DNA yield (top row), DNA quality (second and third rows), index of success (fourth row) as well as total yield (fifth row) across triplicate PCRs targeting the bacterial 16S rRNA locus. For PS, 18 of each sample type were processed whereas for SPRI, 13 fecal, 23 cloacal swabs, and 18 oral swabs were analyzed. Asterisks indicate significant differences between methods.

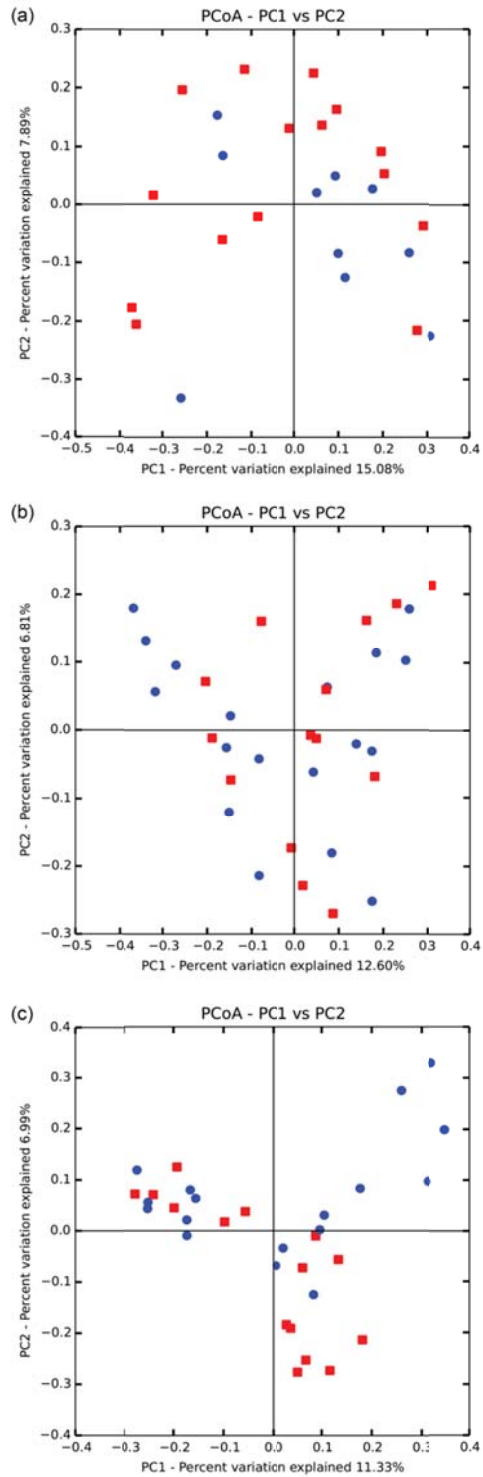


Fig. 2.3. Principal coordinate analysis of 16S rRNA microbial diversity in metagenomic DNA extracts from avian fecal (a), cloacal swab (b), and oral swab (c) samples using the SPRI-based extraction method (squares) and the Mo Bio PowerSoil Kit (circles).

Supporting Information

Materials and methods

Pilot study comparison of sample preservation and commercial metagenomic DNA extraction methods for avian fecal and swab samples

Literature review and pilot study design

In determining candidate sample preservation and DNA extraction methods for comparison, the primary literature was searched for all articles that matched combinations of the following search terms: avian, cloacal, oral, pharyngeal, feces, and DNA (last search on June 8, 2013 in Web of Science). For each study that involved DNA extraction directly from avian fecal or swab samples, the study goal, DNA extraction method, collected sample type(s), sample storage strategy, target avian species, and sequence analysis (if applicable) were noted. The most commonly used DNA storage and extraction methods across all studies in the search were identified to be evaluated on avian fecal and swab samples collected in this study (Table S2.2). Additionally, methods that were either commercially recommended or published for mammalian fecal samples were included (Salonen *et al.* 2010; Ariefdjohan *et al.* 2010). In total, six different sample storage conditions, including various storage temperatures and usage of preservation buffer, as well as eight different metagenomic DNA extraction methods, including kits from Invitrogen, Zymo, Qiagen, and Mo Bio were evaluated. This pilot study included numerous approaches at the expense of a more limited but better replicated subset of approaches.

Sample storage conditions

Storage temperature.—The vineyard and rangeland study sites differed in accessibility to laboratory amenities. The vineyard system was located approximately 93 km from the lab, and therefore, samples from this site could be stored at -80°C in the lab within hours after collection. In contrast, the rangeland system was approximately 296 km from the lab, and consequently, samples from this site were initially stored at -20°C for up to two months before long-term storage in the lab at -80°C. Therefore, samples were collected from both sites to test the effect of immediate storage at -80°C versus temporary initial storage at -20°C.

Sample preservation buffer.— The performance of a lab-made ammonium-sulfate saturated nucleic acid stabilization and storage buffer (25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; hereafter DNA buffer) was evaluated. Fecal samples were stored either dry or in a sample to buffer ratio of 0.15 g/mL. Fecal subsample sizes were standardized using a digital scale. All swab samples were stored in 1 mL of buffer.

Commercial and published metagenomic DNA extraction methods

The following extraction methods were evaluated: 1. Easy DNA Kit (Invitrogen K1800-01) protocol #3 for small amounts of cells, tissues, and leaves, 2. Xpedition Soil/Fecal DNA MiniPrep Kit (Zymo D6202) including 2-mercaptoethanol, 3. QIAamp DNA Stool Mini Kit (Qiagen 51504) human protocol, 4. UltraClean Soil Kit (Mo Bio 12800), 5. UltraClean Fecal Kit (Mo Bio 12811), 6. PowerPlant Pro Kit (Mo Bio 13400), 7. PowerSoil Kit (Mo Bio 12888), and 8. PowerMicrobiome Kit (Mo Bio 26000).

All directions were followed according to manufacturer or literature specifications. However, for some tested methods, several parameters were also modified in attempts to improve DNA quality and yield: 1. starting sample size (adjusted between 0.04 and 0.25 g), 2. addition of 2-mercaptoethanol to the lysis buffer, and 3. sample homogenizer (Mo Bio Vortex adapter vs Precellys 24). For samples stored in DNA buffer, the thawed samples were centrifuged at 10,000g for four minutes and removed the buffer. Fecal samples were subsampled whereas entire swab heads were included in all extraction protocols. For sample lysis in non-commercial protocols, bead tubes were prepared using 2 mL screw-cap o-ring tubes containing 0.5 g of autoclaved zirconia-silica beads (equal parts 0.1 mm and 0.5 mm diameter beads; BioSpec 11079101z and 11079105z, respectively). All elutions were stored at -20°C.

Data analyses

PCR results were assigned one of five ordinal codings for each sample replicate: (0) no amplification, (1) weak amplification, (2) moderate amplification, (3) strong amplification, and (4) very strong amplification. The mean value across triplicate PCRs was calculated for each sample as an index of PCR success. Because this pilot study was not intended to have a full-factorial, balanced, and replicated design, statistical analysis of the collected data was limited. Some samples were homogenized and used among different extraction methods, which supported direct within-sample comparisons. In other comparisons, specific combinations of sample preservation and extraction methods were repeated either in duplicate or in triplicate for independent samples, yielding biological replicates. However, there were several instances in which a particular combination of sample preservation and extraction method was tested only once with one sample, thus possessing neither technical nor biological replicates. The data were analyzed as permissible given these limitations.

SPRI (Sera-Mag) bead solution recipe

1. Use filter pipette tips. Sera-Mag bead solution:

Reagent	Mass	Volume
PEG-8000 powder (Promega V3011)	9 g	8.30 mL
5 M NaCl	-	10 mL
1 M Tris-HCl, pH 8.0	-	500 uL
0.5 M EDTA, pH 8.0	-	100 uL
ddH ₂ O		30.1 mL
Tween-20 (100%)		27.5 uL
Washed Sera-Mag Beads (ThermoScientific, 6515-2105-050250)		1 mL

2. Add 9 g PEG-8000 to 50 mL Falcon tube.
3. To the powder add NaCl, Tris-HCl, EDTA, and ddH₂O.
4. Shake Falcon tube until all PEG has dissolved.
5. Add Tween-20 and mix properly.
6. Resuspend stock solution of SeraMag beads by shaking until "sediment" no longer visible.
7. Transfer 1 mL bead suspension to a 2 mL Eppendorf tube and pellet the beads in a magnetic rack (~3 min). Remove storage buffer and wash beads twice with 1 mL TE (remove tube from magnetic rack and vortex). Fully resuspend the beads in 1 mL TE.
8. Add bead suspension to the Falcon tube and mix immediately. Wrap Falcon tube in aluminum foil and store at 4°C.

SPRI-based metagenomic DNA extraction protocol for avian fecal and swab samples

Materials

Equipment

1. Laminar flow hood
2. Incubator or dry bath
3. Timer
4. Analytical balance
5. Bunsen burner
6. Metal spatulas
7. Microcentrifuge
8. High-powered homogenizer such as the Precellys 24 (Precellys EQ03119.200.RD000.0), PowerLyzer 24 (Mo Bio 13155), or FastPrep-24 (MP Biomedicals 116004500)
9. Magnet rack for 2 mL tubes (Invitrogen DynaMag-2 12321D)
10. Vortex with foam attachment for mixing 2 mL tubes (e.g., VWR 12620-876)

Supplies

1. P20, P200, P1000 pipets and filter tips
2. Bleach
3. Collection tubes: a set of 2 mL snap-cap tubes as A, B, C, and Elution per sample Sterile reagent tubes (e.g., 50 ml Falcon tubes or 125 mL media bottles)
4. Lysis tubes: 2 mL o-ring, screw-cap tubes (Fisher Scientific 02-682-558) filled with 0.5 g of a 1:1 mixture of 0.1 and 0.5 mm diameter, autoclaved, zirconia-silica beads (BioSpec 11079101z and 11079105z, respectively)

Reagents

1. Lysis buffer: 166 mM trisodium phosphate, 111 mM guanidine thiocyanate, 11 mM sodium chloride, 0.3% sodium dodecyl sulfate, 37 mM Tris-HCl pH 8.0, 1% beta-mercaptoethanol
2. 100% ethanol
3. 133 mM ammonium acetate
4. 120 mM aluminum ammonium sulfate
5. SPRI bead solution (see above recipe)
6. 10 mM Tris-HCl, pH 8.0

Fecal samples

1. Thaw samples (preserved in 25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; hereafter DNA buffer) on ice or at 4°C.
2. While the samples are thawing (approximately 20 minutes), bleach and UV-irradiate laminar flow hood workspace. Heat lysis buffer to 60°C for at least 10 minutes. Make a fresh dilution of 70% ethanol in a 50 mL Falcon tube. Label lysis tubes and collection tubes with sample names.
3. Spin down all thawed samples at 10,000g for 4 min.
4. Discard all DNA buffer. *Note:* if remaining sample is anticipated (i.e., >80 mg fecal matter has been collected), keep DNA buffer in a clean tube to transfer back to the remaining sample, after subsampling has been completed, for continued sample preservation and storage.

5. Subsample 80 mg of fecal matter into lysis tubes using an analytical balance and flame-sterilized spatulas. Add 818 uL warmed lysis buffer immediately to tube. Vortex briefly. *Notes:* It is easiest to use the screw cap of the lysis tube as a “weigh boat.” Once the appropriate fecal mass has been subsampled, carefully screw the cap onto the lysis tube. Flick the sample completely into the lysis buffer before vortexing briefly. Additionally, to make this step more time-efficient for high-throughput processing, use as many metal spatulas as possible. Flame-sterilize them all at once. Drop each spatula into a beaker of 10% bleach after use. Rinse them altogether and resterilize when all clean spatulas have been used.
6. Homogenize samples on Precellys 24: two cycles of 6800 rpm for 30 seconds with a 30 second pause between cycles.
7. Centrifuge tubes at 4,500 g for 6 minutes at room temperature.
8. Transfer supernatant to clean 2 mL collection tube A. Be sure to shove pipette tip to bottom of tube, under beads to reach all liquid, without aspirating too many beads.
9. Add 450 uL of 120 mM aluminum ammonium sulfate and vortex for 5 seconds. Incubate at 4C for 5 minutes.
10. Centrifuge tubes at 4,500 g for 6 minutes at room temperature.
11. Avoiding the pellet, transfer entire volume (approximately 850 uL) into a clean 2 mL collection tube B.
12. Centrifuge tubes at 4,500 g for 6 minutes at room temperature.
13. Avoiding the pellet, transfer up to but no more than 450 uL of supernatant into a clean 2 mL collection tube C. Save the rest of the 400 uL at 4°C for future extraction.
14. Add 675 uL (1.5x of 450 uL original sample) of SPRI bead solution to each tube. Pipette to mix 10 times, and incubate tubes at room temperature for 5 minutes to bind DNA. *Notes:* For maximum yield, vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly (tube top hinges facing outward in the centrifuge). Then let sit for another 10 minutes at room temperature.
15. Place the tubes on a magnet rack for 4 minutes to separate.
16. Aspirate and discard the supernatant while the tubes are on the magnet rack.
17. Take the tubes off the magnet. Add 765 uL of 70% ethanol (1.7x sample volume) and pipet 10 times or vortex briefly to mix.
18. Place the tubes back on the magnet for 2 minutes or until the solution clears.
19. Aspirate and discard the supernatant while the tubes are still on the magnet.
20. Repeat steps 19-21 twice more for a total of three ethanol washes.
21. Spin down the tubes for 2 seconds during the final wash before placement on magnet rack. Remove as much of the final ethanol wash as possible. Let tubes air dry for 4 minutes.
22. Take tubes off the magnet. Add 50 uL of 10 mM Tris-HCl, pH 8.0. Pipet to mix 10 times, and incubate at room temperature for 5 minutes to elute DNA. *Notes:* For maximum yield, vortex for 1 minute at 2000 rpm. Spin tubes briefly (tube top hinges facing outward in the centrifuge). Then let sit for another 5 minutes at room temperature to elute DNA.
23. Place the tubes back on the magnet for 4 minutes or until the solution clears.
24. Store eluted DNA at -20°C.

Cloacal swab samples

1. Thaw samples (preserved in 25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; hereafter DNA buffer) on ice or at 4°C.
2. While the samples are thawing (approximately 20 minutes), bleach and UV-irradiate laminar flow hood workspace. Heat lysis buffer to 60°C for at least 10 minutes. Make a fresh dilution of 70% ethanol in a 50 mL Falcon tube. Label lysis tubes and collection tubes with sample names.
3. Add 900 uL PBS to each swab sample.
4. Spin down all samples at 10,000g for 4 min.
5. Discard all DNA buffer.
6. Add 450 uL warmed lysis buffer immediately to each swab sample. Vortex briefly.
7. Homogenize samples on Precellys 24: two cycles of 6800 rpm for 30 seconds with a 30 second pause between cycles.
8. Centrifuge tubes at 10,000g for 1 minute at room temperature.
9. Transfer supernatant to clean 2 mL collection tube A. Be sure to shove pipette tip to bottom of tube, under beads to reach all liquid, without aspirating too many beads.
10. Add 100 uL of 133 mM ammonium acetate and vortex for 5 seconds. Incubate at 4°C for 5 minutes.
11. Centrifuge tubes at 10,000g for 1 minute at room temperature.
12. Avoiding the pellet, transfer up to 600 uL supernatant to collection tube B.
13. Add 100 uL of 120 mM aluminum ammonium sulfate and vortex for 5 seconds. Incubate at 4°C for 5 minutes.
14. Centrifuge tubes at 10,000g for 1 minute at room temperature.
15. Avoiding the pellet, transfer up to 600 uL supernatant to collection tube C.
16. Add 900 uL (1.5x of 600 uL original sample) of SPRI bead solution to each tube. Pipette to mix 10 times, and incubate tubes at room temperature for 5 minutes to bind DNA. *Notes:* For maximum yield, vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly (tube top hinges facing outward in the centrifuge). Then let sit for another 10 minutes at room temperature.
17. Place the tubes on a magnet rack for 4 minutes to separate.
18. Aspirate and discard the supernatant while the tubes are on the magnet rack.
19. Take the tubes off the magnet. Add 1025 uL of 70% ethanol (1.7x sample volume) and pipet 10 times or vortex briefly to mix.
20. Place the tubes back on the magnet for 2 minutes or until the solution clears.
21. Aspirate and discard the supernatant while the tubes are still on the magnet.
22. Repeat steps 19-21 twice more for a total of three ethanol washes.
23. Spin down the tubes for 2 seconds during the final wash before placement on magnet rack. Remove as much of the final ethanol wash as possible. Let tubes air dry for 4 minutes.
24. Take tubes off the magnet. Add 60 uL of 10 mM Tris-HCl, pH 8.0. Pipet to mix 10 times, and incubate at room temperature for 5 minutes to elute DNA. *Notes:* For maximum yield, vortex for 1 minute at 2000 rpm. Spin tubes briefly (tube top hinges facing

outward in the centrifuge). Then let sit for another 5 minutes at room temperature to elute DNA.

25. Place the tubes back on the magnet for 4 minutes or until the solution clears.

26. Store eluted DNA at -20°C.

Oral swab samples

1. Thaw samples (preserved in 25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; hereafter DNA buffer) on ice or at 4°C.
2. While the samples are thawing (approximately 20 minutes), bleach and UV-irradiate laminar flow hood workspace. Heat lysis buffer to 60°C for at least 10 minutes. Make a fresh dilution of 70% ethanol in a 50 mL Falcon tube. Label lysis tubes and collection tubes with sample names.
3. Add 900 uL PBS to each swab sample.
4. Spin down all samples at 10,000g for 4 min.
5. Discard all DNA buffer.
6. Add 450 uL warmed lysis buffer immediately to each swab sample. Vortex briefly.
7. Homogenize samples on Precellys 24: two cycles of 6800 rpm for 30 seconds with a 30 second pause between cycles.
8. Centrifuge tubes at 10,000 g for 1 minute at room temperature.
9. Transfer supernatant to clean 2 mL collection tube A. Be sure to shove pipette tip to bottom of tube, under beads to reach all liquid, without aspirating too many beads
10. Centrifuge tubes at 10,000 g for 1 minute at room temperature.
11. Avoiding the beads, transfer up to 600 uL supernatant to collection tube B.
12. Add 675 uL (1.5x of 450 uL original sample) of SPRI bead solution to each tube. Pipette to mix 10 times, and incubate tubes at room temperature for 5 minutes to bind DNA. *Notes:* For maximum yield, vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly (tube top hinges facing outward in the centrifuge). Then let sit for another 10 minutes at room temperature.
13. Place the tubes on a magnet rack for 4 minutes to separate.
14. Aspirate and discard the supernatant while the tubes are on the magnet rack.
15. Take the tubes off the magnet. Add 765 uL of 70% ethanol (1.7x sample volume) and pipet 10 times or vortex briefly to mix.
16. Place the tubes back on the magnet for 2 minutes or until the solution clears.
17. Aspirate and discard the supernatant while the tubes are still on the magnet.
18. Repeat steps 19-21 twice more for a total of three ethanol washes.
19. Spin down the tubes for 2 seconds during the final wash before placement on magnet rack. Remove as much of the final ethanol wash as possible. Air dry tubes for 4 minutes.
20. Take tubes off the magnet. Add 60 uL of 10 mM Tris-HCl, pH 8.0. Pipet to mix 10 times, and incubate at room temperature for 5 minutes to elute DNA. *Notes:* For maximum yield, vortex for 1 minute at 2000 rpm. Spin tubes briefly (tube top hinges facing outward in the centrifuge). Then let sit for another 5 minutes at room temperature to elute DNA.
21. Place the tubes back on the magnet for 4 minutes or until the solution clears.
22. Store eluted DNA at -20°C.

Amplicon library preparation protocol

Materials

1. Sera-Mag SPRI bead solution (recipe under pilot study methods above)
2. PEG/NaCl solution (recipe below)
3. 0.6 mL and 2 mL microcentrifuge tubes
4. 50 mL Falcon tubes
5. Microcentrifuge
6. Magnetic rack for 2 mL tubes (Invitrogen DynaMag-2 12321D)
7. Ethanol (100%)
8. UltraPure DNase/RNase-free distilled water(Invitrogen 10977-023)
9. P20, P200, and P1000 pipets and filter tips
10. 10 mM Tris-HCl pH 8
11. Qubit dsDNA High Sensitivity Assay Kit (Invitrogen Q32854)
12. NEBNext End Repair Module (NEB E6050S)
13. Cold centrifuge
14. Dry bath
15. NEBNext dA-Tailing Module (NEB E6053S)
16. TruSeq adapters (FC-121-2001 and FC-121-2002)
17. T4 DNA ligase (5 U/uL, Fermentas EL0011, supplied with 10X T4 DNA ligase buffer and 50% PEG-4000 solution)
18. Tween-20 (100%)
19. SYBR Gold nucleic acid gel stain (Invitrogen S-11494)
20. BenchTop 100 bp DNA Ladder (Promega G8291)
21. Agarose
22. Gel electrophoresis rig

PEG/NaCl solution recipe

1. Use filter pipette tips. Sera-Mag bead solution recipe:

Reagent	Mass	Volume
PEG-8000 powder (Promega V3011)	4.5 g	4.15 mL
5 M NaCl	-	5 mL
1 M Tris-HCl, pH 8.0	-	250 uL
0.5 M EDTA, pH 8.0	-	50 uL
ddH ₂ O		15.05 mL
Tween-20 (100%)		13.75 uL
TE		500 uL
Total		25 mL

2. Add 4.5 g PEG-8000 to 50 mL Falcon tube.
3. To the powder add NaCl, Tris-HCl, EDTA, and ddH₂O.
4. Shake Falcon tube until all PEG has dissolved.
5. Add Tween-20 and mix properly.
6. Add 500 uL TE.
7. Mix well. Wrap Falcon tube in aluminum foil and store at 4°C.

Amplicon SPRI bead cleanup

1. Pool 20 uL of triplicate PCRs for each sample.
2. Make a fresh dilution of 80% ethanol in a 50 mL Falcon tube.
3. Resuspend SPRI bead solution in Falcon tubes. Aliquot volume of SPRI bead solution required into 50 mL Falcon tube and let sit at room temperature for at least 30 minutes.
4. Aliquot 1.2x SPRI bead solution (bead solution to sample volume ratio) for each sample.
5. Vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly. Then let sit for another 10 minutes at room temperature.
6. Set on magnet rack for 3 minutes to separate SPRI beads from solution. Wait for solution to clear before proceeding.
7. Aspirate the cleared solution from the wells and discard. Do not disturb separated magnetic beads.
8. Add 200 uL of 80% ethanol. Incubate for 30 seconds.
9. Discard supernatant.
10. Add 200 uL of 80% ethanol. Incubate for 30 seconds.
11. Discard supernatant. Remove all traces of ethanol with P20 or P10 tips.
12. Air dry for 10 minutes.
13. Remove strip tubes from the magnetic rack and add 31 uL of 10 mM Tris-HCl pH 8.
14. Vortex for 1 minute at 2000 rpm. Spin tubes briefly. Let incubate for another 2 minutes at room temperature.
15. Place strip tubes onto magnetic rack for 1 minute.
16. Transfer elutant to storage tubes.

Pool amplicons into libraries

1. Qubit each sample using the Qubit dsDNA High Sensitivity Assay.
2. Pool samples into libraries, such that each library contains uniquely indexed samples.
3. Repeat SPRI bead cleanup if necessary to standardize each library volume to 30 uL 10 mM Tris-HCl pH 8.

End repair

1. Label 0.6 mL tube for master mix.
2. Flick and spin all reagents before use.
3. Mix the following components into the labelled 0.6 mL tube per library:

Reagent	Volume
Template DNA	30 uL
10X End Repair Reaction Buffer	10 uL
End Repair Enzyme Mix	5 uL
Nuclease-free water	65 uL
Total Volume	100 uL

4. Briefly flick and spin down master mix.
5. Aliquot 15 uL of master mix into each library.
6. Briefly flick and spin down library tubes.

7. Incubate in cold centrifuge for 30 minutes at 20°C.
8. Afterward, set temperature to 22°C for ligation reaction later.

SPRI bead cleanup

1. Add 300 uL (3x) of room temperature SPRI bead solution.
2. Vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly. Then let sit for another 10 minutes at room temperature.
3. Set on magnet rack for 3 minutes. Meanwhile, bring out A-tail reaction buffer to thaw on ice.
4. Discard supernatant.
5. Add 400 uL of 80% ethanol. Incubate for 30 seconds.
6. Discard supernatant.
7. Add 400 uL of 80% ethanol. Incubate for 30 seconds.
8. Discard supernatant.
9. Air dry for 10 minutes. Preheat digital dry bath to 37°C. Meanwhile, bring PEG/NaCl solution to warm to room temperature.
10. Add 42 uL of 10 mM Tris-HCl pH 8.
11. Vortex for 1 minute at 2000 rpm. Spin tubes briefly. Let incubate for another 2 minutes at room temperature.

dA-Tail

1. Label 0.6 mL tube for master mix.
2. Flick and spin down all reagents before use.
3. Mix the following components into the labelled 0.6 mL tube per library:

Reagent	Volume	MM ?x
End Repaired, Blunt DNA	42 uL	
NEBNext dA-Tailing Reaction Buffer (10X)	5 uL	uL
Klenow Fragment (3' → 5' exo ⁻)	3 uL	uL
Total volume	50 uL	

4. Briefly flick and spin down master mix.
5. Aliquot 8 uL of master mix per library.
6. Briefly flick and spin down library.
7. Incubate in a dry bath for 30 minutes at 37°C.

SPRI bead cleanup

1. Aliquot (2.2x) 110 uL PEG/NaCl solution.
2. Vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly. Then let sit for another 10 minutes at room temperature.
3. Set on magnet rack for 3 minutes. Meanwhile, bring out T4 DNA ligase buffer, PEG-4000, and adapter oligo mix out to thaw on ice.
4. Discard supernatant.
5. Add 400 uL of 80% ethanol. Incubate for 30 seconds.
6. Discard supernatant.

7. Add 400 uL of 80% ethanol. Incubate for 30 seconds.
8. Discard supernatant.
9. Air dry for 10 minutes.
10. Add 29.3 uL of 10 mM Tris-HCl pH 8.
11. Vortex for 1 minute at 2000 rpm. Spin tubes briefly. Let incubate for another 2 minutes at room temperature.

Adapter ligation

1. Label 0.6 mL tube for master mix.
2. This procedure uses a 10:1 molar ratio of adapter to DNA insert.
3. Prepare the following reaction mix on ice in the following order. If white precipitate is present in the 10X DNA ligase buffer after thawing, warm the buffer to 37°C and vortex until the precipitate has dissolved. Since PEG is highly viscous, vortex the master mix before adding T4 DNA ligase and mix gently thereafter. Mix adapters well (flick and spin down).

Reagent	Volume	Final conc.
DNA sample	29.3 uL	
T4 DNA ligase buffer, 10X	4 uL	1x
PEG-4000, 50%	4 uL	5%
PE adapter oligo mix, 15 uM each	[x uL]	10:1 molar ratio adapter:insert
DNA ligase, 5U/uL	1 uL	0.125 U/uL
Water	[x uL]	
Total volume	40 uL	

4. Briefly flick and spin down master mix.
5. Aliquot 9 uL of master mix per library.
6. Add unique adapters
7. Briefly flick and spin down.
8. Incubate for 30 minutes at 22°C.

SPRI bead cleanup

1. Add 10 uL of 10 mM Tris-HCl pH 8.0 to each library.
2. Aliquot (0.9x) 45 uL PEG/NaCl solution.
3. Vortex at 2000 rpm for 10 minutes at room temperature. Spin tubes briefly. Then let sit for another 10 minutes at room temperature.
4. Set on magnet rack for 3 minutes. Meanwhile, prep Qubit reagents (tubes, working solution, aliquot).
5. Discard supernatant.
6. Add 400 uL of 80% ethanol. Incubate for 30 seconds.
7. Discard supernatant.
8. Add 400 uL of 80% ethanol. Incubate for 30 seconds.
9. Discard supernatant.
10. Air dry for 10 minutes.

11. In the meanwhile, make 25 mL of Elution Buffer + Tween (10 mM Tris-HCl pH 8 + 0.05% Tween-20; EBT) in a clean 50 mL tube for qPCR later:

Reagent	Stock Conc.	Final Conc.	Volume
Tris-HCl pH 8	1 M	10 mM	250 uL
Tween-20	100%	0.05%	12.5 uL
ddH ₂ O			24.7375 mL

12. Add 30 uL of 10 mM Tris-HCl pH 8.
13. Vortex for 1 minute at 2000 rpm. Briefly spin down. Let incubate for another 2 minutes at room temperature.
14. Place on magnet rack for 4 minutes.
15. Transfer 29 uL library to new storage tube. Use the remaining 1 uL for Qubit quantification.

Gel electrophoresis

1. 3 g agarose in 150 mL TAE buffer, microwave until powder is completely dissolved, then cool on bench top for 5 minutes.
2. Add 15 uL SYBR Gold and swirl to mix.
3. Pour gel.
4. Use 3 uL BenchTop 100 bp DNA Ladder in 5 uL water and 2 uL 6x Loading Buffer.
5. Use 2 uL of 6x Loading Buffer for 0.5-1 uL of sample.
6. Load samples and ladder, leaving a lane between each.
7. 120 V for 120 min for best separation between 300, 360, and 420 bp (that is unligated amplicon, single adapter ligated amplicon, and double adapter ligated amplicon, respectively).

Results

Pilot study comparison of sample preservation and metagenomic DNA extraction methods

Literature review of DNA extraction from avian fecal and swab samples

A total of 30 studies fell within the criteria of the literature review for candidate extraction methods. The study goals ranged from avian genotyping (6/30) to dietary analysis (5/30), pathogen detection (12/30), and microbiota characterization (7/30; Table S2.1). Many studies focused on samples from chickens (8/30), and interestingly, few studies included Passerine birds (5/30). Among studies on fecal samples, the most common DNA extraction method employed was the Qiagen QIAamp DNA Stool Mini Kit (12/23), and the second most common method was the Qbiogene FastDNA Kit (3/23). Most fecal samples were stored frozen (8/23) and dry (i.e., without preservation buffer; 9/23). Studies on swab samples exhibited high variation in extraction methods, and only the Qiagen DNeasy Blood and Tissue Kit was observed in more than one study (2/7). Swabs were most commonly stored dry (4/7) and sometimes frozen (3/7).

Sample preservation and metagenomic DNA extraction of fecal samples

Methods evaluation across intra-sample replicates.—Among samples tested across the UltraClean Soil, UltraClean Fecal, and PowerSoil kits used with Vortex homogenization, Sample 12 stored dry at -20°C amplified in neither 16S nor COI PCRs, but Sample 13 stored at -80°C amplified in 16S PCRs (Table S2.2a). Additionally, among the Sample 13 extracts, only that of the PowerSoil kit yielded amplicons in COI PCRs. For the SPRI-based extraction method with the Precellys homogenizer on Sample 25 stored in DNA buffer, both extracts amplified strongly in 16S PCRs, and the addition of a flocculation step substantially improved amplification success in COI PCRs, though at the cost of decreased DNA yield.

When varying sample masses within method, sample 6 stored dry at -80°C and extracted with the Xpedition Soil/Fecal MiniPrep kit using the Vortex adapter performed better in both 16S and COI PCRs using a smaller starting sample mass in the extraction. Interestingly, nearly an order of magnitude less starting material only decreased the DNA yield by about 47%. In contrast, for Sample 19 stored dry at -20°C and extracted using the PowerPlant Pro Kit with pre-beat heated incubation and the Vortex adapter, nearly five times more DNA was extracted from a 0.02 g increase starting sample, and a concomitant increase in 16S amplification success was observed.

Methods evaluation with unreplicated fecal samples.—With only the Vortex adapter as the homogenizer, the PowerSoil and PowerPlant Pro kits appeared to perform the best in terms of amplification success in both 16S and COI PCRs for fecal samples frozen dry. All other kits (Xpedition Soil/Fecal, PowerSoil, PowerMicrobiome, and Qiagen Stool) yielded fair amplification in 16S PCRs yet sporadic or typically no amplification in COI PCRs for samples frozen dry or with DNA buffer (Table S2.2b). Of note, SPRI-based fecal extracts provided high DNA yields and with flocculation, exhibited the strongest amplification success relative to all other methods (Table S2.2a).

Methods evaluation with independent sample replicates.—Between the Easy DNA, Xpedition Soil/Fecal, and PowerSoil, only the Xpedition and PowerSoil extracts yielded amplification success, and only the PowerSoil extracts amplified in both 16S and COI PCRs. These results may have been in part due to differences in storage conditions and homogenizer

use, as the Easy DNA and Xpedition samples were stored dry at -20°C whereas the PowerSoil samples were frozen in DNA buffer. Similarly, the Easy DNA samples were not homogenized whereas the Xpedition samples were Vortexed and the PowerSoil samples were processed on the Precellys.

Sample preservation and metagenomic DNA extraction of cloacal samples

All cloacal swab samples were stored under the same conditions (in DNA buffer initially at -20°C and then at -80°C), thereby precluding evaluation of alternative storage conditions. However, for extraction between PowerSoil, SPRI-based with flocculation, and PowerMicrobiome extractions, the first two methods yielded high DNA quantities and amplification in both 16S and COI PCRs (Table S2.2c). Notably, all positive COI PCRs for swab sample extracts exhibited non-specific amplification. Therefore, although they received high PCR index ratings for amplifying, these COI PCRs were not considered successful. The PowerMicrobiome kit was paired with the Vortex homogenizer whereas the other two methods were paired with the Precellys, which may explain its poorer extraction performance as it only very weakly amplified in 16S PCRs and did not amplify in COI PCRs.

Table S2.1. Literature review of avian studies that have extracted DNA from either fecal or swab samples.

Purpose	Method	Samples	Sample Storage	Bird Species	Sequence Analysis	References
Diet						
	QIAamp DNA Stool Mini Kit (Qiagen) and Xpedition Soli/Fecal DNA MiniPrep Kit (Zymo)	Feces	Dry, -80°C	Western Bluebird	Sanger Sequencing	Jedlicka <i>et al.</i> 2013
	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	n/a	Cockatiels	Sanger Sequencing	Joo & Park 2012
	CTAB, UltraClean Fecal DNA Kit (MoBio), ExtractMaster Fecal DNA Extraction Kit (Epicentre), QIAamp DNA Stool Mini Kit (Qiagen)	Feces	frozen or 70% ethanol	Carrion Crows	None	Oehm <i>et al.</i> 2011
	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	95% ethanol	Little Penguin	Pyrosequencing	Deagle <i>et al.</i> 2010
	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	70% ethanol, 4°C	Macaroni Penguin	Sanger Sequencing	Deagle <i>et al.</i> 2007
Target Parasite						
Mycobacterium avium	QIAamp Blood Kit (Qiagen)	Feces	n/a	Cranes, geese, penguins, flamingos, Galliforms, Ciconiids, emus, parrots	Sanger Sequencing	Münster <i>et al.</i> 2013
Microsporidia	DNA Sorb B	Feces	Dry, -18°C	Great cormorant, great crested grebe, white stork	Sanger Sequencing	Malčková <i>et al.</i> 2013
Chlamydia psittaci	High Pure PCR Template Preparation Kit (Roche)	Feces and cloacal swabs		Pigeons	None	Sachse <i>et al.</i> 2012
Brachyspira intermedia and pilosicoli	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	n/a	Chickens	None	Myers <i>et al.</i> 2010
Cryptosporidium spp.	NucliSens Isolation Kit	Feces	n/a	Chicken, duck, "other birds"	Sanger Sequencing	Hsu <i>et al.</i> 2008
Brachyspira intermedia and pilosicoli	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	Dry, <6 hrs at 4°C	Chicken	Sanger Sequencing	Phillips <i>et al.</i> 2006

Salmonella	DTT boil, Pro-K/phenol/chloroform, GeneClean	Feces	n/a	Chicken	None	Mahon & Lax 1993
Chlamydia psittaci	Instagene matrix (Bio-Rad)	Cloacal swab	dry	Bellbird, rifleman, hihi, whitehead	Sanger Sequencing	Gartrell <i>et al.</i> 2013
<i>Chlamydiaceae</i>	DNeasy Blood & Tissue Kit (Qiagen)	Cloacal swab	Dry, 4°C	12 spp raptors	Sanger Sequencing	Ortega <i>et al.</i> 2012
Psittacine beak and feather disease	Chelex 100 Resin	Cloacal swab	-20°C	Budgerigars	Sanger Sequencing	Hess <i>et al.</i> 2004
Duck plague virus	Phenol-ether	Cloacal and oral-pharyngeal swabs	Viral transport medium	Waterfowl (27 spp)	None	Hansen <i>et al.</i> 2000
Avian polyomavirus	Boil and Sephadex G-50 minicolumn	Cloacal swab	n/a	Lovebird, white-capped pionus parrot, Indian ring-neck parakeet, cockatoo	None	Phalen <i>et al.</i> 1997
Host DNA						
Sex determination	DNeasy Blood & Tissue Kit (Qiagen)	Feces		Cockatiel	Sanger Sequencing	Amada 2012
Genotyping	QIAamp DNA Stool Mini Kit (Qiagen)	Dried feces	Dry, -40°C	Red-legged Partridge	Sanger Sequencing	Guerrini & Barbanera 2009
Genotyping	QIAamp DNA Stool Mini Kit (Qiagen)	Dried feces	Dry, room temperature	capercaillie	None	Regnaut <i>et al.</i> 2006
Genotyping	Phenol/chloroform, DNAzol, GuSCN/silica membrane	Feces	90% EtOH, -20°C	Great bustard, bald ibis	None	Idaghdour <i>et al.</i> 2003
Host identification	QIAamp DNA Stool Mini Kit (Qiagen)	Feces, cloacal swabs	Viral transport medium, -80°C	Ducks, shorebirds, seabirds, pigeon, chicken	None	Cheung <i>et al.</i> 2009
Sex determination	DNeasy Blood & Tissue Kit (Qiagen)	Buccal swabs	-20°C	Red junglefowl, griffon vulture, booted eagle, Northern goshawk, tawny owl, common swift, red-billed chough, carrion crow, Eurasian skylark, dartford warbler, European robin, zebra finch, star finch, Gouldian finch	Sanger Sequencing	Morinha <i>et al.</i> 2013
Microbial						

Community						
	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	No storage	Chicken	Pyrosequencing	Singh <i>et al.</i> 2012
	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	Dry, 2-4°C for < 24 hrs	Zebra Finch	Sanger Sequencing	Benskin <i>et al.</i> 2010
	FastDNA Kit (Qbiogene)	Feces	Dry, -80°C	Canada Goose	Sanger Sequencing	Lu <i>et al.</i> 2009
	FastDNA Kit (Qbiogene)	Feces	Dry, -80°C	Turkey	Sanger Sequencing	Lu & Domingo 2008
	Fecal Kit (MoBio) or FastDNA (Q-Biogene)	Feces	Dry, -80°C	Chicken	Sanger Sequencing	Lu <i>et al.</i> 2007
	QIAamp DNA Stool Mini Kit (Qiagen)	Feces	Frozen	Chicken	None	Chambers <i>et al.</i> 2001
	UltraClean Soil DNA Kit (MoBio)	Cloacal swab	Sucrose lysis buffer, -20°C	Spotted towhee	None	Klomp <i>et al.</i> 2008

Table S2.2. Avian fecal and swab sample storage and DNA extraction parameters evaluated along with their corresponding performance indices of DNA quality, yield, and PCR success at two genetic loci, microbial 16S rRNA and arthropod COI (where 4 indicates the most robust amplification score) for the comparative pilot study. Results are organized by comparisons of specific combinations of storage and extraction methods evaluated (a) across intra-sample replicates, (b) only once, and (c) with independent sample replicates. SPRI = solid phase reversible immobilization method developed in this study. BME = recommended protocol with beta-mercaptoethanol added to the lysis buffer. Beater = homogenizer used for mechanical lysis of sample in lysis buffer. Pre-beat incubation = 70°C for 15 minutes before mechanical lysis.

Comparison	Sample ID	Extraction Method	Prep Mass (g)	Storage	Beater	A _{260/280}	A _{260/230}	DNA Yield (ng/g feces or ng/swab)	16S PCR	COI PCR
(a) Methods evaluated across intra-sample fecal replicates										
Fecal extraction methods	12	Mo Bio UltraClean Soil Kit	0.25	dry, -80°C	Vortex	0.38	0.36	410.00	1	0
		Mo Bio UltraClean Fecal Kit	0.25	dry, -80°C	Vortex	0.39	0.35	34.35	1	0
		Mo Bio PowerSoil Kit	0.25	dry, -80°C	Vortex	-14.54	0.43	2.10	2	4
	13	Mo Bio UltraClean Soil Kit	0.1	dry, -20°C	Vortex	0.66	0.29	4.22	0	0
		Mo Bio UltraClean Fecal Kit	0.1	dry, -20°C	Vortex	0.61	0.65	9.65	0	0
		Mo Bio PowerSoil Kit	0.1	dry, -20°C	Vortex	0.36	3.34	0.28	0	0
	25	SPRI with flocculation	0.08	DNA buffer, -20°C, -80°C	Precellys	1.18	0.21	166.00	4	3
		SPRI	0.08	DNA buffer, -20°C, -80°C	Precellys	1.18	0.38	620.00	4	0
	Fecal prep masses	6	Zymo Xpedition Soil/Fecal MiniPrep Kit (with	0.173	dry, -80°C	Vortex	1.13	0.39	1.87	1

		BME in lysis buffer)	0.029	dry, -80°C	Vortex	1.11	0.6	1.00	2	1
		Mo Bio PowerPlant Pro Kit (with pre-beat incubation)	0.1	dry, -20°C	Vortex	0.96	0.4	25.45	1	0
	19		0.08	dry, -20°C	Vortex	0.93	0.63	4.70	0	0
(b) Methods evaluated with a single fecal sample										
Storage conditions and prep masses	4	Zymo Xpedition Soil/Fecal MiniPrep Kit, with BME	0.23	dry, -80°C	Vortex	1.69	0.4	2.26	2	0
	5		0.23	dry, -20°C	Vortex	1.18	0.61	1.54	2	0
	7		0.1	DNA buffer, -80°C	Vortex	1.61	0.31	1.92	1	0.5
	8		0.031	DNA buffer, -80°C	Vortex	1.02	0.59	1.02	0.5	0
Storage conditions and prep masses	14	Mo Bio PowerSoil Kit	0.2	dry, -20°C	Vortex	1.45	0.36	1.365	1	3
	15		0.07	dry, -80°C	Vortex	2.67	0.09	3.565	3	3
With pre-lysis incubation at 70C for 15 min	18	Mo Bio PowerPlant Pro Kit	0.05	dry, -20°C	Vortex	1.18	0.8	63.50	2	1
Storage conditions	20	Mo Bio PowerMicrobiome Kit	0.14	dry, -20°C	Vortex	0.36	0.5	2.19	0	0
	21		0.14	dry, -80°C	Vortex	0.31	0.45	17.25	1	0
Recommended parameters	22	Qiagen Stool Mini Kit, human protocol	0.16	DNA buffer, -20°C	None	1.45	0.51	5.80	3	1
With mechanical lysis and BME, different prep masses	23		0.18	DNA buffer, -20°C	Vortex	1.25	0.43	11.50	3	0
	24		0.28	DNA buffer, -20°C	Vortex	1.18	0.36	16.90	3	0
(c) Methods evaluated with independent sample replicates										

Fecal samples	9	Invitrogen Easy DNA Kit	0.1	dry, -20°C	None	1.09	0.38	151.00	0	0
	10		0.08	dry, -20°C	None	0.03	-0.01	251.00	0	0
	11		0.1	dry, -20°C	None	0.83	0.73	56.50	0	0
	1	Zymo Xpedition Soil/Fecal MiniPrep Kit	0.25	dry, -20°C	Vortex	-21.07	-0.29	1.67	2	0
	2		0.25	dry, -20°C	Vortex	1.35	0.61	18.20	2	0
	3		0.25	dry, -20°C	Vortex	1.36	0.67	7.73	2	0
	16	Mo Bio PowerSoil Kit with BME	0.04	DNA buffer, -20°C, -80°C	Precellys	1.01	0.36	4.08	2	3
17	0.04		DNA buffer, -20°C, -80°C	Precellys	0.93	0.35	31.4	1	0	
Cloacal swab samples	26	Mo Bio PowerSoil Kit with BME	None	DNA buffer, -20°C, -80°C	Precellys	1.46	1.09	705	2	3
	27		None	DNA buffer, -20°C, -80°C	Precellys	1.36	0.55	43.6	2	3
	28	Mo Bio PowerMicrobiome Kit with BME	None	DNA buffer, -20°C, -80°C	Vortex	1.35	0.36	6.1	1	0
	29		None	DNA buffer, -20°C, -80°C	Vortex	1.4	0.42	102.5	2	3
	30	SPRI with flocculation	None	DNA buffer, -20°C, -80°C	Precellys	1.43	0.52	183.6	2	3
	31		None	DNA buffer, -20°C, -80°C	Precellys	1.74	1.64	1524	2	3

Table S2.3. All barcoded primers used in this study.

Oligo sequence (5' to 3')	Barcode (first 6 bases at 5' end of primer)	Oligo name
Arthropod COI PCR Primers		
AGAGAGAGATATTGGAACWTTATATTTTATTTTGG	AGAGAG	ZAF1
AGAGAGWACTAATCAATTWCCAAATCCTCC	AGAGAG	ZAR1
TGCTACAGATATTGGAACWTTATATTTTATTTTGG	TGCTAC	ZAF2
TGCTACWACTAATCAATTWCCAAATCCTCC	TGCTAC	ZAR2
GACTACTAGATATTGGAACWTTATATTTTATTTTGG	GACTACT	ZAF3
GACTACTWACTAATCAATTWCCAAATCCTCC	GACTACT	ZAR3
GTGCATAGATATTGGAACWTTATATTTTATTTTGG	GTGCAT	ZAF4
GTGCATWACTAATCAATTWCCAAATCCTCC	GTGCAT	ZAR4
ACACATAGATATTGGAACWTTATATTTTATTTTGG	ACACAT	ZAF5
ACACATWACTAATCAATTWCCAAATCCTCC	ACACAT	ZAR5
ACCTTGAGATATTGGAACWTTATATTTTATTTTGG	ACCTTG	ZAF6
ACCTTGWACTAATCAATTWCCAAATCCTCC	ACCTTG	ZAR6
AATGCCAGATATTGGAACWTTATATTTTATTTTGG	AATGCC	ZAF7
AATGCCWACTAATCAATTWCCAAATCCTCC	AATGCC	ZAR7
GTACGCAGATATTGGAACWTTATATTTTATTTTGG	GTACGC	ZAF8
GTACGCWACTAATCAATTWCCAAATCCTCC	GTACGC	ZAR8
GCGTACAGATATTGGAACWTTATATTTTATTTTGG	GCGTAC	ZAF9
GCGTACWACTAATCAATTWCCAAATCCTCC	GCGTAC	ZAR9
TACCATAGATATTGGAACWTTATATTTTATTTTGG	TACCAT	ZAF10
TACCATWACTAATCAATTWCCAAATCCTCC	TACCAT	ZAR10
Bacterial 16S rRNA PCR Primers		
CGTCATGTGCCAGCMGCCGCGGTAA	CGTCAT	515F1
CGTCATGGACTACHVGGGTWTCTAAT	CGTCAT	806R1
AGAGAGGTGCCAGCMGCCGCGGTAA	AGAGAG	515F2
AGAGAGGGACTACHVGGGTWTCTAAT	AGAGAG	806R2
CACTGTGTGCCAGCMGCCGCGGTAA	CACTGT	515F3
CACTGTGGACTACHVGGGTWTCTAAT	CACTGT	806R3

GTCTCAGTGCCAGCMGCCGCGGTAA	GTCTCA	515F4
GTCTCAGGACTACHVGGGTWTCTAAT	GTCTCA	806R4
AGCATCGTGCCAGCMGCCGCGGTAA	AGCATC	515F6
AGCATCGGACTACHVGGGTWTCTAAT	AGCATC	806R6
GACACTGTGCCAGCMGCCGCGGTAA	GACACT	515F7
GACACTGGACTACHVGGGTWTCTAAT	GACACT	806R7
ACACATGTGCCAGCMGCCGCGGTAA	ACACAT	515F8
ACACATGGACTACHVGGGTWTCTAAT	ACACAT	806R8
ATGTGTGTGCCAGCMGCCGCGGTAA	ATGTGT	515F9
ATGTGTGGACTACHVGGGTWTCTAAT	ATGTGT	806R9
AGTACTGTGCCAGCMGCCGCGGTAA	AGTACT	515F10
AGTACTGGACTACHVGGGTWTCTAAT	AGTACT	806R10
GATATCGTGCCAGCMGCCGCGGTAA	GATATC	515F11
GATATCGGACTACHVGGGTWTCTAAT	GATATC	806R11
CAACTAGTGCCAGCMGCCGCGGTAA	CAACTA	515F12
CAACTAGGACTACHVGGGTWTCTAAT	CAACTA	806R12
GCATGTGTGCCAGCMGCCGCGGTAA	GCATGT	515F13
GCATGTGGACTACHVGGGTWTCTAAT	GCATGT	806R13
TGCTACGTGCCAGCMGCCGCGGTAA	TGCTAC	515F14
TGCTACGGACTACHVGGGTWTCTAAT	TGCTAC	806R14
GTGCATGTGCCAGCMGCCGCGGTAA	GTGCAT	515F15
GTGCATGGACTACHVGGGTWTCTAAT	GTGCAT	806R15
CTGACTGTGCCAGCMGCCGCGGTAA	CTGACT	515F16
CTGACTGGACTACHVGGGTWTCTAAT	CTGACT	806R16
CGATTGGTGCCAGCMGCCGCGGTAA	CGATTG	515F17
CGATTGGGACTACHVGGGTWTCTAAT	CGATTG	806R17
CCAGTTGTGCCAGCMGCCGCGGTAA	CCAGTT	515F18
CCAGTTGGACTACHVGGGTWTCTAAT	CCAGTT	806R18
CTTGAGGTGCCAGCMGCCGCGGTAA	CTTGAG	515F19
CTTGAGGGACTACHVGGGTWTCTAAT	CTTGAG	806R19

AATGCCGTGCCAGCMGCCGCGGTAA	AATGCC	515F20
AATGCCGGACTACHVGGGTWTCTAAT	AATGCC	806R20
GGTGTTGTGCCAGCMGCCGCGGTAA	GGTGTT	515F21
GGTGTTGGACTACHVGGGTWTCTAAT	GGTGTT	806R21

Table S2.4. Avian fecal and swab sample DNA extraction parameters evaluated along with their corresponding performance indices of DNA quality, yield, and PCR success at the microbial 16S rRNA (where 0 indicates no band and 4 indicates the most robust amplification score) for the PowerSoil Kit (Mo Bio) and the SPRI-based method. All samples were stored frozen in DNA buffer and homogenized on the Precellys 24. SPRI = solid phase reversible immobilization method developed in this study. BME = beta-mercaptoethanol added in lysis buffer.

Protocol	Sample Type	Fecal Mass (g)	Extraction Parameters	A _{260/280}	A _{260/230}	Total DNA Yield (ng/g or swab)	Triplicate PCR yield (ng)	16S PCR Index
SPRI	13 fecal	0.08 ± 0.002	with inhibitor removal	1.13 ± 0.10	0.46 ± 0.05	35.53 ± 15.03	232.15 ± 161.18	3.08 ± 1.75
	5 cloacal swab with fecal debris	None		1.05 ± 0.52	0.8 ± 0.67	51.20 ± 62.78	75.99 ± 36.26	3.20 ± 1.79
	18 cloacal swab	None		1.10 ± 0.43	0.59 ± 0.45	261.12 ± 390.08	143.96 ± 102.68	3.33 ± 1.33
	18 oral swab	None	no inhibitor removal	1.86 ± 0.25	0.84 ± 0.30	233.52 ± 142.83	109.03 ± 139.98	2.67 ± 1.46
PowerSoil	18 fecal	0.08 ± 0.005	with BME	1.33 ± 0.26	0.42 ± 0.07	165.96 ± 183.49	554.4 ± 333.58	3.30 ± 0.76
	18 cloacal swab	None		0.98 ± 0.33	0.37 ± 0.15	8.31 ± 11.66	116.97 ± 83.96	1.61 ± 0.87
	18 oral swab	None		0.79 ± 1.94	0.07 ± 0.13	3.38 ± 2.04	108.30 ± 82.49	1.56 ± 0.77

Table S2.5. Summary of sequencing performance across all three paired-end 250 bp MiSeq runs.

Analytical Step	SPRI	PowerSoil	Xpedition
Total raw reads	12.4 million	13.4 million	17.4 million
Reads passed filter	10.9 million	11.9 million	15.3 million
Reads assembled with correct barcodes and primers	6.4 million	7.7 million	8.4 million
Reads after QIIME split libraries step	6.1 million	7.3 million	8.2 million

Table S2.6. Recovery of starting material after application of SPRI beads used at three times the sample volume to concentrate amplicons with minimum sample loss as well as after using the concentrated amplicons in the modified with-bead, PCR-free library preparation protocol developed in this study (adjusted for adapter ligation).

Sample	Input DNA (ng)	After concentration		After library prep	
		Eluted DNA (ng)	% Yield	Eluted DNA (ng)	% Yield
1	489.09	402.48	82.29	325.71	80.93
2	416.99	350.88	84.15	265.71	75.73
3	415.05	407.64	98.21	282.86	69.39
4	493.2	383.56	77.77	293.57	76.54
5	493.2	466.12	94.51	327.86	70.34
6	493.2	466.12	94.51	355.71	76.31
7	425.14	356.04	83.75	282.86	79.45
8	493.2	397.32	80.56	295.71	74.43
9	460.48	412.8	89.65	293.57	71.12

Figures

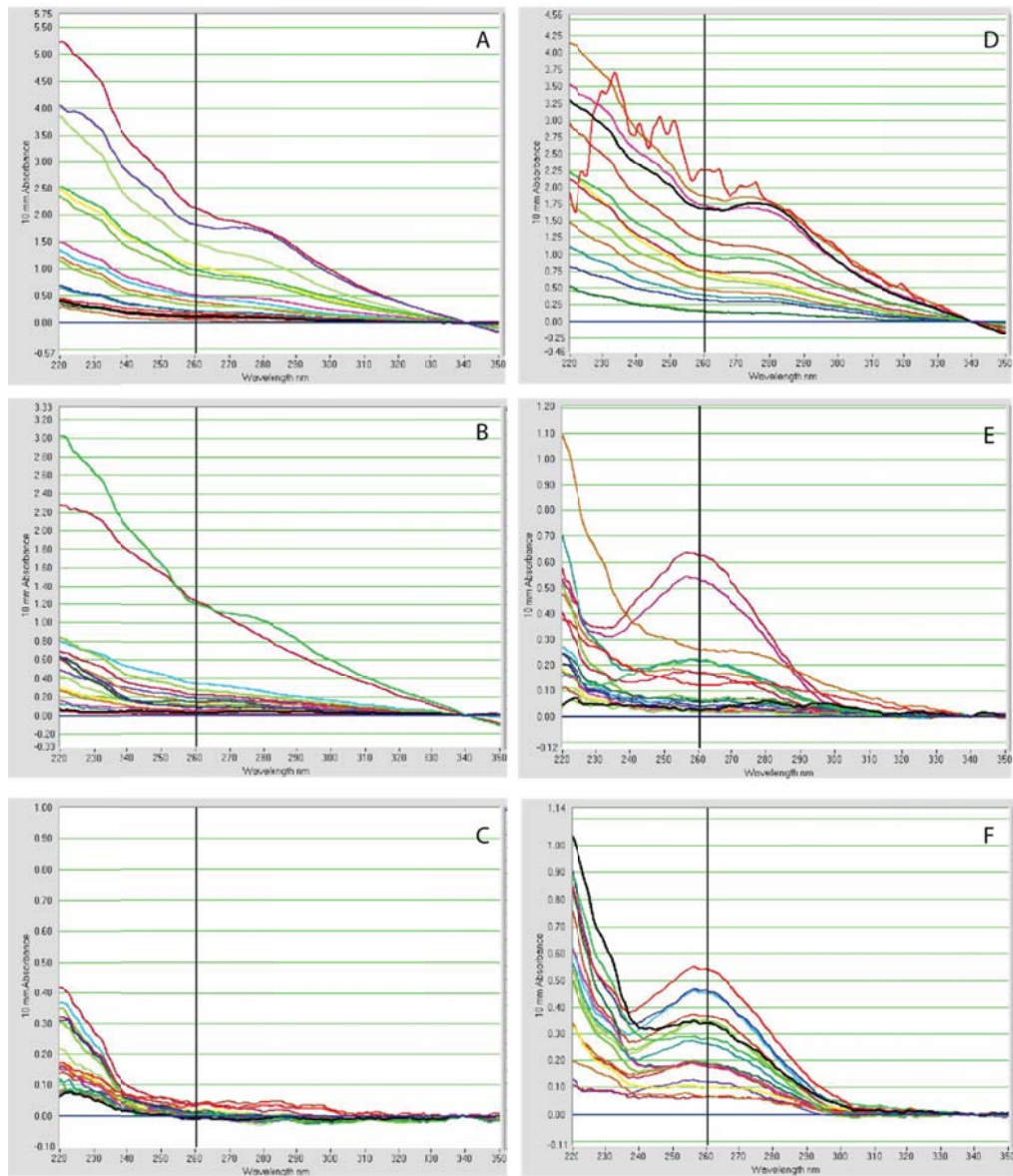


Fig. S2.1. NanoDrop results of metagenomic DNA quality for fecal (A, D), cloacal swab (B, E), and oral swab (C, F) samples extracted using the Mo Bio PowerSoil (A-C) and SPRI-based (D-F) methods. Each line represents a different sample.

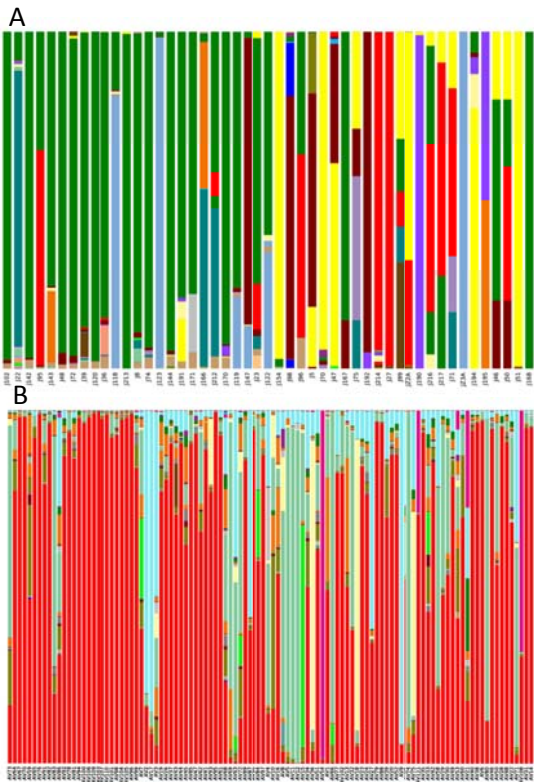


Fig. S2.2. Amplicon diversity for A) the arthropod COI locus in avian fecal samples extracted using the Zymo Xpedition Fecal/Soil Miniprep Kit and B) the bacterial 16S rRNA locus in avian fecal, cloacal swab, and oral swab samples extracted using the Mo Bio PowerSoil Kit and the SPRI-based method developed in this study. Each bar represents a sample, and each color represents a different taxon.

References

- Amada K (2012) Sex determination using *chd1* genes in Cockatiel *Nymphicus hollandicus* by PCR amplification of fecal DNA. *Ornithological Science*, **11**, 65–68.
- Benskin CMH, Rhodes G, Pickup RW, Wilson K, Hartley IR (2010) Diversity and temporal stability of bacterial communities in a model passerine bird, the zebra finch. *Molecular Ecology*, **19**, 5531–5544.
- Chambers PA, Duggan PS, Forbes JM, Heritage J (2001) A rapid, reliable method for the extraction from avian faeces of total bacterial DNA to be used as a template for the detection of antibiotic resistance genes. *Journal of Antimicrobial Chemotherapy*, **47**, 241–243.
- Cheung PP, Leung YHC, Chow C-K *et al.* (2009) Identifying the species-origin of faecal droppings used for avian influenza virus surveillance in wild-birds. *Journal of Clinical Virology*, **46**, 90–93.
- Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics*, **11**, 2039–2048.
- Deagle BE, Gales NJ, Evans K *et al.* (2007) Studying seabird diet through genetic analysis of faeces: a case study on Macaroni Penguins (*Eudyptes chrysolophus*). *PLoS ONE*, **2**, e831.
- Gartrell BD, French NP, Howe L *et al.* (2013) First detection of *Chlamydia psittaci* from a wild native passerine bird in New Zealand. *New Zealand Veterinary Journal*, **61**, 174–176.
- Guerrini M, Barbanera F (2009) Noninvasive genotyping of the Red-Legged Partridge (*Alectoris rufa*, Phasianidae): semi-nested PCR of mitochondrial DNA from feces. *Biochemical Genetics*, **47**, 873–883.
- Hansen WR, Nashold SW, Docherty DE, Brown SE, Knudson DL (2000) Diagnosis of duck plague in waterfowl by polymerase chain reaction. *Avian Diseases*, **44**, 266–274.
- Hess M, Scope A, Heincz U (2004) Comparative sensitivity of polymerase chain reaction diagnosis of psittacine beak and feather disease on feather samples, cloacal swabs and blood from budgerigars (*Melopsittacus undulates*, Shaw 18005). *Avian Pathology*, **33**, 477–481.
- Hsu B-M, Wun H-Y, Hsu C-LL (2008) Detection and species identification of *Cryptosporidium* from Taiwan feeding animals. *Journal of Parasitology*, **94**, 252–256.
- Idaghdour Y, Broderick D, Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustards. *Conservation Genetics*, **4**, 789–792.
- Jedlicka JA, Sharma AM, Almeida RPP (2013) Molecular tools reveal diets of insectivorous birds from predator fecal matter. *Conservation Genetics Resources*, **5**, 879–885.
- Joo S, Park S (2012) Identification of bird species and their prey using DNA barcode on feces from Korean traditional village groves and forests (maeulsoop). *Animal Cells and Systems*, **16**, 488–497.
- Klomp JE, Murphy MT, Smith SB *et al.* (2008) Cloacal microbial communities of female spotted towhees *Pipilo maculatus*: microgeographic variation and individual sources of variability. *Journal of Avian Biology*, **39**, 530–538.
- Lu J, Domingo JS (2008) Turkey fecal microbial community structure and functional gene diversity revealed by 16S rRNA gene and metagenomic sequences. *The Journal of Microbiology*, **46**, 469–477.

- Lu JR, Domingo JWS, Hill S, Edge TA (2009) Microbial diversity and host-specific sequences of Canada goose feces. *Applied and Environmental Microbiology*, **75**.
- Lu J, Santo Domingo J, Shanks OC (2007) Identification of chicken-specific fecal microbial sequences using a metagenomic approach. *Water Research*, **41**, 3561–3574.
- Mahon J, Lax AJ (1993) A quantitative polymerase chain reaction method for the detection in avian faeces of Salmonellas carrying the spvR Gene. *Epidemiology and Infection*, **111**, 455–464.
- Malčeková B, Valenčáková A, Molnár L, Kočíšová A (2013) First detection and genotyping of human-associated microsporidia in wild waterfowl of Slovakia. *Acta Parasitologica*, **58**, 13–17.
- Morinha F, Travassos P, Seixas F *et al.* (2013) High-resolution melting analysis for bird sexing: a successful approach to molecular sex identification using different biological samples. *Molecular Ecology Resources*, **13**, 473–483.
- Münster P, Fechner K, Völkel I, von Buchholz A, Czerny C-P (2013) Distribution of Mycobacterium avium ssp. paratuberculosis in a German zoological garden determined by IS900 semi-nested and quantitative real-time PCR. *Veterinary Microbiology*, **163**, 116–123.
- Myers SE, Dunn PA, Phillips ND, La T, Hampson DJ (2010) *Brachyspira intermedia* and *Brachyspira pilosicoli* are commonly found in older laying flocks in Pennsylvania. *Avian Diseases*.
- Oehm J, Juen A, Nagiller K, Neuhauser S, Traugott M (2011) Molecular scatology: how to improve prey DNA detection success in avian faeces? *Molecular Ecology Resources*, **11**, 620–628.
- Ortega N, Apaza D, Gonzalez F, Salinas J, Caro MR (2012) Occurrence of Chlamydiaceae in non-symptomatic free-living raptors in Spain. *European Journal of Wildlife Research*, **58**, 351–355.
- Phalen DN, Wilson VG, Graham DL (1997) Prevalence of neutralizing antibody and virus shedding in Psittacine birds infected with avian polyomavirus. *Journal of Avian Medicine and Surgery*, **11**, 98–104.
- Phillips ND, La T, Hampson DJ (2006) Development of a two-step nested duplex PCR assay for the rapid detection of *Brachyspira pilosicoli* and *Brachyspira intermedia* in chicken faeces. *Veterinary Microbiology*, **116**, 239–245.
- Regnaut S, Lucas FS, Fumagalli L (2006) DNA degradation in avian faecal samples and feasibility of non-invasive genetic studies of threatened capercaillie populations. *Conservation Genetics*, **7**, 449–453.
- Sachse K, Kuehlewind S, Ruetzger A, Schubert E, Rohde G (2012) More than classical *Chlamydia psittaci* in urban pigeons. *Veterinary Microbiology*, **157**, 476–480.
- Singh KM, Shah T, Deshpande S *et al.* (2012) High through put 16S rRNA gene-based pyrosequencing analysis of the fecal microbiota of high FCR and low FCR broiler growers. *Molecular Biology Reports*, **39**, 10595–10602.

Chapter 3: Host-microbial relationships in a changing world: determinants of natural variation in the avian gut microbiota

Abstract

Understanding the role of environmental heterogeneity in mediating ecological processes is becoming increasingly important to both recognize and predict the impacts of global change. As a growing body of work has demonstrated the significant interdependence of gut microbiota and host health, this study investigated the process of gut microbial community assembly in a natural population of avian hosts. To identify relevant mechanisms underlying variation in the avian gut microbiota, this study pursued longitudinal sampling and deep sequencing of 16S rRNA amplicons to follow microbial shifts in the developing gastrointestinal tract of Western Bluebirds *Sialia mexicana*. Samples for microbiota characterization were collected from nestlings aged 3–19 days post-hatching as well as female breeding adults over the course of three breeding seasons (2011–2013). The results showed high similarity between oral and cloacal microbiota relative to fecal microbiota as well as increased similarity among all gut sites over the nestling period. Shared nest environment and geographic distance only weakly explained variation in microbiota between individuals. Instead, ambient temperature and sample year were significant predictors of variation in microbiota across all gut sites. Environmental fluctuations appeared to drive bottom up trophic effects through decreased precipitation, primary productivity, and arthropod prey availability from 2011 to 2013. The lowest reproductive success among Western Bluebirds was observed during 2013, the most extreme drought year. Correspondingly, signatures in the microbiota unique to the 2013 breeding season were recovered, likely influenced by both temperature and dietary shifts. Overall, the patterns were consistent with the major role of environmental selection in driving variation in microbiota between gut sites within individuals as well as between microhabitats among individuals. These results point to the likelihood that global change will impact avian-microbial interactions, and further studies are warranted to clarify whether environmentally-mediated shifts in gut microbial composition harbor biological significance for host physiology or fitness.

Introduction

A growing appreciation for the enormous numbers of bacteria that reside in host organisms has sparked inquiry into the processes mediating assembly of the vertebrate gut microbiota (Ley *et al.* 2008; McFall-Ngai *et al.* 2013). Much of our current knowledge on the mechanisms underlying gut microbial assembly has been derived from association studies in humans and controlled experiments in laboratory rodents, including the role of diet (Turnbaugh *et al.* 2009b), shared environment (Spor *et al.* 2011), and host genetics (Benson *et al.* 2010). However, it is difficult to translate the implications of these studies more broadly to free-living vertebrate populations, as the spatiotemporal heterogeneity encompassed in natural environments remains unaddressed in previous studies of gut microbiota (but see Carey *et al.* 2013).

Understanding the role of environmental heterogeneity in structuring gut microbiota among free-living organisms is becoming increasingly important in light of the role gut microbes play in host fitness (Clemente *et al.* 2012). Specifically, the existence of extensive interactions

between hosts, microbes, and their environment raises compelling questions regarding the potential for instability in abiotic and biotic environmental conditions over developmental, seasonal, and interannual time scales to impact host-microbial relationships. A greater understanding of the natural variation and response of host-associated microbiota to environmental fluctuations is warranted to support both recognition and prediction of the extent to which natural versus anthropogenic environmental change may disrupt host-symbiont interdependencies and thus host fitness, host population dynamics, and ecosystem function.

Island biogeography and metacommunity ecology provide useful frameworks for considering assembly of the gut microbiota (Mihaljevic 2012). Within host organisms, separate body sites may be considered static patches of island habitats, wherein each site can exhibit strong environmentally selective regimes. Within a population of hosts, each individual can likewise be regarded as a habitat patch except that it is capable of movement and transient interactions with other habitat patches throughout the landscape. Furthermore, many populations of hosts consist of multiple levels of mobile habitat patches within the metapopulation. At all spatial scales, the mechanisms of speciation, dispersal, selection, and drift can mediate colonization of available habitat patches by microbes from the collective species pool in the intrinsic and extrinsic host environment (Costello *et al.* 2012). Thereby, host-associated microbiota form complex systems in which to study assembly processes across multiple levels of biological organization.

To address the dearth of studies on the assembly of host-associated microbiota in natural vertebrate populations, this research evaluates the role of dispersal limitation versus environmental selection as the mechanistic drivers of avian gut microbial assembly. Dispersal limitation predicts a negative association between geographic distance and community similarity whereas environmental selection predicts a positive association between environmental similarity and community similarity (Fig. 3.1). Few studies exist on the processes shaping avian microbiota and have mainly included precocial species (Vo, Chapter 1), but recent advancements in molecular methods and sequencing technology have enhanced the feasibility of studying avian microbiota using minimally invasive sampling (Caporaso *et al.* 2012; Vo & Jedlicka 2014). The goals of this research were to apply massively parallel sequencing towards characterization of spatiotemporal variation in the gut microbiota of an altricial avian species and to compare observed patterns against those expected based on ecological theory to determine the primary processes involved. Environmental selection was hypothesized to play a greater role structuring within-individual variation whereas both environmental selection and dispersal limitation were hypothesized to be relevant in mediating between-individual variation in microbiota (Fig. 3.1). Sampling at broad spatiotemporal scales was prioritized to capture ecological fluctuations, including within-individual, between-individual, and between-nest comparisons as well as comparisons across host age, season, and year. The results of this study inform whether mechanistic generalities underlie assembly of gut microbiota in free-living hosts and whether ecologically relevant levels of environmental heterogeneity significantly impact the assembly process.

Materials and methods

Study system

Bluebirds (genus *Sialia*) comprise an ideal study system in which to interrogate avian-associated microbiota. These birds occur in high abundance throughout North America and readily nest in manmade boxes, which enhance the feasibility of repeated sampling of the same individuals. This study utilized a system of over 100 nest boxes at the San Joaquin Experimental Range (Madera County, California, Fig. S3.1). The study site consists of annual grassland and foothill oak-pine woodlands ranging in elevation from 215 to 520 m. Nest boxes were constructed from non-treated redwood with interior dimensions of 14 x 10 cm, a height of 23 cm, and an entrance hole size of 3.8 cm for Western Bluebirds *Sialia mexicana*. Boxes were erected on Blue Oak *Quercus douglasii* and Interior Live Oak *Quercus wislizenii* trees, at least 1.5 m above ground, spaced 100 to 200 m apart, and generally facing northward or eastward.

In California, Western Bluebird breeding activity typically begins in March and extends through July. Clutch sizes range from three to eight eggs with a mean of five eggs. The female incubates and broods alone, with incubation periods lasting 12–14 days and brooding periods up to seven days post-hatching. Nestlings fledge within 18–21 days post-hatching. The entire nest box system at the San Joaquin Experimental Range was monitored weekly throughout the breeding season over three years (2011–2013) to support sampling for this study. Thermochron iButton data loggers (Maxim DS1921G) were installed on the interior backboard (2 inches from the ceiling) of every active nest box to record temperature once every hour during the nestling period.

To collect age-specific data for all nestlings, hatch dates were determined for each nest. In asynchronously hatched nests, all nestlings hatched on a given date were marked by superciliary down feather clipping, such that cohorts of nestlings from different hatch dates within the same nest could be distinguished (Zimmerling *et al.* 2004). New gloves and isopropanol-sterilized scissors were used for all nestling handling procedures to prevent cross-contamination of microbes between nests. By the fifth day post-hatching, nestlings were large enough to receive a single unique leg band, and by the tenth day post-hatching, each nestling was banded with a distinct color band combination, consisting of a uniquely numbered, metal band (USGS Bird Banding Laboratory) and three colored plastic bands (CA Fish and Game Permit SC-11869 and USGS Permit 21859).

Sample collection and metagenomic DNA extraction

Nestlings were sampled every five days, beginning on the fifth day post-hatching for nearly all nestlings. Fecal, cloacal swab, and oral swab samples were collected following Vo and Jedlicka (2014). Fresh fecal samples were stored dry in 2011 whereas they were stored in 1.5 mL of a lab-made ammonium-sulfate saturated nucleic acid stabilization and storage buffer (25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; Vo & Jedlicka 2014) in 2012 and 2013. If fecal sacs were larger than 0.22 grams, they were subsampled to 0.22 g for storage to maintain a minimum sample to buffer ratio of approximately 0.15 g/mL. Cloacal swab heads were stored in 1 mL of sucrose lysis buffer (20 mM EDTA, 50 mM Tris-HCl pH 9.0, 200 mM NaCl, 0.75 M sucrose; Mitchell & Takacs-Vesbach 2008) in 2011 whereas cloacal and oral swab heads were stored in DNA buffer in 2012 and 2013. All samples were

stored on ice in the field and at -20°C for up to three months before long term storage at either -20°C (2011 samples) or -80°C (2012 and 2013 samples). All protocols were approved by the UC Berkeley Animal Care and Use Committee (MAUP R317-0913).

Initially, metagenomic DNA from fecal (n = 18), cloacal swab (n = 18), and oral swab (n = 18) samples was extracted using the PowerSoil Kit (Mo Bio 12888-50). However, a more effective method based on solid phase reversible immobilization (SPRI) beads was used to extract metagenomic DNA from the remainder of the avian fecal and swab samples (n = 218) following Vo and Jedlicka (2014). Briefly, samples were processed with zirconia-silica beads (0.1 and 0.5 mm diameter) in lysis buffer (166 mM trisodium phosphate, 111 mM guanidine thiocyanate, 11 mM sodium chloride, 0.3% sodium dodecyl sulfate, 37 mM Tris-HCl pH 8.0, 1% beta-mercaptoethanol) on a high-powered tissue homogenizer, Precellys 24 (Precellys EQ03119.200.RD000.0) or PowerLyzer 24 (Mo Bio 13255). Sample lysates were incubated with ammonium acetate and aluminum ammonium sulfate to precipitate proteins and flocculate PCR inhibitors, respectively (Braid *et al.* 2003). SPRI bead solution, using Sera-Mag beads following Rohland and Reich (2012), were then employed to capture, wash, and elute sample DNA. All extracts were quantified on the Nanodrop spectrophotometer and Qubit fluorometer using the dsDNA HS (high sensitivity) Assay Kit (Invitrogen Q32854). Elutions were stored at -20°C until use.

16S rRNA amplicon generation, library preparation, and sequencing

Following Vo and Jedlicka (2014), metagenomic DNA extracts from each sample were amplified in triplicate using degenerate, barcoded primers targeting the bacterial 16S rRNA locus (515F and 806R; Caporaso *et al.* 2012) in 25 µL reactions containing 1X HotMasterMix (5 PRIME 2200400), 0.2 µM of each primer, 0.5 µg/µL BSA (NEB B9000S), 5% DMSO (ACS grade), and at least 0.1 ng of DNA. Thermocycling conditions followed an initial denaturation of 94°C for 2 minutes and 35 cycles of 94°C for 8 seconds, 50°C for 20 seconds, and 72°C for 30 seconds. PCR set-up was performed in an UV-irradiated clean hood, and thermocycling occurred in a separate, isolated room to minimize risk of cross-contamination. PCR reactions were electrophoresed on 1.5% agarose gels and stained with SYBR-Gold (Invitrogen S-11494) to detect amplification success.

For each sample, the triplicate PCR reactions were pooled and SPRI-cleaned using Sera-Mag beads following Rohland and Reich (2012), except that a 1.2 volume ratio of SPRI bead solution to sample was used to discard fragments smaller than 200 bp (e.g., primer dimers). The cleaned samples were subsequently quantified using the Qubit dsDNA HS Assay Kit and pooled equimolarly to form either nine libraries of six samples and 300-400 ng total amplicon DNA per library for a preliminary MiSeq run or 24 libraries of 18-21 samples and 81-507 ng of total amplicon DNA per library for the main MiSeq run. The NEBNext End Repair Module, NEBNext dA-Tailing Module, and Meyer and Kircher (2010) adapter ligation protocols were used with TruSeq Illumina indexed adapters to prepare the libraries using a “with-bead” and “PCR-free” approach (Vo & Jedlicka 2014). For SPRI cleanups between library reactions, Fisher *et al.*'s (2011) SPRI cleanup protocol was performed with modifications following Vo and Jedlicka (2014) to maximize DNA recovery.

The nine libraries for the initial MiSeq run and the 24 libraries for the main MiSeq run were quantified via qPCR using the KAPA Library Quantification Kit (KK4824), pooled

equimolarly, and analyzed on the Bioanalyzer (Agilent Technologies G2940CA). The resultant pooled libraries were submitted to the University of California Davis Genome Center for two 250 bp, paired-end runs on the Illumina MiSeq platform.

Data analysis

The Illumina reads were trimmed of adapters and demultiplexed into the original libraries using CASAVA 1.8. Reads that did not pass filter were removed, and Trimmomatic was used to trim the trailing edge of each sequence to a minimum quality score of Q20 while keeping paired-end reads in sync (Lohse *et al.* 2012). Barcodes were error-corrected following Bystrykh (2012), and those that could not be resolved (i.e., containing more than one error) were discarded. If one read was discarded, its paired read was also discarded. PANDAseq was used to merge paired-end reads, with a quality threshold of 0.6 (Masella *et al.* 2012). Merged reads that did not contain matching barcodes at both ends or that did not contain the correct primer sequences were also discarded.

In QIIME 1.8.0, each library was demultiplexed by barcode, and all barcode and primer sequences from the reads were trimmed before downstream analyses (Caporaso *et al.* 2010). With the `usearch_qf` (usearch quality filtering) protocol, the Greengenes v. 13_8 reference database and clustering at 97% identity were used to remove chimeras, pick 16S rRNA operational taxonomic units (OTUs), and assign taxonomies to all OTUs. A maximum-likelihood phylogenetic tree of all OTUs was generated using FastTree 2.1.3 as implemented within QIIME (Price *et al.* 2010).

For alpha diversity analyses, rarefaction was conducted using the metrics “PD_whole_tree” for phylogenetic distance (Faith & Baker 2007) and “shannon” for Shannon’s diversity index. The `nmle` package (Pinheiro *et al.* 2013) in R 3.01 (R Core Team 2013) was used to perform generalized linear mixed effects analysis of the relationship between alpha diversity (PD_whole_tree and shannon) and different spatial (latitude, longitude), temporal (age, sampling Julian day, year), and environmental (sample type as well as maximum, minimum, mean, standard deviation, and range of temperature from hatch date to sample date) factors as fixed effects with individual, host age, and host nest as nested random effects (using the random intercepts model). Visual inspection of residual plots did not reveal obvious deviations from homoscedasticity or normality. P-values for each effect were obtained using likelihood ratio tests of the full model versus the model without the effect in question. Models were evaluated using Akaike’s Information Criterion (AIC).

To standardize comparisons of microbiota between samples in beta diversity analyses, 16S rRNA OTU relative abundance tables were rarefied to 2,312 sequences per sample, the sequencing depth at which Shannon’s diversity index and phylogenetic distance plateaued for generally all samples (Fig. S3.2). The rarefied OTU tables were then used to generate summaries of taxonomic composition as well as UniFrac distance matrices (Lozupone & Knight 2005) to characterize similarity among samples. Unweighted and weighted UniFrac distance matrices were developed based on the exclusion and inclusion of OTU abundance when calculating distances between communities, respectively. Distance matrices were also created for continuous spatial (Vincenty formula for latitude and longitude data), temporal (age, sampling Julian day, year), and environmental (maximum, minimum, mean, median, standard deviation,

and range of temperatures between hatch date and sample date) factors within QIIME. The significance of each factor for explaining variation in unweighted UniFrac metrics among samples was determined using five different approaches: principal coordinate analysis (PCoA), parametric *t*-tests comparing all within-level and between-level distances of a given factor, nonparametric *t*-tests comparing all within-level and between-level distances of a given factor using 1000 Monte Carlo permutations of sample assignments for OTUs to calculate p-values, partial Mantel tests, and multiple matrix regression with randomization (Wang 2013). Bonferroni corrections were applied to all derived p-values. To measure the robustness of clusters in PCoA plots, repeated resampling of a subset of data from each sample was performed in jackknife analyses, and ellipsoids representing the interquartile ranges of UniFrac distances for each sample were depicted.

Results

Library summary

Amplicons from a total of 346 samples were sequenced across both MiSeq runs. The runs included 90 fecal, 136 cloacal, and 99 oral samples from 41 nestlings across 25 nests from 2011 through 2013. Additionally, one fecal, 10 cloacal, and 10 oral samples from 10 adult female Western Bluebirds across 10 different nests in 2013 were analyzed. The combined sequencing runs generated 25.8 million raw reads and retained 13.4 million reads after quality filtering. Across all samples, 7,692 unique OTUs were observed, which spanned 44, 119, 249, 456, and 856 different bacterial phyla, classes, orders, families, and genera, respectively.

Within-individual variation in gut microbiota

Spatial variation. Across all fecal samples, 3,371 unique OTUs were observed, with a mean of 173 (± 105 SD) OTUs per sample. Fecal samples were dominated in relative abundance by Firmicutes (42.5 ± 32.1 SD%) and Proteobacteria (19.7 ± 26.8 %). Cloacal samples collectively contained 4,804 unique OTUs, with a mean of 182 (± 96 SD) OTUs per sample, and were dominated by Proteobacteria (37.3 ± 27.7 %) and Firmicutes (24.5 ± 24.2 %). Furthermore, among all oral samples, 4,960 unique OTUs were observed, with a mean of 198 (± 123 SD) OTUs per sample. Oral samples were dominated by unassigned taxa (29.5 ± 30.1 %) and Proteobacteria (27.8 ± 23.7 %). PCoA of unweighted UniFrac distances revealed high similarity between oral and cloacal swab samples and substantial segregation of fecal samples from swab samples (Fig. 3.2). Both parametric and nonparametric *t*-tests identified a significant effect of gut site (i.e., cloacal, fecal, oral) for explaining variation in unweighted UniFrac distances among samples ($p < 0.001$ for both tests).

Temporal variation. PCoA plots of sample type by age revealed high differentiation between fecal and swab samples with unweighted UniFrac as well as distinct clusters of all sample types with weighted UniFrac early in the nestling period (4–7 days post-hatching; Fig. 3.3a). During this period, unassigned taxa and Proteobacteria dominated both cloacal and oral samples whereas Firmicutes were most abundant in fecal samples (Fig. S3.3a). By the middle of the nestling period (8–12 days post-hatching), fecal and swab samples continued to segregate distinctly by unweighted UniFrac, but all samples were notably more similar by weighted UniFrac (Fig. 3.3b). In fecal samples, Firmicutes remained most abundant whereas the abundance of unassigned taxa dropped substantially from 23% during the early nestling period

to only 9% by the middle of the nestling period (Fig. S3.3a). Cloacal samples became dominated by Proteobacteria while unassigned taxa remained most abundant in oral samples (Fig. S3.3b-c).

Near the end of the nestling period (13–19 days post-hatching), greater overlap between cloacal and fecal samples was observed by unweighted UniFrac, and all sample types overlapped to a high degree by weighted UniFrac (Fig. 3.3c). During this period, Proteobacteria, Firmicutes, and unassigned taxa dominated cloacal, fecal, and oral samples, respectively (Fig. S3.3). In contrast, greater differentiation was observed between cloacal and oral samples from adults by unweighted UniFrac (Fig. 3.3d). Interestingly, cloacal swabs from adults contained a higher abundance of Cyanobacteria and substantially less Firmicutes than cloacal samples from nearly-fledged nestlings (Fig. S3.3). Likewise, adult fecal samples contained notably higher abundance of Actinobacteria and no detectable Firmicutes relative to nestling fecal samples. Adult oral samples contained Proteobacteria in highest abundance, and unassigned bacteria declined from being most abundant at 29% during the late nestling period to only 15% in adults. Overall, parametric and nonparametric *t*-tests revealed a significant effect of host age for unweighted UniFrac distances among samples ($p = 0.011$ and 0.015 , respectively). However, in contrast to analyses of beta diversity, comparisons of alpha diversity revealed no significant relationships between host age and number of OTUs or phylogenetic diversity for all sample types (Fig. 3.4).

Between-individual variation in gut microbiota

Spatial variation. The relationship between environmental similarity and similarity of microbiota was evaluated by nest site and by nest box temperature. PCoA of unweighted UniFrac of all samples indicated little to no segregation by nest site (Fig. S3.4a). When reanalyzed by sample type, oral samples exhibited minimal grouping by nest whereas no clustering by nest was observed among fecal and cloacal samples (Fig. S3.4b-d). Nevertheless, nest was a significant factor in explaining variation in unweighted UniFrac distances by parametric ($p < 0.001$) and nonparametric ($p = 0.001$) *t*-tests. For maximum nest box temperature, PCoA of unweighted UniFrac revealed no pattern among all samples combined but suggested patterns upon analysis by sample type (Fig. S3.5). Although Partial Mantel tests were non-significant (Table 3.1), multiple matrix regression with randomization (MMRR) identified distances in maximum temperature between hatch date and sampling date to be a significantly associated ($p = 0.043$) with UniFrac distances among samples.

The relationship between geographical distance and similarity of microbiota was assessed using latitude and longitude data for all collected samples. PCoA of unweighted UniFrac of samples altogether revealed no pattern by longitude (Fig. S3.6a). After reanalysis by sample type, increased dissimilarity among samples was somewhat observed with increased longitudinal distance, particularly for oral samples (Fig. S3.6b-d). This result was recapitulated in parametric ($p < 0.001$) and nonparametric ($p = 0.001$) *t*-tests. However, neither partial Mantel tests nor MMRR analysis identified latitude and longitude distances as significant correlates of unweighted UniFrac distances.

Temporal variation. Sampling dates were converted to Julian days to analyze seasonal associations with similarity in microbiota. PCoA of unweighted UniFrac of all samples suggested a pattern of increased dissimilarity among samples with increased distance in sampling date

within a season (Fig. 3.5a). This pattern became more apparent when samples were analyzed by sample type, particularly for fecal and cloacal samples (Fig. 3.5b-d). Furthermore, parametric ($p < 0.001$) and nonparametric ($p = 0.001$) t -tests were consistent with observed PCoA patterns, but neither partial Mantels nor MMRR revealed a significant association between distances in sampling date and unweighted UniFrac distances.

In analyzing interannual trends, PCoA of unweighted UniFrac of all samples suggested greater differentiation of 2013 samples from 2011 and 2012 samples (Fig. 3.6a). When analyzed by sample type, this pattern was more robust (Fig. 3.6b-d). Among fecal samples, all three years appeared to differentially segregate, with greater separation of the 2013 samples from the 2011 and 2012 samples. For cloacal samples, 2013 samples clearly segregated from 2011 and 2012 samples, and among oral samples, which were only collected in 2012 and 2013, distinct clustering by year was observed. The effect of year on unweighted UniFrac distances was significant in parametric ($p < 0.001$) and nonparametric ($p = 0.001$) t -tests. Furthermore, partial Mantel tests revealed distances in sample year to significantly correlate with UniFrac distances among samples after controlling for host age, sampling date within season, latitude and longitude, maximum temperature between hatch date and sampling date as well as minimum temperature between hatch date and sampling date (Table 3.1). Likewise, MMRR analysis indicated distances in sampling year as a significant predictor in UniFrac distances ($p = 0.017$). Analysis of taxonomic composition of each sample type by year revealed representation of Chlamydiae in fecal and cloacal samples only during 2012 as well as notably higher abundance of Actinobacteria and Cyanobacteria in all samples during 2013 relative to 2011 and 2012 (Fig. S3.7).

Determinants of alpha diversity among all samples

Generalized linear mixed models revealed sample type and standard deviation of temperatures (between hatch date and sampling date) as significant predictors of phylogenetic diversity of the microbiota among samples ($p = 0.021$ and $\beta = -1.70$, $p = 0.012$, respectively). However, the model containing year ($p = 0.309$) in addition to sample type and standard deviation of temperatures had the lowest AIC. For variation in Shannon's diversity index among samples, the model containing sample type and standard deviation in temperature had the lowest AIC, and both were significant predictors ($p = 0.006$ and $\beta = -0.506$, $p = 0.036$, respectively).

Discussion

To investigate the mechanisms underlying assembly of the vertebrate gut microbiota, this study applied massively parallel sequencing to characterize samples collected across broad spatiotemporal scales in a natural avian population. With regard to within-individual variation in microbiota, gut site was significantly associated with microbial diversity, whereby cloacal and oral microbiota were more similar to each other than either were to fecal microbiota. Although alpha diversity metrics of the microbiota did not correlate with nestling age, significant associations between beta diversity metrics and host age were observed through the nestling developmental period. Variation in microbiota between individuals was correlated with maximum temperature, standard deviation of temperature, and sampling year but not sampling date or geographic distance. All observed patterns in the gut microbiota were consistent with environmental selection as the main mediator of microbial assembly. No

evidence was obtained for the effect of dispersal limitation, either at the within-individual or the between-individual levels.

Prior studies on mammalian and avian gut microbiota

Before massively parallel sequencing capabilities, extensive efforts with sequencing clone libraries generally yielded a maximum of just over 100 unique OTUs in microbial samples from ostriches, turkeys, gulls, and penguins (Scupham *et al.* 2007; Lu *et al.* 2008; Banks *et al.* 2009; Matsui *et al.* 2010). Pyrosequencing (i.e., 454 sequencing) enabled recovery of six-fold greater microbial diversity in hoatzins and chickens (Godoy-Vitorino *et al.* 2008; Singh *et al.* 2012), and in this study, Illumina sequencing of 16S rRNA amplicons revealed staggering microbial diversity, with over 7,500 unique OTUs observed across all samples in Western Bluebirds. The increased sensitivity afforded by high-throughput sequencing has empowered cost-effective, higher resolution analyses of host-microbial associations and dynamics.

Comparisons of gut microbiota across 60 mammalian species have shown two consistently dominant bacterial phyla, Firmicutes and Bacteroidetes (Ley *et al.* 2008). In contrast, studies in birds have revealed wide variation in the dominant phyla, from Firmicutes and Bacteroidetes in ostriches (Matsui *et al.* 2010), turkeys (Scupham *et al.* 2007), and hoatzins (Godoy-Vitorino *et al.* 2008) to Firmicutes and Proteobacteria in gulls (Lu *et al.* 2008), chickens (Zhu *et al.* 2002; Singh *et al.* 2012), and Canada geese (Lu *et al.* 2009) as well as Firmicutes and Actinobacteria in Adelie penguins (Banks *et al.* 2009) and migratory shorebirds (Santos *et al.* 2012). This study revealed a very low abundance of Bacteroidetes and rather a predominance of Firmicutes, Proteobacteria, and unassigned taxa across all Western Bluebird samples (Fig. S3.3).

Environmental selection through diet has been shown to explain interspecific host variation in mammalian gut microbiota (Ley *et al.* 2008). Bacteroidetes abundance in the human and mouse gut is associated with high carbohydrate, low fat diets and a lean host phenotype whereas abundance of Firmicutes correlates with low carbohydrate, high fat diets and obesity (Turnbaugh *et al.* 2009a; Ridaura *et al.* 2013). Additionally, comparative genomic analyses of human gut-associated Bacteroidetes and Firmicutes have revealed significantly more genes encoding glycoside hydrolases and polysaccharide lyases in Bacteroidetes versus more ATP-binding cassette transporters and phosphotransferase systems in Firmicutes (Mahowald *et al.* 2009). Accordingly, given the specialized enzymes of gut-associated Bacteroidetes for utilizing carbohydrate energy sources, the relative absence of Bacteroidetes in the Western Bluebirds may stem from the absence of carbohydrates in their diet, which typically includes only insects during the breeding season to support the energetics of reproduction and nestling growth (Mock 1991).

Intra-individual variation in the Western Bluebird gut microbiota

Within-individual variation in microbiota was consistent with environmental selection as the main driver of community assembly whereas no relationship was observed between geographic distance and microbial similarity between gut sites. Specifically, fecal and cloacal samples were more proximately located than cloacal and oral samples. However, cloacal and oral samples exhibited higher similarity in microbial diversity. The continuous passage of digesta through the gut lumen is expected to support transient microbiota in the fecal matter, subject to turnover

based on host dietary choices. In contrast to the gut lumen, environmental similarity between the oral and cloacal mucosae likely contributed to their similar microbiota. The epithelial surfaces of the gut are covered by a layer of mucus gel that is comprised mainly of water and glycoproteins (Florey 1955). This mucus deters most microorganisms from contacting and colonizing the underlying surface. Nevertheless, some bacteria are capable of adhering to the mucus layer as well as utilizing its glycoproteins as energy sources (Cohen *et al.* 1983; McSweeney & Walker 1986; McCormick *et al.* 1988; Macfarlane & Gibson 1991; Kirjavainen *et al.* 1998; Macfarlane *et al.* 2001). As swabbing effectively samples the mucosa, the microbiota in both oral and cloacal swabs would be expected to be dominated by bacteria that specialize in invading and persisting in gastrointestinal mucus. Indeed, Proteobacteria are known to produce adhesins specific to gastrointestinal mucins and other glycoproteins (Boedeker 1984; Gerlach & Hensel 2007), and this bacterial phylum was most abundant in both the oral and cloacal samples from Western Bluebirds.

Further differentiation between oral and cloacal microbiota likely arose from the physicochemical differences among the mucosal sites. The avian cloaca receives digesta as well as nitrogenous waste, rendering its pH to be significantly lower than that of the oral cavity (Evans & Heiser 2004). Additionally, given its distal location in the gastrointestinal tract, the cloaca is substantially more anoxic. These environmental differences are expected to support a higher abundance of anaerobes as well as microbes that can tolerate low pH in the cloaca. Most members of the Proteobacteria phylum are either facultatively or obligately anaerobic. Interestingly, although Proteobacteria were abundant in both oral and cloacal microbiota, the oral mucosa was dominated by unassigned taxa whereas the cloacal mucosa contained predominantly Proteobacteria. Unlike novel taxa in the human microbiome (Wylie *et al.* 2012), unassigned taxa in the avian samples were often found in high abundance across samples and were unidentifiable even to the phylum level, thereby continuing the pattern of microbial discovery as increasingly more diverse samples are analyzed with massively parallel sequencing.

Intriguingly, the similarity of microbiota across body sites appeared to increase with host age in this study (Fig. 3.3). The microbiota were most dissimilar during the early nestling period and were most similar by pre-fledging ages, especially by weighted UniFrac. This observation is unexpectedly inconsistent with prior studies in humans, mice, and birds. In humans, subsets of the species pool of environmental microbes colonize different body sites, and over time, the microbiota among sites become highly differentiated (Costello *et al.* 2009). Specifically in the intestines, facultative anaerobes establish first and reduce the environment such that obligate anaerobes subsequently colonize the site (Favier *et al.* 2002; Fanaro *et al.* 2003). Likewise, in chickens, ceca contain a subset of ileal microbiota until 14 days post-hatching when both develop into significantly different microbiota (Lu *et al.* 2003). A factor that may contribute to the disparity in results between this and prior studies involves the development of thermoregulation in precocial species, including chickens, versus altricial species, including Western Bluebirds. The body temperatures of altricial nestlings can be subject to ambient temperature for nearly a week post-hatching before complete homeothermy (Whittow & Tazawa 1991). During this time, a greater temperature gradient likely exists within the nestling gastrointestinal tract such that more distinct microbiota might assemble at different sites. At the onset of thermoregulation, homogenization of the entire

body to 41°C would then comprise selective pressure for bacteria that proliferate successfully within the narrow temperature regime, thus contributing to more similar microbiota across all body sites over the course of the nestling period. As this study is the first to compare microbiota from multiple body sites within the same individual in an altricial avian species, more research in other altricial birds is necessary to evaluate the prevalence of this observed temporal pattern over the nestling period.

Also surprisingly, the absence of a relationship between OTU richness or phylogenetic diversity and host age in this study is inconsistent with findings of previous microbial studies in developing birds (Fig. 3.4). In chickens, increasing microbial complexity is observed in cecal and ileal microbiota between younger (3–7 days post-hatching), middle-aged (14–21 days post-hatching), and older (49 days post-hatching) chicks (Lu *et al.* 2003). Likewise in Tree Swallows, cloacal colonization is characterized by higher species richness and abundance from two to 19 days post-hatching (Mills *et al.* 1999), and in Black-legged Kittiwakes, an increase in microbial diversity occurs with chick age, wherein chick microbiota actually exhibit higher diversity than adult microbiota (Dongen *et al.* 2013). In this study, the establishment of a diverse microbiota was typically achieved by five days post-hatching, with high variance among samples at any given age. Although a clear relationship in microbial diversity with age was not detected, shifts in the community structure occurred over the nestling period and presumably continued to occur after fledging, given the dissimilarity of pre-fledging versus adult microbiota (Fig. S3.3).

Inter-individual variation in the Western Bluebird gut microbiota

Between-individual comparisons of gut microbiota revealed limited evidence for an effect of either geographic distance (i.e., dispersal limitation) or environmental selection by nest site. This latter observation was very surprising as past studies have repeatedly shown that shared environments promote higher similarity in microbiota, such that samples from different host individuals in the same environment harbor more similar microbiota than those of host individuals living in different environments. For example, among humans, couples or family members of the same household exhibit more similar skin, oral, and fecal microbiota than members of different households (Song *et al.* 2013). Likewise, cohabitation of mice give rise to more similar microbiota (Campbell *et al.* 2012), and birds from the same nest share higher similarity of cloacal microbiota in Tree Swallows (Lombardo *et al.* 1996), Spotted Towhees (Klomp *et al.* 2008), as well as cross-fostered Great Tits and Blue tits (Lucas & Heeb 2005). In considering the weak effect of nest site in the Western Bluebird data, general nesting material, nest site habitat, and prey items may have been so homogenous among nests that environmental selection by these factors across all nests was highly comparable. Furthermore, rates of extra-pair paternity among Western Bluebirds are known to be as high as 19% among all nestlings and 45% among all nests (Dickinson & Akre 1998), which would add another dimension of environmental homogenization, in this case the genetic background of host individuals in the nests.

Despite the weak signature of nest site on the observed microbial patterns, there was support for environmental selection by temperature. More specifically, standard deviation of temperatures as well as maximum temperature between hatch date and sampling date were significant predictors of variation of microbiota among samples in the generalized linear models and MMRR analyses, respectively. Some bacteria invariably tolerate wider temperature

fluctuations better than others (Ratkowsky *et al.* 1982; Small *et al.* 1986), and given the dependence of nestlings on ambient temperature for thermoregulation during the first week post-hatching, it may be expected that temperature would comprise a stronger source of environmental selective pressure on the microbiota than nest site alone. The implications of temperature dependence for microbial assembly in altricial versus precocial species comprise compelling directions for future research, and experimental work is necessary to clarify to what extent temperature-mediated shifts in the microbiota affect host metabolic or other physiological functions.

Of particular interest among inter-individual comparisons, multiple lines of evidence—including principal coordinate analysis, parametric and nonparametric *t*-tests, partial Mantel tests, multiple matrix regression with randomization, and generalized linear mixed models—pointed to the significance of sampling year for structuring variation in microbiota among samples. Although sample storage conditions in 2011 differed somewhat from those for samples collected in 2012 and 2013, the largest effect of year was observed between 2012 and 2013, when implemented sample storage conditions were identical. A number of reasons related to environmental selection likely explain the outlier nature of the 2013 sampling year. Overall, precipitation declined drastically from 2011 to 2013. In 2011 and 2012, total precipitation recorded was 17.49 and 18.96 inches, respectively. However, 2013 was a particularly bad drought year during which precipitation totaled only 2.12 inches. Correspondingly, peak standing crop was calculated as 1918 and 2110 pounds per acre for 2010–11 and 2011–12, respectively. In contrast, only 1463 pounds per acre were produced in 2012–13 (San Joaquin Experimental Range unpubl. data).

These precipitation and primary productivity fluctuations appeared to drive bottom-up trophic cascades. Grasshoppers were the most abundant arthropods present at the San Joaquin Experimental Range in 2011, and they continued to be common albeit less abundant in 2012. However, not a single grasshopper was observed at the study site in 2013, and overall arthropod abundance in 2013 was notably low. Predation-independent mortality of nestlings in the nest boxes was nearly non-existent in 2011 and 2012 but remarkably prevalent in 2013. Furthermore, although the typical Western Bluebird breeding season extends from March through early August at the study site, all observable breeding activity ceased in June of 2013, 1.5 months earlier than expected. As aforementioned, nestlings require a high fat and protein diet to support the metabolic requirements of growth and development. However, observations of mistletoe berries in the nest material as well as nestling fecal matter (Fig. 3.7) indicated that the nestlings were being fed non-insect diet items during their energetically demanding nestling growth period. To the best of current knowledge, feeding of berries to nestlings in Western Bluebirds has not been documented. The early curtailment of the 2013 breeding season therefore seemed to be due to a ubiquitous lack of arthropod prey resources at the site to support further reproductive efforts. The shift in diet to berries may have contributed to the higher abundance of Cyanobacteria and Actinobacteria in all 2013 samples (Fig. S3.7). Altogether, the variation in microbiota between years appeared to reflect heterogeneity in resource availability. Variation between seasons and between years in avian microbiota has been observed previously in Alpine Accentors *Prunella collaris* (Janiga *et al.* 2007). However, given the limited breadth of the annual data obtained in this study as well as

prior studies, further research to characterize interannual fluctuations in microbiota and environmental change would comprise informative future research.

Conclusion

Variation in the microbiota of Western Bluebirds was consistent with environmental selection as the main driver of microbial assembly across body sites and host age within individuals as well as between individuals across space and time. The significance of ambient temperature and interannual fluctuations in structuring variation in the microbiota are particularly relevant to concerns over the ecological implications of global change, especially for altricial host species in which a lack of homeothermy during early life render them more susceptible to environmental conditions. Over the next century, higher temperatures, more intense precipitation, greater drought regions, higher interannual variability, and longer growing seasons are projected for North America (Duffy *et al.* 2014). Climatic shifts in temperature and precipitation are known to significantly impact organismal phenology, species ranges, species interactions, and ecosystem dynamics (Walther *et al.* 2002). The results of this study suggest that significant seasonality and shifts in environmental abiotic factors can incur concomitant impacts on host-microbial interactions. Given the robust microbial signatures recovered during the drought year of this study, additional research is needed to connect host-associated microbial form and function. Clarification of the functional implications of the apparent environmentally-driven microbial shifts and their potential short- or long-term effects on host fitness are of utmost interest.

References

- Banks JC, Cary SC, Hogg ID (2009) The phylogeography of Adelie penguin faecal flora. *Environmental Microbiology*, **11**, 577–588.
- Benson AK, Kelly SA, Legge R *et al.* (2010) Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 18933–18938.
- Boedeker EC (1984) Attachment of organisms to the gut mucosa. Volume I and II. CRC Press, Boca Raton.
- Braid MD, Daniels LM, Kitts CL (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. *Journal of Microbiological Methods*, **52**, 389–393.
- Bystrykh LV (2012) Generalized DNA Barcode Design Based on Hamming Codes. *PLoS ONE*, **7**, e36852.
- Campbell JH, Foster CM, Vishnivetskaya T *et al.* (2012) Host genetic and environmental effects on mouse intestinal microbiota. *The ISME Journal*, **6**, 2033–2044.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, **7**, 335–336.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.
- Carey HV, Walters WA, Knight R (2013) Seasonal restructuring of the ground squirrel gut microbiota over the annual hibernation cycle. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, **304**, R33–R42.

- Clemente JC, Ursell LK, Parfrey LW, Knight R (2012) The impact of the gut microbiota on human health: an integrative view. *Cell*, **148**, 1258–1270.
- Cohen PS, Rossoll R, Cabelli VJ, Yang SL, Laux DC (1983) Relationship between the mouse colonizing ability of a human fecal *Escherichia coli* strain and its ability to bind a specific mouse colonic mucous gel protein. *Infection and Immunity*, **40**, 62–69.
- Costello EK, Lauber CL, Hamady M *et al.* (2009) Bacterial community variation in human body habitats across space and time. *Science*, **326**, 1694–1697.
- Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA (2012) The application of ecological theory toward an understanding of the human microbiome. *Science*, **336**, 1255–1262.
- Dickinson JL, Akre JJ (1998) Extrapair paternity, inclusive fitness, and within-group benefits of helping in western bluebirds. *Molecular Ecology*, **7**, 95–105.
- Dongen WF van, White J, Brandl HB *et al.* (2013) Age-related differences in the cloacal microbiota of a wild bird species. *BMC Ecology*, **13**, 11.
- Duffy PB, Maloney E, Sheffield J (2014) Global climate model simulations of North America. In: *Climate Change in North America Regional Climate Studies*. (ed Ohring G), pp. 167–200. Springer International Publishing.
- Evans HE, Heiser JB (2004) What's inside: Anatomy and physiology. In: *Handbook of bird biology*. Cornell Lab of Ornithology, Ithaca.
- Faith DP, Baker AM (2007) Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. *Evolutionary Bioinformatics Online*, **2**, 121–128.
- Fanaro S, Chierici R, Guerrini P, Vigi V (2003) Intestinal microflora in early infancy: composition and development. *Acta Pædiatrica*, **92**, 48–55.
- Favier CF, Vaughan EE, Vos WMD, Akkermans ADL (2002) Molecular monitoring of succession of bacterial communities in human neonates. *Applied and Environmental Microbiology*, **68**, 219–226.
- Fisher S, Barry A, Abreu J *et al.* (2011) A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biology*, **12**, R1.
- Florey H (1955) The croonian lecture: mucin and the protection of the body. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, **143**, 147–158.
- Gerlach RG, Hensel M (2007) Protein secretion systems and adhesins: The molecular armory of Gram-negative pathogens. *International Journal of Medical Microbiology*, **297**, 401–415.
- Godoy-Vitorino F, Ley RE, Gao Z *et al.* (2008) Bacterial community in the crop of the hoatzin, a neotropical folivorous flying bird. *Applied and Environmental Microbiology*, **74**, 5905–5912.
- Janiga M, Sedlářová A, Rigg R, Novotná M (2007) Patterns of prevalence among bacterial communities of alpine accentors (*Prunella collaris*) in the Tatra Mountains. *Journal of Ornithology*, **148**, 135–143.
- Kirjavainen PV, Ouwehand AC, Isolauri E, Salminen SJ (1998) The ability of probiotic bacteria to bind to human intestinal mucus. *FEMS Microbiology Letters*, **167**, 185–189.
- Klomp JE, Murphy MT, Smith SB *et al.* (2008) Cloacal microbial communities of female spotted towhees *Pipilo maculatus*: microgeographic variation and individual sources of variability. *Journal of Avian Biology*, **39**, 530–538.

- Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of mammals and their gut microbes. *Science*, **320**, 1647–1651.
- Lohse M, Bolger AM, Nagel A *et al.* (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, **40**, W622–W627.
- Lombardo MP, Thorpe PA, Cichewicz R *et al.* (1996) Communities of cloacal bacteria in tree swallow families. *Condor*, **98**, 167–172.
- Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228–8235.
- Lu JR, Domingo JWS, Hill S, Edge TA (2009) Microbial diversity and host-specific sequences of Canada goose feces. *Applied and Environmental Microbiology*, **75**, 5919–5926.
- Lu J, Domingo JWS, Lamendella R, Edge T, Hill S (2008) Phylogenetic diversity and molecular detection of bacteria in gull feces. *Applied and Environmental Microbiology*, **74**, 3969–3976.
- Lu J, Idris U, Harmon B *et al.* (2003) Diversity and succession of the intestinal bacterial community of the maturing broiler chicken. *Applied and Environmental Microbiology*, **69**, 6816.
- Lucas FS, Heeb P (2005) Environmental factors shape cloacal bacterial assemblages in great tit *Parus major* and blue tit *P. caeruleus* nestlings. *Journal of Avian Biology*, **36**, 510–516.
- Macfarlane GT, Gibson GR (1991) Formation of glycoprotein degrading enzymes by *Bacteroides fragilis*. *FEMS Microbiology Letters*, **61**, 289–293.
- Macfarlane S, Hopkins MJ, Macfarlane GT (2001) Toxin synthesis and mucin breakdown are related to swarming phenomenon in *Clostridium septicum*. *Infection and Immunity*, **69**, 1120–1126.
- Mahowald MA, Rey FE, Seedorf H *et al.* (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 5859–5864.
- Martiny JBH, Bohannan BJM, Brown JH *et al.* (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102–112.
- Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for Illumina sequences. *BMC Bioinformatics*, **13**, 31.
- Matsui H, Kato Y, Chikaraishi T *et al.* (2010) Microbial diversity in ostrich ceca as revealed by 16S ribosomal RNA gene clone library and detection of novel *Fibrobacter* species. *Anaerobe*, **16**, 83–93.
- McCormick BA, Stocker BA, Laux DC, Cohen PS (1988) Roles of motility, chemotaxis, and penetration through and growth in intestinal mucus in the ability of an avirulent strain of *Salmonella typhimurium* to colonize the large intestine of streptomycin-treated mice. *Infection and Immunity*, **56**, 2209–2217.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3229–3236.
- McSweegan E, Walker RI (1986) Identification and characterization of two *Campylobacter jejuni* adhesins for cellular and mucous substrates. *Infection and Immunity*, **53**, 141–148.
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, **2010**, pdb.prot5448.

- Mihaljevic JR (2012) Linking metacommunity theory and symbiont evolutionary ecology. *Trends in Ecology & Evolution*, **27**, 323–329.
- Mills TK, Lombardo MP, Thorpe PA (1999) Microbial colonization of the cloacae of nestling tree swallows. *The Auk*, **116**, 947–956.
- Mitchell KR, Takacs-Vesbach CD (2008) A comparison of methods for total community DNA preservation and extraction from various thermal environments. *Journal of Industrial Microbiology & Biotechnology*, **35**, 1139–1147.
- Mock PJ (1991) Daily allocation of time and energy of Western Bluebirds feeding nestlings. *The Condor*, **93**, 598–611.
- Pinheiro J, Bates D, DebRoy S, Sarker D, the R Development Core Team (2013) nlme: Linear and Nonlinear Mixed Effects Models.
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, **5**, e9490.
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ratkowsky DA, Olley J, McMeekin TA, Ball A (1982) Relationship between temperature and growth rate of bacterial cultures. *Journal of Bacteriology*, **149**, 1–5.
- Ridaura VK, Faith JJ, Rey FE *et al.* (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, **341**, 1241214.
- Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, **22**, 939–946.
- Santos SS, Pardal S, Proença DN *et al.* (2012) Diversity of cloacal microbial community in migratory shorebirds that use the Tagus estuary as stopover habitat and their potential to harbor and disperse pathogenic microorganisms. *FEMS Microbiology Ecology*, **82**, 63–74.
- Scupham A j., Jones J a., Wesley I v. (2007) Comparison of DNA extraction methods for analysis of turkey cecal microbiota. *Journal of Applied Microbiology*, **102**, 401–409.
- Singh KM, Shah T, Deshpande S *et al.* (2012) High through put 16S rRNA gene-based pyrosequencing analysis of the fecal microbiota of high FCR and low FCR broiler growers. *Molecular Biology Reports*, **39**, 10595–10602.
- Small PM, Täuber MG, Hackbarth CJ, Sande MA (1986) Influence of body temperature on bacterial growth rates in experimental pneumococcal meningitis in rabbits. *Infection and Immunity*, **52**, 484–487.
- Song SJ, Lauber C, Costello EK *et al.* (2013) Cohabiting family members share microbiota with one another and with their dogs. *eLife*, **2**, 1–2.
- Spor A, Koren O, Ley R (2011) Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology*, **9**, 279–290.
- Turnbaugh PJ, Hamady M, Yatsunenko T *et al.* (2009a) A core gut microbiome in obese and lean twins. *Nature*, **457**, 480–484.
- Turnbaugh PJ, Ridaura VK, Faith JJ *et al.* (2009b) The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine*, **1**, 6ra14–6ra14.
- Vo AE, Jedlicka JA (2014) Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for fecal and swab samples. *Molecular Ecology Resources*, In press.

- Walther G-R, Post E, Convey P *et al.* (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.
- Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, **67**, 3403–3411.
- Whittow GC, Tazawa H (1991) The early development of thermoregulation in birds. *Physiological Zoology*, **64**, 1371–1390.
- Wylie KM, Truty RM, Sharpton TJ *et al.* (2012) Novel bacterial taxa in the human microbiome. *PLoS ONE*, **7**, e35294.
- Zhu XY, Zhong T, Pandya Y, Joerger RD (2002) 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. *Applied and Environmental Microbiology*, **68**, 124–137.
- Zimmerling JR, Craigie GE, Robinson AE (2004) A comparison of techniques for marking passerine nestlings. *The Wilson Bulletin*, **116**, 240–245.

Tables

Table 3.1. Partial mantel correlation coefficient (Pearson's) and p-values from causal models of variation in unweighted UniFrac among all fecal, cloacal, and oral samples collected in this study. Asterisk indicates statistically significant p-value ($\alpha = 0.05$).

Dependent variable	Independent variable	Controlled covariable	Pearson's Correlation	p-value
unweighted UniFrac	Age	Jday	-0.011	0.701
unweighted UniFrac	Age	Year	-0.008	0.658
unweighted UniFrac	Age	LatLong	-0.010	0.708
unweighted UniFrac	Age	MaxTemp	-0.005	0.561
unweighted UniFrac	Age	MinTemp	-0.010	0.710
unweighted UniFrac	Jday	Age	0.030	0.187
unweighted UniFrac	Jday	Year	-0.015	0.628
unweighted UniFrac	Jday	LatLong	0.030	0.185
unweighted UniFrac	Jday	MaxTemp	0.048	0.118
unweighted UniFrac	Jday	MinTemp	0.012	0.319
unweighted UniFrac	Year	Age	0.079	0.006*
unweighted UniFrac	Year	Jday	0.075	0.004*
unweighted UniFrac	Year	LatLong	0.080	0.005*
unweighted UniFrac	Year	MaxTemp	0.084	0.002*
unweighted UniFrac	Year	MinTemp	0.070	0.001*
unweighted UniFrac	LatLong	Age	-0.004	0.540
unweighted UniFrac	LatLong	Jday	-0.006	0.586
unweighted UniFrac	LatLong	Year	-0.010	0.625
unweighted UniFrac	LatLong	MaxTemp	0.003	0.453
unweighted UniFrac	LatLong	MinTemp	-0.004	0.571
unweighted UniFrac	MaxTemp	Age	-0.033	0.866
unweighted UniFrac	MaxTemp	Jday	-0.051	0.965
unweighted UniFrac	MaxTemp	Year	-0.043	0.935
unweighted UniFrac	MaxTemp	LatLong	-0.034	0.896
unweighted UniFrac	MaxTemp	MinTemp	-0.039	0.925
unweighted UniFrac	MinTemp	Age	0.044	0.085
unweighted UniFrac	MinTemp	Jday	0.035	0.067
unweighted UniFrac	MinTemp	Year	0.024	0.194
unweighted UniFrac	MinTemp	LatLong	0.044	0.088
unweighted UniFrac	MinTemp	MaxTemp	0.048	0.053

Figures

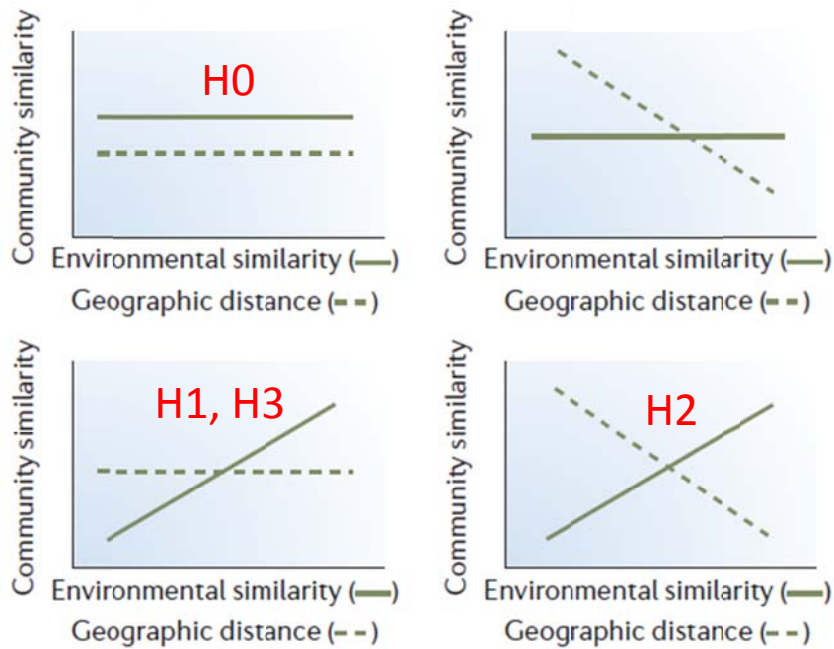


Fig. 3.1. Predictions expected from all hypotheses addressed in this study. H0 represents the null hypothesis. H1 pertains to within-individual variation in microbiota (i.e., across gut sites) whereas H2 concerns between-individual spatial variation in microbiota, and H3 considers between-individual temporal variation in microbiota (i.e., individuals within- and between-nests). Figure modified from (Martiny *et al.* 2006).

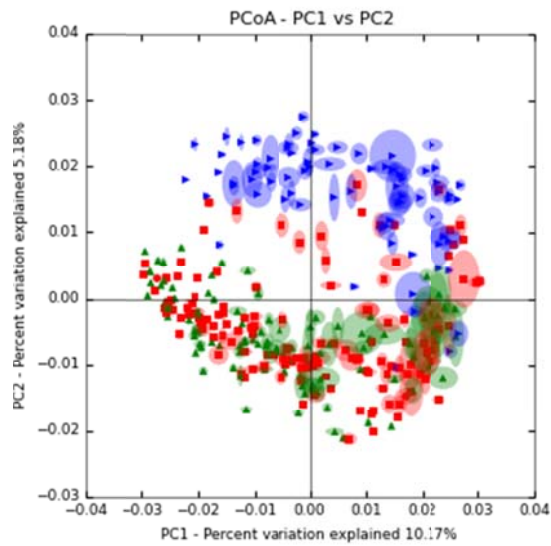


Fig. 3.2. Principal coordinate analysis of unweighted UniFrac for all samples combined (n = 346). Fecal, cloacal, and oral samples are in blue, red, and green, respectively. Ellipsoids indicate interquartile ranges for each sample from jackknifed analysis of beta diversity. High similarity between cloacal and oral microbiota as well as distinct segregation of fecal microbiota depicted.

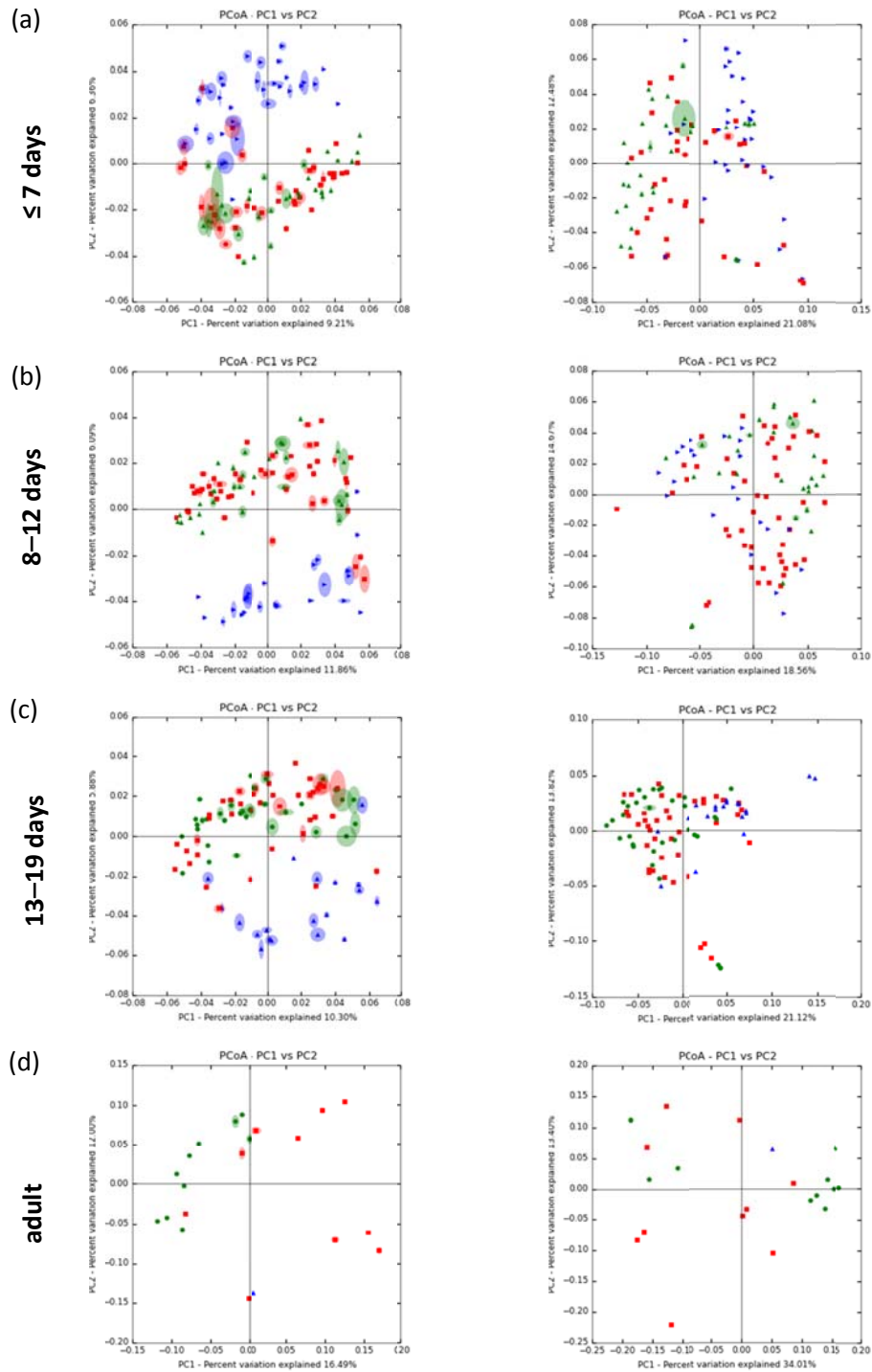


Fig. 3.3. Principal coordinate analysis of UniFrac, unweighted (left) and weighted (right) by OTU abundance, for samples at ≤ 7 days post-hatching (a, $n = 113$), 8-12 days post-hatching (b, $n = 110$), 13-19 days post-hatching (c, $n = 102$), and adulthood (d, $n = 21$). Fecal, cloacal, and oral samples are in blue, red, and green, respectively. Ellipsoids indicate interquartile ranges for each sample from jackknifed analysis of beta diversity. Increase in similarity of microbiota across gut sites with nestling age depicted.

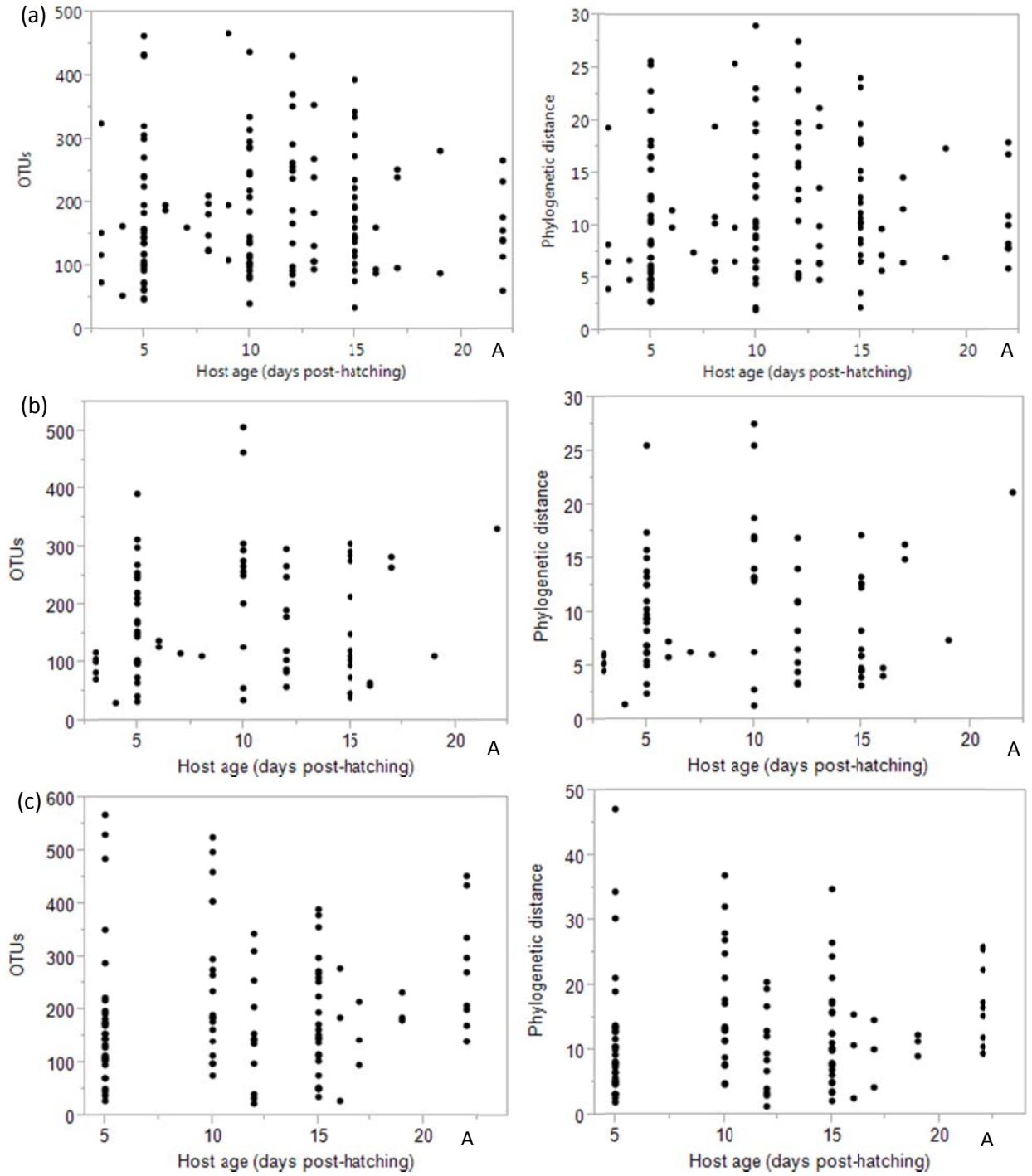


Fig. 3.4. Relationship between host age and number of OTUs (left) or phylogenetic distance (right) in Western Bluebird cloacal swab (a), fecal (b), and oral (c) swab samples. A denotes adult under host age on the x-axis of each plot. No significant associations recovered.

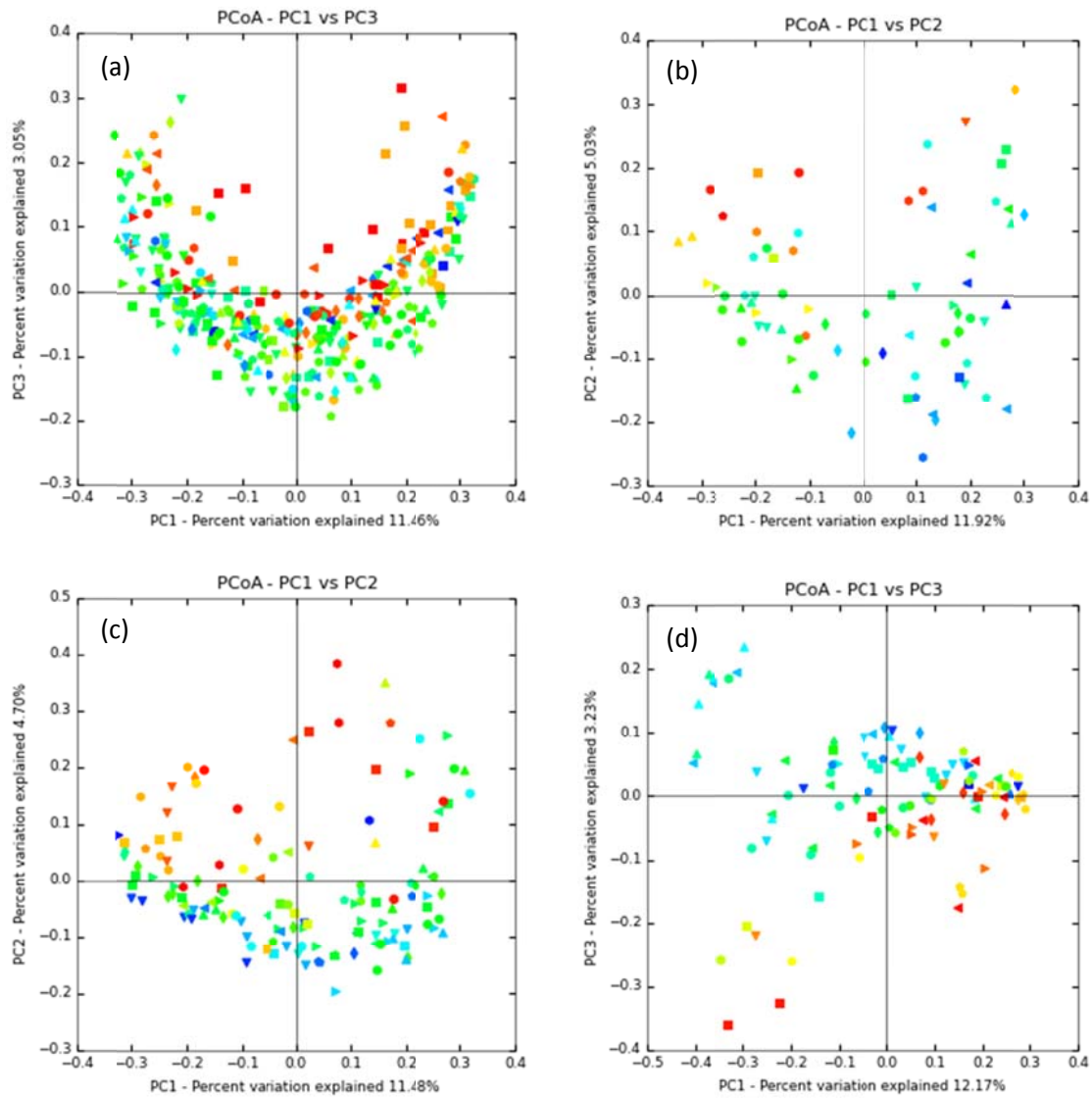


Fig. 3.5. Principal coordinate analysis of unweighted UniFrac for all samples combined (a, $n = 346$) as well as separated by fecal (b, $n = 91$), cloacal (c, $n = 146$), and oral (d, $n = 109$) sample types. Sample date within season depicted by color gradient with earlier dates in cooler colors and later dates in warmer colors. No consistent pattern between sampling date and similarity of microbiota recovered.

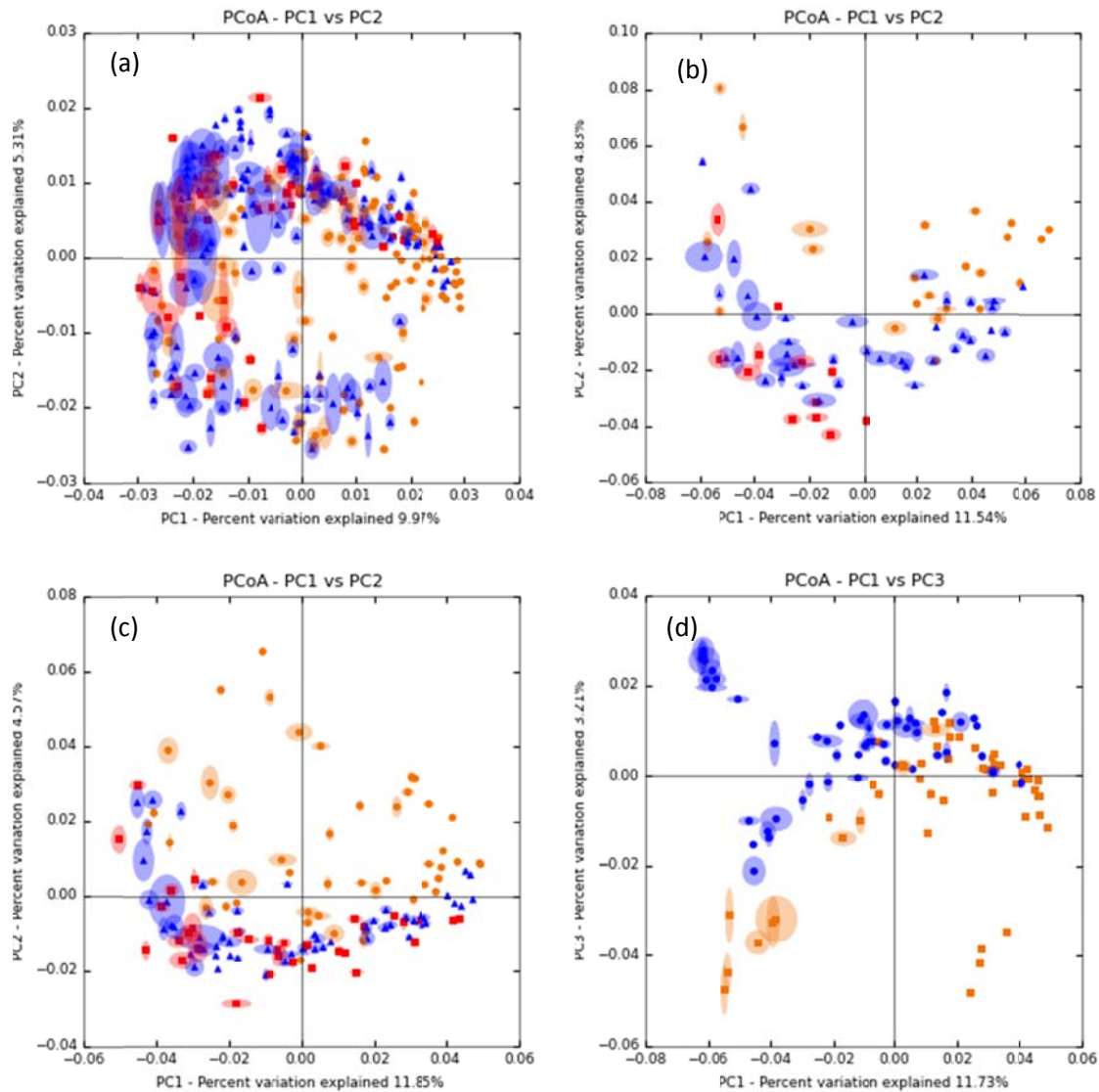


Fig. 3.6. Principal coordinate analysis of unweighted UniFrac for all samples combined (a, n = 346) as well as separated by fecal (b, n = 91), cloacal (c, n = 146), and oral (d, n = 109) sample types. Sample year of 2011, 2012, and 2013 depicted in red, blue, and orange, respectively. Ellipsoids indicate interquartile ranges for each sample from jackknifed analysis of beta diversity. Distinct segregation of 2013 samples depicted.



Fig. 3.7. Photos of unripe mistletoe berries in a blue oak tree (a), close-up of unripe mistletoe berries with a split pericarp in the bottom right berry (b), and abundant dried, split mistletoe pericarps as well as visibly purple fecal matter in a Western Bluebird nest on June 19, 2013 (c) at the San Joaquin Experimental Range, (Madera Co., CA).

Supporting Information

Figures

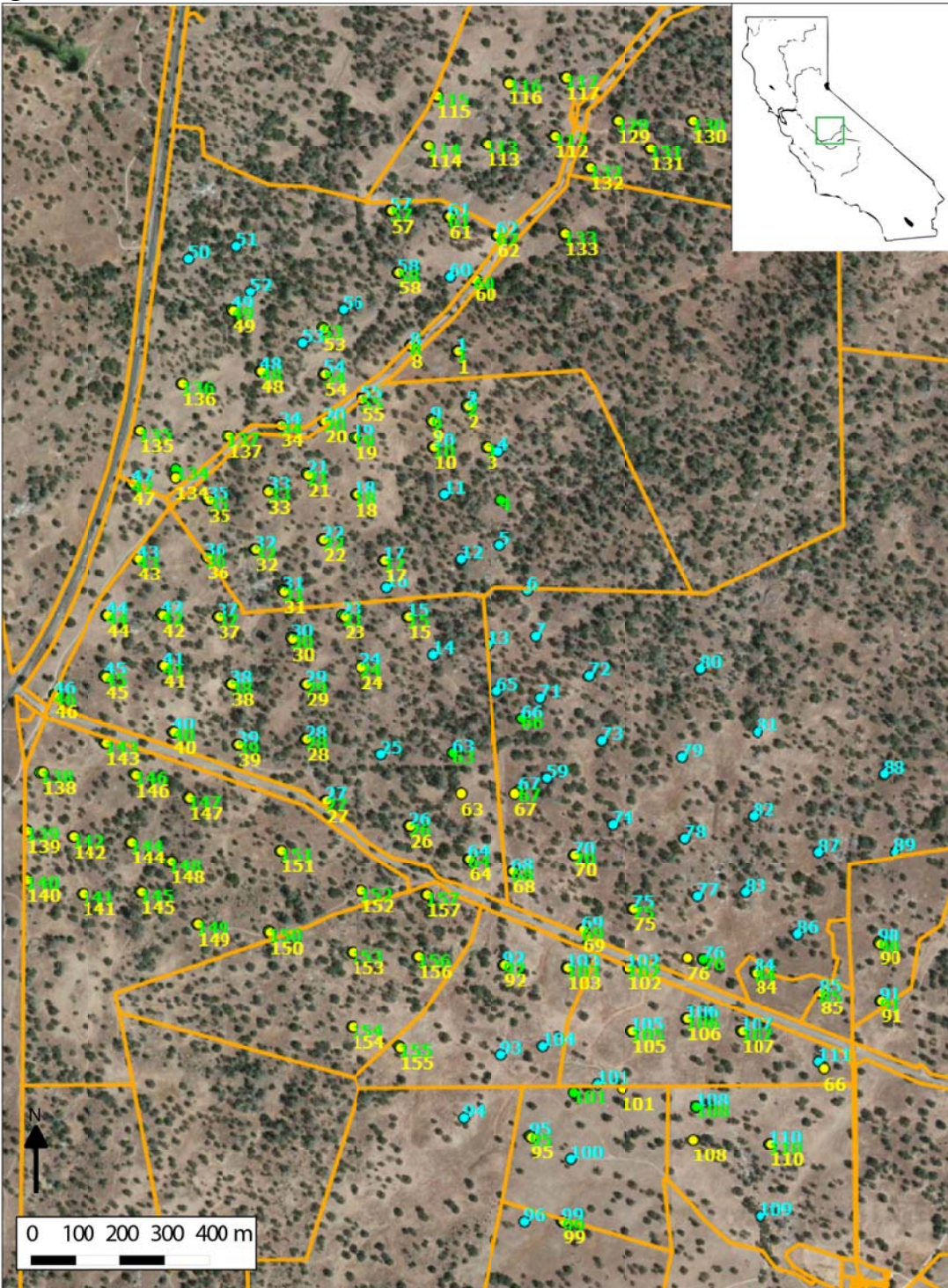


Fig. S3.1. Map of the Western Bluebird nest box system at the San Joaquin Experimental Range (Madera County, CA). Teal, green, and yellow circles and fonts indicate boxes erected in 2011, 2012, and 2013, respectively. Orange lines depict fences.

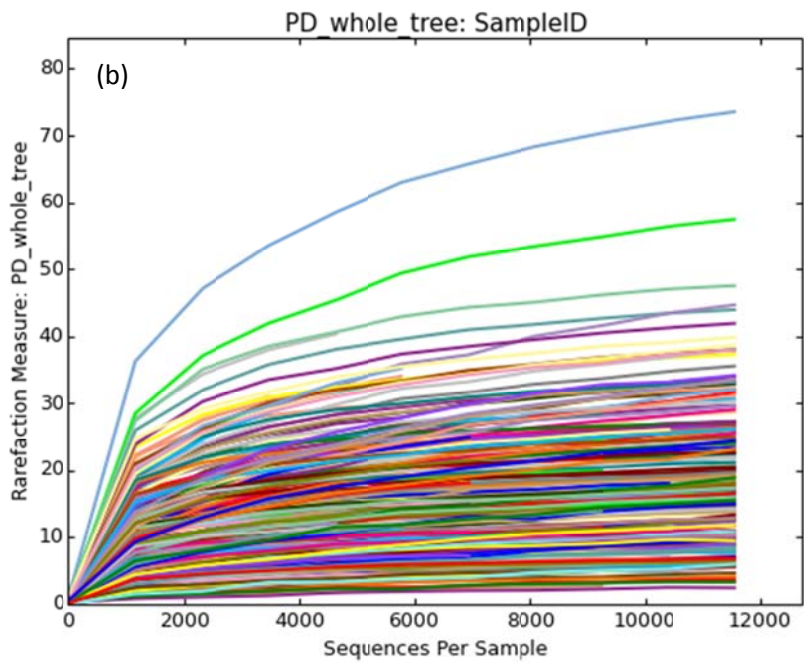
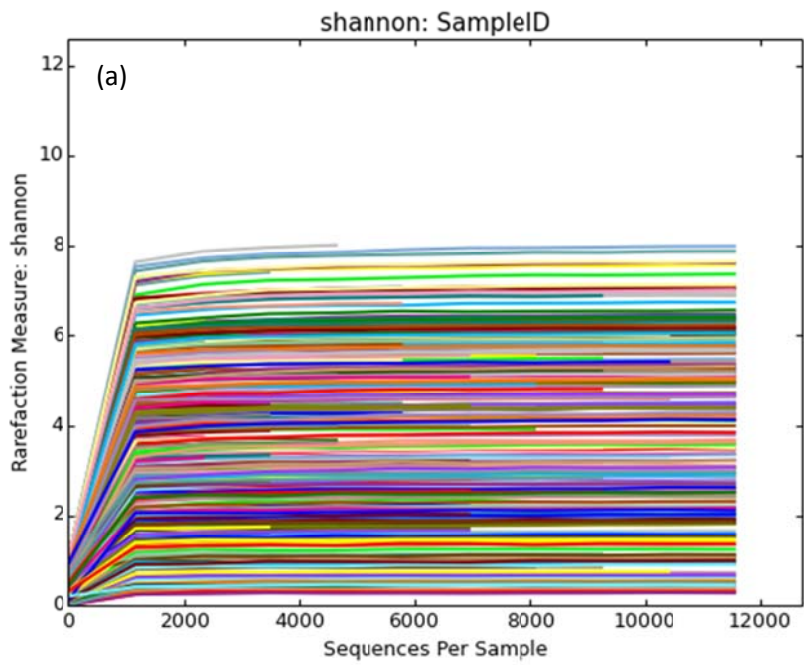


Fig. S3.2. Rarefaction analysis of Shannon's diversity index (a) and phylogenetic distance (b) for microbiota in all samples collected in study (n = 346). Each line and color depicts a different sample.

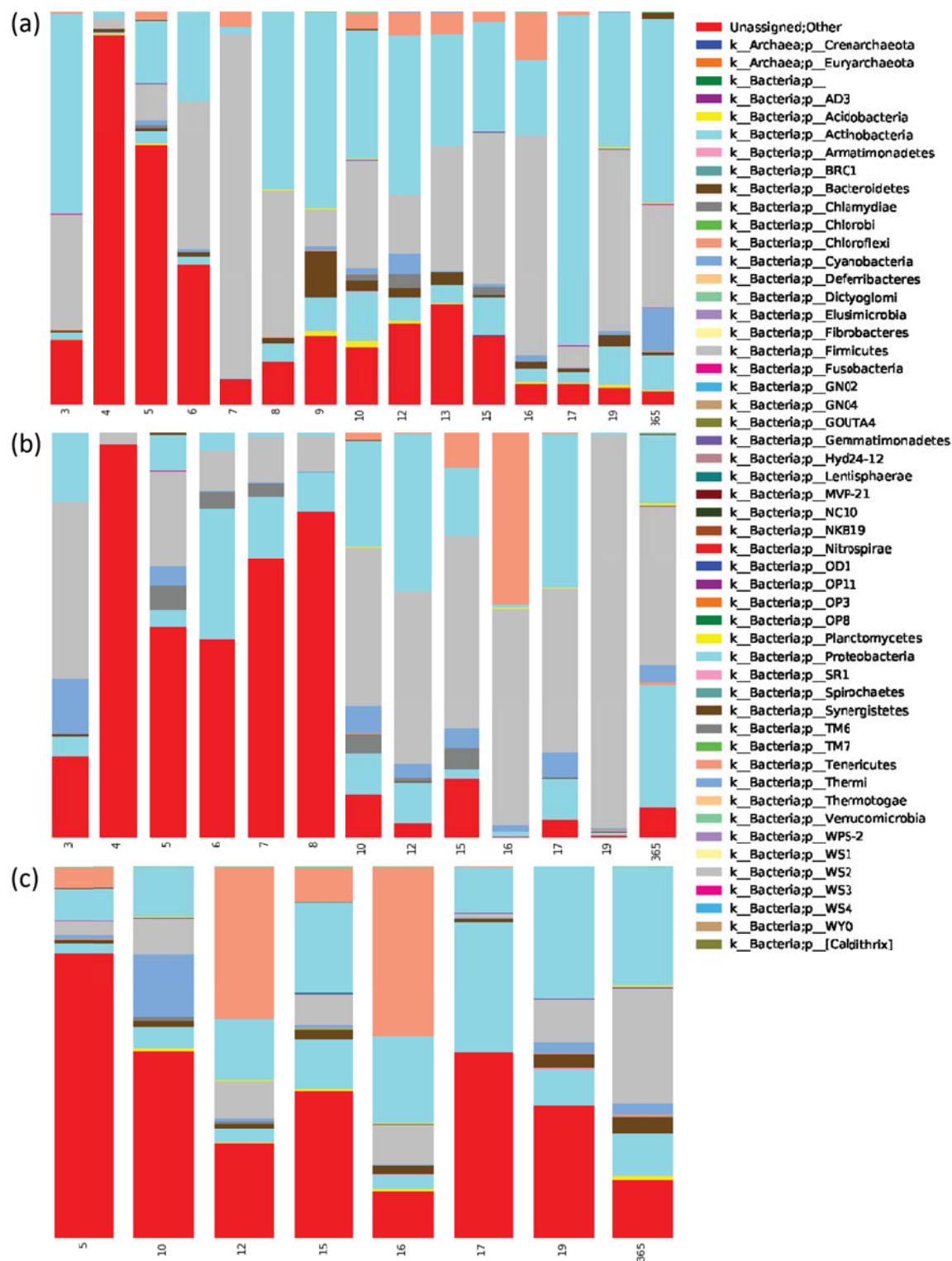


Fig S3.3. Relative abundance of bacterial phyla in avian cloacal (a), fecal (b), and oral (c) samples at various host ages (x-axis indicates days post-hatching, where 365 indicates adults).

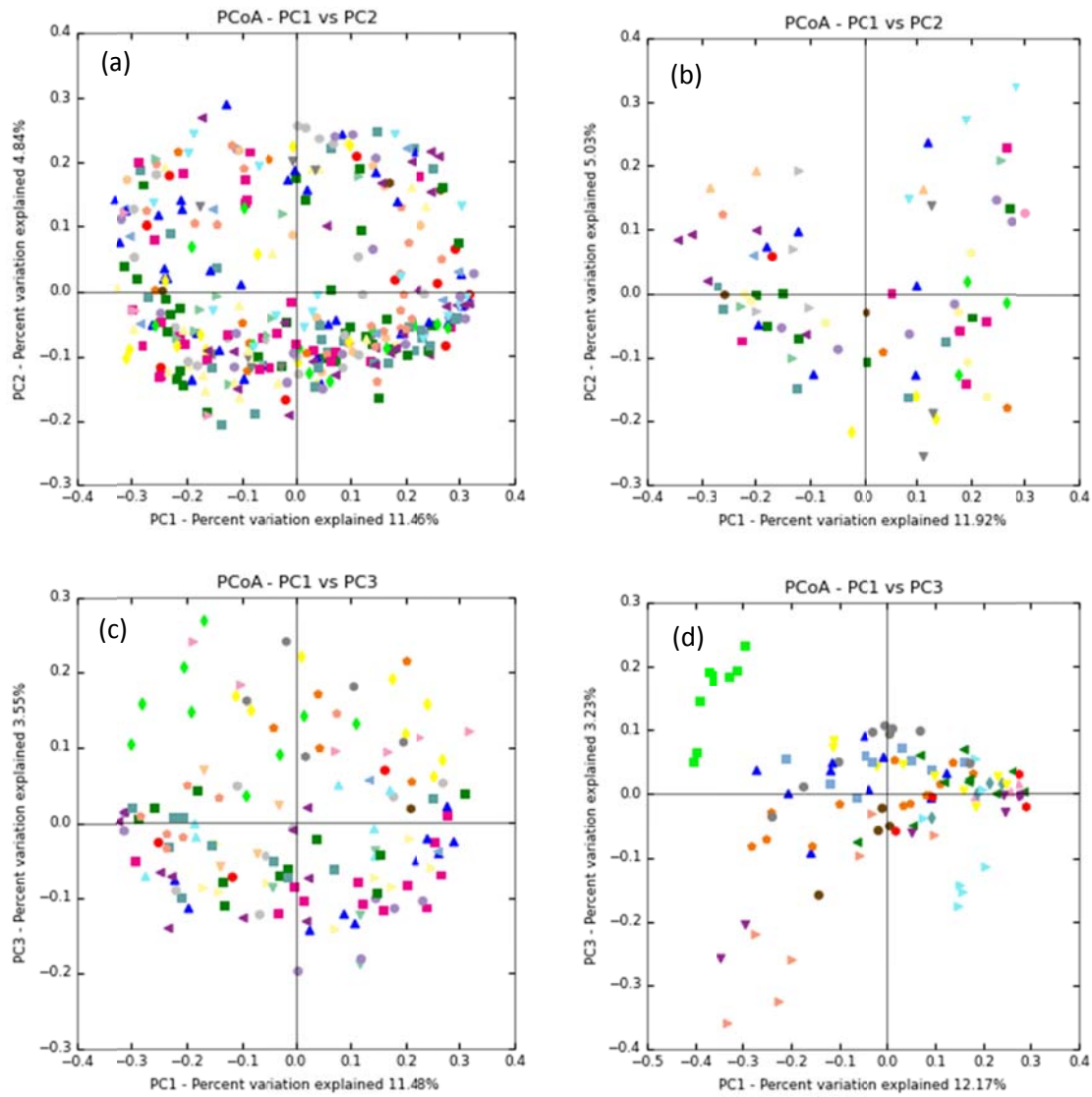


Fig. S3.4. Principal coordinate analysis of unweighted UniFrac for all samples combined (a, $n = 346$) as well as separated by fecal (b, $n = 91$), cloacal (c, $n = 146$), and oral (d, $n = 109$) sample types. Nests depicted by different colors. No consistent pattern recovered between shared nest and similarity of microbiota samples.

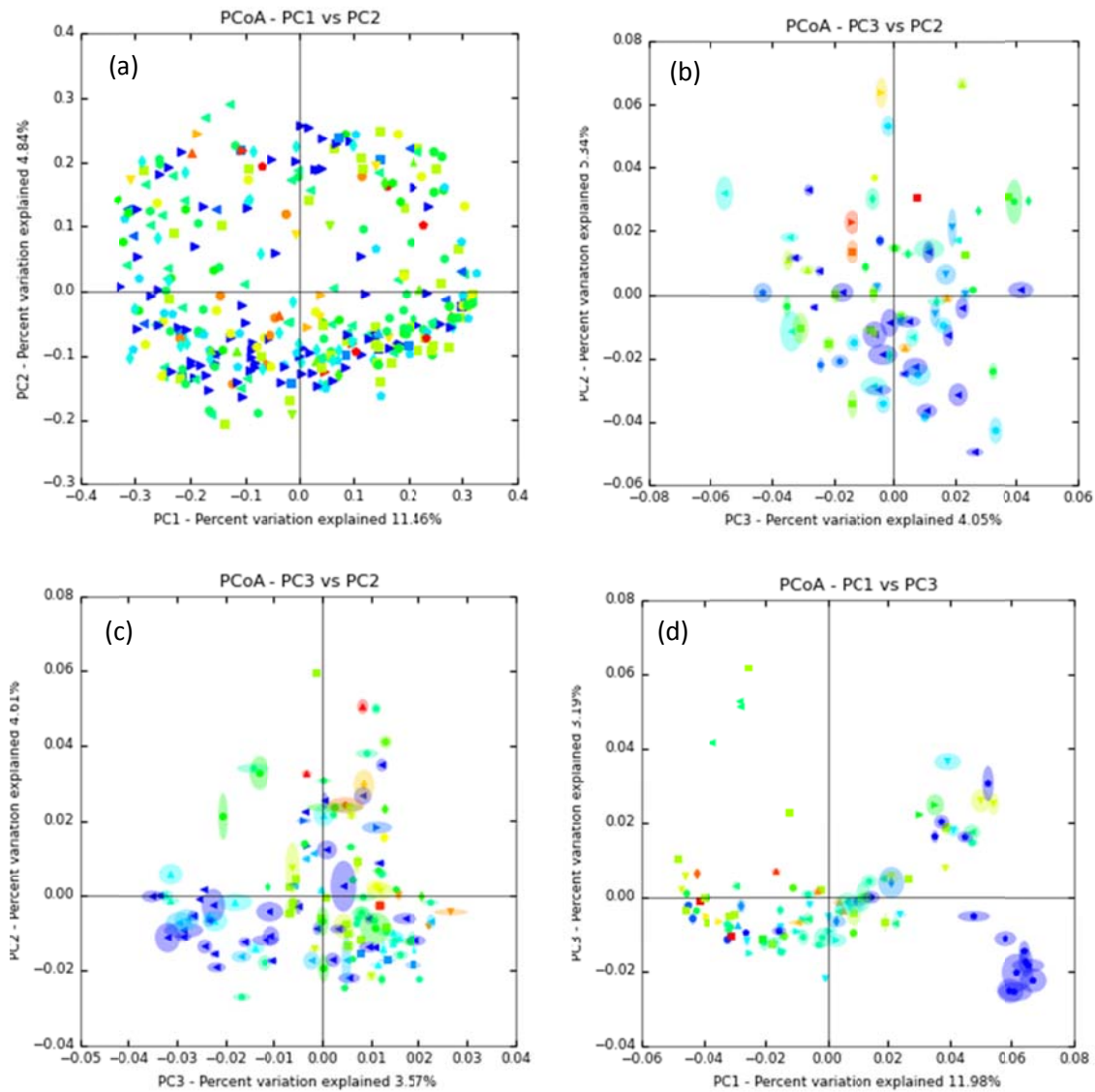


Fig. S3.5. Principal coordinate analysis of unweighted UniFrac for all samples combined (a, $n = 346$) as well as separated by fecal (b, $n = 91$), cloacal (c, $n = 146$), and oral (d, $n = 109$) sample types. Maximum nest box temperature (between hatch date and sampling date) depicted by color gradient with lower values in cooler colors and higher values in warmer colors. Ellipsoids indicate interquartile ranges for each sample from jackknifed analysis of beta diversity. No consistent pattern recovered between maximum temperatures and similarity of microbiota samples.

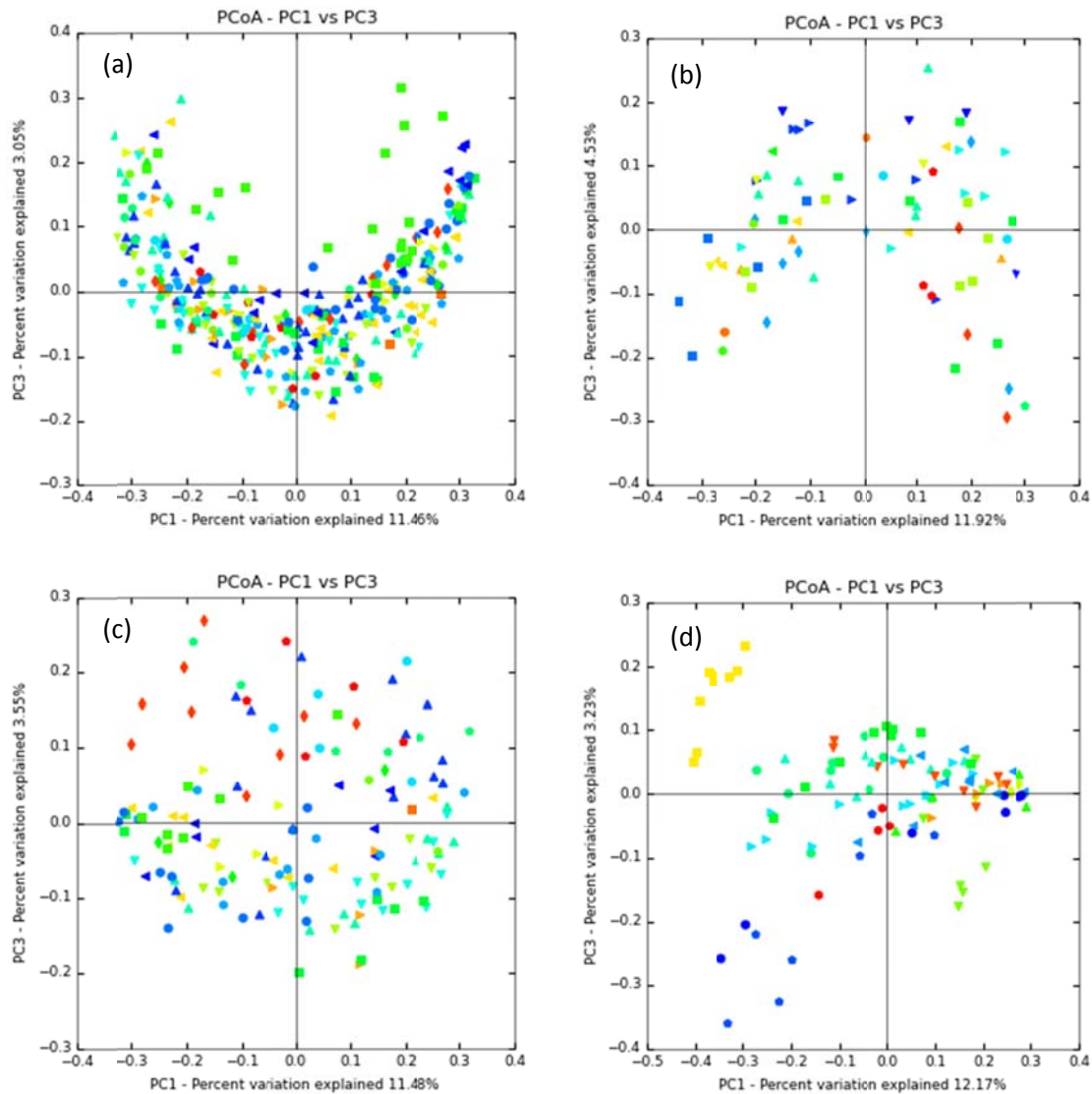


Fig. S3.6. Principal coordinate analysis of unweighted UniFrac for all samples combined (a, $n = 346$) as well as separated by fecal (b, $n = 91$), cloacal (c, $n = 146$), and oral (d, $n = 109$) sample types. Longitude depicted by color gradient with lower values in cooler colors and higher values in warmer colors. No consistent pattern recovered between geographic proximity and similarity of microbiota samples.

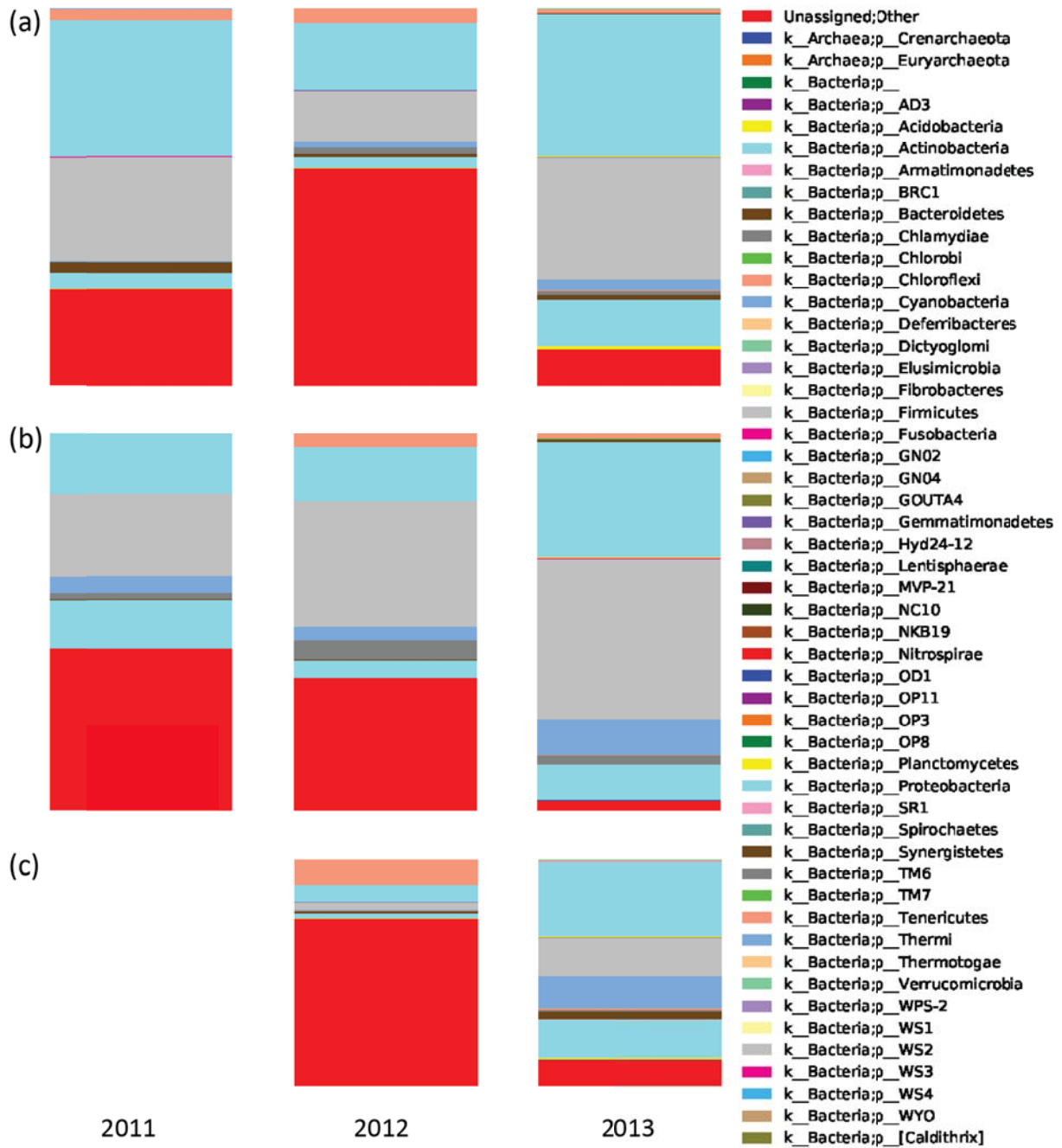


Fig. S3.7. Relative abundance of bacterial phyla in avian cloacal (a), fecal (b), and oral (c) samples across different sampling years (2011–2013). Distinct taxonomic profiles observed for 2013 samples relative to prior years.

Chapter 4: Relationships between the avian gut microbiota and nestling development

Abstract

Recently, a growing body of work has demonstrated the pervasive effects that microbiota can have on the physiology, ecology, and evolution of vertebrate organisms, but to date, most studies have focused on humans and rodent model systems. To address the potential for microbial mediation of phenotypic diversity in wild populations of non-model vertebrates, this study utilized deep sequencing of bacterial 16S rRNA amplicons to characterize the relationship between gut microbial diversity and host morphology in a natural population of nestling Western Bluebirds *Sialia mexicana* at the San Joaquin Experimental Range (Madera County, California) over 2011 through 2013. Cloacal, oral, and fecal samples were collected from nestlings at five, ten, and fifteen days post-hatching. Morphological measurements included nestling mass, mean length and asymmetry of the tarsi as well as the ninth primary flight feathers. To evaluate the relative contribution of gut microbiota in explaining variation in morphological metrics, additional spatiotemporal factors (ambient temperature, brood size, and ectoparasite load) were incorporated into analyses (mixed effects models with microbial alpha diversity as well as parametric and nonparametric *t*-tests, partial Mantel, and multiple matrix regression with randomization analyses with microbial beta diversity). After controlling for the spatiotemporal factors, microbial diversity across gut sites was significantly and differentially correlated with nestling morphological variation. Nevertheless, nestling age, sampling year, and temperature were more robust predictors of nestling morphology than gut microbiota. The results suggest a role for microbiota, in conjunction with spatiotemporal variables, in mediating nestling growth and development. Taken together with the findings of previous research (Chapter 3), environmentally mediated shifts in gut microbiota may in turn impact host condition, and additional experimental work to evaluate this potential is warranted to understand the implications of global change.

Introduction

In recent years, research on host-associated microbiota has flourished under the technological advances of massively parallel sequencing (Weinstock 2012), and arguably the most exciting findings have centered on the extensive impacts that microbiota can have on host phenotypic diversity. The use of germ-free mice in microbial colonization experiments has been especially powerful in determining the transmissibility of putative, microbial-mediated phenotypes and thus, causal relationships. From these studies, microbiota are known to decrease bone-mass density (Sjögren *et al.* 2012), enhance postnatal gut angiogenesis (Stappenbeck *et al.* 2002), and support development of robust immune responses to pathogen infection (Round & Mazmanian 2009). Microbial production of short-chain fatty acids from dietary carbohydrates modulates energy availability (Bergman 1990), water absorption (Hume 1997), gut motility, wound healing (Scheppach 1994), and fat storage (Wostmann *et al.* 1983; Bäckhed *et al.* 2004). Furthermore, obesity is transmissible through transplant of gut microbiota from obese mice to germ-free mice, and leanness is inducible through co-housing and sharing of microbiota between lean and obese mice on the same diet (Ridaura *et al.* 2013).

Given the demonstrated pervasiveness of microbial effects on host biology in humans and mice, there is a need to understand whether such effects are relevant to host condition

and fitness in free-living populations of non-model vertebrates. That is, to what extent does variation in microbiota impact host ecology and evolution in dynamic, natural settings? The neonatal period comprises a particularly vulnerable developmental period during which metabolism, and thus gut microbial function, is likely to exert a large effect on host condition. Nestlings afford a tractable model system for repeated sampling of free-living vertebrates to characterize relationships between microbiota and host condition, yet limited studies exist and have focused predominantly on cloacal microbiota. Specifically, abundance of gram-negative enteric lactose fermenters positively correlates with wing asymmetry among Tree Swallow *Tachycineta bicolor* nestlings (Mills *et al.* 1999). In Pied Flycatcher *Ficedula hypoleuca* nestlings, presence of enterococci explains 10%, 12%, and 22% of the variation in wing length, tarsus length, and mass, respectively (Moreno *et al.* 2003). Interestingly, antibiotic treatment supports faster growth rates in Magellanic Penguin *Spheniscus magellanicus* chicks (Potti *et al.* 2002). Thus, positive and negative correlations have been observed between morphological variation and diversity of the nestling cloacal microbiota.

This study applies massively parallel sequencing towards interrogation of avian-microbial relationships across broader spatiotemporal scales than previous studies to gain greater insights into microbial mediation of nestling growth performance. Specifically, concordance of relationships between microbiota and morphology of Western Bluebird *Sialia mexicana* nestlings are evaluated across several gut sites, nestling periods, and breeding seasons. Given that prior research revealed higher similarity between cloacal and oral microbiota than either with fecal microbiota (Chapter 3), it is predicted that oral and cloacal microbiota may exhibit similar associations whereas fecal microbiota may exhibit unique associations with morphological metrics. Additionally, relationships between microbiota and host condition are analyzed in the context of other temporal and environmental factors with known relevance for nestling growth in order to assess the relative significance of the microbiota. It is predicted that models including microbial diversity in conjunction with temporal and environmental factors are better predictors of variation in nestling morphology than models that do not account for host-associated microbiota. The results of this study are critical to interpreting the implications of significant associations between microbial diversity and ambient temperature as well as sampling year identified in previous research (Chapter 3). That is, this study will provide insights into the extent to which variation in microbiota in response to environmental heterogeneity may harbor ecologically significant implications for host fitness.

Materials and methods

Study site

Over 100 nest boxes were established at the San Joaquin Experimental Range (Madera County, CA) to study host-associated microbiota in Western Bluebirds (Fig. S4.1). The Western Bluebird is a cavity-nesting passerine that inhabits oak woodlands, and it readily breeds in nest boxes. Breeding activity typically begins in March and extends through July (Guinan *et al.* 2008). Clutch sizes range from three to eight eggs with a mean of five eggs. The female incubates and broods alone, with incubation periods of 12–14 days and brooding periods up to seven days post-hatching. Nestlings fledge within 18–21 days post-hatching.

Sampling and Illumina sequencing of bacterial 16S rRNA amplicons

From 2011–2013, the boxes were monitored closely throughout the breeding season to locate active nests and determine hatch dates of all nestlings. Distinct superciliary feather clipping patterns were used to distinguish same-aged nestlings within asynchronously hatched nests. To identify individual nestlings, each nestling was banded with a uniquely numbered aluminum band (USGS Bird Banding Laboratory) at five days post-hatching as well as a unique combination of three additional color bands at 10 days post-hatching (CA Fish and Game Permit SC-11869 and USGS Permit 21859). Nestlings were generally sampled at five, 10, and 15 days post-hatching. On each sample date, oral swab, cloacal swab, and fresh fecal samples were collected and stored following Vo and Jedlicka (2014).

In addition to microbial samples, morphological measurements were obtained for all nestlings on each sampling date. Body size is known to correlate with fledging success and survival of nestlings (Magrath 1991; Arnold & Griffiths 2003) whereas body asymmetry is thought to indicate developmental stress and environmental instability (Eeva *et al.* 2000; Grieco 2003). Furthermore, primary flight feathers, particularly the outer three primaries, are critical for flight performance (Swaddle *et al.* 1996; Chai 1997). Therefore, measurements of nestling mass, tarsus lengths, and ninth primary lengths were collected to determine body size and asymmetry as condition indices relevant for host fitness. The length of each tarsus was measured in duplicate or triplicate using a digital caliper (± 0.03 mm). The length of each ninth primary feather was determined using a stopped ruler (± 0.5 mm), and nestling mass was recorded using an electronic balance (± 0.01 g).

Concurrently, candidate abiotic and biotic environmental factors known to affect avian growth and development were quantified throughout the nestling period. Embryos and chicks exhibit depressed growth rates with extreme ambient temperatures (Donkoh 1989; Olson *et al.* 2006). Likewise, larger brood sizes impact nestling growth through limitations on parental care and feeding rates (Dijkstra *et al.* 1990), and hematophagous (i.e., blood-feeding) ectoparasites, commonly found in avian nests, are expected to deplete host resources for growth (Merino & Potti 1995). Thermochron iButton data loggers (Maxim DS1921G) were installed on the interior backboard (2 inches from the ceiling) of every active box to record temperature once every hour. Initial brood size was recorded as the number of nestlings in a given nest upon hatching whereas the number of nestlings in a nest immediately before fledging comprised the final brood size. Ectoparasite load was determined for each nestling per sampling date through counts of blowfly larva, mites, and lice, performed immediately upon removal of nestlings from the nest to minimize ectoparasite loss from handling. All protocols were approved by the UC Berkeley Animal Care and Use Committee (MAUP R317-0913).

Following Vo and Jedlicka (2014), metagenomic DNA extraction from all fecal and swab samples was performed using solid phase reversible immobilization beads (SPRI). Degenerate, barcoded primers targeting the bacterial 16S rRNA locus (515F and 806R; Caporaso *et al.* 2012) were used to amplify each sample extract in triplicate. Amplicons were pooled and prepared into sequencing libraries using a “with-bead” and PCR-free (Kozarewa *et al.* 2009) approach. The final libraries were pooled and submitted to the University of California Davis Genome Center for two separate 250 bp, paired-end runs on the Illumina MiSeq platform (see Supporting Information for detailed description of laboratory methods).

Data analyses

Illumina reads were trimmed, demultiplexed, quality filtered, and merged before analyses in QIIME 1.8.0 (see Supporting Information for details). Using the Greengenes v. 13_8 database and clustering at 97% identity, reference-based and de novo chimeras were removed prior to operational taxonomic units (OTUs) picking and taxonomy assignment. FastTree 2.1.3, implemented within QIIME, was used to build a maximum-likelihood phylogenetic tree containing all OTUs (Price *et al.* 2010). To obtain alpha diversity metrics for each sample, rarefaction analysis was performed to yield phylogenetic distance (Faith & Baker 2007) and Shannon's diversity index at various sequencing coverage depths. For derivation of beta diversity metrics, 16S rRNA OTU relative abundance tables were first rarefied to 2,312 sequences per sample, the sequencing depth at which phylogenetic distance and Shannon's diversity index plateaued for generally all samples (Vo Chapter 3, Fig. S3.2). Unweighted and weighted UniFrac (Lozupone & Knight 2005) distance matrices, calculated with the exclusion and inclusion of OTU abundance data respectively, were then generated from the rarefied OTU tables.

To prepare morphological data for downstream analyses, mean tarsus and ninth primary lengths were calculated from measurements for both legs and wings, respectively. Fluctuating asymmetry (Lens *et al.* 2002) was also derived for tarsus and ninth primary measurements as $A = (R_i - L_i) / ((R_i + L_i) / 2)$, where R_i is the right side measurement and L_i the left side measurement. Finally, mass normalized by mean tarsus was used as a growth performance index.

Mixed effects analysis with the *nmle* (Pinheiro *et al.* 2013) and *mgcv* (Wood 2011) packages in R 3.01 (R Core Team 2013) was applied to model variation in morphological metrics with microbial alpha diversity metrics, Julian date, year, brood size, temperature, and ectoparasite load as fixed effects as well as individual, host age, and host nest as random effects (Table 4.1). For each morphological metric, a global model including all fixed and random effects was developed for each microbial alpha diversity metric per body site (Table S4.1). Visual inspection of residual plots was used to detect deviations from homoscedasticity, linearity, or normality. Violations of assumptions were addressed using different random structures, variance structures, or generalized additive mixed modeling. Models were evaluated using Akaike's Information Criterion (AIC). P-values for each effect in a given model were obtained using either likelihood ratio tests of the full model versus the model without the effects in question for balanced linear models, Wald tests for effects in linear models with varying levels of missing data, Bayesian estimated covariance matrices for parametric terms in additive models, or frequentist properties of Bayesian confidence intervals for smooth terms in additive models.

For analyzing associations between microbial beta diversity and morphological distances, principal coordinate analysis (PCoA), parametric *t*-tests, and nonparametric *t*-tests with 1000 Monte Carlo randomizations were performed using microbial unweighted UniFrac distance matrices (Table S4.2). Parametric and nonparametric *t*-tests solely detected whether UniFrac distances differed within-level versus between-levels of morphological metrics. For example, a significantly higher UniFrac metric associated with between-level comparisons of tarsus lengths indicates a higher similarity of microbiota among individuals with a 10.0 mm

tarsus length relative to microbiota between individuals with tarsus lengths of 10.0 and 16.0 mm.

To further test the significance of associations between unweighted UniFrac distances and nestling morphology while controlling for the effects of other variables, partial Mantel tests and multiple matrix regressions with randomization (Wang 2013) were performed. QIIME was used to create distance matrices for all continuous factors (Table 4.1). For each morphological metric, a partial Mantel test was performed with all independent factors coded in every possible combination of explanatory or controlled effect per body site (Table S4.3). Significance of an association between an independent and dependent variable was accepted if tests with all confounding variables returned Bonferroni-corrected p-values less than or equal to 0.05. Likewise in MMRRs, a model including all independent factors was evaluated for each body site (Table S4.4). The significance of each independent factor for explaining variation in morphological metrics among samples was determined using Bonferroni corrected p-values. To interpret the results of partial Mantels and MMRRs, a significantly positive association between distances in nestling mass and for example distances in maximum temperature reveals decreased similarity in mass with greater differences in maximum temperature of the nest environment.

Causal relationships cannot be derived from associations. Therefore, significant associations from any test may indicate an effect of the microbiota on host morphology, an effect of host morphology on the microbiota, or a common effect of unmeasured variable(s) on both the microbiota and host morphology. However, consistency of associations between microbiota and several measures of host morphology across approaches suggests an effect of microbiota that warrants further experimental validation.

Results

PCoAs did not reveal visible patterns of microbiota clustering along gradients of the continuous morphological variables, even in sub-analyses by sample type or sampling year (Figs. S4.2-S4.7). Nevertheless, parametric and nonparametric *t*-tests recovered the highest number of significant associations between all six morphological metrics and the total microbiota as well as specific microbiota among gut sites (Table 4.2). Analyses that accounted for the confounding effects of 12 temporal and environmental variables (i.e., mixed models, partial Mantels, and MMRRs) recovered fewer significant relationships (Table 4.2). Among the latter analyses, the total ($n = 325$) and specifically the fecal ($n = 90$) microbiota varied significantly with growth performance whereas both the total and specifically the cloacal ($n = 136$) microbiota associated significantly with mean length of the ninth primaries (Table 4.2). Additionally, the fecal microbiota uniquely correlated with variation in nestling mass, and likewise, only the oral microbiota ($n = 99$) significantly varied with asymmetry of the ninth primaries (Table 4.2).

Regarding the temporal and environmental factors analyzed in conjunction with the microbiota, analyses consistently revealed nestling age and sampling year as significant predictors of variation for all morphological metrics (Fig. 4.1). The significance of sampling day within the season was observed for all morphological metrics in best fit mixed models and for nearly all metrics in MMRRs but was not recapitulated in partial Mantel tests (Fig. 4.1). Maximum, minimum, and standard deviation of temperatures were significant predictors of morphological variation in mixed model analyses whereas mean, median, and temperature

range significantly correlated with nestling morphology in distance-based analyses (i.e., MMRRs, Fig. 4.1). Relative to the temporal and temperature-related variables, brood size, ectoparasite load, and gut microbiota were less frequently recovered as significant predictors of nestling morphology (Fig. 4.1).

Discussion

This study characterized the gut microbiota of free-living birds to investigate the relationship between composition of the microbiota and nestling morphology in the context of co-occurring factors that may also mediate host development. In doing so, the microbiota at three gut sites and the development of six morphological indices were analyzed. After accounting for 12 measured temporal and environmental factors, a significant association was observed between i) the total microbiota and mean length of the ninth primaries as well as growth performance, ii) cloacal microbiota and mean length of the ninth primaries, iii) fecal microbiota and nestling mass as well as growth performance, and iv) oral microbiota and asymmetry of the ninth primaries. Regarding the relative significance of the microbiota for nestling growth, nestling age, sampling year, and ambient temperature tended to figure more prominently than the gut microbiota as significant predictors of nestling morphology.

Drawing connections between the structure and function of host-associated microbiota remains a challenge in non-model organisms, but the results of this study provide insights into relationships between taxonomic and functional diversity of the avian gut microbiota. Prior research on variation in the gut microbiota of nestling Western Bluebirds revealed high similarity of oral and cloacal microbiota, with a predominance of novel OTUs and Proteobacteria, whereas the fecal microbiota were uniquely dominated by Firmicutes (Chapter 3). In the analyses of this study, both cloacal and oral microbiota exhibited significant associations with feather growth whereas the fecal microbiota were associated with skeletal (tarsus) development and mass gains. The concordant morphological associations of the mucosal microbiota suggest correspondence between microbial taxonomic composition and functional capacity across two distant gut sites.

A greater influence of mucosal microbiota relative to gut lumen microbiota on host physiology is commonly hypothesized due to the opportunity for microbiota or their metabolites in the mucosa to directly interact with host cells (Koropatkin *et al.* 2012). Microbes that colonize the gut epithelium can bind of host immune receptors that affect proinflammatory signaling cascades, gut permeability, and systemic inflammation (Cani *et al.* 2009; Amar *et al.* 2011; Schertzer *et al.* 2011). Likewise, mucosal bacteria can generate metabolites that are beneficial to the host, including butyrate, propionate, and acetate which contribute to lipogenesis and gluconeogenesis (Wong *et al.* 2006; Velagapudi *et al.* 2010; Koropatkin *et al.* 2012). Nevertheless, few comparisons of associations between microbiota from various vertebrate gut sites and host physiological measures have been pursued to empirically evaluate the purported greater significance of mucosal microbiota (but see (Koren *et al.* 2011). This research contributes to the current body of work through suggesting substantial roles for both mucosal and luminal gut microbiota in mediating variation in host morphology.

Independent of the microbiota, many temporal and environmental factors have been shown in previous studies to play relevant roles in the process of nestling growth and

development. This study included analysis of the relative significance of three temporal factors (host age, nestling period within season, and breeding year). Host age correlated significantly with host development, as may be expected, and nestling morphology exhibited higher variation between-seasons than within-season. Regarding interannual variation, total precipitation ranged from 17.49 inches in 2011 to 18.96 inches in 2012 to 2.12 inches in 2013. Likewise, primary productivity varied from 1918 pounds per acre in 2011 to 2110 pounds per acre in 2012 to 1463 pounds per acre in 2013 (San Joaquin Experimental Range unpublished data). Nestlings require a high-protein diet to support the energetics of somatic growth (Mock *et al.* 1991), and a marked absence of arthropods corresponding to drought conditions in 2013 appeared to contribute to a food-limited breeding season. Food restriction has been shown to decrease growth rates, reduce body size and mass, and increase morphological asymmetry in young birds of many species (Øyan & Anker-Nilssen 1996; Ziel & Visser 2001; Searcy *et al.* 2004; Pravosudov & Kitaysky 2006). Interestingly, some morphological patterns were consistent with these expectations (i.e., a positive association between year and tarsus asymmetry, a negative association between year and tarsus length) whereas others ran counter to predictions (i.e., a positive association between year and mass as well as ninth primary length, a negative association between year and tarsus asymmetry as well as ninth primary asymmetry).

Additionally, this study analyzed the relative significance of nine environmental variables, related to brood size, temperature, and ectoparasites, for nestling growth. Brood size has been shown to affect nestling growth given increased limitations on parental feeding rates with larger brood sizes (Mock *et al.* 1991) as well as mass-dependent scaling of nestling energy requirements over the nestling period (Weathers 1992). The results revealed decreased differences in growth performance among nestlings with increased initial brood sizes but greater differences in growth performance among nestlings with larger final brood sizes. Ambient temperature is also known to affect growth of altricial nestlings, as they are minimally capable of thermoregulation during the first week post-hatching and thus subject to effects of environmental temperature on metabolic function (Whittow & Tazawa 1991). In this study, extremes in temperature and temperature heterogeneity correlated with decreased nestling growth, which is consistent with observations in nestling Eastern Kingbirds (*Tyrannus tyrannus*; Murphy 1985) but interestingly inconsistent with relationships between maximum temperature and nestling growth in Tree Swallows (McCarty & Winkler 1999) and Black Kites (*Milvus migrans*; Hiraldo *et al.* 1990). Differences in absolute values of temperatures between these studies may explain the discordant results.

Finally, a great body of work has been completed to date on the relationship between ectoparasites and host growth and development in birds. In bluebirds, hematophagous blow fly (*Protocalliphora sialia*) larvae frequently parasitize nestlings. Although parasitization rates can reach hundreds of larvae per nest (Wittmann & Beason 1992), effects on host fitness have been equivocal. Some studies reported effects on blood hematocrit and hemoglobin levels among parasitized nestlings (e.g., Whitworth & Bennett 1992), but most studies observed no effect on nestling morphology or fledging success (e.g., Gold & Dahlsten 1983; Wittmann & Beason 1992; Johnson & Albrecht 1993). To the best of knowledge, no study has investigated the role of gut microbiota versus ectoparasites in host growth and development among wild nestlings. In this study, ectoparasite load was rarely recovered as a significant predictor of nestling morphology but was a significant predictor of tarsus asymmetry in mixed models. The latter finding supports

the possibility of ectoparasite effects on avian morphology in a field that generally encounters non-significant associations. Relative to ectoparasites, gut microbiota was comparably detected as either a significant predictor or important contributing factor to the best fit models of nestling morphological metrics.

By considering microbiota in conjunction with temporal and environmental factors, this study reveals a more dynamic, interactive network of potential processes underlying variation in host morphology. Importantly, associations between microbiota and morphological metrics did not simply co-vary with other temporal and environmental variables. For example, although the drought-stricken 2013 breeding season was associated with depressed growth in certain morphological measures, only the oral microbiota likewise exhibited significant associations with nestling morphology in 2013 whereas the cloacal and fecal microbiota significantly correlated with morphological variation in 2012. Although the gut microbiota less robustly predicted nestling morphology relative to other spatiotemporal factors, the results of this study still underscore the need for a broadening of conceptual frameworks to account for the “unseen majority” (i.e., host-associated microbiota) in mediating host phenotypic diversity.

Conclusion

Diversity of the total gut microbiota as well as microbiota in each gut site was significantly associated with nestling morphology, particularly with mass, growth of the ninth primaries, and growth performance. In addition, nestling age, ambient temperature, and brood size significantly correlated with nestling morphological variation, and interannual heterogeneity in relationships was observed. The potential for microbial impacts on host physiology is supported given the recovery of significant associations across multiple analytical approaches and given that the inclusion of microbiota in the best fit models did not simply co-vary with other spatiotemporal variables. The results suggest the combined relevance of variation in microbiota and spatiotemporal factors for host development. Taken together with the findings of previous research (Chapter 3), it is possible for microbial shifts in response to environmental heterogeneity to mediate effects on host physiology, independent of direct effects of the environment on host condition. Experimental work coupled with metatranscriptomic data would be invaluable for clarifying this possibility, with ecologically relevant implications in the context of global change. Additionally, although prior avian studies have reported associations between cloacal microbiota and asymmetry in flight feathers, this study recovered an unexpected association with the oral microbiota. Asymmetry in flight feathers can significantly impact flight speed, maneuverability, and performance among birds (Swaddle *et al.* 1996). Thus, future research encompassing gut microbiota, feather asymmetry, and biomechanical efficiency would comprise an exciting, novel avenue for bridging form, function, and fitness in metaorganismal biology.

References

- Amar J, Chabo C, Waget A *et al.* (2011) Intestinal mucosal adherence and translocation of commensal bacteria at the early onset of type 2 diabetes: molecular mechanisms and probiotic treatment. *EMBO Molecular Medicine*, **3**, 559–572.
- Arnold KE, Griffiths R (2003) Sex-specific hatching order, growth rates and fledging success in jackdaws *Corvus monedula*. *Journal of Avian Biology*, **34**, 275–281.

- Bäckhed F, Ding H, Wang T *et al.* (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15718–15723.
- Bergman EN (1990) Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiological Reviews*, **70**, 567–590.
- Cani PD, Possemiers S, Wiele TV de *et al.* (2009) Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. *Gut*, **58**, 1091–1103.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.
- Chai P (1997) Hummingbird hovering energetics during moult of primary flight feathers. *The Journal of Experimental Biology*, **200**, 1527–1536.
- Dijkstra C, Bult A, Bijlsma S *et al.* (1990) Brood Size Manipulations in the Kestrel (*Falco tinnunculus*): Effects on Offspring and Parent Survival. *Journal of Animal Ecology*, **59**, 269–285.
- Donkoh A (1989) Ambient temperature: a factor affecting performance and physiological response of broiler chickens. *International Journal of Biometeorology*, **33**, 259–265.
- Eeva T, Tanhuanpää S, Råbergh C *et al.* (2000) Biomarkers and fluctuating asymmetry as indicators of pollution-induced stress in two hole-nesting passerines. *Functional Ecology*, **14**, 235–243.
- Faith DP, Baker AM (2007) Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. *Evolutionary Bioinformatics Online*, **2**, 121–128.
- Gold CS, Dahlsten DL (1983) Effects of Parasitic Flies (*Protocalliphora* spp.) on Nestlings of Mountain and Chestnut-Backed Chickadees. *The Wilson Bulletin*, **95**, 560–572.
- Grieco F (2003) Greater food availability reduces tarsus asymmetry in nestling blue tits. *The Condor*, **105**, 599–603.
- Guinan JA, Gowaty PA, Eltzroth EK (2008) Western Bluebird (*Sialia mexicana*) (A Poole, Ed.). *The Birds of North America Online*.
- Hiraldo F, Veiga JP, Máñez M (1990) Growth of nestling black kites *Milvus migrans*: effects of hatching order, weather and season. *Journal of Zoology*, **222**, 197–214.
- Hume ID (1997) Fermentation in the Hindgut of Mammals. In: *Gastrointestinal Microbiology* Chapman & Hall Microbiology Series. (eds Mackie RI, White BA), pp. 84–115. Springer US.
- Johnson LS, Albrecht DJ (1993) Effects of Haematophagous Ectoparasites on Nestling House Wrens, *Troglodytes aedon*: Who Pays the Cost of Parasitism? *Oikos*, **66**, 255–262.
- Koren O, Spor A, Felin J *et al.* (2011) Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 4592–4598.
- Koropatkin NM, Cameron EA, Martens EC (2012) How glycan metabolism shapes the human gut microbiota. *Nature Reviews Microbiology*, **10**, 323–335.
- Kozarewa I, Ning Z, Quail MA *et al.* (2009) Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nature Methods*, **6**, 291–295.

- Lens L, Van Dongen S, Kark S, Matthysen E (2002) Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies? *Biological Reviews*, **77**, 27–38.
- Lozupone C, Knight R (2005) UniFrac: a New phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228–8235.
- Magrath RD (1991) Nestling weight and juvenile survival in the blackbird, *Turdus merula*. *Journal of Animal Ecology*, **60**, 335–351.
- Mccarty JP, Winkler DW (1999) Relative importance of environmental variables in determining the growth of nestling Tree Swallows *Tachycineta bicolor*. *Ibis*, **141**, 286–296.
- Merino S, Potti J (1995) Mites and blowflies decrease growth and survival in nestling Pied Flycatchers. *Oikos*, **73**, 95–103.
- Mills TK, Lombardo MP, Thorpe PA (1999) Microbial colonization of the cloacae of nestling tree swallows. *The Auk*, **116**, 947–956.
- Mock PJ, Khubesian M, Larcheveque DM (1991) Energetics of growth and maturation in sympatric passerines that fledge at different ages. *The Auk*, **108**, 34–41.
- Moreno J, Briones V, Merino S *et al.* (2003) Beneficial effects of cloacal bacteria on growth and fledging size in nestling Pied flycatchers (*Ficedula hypoleuca*) in Spain. *The Auk*, **120**, 784–790.
- Murphy MT (1985) Nestling Eastern Kingbird growth: effects of initial size and ambient temperature. *Ecology*, **66**, 162–170.
- Olson CR, Vleck CM, Vleck D (2006) Periodic cooling of bird eggs reduces embryonic growth efficiency. *Physiological and Biochemical Zoology*, **79**, 927–936.
- Øyan HS, Anker-Nilssen T (1996) Allocation of growth in food-stressed Atlantic puffin chicks. *The Auk*, **113**, 830–841.
- Pinheiro J, Bates D, DebRoy S, Sarker D, the R Development Core Team (2013) nlme: Linear and Nonlinear Mixed Effects Models.
- Potti J, Moreno J, Yorío P *et al.* (2002) Bacteria divert resources from growth for magellanic penguin chicks. *Ecology Letters*, **5**, 709–714.
- Pravosudov VV, Kitaysky AS (2006) Effects of nutritional restrictions during post-hatching development on adrenocortical function in western scrub-jays (*Aphelocoma californica*). *General and Comparative Endocrinology*, **145**, 25–31.
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE*, **5**, e9490.
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ridaura VK, Faith JJ, Rey FE *et al.* (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, **341**, 1241214.
- Round JL, Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nature Reviews Immunology*, **9**, 313–323.
- Scheppach W (1994) Effects of short chain fatty acids on gut morphology and function. *Gut*, **35**, S35–S38.
- Schertzer JD, Tamrakar AK, Magalhães JG *et al.* (2011) NOD1 activators link innate immunity to insulin resistance. *Diabetes*, **60**, 2206–2215.
- Searcy WA, Peters S, Nowicki S (2004) Effects of early nutrition on growth rate and adult size in song sparrows *Melospiza melodia*. *Journal of Avian Biology*, **35**, 269–279.

- Sjögren K, Engdahl C, Henning P *et al.* (2012) The gut microbiota regulates bone mass in mice. *Journal of Bone and Mineral Research*, **27**, 1357–1367.
- Stappenbeck TS, Hooper LV, Gordon JI (2002) Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 15451–15455.
- Swaddle JP, Witter MS, Cuthill IC, Budden A, McCowen P (1996) Plumage condition affects flight performance in Common Starlings: implications for developmental homeostasis, abrasion and moult. *Journal of Avian Biology*, **27**, 103–111.
- Velagapudi VR, Hezaveh R, Reigstad CS *et al.* (2010) The gut microbiota modulates host energy and lipid metabolism in mice. *Journal of Lipid Research*, **51**, 1101–1112.
- Vo AE, Jedlicka JA (2014) Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for fecal and swab samples. *Molecular Ecology Resources*, In press.
- Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, **67**, 3403–3411.
- Weathers WW (1992) Scaling nestling energy requirements. *Ibis*, **134**, 142–153.
- Weinstock GM (2012) Genomic approaches to studying the human microbiota. *Nature*, **489**, 250–256.
- Whittow GC, Tazawa H (1991) The early development of thermoregulation in birds. *Physiological Zoology*, **64**, 1371–1390.
- Whitworth TL, Bennett GF (1992) Pathogenicity of larval Protocalliphora (Diptera: Calliphoridae) parasitizing nestling birds. *Canadian Journal of Zoology*.
- Wittmann K, Beason RC (1992) The effect of blowfly parasitism on nestling Eastern bluebird development (Efecto del parasitismo de *Protocalliphora sialia* (Diptera) en el desarrollo de pichones de *Sialia sialis*). *Journal of Field Ornithology*, **63**, 286–293.
- Wong JMW, de Souza R, Kendall CWC, Emam A, Jenkins DJA (2006) Colonic health: fermentation and short chain fatty acids. *Journal of Clinical Gastroenterology*, **40**, 235–243.
- Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, **73**, 3–36.
- Wostmann BS, Larkin C, Moriarty A, Bruckner-Kardoss E (1983) Dietary intake, energy metabolism, and excretory losses of adult male germfree Wistar rats. *Laboratory Animal Science*, **33**, 46–50.
- Ziel CE van der, Visser GH (2001) The effect of food restriction on morphological and metabolic development in two lines of growing Japanese Quail chicks. *Physiological and Biochemical Zoology*, **74**, 52–65.

Tables

Table 4.1. Description of all variables used in analytical models in this study.

Variable	Description
<i>Independent</i>	
Age	Number of days post-hatching
Ectos	Total ectoparasite load
FinalBroodSize	Brood size immediately before fledging
InitialBroodSize	Brood size during early nestling period at five days post-hatching
Jday	Sampling date converted to Julian day
MaxTemp	Maximum temperature between hatch and sampling dates
MinTemp	Minimum temperature between hatch and sampling dates
PD	Phylogenetic distance of microbiota
SDTemp	Standard deviation of temperatures between hatch and sampling dates
Sex	Sex of individual assessed by plumage at 15 days post-hatching
Shannon	Shannon's diversity index of microbiota
Year	Sampling year
<i>Dependent</i>	
Avg_9	Mean of measurements of left and right ninth primary feathers
FA_T	Asymmetry of tarsus bones
FA_9	Asymmetry of ninth primary feathers
GrowthPerformance	Index derived from the quotient of Mass over Tarsavg
Mass	Mass of nestling
Tarsavg	Mean of duplicate measurements of left and right tarsus bones

Table 4.2. Summary of results of all five analytical approaches used to investigate the relationships between gut microbiota and nestling morphological metrics in this study (Tables S4.1, S4.3-4.4). “Yes” denotes a significant p-value ($\alpha = 0.05$), highlighted in red. “Some” indicates inclusion of microbiota in the best fit model despite a nonsignificant p-value for the factor in mixed models or significance of some but not all 12 partial Mantel tests. “No” denotes p-values higher than 0.05. Results are presented in order of significance in controlled analyses (i.e., mixed models, partial Mantels, and MMRRs).

Microbiota	Morphological metric	Parametric t-test	Nonparametric t-test	Mixed models	Partial Mantel	MMRR
All	Growth Performance	Yes	Yes	No	Some	Yes
	Mean ninth primary length	Yes	Yes	Some	No	Yes
	Mass	Yes	Yes	Some	Some	No
	Mean tarsus length	Yes	Yes	Some	No	No
	Ninth primary asymmetry	Yes	Yes	Some	No	No
	Tarsus asymmetry	Yes	Yes	No	No	No
Cloacal	Mean ninth primary length	Yes	Yes	Yes	No	No
	Mean tarsus length	Yes	Yes	Some	No	No
	Mass	Yes	Yes	No	No	No
	Tarsus asymmetry	Yes	Yes	No	No	No
	Ninth primary asymmetry	Yes	Yes	No	No	No
	Growth Performance	Yes	Yes	No	No	No
Fecal	Growth Performance	No	No	No	Yes	Yes
	Mass	No	No	Some	Yes	No
	Ninth primary asymmetry	Yes	Yes	Some	No	No
	Mean tarsus length	No	No	Some	No	No
	Mean ninth primary length	No	No	Some	No	No
	Tarsus asymmetry	No	No	No	No	No
Oral	Ninth primary asymmetry	Yes	Yes	Yes	No	No
	Mean ninth primary length	Yes	Yes	Some	No	No
	Mass	No	No	Some	No	No
	Mean tarsus length	No	No	Some	No	No
	Tarsus asymmetry	No	No	No	No	No
	Growth Performance	No	No	No	No	No

Figures

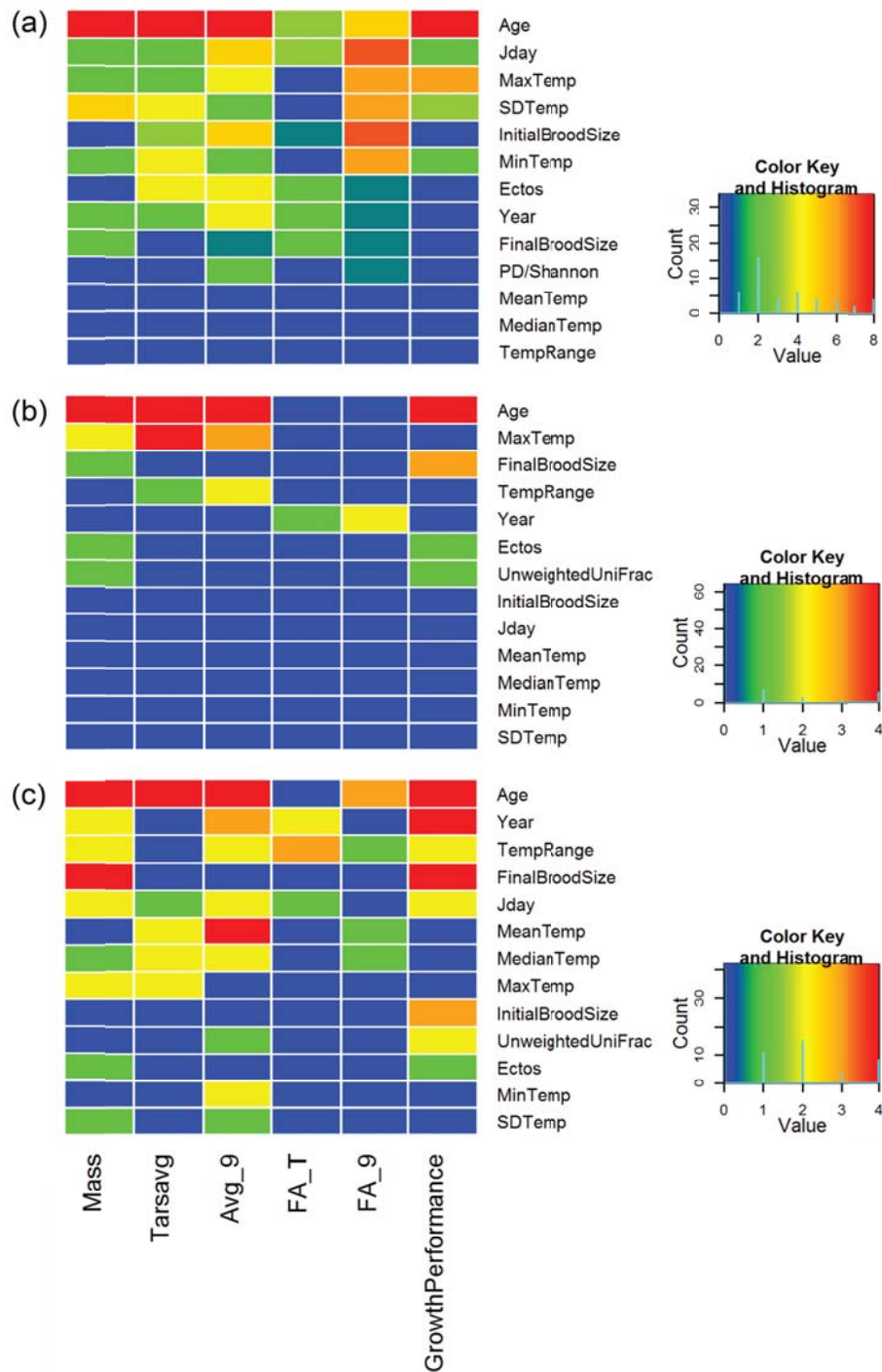


Fig. 4.1. Heatmaps depicting independent variables (right) ranked according to the number of times they were significantly associated with nestling morphological metrics (bottom) using three analytical approaches: mixed models (a), partial Mantel tests (b), and multiple matrix regressions with randomization (c). With each analytical approach, four analyses were performed including either cloacal, fecal, oral, or all samples. Variable codes are described in Table 4.1, and models are described in Tables S4.1, S4.3-4.4.

Supporting Information

Materials and methods

Sampling and Illumina sequencing of bacterial 16S rRNA amplicons

Fresh fecal samples were stored dry at -20°C in 2011 or in an ammonium-sulfate saturated nucleic acid stabilization and storage buffer (25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; Vo & Jedlicka 2014), not exceeding a sample to buffer ratio of 0.15 g/mL, at -20°C temporarily and then -80°C permanently in both 2012 and 2013. Cloacal swab heads (Puritan Medical PurFlock 25-3316 1PN) were stored in 1 mL of sucrose lysis buffer (20 mM EDTA, 50 mM Tris-HCl pH 9.0, 200 mM NaCl, 0.75 M sucrose; Mitchell & Takacs-Vesbach 2008) in 2011 whereas cloacal and oral swab heads were stored in DNA buffer during both 2012 and 2013.

For metagenomic DNA extraction, samples were homogenized with zirconia-silica beads (0.1 and 0.5 mm diameter) in lysis buffer (166 mM trisodium phosphate, 111 mM guanidine thiocyanate, 11 mM sodium chloride, 0.3% sodium dodecyl sulfate, 37 mM Tris-HCl pH 8.0, 1% beta-mercaptoethanol) on a Precellys 24 (Precellys EQ03119.200.RD000.0) or PowerLyzer 24 (Mo Bio 13255). Sample lysates were incubated with ammonium acetate to precipitate proteins and aluminum ammonium sulfate to flocculate PCR inhibitors (Braid *et al.* 2003). SPRI beads were then used to capture, wash, and elute sample DNA. The Nanodrop spectrophotometer and Qubit fluorometer dsDNA HS (high sensitivity) Assay Kit (Invitrogen Q32854) were used to quantify all extracts, and elutions were stored at -20°C until use.

The metagenomic DNA extracts were amplified following Vo and Jedlicka (2014) using degenerate, barcoded primers targeting the bacterial 16S rRNA locus (515F and 806R; Caporaso *et al.* 2012) in triplicate 25 µL reactions containing 1X HotMasterMix (5 PRIME 2200400), 0.2 µM of each primer, 0.5 µg/µL BSA (NEB B9000S), 5% DMSO (ACS grade), and at least 0.1 ng of DNA. Thermocycling conditions followed an initial denaturation of 94°C for 2 minutes and 35 cycles of 94°C for 8 seconds, 50°C for 20 seconds, and 72°C for 30 seconds. To minimize risk of cross-contamination, PCR reactions were set-up in an UV-irradiated clean hood and subsequently thermocycled in a separate, isolated room. Amplicons were electrophoresed on 1.5% agarose gels stained with SYBR-Gold (Invitrogen S-11494) to determine amplification success.

To prepare amplicon libraries for Illumina sequencing, triplicate PCR reactions were pooled and cleaned using a 1.2 volume ratio of SPRI bead solution to sample to discard fragments smaller than 200 bp (e.g., primer dimers). After quantification using the Qubit dsDNA HS Assay Kit, the cleaned samples were pooled equimolarly to form either nine libraries of six samples and 300-400 ng total amplicon DNA per library for a preliminary MiSeq run or 24 libraries of 18-21 samples and 81-507 ng of total amplicon DNA per library for the main MiSeq run. Following Vo and Jedlicka (2014), the NEBNext End Repair Module, NEBNext dA-Tailing Module, and Meyer and Kircher (2010) adapter ligation protocols were used with TruSeq Illumina indexed adapters to prepare the libraries using a “with-bead” and “PCR-free” approach. After quantification of each of the nine libraries for the initial MiSeq run or each of the 24 libraries for the main MiSeq run via qPCR using the KAPA Library Quantification Kit

(KK4824), libraries were pooled equimolarly and analyzed on the Bioanalyzer (Agilent Technologies G2940CA).

Data analyses

CASAVA 1.8 was used to trim the Illumina reads of adapters and to demultiplex the reads into the original libraries. Reads that failed to pass filter were removed, and Trimmomatic trimmed the trailing edges of sequences to a minimum quality score of Q20 while keeping paired-end reads in sync (Lohse *et al.* 2012). Following Bystrykh (2012), barcodes were error-corrected and those containing more than one error were discarded. PANDAseq merged paired-end reads with a quality threshold of 0.6 (Masella *et al.* 2012). Merged reads without matching barcodes at both ends or without correct primer sequences were also discarded. Using QIIME 1.8.0, each library was then demultiplexed by barcode, and all barcode and primer sequences were removed from each read before downstream analyses (Caporaso *et al.* 2010).

References

- Braid MD, Daniels LM, Kitts CL (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. *Journal of Microbiological Methods*, **52**, 389–393.
- Bystrykh LV (2012) Generalized DNA Barcode Design Based on Hamming Codes. *PLoS ONE*, **7**, e36852.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.
- Lohse M, Bolger AM, Nagel A *et al.* (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, **40**, W622–W627.
- Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, **13**, 31.
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, **2010**, pdb.prot5448.
- Mitchell KR, Takacs-Vesbach CD (2008) A comparison of methods for total community DNA preservation and extraction from various thermal environments. *Journal of Industrial Microbiology & Biotechnology*, **35**, 1139–1147.
- Vo AE, Jedlicka JA (2014) Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for fecal and swab samples. *Molecular Ecology Resources*, In press.

Tables

Table S4.1. Generalized least squares, linear mixed effects, generalized additive, and generalized additive mixed effects models analyzing associations between alpha diversity of gut microbiota and nestling morphological metrics in the context of several spatiotemporal factors that may mediate nestling growth and development. Inclusion of microbiota in best fit models and significant p-values are highlighted in yellow.

Body site	Model	Global model	Model with lowest AIC (significant factors bolded)	Variance structure	Random structure	Alpha diversity in final model?	Alpha diversity coefficient	Alpha diversity pvalue
All samples	lme	Mass ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + PD + MinTemp + SDTemp + Jday + Sex + FinalBroodSize + Ectos	-	~1 fNest/fln individual	Y	-0.015	0.495
	lme	Mass ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + Shannon + MinTemp + SDTemp + Jday + Sex + FinalBroodSize + Ectos	-	~1 fNest/fln individual	Y	-0.015	0.851
	gamm	Tarsavg ~ s(Age) + s(PD_whole_tree) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp) + factor(Ectos)	Tarsavg ~ s(Age) + s(PD_whole_tree) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp) + factor(Ectos)	-	fNest = ~ 1, fln individual = ~1	Y		0.650

	gamm	Tarsavg ~ s(Age) + s(Shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize))+ s(Jday) + s(MinTemp) + factor(Ectos)	Tarsavg ~ s(Age) + s(Shannon) + s(SDTemp) + factor(InitialBroodSiz e) + s(Jday) + s(MinTemp) + factor(Ectos)	-	fNest =~ 1, fIndividual =~1	Y		0.672
	gamm	Avg_9 ~ s(Age) + s(PD_whole_tree) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize))+ s(Jday) + s(MinTemp) + factor(Ectos)	Avg_9 ~ s(Age) + s(PD_whole_tree) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSiz e) + s(Jday) + s(MinTemp) + factor(Ectos)	-	fNest =~ 1	Y		0.875
	gamm	Avg_9 ~ s(Age) + s(shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize))+ s(Jday) + s(MinTemp) + factor(Ectos)	Avg_9 ~ s(Age) + s(shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSiz e) + s(Jday) + s(MinTemp) + factor(Ectos)	-	fNest =~ 1	Y		0.793
	lme	FA_T ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Age + InitialBroodSize + Ectos	-	~1 fNest/fIn dividual	N		

	lme	FA_T ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Age + FinalBroodSize + Ectos	-	~1 fNest/fIn dividual	N		
	gam	FA_9 ~ factor(Age) + s(PD_whole_tree) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp) + factor(Ectos)	FA_9 ~ factor(Age) + s(PD_whole_tree) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp) + factor(Ectos)	-	-	Y		0.184
	gamm	FA_9 ~ s(Age) + s(shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp) + factor(Ectos)	FA_9 ~ s(Age) + s(shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp) + factor(Ectos)	-	fNest = ~ 1	Y		0.859
	lme	GrowthPerformance ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + SDTemp + Jday	-	~1 fNest/fIn dividual	N		

	lme	GrowthPerformance ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + SDTemp	-	~1 fNest/fln dividual	N		
Cloacal	lme	Mass ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + PD + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	-	~1 Nest	Y	0.0046	0.918
	lme	Mass ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Sex + InitialBroodSize + FinalBroodSize + Ectos	-	~1 Nest	Y	0.0738	0.941
	gamm	Tarsavg ~ s(Age) + s(PD) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp)	Tarsavg ~ s(Age) + s(PD) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp)	-	Nest =~ 1	Y		0.005
	gamm	Tarsavg ~ s(Age) + s(Shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSize) + s(Jday) + s(MinTemp)	Tarsavg ~ s(Age) + s(Shannon) + factor(Year) + s(MaxTemp) + s(SDTemp) + factor(InitialBroodSiz e) + s(Jday) + s(MinTemp)	-	Nest =~ 1, Individual =~ 1	N		

lme	Avg_9 ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Avg_9 ~ Age + PD + MaxTemp + SDTemp + InitialBroodSize + Ectos	varPower(form=~Age)	~1 Nest	Y	-0.208	0.012
lme	Avg_9 ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Avg_9 ~ Age + Shannon + MaxTemp + Year + Jday + FinalBroodSize + Ectos	varPower(form=~Age)	~1 Nest	Y	-0.058	0.010
lme	FA_T ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	-	~1 Nest	N		
lme	FA_T ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Age + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	-	~1 Nest	N		
gls	FA_9 ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_9 ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + InitialBroodSize + FinalBroodSize	varExp(form =~Age)	-	N		

	gls	FA_9 ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_9 ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Jday + InitialBroodSize + FinalBroodSize + Ectos	varExp(form = ~Age)	-	N		
	lme	GrowthPerformance ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + Jday	-	~1 Nest	N		
	lme	GrowthPerformance ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + Jday	-	~1 Nest	N		
Fecal	gls	Mass ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + PD + MaxTemp + SDTemp + Year + Jday + FinalBroodSize	-	-	Y	-0.067	0.272
	gls	Mass ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + Shannon + MaxTemp + SDTemp + Year + Jday + FinalBroodSize	-	-	Y	-0.227	0.318

gam	Tarsavg ~ s(Age) + s(PD_whole_tree) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Tarsavg ~ s(Age) + s(PD_whole_tree) + factor(Ectos)	-	-	Y	0.091
gam	Tarsavg ~ s(Age) + s(Shannon) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Tarsavg ~ s(Age) + s(Shannon) + s(Jday) + factor(Ectos)	-	-	Y	0.653
gam	Avg_9 ~ s(Age) + s(PD_whole_tree) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Avg_9 ~ s(Age) + s(PD_whole_tree) + factor(Year) + factor(InitialBroodSiz e) + s(Jday) + factor(Ectos)	-	Nest =~ 1	Y	0.222
gam	Avg_9 ~ s(Age) + s(shannon) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Avg_9 ~ s(Age) + s(shannon) + factor(Year) + factor(InitialBroodSiz e) + s(Jday) + factor(Ectos)	-	Nest=~ 1	Y	0.160
lme	FA_T ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Year + Jday	-	~1 Nest	N	

	lme	FA_T ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Year + Jday	-	~1 Nest	N		
	gam	FA_9 ~ s(Age) + s(PD_whole_tree) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	FA_9 ~ s(Age) + s(PD_whole_tree) + factor(Year) + factor(InitialBroodSize) + s(Jday)	-	-	Y		0.059
	gam	FA_9 ~ s(Age) + s(Shannon) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	FA_9 ~ s(Age) + factor(InitialBroodSize)	-	-	N		
	lme	GrowthPerformance ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + Year + Jday	-	~1 Nest	N		
	lme	GrowthPerformance ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + Year + Jday	-	~1 Nest	N		

Oral	lme	Mass ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + PD + MinTemp + SDTemp	-	~1 Nest	Y	-0.056	0.092
	lme	Mass ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Mass ~ Age + Shannon + MinTemp + SDTemp + Jday + FinalBroodSize + Ectos	-	~1 Nest	Y	-0.157	0.297
	gamm	Tarsavg ~ s(Age) + s(PD_whole_tree) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Tarsavg ~ s(Age) + s(PD_whole_tree)	-	Nest =~ 1, Individual =~ 1	Y		0.136
	lme	Tarsavg ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	Tarsavg ~ Age + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + FinalBroodSize	varExp(form =~Age)	~1 Nest	N		
	gamm	Avg_9 ~ factor(Age) + s(PD) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Avg_9 ~ factor(Age) + s(PD_whole_tree) + factor(Year) + s(Jday) + factor(Ectos)	-	Nest =~ 1	Y		0.877

	gamm	Avg_9 ~ factor(Age) + s(shannon) + factor(Year) + factor(InitialBroodSize) + s(Jday) + factor(Ectos)	Avg_9 ~ factor(Age) + s(shannon) + factor(Year) + s(Jday) + factor(Ectos)	-	Nest =~ 1	Y		0.953
	gls	FA_T ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Age + InitialBroodSize + Ectos	varExp(form =~Nest)	-	N		
	gls	FA_T ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_T ~ Age + MaxTemp + MinTemp + Jday + InitialBroodSize + Ectos	varConstPo wer(form=~ Nest)	-	N		
	gls	FA_9 ~ Age + PD_whole_tree + MaxTemp + MinTemp + SDTemp + Year + Jday + InitialBroodSize + FinalBroodSize + Ectos	FA_9 ~ Age + MaxTemp + MinTemp + SDTemp + Jday + InitialBroodSize + FinalBroodSize	-	varFixed(~Ne st)	N		
	gls	FA_9 ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	FA_9 ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + InitialBroodSize + FinalBroodSize	-	varFixed(~Ne st)	Y	-0.0093	0.000

	lme	GrowthPerformance ~ Age + PD + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MaxTemp + MinTemp + SDTemp	-	~1 Nest	N		
	lme	GrowthPerformance ~ Age + Shannon + MaxTemp + MinTemp + SDTemp + Year + Jday + Sex + InitialBroodSize + FinalBroodSize + Ectos	GrowthPerformance ~ Age + MinTemp + SDTemp	-	~1 Nest	N		

Table S4.2. Bonferonni-corrected p-values of parametric (P) and nonparametric (NP) t-tests of whether unweighted UniFrac distances within levels of a given morphological metric are different than those between levels of the morphological metric. N/A indicates no variance between groups. Significant p-values ($\alpha = 0.05$) are highlighted in yellow.

SampleType	Mass		Tarsavg		Avg_9		FA_T		FA_9		GrowthPerformance	
	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP
All samples	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.001
All cloacal	0.000	0.001	0.002	0.003	0.000	0.001	0.001	0.002	0.000	0.001	0.001	0.001
2011 cloacal	0.414	0.426	0.414	0.415	0.201	0.211	0.414	0.440	0.001	0.001	0.414	0.421
2012 cloacal	0.016	0.018	0.009	0.006	0.001	0.003	0.009	0.018	0.000	0.001	0.009	0.004
2013 cloacal	0.073	0.093	0.073	0.084	0.301	0.314	0.073	0.070	0.093	0.089	0.073	0.077
All fecal	0.446	0.452	0.720	0.727	0.315	0.320	N/A	N/A	0.327	0.320	N/A	N/A
2011 fecal	N/A	N/A	N/A	N/A	0.513	0.499	N/A	N/A	0.513	0.513	N/A	N/A
2012 fecal	0.173	0.181	N/A	N/A	0.550	0.556	N/A	N/A	0.012	0.013	N/A	N/A
2013 fecal	N/A	N/A	N/A	N/A	0.927	0.930	N/A	N/A	0.171	0.180	N/A	N/A
All oral	0.135	0.128	0.262	0.264	0.000	0.001	N/A	N/A	0.008	0.011	N/A	N/A
2012 oral	0.440	0.525	0.487	0.559	0.023	0.023	N/A	N/A	0.973	0.976	N/A	N/A
2013 oral	N/A	N/A	N/A	N/A	0.133	0.129	N/A	N/A	0.001	0.002	N/A	N/A

Table S4.3. Summary of all partial Mantel tests run to analyze the relationship between each dependent variable and independent variable with all other independent variables alternately held as a control variable. Each specific combination of dependent and independent variable thereby had 12 partial Mantels run to evaluate the significance of the focal independent variable. Only independent variables with significant p-values ($\alpha = 0.05$) across all 12 tests are reported, whereby Cor indicates the range of Pearson’s correlations and P denotes the range of P-values observed across all 12 significant tests. NS indicates non-significance. Significant p-values and correlations for independent variables are highlighted in yellow.

125

Sample	Independent variable	Dependent_variable											
		Mass		Tarsavg		Avg_9		FA_T		FA_9		GrowthPerf	
		Cor	P	Cor	P	Cor	P	Cor	P	Cor	P	Cor	P
All	Age	0.730-0.741	0.001	0.663-0.674	0.001	0.828-0.833	0.001	NS	NS	NS	NS	0.570-0.584	0.001
	Ectos	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	FinalBroodSize	0.040-0.138	0.001-0.031	NS	NS	NS	NS	NS	NS	NS	NS	0.103-0.167	0.001-0.002
	InitialBroodSize	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	Jday	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	MaxTemp	0.088-0.186	0.001	0.176-0.239	0.001	0.106-0.200	0.001	NS	NS	NS	NS	NS	NS
	MeanTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	MedianTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	MinTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	SDTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	TempRange	NS	NS	NS	NS	0.080-0.210	0.001-0.002	NS	NS	NS	NS	NS	NS
	UnweightedUniFrac	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	Year	NS	NS	NS	NS	NS	NS	0.047-0.066	0.015-0.046	0.068-0.104	0.002-0.043	NS	NS
Cloacal	Age	0.723-0.736	0.001	0.658-0.670	0.001	0.820-0.825	0.001	NS	NS	NS	NS	0.555-0.571	0.001

	SDTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	TempRange	NS	NS	0.123- 0.257	0.001- 0.006	NS	NS	NS	NS	NS	NS	NS	NS	
	UnweightedUniFrac	0.091- 0.122	0.008- 0.018	NS	NS	NS	NS	NS	NS	NS	NS	0.157- 0.200	0.00 1- 0.00 7	
	Year	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
Oral	Age	0.738- 0.748	0.001	0.674- 0.686	0.001	0.834- 0.840	0.001	NS	NS	NS	NS	0.599- 0.609	0.00 1	
	Ectos	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	FinalBroodSize	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.089- 0.147	0.00 6- 0.03 8	
	InitialBroodSize	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	Jday	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	MaxTemp	NS	NS	0.133- 0.218	0.001- 0.005	0.112- 0.215	0.001- 0.002	NS	NS	NS	NS	NS	NS	NS
	MeanTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	MedianTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	MinTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	SDTemp	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	TempRange	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
	UnweightedUniFrac	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	
Year	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.044- 0.067	0.001- 0.034	NS	NS	

Table S4.4. Summary of all multiple matrix regressions with randomizations analyzing the relative significance of several spatiotemporal variables including gut microbiota in explaining variation of various nestling morphological metrics. The overall fit of each model is reported using the F statistic (Fstat) and its corresponding p-value (Fp). Parameters (Par) reported for each independent variable in each model include regression coefficients (coef), *t*-test statistics (tstat), and corresponding *t*-statistic p-values. Significant p-values for independent variables are highlighted in yellow.

Sample	Dependent	R ²	Fstat	Fp	Par	Int	Age	Jday	Year	Max Temp	Min Temp	Median Temp	Mean Temp	SD Temp	Temp Range	Ectos	Initial Brood	Final Brood	Unwt Uni Frac
All	Mass	0.557	2501.7	0.001	coef	-0.025	0.991	0.011	-0.283	0.167	0.065	-0.121	0.030	0.310	-0.127	0.026	-0.220	0.549	1.686
					tstat	-0.107	172.7	6.942	-5.717	9.978	3.871	-3.240	0.662	5.196	-9.035	4.255	-5.844	18.77	5.675
					tp	1.000	0.001	0.025	0.018	0.001	0.076	0.216	0.807	0.043	0.001	0.304	0.093	0.001	0.074
	Tarsa vg	0.475	1798.0	0.001	coef	0.009	0.482	-0.001	0.103	0.182	-0.026	-0.227	0.227	0.111	-0.025	-0.007	-0.053	0.068	0.654
					tstat	0.065	140.5	-1.150	3.484	18.27	-2.556	10.180	8.399	3.101	-2.965	-1.945	-2.346	3.859	3.683
					tp	1.000	0.001	0.758	0.225	0.001	0.349	0.002	0.016	0.285	0.236	0.713	0.568	0.319	0.361
	Avg_9	0.705	4755.7	0.001	ceff	6.353	2.653	0.034	-1.242	-0.090	-0.409	-0.927	1.478	-0.870	0.296	0.036	0.031	-0.130	-4.962
					tstat	13.61	234.7	10.76	-12.73	-2.750	-12.44	12.638	16.59	-7.399	10.68	2.946	0.414	-2.257	-8.481
					tp	0.918	0.001	0.001	0.001	0.267	0.001	0.001	0.001	0.001	0.006	0.001	0.536	0.910	0.504
	FA_T	0.023	47.3	0.095	coef	0.033	0.000	0.000	0.004	0.000	0.000	0.001	0.000	-0.002	0.001	0.000	0.001	-0.001	-0.014
					tstat	23.41	4.619	-11.24	13.83	-0.666	1.554	2.672	-0.513	-4.954	10.93	-10.04	5.546	-8.009	-7.812
					tp	0.146	0.150	0.038	0.003	0.870	0.664	0.526	0.891	0.211	0.001	0.138	0.327	0.127	0.173
	FA_9	0.058	121.9	0.042	coef	0.052	-0.002	0.000	0.006	0.003	0.001	-0.012	0.010	0.000	-0.003	-0.002	0.006	0.004	-0.002
					tstat	10.88	-16.80	4.305	6.443	9.923	2.265	16.059	11.25	-0.375	-10.08	-18.39	8.228	6.980	-0.357
					tp	0.495	0.001	0.405	0.144	0.060	0.635	0.004	0.047	0.947	0.022	0.057	0.277	0.317	0.965
Growth Performance	0.363	1136.1	0.001	coef	0.009	0.024	0.001	-0.028	-0.001	0.003	0.002	-0.001	0.003	-0.005	0.001	-0.019	0.035	0.095	
				tstat	1.014	115.4	10.89	-15.42	-1.004	4.295	1.584	-0.380	1.458	-9.306	4.453	-13.34	32.38	8.655	
				tp	1.000	0.001	0.017	0.001	0.758	0.135	0.634	0.910	0.649	0.001	0.452	0.004	0.001	0.036	
Cloacal	Mass	0.558	413.9	0.001	Coef	1.330	0.985	0.007	-0.159	0.295	0.074	-0.420	0.308	0.198	-0.153	0.072	-0.303	0.492	-0.117
					Tstat	2.494	68.47	1.849	-1.338	7.816	1.872	-4.388	2.713	1.363	-4.710	4.526	-3.248	6.596	-0.168

	Tarsa vg	0.486	310.2	0.001	Tp	0.993	0.001	0.413	0.452	0.002	0.263	0.033	0.179	0.467	0.007	0.119	0.189	0.008	0.946	
					coef	0.180	0.493	-0.004	0.108	0.260	-0.024	-0.402	0.424	0.145	-0.044	0.020	-0.100	0.034	0.238	
					tstat	0.552	55.95	-1.898	1.480	11.24	-0.992	-6.866	6.092	1.628	-2.200	2.103	-1.740	0.742	0.560	
	Avg_9	0.691	734.4	0.001	Tp	0.994	0.001	0.470	0.458	0.001	0.602	0.006	0.011	0.449	0.201	0.645	0.554	0.817	0.837	
					coef	3.821	2.599	0.022	-1.080	-0.116	-0.486	-0.684	1.197	-0.953	0.390	0.106	0.283	-0.168	-1.389	
					tstat	3.625	91.39	2.967	-4.606	-1.558	-6.244	-3.619	5.327	-3.324	6.090	3.390	1.532	-1.139	-1.011	
	FA_T	0.023	7.8	0.676	Tp	0.969	0.001	0.118	0.005	0.330	0.001	0.054	0.006	0.050	0.001	0.180	0.469	0.571	0.641	
					coef	0.023	0.000	0.000	0.004	0.000	0.000	0.001	0.000	-0.002	0.000	0.000	0.001	-0.001	0.000	
					tstat	7.058	2.995	-5.428	5.008	0.958	2.052	1.505	-0.540	-2.856	2.384	-3.918	1.536	-2.577	0.084	
	FA_9	0.078	27.6	0.161	Tp	0.630	0.175	0.070	0.037	0.732	0.384	0.604	0.864	0.296	0.245	0.339	0.659	0.476	0.979	
					coef	0.023	-0.002	0.000	0.006	0.004	0.001	-0.012	0.009	-0.004	-0.002	-0.002	0.004	0.008	0.035	
					tstat	2.248	-8.877	-0.650	2.740	5.688	0.946	-6.354	4.353	-1.481	-3.144	-6.547	2.481	5.378	2.652	
	Growt h Perfor manc e	0.349	176.1	0.001	Tp	0.817	0.004	0.819	0.286	0.120	0.761	0.105	0.262	0.678	0.263	0.124	0.576	0.198	0.530	
					coef	0.109	0.024	0.001	-0.025	0.003	0.003	-0.009	0.008	-0.003	-0.005	0.002	-0.024	0.034	-0.030	
					tstat	5.539	44.58	3.873	-5.674	1.955	2.215	-2.496	1.959	-0.481	-4.239	3.683	-7.030	12.48	-1.147	
	fecal	Mass	0.549	137.9	0.001	Tp	0.874	0.001	0.120	0.007	0.392	0.267	0.270	0.416	0.839	0.022	0.277	0.029	0.001	0.667
						coef	-1.030	0.946	0.026	-0.670	0.046	0.095	0.122	-0.254	0.395	-0.116	0.321	-0.473	0.781	3.092
						tstat	-1.088	39.41	2.711	-2.642	0.635	1.346	0.818	-1.384	1.608	-1.987	5.772	-2.978	5.802	2.601
		Tarsa vg	0.459	96.1	0.001	Tp	0.991	0.001	0.033	0.029	0.590	0.232	0.530	0.285	0.176	0.088	0.005	0.060	0.003	0.068
						coef	0.141	0.449	0.030	-0.362	0.090	-0.002	-0.092	0.009	-0.018	0.016	0.044	0.022	0.071	0.689
						tstat	0.254	31.83	5.288	-2.431	2.095	-0.046	-1.049	0.085	-0.123	0.459	1.333	0.237	0.903	0.987
Avg_9		0.703	268.0	0.001	Tp	0.996	0.001	0.005	0.063	0.135	0.972	0.513	0.961	0.940	0.721	0.572	0.901	0.572	0.572	
					coef	8.283	2.774	0.054	-2.433	-0.058	-0.443	-0.994	1.605	-1.587	0.368	0.143	-0.274	0.323	-8.589	
					tstat	4.128	54.50	2.607	-4.529	-0.372	-2.947	-3.142	4.133	-3.049	2.981	1.213	-0.815	1.133	-3.409	
FA_T		0.032	3.8	0.785	Tp	0.544	0.001	0.127	0.004	0.823	0.065	0.085	0.044	0.059	0.051	0.690	0.712	0.536	0.091	
					coef	0.017	0.000	0.000	0.004	0.000	0.001	0.001	-0.001	-0.003	0.001	-0.001	0.002	-0.002	0.008	
					tstat	2.968	0.120	-0.980	2.401	-0.733	1.379	0.699	-0.713	-2.319	3.580	-2.568	2.026	-1.939	1.159	
FA_9		0.099	12.4	0.234	coef	0.770	0.947	0.648	0.169	0.727	0.444	0.775	0.777	0.261	0.047	0.398	0.459	0.420	0.662	
						coef	-0.004	-0.001	0.000	0.019	0.000	-0.002	-0.015	0.017	-0.002	-0.003	-0.006	0.024	-0.007	0.080

					tstat	-0.180	-2.413	1.510	3.059	0.260	-1.075	-4.280	3.936	-0.385	-1.960	-4.502	6.306	-2.287	2.821
					tp	0.973	0.197	0.568	0.143	0.916	0.623	0.141	0.200	0.865	0.340	0.206	0.096	0.415	0.421
Growth Performance	0.386	71.2	0.001	coef	-0.098	0.022	0.000	-0.027	-0.004	0.002	0.010	-0.005	0.012	-0.006	0.013	-0.034	0.047	0.237	
				tstat	-2.815	25.52	0.976	-2.937	-1.334	0.586	1.788	-0.751	1.373	-2.729	6.335	-5.887	9.595	5.433	
					tp	0.844	0.001	0.614	0.039	0.424	0.692	0.298	0.688	0.408	0.065	0.034	0.013	0.001	0.014
Mass	0.565	314.5	0.001	coef	0.427	1.003	0.010	-0.195	0.035	0.043	0.043	0.029	0.332	-0.065	-0.034	0.073	0.438	0.926	
				tstat	0.631	61.83	2.096	-1.316	0.646	0.853	0.417	0.220	1.881	-1.432	-2.252	0.674	5.475	1.054	
					tp	0.992	0.001	0.254	0.260	0.710	0.524	0.805	0.903	0.248	0.347	0.390	0.721	0.014	0.630
Tarsavg	0.485	228.0	0.001	coef	-0.034	0.484	-0.005	0.085	0.101	-0.038	-0.093	0.183	0.035	0.023	-0.026	0.026	0.116	0.653	
				tstat	-0.086	50.77	-1.912	0.984	3.193	-1.263	-1.525	2.368	0.341	0.873	-2.992	0.412	2.470	1.265	
					tp	0.998	0.001	0.394	0.454	0.089	0.380	0.459	0.257	0.866	0.585	0.322	0.828	0.265	0.639
Avg_9	0.717	612.4	0.001	coef	0.459	2.615	0.047	-0.404	-0.038	-0.198	-1.121	1.726	-0.748	0.066	-0.016	-0.291	-0.149	2.735	
				tstat	0.357	84.79	5.403	-1.438	-0.365	-2.048	-5.656	6.919	-2.231	0.767	-0.559	-1.406	-0.982	1.637	
					tp	0.997	0.001	0.003	0.192	0.800	0.126	0.004	0.002	0.122	0.570	0.823	0.395	0.587	0.434
FA_T	0.022	5.4	0.782	coef	0.031	0.000	0.000	0.001	0.000	-0.001	0.000	0.000	-0.002	0.001	0.000	0.001	-0.001	-0.010	
				tstat	8.381	0.099	-3.870	1.489	-0.189	-2.653	0.042	0.146	-1.825	5.111	-3.339	1.860	-3.138	-1.979	
					tp	0.295	0.956	0.172	0.327	0.944	0.225	0.991	0.940	0.503	0.025	0.379	0.536	0.306	0.590
FA_9	0.057	14.6	0.313	coef	0.091	-0.002	0.000	0.001	0.003	0.000	-0.011	0.007	0.004	-0.004	-0.002	0.001	0.005	-0.045	
				tstat	7.086	-6.314	3.704	0.442	3.177	0.164	-5.797	2.784	1.298	-4.703	-6.222	0.562	3.361	-2.733	
					tp	0.266	0.027	0.179	0.773	0.323	0.936	0.101	0.415	0.666	0.101	0.147	0.873	0.332	0.550
Growth Performance	0.390	154.6	0.001	coef	0.050	0.025	0.001	-0.019	-0.005	0.002	0.008	-0.001	0.007	-0.004	-0.001	-0.002	0.025	0.033	
				tstat	2.048	43.36	4.822	-3.495	-2.615	1.134	2.041	-0.166	1.049	-2.489	-2.366	-0.419	8.540	1.044	
					tp	0.963	0.001	0.049	0.031	0.189	0.470	0.360	0.944	0.616	0.147	0.484	0.860	0.003	0.715

Figures

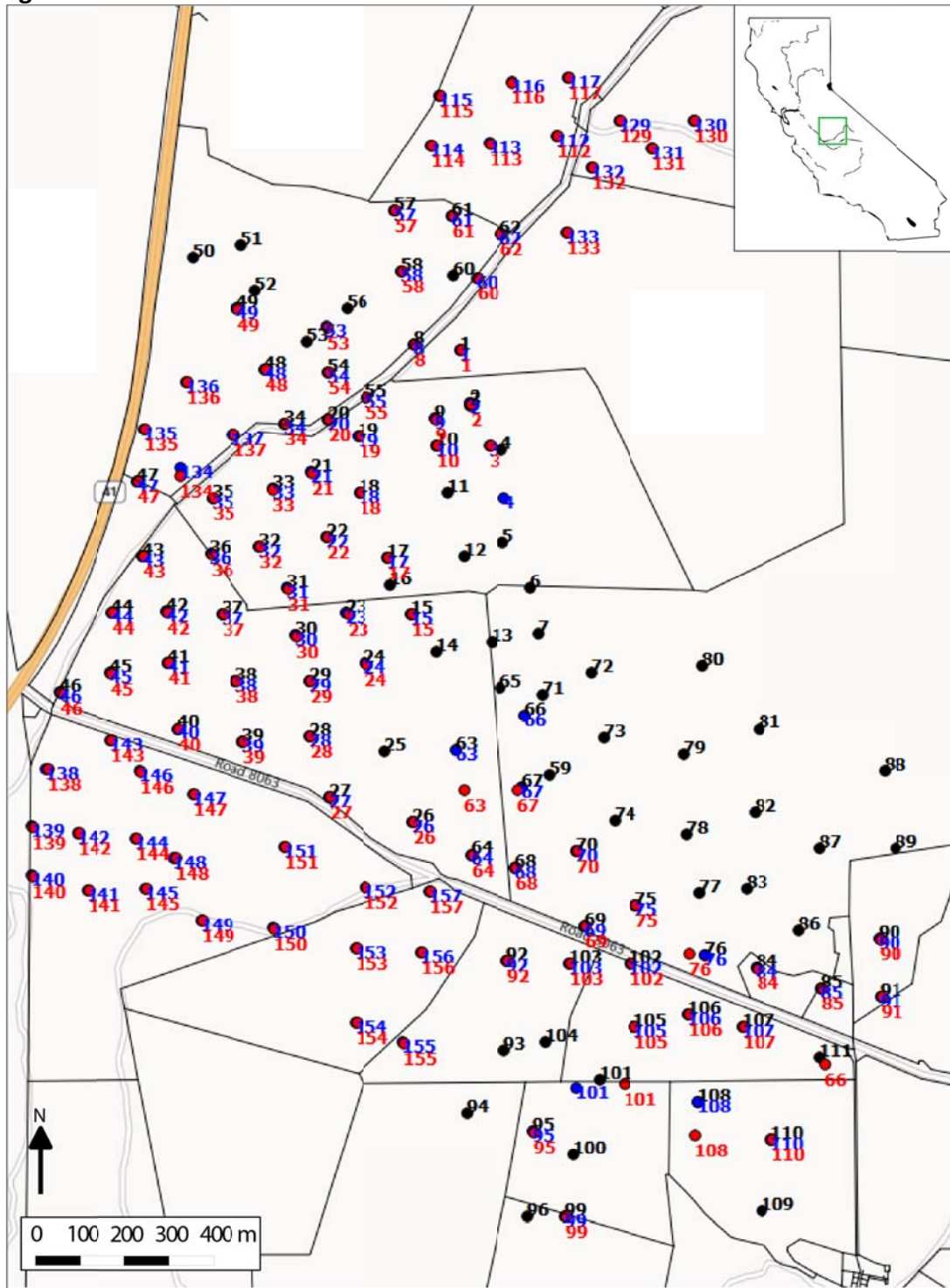


Fig. S4.1. Map of the Western Bluebird nest box study system at the San Joaquin Experimental Range (Madera County, CA) used in this study from 2011 through 2013. Black, blue, and red circles and fonts indicate boxes erected in 2011, 2012, and 2013, respectively. Lines depict fences.

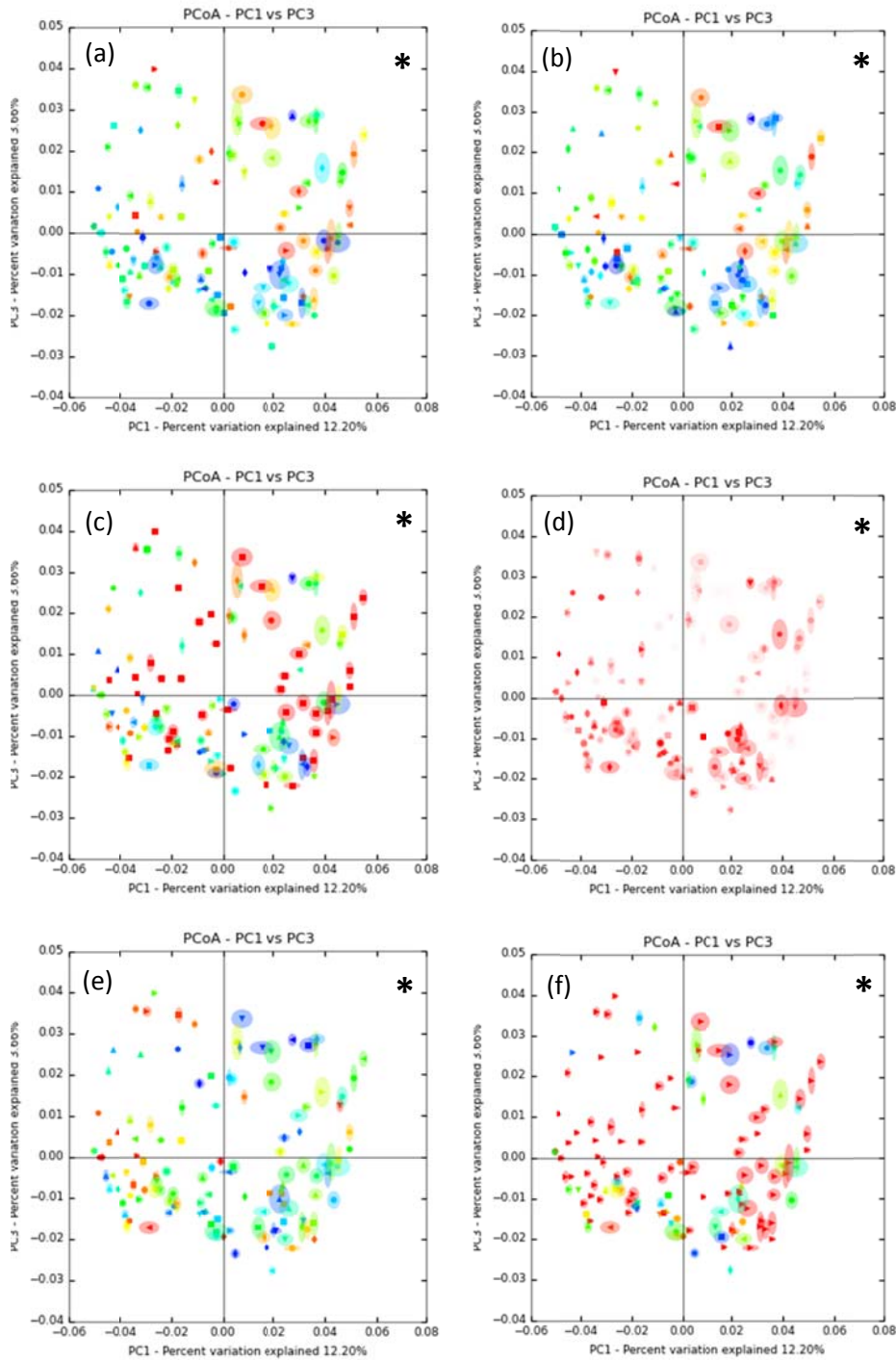
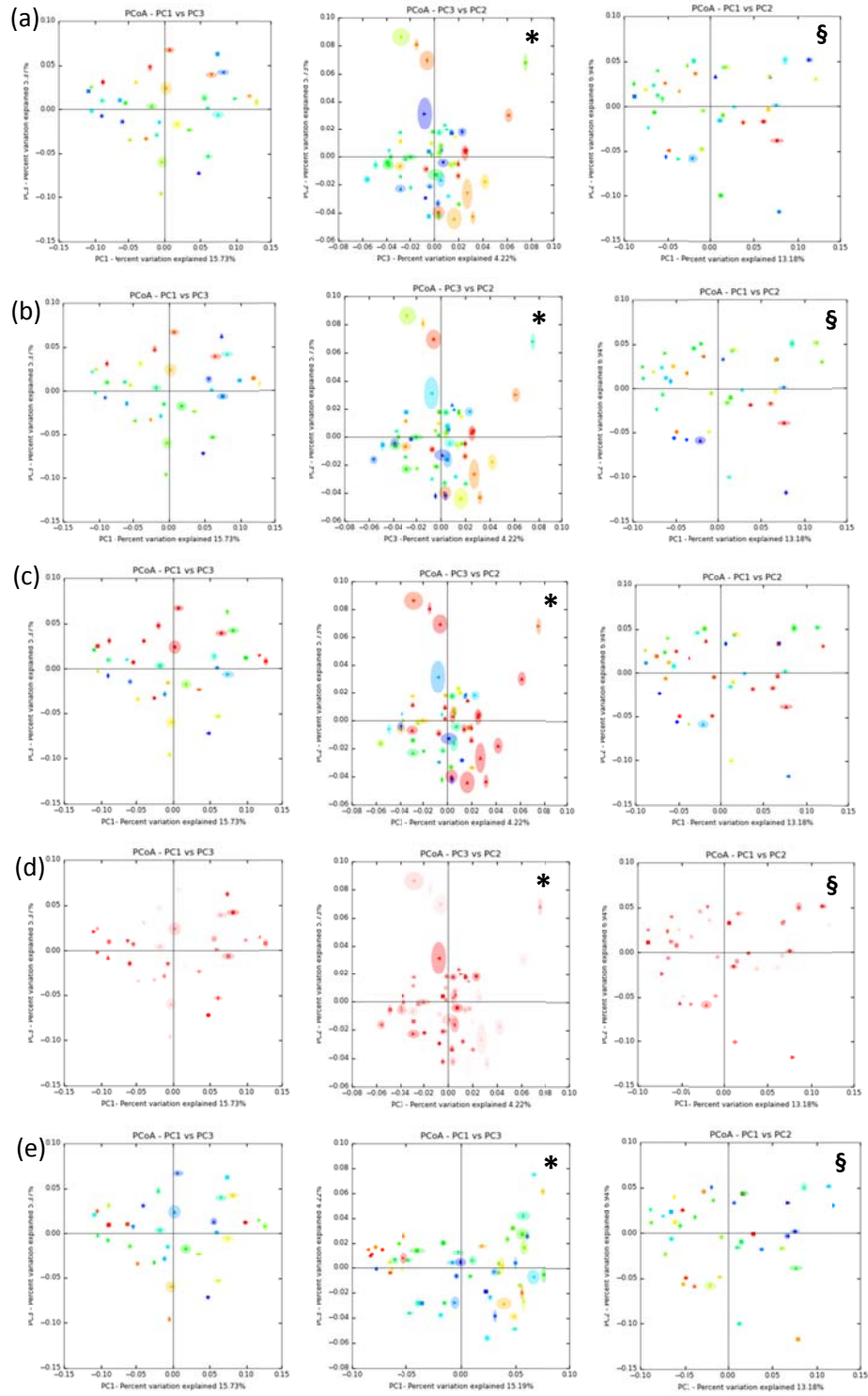


Fig. S4.2. Principal coordinate analysis of unweighted UniFrac for cloacal samples ($n = 136$) colored by mass (a), average tarsus length (b), average ninth primary length (c), growth performance (d), tarsus asymmetry (e), and ninth primary asymmetry (f). Asterisk indicates significant p -values in both parametric and nonparametric t -tests comparing UniFrac within versus between levels of host condition.



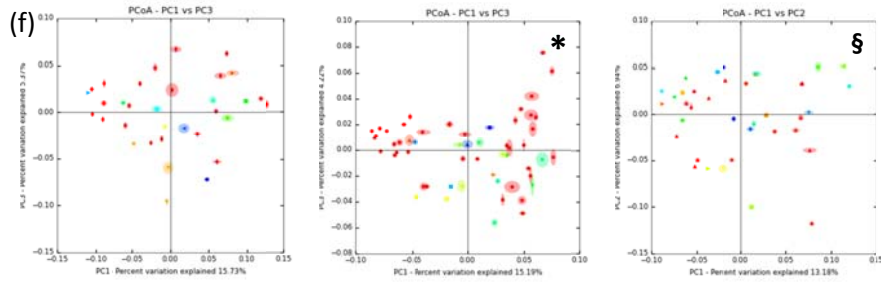


Fig. S4.3. Principal coordinate analysis of unweighted UniFrac for cloacal samples colored by mass (a), average tarsus length (b), average ninth primary length (c), growth performance (d), tarsus asymmetry (e), and ninth primary asymmetry (f) for the 2011 (first column, $n = 34$), 2012 (middle column, $n = 60$), and 2013 (right column, $n = 42$) sampling years. Color gradient shows lower values with cooler colors and higher values in warmer colors. * and § indicate significant and nearly significant p -values, respectively, in both parametric and nonparametric t -tests comparing UniFrac within versus between levels of host condition.

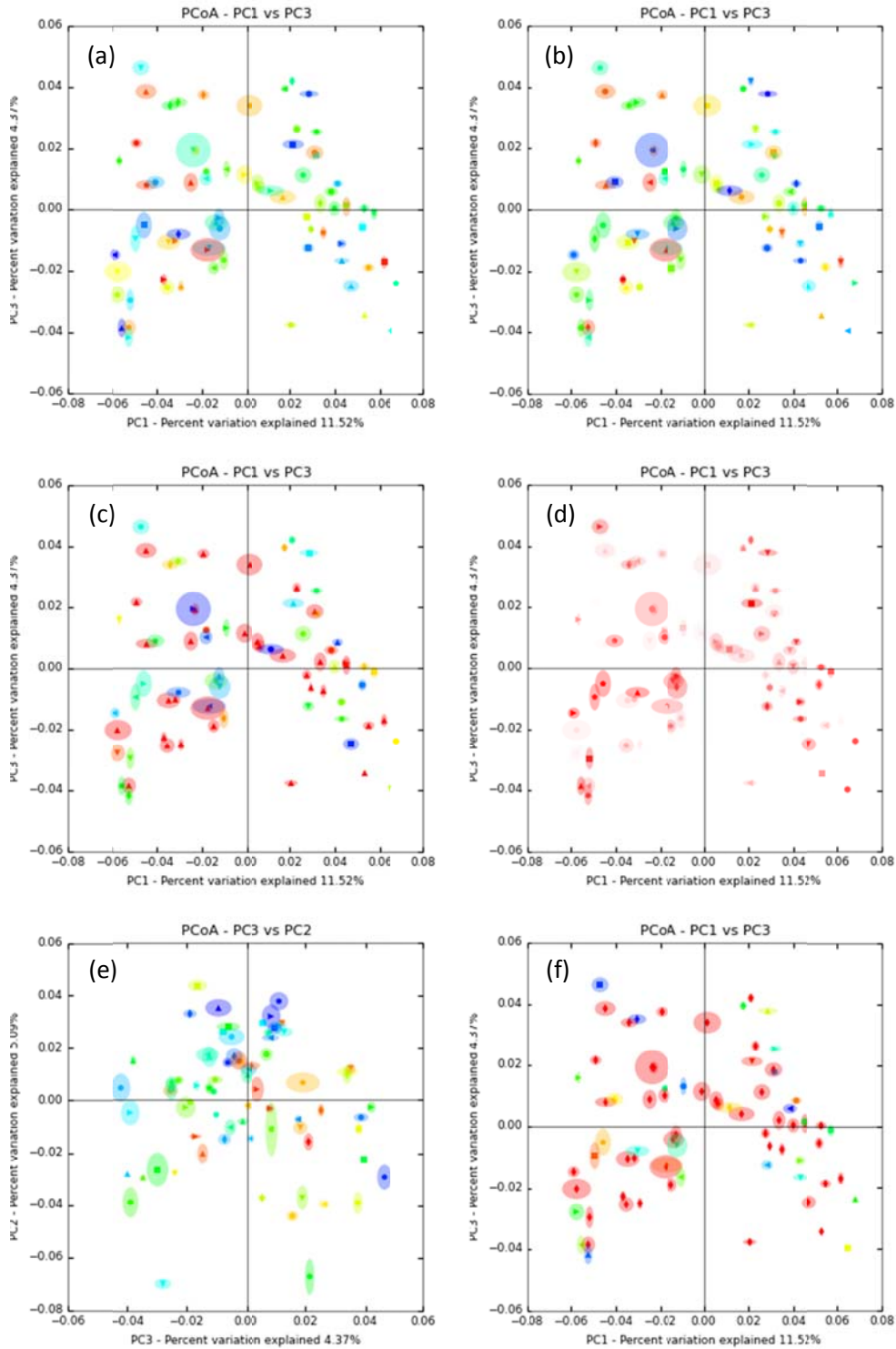
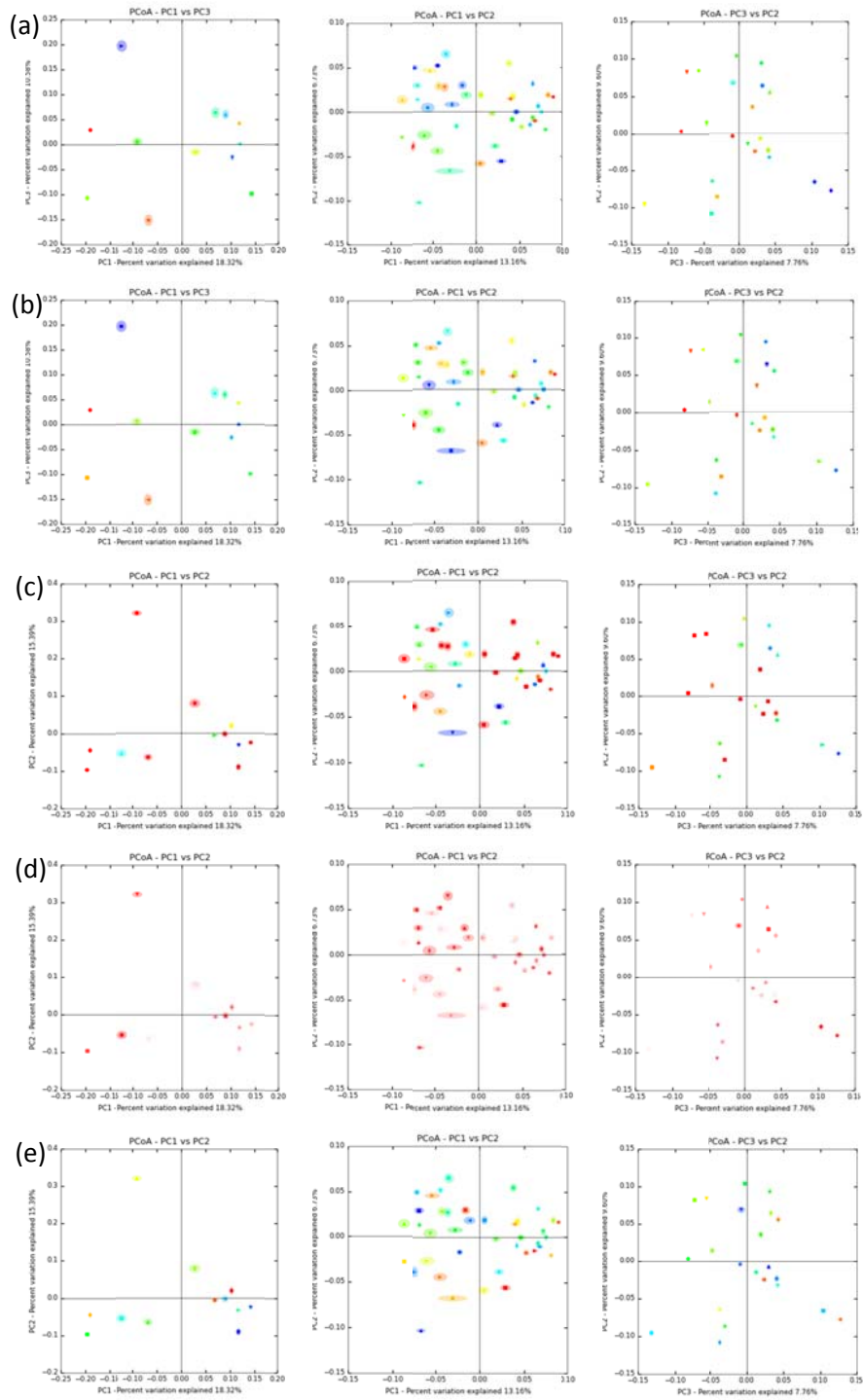


Fig. S4.4. Principal coordinate analysis of unweighted UniFrac for fecal samples ($n = 90$) colored by mass (a), average tarsus length (b), average ninth primary length (c), growth performance (d), tarsus asymmetry (e), and ninth primary asymmetry (f).



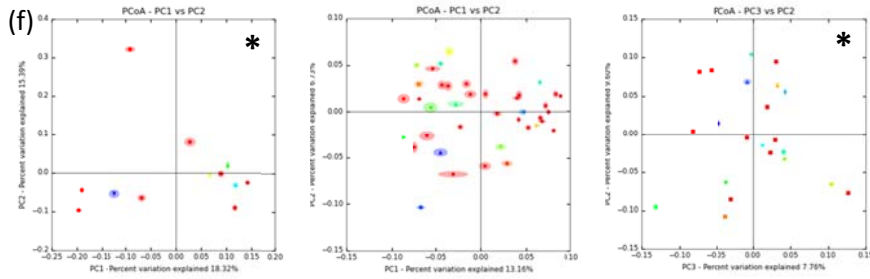


Fig. S4.5. Principal coordinate analysis of unweighted UniFrac for fecal samples colored by mass (a), average tarsus length (b), average ninth primary length (c), growth performance (d), tarsus asymmetry (e), and ninth primary asymmetry (f) for the 2011 (first column, $n = 12$), 2012 (middle column, $n = 46$), and 2013 (right column, $n = 32$) sampling years. Color gradient shows lower values with cooler colors and higher values in warmer colors. Asterisk indicates significant p -values in both parametric and nonparametric t -tests comparing UniFrac within versus between levels of host condition.

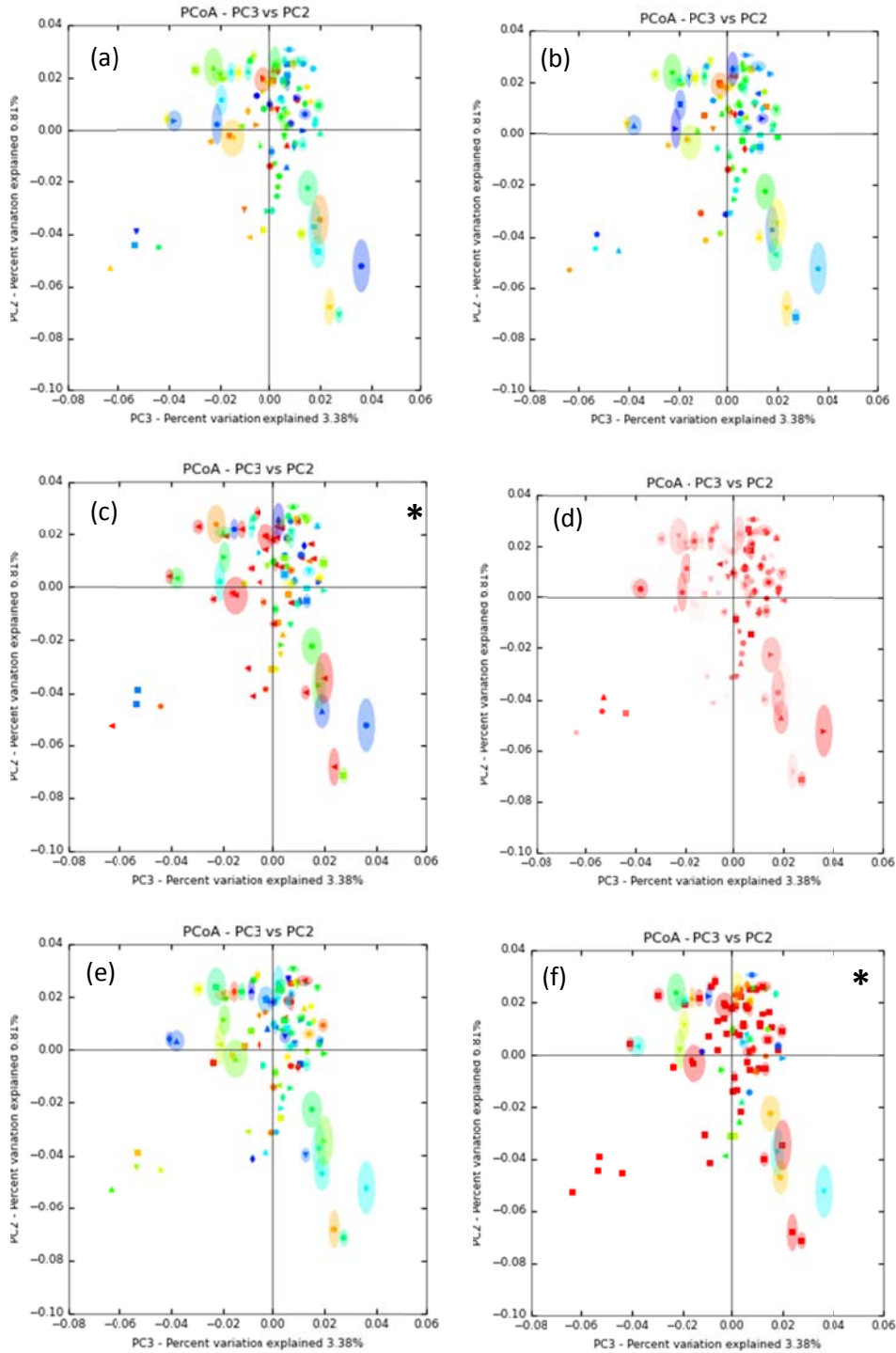
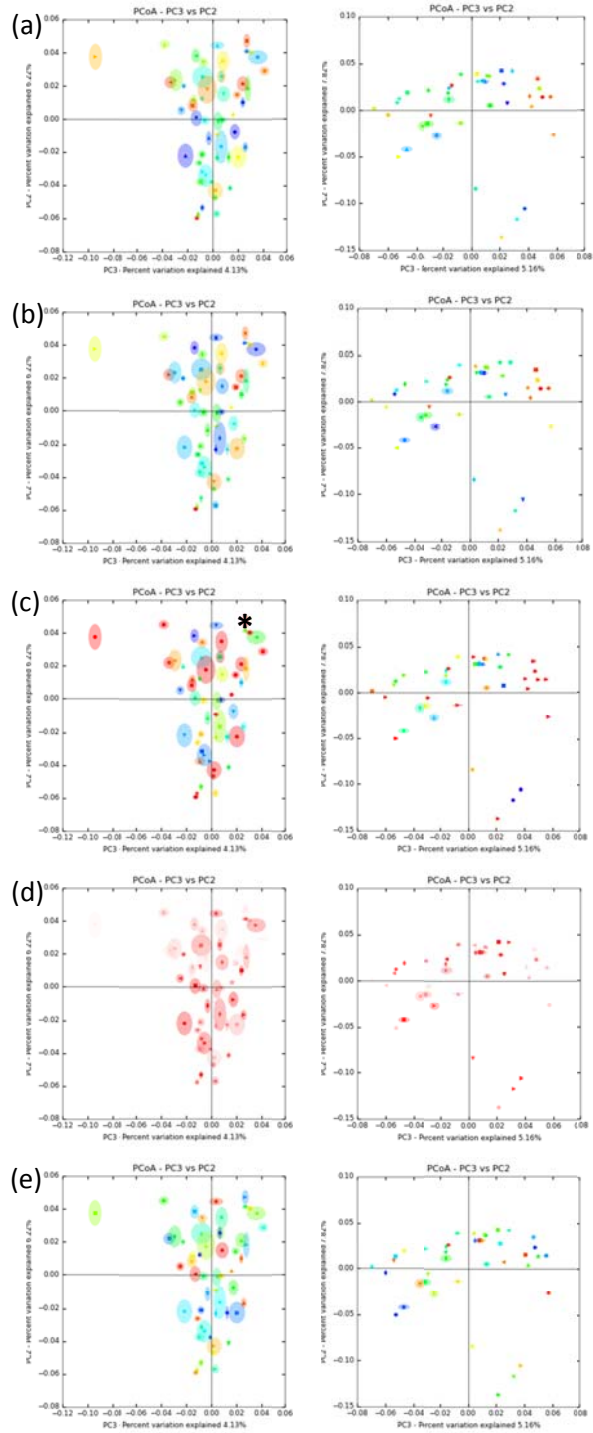


Fig. S4.6. Principal coordinate analysis of unweighted UniFrac for oral samples ($n = 99$) colored by mass (a), average tarsus length (b), average ninth primary length (c), growth performance (d), tarsus asymmetry (e), and ninth primary asymmetry (f). Asterisk indicates significant p -values in both parametric and nonparametric t -tests comparing UniFrac within versus between levels of host condition.



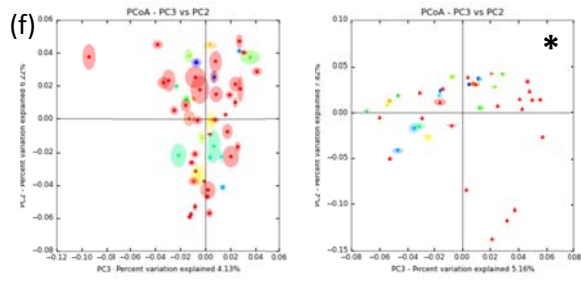


Fig. S4.7. Principal coordinate analysis of unweighted UniFrac for oral samples colored by mass (a), average tarsus length (b), average ninth primary length (c), growth performance (d), tarsus asymmetry (e), and ninth primary asymmetry (f) for the 2012 (middle column, $n = 57$) and 2013 (right column, $n = 42$) sampling years. Color gradient shows lower values with cooler colors and higher values in warmer colors. Asterisk indicates significant p -values in both parametric and nonparametric t -tests comparing UniFrac within versus between levels of host condition.

Chapter 5: Impacts of an immune challenge on the gut microbiota and growth performance of neonatal hosts: insights from an avian model system

Abstract

This study investigated the potential for immunological challenges to impact the gut microbiota of vertebrate organisms. Given recent documentation of the widespread role of gut bacteria in host health and disease, the focus of this study is especially warranted for early life stages during which underdeveloped immune systems, pathogen exposure, and gut microbial assembly can all affect host condition. In this research, an inactivated Newcastle Disease virus vaccine was employed in a tractable avian model system of Western Bluebirds *Sialia mexicana* at the San Joaquin Experimental Range (Madera County, CA), whereby different nestlings within each nest were assigned to either vaccine or control treatments at five days post-hatching. Illumina sequencing of bacterial 16S rRNA amplicons was used to characterize bacterial assemblages in the oral cavity, fecal matter, and cloaca of the nestlings, and comparisons were made between experimental groups before, five days after, and 10 days after vaccination. Significant differences in the alpha and beta diversity of microbiota across all three gut sites were consistent with the hypothesis that use of a viral vaccine can incur shifts in the composition of gut bacteria at both local and systemic levels. Future research encompasses understanding the mechanisms underlying the observed results as well as informing the extent to which the gut microbial shifts may impact host fitness.

Introduction

Microbiota are increasingly recognized as pervasive symbionts in animal hosts (Ley *et al.* 2008; McFall-Ngai *et al.* 2013), and research to date has demonstrated the widespread role that gut microbiota play in the physiology of host organisms, including metabolism (Tremaroli & Bäckhed 2012), gut development (Sommer & Bäckhed 2013), adiposity (Turnbaugh *et al.* 2009), behavior (Heijtz *et al.* 2011), and immune function (Hooper *et al.* 2012). Interactions between the host immune system and the gut microbiota are especially dynamic, as a fine balance is maintained between immunological ignorance of the enormous numbers of bacteria in the gut lumen and immunological responsiveness towards bacteria that transgress host tissues (Hooper *et al.* 2012). To achieve this balance, several regulatory mechanisms, including physical mucus barriers (Ouwkerk *et al.* 2013), general antimicrobial proteins (Salzman *et al.* 2010), and specific antibodies (Macpherson & Uhr 2004), shape the localization and composition of gut microbial assemblages (see Supporting Information for more details). Given that the structure and function of the gut microbiota are intimately linked with host physiology and phenotypic diversity (Ley *et al.* 2008; Ridaura *et al.* 2013), a greater understanding of the extent to which immunological perturbations may impact the gut microbiota is warranted to derive implications for host condition. Yet despite the extensive interconnections in the relationship between the host immune system and gut microbiota, surprisingly no study to date has investigated how activation of a host immune response to an antigenic threat (i.e., an immune challenge) may affect gut microbial ecology.

Immune challenges, such as pathogen infections or allergen exposure, can frequently occur throughout the lifetime of organisms. However, the neonatal host life stage is a particularly interesting developmental period during which the vertebrate immune system is

underdeveloped (Barrios *et al.* 1996) and thus susceptible to microbial infections. At the same time, early life events can impact assembly of the microbiota and thereby incur life-long effects on host metabolic programming (Harmsen *et al.* 2000; Dominguez-Bello *et al.* 2010; Koenig *et al.* 2011). Furthermore, ecoimmunological theory predicts that under finite resources, trade-offs occur between immune function and somatic growth within individuals (Lee 2006). Therefore, investigating the potential for an early life immune challenge to alter host gut microbiota comprises a valuable opportunity to clarify the role of microbiota as intermediaries of the outcome of pathogen exposure, immunological responsiveness, and host development.

Using a viral vaccine in a free-living population of nestling birds, this research generates ecologically relevant insights into natural variation in gut bacteria after immunological challenge. Massively parallel sequencing (Schuster 2008) is applied to characterize gut bacteria throughout the nestling period, and morphological metrics are obtained to assess nestling growth performance. This research addresses the hypothesis that vaccine-administered viral antigens produce a specific immune response (Barrios *et al.* 1996) that skews the subsets and function of immune cells in the host, with the prediction that differences in assembly and composition of the gut microbiota are detectable between vaccinated and unvaccinated individuals. Evaluation of the effects of a viral challenge on the bacterial microbiota is of specific interest given that many pathogens of public concern are viruses (Reed *et al.* 2003; CDC 2014) and that bacteria play an integral role in host metabolism across all life stages (Nicholson *et al.* 2012). Thus, this study integrates ecoimmunology and microbial ecology to uniquely address a gap in current knowledge on immunological mediation of gut bacteria, with important implications for disease ecology as well as veterinary and human medicine.

Materials and methods

Study site and system

A system of over 100 nest boxes was established at the San Joaquin Experimental Range (Madera County, CA) to support this study (Fig. S5.1). The Western Bluebird *Sialia mexicana* inhabits oak woodlands and is a frequent nest box occupant during its breeding season from March through July in California. Western Bluebirds lay three to eight eggs, with a mean of five eggs, per clutch. Incubation periods range from 12 to 14 days. Nestlings fledge from the nest after 18 to 21 days post-hatching.

Experimental design, sampling, and Illumina sequencing of bacterial 16S rRNA amplicons

In 2013, the boxes were closely monitored throughout the breeding season to locate active nests, ascertain hatch dates, and thus determine specific ages for each nestling. Superciliary feather clipping was used to track nestling age within asynchronously hatched nests. At five days post-hatching, oral swab, cloacal swab, and fresh fecal samples were aseptically collected from each nestling following Vo and Jedlicka (2014). Briefly, fecal samples were stored in a lab-made ammonium-sulfate saturated nucleic acid stabilization and storage buffer (25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2; Vo & Jedlicka 2014), not exceeding a sample to buffer ratio of 0.15 g/mL. Cloacal and oral swab heads (Puritan Medical PurFlock 25-3316 1PN) were stored in 1 mL of DNA buffer (25 mM sodium citrate, 10 mM EDTA, 70 g ammonium sulfate/100 mL solution, pH 5.2). All samples were maintained at -20°C in the field and at -80°C long-term. Morphological measurements were also obtained for each

nestling. The length of both the right and left tarsus was measured in duplicate using a digital caliper (± 0.03 mm). The length of both ninth primary feathers was determined using a stopped ruler (± 0.5 mm), and nestling mass was recorded using an electronic balance (± 0.01 g). A uniquely numbered aluminum leg band (USGS Bird Banding Laboratory) was attached to each nestling for subsequent identification.

Half of the nestlings in each nest were assigned to a challenge group ($n = 15$ individuals total) whereas the other half was assigned to a control group ($n = 14$ individuals total), which controlled for differences between nests. For the immune challenge, 75 μ L of an inactivated Newcastle Disease virus vaccine (AviPro 105 ND, Lohmann Animal Health International) was administered intraperitoneally with sterile, single use 1CC syringes (Small Parts SYR-001CC-10) and 22 gauge needles (BD 305155). The vaccination site was cleansed with 70% isopropanol before injection. This specific vaccine and dosage is safe for administration to neonatal Western Bluebirds in the United States and produces a robust antibody response in Western Bluebird nestlings after a single vaccination (Fair & Myers 2002; Fair *et al.* 2003). Newcastle Disease virus is not found in natural Western Bluebird populations (Pearson & McCann 1975), and therefore, exposure history and maternal transmission do not confound the effects of vaccine use as an immune challenge. Microbial samples and morphological measurements were again obtained from each nestling at 10 and 15 days post-hatching. At 10 days post-hatching, an additional unique combination of three colored leg bands was attached to each nestling to facilitate identification among individuals. All protocols were approved by the UC Berkeley Animal Care and Use Committee (MAUP R317-0913), California Department of Fish and Game (Permit SC-11869), and USGS Bird Banding Laboratory (Permit 21859).

Following Vo and Jedlicka (2014), metagenomic DNA was extracted from all fecal and swab samples using solid phase reversible immobilization beads. Subsequently, metagenomic DNA extracts were amplified using degenerate, barcoded primers targeting the bacterial 16S rRNA locus (515F and 806R; Caporaso *et al.* 2012). Illumina amplicon libraries were prepared using a combinatorial tagging approach with a “with-bead” (Fisher *et al.* 2011) and PCR-free (Kozarewa *et al.* 2009) workflow. The final pool of libraries was submitted to the University of California Davis Genome Center for a separate 250 bp, paired-end run on the Illumina MiSeq platform (see Supporting Information for detailed description of laboratory methods).

Data analyses

Illumina reads were trimmed of adapters and demultiplexed by adapter index into the original libraries using CASAVA 1.8. After removing reads that failed to pass filter, Trimmomatic was used to trim the trailing edges of sequences to a minimum quality score of Q20 while keeping paired-end reads in sync (Lohse *et al.* 2012). Following Bystrykh (2012), barcodes were error-corrected, and reads with barcodes containing more than one error were discarded. Paired-end reads were merged using PANDAseq with a quality threshold of 0.6 (Masella *et al.* 2012). Merged reads lacking either matching barcodes at both ends or correct primer sequences were discarded. Using QIIME 1.8.0, each library was demultiplexed by barcode, and all barcode and primer sequences were trimmed from each read before downstream analyses (Caporaso *et al.* 2010). Through the usearch quality filtering (usearch_qf) protocol, the Greengenes v. 13_8 database and clustering at 97% identity were applied to remove chimeras (de novo and reference-based), pick operational taxonomic units (OTUs), and make OTU taxonomy

assignments. FastTree 2.1.3 was implemented in QIIME to build a maximum-likelihood phylogenetic tree containing all OTUs (Price *et al.* 2010).

To derive alpha diversity metrics for each sample, rarefaction analysis was performed to generate phylogenetic distance (Faith & Baker 2007) and Shannon's diversity index at various sequencing coverage depths. For beta diversity metrics, 16S rRNA OTU relative abundance tables were first rarefied to 1,000 sequences per sample, the sequencing depth at which phylogenetic distance and Shannon's diversity index plateaued for generally all samples (Fig. S5.2), to standardize comparisons between samples. Unweighted UniFrac (Lozupone & Knight 2005) matrices, which excludes OTU abundance in the calculation of distances between communities, were generated from the rarefied OTU tables. To prepare morphological data for downstream analyses, mean tarsus length was calculated from all replicate measurements for both right and left legs per sample date. Mass normalized by mean tarsus was used as a growth performance index. Fluctuating asymmetry (Lens *et al.* 2002) of the ninth primaries was derived following $A = (R_i - L_i) / ((R_i + L_i) / 2)$, where R_i and L_i denote right and left side measurements, respectively.

To analyze the effect of experimental treatment on bacterial diversity as well as morphological distances, principal coordinate analysis (PCoA), parametric *t*-tests, and nonparametric *t*-tests with 1000 Monte Carlo randomizations were performed for each sample type and nestling age. Likewise, QIIME was used to compare taxonomic compositions of microbiota between experimental treatments for each sample type and age, wherein G-tests were applied to detect OTUs that differed significantly in abundance between treatments. Bonferroni-corrected *p*-values were used to determine significant differences ($\alpha = 0.05$).

Results

To evaluate the potential for a viral vaccine to impact the bacterial gut microbiota, cloacal ($n = 87$), fecal ($n = 64$), and oral ($n = 87$) samples were characterized through Illumina sequencing of microbial 16S rRNA gene amplicons. Several alpha diversity measures (phylogenetic diversity, Shannon's diversity index, and the relative abundance of shared taxa) were used to assess microbial diversity within each sample for subsequent comparison among samples. Additionally, unweighted UniFrac distances were used to directly compare microbiota among samples in analyses of beta diversity.

Neither phylogenetic diversity nor Shannon's diversity index significantly differed in cloacal, fecal, or oral samples from challenged versus control individuals (Fig. 5.1). However, a notable decrease in the variation of both alpha diversity metrics was observed in fecal microbiota at 10 days post-hatching (Fig. 5.1b). The taxonomic composition of the microbiota across all gut sites diverged between challenged and control individuals over the nestling period (Fig. 5.2). At 10 days post-hatching, fecal microbiota was dominated by Chlamydiae in the vaccinated group, and oral microbiota of vaccinated individuals exhibited a marked lack of Cyanobacteria relative to the control group (Fig. 5.2b-c). By 15 days post-hatching, microbiota across all gut sites in the vaccinated group contained Cyanobacteria whereas the phylum was predominantly absent in the control group (Fig. 5.2).

Closer analysis and significance testing of the differences in taxonomic composition between treatment groups for each gut site at each age was pursued. At five days post-hatching, wherein microbial samples were collected before immune challenge, significant

differences in the relative abundance of shared bacterial taxa were observed between control and challenged groups across all gut sites (Table 5.2). To gain insight into the effect of experimental treatment on the microbiota, those taxa that differed significantly in abundance between treatment groups before immune challenge were excluded in comparative microbial analyses at 10 and 15 days post-hatching. After this subtractive step, a substantial number of shared taxa still significantly differed in abundance at 10 and 15 days post-hatching (Table 5.2). Furthermore, the proportion of differentially abundant, shared taxa between treatment groups increased from 10 to 15 days post-hatching across all gut sites (Table 5.2).

In analyses of beta diversity across gut sites within each treatment group, parametric and nonparametric *t*-tests revealed significantly higher unweighted UniFrac distances in between-gut-site comparisons versus within-gut-site comparisons at all ages for both treatment groups (Fig. S5.3). Principal coordinate analyses of unweighted UniFrac distances showed no segregation among microbiota between gut sites at five days post-hatching but visible segregation between mucosal (cloacal and oral) and fecal microbiota at 10 days post-hatching for both challenged and control individuals (Fig. 5.3a-b). At 15 days post-hatching, distinct clustering of microbiota by gut site was observed only in the challenged group whereas higher similarity of microbiota between gut sites was seen in the control group (Fig. 5.3c).

Additional analyses of beta diversity across treatment groups within gut sites showed distinct segregation of the fecal microbiota between control and challenged individuals at 10 days post-hatching (i.e., five days post-vaccination; Fig. 5.4). Likewise, both parametric and nonparametric *t*-tests revealed significantly higher unweighted UniFrac distances in between-treatment comparisons relative to within-treatment comparisons for fecal microbiota of 10-day old nestlings (Table 5.1, Fig. S5.4).

Regarding nestling morphology, neither asymmetry of the ninth primaries nor growth performance differed between challenged and control individuals across all ages and gut sites (Figs. S5.5-5.7). However, distances in unweighted UniFrac of the cloacal microbiota at five days post-hatching and of the oral microbiota at 15 days post-hatching significantly associated with distances in asymmetry of the ninth primaries (Table 5.1).

Discussion

In evaluating the potential for viral immune challenges to impact the bacterial gut microbiota of vertebrate organisms, this study found significant differences in both alpha and beta diversity of gut bacteria between challenged and control nestling birds. Specifically, differences in the abundance of shared bacterial taxa between treatment groups were significant across three different gut sites (cloacal, fecal, oral) and two time points (10 and 15 days post-hatching) after immune challenge (Table 5.2). Fecal microbiota were significantly different between vaccinated and control individuals at five days post-challenge (Fig. 5.4b). Furthermore, microbiota across the three different gut sites were notably more distinct among vaccinated individuals than control individuals at 10-days post-challenge (Fig. 5.3c). These results invite consideration of the mechanisms underlying the observed patterns and their implications for host condition.

High within-nest variation in the microbiota among Western Bluebird nestlings has been documented previously (Vo, Chapter 3), and this variation likely contributed to the significant differences in the abundance of shared OTUs between treatment groups before immune challenge. Within-nest variation in microbiota appears unrelated to diet as molecular

characterization of ingested arthropods, the main food source for developing nestlings, revealed no significant differences for individuals within-nest versus between-nest (Vo & Jedlicka 2014). High genetic diversity of loci contributing to microbial community assembly among nestlings within nests may be relevant due to high rates of extra-pair paternity in Western Bluebirds (19% of all nestlings and 45% of all nests; Dickinson & Akre 1998). Nevertheless, given the random assignment of nestlings to treatment groups, it is unlikely that extra-pair paternity would generate systematic differences in microbiota between groups before immune challenge. Furthermore, metagenomic DNA extraction or PCR during sample preparation can impact representation of the true abundances of bacteria in the original sample. However, procedural effects are expected to be systematic such that comparisons of within-study results should be internally consistent. Copy number differences of the 16S rRNA gene between bacterial species also do not affect inferences drawn here from within-taxon comparisons of abundance between treatment groups. Thus, stochasticity may best explain the pre-challenge differences in the abundance of shared OTUs between treatment groups, but after excluding these OTUs from downstream analyses, significant differences in the abundance of numerous other OTUs were still observed post-challenge.

Interestingly, beta diversity analyses revealed significant differences in the fecal but not cloacal or oral microbiota between treatment groups post-challenge (Figs. 5.1b and 5.4b), yet microbiota across all gut sites were notably more segregated among challenged individuals relative to control individuals (Fig. 5.3c). Mechanisms through which the host immune system is known to shape the gut microbiota typically involve secretion of general (e.g., antimicrobial peptides) or specific (e.g., antibodies) regulatory factors (Hooper *et al.* 2012). A skew towards production of specific antibodies may explain the effects of the immune challenge on the alpha and beta diversity of microbiota across all gut sites. In both avian (Fair *et al.* 2003) and mammalian (Barrios *et al.* 1996) neonates, immune responses to inactivated viral vaccine antigens involve the upregulation of CD4 T-helper 2 (Th2) cytokines (e.g., IL-5), the inhibition of CD4 T-helper 1 (Th1) cytokines (e.g., IFN- γ), and thus a Th2 bias that drives production of specific antiviral antibodies. Although immunological immaturity exists among neonates, this robust antibody-mediated immune response is reproducible *in vivo* in mammals and birds during the first week of life, and it is believed that the absence of strong co-stimulatory signals in the underdeveloped neonatal immune system promotes a Th2 bias by default (Barrios *et al.* 1996). Within this Th2-skewed immunological environment, antigen presenting cells (e.g., dendritic cells, macrophages) that display peptides of bacteria sampled from the gut likewise have the opportunity to interact with compatible T and B cells to subsequently induce production of gut bacterial-specific secretory antibodies (IgA).

In the gastrointestinal tract, the layer of mucus between the gut lumen and the intestinal epithelium is thought to comprise the key mechanism for maintaining host immunological ignorance of the vast microbial load present in the intestinal lumen. Dendritic cells typically phagocytose bacteria associated with the mucosal surfaces rather than the gut lumen (Duerkop *et al.* 2009). Likewise, after transcytosis of IgA from the basal to the apical side of the intestinal epithelium, the IgA are usually retained in the mucus layer rather than diffused into the gut lumen. It is therefore particularly interesting that the fecal microbiota exhibited a more robust signature of effects post-challenge than the mucosal microbiota in this study. Flow cytometric analysis of fecal bacteria, which comprises a highly sensitive detection method, has

revealed that 24–74% of *in vivo* fecal anaerobic bacteria can be coated with IgA (Waaij *et al.* 1996). Therefore, efficient uptake of luminal bacteria and high loads of IgA diffusion into the gut lumen is possible, which can feasibly drive changes in the composition of luminal microbiota.

Immunological responses involving specific antibody production usually run the course of a week. After plasma cells have been activated and specific antibodies have sufficiently eliminated the main antigenic threat, the physiological changes responsible for the mounted immune responses are downregulated to prevent sustained immune activation and immunological pathology (Murphy 2011). This time series of events is consistent with the signature of beta diversity differences in the fecal microbiota between treatment groups existing only at 10 and not 15 days post-hatching, which corresponds to five and 10 days post-vaccination, respectively. Additionally, the half-life of avian antibodies is known to span 1.68–3.3 days (Davison *et al.* 2011), which could explain the detectable persistence of effects on the relative abundance and beta diversity of microbiota across all gut sites at 15 days post-hatching.

That the oral microbiota showed effects from an immune challenge administered in the peritoneal cavity (i.e., abdominal cavity) formed another surprising result of this study. The peritoneal cavity is proximal to the lower gut and substantially distal to the oral cavity. These results suggest that both local and systemic microbial effects can result from an immune challenge. This may occur as a result of the similarity of microbiota between the cloacal and oral mucosal environments (Vo, Chapter 3). Activated antigen presenting cells that have phagocytosed bacteria from the intestines as well as activated T cells can travel through the lymph system to other sites in the host body. Thus, if these immune cells recruit to the oral mucosa and if the same bacteria in the cloaca are found in the oral cavity, a similar effect of the immune challenge would be observed between the two distantly located gut sites. The mechanistic details of this potential warrant further investigation and was beyond the scope of this study.

It is conceivable that an immunologically independent mechanism may explain the impact of vaccination on the gut microbiota. If vaccination leads to sickness and behavioral modifications such as lethargy and reduced vocalization, the effects would substantially impact parental feeding decisions, which are typically based on proximity of nestlings to the box hole, height of open gape extension, and amplitude of begging (Leonard *et al.* 1994). Nestlings exhibiting reduced activity and begging behavior may be consistently overlooked, underfed, or fed poorer quality prey items. A microbial signature reflecting this diet shift would be expected more in the fecal microbiota than the mucosal microbiota due to the direct dependence of fecal composition on diet. However, nestling morphological metrics did not significantly differ between treatment groups at any age, which suggests that vaccinated nestlings were being fed sufficiently such that no trade-off was mandated between the energetically demanding activities of somatic growth and immune activation. Thus, although the interaction of nestling behavioral and parental feeding modifications could underlie the observed pattern in fecal microbiota between treatments, it does not appear to be a likely explanatory factor.

Ecoimmunological theory predicts that given limited resources, trade-offs will occur between energetically demanding physiological processes, such as the activation of an immune response versus somatic growth (Martin *et al.* 2011). The 2013 Western Bluebird breeding

season comprised an exceptional drought year whereby 89% and 31% declines in precipitation and primary productivity were respectively observed from 2012 to 2013 at the study site (San Joaquin Experimental Range unpublished data). Similarly, arthropod abundance was visibly diminished in 2013 (pers. obs.). Observations of unprecedented nestling deaths in the nest boxes were presumed to be due to starvation. Nevertheless, the immune challenged nestlings in this study did not exhibit significant morphological differences, either in asymmetry of the ninth primaries or growth performance, relative to control nestlings. These outcomes suggest that despite apparent food shortages, compensatory mechanisms may exist to yield a general lack of tradeoffs between immune function and nestling growth. However, the oral microbiota were responsive to immune challenge and were significantly associated with asymmetry of the flight feathers. Thus, a more dynamic, interactive network between immune function, microbial assembly, and host growth may exist and could be clarified with finer-scale, mechanistic, experimental work.

Conclusions

The goal of this study was to investigate whether immune challenges may incur effects that to date have yet to be investigated: alteration of the vertebrate gut microbiota during early life stages. Given the intricate nature of interactions between the gut microbiota and host immune function, it was hypothesized that an immune challenge could comprise a physiological perturbation with ramifications for gut microbial ecology. Indeed, shifts for challenged individuals relative to control individuals in the alpha and beta diversity of the microbiota in all three studied gut sites were consistent with the proposed hypothesis. Interestingly, the administration of a viral immune challenge appeared capable of inducing changes in the bacterial microbiota. Furthermore, administration of an intraperitoneal challenge seemed to generate effects on the oral microbiota. These patterns might seem counterintuitive but may not be in light of proposed immunological pathways.

Steps for future research include comparative studies on the effects of various vaccines, experiments to evaluate immunological hypotheses for the mechanisms underlying the observed patterns, and elucidation on the extent to which these vaccine-mediated shifts in microbiota may be biologically relevant for host condition. These directions are relevant for understanding the selective pressures posed by early life immune challenges on the ecology and evolution of host-microbial associations and host development. Likewise, further research is especially of interest in the context of public health, wherein the recommended vaccination schedule for infants and children includes 10 different vaccines, covering 10 viral and 14 total pathogens, administered a total of 33 different times during the first six years of life. Ultimately, greater characterization of the gut microbial implications of vaccines would support more informed usage, with widespread application to agricultural industries, veterinary practice, and human medicine.

References

- Atarashi K, Tanoue T, Shima T *et al.* (2011) Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science*, **331**, 337–341.
- Barrios C, Brawand P, Berney M *et al.* (1996) Neonatal and early life immune responses to various forms of vaccine antigens qualitatively differ from adult responses:

- predominance of a Th2-biased pattern which persists after adult boosting. *European Journal of Immunology*, **26**, 1489–1496.
- Braid MD, Daniels LM, Kitts CL (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. *Journal of Microbiological Methods*, **52**, 389–393.
- Bystrykh LV (2012) Generalized DNA barcode design based on Hamming codes. *PLoS ONE*, **7**, e36852.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.
- CDC (2014) Recommended Immunizations for Children (Birth through 6 years). Centers for Disease Control and Prevention. <http://www.cdc.gov/vaccines/schedules/easy-to-read/child.html>.
- Davison F, Kaspers B, Schat KA, Kaiser P (2011) *Avian Immunology*. Academic Press.
- Dickinson JL, Akre JJ (1998) Extrapair paternity, inclusive fitness, and within-group benefits of helping in western bluebirds. *Molecular Ecology*, **7**, 95–105.
- Dominguez-Bello MG, Costello EK, Contreras M *et al.* (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 11971–11975.
- Duerkop BA, Vaishnava S, Hooper LV (2009) Immune responses to the microbiota at the intestinal mucosal surface. *Immunity*, **31**, 368–376.
- Fair JM, Myers OB (2002) The ecological and physiological costs of lead shot and immunological challenge to developing Western Bluebirds. *Ecotoxicology*, **11**, 199–208.
- Fair JM, Myers OB, Ricklefs RE (2003) Immune and growth response of Western Bluebirds and Ash-throated Flycatchers to soil contaminants. *Ecological Applications*, **13**, 1817–1829.
- Faith DP, Baker AM (2007) Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. *Evolutionary Bioinformatics Online*, **2**, 121–128.
- Fisher S, Barry A, Abreu J *et al.* (2011) A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biology*, **12**, R1–15.
- Harmsen HJM, Wildeboer-Veloo ACM, Raangs GC *et al.* (2000) Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *Journal of Pediatric Gastroenterology and Nutrition*, **30**, 61–67.
- Heijtz RD, Wang S, Anuar F *et al.* (2011) Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 3047–3052.
- Hooper LV, Littman DR, Macpherson AJ (2012) Interactions between the microbiota and the immune system. *Science*, **336**, 1268–1273.
- Ivanov II, Atarashi K, Manel N *et al.* (2009) Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*, **139**, 485–498.
- Jedlicka JA, Sharma AM, Almeida RPP (2013) Molecular tools reveal diets of insectivorous birds from predator fecal matter. *Conservation Genetics Resources*, **5**, 879–885.

- Koenig JE, Spor A, Scalfone N *et al.* (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 4578–4585.
- Kozarewa I, Ning Z, Quail MA *et al.* (2009) Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nature Methods*, **6**, 291–295.
- Lee KA (2006) Linking immune defenses and life history at the levels of the individual and the species. *Integrative and Comparative Biology*, **46**, 1000–1015.
- Lens L, Van Dongen S, Kark S, Matthysen E (2002) Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies? *Biological Reviews*, **77**, 27–38.
- Leonard ML, Teather KL, Horn AG, Koenig WD, Dickinson JL (1994) Provisioning in western bluebirds is not related to offspring sex. *Behavioral Ecology*, **5**, 455–459.
- Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of mammals and their gut microbes. *Science*, **320**, 1647–1651.
- Lohse M, Bolger AM, Nagel A *et al.* (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, **40**, W622–W627.
- Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228–8235.
- Macpherson AJ, Uhr T (2004) Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. *Science*, **303**, 1662–1665.
- Martin LB, Hawley DM, Ardia DR (2011) An introduction to ecological immunology. *Functional Ecology*, **25**, 1–4.
- Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for Illumina sequences. *BMC Bioinformatics*, **13**, 31.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3229–3236.
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, **2010**, pdb.prot5448.
- Murphy K (2011) *Janeway's Immunobiology, Eighth Edition*. Garland Science.
- Nicholson JK, Holmes E, Kinross J *et al.* (2012) Host-gut microbiota metabolic interactions. *Science*, **336**, 1262–1267.
- Ouwerkerk JP, de Vos WM, Belzer C (2013) Glycobiome: bacteria and mucus at the epithelial interface. *Best Practice & Research Clinical Gastroenterology*, **27**, 25–38.
- Pearson G, McCann M (1975) The role of indigenous wild, semidomestic, and exotic birds in the epizootiology of velogenic viscerotropic Newcastle disease in southern California, 1972–1973. *Journal of the American Veterinary Medical Association*, **167**, 610–614.
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE*, **5**, e9490.
- Pull SL, Doherty JM, Mills JC, Gordon JI, Stappenbeck TS (2005) Activated macrophages are an adaptive element of the colonic epithelial progenitor niche necessary for regenerative responses to injury. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 99–104.

- Reed KD, Meece JK, Henkel JS, Shukla SK (2003) Birds, migration and emerging zoonoses: West Nile virus, Lyme disease, influenza A and enteropathogens. *Clinical Medicine & Research*, **1**, 5–12.
- Ridaura VK, Faith JJ, Rey FE *et al.* (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, **341**, 1241214.
- Salzman NH, Hung K, Haribhai D *et al.* (2010) Enteric defensins are essential regulators of intestinal microbial ecology. *Nature Immunology*, **11**, 76–82.
- Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nature Methods*, **5**, 16–18.
- Sommer F, Bäckhed F (2013) The gut microbiota — masters of host development and physiology. *Nature Reviews Microbiology*, **11**, 227–238.
- Tremaroli V, Bäckhed F (2012) Functional interactions between the gut microbiota and host metabolism. *Nature*, **489**, 242–249.
- Turnbaugh PJ, Hamady M, Yatsunenko T *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature*, **457**, 480–484.
- Vaishnava S, Behrendt CL, Ismail AS, Eckmann L, Hooper LV (2008) Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 20858–20863.
- Vo AE, Jedlicka JA (2014) Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for fecal and swab samples. *Molecular Ecology Resources*, In press.
- Waaij LA van der, Limburg PC, Mesander G, Waaij D van der (1996) In vivo IgA coating of anaerobic bacteria in human faeces. *Gut*, **38**, 348–354.
- Wu H-J, Ivanov II, Darce J *et al.* (2010) Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity*, **32**, 815–827.

Tables

Table 5.1. Results of parametric and nonparametric *t*-tests assessing whether unweighted UniFrac distances within versus between levels of treatment, asymmetry of the ninth primaries, or growth performance were significantly different. Bonferonni-corrected p-values are reported, and asterisks indicate significant p-values ($\alpha = 0.05$). “NO VAR” indicates no variance between groups.

Gut Site	Age (days post-hatching)	Treatment		Asymmetry 9th primaries		Growth Performance	
		Parametric p	Nonparametric p	Parametric p	Nonparametric p	Parametric p	Nonparametric p
cloacal	5	0.364	0.347	0.017*	0.018*	NO VAR	NO VAR
	10	0.182	0.197	0.337	0.338	NO VAR	NO VAR
	15	0.579	0.562	0.035	0.024	NO VAR	NO VAR
fecal	5	0.534	0.509	NO VAR	NO VAR	NO VAR	NO VAR
	10	0.008*	0.007*	0.253	0.270	NO VAR	NO VAR
	15	0.760	0.756	0.258	0.247	NO VAR	NO VAR
oral	5	0.369	0.346	NO VAR	NO VAR	NO VAR	NO VAR
	10	0.891	0.894	0.162	0.159	NO VAR	NO VAR
	15	0.523	0.522	0.052*	0.057*	NO VAR	NO VAR

Table 5.2. Total shared OTUs and OTUs that significantly differed in abundance ($\alpha = 0.05$ for Bonferonni-corrected G-test p-values) in the microbiota of individuals in the control and vaccinated groups for each sampled gut site and host age. Asterisk indicates value after subtracting all OTUs that significantly differed in abundance at the age of five days post-hatching.

Gut site	Age	# differentially abundant OTUs	# total shared OTUs	% differentially abundant of total shared OTUs
cloacal	5	51	92	55
	10	20*	79*	25*
	15	91*	273*	33*
fecal	5	78	215	36
	10	11*	82*	13*
	15	38*	86*	44*
oral	5	115	238	48
	10	42*	218*	19*
	15	61*	253*	24*

Figures

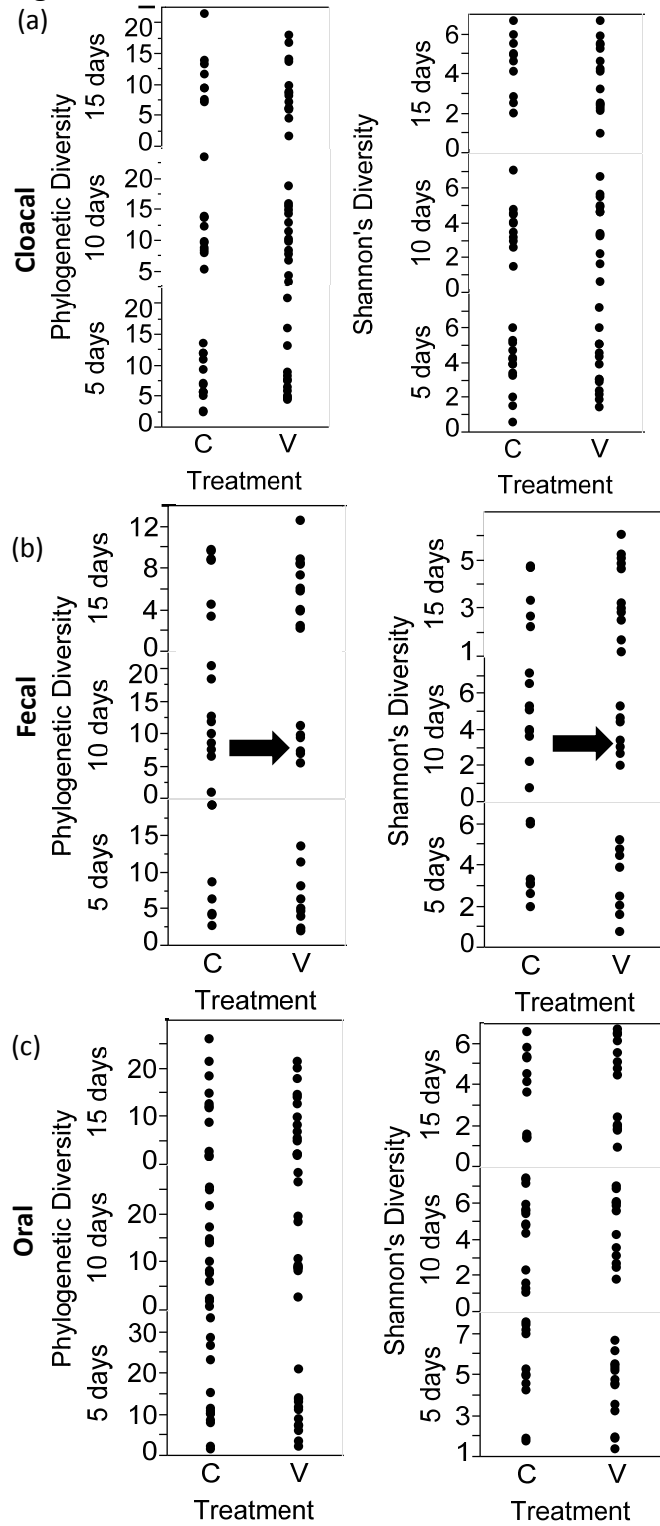


Fig. 5.1. Comparison of phylogenetic diversity and Shannon's diversity index of cloacal (a), fecal (b), and oral (c) microbiota between experimental treatments (C for control, V for vaccinated) at ages five, 10, and 15 days post-hatching. Notable differences are highlighted with arrows.

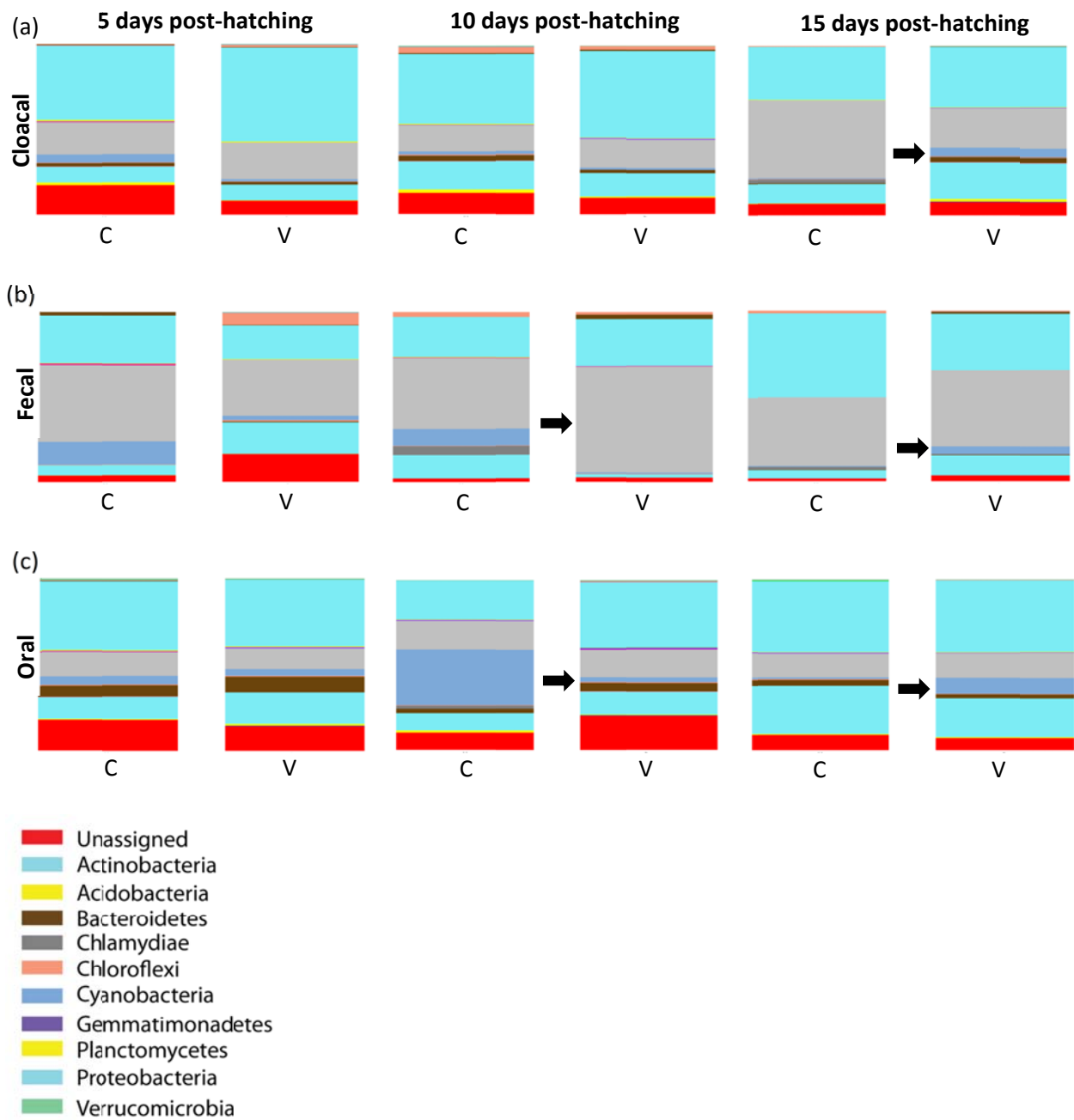


Fig. 5.2. Comparison of taxonomic composition (bacterial phyla) of cloacal (a), fecal (b), and oral (c) microbiota between experimental treatments (C for control, V for vaccinated) at five, 10, and 15 days post-hatching. Taxa are colored in alphabetical order by phylum name from the bottom to the top of each column. Notable differences are highlighted with arrows.

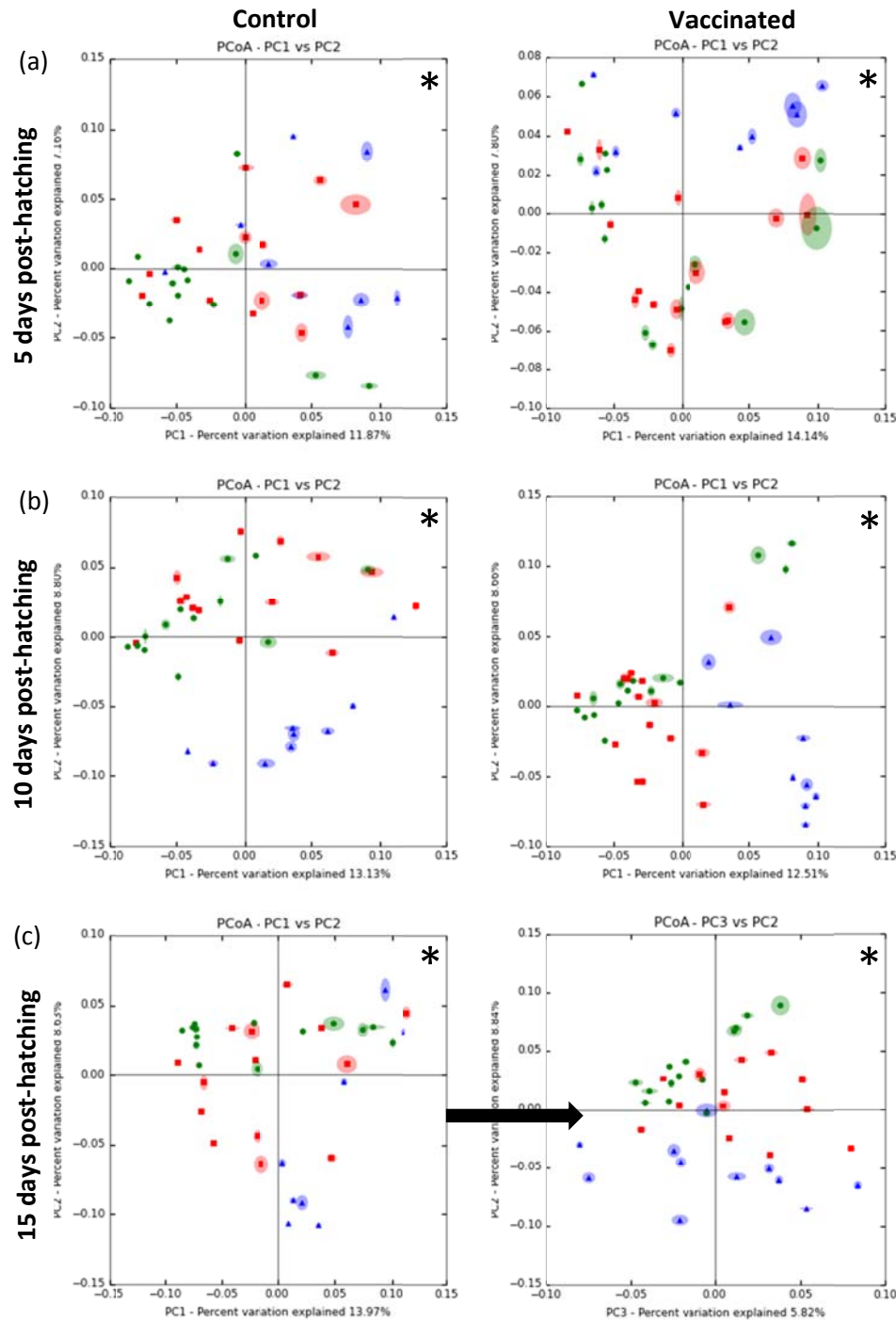


Fig. 5.3. Principal coordinate analysis of bacterial unweighted UniFrac for cloacal (red), fecal (blue), and oral (green) samples at five (a), 10 (b), and 15 (c) days post-hatching for vaccinated (right) and control (left) individuals. Asterisk indicates significant differences in unweighted UniFrac distances within versus between gut sites, using both parametric and nonparametric t -tests. Notable differences in segregation of microbiota between gut sites highlighted with arrow.

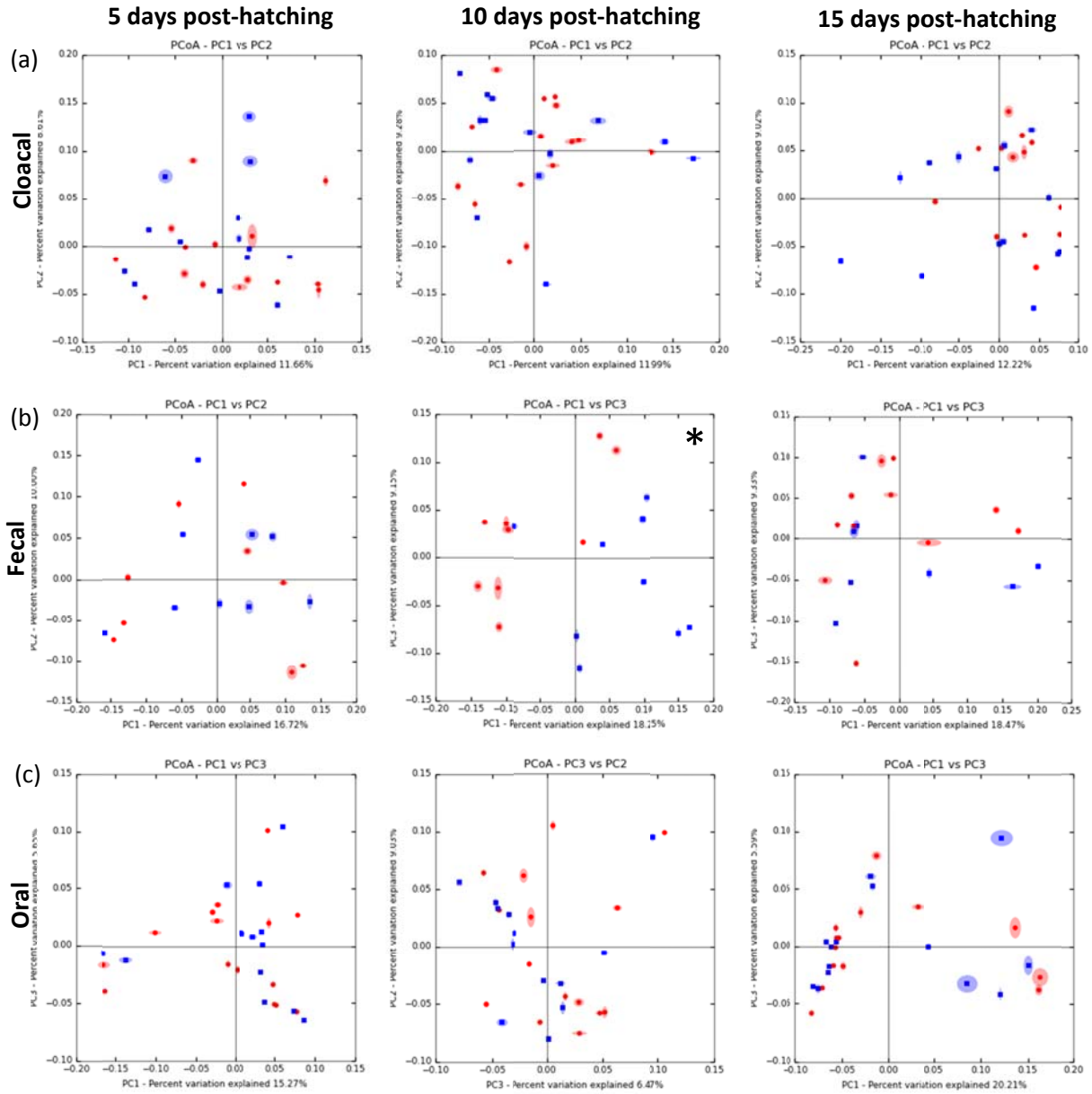


Fig. 5.4. Principal coordinate analysis of bacterial unweighted UniFrac for cloacal (a), fecal (b), and oral (c) samples at five (left), 10 (middle), and 15 (right) days post-hatching, colored by vaccination (red) or control (blue) treatment. Asterisk indicates significant differences in unweighted UniFrac distances within versus between treatment groups, using both parametric and nonparametric *t*-tests.

Supporting Information

Introduction

Interactions between the gut microbiota and immune function

Through Toll-like receptors (TLRs) and MyD88 activation, intestinal epithelial cells detect any bacteria that invade the mucosal surfaces, and subsequent production of antimicrobial factors such as RegIII γ generally limits microbial incursions into host tissues (Vaishnava *et al.* 2008). Dendritic cells in the lamina propria, the layer beneath the intestinal epithelium, additionally phagocytose bacteria that penetrate the mucosa. Interaction with these bacteria-loaded dendritic cells and helper T cells activates B cells to produce secretory immunoglobulin A (IgA) that are transported into intestinal lumen to bind and prevent microbial translocation across the epithelium into host tissues (Macpherson & Uhr 2004). Antibacterial peptides such as defensins are furthermore produced by the intestinal epithelium and diffuse into the lumen to mediate overall community composition of the gut microbiota (Salzman *et al.* 2010).

The effects of interactions between the host and microbiota are also bidirectional. Sensing of commensal bacteria through TLRs and MyD88 signaling is necessary for cell proliferation and repair of damaged host intestinal epithelium (Pull *et al.* 2005). Different taxa within the microbial consortium polarize different T cell subsets in the lamina propria, with ramifications at both the local and systemic levels. Specifically, skewed induction of pro-inflammatory T_H17 cells (Ivanov *et al.* 2009) versus anti-inflammatory T_{reg} cells (Atarashi *et al.* 2011) significantly mediates the trajectory of local immune responses to enteric pathogenic threats and the incidence of collateral damage from immunological hyper-responsiveness. Beyond the lamina propria, induction of intestinal T_H17 cells contributes to the development of systemic autoimmune diseases such as arthritis (Wu *et al.* 2010) whereas induction of intestinal T_{reg} reduces systemic inflammation (Atarashi *et al.* 2011).

Numerous pathways exist for an immunological perturbation such as a vaccination to subsequently impact the structure and composition of the gut microbiota, with dynamic sequelae that affect systemic host condition. For example, a vaccine can produce a skew in systemic cytokine profiles and availability of T-cell subsets, which would then affect recruitment of immune cells to the lamina propria and effector activities in response to gut microbial insults. The dearth of research on this and other mechanistic possibilities that may underlie the intersection of immune challenges, immune function, and gut microbial ecology prompted the current study, which investigates the effects of a viral vaccine on the bacterial microbiota within the gastrointestinal tract of neonates.

Materials and methods

Metagenomic DNA extraction, 16S rRNA amplicon library preparation, and sequencing

Samples were homogenized with zirconia-silica beads (0.1 and 0.5 mm diameter) in lysis buffer (166 mM trisodium phosphate, 111 mM guanidine thiocyanate, 11 mM sodium chloride, 0.3% sodium dodecyl sulfate, 37 mM Tris-HCl pH 8.0, 1% beta-mercaptoethanol) on a PowerLyzer 24 (Mo Bio 13255). Sample lysates were incubated with ammonium acetate to precipitate proteins and aluminum ammonium sulfate to flocculate PCR inhibitors (Braid *et al.* 2003). SPRI beads were used to capture, wash, and elute sample DNA. The Nanodrop spectrophotometer and

Qubit fluorometer dsDNA HS (high sensitivity) Assay Kit (Invitrogen Q32854) quantified all extracts, and elutions were stored at -20°C until use.

Metagenomic DNA extracts were amplified using degenerate, barcoded primers targeting the bacterial 16S rRNA locus (515F and 806R; Caporaso *et al.* 2012) in triplicate 25 µL reactions containing 1X HotMasterMix (5 PRIME 2200400), 0.2 µM of each primer, 0.5 µg/µL BSA (NEB B9000S), 5% DMSO (ACS grade), and at least 0.1 ng of DNA. Thermocycling conditions followed an initial denaturation of 94°C for 2 minutes and 35 cycles of 94°C for 8 seconds, 50°C for 20 seconds, and 72°C for 30 seconds. PCR reactions were set-up in an UV-irradiated clean hood and thermocycled in a separate, isolated room to minimize risk of cross-contamination. Amplicons were electrophoresed on 1.5% agarose gels stained with SYBR-Gold (Invitrogen S-11494) to verify amplification success.

For Illumina amplicon library preparation, triplicate PCR reactions were pooled and cleaned using a 1.2 volume ratio of SPRI bead solution to sample to discard fragments smaller than 200 bp (e.g., primer dimers). After quantification using the Qubit dsDNA HS Assay Kit, the cleaned samples were pooled equimolarly to form 24 libraries of 18-21 samples and 81-507 ng of total amplicon DNA per library. Following Vo and Jedlicka (2014), the NEBNext End Repair Module, NEBNext dA-Tailing Module, and Meyer and Kircher (2010) adapter ligation protocols were used with TruSeq Illumina indexed adapters to prepare the libraries using a “with-bead” and “PCR-free” approach. After quantification of each of the 24 libraries via qPCR using the KAPA Library Quantification Kit (KK4824), libraries were pooled equimolarly and analyzed on the Bioanalyzer (Agilent Technologies G2940CA) before sequencing on the Illumina MiSeq platform.

References

- Atarashi K, Tanoue T, Shima T *et al.* (2011) Induction of colonic regulatory T cells by indigenous Clostridium species. *Science*, **331**, 337–341.
- Braid MD, Daniels LM, Kitts CL (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. *Journal of microbiological methods*, **52**, 389–393.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.
- Ivanov II, Atarashi K, Manel N *et al.* (2009) Induction of Intestinal Th17 Cells by Segmented Filamentous Bacteria. *Cell*, **139**, 485–498.
- Macpherson AJ, Uhr T (2004) Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. *Science*, **303**, 1662–1665.
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, **2010**, pdb.prot5448.
- Pull SL, Doherty JM, Mills JC, Gordon JI, Stappenbeck TS (2005) Activated macrophages are an adaptive element of the colonic epithelial progenitor niche necessary for regenerative responses to injury. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 99–104.
- Salzman NH, Hung K, Haribhai D *et al.* (2010) Enteric defensins are essential regulators of intestinal microbial ecology. *Nature Immunology*, **11**, 76–82.
- Vaishnava S, Behrendt CL, Ismail AS, Eckmann L, Hooper LV (2008) Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface.

Proceedings of the National Academy of Sciences of the United States of America, **105**, 20858–20863.

Vo AE, Jedlicka JA (2014) Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for fecal and swab samples. *Molecular Ecology Resources*, In press.

Wu H-J, Ivanov II, Darce J *et al.* (2010) Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity*, **32**, 815–827.

Results

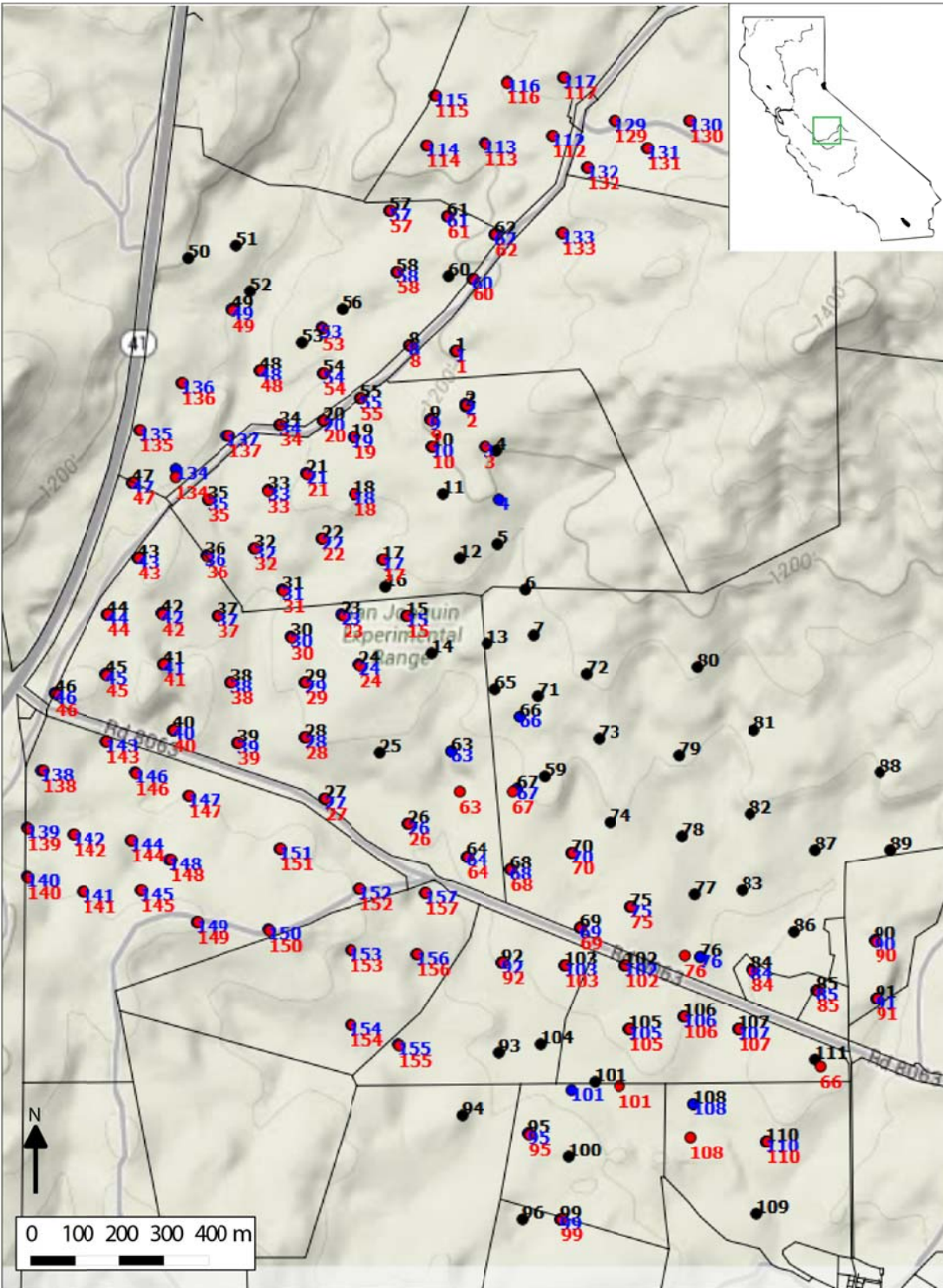
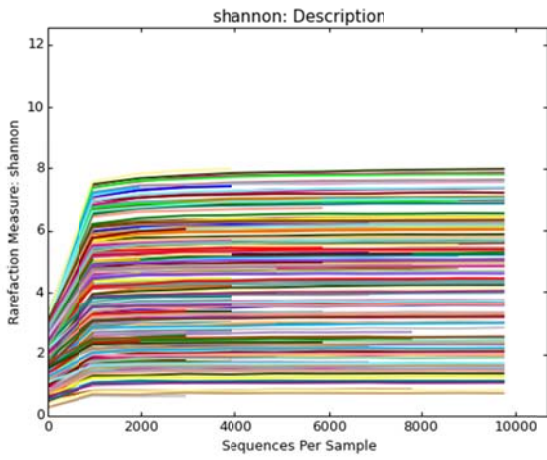


Fig. S5.1. Map of the Western Bluebird nest box study system at the San Joaquin Experimental Range (Madera County, CA) used in this study from 2011 through 2013. Black, blue, and red circles and fonts indicate boxes erected in 2011, 2012, and 2013, respectively. Lines depict fences.

(a)



(b)

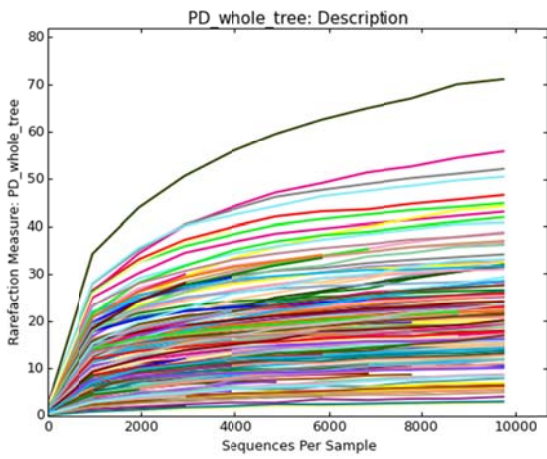


Fig. S5.2. Rarefaction analysis of Shannon's diversity index (a) and phylogenetic distance (b) for microbiota in all samples collected in study ($n = 238$). Each line depicts a different sample.

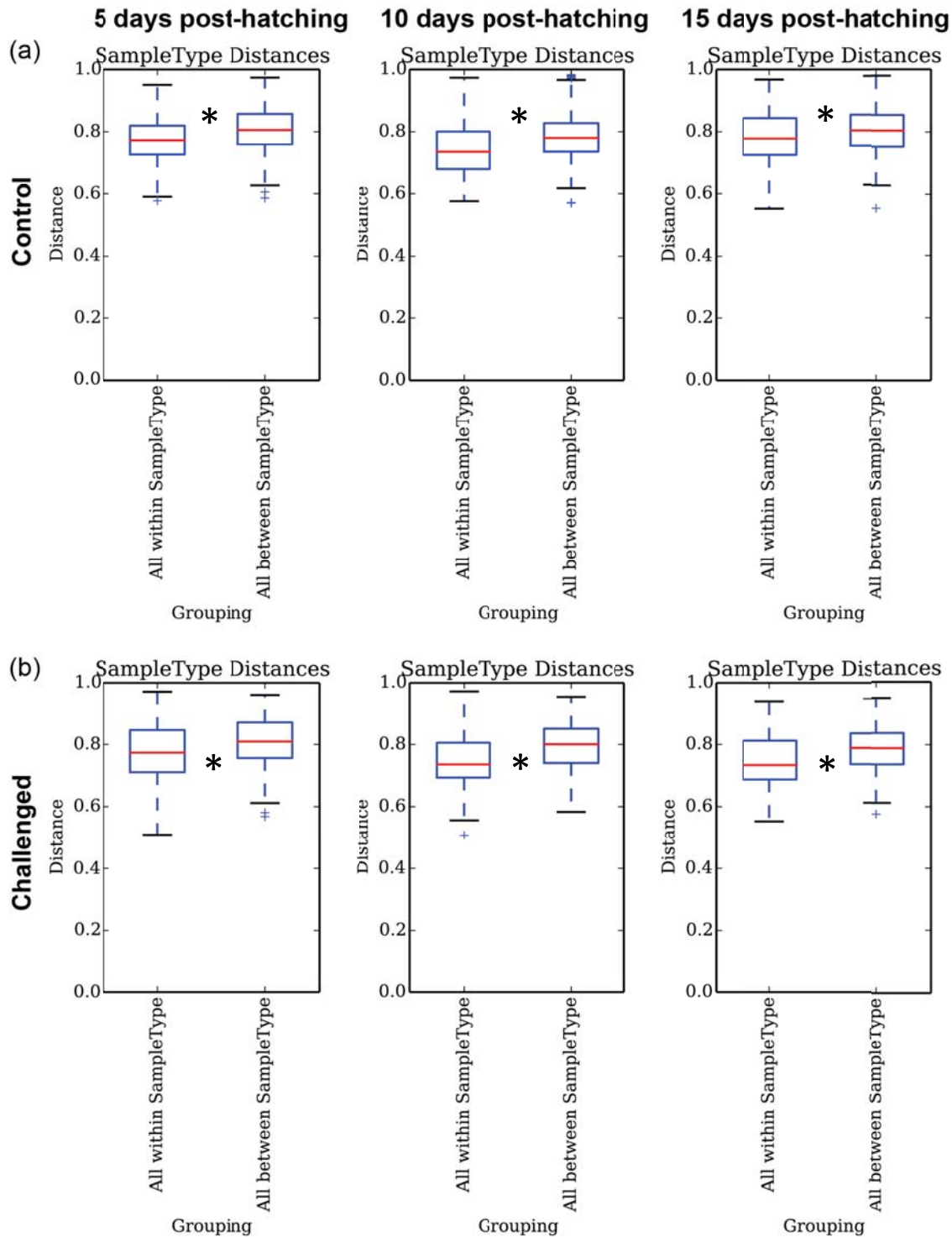


Fig. S5.3. Boxplots comparing bacterial unweighted UniFrac distances within versus between sample types (i.e., gut site: cloacal, fecal, and oral) at five (left), 10 (middle) and 15 (right) days post-hatching of challenged (a) and control (b) individuals. Asterisk indicates significant differences in unweighted UniFrac distances within versus between gut sites, using both parametric and nonparametric *t*-tests.

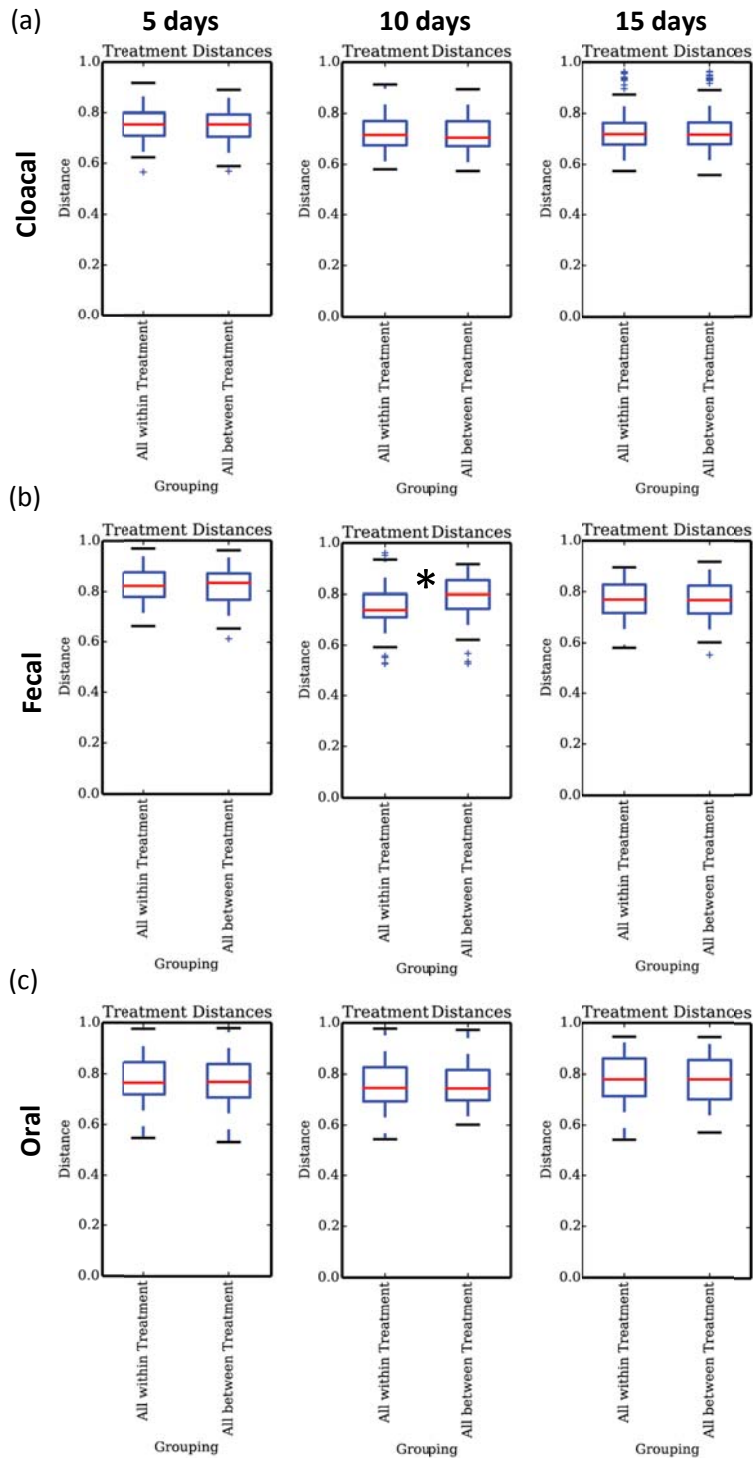


Fig. S5.4. Boxplots comparing bacterial unweighted UniFrac distances within versus between treatment groups for cloacal samples (a), fecal samples (b), and oral samples (c) at five (left), 10 (middle) and 15 (right) days post-hatching. Asterisk indicates significant differences in unweighted UniFrac distances within versus between treatment groups, using both parametric and nonparametric *t*-tests.

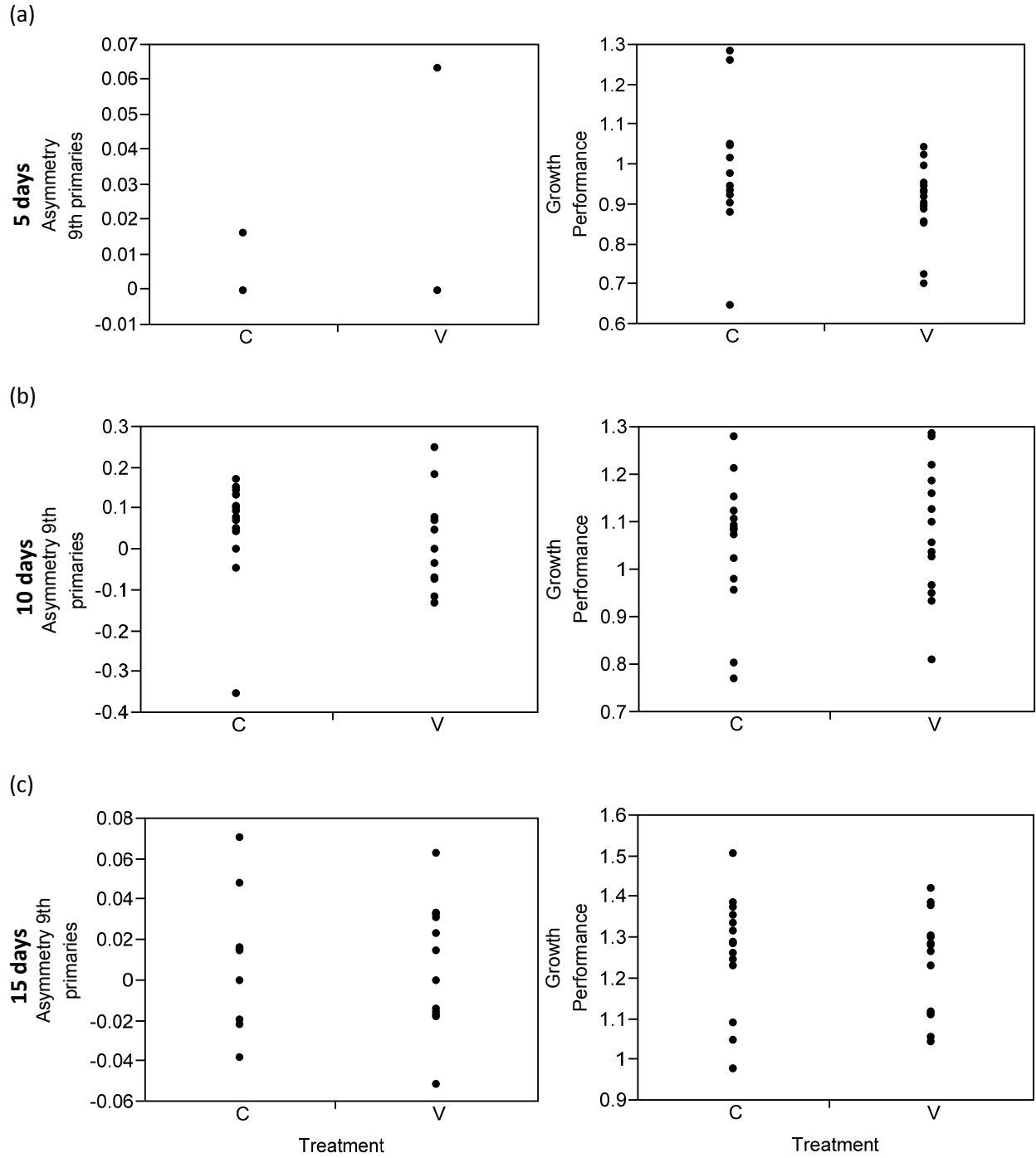


Fig. S5.5. Comparison of asymmetry in the ninth primaries and growth performance of control versus vaccinated nestlings in the cloacal sample group at ages five (a), 10 (b), and 15 (c) days post-hatching.

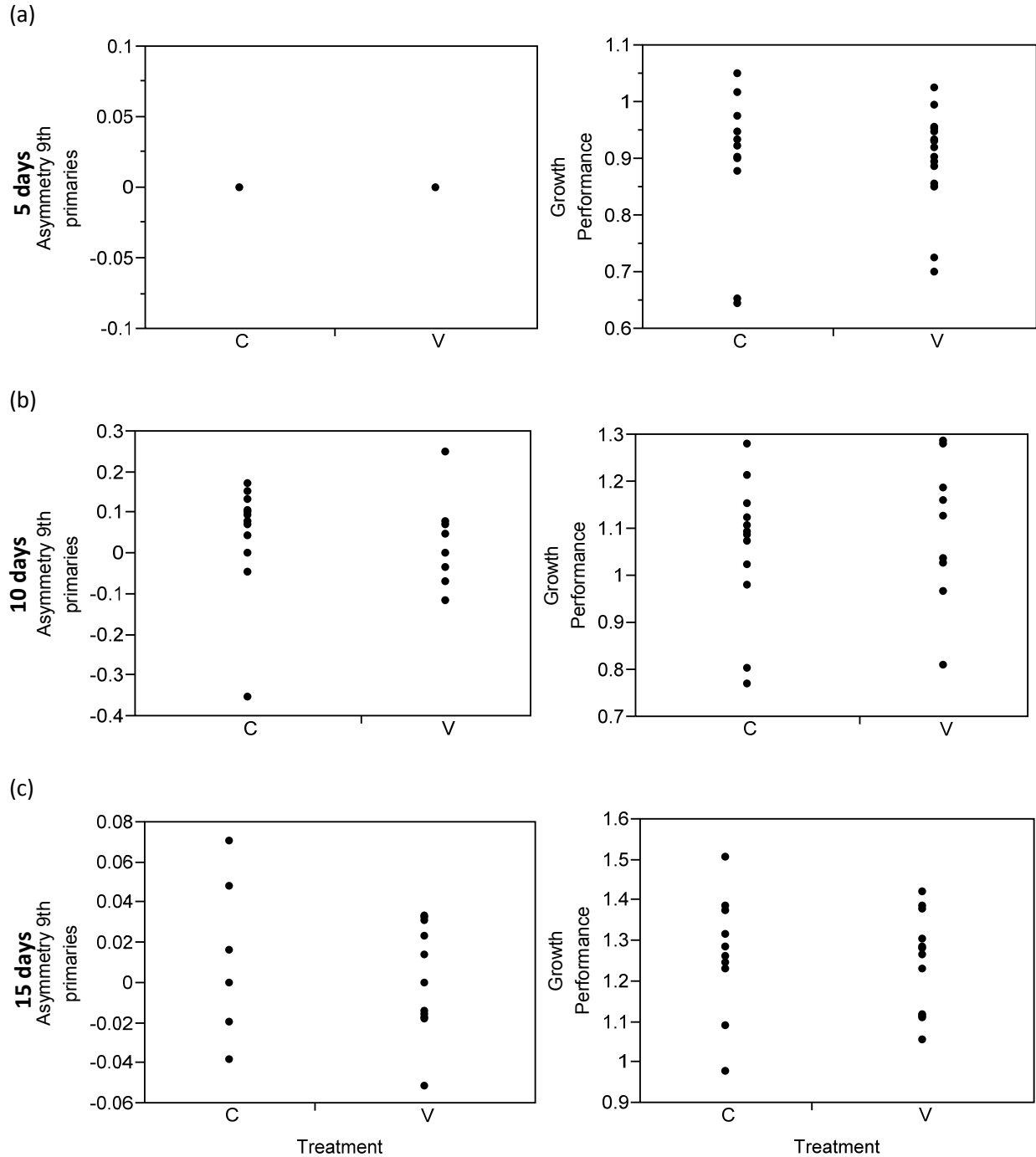


Fig. S5.6. Comparison of asymmetry in the ninth primaries and growth performance of control versus vaccinated nestlings in the fecal sample group at ages five (a), 10 (b), and 15 (c) days post-hatching.

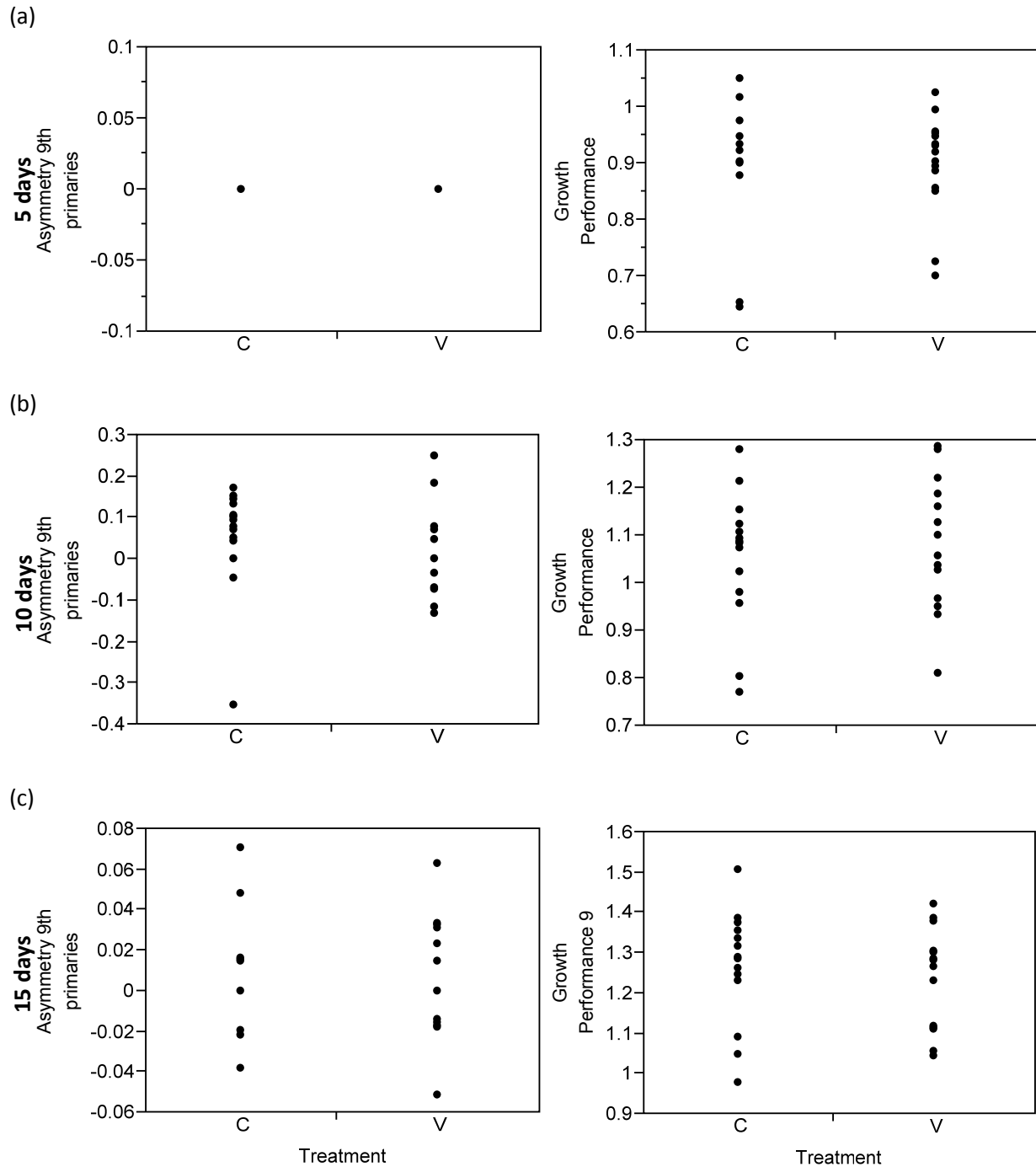


Fig. S5.7. Comparison of asymmetry in the ninth primaries and growth performance of control versus vaccinated nestlings in the oral sample group at ages five (a), 10 (b), and 15 (c) days post-hatching.

Synthesis

Introduction

The numerical and genetic predominance of microbes found in host systems has motivated a great body of work to date investigating how microbiota assemble, interact with host organisms, and influence host biology (Ley *et al.* 2008). From this work, the broader “metaorganismal” unit of biological organization was born and developed (Bosch & McFall-Ngai 2011). Metaorganismal research in humans and mouse models revealed extraordinary microbial diversity in the vertebrate gut as well as pervasive effects of the gut microbiota on host condition (Turnbaugh *et al.* 2009; Ridaura *et al.* 2013). Yet, the causes and consequences of variation in the microbiota of free-living hosts remain exceedingly understudied (McFall-Ngai *et al.* 2013). Understanding the role of microbiota in mediating host physiology and fitness in natural systems is critical to determining the relevance of microbes for studying vertebrate ecology and evolution in non-model taxa. Technological advances have now opened the doors for researchers to achieve more comprehensive studies on host-associated microbiota. Accordingly, this dissertation focused on the field of avian microbiology and addressed the hypothesis that *variation in the avian gut microbiota incurs biologically relevant impacts for host birds in free-living populations*. To evaluate this hypothesis, several sub-questions were pursued:

1. How can minimally invasive samples and massively parallel sequencing be integrated into the study of avian microbiology?
2. How much standing variation exists in the avian gut microbiota, and what are the spatiotemporal drivers?
3. How do avian gut microbiota vary with host morphology?
4. How do avian-associated microbial systems respond to perturbation?

Empirical findings

This research utilized nesting Western Bluebirds *Sialia mexicana* in a nest box system at the San Joaquin Experimental Range (Madera County, CA) to study gut microbial ecology in a natural population of birds. In particular, studies focused on the nestling period as a vulnerable developmental stage during which birds undergo substantial somatic growth and microbial colonization of the gut. These co-occurring processes present the ideal opportunity to obtain longitudinal samples to analyze associations between spatiotemporal factors, assembly of the microbiota, and nestling development.

To enhance the feasibility of avian microbiological research that exploits the availability of high-throughput sequencing technology, **Chapter Two** provided much needed protocols to incorporate the effective use of minimally invasive avian samples. Practical fecal and swab preservation protocols were validated, including sample storage in a DNA stabilization buffer on ice and subsequently at -20°C in a standard household freezer for several months in the field before transfer of the samples to -80°C in the laboratory for long-term storage. Likewise, the development of a novel metagenomic DNA extraction method based on solid phase reversible immobilization (SPRI) beads produced significantly higher yields and quality of DNA than the standard extraction kits recommended by both the Human and Earth Microbiome Projects. Furthermore, a cost-effective library preparation protocol was developed for amplicon

sequencing on the Illumina platform and validated using bacterial 16S rRNA amplicons generated from SPRI-based extracts of minimally invasive avian samples.

Applying the methods developed in **Chapter Two**, **Chapter Three** revealed high intra- and inter-individual variation in the gut microbiota of Western Bluebirds. Intra-individual differences among microbiota across gut sites appeared to diminish with host age. Inter-individual variation in microbiota was not strongly predicted by shared nest environment or geographic distance between nests but rather was predicted by ambient temperature and inter-annual variation. Contextualization of the results of **Chapter Three** with nestling morphological data in **Chapter Four** revealed associations between the fecal, cloacal, and oral microbiota with nestling mass, mean length of the ninth primaries, and asymmetry of the ninth primaries, respectively. The significance of these associations was recovered even after accounting for the effects of 12 different temporal and environmental factors. Although variation in the microbiota comprised a less robust predictor than host age, sampling year, and ambient temperature for nestling morphology, inclusion of microbiota as an explanatory factor improved models describing morphological variation. Finally, **Chapter Five** analyzed the impact of a highly relevant ecological perturbation, an immune challenge, on the assembly of the gut microbiota and nestling development. Differences in both alpha and beta diversity of the gut microbiota were observed post-challenge with an inactivated viral vaccine, but no differences were observed in nestling morphology between treatment groups.

Implications

Prior to the work presented in **Chapter Two**, the general standard for sample preservation and long-term storage of fecal and swab samples involved dry storage immediately at -80°C . However, long field work seasons at sites with limited local amenities frequently make immediate sample storage at -80°C logistically prohibitive. The alternatives of bringing a vat of liquid nitrogen on all field endeavors or choosing only study sites with ultracold freezers comprise disruptive restrictions that are counter to the spirit of scientific exploration and can limit both research questions posed as well as the discovery of species studied. The sample preservation approach developed in **Chapter Two** can be streamlined through the production of aliquots of DNA stabilization buffer in sample tubes before beginning field research. Furthermore, making use of more generally available equipment (i.e., household freezers) opens the possibility of working at study sites that would otherwise be intractable.

Additionally, previously recommended commercial kits for the extraction of DNA from swab samples consistently underperformed with regard to both DNA yield and quality. However, swab samples are the most reliable sample types that may be obtained from birds, as fecal voidance cannot be dependably elicited by researchers with a bird in hand. **Chapter Two** contributed a critically effective extraction method, essential to making swab use widely possible in this dissertation and likely useful for many other researchers who have encountered difficulty with DNA extraction from swab samples. Alongside detailed protocols for successful MPS characterization of microbiota in minimally invasive avian samples, **Chapter Two** offered comprehensive methodological groundwork to support feasible research in avian microbiology.

Insights encountered in **Chapter Three** regarding the spatiotemporal variation of avian-associated microbiota were surprisingly inconsistent with prior research in both avian and mammalian systems. Although much support for the primary role of environmental selection or

filtering in shaping variation in the avian gut microbiota was observed, the magnitude and directionality of effects for various environmental factors or shifts were not always predictable. For example, differentiation of the microbiota across sites throughout the gastrointestinal tract was more apparent during the early nestling period and declined with nestling age. Shared nest environment did not substantially contribute to higher similarity in microbiota among nestlings. Increased variation in ambient temperature during the nestling period was negatively associated with phylogenetic diversity and Shannon's diversity of the gut microbiota, and interannual variation in environmental condition appeared to be the most robust mediator of the composition of microbiota between individuals. The results suggest that variation in life-history strategies, such as those related to development (precocial versus altricial) and reproduction (monogamy versus extra-pair copulations), may play pertinent roles in shaping patterns of gut microbial variation both within-individuals with age and between-individuals among nests. More broadly, the findings point to the relevance of studying host-microbial relationships in natural systems to capture the wider scope of intra- and inter-individual variation in microbiota among various taxa.

Although microbial associations with nestling condition have been reported in previous studies, **Chapter Four** uniquely contextualized these associations among a number of additional relevant factors that are known to affect nestling growth. Gut microbiota appeared to play a significant but less robust role than other temporal (i.e., host age, sampling year) and environmental (i.e., ambient temperature) variables in mediating nestling morphology. Taken together with the results of **Chapter Three**, interannual variation and ambient temperature associate with both avian gut microbial diversity and nestling development. However, the potential exists for the avian gut microbiota to respond to host extrinsic environmental change and for these responses to have subsequent effects on host condition, independent of the direct effects of other spatiotemporal factors.

The results of **Chapter Five** suggest that although the avian gut microbiota respond to host intrinsic environmental change (i.e., immunological perturbation), the shifts appear either transitive or subtle with no measured effect on host development. These latter findings are consistent with either i) too short of a duration of an effect of the immune challenge on the microbiota for the manifestation of host morphological change and/or ii) functional redundancy in the gut microbiota whereby immunologically-mediated shifts in taxonomic composition or relative abundance of OTUs generate no detectable morphological effects. Interestingly, given the range of environmental selective pressures covered in **Chapters Three through Five**, host extrinsic environmental change at the San Joaquin Experimental Range appeared to harbor a greater capacity for impacting the gut microbiota and nestling development than host intrinsic environmental change through administration of the viral immune challenge.

Chapter Five also demonstrated the value of avian-associated microbiota as a tractable model system to pursue ecological questions with significantly broader relevance to society. Although centuries of research exist on the physiological effects of vaccine use, curiously no study (to the best of current knowledge) has investigated the impact of vaccines on host-associated microbiota. Therefore, **Chapter Five** makes a unique contribution as the first study to document shifts in the bacterial microbiota of multiple sites along the vertebrate gastrointestinal in response to the administration of a viral vaccine. This finding warrants more detailed studies on the mechanisms underlying the results and on the reproducibility of the

results in other vertebrate host taxa. Furthermore, given the ubiquitous use of vaccines in both human and veterinary preventative medicine, it is of interest to understand the extent to which these vaccine-mediated microbial shifts can lead to functional and biologically meaningful effects on host health.

Limitations

Although **Chapter Two** offered valuable methods to support massively parallel sequencing analyses of minimally invasive samples in avian microbiological research, the greatest limitations to the pipelines developed involve the use of singular 2 mL tubes during DNA extraction and 1.5 mL tubes during library preparation. It is possible to process 48 extraction samples or 24 libraries at a time using these tubes. However, there is room for further optimization using large volume, 96-well plates as well as expandable multichannel pipettes for increased high-throughput capacity and scalability.

Chapter Three incorporated broad spatiotemporal sampling and non-microbial environmental factors to evaluate the relative significance of drivers structuring avian-associated microbial assemblages. Sampling year and temperature fluctuations figured prominently among significant explanatory variables, which is both intriguing and concerning given the increasing prevalence of human environmental impacts and global change. Although the interannual comparisons generated some of the most interesting results of the chapter, they were limited by the three-year scope of the dissertation work. Longer-term datasets that continuously record variation in environmental parameters combined with repeated sampling of the same hosts are needed to better clarify whether microbial signatures associated with different years persist throughout the lifetime of the host individuals or rather track volatile seasonal and annual shifts in the environment.

Much work also remains to be done on the potential for gut microbiota to affect host condition and fitness. **Chapter Four** was limited as an observational study, and thus all results were purely correlative rather than indicative of causality. Additionally, the immune challenge in **Chapter Five** comprised just one of many more robust environmental factors of interest that may impact gut microbiota and host condition. In considering the assessment of host condition, growth performance remains controversial with regard to what values form the “ideal” growth rate for maximum fitness. For example, does maximal weight gain or adiposity during the nestling period indicate greater survival, or does allocation of energy to other developmental processes such as immune function also contribute to maximal survival likelihood and thus confound the assumption that a higher growth performance index is advantageous? Nevertheless, the results of this dissertation generate interesting hypotheses on the relevance of gut microbiota in years of favorable versus adverse environmental conditions for host growth performance. Experimental studies with food supplementation, temperature manipulations, and physiological tests of flight performance with simultaneous characterization of the gut microbiota would be fascinating to parse the implications of this research.

Exciting avenues for future research

This dissertation has revealed high variation in gut microbiota among nestlings, whereby the microbiota appear to respond to environmental heterogeneity and perturbation as well as associate with nestling morphological variation. Despite these contributions, large gaps remain

in our knowledge on the connectivity between natural variation in avian microbiota, their physiological impacts and their broader implications for ecological and evolutionary processes. To systematically discuss these open areas for research, an intriguing, recently published list of 100 fundamental yet unanswered ecological questions (Sutherland *et al.* 2013) is referenced. This list constitutes an ideal scaffold against which to consider the relevance of host-associated microbiota and the potential for avian study systems to address these outstanding questions.

The first category of questions in The List covers ecology and evolution. The most relevant question involves how “evolutionary and ecological theory [should] be modified for organisms where the concepts of individual and fitness are not easily defined” (Sutherland *et al.* 2013). For example, 100 trillion bacteria inhabit humans, and the collective bacterial genome, also known as the microbiome, contains 100 times more genes than the human genome (Gill *et al.* 2006). This metaorganismal or holobiont biological unit has increasingly taken precedence in the concept of the 'individual'. Zilber-Rosenberg and Rosenberg (Zilber-Rosenberg & Rosenberg 2008) have posited the hologenome theory of evolution in which microbiota are integrated into the evolution of animals and plants. That is, the host and all of its microbiota comprise a unit of selection in evolution. The theory incorporates Lamarckian aspects in that the microbes are acquired and lost within the host organism’s life through use and disuse. Furthermore, vertical transmission of microbiota supports inheritance of the acquired microbiota between generations (Rosenberg *et al.* 2009). However, the theory is contextualized within a Darwinian framework and group selection; cooperation between all members of a holobiont contributes to the fitness of the host organism in its environment. To address the core of this theoretical consideration, much work remains to be done to clarify the “relative contributions of different levels of selection (gene, individual, group) to life-history evolution” (Sutherland *et al.* 2013). Birds exhibit fascinatingly diverse life-history traits that span various dietary (e.g. Snow 1981; Robinson & Holmes 1982), developmental (Starck & Ricklefs 1998), physiological (e.g. Olson & Owens 2005), reproductive (Jetz *et al.* 2008), migratory (Greenberg & Marra 2005) and behavioral (Revis *et al.* 2004) strategies. Accordingly, through comparative studies including multiple species, birds comprise a fabulous study system to begin understanding the role that microbiota and various levels of selection play in life-history evolution.

The second category of questions on The List encompasses populations. Very little is known regarding the mechanisms that mediate how processes occurring at the level of individual microbes generate patterns observed at the microbial or host population scale. For instance, microbiota are known to play an influential role in the epigenetics of host intestinal epithelial cells (IECs). Although trillions of bacteria come in direct contact with the IECs, excessive immunological response to these bacteria is not observed due in part to methylation and decreased expression of Toll-like receptor (TLR) genes in the IECs (Takahashi *et al.* 2011). IECs in germ-free mice exhibit significantly lower levels of TLR methylation and thus higher expression of TLRs than conventional mice (Takahashi *et al.* 2011). Similarly, microbial metabolites such as short-chain fatty acids have been shown to act as histone deacetylase inhibitors, involved in epigenetic regulation of gene expression (Davie 2003; Waldecker *et al.* 2008). However, we currently do not know how intra- and inter-individual variation in microbial mediation of epigenetic effects may impact host population dynamics. Ornithologists have long investigated the genetic basis of complex traits such as song and plumage polymorphism, which often mediate social interactions and reproductive success in birds (Podos *et al.* 2004; Hill &

McGraw 2006). Given the numerical and genetic predominance of microbes within birds, the possibility that loci interact, both within the avian genome as well as among avian and microbial genomes, warrants greater investigation. Epistatic as well as direct interactions between birds and microbes may account for much of the variation in complex traits that drive dynamics at the avian population level.

A third set of questions on The List focuses specifically on disease and microorganisms. After considerable technological advances in characterizing genotypic and taxonomic diversity in host-associated microbiota, the relationship between genotypic, taxonomic, and functional diversity is still strikingly underexplored. Understanding these relationships can be complicated by the prevalence of horizontal gene transfer among bacterial lineages and the technical challenges involved with linking specific gene expression and function to individual taxa within complex microbial assemblages. However, methods in single-cell genomics and transcriptomics are currently being developed to facilitate such research (Shapiro *et al.* 2013; Streets *et al.* 2014). With regard to function, microbiota present a unique opportunity to vet whether ecological principles derived from macroorganismal studies may operate on the microorganismal scale or vice versa. As an example, niche conservatism has been invoked to explain a number of species distribution patterns for macroorganisms (Wiens *et al.* 2010). Gut microbiota are known to be composed of a few deep lineages that diversify prolifically at phylogenetically shallow levels. It would be interesting to analyze functional data in conjunction with phylogenetic data for a given host-associated microbial assemblage to gain insight on the relevance of niche conservatism for the assembly and maintenance of the microbiota. Furthermore, functional data will help ornithologists interpret whether variation in microbial genes and taxa among birds can indicate biologically significant processes underlying variation in host susceptibility to disease or reservoir competence for zoonotic pathogens.

A fourth category of questions on The List encompasses communities and diversity. The relative extent of stochastic versus deterministic processes in driving observed patterns of microbial diversity remains to be clarified. Even among potential deterministic factors, much emphasis has been placed on certain factors such as diet whereas the combined effects of weak or complex interactions, such as interspecific interactions among microbes or between hosts and microbes, are more difficult to measure and study. Furthermore, environmental variation across various spatial and temporal scales abounds in natural settings. Nevertheless, little work has been done to clarify the mechanisms that may mediate how host extrinsic or intrinsic environmental heterogeneity influences microbial diversity and composition across various body sites and time points within individuals, within species or between species. Such knowledge is increasingly important in the face of global change and ties into a related section on The List concerning human impacts. Given rapid environmental change, “what determines whether species adapt, shift their ranges or go extinct” (Sutherland *et al.* 2013)? Are these concerns relevant to host-associated microbiota? It might be anticipated that microbiota occurring on the host exterior, such as skin or feathers, may be more susceptible to rapid environmental change than those residing in the host interior, such as the gut. Likewise, microbiota of altricial nestlings during the early nestling period may experience environmental fluctuations more than that of precocial nestlings. The extent to which anthropogenic impacts may affect the composition of avian-associated microbiota and whether such effects may have

functional implications for avian fitness remain highly relevant but presently unexplored areas for research.

Lastly, another set of questions on The List considers ecosystems and their functioning. The importance of rare species in ecosystem function is an interesting issue. Within microbial metagenetic datasets obtained through massively parallel sequencing, it has become the norm to discover rare species, often represented by only a single sequence and referred to as 'singletons' (Huse *et al.* 2013). These singletons are usually assumed to be technical artifacts and therefore typically removed from downstream data analyses. However, rare species can sometimes play a significant role in ecosystem function, such as superspreaders in disease ecology (Paull *et al.* 2012) or contributors to ecosystem stability (Tilman *et al.* 2006). It raises the question of whether greater effort should be expended to explore the functional implications of rare members of the microbiota, especially in the context of host fitness. The diversity of low-abundance microbes associated with birds combined with the largely unknown fitness consequences of these avian-microbial interactions demand greater research attention to build a more integrative understanding of avian biology, ecology and evolution.

To address the above unresolved research directions, an integrative approach could be developed to characterize taxonomic and functional diversity of host-associated microbiota combined with phenotypic and fitness measures across several populations, host species and spatiotemporal scales. Many ornithologists have in place extensive ongoing research efforts, including long-term ecological research sites or annual museum collection trips, in which phenotypic and fitness data are already being collected. Supplementing these efforts with microbial sampling would require some changes to standard operating procedures but would yield a wealth of valuable data to robustly investigate the implications of avian-microbial interactions as well as to inform broader ecological and evolutionary theory. The avenues presented in this section only just begin to cover the plethora of open research areas in microbial ornithology. Despite the brevity, it is hoped that the ideas will inspire additional perspectives and research endeavors.

Conclusion

Avian microbiology presently sits at an exciting precipice: the tools are available to make great strides in advancing current knowledge, yet unaddressed questions and unexplored research directions abound. This dissertation makes a number of contributions to increase awareness of avian microbiology and to expand upon current knowledge grounding the field. Namely, high variation in gut microbial composition exists both within and among individuals. The gut microbiota appear responsive to environmental change, and they are significantly associated with host morphological variation. These results raise compelling questions regarding how this standing yet somewhat plastic variation in gut microbiota may contribute to differences in host fitness among individuals, especially in the face of global change, and mediate the evolution of host species.

Avian species, especially those that utilize nest boxes such as bluebirds, swallows, flycatchers, titmice, wrens, sparrows, and finches, make excellent systems to carry out experimental manipulations in free-living populations of vertebrates. They span a wide range of interesting life-history traits, and importantly, they enable the collection of repeated measures or the administration of repeated manipulations to the same individuals, which is often

prohibitively difficult in other study systems and which would often substantially restrict the scope of experimental designs implemented. Although avian physiology differs in marked ways from mammalian physiology, these differences also provide the valuable opportunity to vet hypotheses and generalities derived from mammalian systems. In doing so, avian study systems present the ability to broaden standing conceptual frameworks to include greater taxonomic coverage. Altogether, the results of this dissertation emphasize the power of the field of avian microbiology as a novel system to investigate questions in ecology and evolution, and these findings portend an exciting future for this rapidly emerging field.

References

- Bosch TCG, McFall-Ngai MJ (2011) Metaorganisms as the new frontier. *Zoology (Jena, Germany)*, **114**, 185–190.
- Snow DW (1981) Tropical frugivorous birds and their food plants: a world survey. *Biotropica*.
- Davie JR (2003) Inhibition of histone deacetylase activity by butyrate. *The Journal of Nutrition*, **133**, 2485S–2493S.
- Gill SR, Pop M, DeBoy RT *et al.* (2006) Metagenomic analysis of the human distal gut microbiome. *Science*, **312**, 1355.
- Greenberg R, Marra PP (2005) *Birds of Two Worlds: The Ecology and Evolution of Migration*. JHU Press.
- Hill GE, McGraw KJ (2006) *Bird Coloration: Function and evolution*. Harvard University Press.
- Huse HM, Welch DBM, Sogin ML (2013) Sequencing errors, diversity estimates and the rare biosphere. In: *The Science and Applications of Microbial Genomics*, pp. 188–207.
- Jetz W, Sekercioglu CH, Böhning-Gaese K (2008) The worldwide variation in avian clutch size across species and space. *PLoS Biology*, **6**, e303.
- Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of Mammals and Their Gut Microbes. *Science*, **320**, 1647–1651.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3229–3236.
- Olson VA, Owens IPF (2005) Interspecific variation in the use of carotenoid-based coloration in birds: diet, life history and phylogeny. *Journal of Evolutionary Biology*, **18**, 1534–1546.
- Paull SH, Song S, McClure KM *et al.* (2012) From superspreaders to disease hotspots: linking transmission across hosts and space. *Frontiers in Ecology and the Environment*, **10**, 75–82.
- Podos J, Huber SK, Taft B (2004) Bird song: the interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics*, **35**, 55–87.
- Revis HC, Waller DA, Brittingham M (2004) Bactericidal and fungicidal activity of ant chemicals on feather parasites: an evaluation of anting behavior as a method of self-medication in songbirds. *The Auk*, **121**, 1262–1268.
- Ridaura VK, Faith JJ, Rey FE *et al.* (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, **341**, 1241214.
- Robinson SK, Holmes RT (1982) Foraging behavior of forest birds: the relationships among search tactics, diet, and habitat structure. *Ecology*, **63**, 1918.

- Rosenberg E, Sharon G, Zilber-Rosenberg I (2009) The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework. *Environmental Microbiology*, **11**, 2959–2962.
- Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics*, **14**, 618–630.
- Starck JM, Ricklefs RE (1998) *Avian Growth and Development: Evolution Within the Altricial-precocial Spectrum*. Oxford University Press.
- Streets AM, Zhang X, Cao C *et al.* (2014) Microfluidic single-cell whole-transcriptome sequencing. *Proceedings of the National Academy of Sciences*, **111**, 7048–7053.
- Sutherland WJ, Freckleton RP, Godfray HCJ *et al.* (2013) Identification of 100 fundamental ecological questions. *Journal of Ecology*, **101**, 58–67.
- Takahashi K, Sugi Y, Nakano K *et al.* (2011) Epigenetic control of the host gene by commensal bacteria in large intestinal epithelial cells. *Journal of Biological Chemistry*, **286**, 35755–35762.
- Tilman D, Reich PB, Knops JMH (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature*, **441**, 629–632.
- Turnbaugh PJ, Hamady M, Yatsunenko T *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature*, **457**, 480–484.
- Waldecker M, Kautenburger T, Daumann H, Busch C, Schrenk D (2008) Inhibition of histone-deacetylase activity by short-chain fatty acids and some polyphenol metabolites formed in the colon. *The Journal of Nutritional Biochemistry*, **19**, 587–593.
- Wiens JJ, Ackerly DD, Allen AP *et al.* (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.
- Zilber-Rosenberg I, Rosenberg E (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews*, **32**, 723–735.