ABSTRACT

THE EFFECTS OF STRESS ON AVIAN GENE EXPRESSION AND MICROBIOMES

by

Daniel J. Newhouse

July 2019

Director of Dissertation: Christopher N. Balakrishnan, Department of Biology

Major Department: Interdisciplinary Doctoral Program in Biological Sciences

ABSTRACT

Organisms must cope with stressors throughout their lifetime. Stressors are broad and come from both intrinsic and extrinsic sources. In this era of rapid global change, exposure to stressors will become more unpredictable and frequent. Thus, it is imperative we begin to understand how organisms respond to these stimuli. Birds are useful models to investigate responses to stress, as they are a diverse taxonomic group encompassing both well studied ecological and laboratory models. Here, I use several bird species and sources of stress to investigate how birds respond from a transcriptomic and microbiome perspective.

The first two chapters investigate how infection and androgens impact gene expression of the avian immune system. In Chapter 1, I explore the avian immune transcriptomic response to West Nile virus, a common avian pathogen that has had devastating effects on birds since its emergence in North America. In Chapter 2, I test the Immunocompetence Handicap Hypothesis, which states that androgen-dependent sexually selected traits are costly to produce because

androgens are simultaneously immunosuppressive. I used previously published transcriptome data and found support for the hypothesis.

In Chapters 3 & 4, I turn my focus to early-life stress and white-throated sparrows (WTSPs). WTSPs exhibit two stable alternative parental care strategies, biparental care and female-biased parental care. In Chapter 3, I show that female-biased care induces a transcriptomic stress response in nestlings. In Chapter 4, I investigate the microbiome of WTSP nestlings and find no differences between individuals raised under the different parental care strategies. However, I do find some evidence of host genetic control of the microbiome.

Overall, I have begun to explore how birds cope with various stressors, including infection, androgen induced immunosuppression, and developmental stress. Through primarily gene expression approaches, I uncover the molecular pathways affected by these stimuli. Each chapter will set the stage for future integrative work to explore organismal responses to their environment.

THE EFFECTS OF STRESS ON AVIAN GENE EXPRESSION AND MICROBIOMES

A Dissertation

The Faculty of the Interdisciplinary Doctoral Program in Biological Sciences,

The Brody School of Medicine

In Association with the Department of Biology,

Thomas Harriot College of Arts and Sciences

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Interdisciplinary Doctoral Program in Biological Sciences

by

Daniel J. Newhouse

July, 2019

THE EFFECTS OF STRESS ON AVIAN GENE EXPRESSION AND MICROBIOMES

by

Daniel J. Newhouse

APPROVED BY:

DIRECTOR OF DISSERTATION: _____
Christopher N. Balakrishnan, Ph.D.

COMMITTEE MEMBER: _____
Michael S. Brewer, Ph.D.

COMMITTEE MEMBER: _____
Susan B. McRae, Ph.D.

COMMITTEE MEMBER: _____
Ariane L. Peralta, Ph.D.

COMMITTEE MEMBER: _____
Garth Spellman, Ph.D.

CHAIR OF THE DEPARTMENT OF BIOLOGY: _____
Cindy Putnam-Evans, Ph.D.

DEAN OF THE GRADUATE SCHOOL: _____
Paul J. Gemperline, PhD

ACKNOWLEDGEMENTS

This work would not have been possible without the help of many colleagues, collaborators, and co-authors. I could not have done much of the white-throated sparrow work without substantial help from Elaina Tuttle, Rusty Gonser, Margarida Barcelo-Serra, and Lindsay Forrette. Thank you for answering my many emails, sending me samples, teaching me the ways of the sparrow, and for being excellent hosts and friends during fieldwork at Cranberry Lake. I hope to never step foot in that bog again. I also must thank Erik Hofmeister, Marty Martin, and Meredith Kernbach for being amazing collaborators on the West Nile virus projects. Lastly, thank you to Ben Vernasco for turning a lunchtime discussion into a manuscript.

I am also grateful for my committee members and their time and effort devoted towards my research. I thank Sue McCrae for teaching me much of what I know about birds, Ariane Peralta for imparting microbiome wisdom, and Michael Brewer for substantial bioinformatics help over the years. Lastly thanks to Garth Spellman, for taking the time to Skype and provide feedback as my external committee member.

The ECU Department of Biology has been a wonderful home. Many faculty, particularly Terry West, Ed Stellwag, Xiaoping Pan, Margit Schmidt, and Kyle Summers, have been extremely helpful in my development as a scientist. Many graduate students made this journey much more enjoyable and I am thankful for all the friendships. In particular, I must thank past and present Balalab members Allison Lansverk, John Davidson, Dustin Foote, and Robert Driver. Lastly, I want to acknowledge the ECU Biology staff that really made much of this possible: Joyce Beatty, Tammy McKinney, and Jen Jacobs. Thank you for always fighting for the students.

Thank you to my family for being so supportive during graduate school. I am especially thankful for my wife, Kelly, and our dog, Beau, for being there through all the ups and downs. Lastly, I want to thank my advisor, Chris Balakrishnan. I cannot imagine going through graduate school without you as my advisor. Thank you for bringing me in as a Master's student and teaching me everything from the start. And thank you again for believing in me and encouraging me to pursue a PhD. You have been supportive since day one, even when I deleted a hard drive and when my eye swelled shut in Taiwan. I am forever grateful for everything you have done to help me over the years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. TRANSCRIPTIONAL RESPONE TO WEST NILE VIRUS INFECTION IN THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*)

Previously published in
Royal Society Open Science, 2017, 4(6), 170296

Abstract

West Nile virus (WNV) is a widespread arbovirus that imposes a significant cost to both human and wildlife health. WNV exists in a bird-mosquito transmission cycle in which passerine birds act as the primary reservoir host. As a public health concern, the mammalian immune response to WNV has been studied in detail. Little, however, is known about the avian immune response to WNV. Avian taxa show variable susceptibility to WNV and what drives this variation is unknown. Thus, to study the immune response to WNV in birds, we experimentally infected captive zebra finches (*Taeniopygia guttata*). Zebra finches provide a useful model, as like many natural avian hosts they are moderately susceptible to WNV and thus provide sufficient viremia to infect mosquitoes. We performed RNAseq in spleen tissue during peak viremia to provide an overview of the transcriptional response. In general, we find strong parallels with the mammalian immune response to WNV, including up-regulation of five genes in the Rig-I-like receptor signaling pathway, and offer insights into avian specific responses. Together with complementary immunological assays, we provide a model of the avian immune response to WNV and set the stage for future comparative studies among variably susceptible populations and species.

Introduction

West Nile virus (WNV) is a single-stranded RNA flavivirus that exists in an avian-mosquito transmission cycle, where birds (typically Passeriformes) act as the primary amplification hosts. In addition to birds, nearly 30 other non-avian vertebrate species have been documented as hosts (1). Although many WNV-infected hosts are asymptomatic, WNV infection can cause severe meningitis or encephalitis in those that are highly susceptible. Avian species for the most part exhibit low to moderate susceptibility. That is, individuals become infected and develop sufficient viremia for transmission via mosquito blood meal, but the hosts recover and avoid significant mortality (reviewed in (2)). First described in 1937, WNV has not resulted in widespread avian decline throughout its historical range (3), perhaps due to host-parasite coevolution. However, the emergence of WNV in North America in 1999 has negatively impacted a wide range of populations (4,5). Surveys of North American wild birds have shown a variety of competent WNV hosts, with varying degrees of susceptibility, morbidity, and pathogenicity (2). American robins (*Turdus migratorius*) appear to be the main host in spreading WNV infection in North America (6), but infection appears most detrimental to members of Family Corvidae (7). Despite great variation in susceptibility, the mechanisms underlying this variation are primarily unknown (2).

Largely due to interest in human health implications, most work describing the host immune response to WNV infection has been performed in mammalian systems (8). From these studies, we know that in mammals, both the innate and adaptive arms are critical for virus detection and clearance (9,10). Within the innate immune response, the retinoic acid-inducible gene 1 (Rig-I)-like receptor (RLR) pathway appears to play a key role in viral clearance. This pathway recognizes viral products and initiates type I interferon expression (11). Mice lacking

the viral recognition RLR genes in this pathway, DDx58 (Rig-I) and IFIH1 (MDA5), become highly susceptible to WNV infection (12). In the adaptive immune system, a broad range of components appear to play important roles in mounting a response, including antibody and CD4+ and CD8+ T cells (9,13,14). Interestingly, major histocompatibility complex (MHC) class I genes are up-regulated post-infection (15,16). Viruses typically evade MHC class I detection (17,18), as MHC class I molecules bind and present viral peptides to CD8+ T cells. However, the purpose of WNV induced MHC expression is unclear.

While the mammalian immune response to WNV infection has been extensively studied, the avian immune response remains mostly unknown. Of the studies in birds, many involve experimentally infecting wild caught birds (reviewed in (2)), or domestic chickens (*Gallus gallus*) (19). These studies primarily focus on viral detection, tissue tropism, antibody production, or lymphocyte counts (2,19,20). Little is known about the molecular mechanisms driving the immune response to WNV infection (but see (21)). Furthermore, current avian WNV studies suffer many challenges. Wild caught birds may be co-infected with other parasites (e.g. avian malaria) and are difficult to maintain in captivity for experimental infection studies. Chickens, although an avian model species, are uncommon hosts and highly resistant to WNV infection (22). Therefore, chickens are not ideal to describe the avian immune response to WNV infection. Passeriformes and Galliformes are also highly divergent bird lineages, with distinctive immune gene repertoires and architecture (23).

As passerine birds are the main hosts for WNV, we have sought to develop a passerine model to study the impacts of WNV infection on a taxonomically appropriate host (24). We have recently shown that zebra finches, *Taeniopygia guttata*, are moderately susceptible hosts for WNV (25). That is, WNV rapidly disseminates to a variety of tissues and is detectable in most

samples by four days post-inoculation (dpi). Despite rapid development of sufficient viremia for arthropod transmission, zebra finches develop anti-WNV anitbodies, clear WNV by 14dpi, and avoid significant mortality (25). This moderate disease susceptibility is similar to what is observed in many natural WNV hosts. Zebra finches are also an established biomedical model system with a suite of genetic and genomic tools available (26).

In this study, we experimentally infected zebra finches and performed RNAseq to describe their transcriptional response up to the point of peak viremia. In doing so, we characterize the zebra finch immune response to WNV infection, explore expression of the avian RLR pathway in response to WNV, gain insights into the avian immune response to this widespread infectious disease, and uncover conserved evolutionary responses in avian and mammalian systems.

Results

Experimental infection

We challenged six individuals with 10^5 plaque forming units (PFU) WNV and sequenced RNA (Illumina RNAseq) isolated from spleens, an organ critical to the avian immune response. Three birds served as procedural controls and on day 0 were injected subcutaneously with 100 µL of BA1 media, as previously described (27). Peak viremia occurs at 4.6 ±1.7 dpi as quantified via RT-PCR (25) and thus, we characterized the transcriptional response leading to (2dpi, n = 3) and at peak viral load (4dpi, n = 3) in the present study. WNV RNA was detected by culture in lung and kidney RNA pools of 2 out of 3 birds sampled at day 2, and all 3 birds sampled at 4dpi. These findings were verified by semi-quantitative RT-PCR. Because WNV is rarely detected in spleen by 2dpi, but all birds previously inoculated at 10^5 PFU developed WNV antibodies [25], we treated all six birds inoculated with WNV as being infected.

Sequencing results & read mapping

We obtained 18-30 million paired-end, 100bp reads for each sample and removed 0.57-1.24% of the total bases after adapter trimming (Supplementary Table S1). On average, 79.0-80.8% total trimmed reads mapped to the zebra finch reference genome (Supplementary Table S2), corresponding to 18,618 Ensembl-annotated genes (28). Of these, 14,114 genes averaged at least five mapped reads across all samples and were used for differential expression (DE) analyses.

Sample clustering & differential expression

We tested for DE two ways: as pairwise comparisons between treatments to identify specific genes with *DEseq2* (29) and as a time-course grouping genes into expression paths with *EBSeqHMM* (30). To visualize patterns of expression variation among samples, we conducted principal component analysis (PCA) and distance-based clustering (Supplemental Figures S1 & S2). The first three principal components explained 93.04% of the variance in gene expression, but none of the PCs were significantly correlated with treatment (ANOVA, PC1: $p = 0.288$, PC2: $p = 0.956$, PC3: $p = 0.202$).

Although clustering analyses suggest that across the genome, much of the variation in expression was independent of the experimental treatment, pairwise comparisons revealed many genes that were regulated in response to infection (Supplementary Table S3). When comparing Control vs. 2dpi, we found 161 differentially expressed genes (FDR < 0.10, average $\log_2$ fold-change (FC) = 1.74). This gene list includes several immune related genes associated with the innate (e.g. IL18) and adaptive (e.g. MHC IIB) immune system (Table 1.1, Figure 1.1). Sixty-

five genes were differentially expressed between Control and 4dpi (average $log_2FC = 1.61$), also with several immune relevant genes including five genes in the RLR pathway (Table 1.1, Figure 1.2, Figure 1.3). Lastly, we observed 44 DE genes between 2dpi vs. 4dpi individuals (average $log_2FC = 1.56$). Three of these have described functions in immunity. We also combined 2dpi and 4dpi cohorts and compared with control, but due to high variation in gene expression between days 2 and 4 dpi, we only found 16 DE genes (average $log_2FC = 1.64$) between Control and Infected cohorts, one of which was associated with immunity.

Table 1.1 **Candidate immune genes differentially expressed in the present study and comparisons with mammals.**

Ensembl ID	Gene Name	Log_2 Fold Change	FDR	Regulation Pattern Observed	Regulation Pattern in Mammals	Reference
Control vs Infected						
ENSTGUG00000013615	NFKBIZ	0.73	0.064	Up	Up	43
Control vs 2dpi						
ENSTGUG00000000297	IL18	1.01	0.010	Up	No change	49
ENSTGUG00000000678	TIM1	1.49	7.99E-05	Up	Up	39,40
ENSTGUG00000001485	IRF6	-2.09	0.037	Down	Up	41
ENSTGUG00000003354	NKRF	-2.35	4.85E-05	Down	Unknown	
ENSTGUG00000005295	C-C motif chemokine	2.08	0.007	Up	Up	41,42,43
ENSTGUG00000008638	UCHL1	-1.91	0.029	Down	Unknown	
ENSTGUG00000008991	APOD	1.76	0.053	Up	Up	44
ENSTGUG00000009454	IFITM10	1.24	2.71E-04	Up	Unknown	
ENSTGUG00000009769	TNFRSF13C	0.89	0.010	Up	Unknown	
ENSTGUG00000015634	Novel gene (MHC IIB)	2.23	0.001	Up	Up	46
ENSTGUG00000016383	SIGLEC1	1.39	0.046	Up	Up	43
ENSTGUG00000017149	Novel gene (MHC IIB)	1.57	0.099	Up	Up	46
Control vs 4dpi						
ENSTGUG00000001516	DDx58	1.50	1.45E-08	Up	Up	43
ENSTGUG00000002144	IRF4	1.39	0.022	Up	Up	41
ENSTGUG00000002305	LY86	-1.07	1.70E-05	Down	Unknown	
ENSTGUG00000002516	DHx58	1.67	4.05E-06	Up	Up	43
ENSTGUG00000004105	ADAR	1.13	1.70E-05	Up	Up	45
ENSTGUG00000006914	IFIH1	0.95	0.093	Up	Up	43
ENSTGUG00000007454	TNFRSF13B	1.61	0.010	Up	Unknown	
ENSTGUG00000008354	IFIT5	2.97	1.07E-09	Up	Unknown	
ENSTGUG00000008788	EIF2AK2	2.15	9.86E-07	Up	Up	41,76
ENSTGUG00000009162	PXK	1.19	0.011	Up	Unknown	
ENSTGUG00000009536	TRIM25	1.24	0.001	Up	Up	43
ENSTGUG00000009838	IRF7	0.80	0.047	Up	Up	43
ENSTGUG00000011784	ZC3HAV1	1.31	7.09E-08	Up	Up	43
ENSTGUG00000017534	MOV10	1.29	0.017	Up	Up	45
2dpi vs 4dpi						
ENSTGUG00000003354	NKRF	1.70	0.033	Up	Unknown	
ENSTGUG00000005206	ADA	-1.48	0.001	Down	Unknown	
ENSTGUG00000011784	ZC3HAV1	0.82	0.058	Up	Up	43

Figure 1.1. **Immune genes differentially expressed between day 2 post-inoculation and control** A) Heatmap of expression levels (log transformed read counts) across all treatments of immune genes differentially expressed at 2dpi relative to control. B-D) Expression values (normalized read counts) for three key immune genes and their regulation pattern classification by *EBSeqHMM*. Asterisks represent statistical significance in *DEseq2* analysis after FDR correction (* p<0.10, ** p<0.05, *** p<0.01).

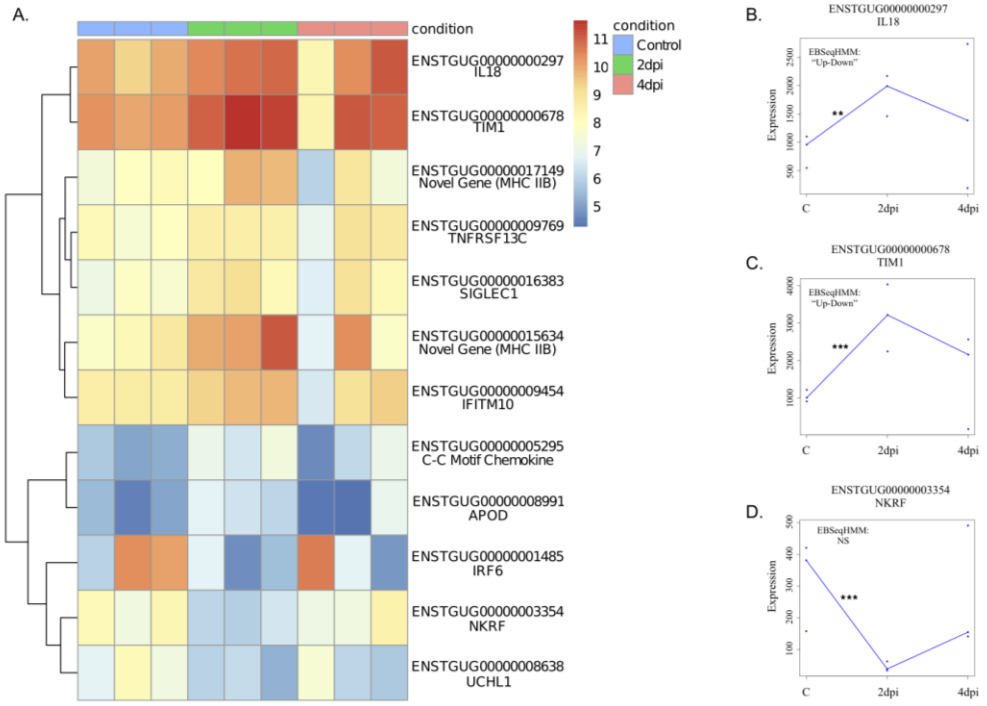
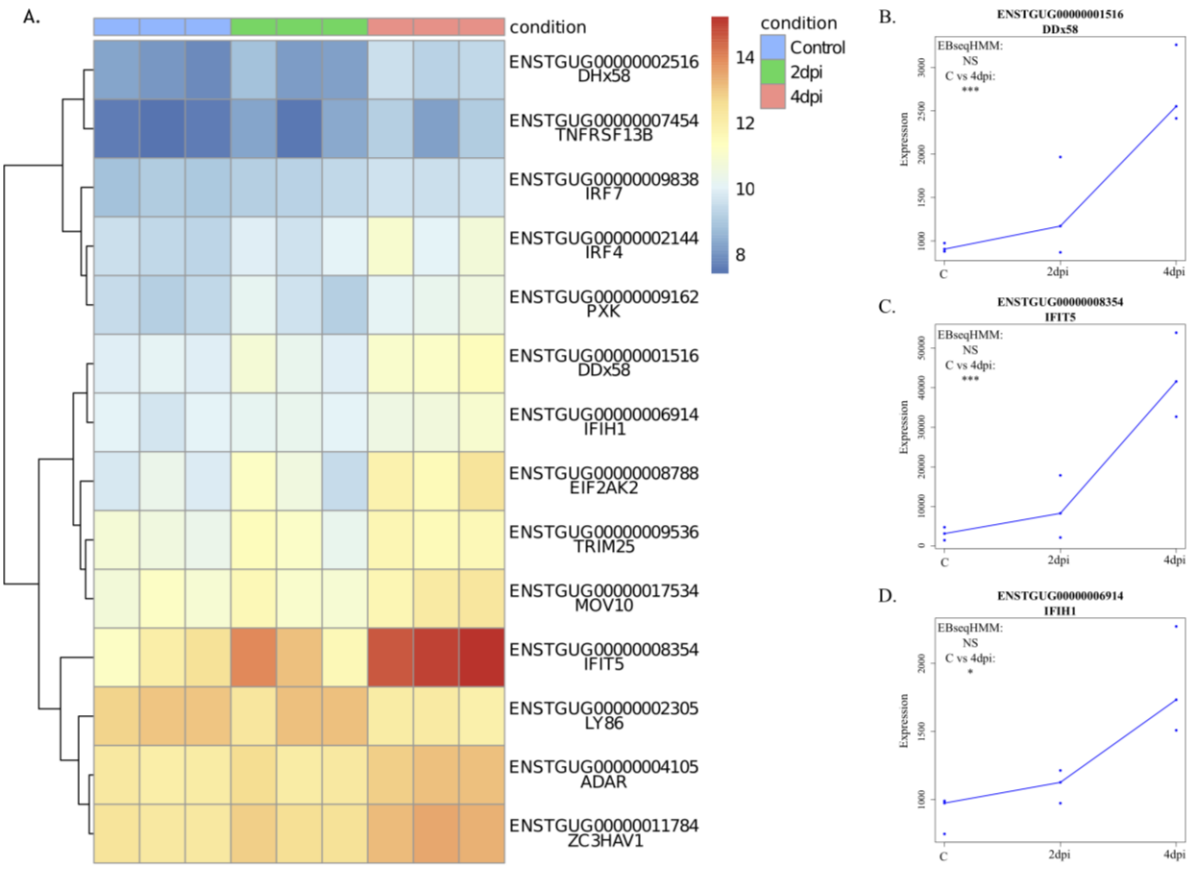


Figure 1.2. **Immune genes differentially expressed between day 4 post-inoculation and control A)** Heatmap of expression levels (log transformed read counts) across all treatments of immune genes differentially expressed at 4dpi relative to control. B-D) Expression values (normalized read counts) for three key immune genes and their regulation pattern classification by *EBSeqHMM*. Asterisks represent statistical significance in *DEseq2* analysis after FDR correction (* p<0.10, ** p<0.05, *** p<0.01).



When analyzed for DE as a time course in *EBSeqHMM*, 686 genes showed evidence of differential expression (posterior probability > 0.99, FDR < 0.01). Most DE genes (n = 561) were suppressed relative to controls a days 2 and 4 post infection ("Down-Down"). Seventy-five genes were "Up-Down", 49 were "Down-Up" and one was "Up-Up". As expected, we found overlap of several immune genes between the two analyses. For example, IL18, APOD and IFITM10 are "Up-Down" and this trend is reflected in the *DEseq2* Control vs 2dpi analysis (Figure 1.1).

Functional annotation of differentially expressed genes

To place differentially expressed genes into groups based on their biological function, we performed a gene ontology (GO) analysis using the *GOrilla* tool (31,32). GOrilla utilizes the

ranked order of genes from *DEseq2* based on FDR adjusted p-values. An enrichment score is calculated based on the number of genes in the top of the list that belong to a particular GO category relative to the expected number based on the frequency of functionally related genes in the total list. As above, we conducted GO analyses based on multiple pairwise analyses of gene expression. We found five significantly enriched GO categories between Control vs. Infected (2dpi and 4dpi) cohorts, of which "response to virus" is the most significant (FDR = 0.008, Enrichment = 5.34) (Table 2). We observed the strongest evidence of functional enrichment in the Control vs. 2dpi (n=120) and Control vs. 4dpi (n=36) contrasts (FDR < 0.05) (Supplemental Table S4). Many enriched GO terms in the Control vs 2dpi contrast are involved in membrane components, metabolism, and cellular processes. Four GO categories were immune relevant, including "inflammatory response" and "positive regulation of cytokine biosynthetic process" (Supplementary Table S4). The immune response manifests itself most strongly in the Control vs 4dpi contrast with many enriched GO terms being immune-related categories (Table 1.2, Supplementary Table S4) and a broad range of differentially expressed immune genes (n = 14, Table 1.1). Only two enriched GO categories are enriched between 2 and 4dpi: "inner mitochondrial membrane protein complex" and "mitochondrial protein complex" (Table 1.2, Supplementary Table S4).

Table 1.2. **Top five most significant gene ontology (GO) categories, FDR adjusted p-value, and GOrilla enrichment score, among *DEseq2* pairwise comparisons.** Enrichment is calculated as (b/n)/(B/N), where N is the total number of genes, B is the total number of genes associated with a specific GO term, n is the number of genes in the top of the input list, and b is the number of GO-term-associated genes in the top of the list (31,32).

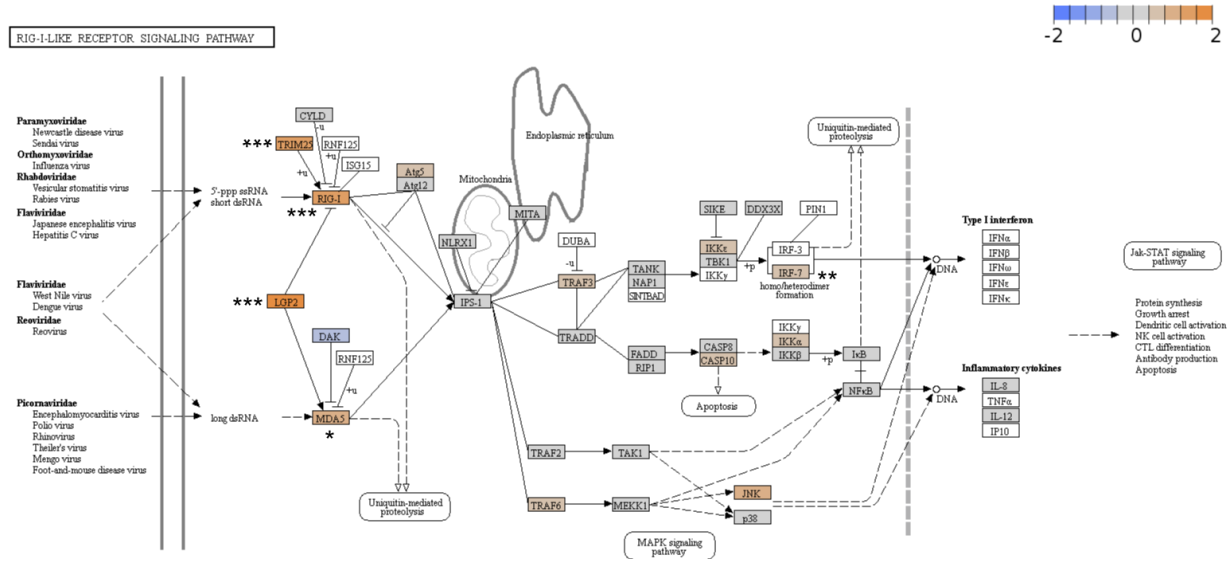
GO ID	Description	FDR	Enrichment
Control vs WNV			
GO:0009615	response to virus	0.008	5.34
GO:0051276	chromosome organization	0.019	1.97
GO:1903047	mitotic cell cycle process	0.035	1.69
GO:0034723	DNA replication-dependent nucleosome assembly	0.039	14.27
GO:0006335	DNA replication-dependent nucleosome organization	0.048	14.27
Control vs 2dpi			
GO:0044425	membrane part	1.77E-07	1.29
GO:0098800	inner mitochondrial membrane protein complex	1.15E-06	3.73
GO:0044459	plasma membrane part	2.66E-06	1.98
GO:0031224	intrinsic component of membrane	4.04E-06	1.35
GO:0098798	mitochondrial protein complex	4.06E-06	3.32
Control vs 4dpi			
GO:0009615	response to virus	0.001	19.23
GO:0051607	defense response to virus	0.002	51.27
GO:0060337	type I interferon signaling pathway	0.002	24.25
GO:0051707	response to other organism	0.002	11.73
GO:0098586	cellular response to virus	0.003	70.99
2dpi vs 4dpi			
GO:0098800	inner mitochondrial membrane protein complex	0.01	3.09
GO:0098798	mitochondrial protein complex	0.04	2.69

We also conducted a similar analysis of genes identified as DE by *EBseqHMM*, which revealed one (Up-Up), 199 (Up-Down), 69 (Down-Up) and 527 (Down-Down) significantly enriched GO categories (FDR < 0.05) (Supplementary Table S5). Interestingly, Up-Down GO categories had the strongest representation of immune related GO terms, including “immune

response" (FDR = 4.85 x 10^{-4}) and "negative regulation of immune system process" (FDR = 6.01 x 10^{-4}). Among Down-Down genes, we observed enrichment of many metabolic and membrane processes and only one immune related category ("positive regulation of innate immune response", FDR = 0.01). We find enrichment of mitochondrial components and processes among Down-Up genes, similar to the 2dpi vs 4dpi contrast in the *DEseq2* analysis. Additionally, ten categories involved in immunoglobulin processes were significantly enriched among "Down-Up" genes, driven by the presence of the joining chain of multimeric IgA and IgM (JCHAIN) gene. Lastly, as in the *DEseq2*-based analysis, we also detected a strong enrichment signature of membrane proteins. Genes annotated as "plasma membrane part" were highly enriched among those showing an Up-Down pattern (FDR = 1.61 x 10^{-12}, Supplementary Table S5). Combined, we find broad overlap in GO representation between the *EBseqHMM* and *DEseq2* approaches.

In addition to placing genes into broad systematic functions in the GO analysis, we were also interested in placing our gene expression results in the context of immune pathways of interest. The RLR antiviral pathway is critical to WNV clearance in mammals (12) and appears important in mounting an immune response to avian influenza in ducks (33-35). Utilizing *Pathview v1.8.0* (36), we find that WNV infection induces the RLR pathway. Five genes, including the two RLR genes, DDx58 and IFIH1, which encode the Rig-I and MDA5 viral detection molecules, are significantly up-regulated (Table 1.1, Figure 1.2, Supplementary Figure S3). We detect expression of 36/37 genes in the pathway, many of which are also up-regulated, though not always significantly (Figure 1.3).

Figure 1.3. Regulation of the zebra finch RLR pathway. Color represents $\log_2$ fold change between Control and 4dpi. Asterisks represent statistical significance in *DEseq2* analysis after FDR correction (* p<0.10, ** p<0.05, *** p<0.01).



Discussion

We have characterized the zebra finch transcriptional response to WNV infection. Overall, we find that as in mammalian systems, components of both the adaptive and innate immune pathways are activated following infection. While WNV is primarily an avian specific infectious disease, most work describing the host immune response to infection has been performed in mammals. Despite genomic, physiological and evolutionary differences between birds and mammals, the host immune response shows broad similarity between taxa (Table 1.1).

We were particularly interested in the role of the innate RLR pathway. This pathway mounts an antiviral innate immune response and is critical for WNV detection and clearance in mammals (12). We have shown here that the RLR pathway in zebra finches is induced by WNV infection. Five genes in this pathway are significantly up-regulated at 4dpi (Figure 1.3, Supplementary Figure S3), including DDx58 and IFIH1 (Figure 1.2B,D), which encode molecules that recognize WNV particles in mammals (37). This results in a corresponding over-

representation of genes in the interferon signaling and regulation GO categories (Table 1.2, Supplemental Table S4). While no studies have investigated the role of the RLR following WNV infection in birds, this pathway appears important for avian influenza clearance in ducks (33-35), Buggy Creek virus clearance in house sparrows (38), and likely for the broad avian antiviral immune response, including WNV. Interestingly, chickens (*Gallus gallus*), which are often used as sentinels for WNV, have lost the gene encoding the DDx58 RLR during their evolution (33) yet do not develop disease post WNV infection (22). This suggests that chickens respond to WNV using a Rig-I independent mechanism and highlights the importance of future work targeting the evolution of avian innate immunity.

We observed other parallels with mammals as well (Table 1.1). For example, T-Cell Immunoglobulin Mucin Receptor 1 (TIM1) is up-regulated at 2dpi in zebra finches (Figure 1.1A, C). In human cell lines, expression of TIM1 promotes infection of WNV virus like particles (VLPs) (39,40), suggesting that the up-regulation of TIM1 seen in zebra finches may promote viral entry as well. Similarly, C-C motif chemokine (ENSTGUG00000005295) is up-regulated in our study at 2dpi and in previous human cell line and mouse experiments, suggesting a conserved role in chemokine production following WNV infection (41-43). Apolipoprotein D (APOD), a gene typically involved in brain injury and potentially responding to the neurodegenerative nature of WNV, is up-regulated in WNV infected mice (44), as well as in our study. Two interferon stimulated genes (ISGs), ADAR and MOV10 are both significantly up-regulated at 4dpi relative to control. Schoggins et al. (45) showed ADAR expression to enhance WNV replication and MOV10 expression to have antiviral activity. While further testing of these genes is needed to validate their roles in avian WNV infection, they nonetheless offer insights into a broad range of conserved responses between mammals and birds.

Within the adaptive immune response, the role of the MHC in the host response to WNV is also particularly interesting. The MHC plays a key role in antigen processing and presentation. The MHC comprises two main gene families (Class I & II) and both are up-regulated in mammals following WNV infection (15,16,46). Similarly, two genes encoding MHC class IIB proteins are significantly up-regulated in zebra finches at 2dpi (Figure 1.1). Unlike mammals, however, we found that MHC class I is not significantly DE in any comparison (e.g. C vs 4dpi, $\log_2$FC = 0.001, FDR = 0.99). In mammals, upregulation of MHC class I may not be adaptive for the host, as upregulation may be a mechanism by which the virus evades Natural Killer (NK) cell detection by the innate immune system (15). It has also been suggested that MHC up-regulation is a byproduct of flavirus assembly (47). Interestingly, at 2dpi, interleukin-18 (IL18) is significantly up-regulated (Table 1.1, Figure1.1A, B). IL18 can enhance NK cell activity (48) and is potentially a mechanism by which the immune system can counteract WNV evasion strategies via NK cell activation, although further testing is needed to quantify NK cell activity in zebra finches to support this hypothesis.

Despite many similarities, several immune genes differentially expressed in our analyses have not been previously reported in the mammalian WNV literature or are expressed differently in zebra finches (Table 1.1). For example, at 2dpi, the proinflammatory cytokine IL18 was significantly up-regulated in zebra finches (Figure 1.1), contrasting a previous study in human cell lines, which show no difference in IL18 expression following WNV infection (49). Furthermore, interferon regulatory factor 6 (IRF6) was down-regulated at 2dpi, but up-regulated in human macrophages following infection (41). Another significantly down-regulated gene at 2dpi, ubiquitin carboxyl-terminal hydrolase L1 (UCHL1), has been previously associated with pattern recognition receptor (PRR) pathway (e.g. RLR) function in human cell lines infected

with high-risk human papilloma virus (50). When up-regulated, UCHL1 supresses PRR expression leading to viral evasion of the host immune response. However, down-regulation of UCHL1 restores functional PRR pathways (50). Thus, the down-regulation of UCHL1 2dpi in zebra finches may be associated with the up-regulation of the PRR RLR pathway in this study (Table 1.1, Figure 1.3). Lastly, interferon-induced protein with tetratricopeptide repeats (IFIT) and interferon-inducible transmembrane proteins (IFITM) gene families are known innate antiviral proteins and have been shown to restrict WNV entry in human cells lines (45, 51). Both IFIT5 and IFITM10 are up-regulated (Figure 1.2A, C) in our study and yet, to our knowledge, neither have previously been implicated in the WNV immune response. This potentially reveals an avian-specific function of IFIT5 and IFITM10. Lastly, several genes involved in metabolic and mitochondrial processes were DE in our analyses. Viral alteration of host metabolism typically benefits viral replication (52,53) and highlights the need for future work investigating the role of WNV on host physiology.

Functional enrichment of immune-related GO terms primarily appears in Up-Down path defined by *EBseqHMM* (Supplementary Table S5), as many genes in the immune system are up-regulated post-infection (Table 1.1, Figure 1.1, Figure 1.2). In both the *EBseqHMM* and *DEseq2* analyses, most of the significant immune GO categories are innate immune responses, although adaptive immune categories involved in immunoglobulin complexes and B & T cell proliferation appear in the EBseqHMM analysis (Table 1.2, Supplementary Tables S4 & S5). Similar to the mammalian model, broad organismal processes, encompassing both innate and adaptive immunity, are represented in the zebra finch response to WNV.

Like many passerine birds infected in nature, zebra finches are moderately susceptible to WNV, developing sufficient viremia to serve as competent hosts, but generally resist mortality

due to infection (25). While there are clear differences among treatments in terms of differentially expressed genes (Table 1.1), the modest effect of treatment on overall expression profile (Supplemental Figures S1 & S2) may be a reflection of this moderate susceptibility. Most zebra finches are able to clear WNV inflection by 14 dpi (25). WNV infection intensity varies among tissues (20), but due to the spleen's important role in the avian immune system (54,55) we expect the results presented here to be representative of the overall immune response. Although we expect to have missed some genes that are regulated in response to infection, *DEseq2* has been shown to perform very well (low false positive rate) in experiments with a sample size of three (56). Further studies will also be required to document more subtle, and tissue-specific patterns of gene regulation in response to infection. We note that we only sampled our control group at 4dpi and thus, do not have a direct procedural control at 2dpi. Changes in gene expression at 2dpi therefore could be in part due to the injection itself. Pronounced DE of immune-related genes at 2dpi, however, suggests that changes in gene expression were driven by WNV infection rather than by the injection, which might be predicted to trigger a more general stress response.

We have begun to develop the zebra finch as an avian model for the host response to WNV infection. We show here that in terms of gene expression, the zebra finch immune response is largely conserved with that seen in mammalian-based studies (Table 1.1). Additionally, we identify many components of the immune system that have not been previously implicated in the host immune response to WNV. This potentially reveals an avian-specific immune response and highlights avenues for future research. Combined with our recent immunological characterization (25), we have broadly described the immune response of a

moderately susceptible avian host for WNV. This sets the stage for future comparative work to uncover the genetic basis of variable avian susceptibility to WNV infection.

Methods

Experimental setup

All animal use was approved by the USGS National Wildlife Health Center Institutional Animal Care and Use Committee (IACUC Protocol: EP120521) and this study was performed in accordance with USGS IACUC guidelines. The experimental infection setup is described in detail in (25). Briefly, nine female zebra finches were randomly divided into three cohorts, one unchallenged and two challenged (n = 3 each). Birds were challenged subcutaneously with 100ul BA1 media containing 10^5 plaque-forming units (PFU) of the 1999 American crow isolate of WNV (NWHC 16399-3) and sacrificed at 2 and 4 dpi, corresponding to peak viremia. Uninfected individuals were injected with 100ul BA1 media and sacrificed at 4dpi. WNV infection was confirmed by RT-PCR, as previously described (26), in lung and kidney pooled tissue (25). Due to the critical role of the spleen in the initation of the immune response, and its common use in experimental infection gene expression studies (57-60), we focused our study on gene expression in the spleen. Spleens from each individual were removed, placed into RNAlater (Qiagen, Valencia, CA USA), and frozen at -80 °C until RNA extraction.

RNA extraction & sequencing

Whole spleen tissue was homogenized in Tri-Reagant (Molecular Research Company) and total RNA was purified with a Qiagen RNeasy (Valencia, CA USA) mini kit following the manufacturer's protocol. RNA was DNAse treated and purified. Purified RNA was quality

assessed on a Bioanalyzer (Agilent, Wilmington, DE USA) to ensure RNA quality before sequencing (RIN = 6.6-8.1). All library prep and sequencing was performed at the University of Illinois Roy J. Carver Biotechnology Center. A library for each sample was prepared with an Illumina TruSeq Stranded RNA sample prep kit. All libraries were pooled, quantitated by qPCR, and sequenced on one lane of an Illumina HiSeq 2000 with a TruSeq SBS Sequencing Kit producing paired-end 100nt reads. Reads were analyzed with Casava 1.8.2 following manufacturer's instructions (Illumina, San Diego, CA). Sequencing data from this study have been deposited in the NCBI Sequence Read Archive (BioProject: PRJNA352507).

Adapter trimming & read mapping

We removed Illumina adapters from reads with *Trim Galore! v0.3.7* (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) which makes use of *Cutadapt v1.7.1* (61). Reads were then mapped to the zebra finch genome (*v3.2.74*,26) using *TopHat v2.0.13* (62), which utilizes the aligner *Bowtie v2.2.4* (63). We specified the library type as fr-firststrand in *TopHat2*. Successfully mapped reads were converted from SAM to BAM format with *SAMtools View v1.2* (64,65) and counted in *htseq-count v0.6.0* specifying '-s rev' (66). This assigned zebra finch Ensembl gene IDs and we only retained genes with >5X mapping across each sample.

Differential expression

Gene counts were then normalized for read-depth and analyzed for DE in *DEseq2 v1.8.1* (29). We analyzed DE across four comparisons: Control vs. Infected, Control vs. 2dpi, Control vs. 4dpi, and 2dpi vs. 4dpi. We visualized expression profiles in *R v3.3.0* (67) by PCA with the R

package pcaExplorer (68), and hierarchical clustering heat maps with the ggplot2 library (69) following the *DEseq2* manual. *DEseq2* tests for DE with a Wald test and genes were considered differentially expressed if the Benjamini & Hochberg (70) false discovery rate (FDR) correction for multiple testing p value < 0.10. We chose this significance threshold as *DEseq2* is generally conservative in classifying DE (71). Furthermore, this cutoff is used by the DEseq2 authors (29) and has been used in other RNAseq experimental infection studies (72). We plotted genes of interest individually with the plotCounts function in *DEseq2* and clustered expression profiles of these genes with the pheatmap R library to view expression levels across samples and treatments.

We tested DE genes for enriched gene ontology (GO) categories with *GOrilla* (31,32). *GOrilla* does not perform analyses with zebra finch Ensembl IDs, so we converted zebra finch Ensembl IDs to human Ensembl IDs using BioMart (73). We utilized this set of 10,152 genes for analysis. For each pairwise comparison, we used the FDR ranked order DE genes from *DEseq2*. Statistical significance was determined with p-values corrected for multiple hypothesis testing ($p < 0.05$) using the Benjamini & Hochberg method (70). To visualize DE results in the context of the RLR pathway, we utilized *Pathview v1.8.0* (36) to plot the log fold change of each gene detected in our dataset into the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway (KEGG ID = 04622) (74,75).

Time-course gene expression

In addition to the pair-wise comparisons performed in *DEseq2*, we were interested in understanding how clusters of genes are differentially expressed over the time course of infection. Thus, we performed DE analyses in *EBSeqHMM* (30). *EBSeqHMM* utilizes a bayesian

approach with a hidden Markov model to identify DE between ordered conditions. Genes are then grouped into expression paths (i.e. “Up-Down”, “Down-Down”), in which DE occurs when expression paths change between at least one adjacent condition. For example, a gene up-regulated at both 2dpi relative to control and 4dpi relative to 2dpi would be classified as “Up-Up”. We included three time points, with control individuals classified as t1, 2dpi as t2 and 4dpi as t3. Genes were considered DE at posterior probability > 0.99 and FDR < 0.01. We chose a more stringent cutoff in this analysis as *EBseq* can be liberal in classifying differential expression (71) and based on visual inspection of expression profiles. We ordered genes based on posterior-probability for each expression path and performed the GO analysis described above.

Ethics
All animal use was approved by the USGS National Wildlife Health Center Institutional Animal Care and Use Committee (IACUC Protocol: EP120521)

Data Availability
Sequencing data have been deposited in the Sequence Read Archive (SRA) under accession PRJNA352507.

Competing Interests
The authors declare no competing interests.

Author Contributions
EKH & CNB conceived the study. EKH performed experimental infections and dissection. CNB & DJN extracted and sequenced RNA and performed data analysis. All authors wrote and approved the manuscript.

Acknowledgements
The authors wish to thank Melissa Lund for technical assistance.

Funding
Funding provided by U.S. Geological Survey National Wildlife Health Center and East Carolina University. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

1. Chancey C, Grinev A, Volkova E, Rios M. 2015 The Global Ecology and Epidemiology of West Nile Virus. Biomed Res. Int. 2015, 20. (doi:10.1155/2015/376230)

2. Pérez-Ramírez E, Llorente F, Jiménez-Clavero MÁ. 2014 Experimental infections of wild birds with West Nile virus. *Viruses* **6**, 752–781. (doi:10.3390/v6020752)

3. McLean RG, Ubico SR, Docherty DE, Hansen WR, Sileo L, McNamara TS. 2001 West Nile virus transmission and ecology in birds. *Ann N Y Acad Sci* **951**, 54–57. (doi:10.1111/j.1749-6632.2001.tb02684.x)

4. LaDeau SL, Kilpatrick a M, Marra PP. 2007 West Nile virus emergence and large-scale declines of North American bird populations. *Nature* **447**, 710–713. (doi:10.1038/nature05829)

5. George TL, Harrigan RJ, LaManna JA, DeSante DF, Saracco JF, Smith TB. 2015 Persistent impacts of West Nile virus on North American bird populations. *Proc. Natl. Acad. Sci.* **112**, 14290–14294. (doi:10.1073/pnas.1507747112)

6. Kilpatrick AM, Daszak P, Jones MJ, Marra PP, Kramer LD. 2006 Host heterogeneity dominates West Nile virus transmission. *Proc. Biol. Sci.* **273**, 2327–2333. (doi:10.1098/rspb.2006.3575)

7. Komar N, Langevin S, Hinten S, Nemeth N, Edwards E, Hettler D, Davis B, Bowen R, Bunning M. 2003 Experimental infection of North American birds with the New York 1999 strain of West Nile virus. *Emerg. Infect. Dis.* **9**, 311–322. (doi:10.3201/eid0903.020628)

8. Suthar MS, Diamond MS, Gale M. 2013 West Nile virus infection and immunity. *Nat. Rev. Microbiol.* **11**, 115–28. (doi:10.1038/nrmicro2950)

9. Diamond MS, Shrestha B, Mehlhop E, Sitati E, Engle M. 2003 Innate and adaptive immune responses determine protection against disseminated infection by West Nile encephalitis virus. *Viral Immunol.* **16**, 259–278. (doi:10.1089/088282403322396082)

10. Diamond M, Shrestha B, Marri A. 2003 B cells and antibody play critical roles in the immediate defense of disseminated infection by West Nile encephalitis virus. *J. Virol.* **77**, 2578–2586. (doi:10.1128/JVI.77.4.2578)

11. Diamond MS, Gale M. 2012 Cell-intrinsic innate immune control of West Nile virus infection. *Trends Immunol.* **33**, 522–530. (doi:10.1016/j.it.2012.05.008)

12. Errett JS, Suthar MS, McMillan A, Diamond MS, Gale M. 2013 The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection. *J. Virol.* **87**, 11416–25. (doi: 10.1128/JVI.01488-13)

13. Sitati EM, Diamond MS. 2006 CD4+ T-cell responses are required for clearance of West Nile virus from the central nervous system. *J. Virol.* **80**, 12060–12069. (doi:10.1128/JVI.01650-06)

14. Shrestha B, Diamond M. 2004 Role of CD8+ T cells in control of West Nile virus infection. *J. Virol.* **78**, 8312–8321. (doi:10.1128/JVI.78.15.8312)

15. Lobigs M, Müllbacher A, Regner M. 2003 MHC class I up-regulation by flaviviruses: Immune interaction with unknown advantage to host or pathogen. *Immunol. Cell Biol.* **81**, 217–223. (doi:10.1046/j.1440-1711.2003.01161.x)

16. Cheng Y, King NJC, Kesson AM. 2004 Major histocompatibility complex class I (MHC-I) induction by West Nile virus: involvement of 2 signaling pathways in MHC-I up-regulation. *J. Infect. Dis.* **189**, 658–668. (doi:10.1086/381501)

17. Hewitt EW. 2003 The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology* **110**, 163–169. (doi:10.1046/j.1365-2567.2003.01738.x)

18. Petersen JL, Morris CR, Solheim JC. 2003 Virus evasion of MHC class I molecule presentation. *J. Immunol.* **171**, 4473–4478. (doi:10.4049/jimmunol.171.9.4473)

19. Fair JM, Nemeth NM, Taylor-McCabe KJ, Shou Y, Marrone BL. 2011 Clinical and acquired immunologic responses to West Nile virus infection of domestic chickens (*Gallus gallus domesticus*). *Poult. Sci.* **90**, 328–336. (doi:10.3382/ps.2010-00809)

20. Gamino V, Höfle U. 2013 Pathology and tissue tropism of natural West Nile virus infection in birds: A review. *Vet. Res.* **44**. (doi:10.1186/1297-9716-44-39)

21. Tag-El-Din-Hassan HT, Sasaki N, Moritoh K, Torigoe D, Maeda A, Agui T. 2012 The chicken 2'-5' oligoadenylate synthetase A inhibits the replication of West Nile virus. *Jpn. J. Vet. Res.* **60**, 95–103.

22. Langevin SA, Bunning M, Davis B, Komar N. 2001 Experimental infection of chickens as candidate sentinels for West Nile virus. *Emerg. Infect. Dis.* **7**, 726–729. (doi:10.3201/eid0704.017422)

23. Balakrishnan CN *et al.* 2010 Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC Biol.* **8**, 29. (doi:10.1186/1741-7007-8-29)

24. Bean AGD, Baker ML, Stewart CR, Cowled C, Deffrasnes C, Wang L-F, Lowenthal JW. 2013 Studying immunity to zoonotic diseases in the natural host - keeping it real. *Nat. Rev. Immunol.* **13**, 851–61. (doi:10.1038/nri3551)

25. Hofmeister EK, Lund M, Shearn-Bochsler V, Balakrishnan CN. 2017 Susceptibility and antibody response of the laboratory model zebra finch (*Taeniopygia guttata*) to West Nile virus. *PLoS One* **12**, e0167876. (doi:10.1371/journal.pone.0167876)

26. Warren WC *et al.* 2010 The genome of a songbird. *Nature* **464**, 757–62. (doi:10.1038/nature08819)

27. Hofmeister EK, Dusek RJ, Fassbinder-Orth C, Owen B, Franson JC. 2016 Susceptibility and antibody response of vesper sparrows (*Pooecetes gramineus*) to West Nile virus: a potential amplification host in sagebrush-grassland habitat. *J. Wildl. Dis.* **52**, 2015-06–148. (doi:10.7589/2015-06-148)

28. Yates A *et al.* 2016 Ensembl 2016. *Nucleic Acids Res.* **44**, D710-6. (doi:10.1093/nar/gkv1157)

29. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. (doi:10.1186/s13059-014-0550-8)

30. Leng N *et al.* 2015 EBSeq-HMM: A Bayesian approach for identifying gene-expression changes in ordered RNA-seq experiments. *Bioinformatics* **31**, 2614–2622. (doi:10.1093/bioinformatics/btv193)

31. Eden E, Lipson D, Yogev S, Yakhini Z. 2007 Discovering Motifs in Ranked Lists of DNA Sequences. *PLOS Comput. Biol.* **3**, e39.

32. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009 GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48. (doi:10.1186/1471-2105-10-48)

33. Barber MR, Aldridge Jr. JR, Webster RG, Magor KE. 2010 Association of RIG-I with innate immunity of ducks to influenza. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5913–5918. (doi:10.1371/journal.pcbi.0030039)

34. Huang Y *et al.* 2013 The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat. Genet.* **45**, 776–83. (doi:10.1038/ng.2657)

35. Wei L, Cui J, Song Y, Zhang S, Han F, Yuan R, Gong L, Jiao P, Liao M. 2014 Duck MDA5 functions in innate immunity against H5N1 highly pathogenic avian influenza virus infections. *Vet. Res.* **45**, 66. (doi:10.1186/1297-9716-45-66; 10.1186/1297-9716-45-66)

36. Luo W, Brouwer C. 2013 Pathview: An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**, 1830–1831. (doi:10.1093/bioinformatics/btt285)

37. Takeuchi O, Akira S. 2009 Innate immunity to virus infection. *Immunol. Rev.* **227**, 75–86. (doi:10.1111/j.1600-065X.2008.00737.x)

38. Fassbinder-Orth CA, Barak VA, Rainwater EL, Altrichter AM. 2014 Buggy Creek virus (Togaviridae: alphavirus) upregulates expression of pattern recognition receptors and

interferons in house sparrows (*Passer domesticus*). *Vector-Borne Zoonotic Dis.* **14**, 439–446. (doi:10.1089/vbz.2013.1531)

39. Jemielity S *et al.* 2013 TIM-family Proteins Promote Infection of Multiple Enveloped Viruses through Virion-associated Phosphatidylserine. *PLoS Pathog.* **9**. (doi:10.1371/journal.ppat.1003232)

40. Amara A, Mercer J. 2015 Viral apoptotic mimicry. *Nat. Rev. Microbiol.* **13**, 461–469. (doi:10.1038/nrmicro3469)

41. Qian F *et al.* 2013 Identification of Genes Critical for Resistance to Infection by West Nile Virus Using RNA-Seq Analysis. *Viruses*. **5**. (doi:10.3390/v5071664)

42. Hussmann KL, Fredericksen BL. 2014 Differential induction of CCL5 by pathogenic and non-pathogenic strains of West Nile virus in brain endothelial cells and astrocytes. *J. Gen. Virol.* **95**, 862–867. (doi:10.1099/vir.0.060558-0)

43. Kumar M, Belcaid M, Nerurkar VR. 2016 Identification of host genes leading to West Nile virus encephalitis in mice brain using RNA-seq analysis. *Sci. Rep.* **6**, 26350.

44. Venter M, Myers TG, Wilson MA, Kindt TJ, Paweska JT, Burt FJ, Leman PA, Swanepoel R. 2005 Gene expression in mice infected with West Nile virus strains of different neurovirulence. *Virology* **342**, 119–140. (doi:10.1016/j.virol.2005.07.013)

45. Schoggins JJW, Wilson SJS, Panis M, Murphy MMY, Jones CCT, Bieniasz P, Rice CM, CM R. 2011 A diverse array of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485. (doi:10.1038/nature09907.A)

46. Liu Y, King N, Kesson A, Blanden R V., Müllbacher A. 1989 Flavivirus infection up-regulates the expression of class I and class II major histocompatibility antigens on and enhances T cell recognition of astrocytes in vitro. *J. Neuroimmunol.* **21**, 157–168. (doi:10.1016/0165-5728(89)90171-9)

47. Lobigs M, Müllbacher A, Regner M. 2003 MHC class I up-regulation by flaviviruses: Immune interaction with unknown advantage to host or pathogen. *Immunol. Cell Biol.* **81**, 217–223. (doi:10.1046/j.1440-1711.2003.01161.x)

48. Watzl C. 2014 How to trigger a killer: modulation of natural killer cell reactivity on many levels. *Adv. Immunol.* **124**, 137–170. (doi:10.1016/B978-0-12-800147-9.00005-4)

49. Kumar M, Verma S, Nerurkar VR. 2010 Pro-inflammatory cytokines derived from West Nile virus (WNV)-infected SK-N-SH cells mediate neuroinflammatory markers and neuronal death. *J. Neuroinflammation* **7**, 73. (doi:10.1186/1742-2094-7-73)

50. Karim R *et al.* 2013 Human Papillomavirus (HPV) upregulates the cellular deubiquitinase UCHL1 to suppress the keratinocyte's innate immune response. *PLoS Pathog.* **9**. (doi:10.1371/journal.ppat.1003384)

51. Brass AL *et al.* 2009 The IFITM proteins mediate cellular resistance to Influenza A H1N1 virus, West Nile virus, and Dengue virus. *Cell* **139**, 1243–1254. (doi:10.1016/j.cell.2009.12.017)

52. El-Bacha T, Da Poian AT. 2013 Virus-induced changes in mitochondrial bioenergetics as potential targets for therapy. *Int. J. Biochem. Cell Biol.* **45**, 41–46. (doi:http://dx.doi.org/10.1016/j.biocel.2012.09.021)

53. Zeidler DJ, Fernandes-Siqueira OL, Barbosa MG, Da Poian TA. 2017 Non-canonical roles of Dengue virus non-structural proteins. *Viruses* . **9**. (doi:10.3390/v9030042)

54. Jeurissen SH. 1993 The role of various compartments in the chicken spleen during an antigen-specific humoral response. *Immunology* **80**, 29–33.

55. Smith KG, Hunt JL. 2004 On the use of spleen mass as a measure of avian immune system strength. *Oecologia* **138**, 28–31. (doi:10.1007/s00442-003-1409-y)

56. Schurch NJ *et al.* 2016 How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* **22**, 839–851. (doi:10.1261/rna.053959.115)

57. Bonneaud C, Balenger SL, Russell AF, Zhang J, Hill GE, Edwards S V. 2011 Rapid evolution of disease resistance is accompanied by functional changes in gene expression in a wild bird. *Proc. Natl. Acad. Sci.* **108**, 7866–7871. (doi:10.1073/pnas.1018580108)

58. Zhang Q, Hill GE, Edwards S V, Backstrom N. 2014 A house finch (*Haemorhous mexicanus*) spleen transcriptome reveals intra- and interspecific patterns of gene expression, alternative splicing and genetic diversity in passerines. *BMC Genomics* **15**, 305. (doi:10.1186/1471-2164-15-305)

59. Hamzić E *et al.* 2016 RNA sequencing-based analysis of the spleen transcriptome following infectious bronchitis virus infection of chickens selected for different mannose-binding lectin serum concentrations. *BMC Genomics* **17**, 82. (doi:10.1186/s12864-016-2403-1)

60. Van Goor A, Ashwell CM, Persia ME, Rothschild MF, Schmidt CJ, Lamont SJ. 2017 Unique genetic responses revealed in RNA-seq of the spleen of chickens stimulated with lipopolysaccharide and short-term heat. *PLoS One* **12**, e0171414.

61. Martin M. 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10. (doi:10.14806/ej.17.1.200)

62. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013 TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36. (doi:10.1186/gb-2013-14-4-r36)

63. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359. (doi:10.1038/nmeth.1923)

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

65. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)

66. Anders S, Pyl PT, Huber W. 2015 HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169. (doi:10.1093/bioinformatics/btu638)

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

68. Marini F. 2016 pcaExplorer: Interactive Visualization of RNA-seq Data Using a Principal Components Approach. R package version 1.1.2, https://github.com/federicomarini/pcaExplorer.

69. Wickham H. 2009 ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

70. Benjamini Y, Hochberg Y. 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300. (doi:10.2307/2346101)

71. Seyednasrollah F, Laiho A, Elo LL. 2015 Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief. Bioinform.* **16**, 59–70. (doi:10.1093/bib/bbt086)

72. Videvall E, Cornwallis CK, Palinauskas V, Valkiūnas G, Hellgren O. 2015 The avian transcriptome response to malaria infection. *Mol. Biol. Evol.* **32**, 1255–67. (doi:10.1093/molbev/msv016)

73. Kinsella RJ et al. 2011 Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database: J. Biol. Databases Curation.* **2011**, bar030. (doi:10.1093/database/bar030)

74. Kanehisa M, Goto S. 2000 KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30. (doi:10.1093/nar/28.1.27)

75. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016 KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462. (doi:10.1093/nar/gkv1070)

76. Elbahesh H, Scherbik S V., Brinton MA. 2011 West Nile virus infection does not induce PKR activation in rodent cells. *Virology* **421**, 51–60. (doi:10.1016/j.virol.2011.08.008)

CHAPTER 2. TRANSCRIPTOMIC SUPPORT FOR THE IMMUNOCOMPETENCE HANDICAP HYPOTHESIS BUT NOT THE OXIDATION HANDICAP HYPOTHESIS

Abstract

Sexually selected traits are hypothesized to be honest signals of individual quality due to the costs associated with their development or expression. Testosterone, a sex steroid known to influence the development/expression of sexually selected traits, has been proposed to underlie the costs associated with sexually selected traits via its immunosuppressive effects (i.e., the Immunocompetence Handicap Hypothesis) or by influencing an individual's exposure/susceptibility to oxidative stress (i.e., the Oxidation Handicap Hypothesis). Previous work testing these hypotheses has primarily focused on physiological measurements of immunity or oxidative stress, but little is known about the molecular pathways by which testosterone could influence immunity and/or oxidative stress pathways. To measure the direct consequences of experimentally elevated testosterone, we used previously published RNA-seq data from studies that measured the transcriptome of individuals treated with either a testosterone-filled or an empty (i.e., control) implant. Of the two published datasets, we found strong support for the Immunocompetence Handicap Hypothesis. However, we found no support for the Oxidation Handicap Hypothesis. More specifically, testosterone-treated individuals exhibited strong signatures of immunosuppression, encompassing both cell-mediated and humoral immunity. Our results suggest that testosterone enforces the honesty of sexually-selected traits by influencing an individual's immunocompetence rather than their exposure or susceptibility to oxidative stress.

Introduction

There is a long-standing interest in understanding why sexually selected traits have evolved and one hypothesis suggests that mates have selected for traits that are costly to develop or bear (i.e., the handicap hypothesis; Zahavi 1975). An important assumption of the handicap hypothesis is that an individual's investment in sexually selected traits correlates with their investment in other traits that also influence their reproductive success or survival (Grafen 1990, Andersson 1994). Individuals face resource-based tradeoffs when fitness-related traits exhibit negative correlations and, because of these negative correlations, individuals can incur survival costs from their reproductive investments (Stearns 1992). From a proximate perspective, these costs arise because the development and/or expression of traits important for reproduction (e.g., sexually selected traits) and traits important for survival (e.g., immune function) are dependent on the same mechanism (Zera and Harshman 2001). As such, our understanding of the evolution of sexually selected traits is dependent upon our understanding of the mechanisms that underlie their development and/or expression.

Testosterone is a sex steroid that is known to influence the development and/or expression of sexually selected traits (Hau 2007, Fusani 2008, Ball and Balthazart 2009). In combination with its effects on other fitness related traits (e.g., immune function; Segner et al. 2017), testosterone is thought to enforce the honesty of sexually selected traits (Ketterson et al. 1999, Owen-Ashley et al. 2004). Two prominent hypotheses have been proposed to explain how testosterone enforces the honesty of sexually selected traits: the Immunocompetence Handicap Hypothesis (Folstad and Karter 1992) and the Oxidation Handicap Hypothesis (Alonso-Alvarez et al. 2007). The Immunocompetence Handicap Hypothesis proposes that sexually selected traits remain honest because testosterone has antagonistic effects on an individual's immune function.

Therefore, among free-living animals, poor quality or low condition individuals cannot maintain high levels of circulating testosterone due its immunosuppressive effects (Folstad and Karter 1992). A meta-analysis by Roberts et al. (2004) revealed weak support for this hypothesis. However, a more recent meta-analysis found that experimentally increasing testosterone results in suppression of both cell-mediated and humoral immunity (Foo et al. 2017). This same meta-analysis also found positive, but nonsignificant, trends between multiple measures of immune function and naturally occurring levels of circulating testosterone (Foo et al. 2017). These results fit the predictions of the Immunocompetence Handicap Hypothesis because individuals naturally expressing high of testosterone represent high quality or high condition individuals that can invest in sexually selected traits without compromising their immune system (e.g., Peters et al. 2000). The Oxidation Handicap Hypothesis, on the other hand, states that sexually selected traits remain honest because testosterone increases an individual's susceptibility and/or exposure to oxidative stress (Alonso-Alvarez et al. 2007). Of the few studies that have directly tested the Oxidation Handicap Hypothesis, some have found support (Mougeot et al. 2009, Hoogenboom et al. 2012) while results from others did not find support for this hypothesis (Isaksson et al. 2011, Casagrande et al. 2012, Taff and Freeman-Gallant 2014, Baldo et al. 2015). Nonetheless, both hypotheses have been primarily tested using physiological measurements of oxidative stress and immunity, but relatively little is known about the underlying molecular pathways. Given that sex steroids partly function by binding to intracellular receptors and acting as transcription factors (Nelson 2011), measuring the relationship between testosterone and transcription can shed light on the proximate pathways that testosterone influences.

Modern sequencing approaches, like RNA sequencing (RNA-seq), allow for comprehensive measurements of whole transcriptomes and the relative abundance of each transcript (Wang et al. 2009). This approach offers the opportunity to assess coordinated, large-scale transcriptional responses rather than focusing on targeted candidate genes (e.g. via qPCR). Additionally, RNA-seq provides many advantages over microarray-based studies, as RNA-seq provides higher sensitivity and is not subject to hybridization biases (Wang et al. 2009). RNA-seq approaches have been used to investigate the role of testosterone on gene expression, particularly in the context of sex differences (e.g. Gao et al. 2015, Cox et al. 2017) and gonadal development (e.g. Monson et al. 2017, Zheng et al. 2019). Similarly, RNA-seq based studies have been crucial in providing a more comprehensive understanding of the complex and dynamic immune and stress responses (e.g. Barshis et al. 2013, Huang et al. 2013, Kim et al. 2018). In the context of mate choice, measuring the relationship between testosterone and transcription can ultimately shed light on the pathways that testosterone influences to potentially enforce the honesty of sexually selected traits (e.g., immune or oxidative stress pathways). However, our understanding of the pleiotropic nature of testosterone is partly dependent upon our understanding of the direct consequences of circulating testosterone on the transcriptomic signatures of immunity and oxidative stress (e.g., via RNA-seq), something relatively few studies have examined (but see Wenzel et al. 2013).

Here, we used published transcriptome datasets from studies that compared gene expression between testosterone-treatment and control subjects in two species: golden-collared manakin (*Manacus vitellinus*) and Japanese quail (*Coturnix japonica*). Golden-collared manakin males produce brightly colored plumage ornaments and engage in elaborate courtship behaviors during the breeding season, a process that is dependent on the activational effects of testosterone

(Day et al. 2007, Schlinger et al. 2013). Japanese quail males produce brightly colored cheek feather to attract females and, like manakins, this process is dependent on the activational effects of testosterone (Hiyama et al. 2018). Female quail were also recently shown to prefer males with higher testosterone levels (Hiyama et al. 2018). Using transcriptomic data from muscle and the foam gland, tissues that are, respectively, known to be sensitive to testosterone in manakins (Fuxjager et al. 2016) and in quail (Adkins-Regan 1999), we re-analyze the data to explicitly test the Immunocompetence Handicap Hypothesis and the Oxidation Handicap Hypothesis. We constructed co-expression networks to identify gene networks responding to testosterone treatment. If testosterone is immunosuppressive, then we predict that testosterone treatment will cause consistent down-regulation (i.e. suppression) of genes with annotated immune function in both species. Similarly, if testosterone influences an individual's susceptibility or exposure to oxidative stress, then we predict that testosterone treatment will cause a decrease in the expression of genes with annotated functions in antioxidant protection and/or an increase in genes that are expressed in response to oxidative stress.

Methods

Study selection

To identify studies of interest, we first performed a literature search on both Scopus and Google Scholar with the following search terms: "testosterone" AND "RNA-seq" or "transcriptome" or "transcriptomics". To identify unpublished data, we also searched the NCBI Sequence Read Archive (SRA) and Gene Expression Omnibus databases with the search term "testosterone". We extracted experimental design information from the methods sections of published work and from the metadata of the above databases.

To be included, studies must have performed RNA-seq analyses in both an experimentally elevated testosterone and a sham manipulated control group, tested adult male individuals, publicly deposited raw sequencing data, sequenced on a comparable platform (e.g. Illumina), and sequenced to ≥ 5 million reads per sample. We excluded studies that castrated individuals before experimentally increasing testosterone and studies that only included female experimental groups. Lastly, we restricted our analyses to existing bulk RNA-seq data and excluded microarray data, as RNA-seq allows higher sensitivity and is not subject to hybridization biases (Wang et al. 2009). This process resulted in two studies for re-analysis. Fuxjager et al. (2016) experimentally increased testosterone in golden-collared manakins (*Manacus vitellinus*, "manakin") and performed RNA-seq on pectoralis and scapulohumeralis caudalis tissue (n= 3 each testosterone and control for each tissue). Finseth and Harrison (2018) experimentally increased testosterone in Japanese quail (*Coturnix japonica*, "quail") experiencing short days and performed RNA-seq on the foam gland (FG) (n=6 each testosterone and Control).

Data re-analysis

We downloaded the raw sequencing data from SRA with sratoolkit fastq-dump (quail: PRJNA397592; manakin: PRJNA297576). We first adaptor trimmed all reads with Trim Galore! and aligned trimmed reads to the respective reference genome (*M. vitellinus* v2, *C. japonica* v2) for each species with STAR v2.5.3 (Dobin et al. 2013). We then quantified expression with htseq-count v0.6.0 (Anders et al. 2015), specifying strand 'no'. We created a count matrix for each species and kept all genes with an average count ≥ 5 across all samples. Using this count matrix, we first normalized counts to sequencing depth and variance stabilizing transformed

counts with DEseq2 (Love et al. 2014). Transformed counts were visualized with a principal component analysis (PCA) using pcaExplorer v2.8.1 (Marini & Binder 2016).

To test for the effect of testosterone treatment on transcription, we constructed gene co-expression network analysis with the weighted gene co-expression network analysis (WGCNA) tool (Langfelder & Horvath 2008, Langfelder et al. 2011). Briefly, WGCNA constructs modules (i.e. networks) of co-expressed, interacting genes independent of the underlying data structure. These modules are then correlated with a trait of interest, representing coordinated changes in expression within that trait. We created modules independently for each species with the following shared parameters: network type = signed, minimum module size = 30, and module dissimilarity = 0.2. We used $\beta = 12$ for quail and $\beta = 18$ for manakin, which represents the point the network reached scale free topology. We then tested for correlations between modules and testosterone treatment using a $p<0.05$ cutoff. We identified the hub genes of each module by selecting the top five genes with the highest module membership (MM) score. Modules negatively correlated with treatment represent a decrease in expression following testosterone treatment relative to control. Modules positively correlated represent an increase in expression. For manakin, we also tested for correlations with muscle type to see if any modules were correlated with testosterone treatment differed between the tissues sampled.

We were also interested if modules correlated with testosterone treatment in one species were preserved in the other species. Thus, we tested for module preservation in WGCNA between quail and manakin. From the list of genes expressed in each species, we identified 7,366 shared one-to-one orthologs between quail and manakin using Ensembl BioMart (Kinsella et al. 2011). This module preservation approach calculates a Zsummary score, with Zsummary < 0

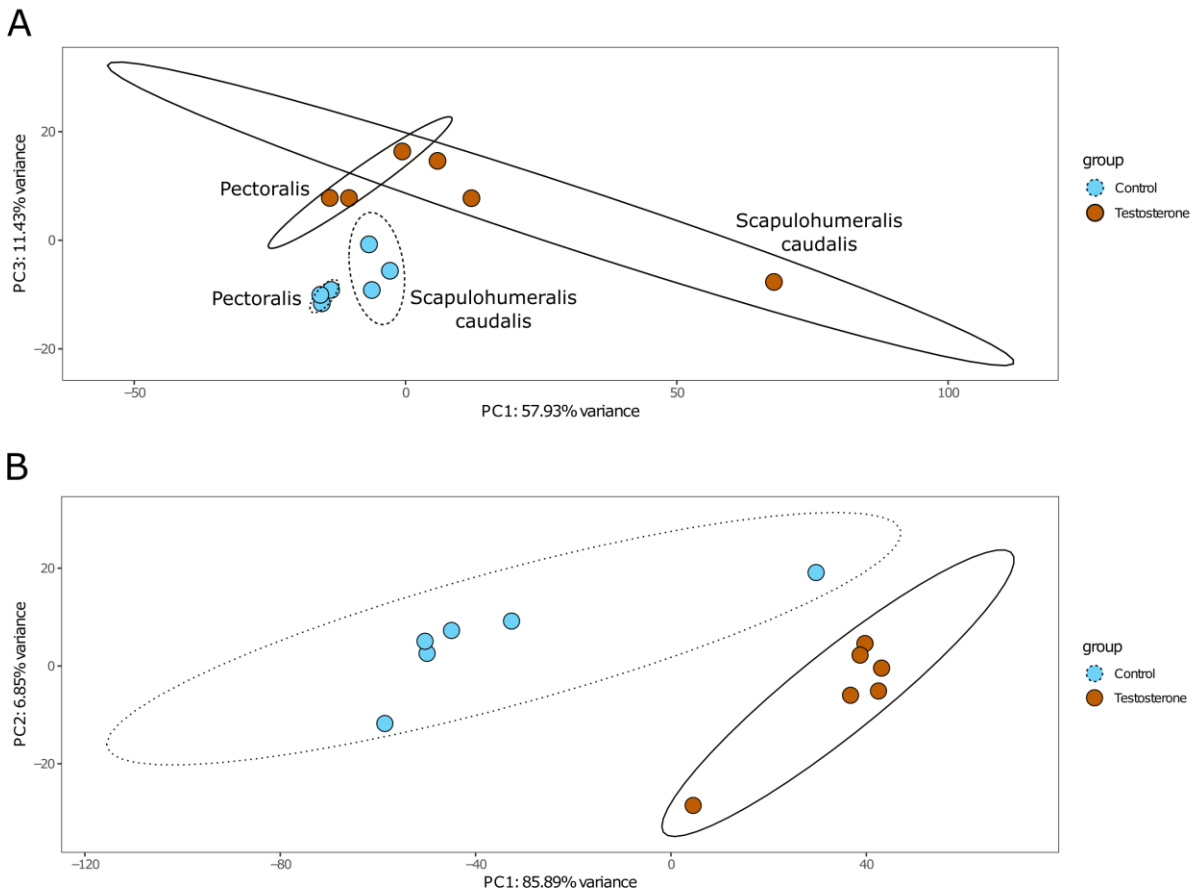
representing no preservation, Zsummary > 2 and < 10 representing moderate preservation, and Zsummary > 10 indicating strong preservation between the two species.

To test the Immunocompetence and Oxidative Stress handicap hypotheses, we performed ranked order gene ontology (GO) analyses with GOrilla (Eden et al. 2007, 2009). In WGCNA, each gene is assigned a MM score for each module. MM represents the correlation of the gene to the module eigengene. Genes with high MM in a module are representative genes for the expression pattern of that module. For each module, we ordered the gene list by descending MM scores for the given module and input this entire list into GOrilla. This ranked approach allowed us to use all genes for GO enrichment in a single list. GOrilla then tests for enrichment and places greater weight on those genes at the top of the list relative to the bottom. GO categories were significantly enriched if the qvalue < 0.05. To find support for the Immunocompetence Handicap Hypothesis, immune related GO categories (e.g., "immune system process) had to be significantly enriched among down-regulated genes. To find support for the Oxidation Handicap hypothesis, oxidative stress related GO categories had to be significantly enriched among up-regulated genes (e.g., "response to oxidative stress") or down-regulated genes (e.g. "antioxidant activity").

Results

After filtering, we used 13,509 manakin genes and 13,946 quail genes for PCA and WGCNA network construction. Testosterone treatment had pronounced effects on gene expression and individuals clustered by treatment in both comparisons (Figure 2.1).

Figure 2.1. Principal component analyses of (A) manakin and (B) quail. Samples separate by treatment along PC3 for manakin and PC1 for quail. Each circle represents a sample and is color-coded by treatment. Manakin samples are labeled by muscle type. Ellipses represent 95% confidence intervals.



WGCNA – Quail

WGCNA constructed 18 modules for quail, six of which were correlated with testosterone treatment (Supplemental Figure 1). These are coded with arbitrary colors for visual presentation.

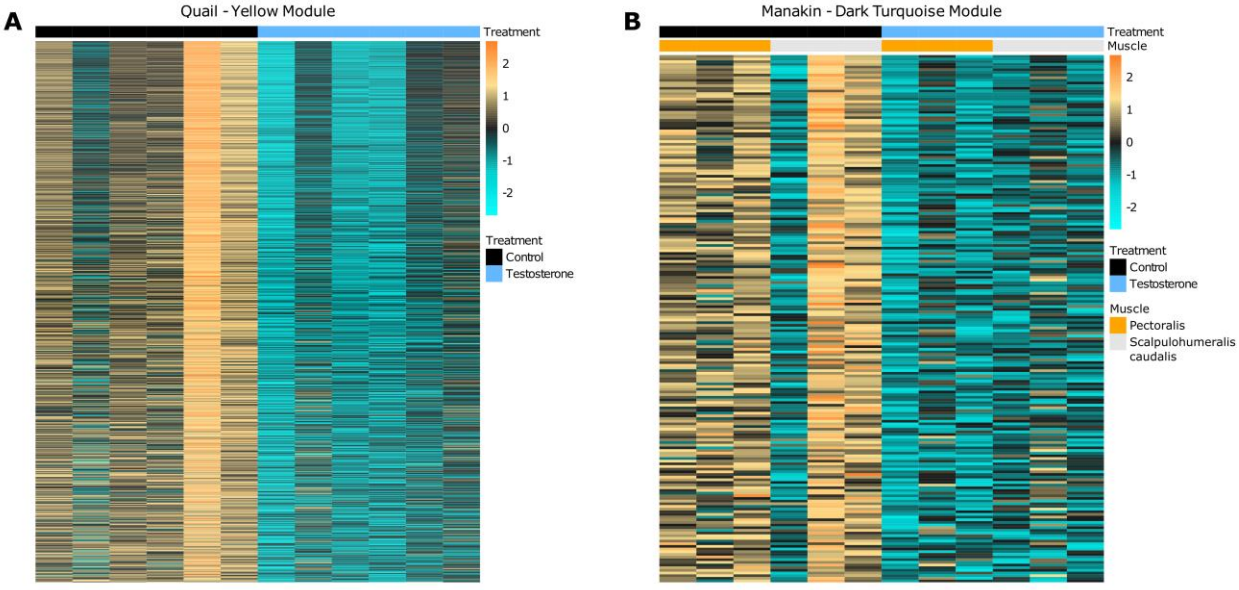
The yellow module (925 genes, r=-0.74) and dark green module (88 genes, r=-0.67) were both strongly enriched for immune related GO categories (Table 2.1). The yellow module was primarily enriched for broad immune categories, e.g. “immune system process” and “immune response”, whereas the dark green module was primarily enriched for lymphocyte and leukocyte related categories. This represents a significant decrease in immune gene expression following treatment (Figure 2.2A). As we were primarily interested in the immune effects of testosterone treatment, we also identified the hub genes of the yellow and dark green modules. The yellow module hubs were SASH3, ITGB2, SLAMF8 (LOC107324444), TRAF3IP3, and EVI2A. The dark green hub genes were FBL, PIK3R6, STOML2, GPR157, and DNAL4.

Table 2.1. Immune module GO enrichment for both species. The top 5 gene ontology (GO) categories are presented, along with FDR adjusted p-value and GOrilla enrichment score. Enrichment is calculated as (b/n)/(B/N), where N is the total number of genes, B is the total number of genes associated with a specific GO term, n is the number of genes in the top of the input list, and b is the number of GO-term-associated genes in the top of the list (Eden et al. 2007, 2009).

GO ID	Description	FDR	Enrichment
Quail, Yellow Module			
GO:0002376	immune system process	3.11E-45	4.53
GO:0006955	immune response	3.47E-35	5.53
GO:0002682	regulation of immune system process	4.30E-35	3.59
GO:0002684	positive regulation of immune system process	1.34E-32	4.25
GO:0046649	lymphocyte activation	1.08E-30	11.1
Quail, Dark Green Module			
GO:0002684	positive regulation of immune system process	4.09E-05	1.86
GO:1903706	regulation of hemopoiesis	5.89E-05	2.17
GO:0046649	lymphocyte activation	6.06E-05	2.45
GO:0038023	signaling receptor activity	6.38E-05	1.83
GO:0002682	regulation of immune system process	6.42E-05	1.61
Manakin, Dark Turquoise Module			
GO:0006955	immune response	8.81E-10	2.48
GO:0046649	lymphocyte activation	3.29E-08	7.44
GO:0042110	T cell activation	3.57E-08	10.58
GO:0002376	immune system process	4.34E-08	3.04

GO:0002521	leukocyte differentiation	3.31E-07	4.09

Figure 2.2. Expression heatmaps of the (A) Yellow Module and (B) Dark Turquoise Module, both of which were significantly correlated with testosterone treatment in quail and manakin respectively and represent down-regulation of the immune system. Each column represents a sample color coded by treatment or muscle type. Each row represents a module gene. High expression is indicated by orange colors and low expression is represented by blue colors.



The black and purple modules were also negatively correlated with testosterone treatment and were enriched for translation and muscle process GO categories respectively. Lastly, we find two modules up-regulated following testosterone treatment. The turquoise module was the most strongly correlated with testosterone treatment (4423 genes, r=0.98). GO enrichment is largely driven by genes involved in the Golgi apparatus and endoplasmic reticulum functions

(Supplement Table 1). The Green module (795 genes, r=0.61) is primarily enriched for broad metabolic activity and protein modification processes.

WGCNA – Manakin

Under our chosen parameters, WGCNA constructed 34 modules for manakin, 12 of which were correlated with testosterone treatment (Supplemental Figure 2). Seven modules were correlated with muscle type. None of these modules were also correlated with testosterone treatment, indicating no tissue specific response at the network level. Of the 12 modules correlated with testosterone treatment, 7 were negatively correlated and 5 positively correlated. Like the quail, manakins also exhibited a significant decrease in immune gene expression following testosterone treatment (Figure 2.2B, Supplemental Table 2). The dark turquoise module (198 genes, r=-0.71) was strongly enriched for a broad range of immune related GO categories (Table 2.1). The dark turquoise hub genes were MHC1A (LOC108639055), INPPL1 (LOC103767762), CCL14 (LOC103758017), CCL3L (LOC103757995), and an uncharacterized non-coding RNA (LOC108640668).

The remaining negatively correlated modules were primarily enriched for metabolism (green, dark olive green), ribosomal components (dark red, pale turquoise), and mitochondria related categories (steel blue, pale turquoise). Among the positively correlated modules, we also found enrichment of cellular metabolism, catabolism, and mitochondrial related GO categories (Supplemental Table 2).

WGCNA – Module Preservation

We performed reciprocal module preservation analyses to test for quail modules preserved in manakin and manakin modules preserved in quail. No modules were strongly preserved between

the two species (Supplemental Figure 3). While we found strong signatures of immune system down-regulation in both species, no immune related modules were preserved.

Discussion

In this study, we tested the Immunocompetence Handicap Hypothesis and the Oxidation Handicap Hypothesis by quantifying transcriptional responses to experimentally increased circulating testosterone in two species of bird. Our gene network analysis revealed that both manakin and quail exhibit immunosuppression following testosterone treatment, supporting the Immunocompetence Handicap Hypothesis. However, we did not find support for the Oxidation Handicap Hypothesis, as there was no enrichment of genes expressed related to oxidative damage, nor was there suppression of genes related to antioxidant defenses in either species. These results suggest that high levels of circulating testosterone can be costly to maintain partly due to their potential negative effects on an individual's immunocompetence and not the individual's susceptibility or exposure to oxidative stress. Oxidative stress could still be involved in enforcing the costs of reproduction or sexually selected signals. However, our results suggest that this cost is not borne out via pathways that are sensitive to testosterone.

Our analyses revealed that immunosuppression was broad, encompassing aspects of both cell-intrinsic innate immunity (e.g. lymphocyte activation and cytokine signaling) as well as adaptive immunity (e.g. antigen processing and presentation) across both species (Table 2.1). Combined, this represents a vulnerable immune state in these tissues. The observed effect of testosterone could occur through both genomic and non-genomic pathways, but regulation of the immune system by androgens receptor likely plays an important role (Trigunaite et al. 2015, Segner et al. 2017, Gubbels Bupp & Jorgensen 2018). More specifically, while testosterone

exposure and subsequent androgen receptor activity can promote innate immune cell differentiation and development, testosterone also reduces activity of these cells (Gubbels Bupp & Jorgensen 2018). Rettew et al. (2008), for example, found that testosterone exposure suppressed TLR4 expression, a key regulator of innate immunity and inflammation. TLR4 was present in non-immune related modules of our study (manakin dark magenta, quail turquoise). Nonetheless, the hub genes of the immune related modules highlight broad suppression of the innate immune signaling (quail yellow: SASH3, SLAMF8, TRAF3IP3; manakin dark turquoise: INPPL1, CCL14, CCL3L, ncRNA) (Beer et al. 2005, Dauphinee et al. 2013, Veillette 2010, Zou et al. 2015, Thomas et al. 2016, Sokol & Luster 2015, Wang et al. 2018). Similarly, testosterone exposure had substantial effects on the regulation of the adaptive immune system. Testosterone exposure greatly reduces T cell activity (Lin et al. 2010, Kissick et al. 2014), which is a prominent signature in both quail (Supplemental Table 1) and manakin (Table 2.1). In addition to suppression of T cell activity in manakin, we also identified MHC class IA as a hub gene in the manakin dark turquoise module. MHC class IA binds and presents viral peptides to CD8+ T cells, which is a critical component of the adaptive immune response (Neefjes et al. 2011). Previous work has shown suppressive effects of testosterone on CD4+ T cells/MHC class IIB (Lin et al. 2010) and CD8+ T cells (Page et al. 2006). However, our study is the first to describe suppression of genes involved in T cell activity as well as MHC class I.

We were also interested in whether the expression response to testosterone was shared between manakin and quail. No modules were preserved between quail and manakin. Thus, despite evidence of immunosuppression in both species, the immune related gene networks are not preserved between the species. These results suggest either a species specific and/or tissue specific response to testosterone treatment, both of which have previously been documented

transcriptomic data (Breschi et al. 2016). We identified these immunosuppression signatures in muscle and foam gland, which are not traditionally studied in avian immunology (Rose 1979, Schat et al. 2014). Even though we did not measure immune tissues, this immunosuppression signature was apparent and highlights the sensitivity of RNA-seq to detect functional signatures in non-traditional tissues (e.g. Louder et al. 2018). In both species, T is necessary to produce secondary sexual characteristics for mating (see Introduction). Thus, our results point to a direct trade-off between expression of sexually selected traits and immune function.

Given that we found strong support for immunosuppression in both studies, additional experiments should be conducted to continue to broaden our understanding of testosterone's immunosuppressive effects. First, studies should focus on performing testosterone manipulations and examining transcriptomic responses in a wider range of tissues and species. Moreover, studying should prioritize conducting experimental infections and/or immune challenges in combination with RNA-seq analyses to examine how transcriptomic signatures relate to immune function. Novel endocrine-based experiments, similar to Goymann et al. (2015) and Goymann and Dávila (2017), paired with RNA-seq analyses can also shed like on how acute changes in testosterone levels influence transcription over shorter timeframes. Overall, these integrative approaches will ultimately provide novel insights into the evolution of sexually selected traits at the mechanistic level.

Acknowledgements

Ignacio Moore, Christopher Balakrishnan, and Dana Hawley provided feedback on an earlier version of this manuscript.

References

Adkins-Regan, E. (1999). Foam Produced by Male Coturnix Quail: What Is Its Function? *The Auk*, *116*(1), 184–193. doi: 10.2307/4089465

Alonso-Alvarez, C., Bertrand, S., Faivre, B., Chastel, O., & Sorci, G. (2007). Testosterone and oxidative stress: the oxidation handicap hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1611), 819–825. doi: 10.1098/rspb.2006.3764

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. doi: 10.1093/bioinformatics/btu638

Andersson, M. B. (1994). *Sexual selection*. Princeton, N.J: Princeton University Press.

Baldo, S., Mennill, D. J., Guindre-Parker, S., Gilchrist, H. G., & Love, O. P. (2015). The Oxidative Cost of Acoustic Signals: Examining Steroid Versus Aerobic Activity Hypotheses in a Wild Bird. *Ethology*, *121*(11), 1081–1090. doi: 10.1111/eth.12424

Ball, G. F., and J. Balthazart. (2009). Neuroendocrine Regulation of Reproductive Behavior in Birds. Pages 855–897 Hormones, Brain and Behavior. Third Edit. Elsevier.

Barshis, D. J., Ladner, J. T., Oliver, T. A., Seneca, F. O., Traylor-Knowles, N., & Palumbi, S. R. (2013). Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences*, *110*(4), 1387–1392. doi: 10.1073/pnas.1210224110

Beer, S., Scheikl, T., Reis, B., Huser, N., Pfeffer, K., & Holzmann, B. (2005). Impaired Immune Responses and Prolonged Allograft Survival in Sly1 Mutant Mice. *Molecular and Cellular Biology*, *25*(21), 9646–9660. doi: 10.1128/MCB.25.21.9646-9660.2005

Breschi, A., Djebali, S., Gillis, J., Pervouchine, D. D., Dobin, A., Davis, C. A., … Guigó, R. (2016). Gene-specific patterns of expression variation across organs and species. *Genome Biology*, *17*(1). doi: 10.1186/s13059-016-1008-y

Casagrande, S., Costantini, D., & Groothuis, T. G. G. (2012). Interaction between sexual steroids and immune response in affecting oxidative status of birds. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *163*(3–4), 296–301. doi: 10.1016/j.cbpa.2012.07.018

Coil W. H., Wetherbee D. K., 1959 Observations on the cloacal gland of the Eurasian quail Coturnix coturnix. Ohio J. Sci. 59: 268–270.

Cox, R. M., Cox, C. L., McGlothlin, J. W., Card, D. C., Andrew, A. L., & Castoe, T. A. (2017). Hormonally Mediated Increases in Sex-Biased Gene Expression Accompany the Breakdown of Between-Sex Genetic Correlations in a Sexually Dimorphic Lizard. *The American Naturalist*, *189*(3), 315–332. doi: 10.1086/690105

Dauphinee, S. M., Clayton, A., Hussainkhel, A., Yang, C., Park, Y.-J., Fuller, M. E., … Karsan, A. (2013). SASH1 Is a Scaffold Molecule in Endothelial TLR4 Signaling. *The Journal of Immunology*, *191*(2), 892–901. doi: 10.4049/jimmunol.1200583

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21.

Eden, E., Lipson, D., Yogev, S., & Yakhini, Z. (2007). Discovering Motifs in Ranked Lists of DNA Sequences. *PLOS Computational Biology*, *3*(3), e39–e39.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, *10*(1), 48–48. doi: 10.1186/1471-2105-10-48

Finseth, F. R., & Harrison, R. G. (2018). Genes Integral to the Reproductive Function of Male Reproductive Tissues Drive Heterogeneity in Evolutionary Rates in Japanese Quail. *G3: Genes|Genomes|Genetics*, *8*(1), 39–51. doi: 10.1534/g3.117.300095

Folstad, I., & Karter, A. J. (1992). Parasites, Bright Males, and the Immunocompetence Handicap. *American Naturalist*, *139*(3), 603–622.

Foo, Y. Z., Nakagawa, S., Rhodes, G., & Simmons, L. W. (2017). The effects of sex hormones on immune function: a meta-analysis: Sex hormones and immune function. *Biological Reviews*, *92*(1), 551–571. doi: 10.1111/brv.12243

Fusani, L. (2008). Testosterone control of male courtship in birds. Hormones and Behavior 54:227–233.

Fuxjager, M. J., Lee, J.-H., Chan, T.-M., Bahn, J. H., Chew, J. G., Xiao, X., & Schlinger, B. A. (2016). Research Resource: Hormones, Genes, and Athleticism: Effect of Androgens on the Avian Muscular Transcriptome. *Molecular Endocrinology*, *30*(2), 254–271. doi: 10.1210/me.2015-1270

Gao, J., Liu, S., Zhang, Y., Yang, Y., Yuan, C., Chen, S., & Wang, Z. (2015). Effects of 17 α-methyltestosterone on transcriptome, gonadal histology and sex steroid hormones in rare minnow Gobiocypris rarus. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *15*, 20–27. doi: 10.1016/j.cbd.2015.05.001

Goymann, W., C. P. Villavicencio, and B. Apfelbeck. (2015). Does a short-term increase in testosterone affect the intensity or persistence of territorial aggression? — An approach using an individual's hormonal reactive scope to study hormonal effects on behavior. Physiology & Behavior 149:310–316.

Goymann, W., and P. F. Dávila. (2017). Acute peaks of testosterone suppress paternal care: Evidence from individual hormonal reaction norms. Proceedings of the Royal Society B: Biological Sciences 284:20170632.

Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, *144*(4), 517–546. doi: 10.1016/S0022-5193(05)80088-8

Gubbels Bupp, M. R., & Jorgensen, T. N. (2018). Androgen-Induced Immunosuppression. *Frontiers in Immunology*, *9*. doi: 10.3389/fimmu.2018.00794

Hau, M. (2007). Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. *BioEssays*, *29*(2), 133–144. doi: 10.1002/bies.20524

Hiyama, G., Mizushima, S., Matsuzaki, M., Tobari, Y., Choi, J.-H., Ono, T., … Sasanami, T. (2018). Female Japanese quail visually differentiate testosterone-dependent male attractiveness for mating preferences. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-018-28368-z

Hoogenboom, M. O., Metcalfe, N. B., Groothuis, T. G. G., de Vries, B., & Costantini, D. (2012). Relationship between oxidative stress and circulating testosterone and cortisol in pre-spawning female brown trout. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *163*(3–4), 379–387. doi: 10.1016/j.cbpa.2012.07.002

Huang, Y., Li, Y., Burt, D. W., Chen, H., Zhang, Y., Qian, W., … Li, N. (2013). The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nature Genetics*, *45*(7), 776–783. doi: 10.1038/ng.2657

Ketterson, E. D., and V. Nolan. (1999). Adaptation, exaptation, and constraint: a hormonal perspective. American Naturalist 154(suppl.):S4-S25.

Kim, J., Kim, H., & Son, C. (2018). Tissue-Specific Profiling of Oxidative Stress-Associated Transcriptome in a Healthy Mouse Model. *International Journal of Molecular Sciences, 19*(10), 3174. doi: 10.3390/ijms19103174

Kinsella, R. J., Kahari, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., … Flicek, P. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database*, *2011*(0), bar030–bar030. doi: 10.1093/database/bar030

Kirkpatrick, M., & Ryan, M. J. (1991). The evolution of mating preferences and the paradox of the lek. *Nature*, *350*(6313), 33–38. doi: 10.1038/350033a0

Kissick, H. T., Sanda, M. G., Dunn, L. K., Pellegrini, K. L., On, S. T., Noel, J. K., & Arredouani, M. S. (2014). Androgens alter T-cell immunity by inhibiting T-helper 1 differentiation. *Proceedings of the National Academy of Sciences*, *111*(27), 9887–9892. doi: 10.1073/pnas.1402468111

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

Langfelder, P., Luo, R., Oldham, M. C., & Horvath, S. (2011). Is My Network Module Preserved and Reproducible? *PLoS Computational Biology*, *7*(1), e1001057. doi: 10.1371/journal.pcbi.1001057

Lin, A. A., Wojciechowski, S. E., & Hildeman, D. A. (2010). Androgens suppress antigen-specific T cell responses and IFN-γ production during intracranial LCMV infection. *Journal of Neuroimmunology*, *226*(1–2), 8–19. doi: 10.1016/j.jneuroim.2010.05.026

Louder, M. I. M., Hauber, M. E., & Balakrishnan, C. N. (2018). Early social experience alters transcriptomic responses to species-specific song stimuli in female songbirds. *Behavioural Brain Research*, *347*, 69–76. doi: 10.1016/j.bbr.2018.02.034

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550–550. doi: 10.1186/s13059-014-0550-8

Marini, F., & Binder, H. (2016). Development of Applications for Interactive and Reproducible Research: a Case Study. *Genomics and Computational Biology*, *3*(1), 39. doi: 10.18547/gcb.2017.vol3.iss1.e39

Monson, C., Forsgren, K., Goetz, G., Harding, L., Swanson, P., & Young, G. (2017). A teleost androgen promotes development of primary ovarian follicles in coho salmon and rapidly alters the ovarian transcriptome†. *Biology of Reproduction*, *97*(5), 731–745. doi: 10.1093/biolre/iox124

Mougeot, F., Martínez-Padilla, J., Webster, L. M. I., Blount, J. D., Pérez-Rodríguez, L., & Piertney, S. B. (2009). Honest sexual signalling mediated by parasite and testosterone effects on oxidative balance. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1093–1100. doi: 10.1098/rspb.2008.1570

Neefjes, J., Jongsma, M. L. M., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*, *11*(12), 823–836. doi: 10.1038/nri3084

Nelson, R. J. (2011). An Introduction to Behavioral Endocrinology. Oxford University Press, Incorporated.

Owen-Ashley, N. T., D. Hasselquist, and J. C. Wingfield. (2004). Androgens and the immunocompetence handicap hypothesis: unraveling direct and indirect pathways of immunosuppression in song sparrows. American Naturalist 164:490–505.

Page, S. T., Plymate, S. R., Bremner, W. J., Matsumoto, A. M., Hess, D. L., Lin, D. W., … Wu, J. D. (2006). Effect of medical castration on CD4 $^{+}$ CD25 $^{+}$ T cells, CD8 $^{+}$ T cell IFN-γ expression, and NK cells: a physiological role for testosterone and/or its metabolites. *American Journal of Physiology-Endocrinology and Metabolism*, *290*(5), E856–E863. doi: 10.1152/ajpendo.00484.2005

Peters, A. (2000). Testosterone treatment is immunosuppressive in superb fairy–wrens, yet free–living males with high testosterone are more immunocompetent. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1446), 883–889. doi: 10.1098/rspb.2000.1085

Rettew, J. A., Huet-Hudson, Y. M., & Marriott, I. (2008). Testosterone Reduces Macrophage Expression in the Mouse of Toll-Like Receptor 4, a Trigger for Inflammation and Innate Immunity. *Biology of Reproduction*, *78*(3), 432–437. doi: 10.1095/biolreprod.107.063545

Roberts, M. L., Buchanan, K. L., & Evans, M. R. (2004). Testing the immunocompetence handicap hypothesis: a review of the evidence. *Animal Behaviour*, *68*(2), 227–239. doi: 10.1016/j.anbehav.2004.05.001

Rose, M. E. (1979). The Immune System in Birds. *Journal of the Royal Society of Medicine*, *72*(9), 701–705. doi: 10.1177/014107687907200914

Schat, K., Kaspers, B., & Kaiser, P. (Eds.). (2014). *Avian Immunology*. doi: 10.1016/C2011-0-07243-5

Schlinger, B. A., Barske, J., Day, L., Fusani, L., & Fuxjager, M. J. (2013). Hormones and the neuromuscular control of courtship in the golden-collared manakin (Manacus vitellinus). *Frontiers in Neuroendocrinology*, *34*(3), 143–156. doi: 10.1016/j.yfrne.2013.04.001

Segner, H., Verburg-van Kemenade, B. M. L., & Chadzinska, M. (2017). The immunomodulatory role of the hypothalamus-pituitary-gonad axis: Proximate mechanism for reproduction-immune trade offs? *Developmental & Comparative Immunology*, *66*, 43–60. doi: 10.1016/j.dci.2016.07.004

Sokol, C. L., & Luster, A. D. (2015). The Chemokine System in Innate Immunity. *Cold Spring Harbor Perspectives in Biology*, *7*(5), a016303. doi: 10.1101/cshperspect.a016303

Stearns, S. C. (1992). *The Evolution of Life Histories*. Oxford, New York: Oxford University Press.

Taff, C. C., & Freeman-Gallant, C. R. (2014). An experimental test of the testosterone mediated oxidation handicap hypothesis in a wild bird. *Hormones and Behavior*, *66*(2), 276–282. doi: 10.1016/j.yhbeh.2014.05.006

Thomas, M. P., Erneux, C., & Potter, B. V. L. (2017). SHIP2: Structure, Function and Inhibition. *ChemBioChem*, *18*(3), 233–247. doi: 10.1002/cbic.201600541

Trigunaite, A., Dimo, J., & Jørgensen, T. N. (2015). Suppressive effects of androgens on the immune system. *Cellular Immunology*, *294*(2), 87–94. doi: 10.1016/j.cellimm.2015.02.004

Veillette, A. (2010). SLAM-Family Receptors: Immune Regulators with or without SAP-Family Adaptors. *Cold Spring Harbor Perspectives in Biology*, *2*(3), a002469–a002469. doi: 10.1101/cshperspect.a002469

Wang, M., Jiang, S., Wu, W., Yu, F., Chang, W., Li, P., & Wang, K. (2018). Non-coding RNAs Function as Immune Regulators in Teleost Fish. *Frontiers in Immunology*, *9*. doi: 10.3389/fimmu.2018.02801

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57–63. doi: 10.1038/nrg2484

Wapstra, E., van de Crommenacker, J., Groothuis, T., Komdeur, J., Olsson, M., While, G., … Isaksson, C. (2011). Aggression, but not testosterone, is associated to oxidative status in a free-living vertebrate. *Behaviour*, *148*(5–6), 713–731. doi: 10.1163/000579511X574204

Wenzel, M. A., Webster, L. M. I., Paterson, S., Mougeot, F., Martínez-Padilla, J., & Piertney, S. B. (2013). A transcriptomic investigation of handicap models in sexual selection. *Behavioral Ecology and Sociobiology*, *67*(2), 221–234. doi: 10.1007/s00265-012-1442-0

Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214. doi: 10.1016/0022-5193(75)90111-3

Zera, A. J., & Harshman, L. G. (2001). The Physiology of Life History Trade-Offs in Animals. *Annual Review of Ecology and Systematics*, *32*(1), 95–126. doi: 10.1146/annurev.ecolsys.32.081501.114006

Zheng, Y., Yuan, J., Meng, S., Chen, J., & Gu, Z. (2019). Testicular transcriptome alterations in zebrafish (Danio rerio) exposure to 17β-estradiol. *Chemosphere*, *218*, 14–25. doi: 10.1016/j.chemosphere.2018.11.092

Zou, Q., Jin, J., Xiao, Y., Hu, H., Zhou, X., Jie, Z., … Sun, S.-C. (2015). T cell development involves TRAF3IP3-mediated ERK signaling in the Golgi. *The Journal of Experimental Medicine*, *212*(8), 1323–1336. doi: 10.1084/jem.20150110

CHAPTER 3. PARENT AND OFFSPRING GENOTYPES INFLUENCE GENE EXPRESSION IN EARLY LIFE

Abstract

Parents can have profound effects on offspring fitness. Little, however, is known about the mechanisms through which parental genetic variation influences offspring physiology in natural systems. White-throated sparrows (*Zonotrichia albicollis*, WTSP) exist in two genetic morphs, tan and white, controlled by a large polymorphic supergene. Morphs mate disassortatively, resulting in two pair types: tan male x white female (TxW) pairs, which provide biparental care and white male x tan female (WxT) pairs, which provide female-biased care. To investigate how parental composition impacts offspring, we performed RNA-seq on whole blood of WTSP nestlings sampled from nests of both pair types. Parental pair type had a large effect on nestling gene expression, with 881 genes differentially expressed (DE) and seven correlated gene co-expression modules. The DE genes and modules up-regulated in WxT nests with female-biased parental care primarily function in metabolism and stress-related pathways resulting from the overrepresentation of proteolysis and stress response genes (e.g. SOD2, NR3C1). These results show that parental genotypes and/or associated behaviors influence nestling physiology, and highlight avenues of further research investigating the ultimate implications for the maintenance of this polymorphism. Nestlings also exhibited morph-specific gene expression, with 92 differentially expressed genes, comprising innate immunity genes and genes encompassed by the supergene. Remarkably, we identified the same regulatory hub genes in these blood-derived expression networks as were previously identified in adult WTSP brains (EPM2A, BPNT1, TAF5L). These hub genes were located within the supergene, highlighting the importance of this gene complex in structuring regulatory networks across diverse tissues.

Keywords

Transcriptome, parental effects, early life stress, nestling, RNAseq, ornithology

Introduction

Parents can have profound impacts on offspring development and fitness. Parental effects can manifest throughout the developmental period, both pre- and post-natally (reviewed in Meaney 2001, Lupien et al. 2009) and can be mediated through parental behaviors, genetics and physiology during early development (Trivers 1972). Parents play a substantial role in establishing the early life environment of offspring. For example in birds, parental decisions on nest placement, incubation behavior, and nest defense could strongly impact developmental conditions of the egg. These parental behaviors will impact exposure to sunlight, humidity, temperature, and other environmental impacts of the eggs, which can influence developmental physiology (e.g. Nord & Nilsson 2011). In addition to parental behaviors, prenatal effects often arise via physiological maternal effects. Developing offspring are susceptible to the maternally created environment (e.g. maternal hormones, immune state, nutrition), which influence offspring physiology (Mousseau & Fox 1998, Jacquin et al. 2012; reviewed in Gluckman et al. 2008, Wolf & Wade 2009, Cottrell & Secki 2009).

The magnitude of parental effects, particularly in altricial species, is likely largest during the postnatal period, when offspring rely entirely on the parents for provisioning and growth (Royle et al. 2012). Provisioning plays a prominent role in offspring development, with the quality and quantity of food items crucial for offspring development (van Oers et al. 2015, Griebel et al. 2019). Similar to the prenatal stage, parental behaviors could also have strong

impacts on offspring physiology. In many species, offspring are left alone during parental foraging trips, increasing environmental exposure (Lloyd and Martin 2004) and predation risk (Lima 2009). Parental separation can also increase offspring anxiety (Millstein & Holmes 2007). Siblings must also compete to optimize provisioning, brooding warmth, and preening (Mock & Parker 1997). Thus, this postnatal environment, largely mediated through parental effects, can be a potential source of early life stress (ELS) in offspring, which may result in life-long fitness effects (reviewed in Monaghan 2014).

ELS has broad effects on organisms, including impaired neural development, neuroendocrine signaling, behavior, and physiology (McEwen 2007, Monaghan 2014). For example, ELS is associated with impaired neuroendocrine function and corresponding impaired hypothalamic-pituitary-adrenal (HPA) development, which leads to increase stress response sensitivity later in life (e.g. Heim et al. 2008, Spencer et al. 2009, Crespi et al. 2012, Spencer 2017). ELS can exacerbate behavioral alterations as organisms develop and mature including symptoms of anxiety and depression in the postnatal environment (Noguera et al. 2017) and result in impaired behavior as reproductive adults (e.g. Krause et al. 2009, reviewed in Bolton et al. 2017). While the organismal effects of ELS are well studied, the genetic underpinnings are relatively underexplored. Much of the genetic work in the context of ELS has focused on gene regulatory impacts, particularly in mammalian biomedical models (reviewed in Szyf et al. 2007, Szyf 2009, Silberman et al. 2016, Alyamani & Murgatroyd 2018). In particular, the quality of parental care can have strong impacts on offspring health resulting from epigenetic modifications (Liu et al. 1997, Meaney 2001, Weaver et al. 2004). These gene regulation studies primarily use changes in DNA methylation as an indicator of ELS (Murgatroyd et al. 2009, Kinnally et al. 2011, Lewis & Olive 2014) and recent work has expanded these approaches into non-

mammalian organisms (e.g. Rubenstein et al. 2016, Moghadam et al. 2017, Pértille et al. 2017, Gott 2018, Sheldon et al. 2018). DNA methylation studies of ELS investigate changes to the structure of DNA, but are often limited in the functional implications of ELS (i.e. transcription and translation). In general, these modifications are thought to alter transcriptional activity of genes in the modified genomic region (Berger 2007, Lowdon et al. 2016). Indeed, several studies have also taken candidate gene approaches to investigating gene expression in the context of ELS (Marco et al. 2014, Diaz-Real et al. 2017, Anastasiadi et al. 2018, Reshetnikov et al. 2018). However, very few studies assess genome-wide transcription under ELS (Moghadam et al. 2017), particularly in the context of parental effects (but see: Weaver et al. 2006).

In this study, we examined the white-throated sparrow (*Zonotrichia albicollis*, WTSP) to assess the role of parental genotype on offspring gene expression. WTSPs exist in two plumage morphs, tan (T) and white (W), that are found in both sexes and in roughly equal frequencies (Lowther 1961). These morphs are genetically determined by alternative alleles of a supergene, a group of linked genes that are inherited together, show limited recombination, and maintain complex behavioral traits (i.e. WTSP morphs; Schwander et al. 2014, Taylor & Campagna 2016). The WTSP supergene resulted from a complex chromosomal rearrangement comprising multiple inversions (hereafter referred to as "inversion" or "inverted"). This inversion contains ~1,100 genes on chromosome two, termed ZAL2^{m} (Throneycroft 1975, Thomas et al. 2008, Romanov et al. 2009, Tuttle et al. 2016). W morphs are nearly always heterozygous for the inversion (ZAL2/ZAL2^{m}) and T morphs are always homozygous (ZAL2/ZAL2; Thorneycroft 1966, 1975).

This unusual polymorphism in WTSPs influences hormonal profiles and the behavior of both sexes, and thus has the potential to influence pre- and post-natal environments for the

offspring of different morphs. W morph males maintain higher levels of testosterone during the pre-laying, incubation, and brooding stages and oestradiol during the laying and brooding stages (Horton et al. 2014). Only oestradiol has been shown to differ between adult female morphs during the breeding season and is higher in W morph females during the pre-laying and laying stages (Horton et al. 2014). These genetic and hormonal differences also translate into striking behavioral differences. W morphs, for example, are highly territorial and sing frequently whereas T morphs are far less territorial and aggressive (Lowther 1962, Kopachena & Falls 1993, Tuttle 2003, Horton & Holberton 2010, Horton et al. 2014). More importantly from the perspective of offspring, males of each morph also differ in paternal investment (Knapton & Falls 1983, Horton et al. 2014). W morph males are promiscuous and provision nestlings very little. T morph males defend their within-pair paternity through mate guarding and are highly paternal. Females tend to provision at intermediate levels, but T morph females may compensate for unassisted care from W morph males and provision more than W morph females (Knapton & Falls 1983). A final wrinkle in this complex mating system is that morphs nearly always mate with the opposite morph (98.5%, Tuttle et al. 2016), resulting in two stable pair types: T male x W female (TxW) and W male x T female (WxT) (Lowther 1961, Tuttle 2003, Tuttle et al. 2016). Because males differ in paternal investment, this results in two distinct parental care strategies. TxW pairs provide biparental care and WxT pairs provide female-biased parental care. In this study we examined gene expression profiles of offspring from both pair-types in order to assess the physiological consequences of variation in parental genotype.

Methods

Field based sample collection

All nestling whole blood samples in this study came from a breeding population of WTSPs at the Cranberry Lake Biological Station in northern New York, USA (SUNY-ESF, 44.15ºN, 74.78ºW) and were collected during the 2015 breeding season. We only utilized samples collected during the first clutch (June 6 - June 14, 2015), as WTSP males may increase paternal investment in replacement broods (Horton et al. 2014). We collected ~80µL blood in capillary tubes via brachial venipuncture on days 5-7 post-hatch. Approximately 60µL blood was preserved in Longmire's lysis buffer (Longmire et al. 1992) for genotyping and ~20µL was immediately placed in RNAlater. Within six hours of collection, samples were placed temporarily into liquid nitrogen, before being shipped overnight on dry ice to -80ºC storage until RNA extraction. All animal sampling protocols were approved by the Indiana State University Institutional Animal Care and Use Committee (IACUC 562158-1:ET/RG, 562192-1:ET/RG).

Molecular sexing & genotyping

Nestling DNA was extracted from erythrocytes using the DNA IQ® magnetic extraction system (Promega Corp, Madison, WI USA). To determine sex and morph, we used PCR to fluorescently label and amplify a region of the chromo-helicase-DNA-binding (CHD) gene, and a region of the vasoactive intestinal peptide following Griffiths et al. (1998) and Michopolous et al. (2007). The PCR products were run and analyzed on an ABI PRISM^{TM} 310 genetic analyzer.

RNA extraction, library preparation, & sequencing

We sampled a total of 52 nestlings for RNA extraction, but due to issues with RNA quality after extraction, only 32 were used for sequencing. These samples represent 23 nestlings

from eight TxW pairs and nine nestlings from three WxT pairs. Additionally, these data represent 18 females, 14 males, 15 T morph, and 17 W morph individuals.

We removed RNAlater and homogenized whole blood tissue samples with Tri-Reagent (Molecular Research Company). Total RNA was purified with a Qiagen RNeasy mini kit (Valencia, CA, USA), followed by DNase treatment and further purification. We quality assessed RNA with an Agilent Bioanalyzer (RIN > 7) (Wilmington, DE, USA). Both library preparation and sequencing were performed at the University of Illinois Roy J. Carver Biotechnology Center. A library was prepared for each RNA sample using the Illumina HT TruSeq (San Diego, CA, USA) stranded RNA sample prep kit. Libraries were distributed into four pools with equimolar concentrations and quantitated via qPCR. Each of the pools was sequenced on an individual lane of an Illumina HiSeq 2500 using the Illumina TruSeq SBS sequencing kit v4 producing 100-nucleotide single-end reads.

Creation of masked reference genome

The WTSP reference genome was generated from a male T morph individual (Tuttle et al. 2016). Thus, the reference genome does not contain any sequence data from the $ZAL2^m$ inversion. To avoid any potential bias in mapping reads derived from W morph individuals onto a T morph genome, we generated a masked reference genome for this study. To do so, we used previously published whole genome sequences from three W morph adults (Tuttle et al. 2016). Reads were adapter trimmed with *Trim Galore!* v0.3.8 (https://github.com/FelixKrueger/TrimGalore) and aligned to the WTSP reference genome with *bwa mem* v 0.7.10-r789 (Li 2013). We converted and sorted the resulting SAM alignment to BAM format with *samtools view* and *samtools sort*, respectively (*samtools* v1.2, Li et al. 2009).

We then merged all genomic scaffolds corresponding to the ZAL2^{m} inversion, as identified in Tuttle et al. (2016), with *samtools merge*. We called SNPs within the inversion using *samtools mpileup* and *bcftools call* v 1.2 (Li et al. 2009, Li 2011). We only kept SNPs that were heterozygous in each of the three individuals with *SnpSift* v 4.3p (Cingolani et al. 2012) and used these SNPs to mask the reference genome with *bedtools maskfasta* v 2.21.0 (Quinlan & Hall 2010).

Quality control, read mapping, differential expression, & gene ontology

We trimmed Illumina sequencing adapters from each of the 32 libraries with *Trim Galore!* v0.3.8 which uses *Cutadapt* v1.7.1 (Martin 2011). Trimmed reads were then mapped to the masked reference genome with *STAR* v2.5.3a (Dobin et al. 2013). The mapping results were then quantified and assigned gene IDs with *htseq-count* v0.6.0 (Anders et al. 2015) specifying '-s reverse' and '-i gene'. We then removed lowly expressed genes by summing the counts for each gene across all 32 samples, dividing by 32 to obtain the study average, and removing genes with an average read count of < 5.

All statistical analyses were performed with R v3.5.0 (R Core Team 2018). We first identified outlier samples based on visual inspection of sample distance in a dendrogram within *WGCNA* (Horvath 2011). Two samples, one T female and one T male representing an entire TxW nest, were identified as outliers and removed from all future analyses (Figure S1). Using the remaining 30 samples, we normalized reads accounting for sequencing depth and assessed differential expression with *DEseq2* (Love et al. 2014). We performed variance stabilizing transformation of reads in *DEseq2* and performed PCA and hierarchical clustering based on Euclidean distance of gene expression profiles with *pcaExplorer* v2.6.0 (Marini & Binder 2016).

Differential expression analyses utilized pairwise comparisons between nestling morph and pair type (i.e. parental morphs). We controlled for sex in morph comparisons and sex, morph, and nest ID for pair type comparisons. To include nest ID in the pair type comparison, we followed the "individuals nested within groups" guide in the *DEseq2* manual. We did not include nestling age in analyses, as most samples were 6 days old (n=21), limiting comparisons with nestlings aged Day 5 (n=3) or Day 7 (n=6). Network analysis (see below) did not reveal any effect of age on variables of interest (morph, pair type; data not shown). We also tested for an interaction between nestling morph and pair type utilizing a grouping variable as outlined in the *DEseq2* manual. *DEseq2* determines differential expression with a Wald test followed by Benjamini & Hochberg (1995) FDR correction. Genes were considered differentially expressed (DE) if the FDR corrected p-value was < 0.10. Details for each model run, including the R code used, can be found at https://github.com/danielnewhouse/wtsp.

We next tested for gene ontology (GO) enrichment among DE genes with *GOrilla* (Eden et al. 2007, 2009). For each *DEseq2* comparison, we ordered the list of genes based on ascending FDR values, excluding any genes in which *DEseq2* did not assign a FDR value. The WTSP genome is not completely annotated, so any loci without a gene symbol were excluded from GO analyses (n=1,926). *GOrilla* places greater weight on genes located at the top of the list (i.e. DE genes), while accounting for the contribution of each gene in the given comparison. GO categories were considered significantly enriched if the FDR corrected p-value <0.05. *GOrilla* does not support WTSP annotation; so, all analyses were based on homology to human gene symbols.

Weighted gene co-expression network analysis (WGCNA)

We used the *WGCNA* package in R (Zhang & Horvath 2005, Langfelder & Horvath 2008) to identify modules of co-expressed genes in our dataset. We first exported variance stabilizing transformed (vst) read counts from *DEseq2*, removed genes with an average vst < 5 averaged across all 30 samples, and imported the subsequent list of 8,982 genes into *WGCNA*. To build the co-expression matrix, we chose a soft thresholding power (β) value of 12, at which the network reaches scale-free topology (Figure S2). We generated a signed network with minimum module size of 30 genes and merged highly correlated modules (dissimilarity threshold = 0.20, which corresponds to $R^2 = 0.80$). We then correlated the eigengene, which is the first principal component of a module, of these merged modules with external traits (pair type, morph, sex, nest ID). Modules with $p < 0.05$ were considered significantly correlated with a given trait. For all morph-specific results, we tested for an enrichment of inversion genes using a Fisher's exact test ($\alpha < 0.05$).

To visualize the interaction of genes within a module, we generated the intramodular connectivity (IM) score for each gene, which represents the interconnection of module genes. We exported all IM scores for modules of interest and imported into *VisAnt* v5.51 (Hu et al. 2013) for visualization. To maximize network clarity, we only plotted the top 300 interactions based on IM scores. Thus, we only visualized the most connected genes. To identify hub genes, we visualized the Degree Distribution (DD) for the network and selected the most connected genes above a natural break in the distribution. This resulted in one to nine hub genes per module.

To understand the biological function of modules correlated with traits of interest, we performed a target vs background GO analysis in *GOrilla*. For each module, we tested the

assigned genes for each module against the entire list of 8,982 genes used for the *WGCNA* analysis. GO categories were significant with a FDR corrected p-value < 0.05.

Results

Sequencing results

We sequenced each sample to an average depth of 29.4 million reads (range = 16.2-58.5 million reads). The 32 libraries were distributed into four pools in equimolar concentration. One pool contained only four samples, which corresponded to the four samples with lowest RNA concentrations. This pool was sequenced to an average depth of 56.17 million reads per library. The remaining three pools were sequenced to an average depth of 25.62 million reads per library. Samples mapped to our masked genome at an average rate of 91.08% (range = 88.19%-92.87%) (Table S1). A total of 8,982 genes had count values ≥ 5 across all samples, which included 641 located in the W morph inversion. Samples did not segregate by pair type or morph in PCA or hierarchical clustering (Figures S3, S4).

Differential expression – morph

Ninety-two genes were differentially expressed between morphs. Sixty-five of these genes (71%) were located in the inversion, representing a significant enrichment (χ^2=553.73, df=1, p<0.00001) (Table S2). The inversion represents only 641 out the 8,892 genes (7%) sampled here. Additionally, expression of many of these 92 genes was elevated in W morph nestlings and a number of these genes had well-known functions in innate immunity (e.g. IFIT5, IL20RA, EIF2AK2, RSAD2). There was GO enrichment of four categories, two of which are

immunity related: “immune response” ($p = 0.019$) and “defense response to virus” ($p = 0.049$) (Table S3).

Differential expression – pair type

Pair type had the largest effect on gene expression, with 881 genes DE between offspring from the two different pair types (FDR < 0.10, Table S2). Many genes associated with stress responses were elevated in nestlings in WxT nests, including the glucocorticoid receptor (NR3C1), superoxide dismutase (SOD)1 & SOD2, DEP domain-containing mTOR-interacting protein (DEPTOR), and several ubiquitin-mediated proteolysis pathway genes (e.g. UBE2D3, PSMD3, PSMD6). Additionally, several immune system related genes were also elevated in WxT nests, including cytokines (e.g. IL2RA, IL7R), suppressor of cytokine signaling 1 (SOCS1), and five putative major histocompatibility complex (MHC) class I loci. No GO categories were significantly enriched, however.

We next tested for a morph-specific response to pair type. Within WxT nests, 40 genes were DE ($p < 0.10$) between T and W morph nestlings. Twelve of these genes (30%) are located within the inversion, again reflecting an enrichment of inversion genes among those differentially expressed between morph ($\chi^2=34.44$, df=1, $p<0.00001$). Only two genes (THSD7B & CFAP44) were DE between morphs within TxW nests, both of which are uniquely DE between morphs in TxW nests. No GO categories were enriched in either comparison.

WGCNA – morph

WGCNA revealed 26 modules, five of which were correlated with morph (Table 3.1, Figure 3.1). The light cyan module (183 genes, $R^2=0.67$, $p=5x10^{-5}$) and ivory module (72 genes,

R^2=-0.66, p=9x10^{-5}) contained genes elevated and suppressed, respectively, in W morph nestlings relative to T morph nestlings. These modules are both enriched for genes located within the chromosomal inversion (light cyan module = 70/183 (38%) genes, χ^2=266.49, df=1, p<0.00001; ivory module = 40/72 (56%), χ^2=261.60, df=1, p<0.00001) (Figure S5). The hubs of each of these modules are also located in the chromosomal inversion (Table 3.1, Figure S5). Additionally, the sky blue module (58 genes, R^2=0.53, p=0.003) and dark red module (102 genes, R^2=0.47, p=0.009) (Figure S6) contained genes elevated in W morph nestlings and many of these genes overlap with the immune related genes described in the morph DE tests above. The hubs of these networks (e.g. sky blue: EIF2AK2, IFIT5, OASL; dark red: TRAF5) (Table 3.1) reflect a conserved innate immunity network structure in avian blood (see also Kernbach et al., in review) (Figure S6).

Module	**R^2**	**p-value**	**Hub genes**	**DD of hub gene(s)**
Dark Red	0.47	0.009	TRAF5	32
Ivory	-0.66	9x10^{-5}	GOPC, HDAC2, HINT3, TAF5L, TRMT61B, MARC2	>29
Light Cyan	0.67	5x10^{-5}	BPNT1, EPM2A, LOC102066536 (GST-like), MAN1A1, MEI4, RNASET2, SLC18B1, TTC32	>27
Salmon	-0.5	0.005	NSL1	39

Sky Blue	0.53	0.003	DTX3L, EIF2AK2, IFIT5, LOC102064521 (OASL), LOC102065196 (IFI27L2), PARP9, PARP14, RSAD2, ZNFX1	>22

Table 3.1. WGCNA modules correlated with morph, strength of correlation (R^2), p-value, hub gene(s) of module, and the degree distribution of hub gene(s).

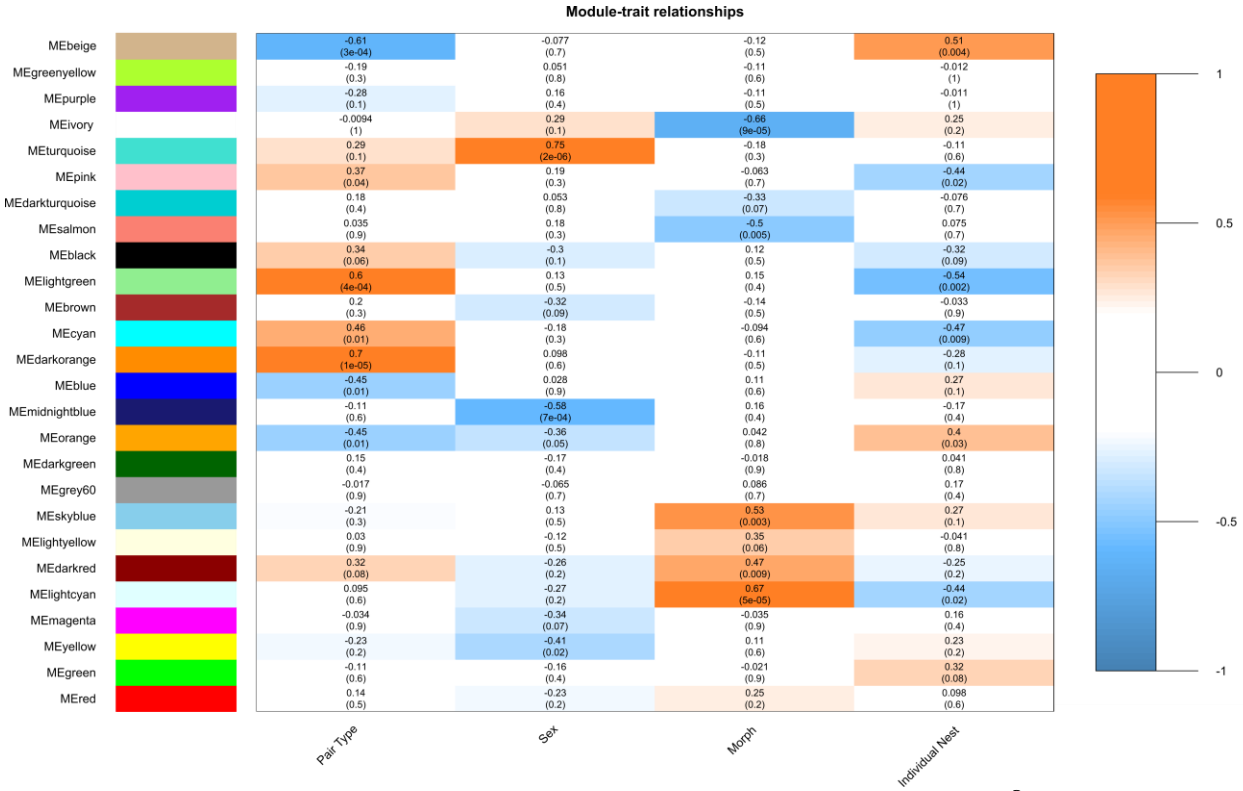


Figure 3.1. WGCNA module-trait correlation matrix. Each box contains the R^2 correlation value followed by p-value in parentheses of a given trait with the module. Correlation values range from -1 to 1, with orange colors representing positive correlation and blue colors representing negative correlation.

WGCNA – pair type

We found seven modules correlated with pair type (Table 3.2, Figure 3.1). The blue module represented genes that are elevated in nestlings from WxT nests (1,142 genes, R^2 = -0.45, p=0.01). This module contained both the largest number of genes and correspondingly strongest functional enrichment. Many of these GO enrichments were related to protein function, resulting from the presence of ribosomal genes. Interestingly, several GO categories for metabolism, catabolism, and proteolysis were also enriched, driven by genes encoding ubiquitin-conjugating enzymes and proteasome subunits (e.g. "proteasomal protein catabolic process", $p=2.34x10^{-4}$; "proteasome-mediated ubiquitin-dependent protein catabolic process", $p=5.32x10^{-4}$) (Table S4). Many of these (e.g. PSMF1, PSMD3, PSMD6, UBE2D2, UBE2D3, UBE3C) were also DE between offspring of the two pair types (Figure 3.2). Lastly, the blue module contains one hub gene, NDUFB3 (DD=42) (Figure 3.2), which is involved in the mitochondrial electron transport chain.

Module	**R^2**	**p-value**	**Hub genes**	**DD of hub gene(s)**
Beige	-0.61	$3x10^{-4}$	DEPTOR	39
Blue	-0.45	0.01	NDUFB3	42
Cyan	0.46	0.01	HELZ	36
Dark Orange	0.7	$1x10^{-5}$	NCOA6	45
Light Green	0.6	$4x10^{-4}$	CDK19, CHD4, EPG5	>28
Orange	-0.45	0.01	ZFX	31
Pink	0.37	0.04	LOC102060916 (C12orf4)	19

Table 3.2. WGCNA modules correlated with pair type, strength of correlation (R^2), p-value, hub gene(s) of module, and the degree distribution of hub gene(s).

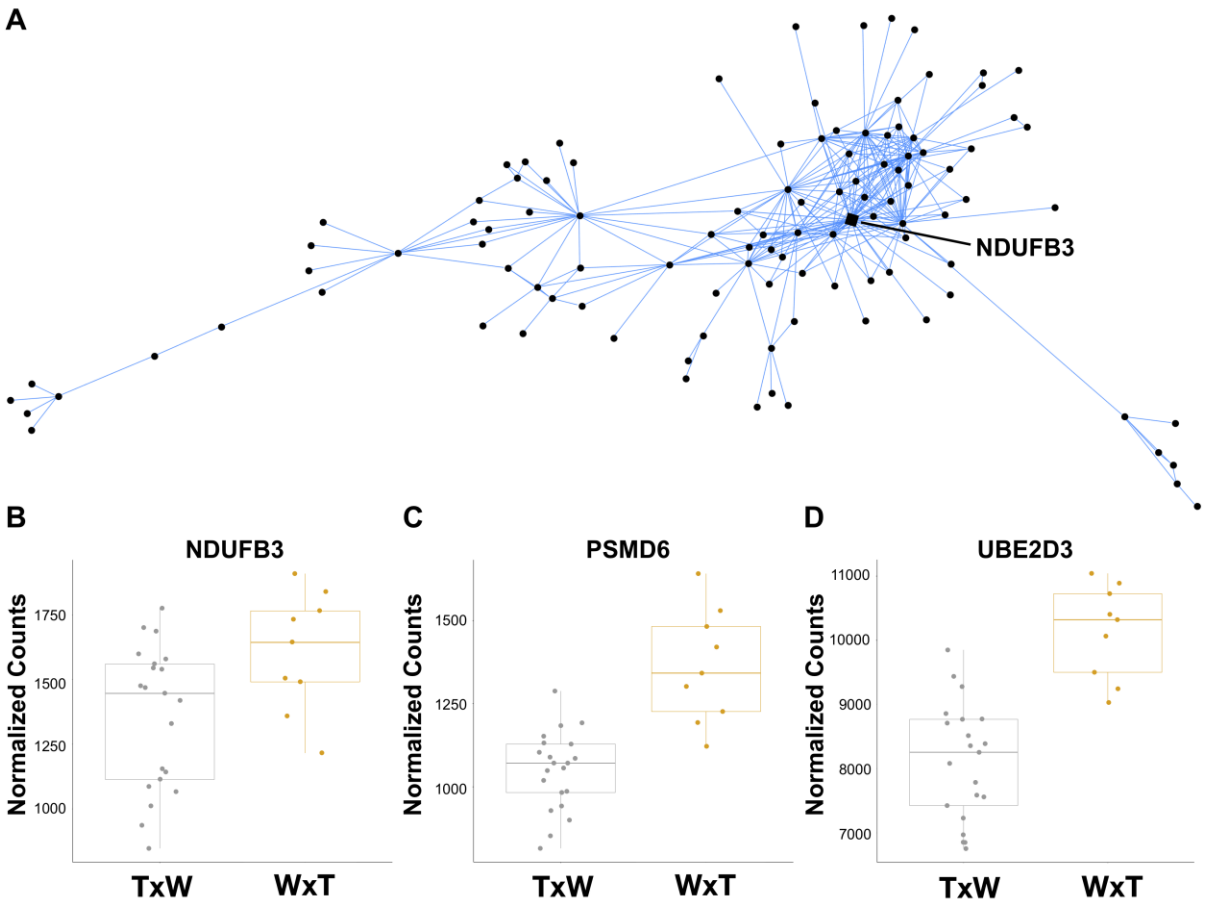


Figure 3.2. (A) Network of blue module, highlighting hub gene NDUFB3, along with normalized expression plots of (B) NDUFB3, as well as ubiquitin-mediated proteolysis-related genes (C) PSMD6 and (D) UBE2D3. TxW represents samples from nests sired by a T male and a W female. WxT represents samples from nests sired by a W male and a T female. Each circle represents a gene and diamonds represent hub genes described in Table 3.2.

The beige and light green modules represented candidate stress response networks. These modules showed contrasting expression patterns in nestlings from WxT nests (Figure 3.3 & 3.4). Although not significantly enriched for any GO categories, the beige module comprised 335 genes that were upregulated in WxT nests relative to TxW nests (R^2=-0.61, p=3x10^{-4}). DEPTOR, which functions as an inhibitor of the mTOR pathway in response to stress (e.g. Desantis et al. 2015), was the single hub in the beige module (DD=39, Figure 3.3). The beige module also

contained NR3C1, which is activated in response to increased glucocorticoid secretion. Lastly, the light green module (116 genes, R^2=0.60, p=4x10^{-4}) contained genes with low expression in TxW nests relative to WxT nests. There were three hub genes (DD > 28), CDK19, CHD4, and EPG5, each with previously described roles in the stress response (Figure 3.4).

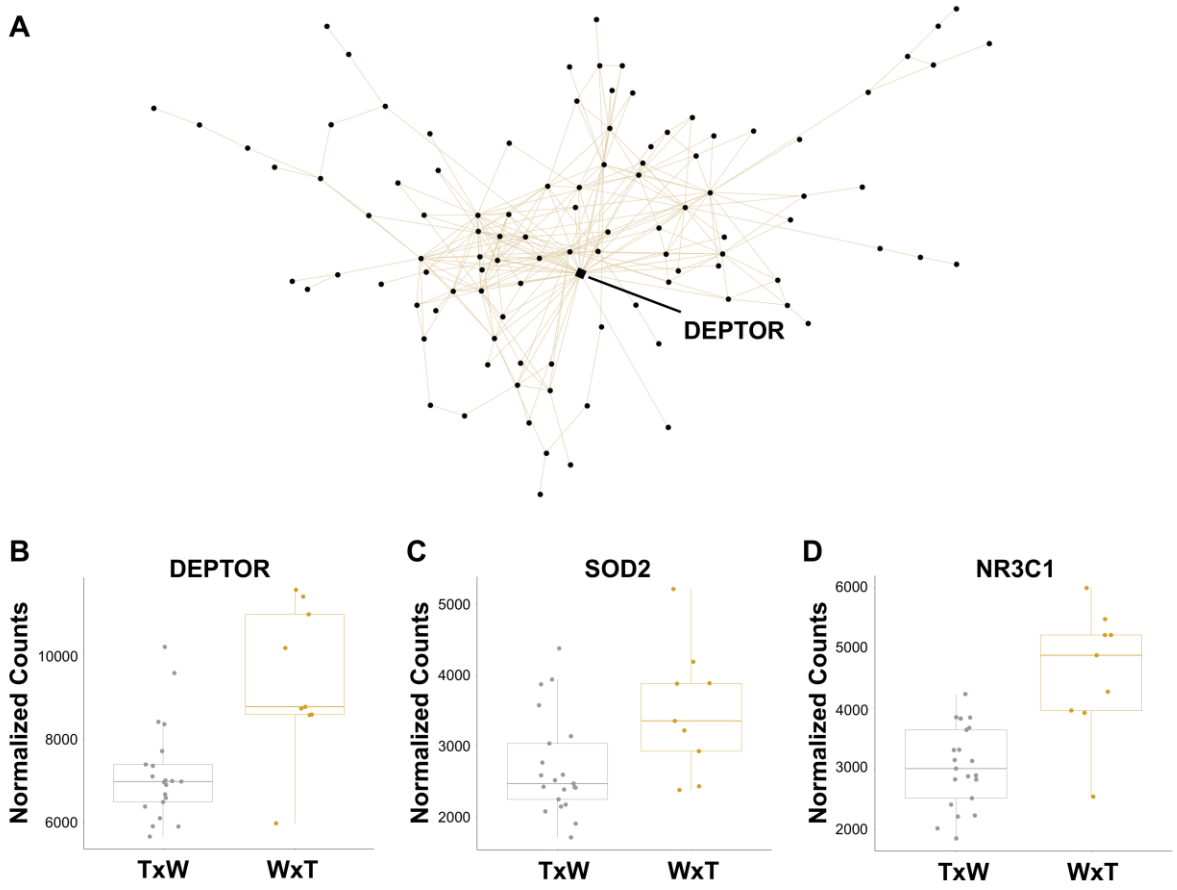


Figure 3.3. (A) Network of beige module, highlighting hub gene DEPTOR, along with normalized expression plots of hub gene (B) DEPTOR, as well as stress responsive genes (C) SOD2 and (D) NR3C1. TxW represents samples from nests sired by a T male and a W female. WxT represents samples from nests sired by a W male and a T female.

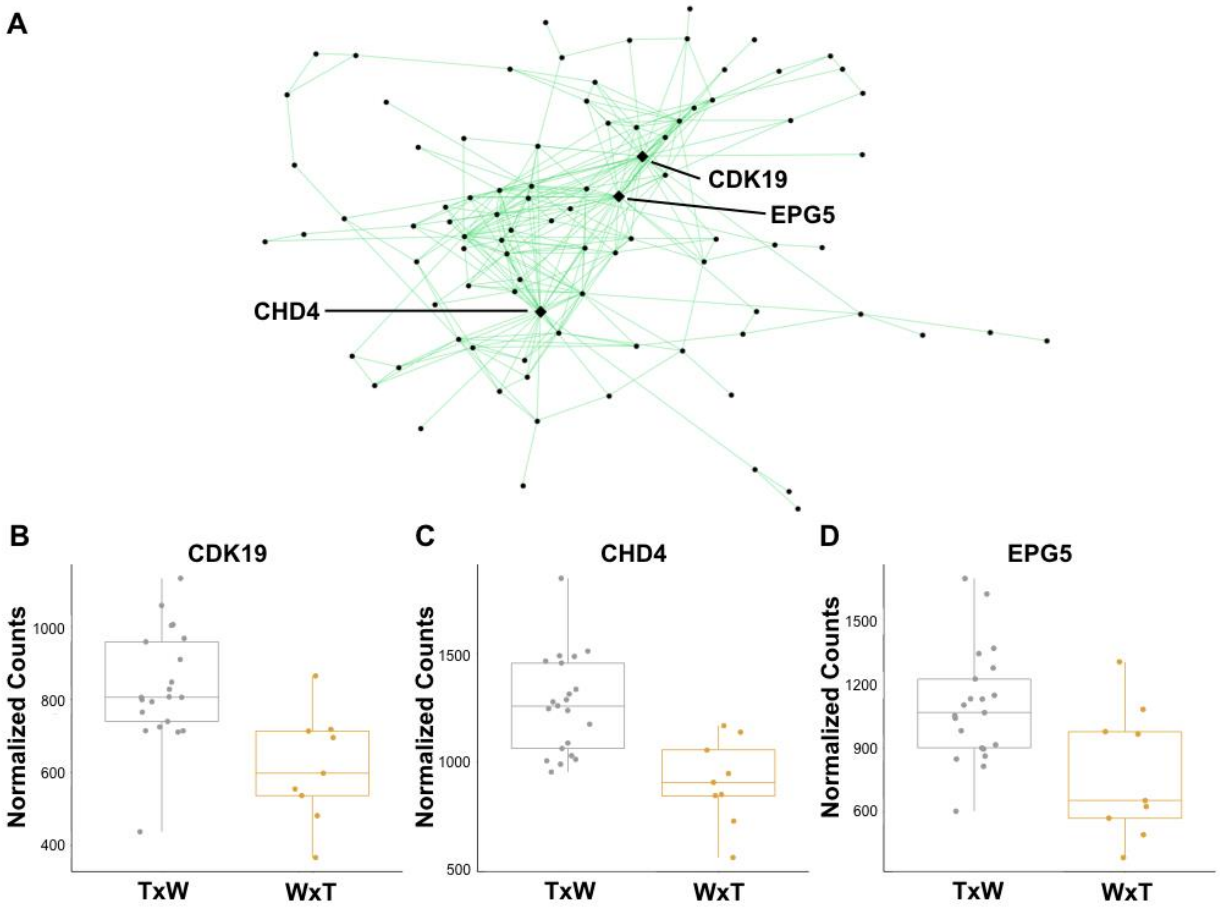


Figure 3.4. (A) Network of light green module and normalized expression plots of hub genes (B) CDK19, (C) CHD4, and (D) EPG5. TxW represents samples from nests sired by a T male and a W female. WxT represent samples from nests sired by a W male and a T female. Each circle represents a gene and diamonds represent hub genes described in Table 3.2.

For each pair type module, the correlation was stronger for the overall effect of pair type than any individual nest, indicating that one nest did not drive the correlation. This trend was reflected in gene expression plots of hub genes and candidate genes described above (Figure S7). We did not observe modules correlated with pair type that were also correlated with nestling morph or sex, suggesting there is no morph or sex-specific response to a given pair type at the network level.

Discussion

By assessing genome-wide transcription in nestlings raised by different WTSP pair types we have identified distinct transcriptomic signatures that suggest WxT pairs induce a stress response in developing nestlings relative to TxW pairs. This is reflected both by differential expression of several genes involved in protein degradation as well as networks of co-expressed genes with stress response hubs. Additionally, we identified morph-specific gene expression driven by innate immunity genes and genes located in the chromosome 2 inversion. As adults, the genes within the inversion strongly influence the WTSP neural transcriptome (Balakrishnan et al. 2014, Zinzow-Kramer et al. 2015). Our results here suggest that parental genotypes and their associated behaviors, rather than nestling genotype, have the strongest influence on the nestling transcriptome.

Gene expression differences resulting from pair type

We find 881 genes DE between nestlings raised under the two pair types. Many of these genes function in the proteasome or ubiquitin-mediated proteolysis. Cells naturally use the proteasome for degradation of proteins targeted by the ubiquitination process, but genes involved in proteasome formation (e.g. PSMD6, PSMD11) and ubiquitination (e.g. UBE2B) are up-regulated in cells experiencing mild oxidative stress (Aiken et al. 2011, Shang & Taylor 2011, Livneh et al. 2016) or organisms experiencing abiotic stress (Dhanasiri et al. 2013, Tomalty et al. 2015). Thus, increased expression of these genes in nestlings from WxT nests suggests they are responding to oxidative stress. As a result, there may be a cost to having a W morph father and T morph mother at the nestling stage.

To complement our differential expression approach, we also constructed co-expression networks with *WGCNA*. *WGCNA* identifies modules of co-regulated genes blind to the experimental design. These modules are then correlated with external traits, offering a systems-level view into how conditions impact transcriptional networks. Within these networks, we can then perform GO analyses as described above and identify network hubs, which are the most highly connected genes within that network. Using this approach, we identified 26 modules of co-regulated genes in this dataset (Figure 3.1), seven of which were significantly correlated with parental pair type. The blue module contains genes that are elevated in nestlings in WxT nests. The blue module hub gene was NDUFB3 (Module Membership [MM]=0.938, DD=42) (Figure 3.2), which encodes a subunit of the mitochondrial membrane respiratory chain. Interestingly, many of the same proteolysis-related genes highlighted in the differential expression results are also present in this module, resulting in the enrichment of several metabolism and stress-related GO categories (Table S4).

Two modules, light green and beige, contained stress responsive hub genes. The light green module contained genes that are suppressed in nestlings in WxT nests, with three hub genes: CDK19, CHD4, and EPG5 (Figure 3.4). The absence of EPG5 expression (via knockout) and reduction in CHD4 expression (via knockdown) has been associated with increased DNA damage (Zhao et al. 2013, Larsen et al. 2010). Similarly, down-regulation of CDK19 following knockdown is associated with an increased stress response (Audetat et al. 2017). Suppression of these genes in these nestlings could be indicative of increased cellular damage. The beige module contains genes whose expression is elevated in nestlings from WxT nests and contains one hub gene, DEPTOR, which is an inhibitor of mTOR signaling (Figure 3.3). The exact role of DEPTOR remains unclear, but up-regulation likely inhibits the mTORC1 pathway to reduce

endoplasmic reticulum stress, promote cell survival, and avoid apoptosis (Peterson et al. 2009, Desantis et al. 2015, Catena et al. 2016).

Increased expression of genes in the beige module in these nestlings and the high connectivity of DEPTOR to other co-expressed genes provide further support for a transcriptional stress response within WxT nests. The beige module also contains two well-studied stress responsive genes, superoxide dismutase 2 (SOD2) and the glucocorticoid receptor (NR3C1). SOD2 mitigates the effects of exposure to reactive oxygen species by scavenging free radicals (Zelko et al. 2002). NR3C1 binds glucocorticoids and has primarily been studied in the context of ELS and methylation of an upstream promoter. NRC3C1 methylation is often associated with down-regulation of NR3C1 (e.g. McGowan et al. 2009) and impairment of the HPA axis, but up-regulation following methylation has also been observed as part of the stress response (Turner et al. 2006, Bockmühl et al. 2015). Increased expression observed here directly implicates the HPA axis and suggests these nestlings may be activating SOD2 and NR3C1 to cope with elevated levels of reactive oxygen species and corticosterone, respectively (Wang et al. 2018, Finsterwald & Alberini 2014). However, further work is needed to investigate stress physiology, corticosterone levels, and uncover the epigenetic state of NR3C1 in these nestlings and how this may relate to ELS (Banerjee et al. 2011, McCoy et al. 2016, Rubenstein et al. 2016, Quirici et al. 2016, Greggor et al. 2017).

How does parental genotype influence offspring gene expression?

In a non-experimental study, we have limited power to make inferences about the mechanism by which parental genotype impacted offspring gene expression. Given the well-studied reproductive biology of WTSPs, however, two mechanisms seem especially likely:

hormone-mediated maternal effects and/or differences in parental provisioning. In weighing the evidence for these two non-mutually exclusive possibilities, we conclude that the difference in parental provisioning is the most plausible explanation for the observed gene expression differences. As described above, WTSP morphs differ in hormone levels. Only oestradiol, however, has been shown to differ between adult female morphs during the breeding season and is higher in W morph females during the pre-laying and laying stages (Horton et al. 2014). No baseline differences in any other hormone measured to date (corticosterone, testosterone, DHEA, DHT) have been described during the breeding season (Spinney et al. 2006, Swett & Breuner 2009, Horton & Holberton 2010, Horton et al. 2014). Taken together this suggests that hormone deposition into eggs may not differ dramatically between the morphs. By contrast, there is strong evidence of differences in provisioning among morph types (Knapton & Falls 1983, Kopachna & Falls 1993, Horton & Holberton 2010, Horton et al. 2014). Reduced provisioning by W morph males appears to be stable across populations resulting in female-biased parental care in WxT nests (Knapton & Falls 1983, Horton et al. 2014). Therefore, parental care variation is a likely source of behaviorally mediated maternal or paternal effects (see Crean & Bonduriansky 2014) that would explain the strong signature of stress exposure in the expression data.

Previous work revealed no difference in clutch size between pair types (Knapton et al. 1984, Formica et al. 2004) and no effect of pair type on nestling mass (Knapton et al. 1984, Tuttle et al. 2017). Also, nestlings did not differ in mass at time of sampling between the TxW and WxT nests used in this study (Smith et al. in review). Increased provisioning by females to compensate for reduced care by males could explain this observation, and this has been observed previously in a separate WTSP population (Knapton & Falls 1983). In this scenario reduced brooding and increased maternal separation could also negatively impact nestling physiology and

act as a source of ELS (reviewed in Ledón-Rettig et al. 2013). Somewhat surprisingly, given the gene expression findings described here, a recent study in our study population did not detect differences in reactive oxygen metabolites in plasma of offspring of the two different pair types (Grunst et al. 2019). ROM, however, only provides a limited overview of the stress response and the RNA-seq response we observed could even mitigate long-term consequences of ELS. The results here further highlight the utility of blood RNA-seq as a highly sensitive measure of environmental exposures (Louder et al. 2018).

The field portion of our study was carried out as part of a long-term study and was limited by the fact that we did not perform a cross-fostering experiment. We aimed to mitigate potential environmental confounds by collecting samples from nestlings during a time period restricted to nine days. Certainly the environment may influence gene expression in our samples, but consistent changes among the samples in the two breeding pair combinations suggest the role of parental genotype by sex is a significant driver of nestling gene expression, rather than temporal or spatial environmental variation.

Morph-specific gene expression

We were also interested in morph-specific gene expression and how nestling morph may respond to differences in parental pair type. WTSPs have been studied extensively as adults, but very rarely in other life stages. W morph males and T morph females exhibit earlier reproductive and actuarial senescence, potentially resulting from the high energy expenditure lifestyle of W morph males and biased parental care given by T morph females (Grunst et al. 2018a, Grunst et al. 2018b). There also appears to be seasonal variation in fitness between the morphs as adults. Following cold, wet winters, W morph males exhibit lower recruitment on the breeding grounds,

leading to an overproduction of W morph male nestlings, potentially to stabilize morph frequencies in the population (Tuttle et al. 2017). Thus, morph specific differences may arise in early life. We found 92 genes DE between morphs, including many innate immune-related genes and genes located within the inversion (65/92 genes, Table S2). *WGCNA* revealed five modules correlated with morph (Figure 3.1). These included two innate immunity-related modules with increased expression in W morphs (Dark Red & Sky Blue) and two modules enriched with genes located in the inversion (Ivory = 40/72, Light Cyan = 70/183) (Figures S5, S6). The sky blue module contains nine hub genes and the dark red module contains one hub gene, both of which include well-studied anti-viral genes (e.g. sky blue: OASL, RSAD2; dark red: TRAF5). These genes also form a co-expression module in avian blood following West Nile virus infection (Kernbach et al., in review). Adult WTSP morphs differ in their ability to clear infection (Boyd et al. 2018), so the immune activation here may be indicative of an increased parasite load in W morph nestlings, although further investigation is required. The light cyan module contains genes elevated in W morph nestlings and contains eight hub genes, each located in the inversion (Table 3.1). Three of these, EPM2A, BPNT1, and TAF5L, were also identified as hub genes in brain tissues of adult W morph males (Zinzow-Kramer et al. 2015). These nestlings thus exhibit expression differences in inversion genes prior to any phenotypic or behavioral differences, revealing the importance of the inversion in maintaining morph phenotypes throughout life. Additionally, the conservation of network hub genes in a different tissue and life stage highlights avenues for further investigation into WTSP gene regulation.

Despite broad gene expression differences between the morphs, within pair types morph-specific expression was limited. In part due to small sample size, nestlings in TxW nests only have two genes DE between morphs. There is a larger effect of morph within WxT nests, where

the number of DE genes increased to 40. These genes encompassed a wide range of gene functions without any obvious stress-related candidate genes. Of these 40 genes, 34 are uniquely DE within WxT nests and do not overlap with the overall list of 92 genes DE between morphs using all samples. Interestingly, glucocorticoid-induced transcript 1 (GLCCI1) is elevated in W morph nestlings in WxT nests. The function of GLCCI1 remains unclear (Kim et al. 2016), but expression differences between morphs observed here implicates the role of glucocorticoids in response to pair type. This suggests that nestling morphs may respond differently to the parental pair type though larger sample sizes will be needed to explore this further.

Conclusions

Using the WTSP, a system with alternative parental care strategies, we show that nestlings in WxT nests (female-biased parental care) have increased expression of stress-related genes, and parental genotypes may act as a source of ELS in the species. Nestling morph also influences transcription, but parental pair type appears to have the greatest effect on their transcriptome. Combined, this supports the parental effects hypothesis (Wade 1998, Schrader et al. 2018), where offspring phenotypes are primarily a result of the nest environment and care received, rather than from offspring genotypes (i.e. T vs. W). Nearly 54% of observed pairs have been WxT (Tuttle et al. 2016). Thus, roughly half of the nestlings in every population will experience female-biased parental care. Our results suggest that these differences in parental pair type have at least short-term consequences on offspring physiology. While we have identified impacts at the level of transcription, an integrative approach assessing nestling WTSP physiology and performing cross-fostering experiments will further elucidate the consequences of variation in parental pair type. Importantly, it remains unclear whether female-biased parental

care or differences in maternal effects translate into long-term fitness consequences for offspring. There appears to be a cost associated with parental genotype, as the less cooperative reproductive strategy (WxT pairs) accelerates actuarial senescence in tan morph females and reproductive senescence in white morph males (Grunst et al. 2018a, Grunst et al. 2018b). We show here that this cost is also translated into nestlings within WxT nests via increased stress-related gene expression. This work sets the stage to further explore morph-specific fitness consequences in nestlings experiencing alternative parental care strategies.

Acknowledgements

We acknowledge Lindsay Forrette, Andrea Grunst, and Melissa Grunst for assistance in the field, Sarah Ford for assistance with molecular work, Rachel Wright for WGCNA code, and Cranberry Lake Biological Station. Funding provided by East Carolina University, Indiana State University, The National Science Foundation (grant no. DUE-0934648) and the National Institutes of Health (grant no. 1R01Gm084229 to E.M.T and R.A.G.) and a Sigma Xi Grants in Aid of Research award to DJN. Birds were banded with color bands and a Fish and Wildlife band (Master Banding Permit 22296 to EMT and permit 24105 to RAG). Dr. Alvaro Hernandez and Chris Wright provided guidance and oversight on sequencing carried out at the University of Illinois. All methods were conducted in accordance with legal and ethical standards and were approved by Indiana State University's Institutional Animal Care and Use Committee (protocols 562158-1:ET/RG and 562192-1:ET/RG).

Data Accessibility

The 32 RNAseq libraries used in this study will be submitted to the NCBI Sequence Read Archive (SRA). All files needed to produce these results, including code and counts files, will be uploaded to this project's GitHub page: https://github.com/danielnewhouse/wtsp

Author Contributions

DJN designed and performed research, analyzed the data, and wrote the paper. MBS performed research, contributed samples, and reviewed drafts of the paper. EMT designed and performed research and contributed samples. RAG designed and performed research, contributed samples, and reviewed drafts of the paper. CNB designed and performed research, contributed reagents, and reviewed drafts of the paper.

References

Aiken, C. T., Kaake, R. M., Wang, X., & Huang, L. (2011). Oxidative Stress-Mediated Regulation of Proteasome Complexes. *Molecular & Cellular Proteomics*, *10*(5). Retrieved from http://www.mcponline.org/content/10/5/R110.006924.abstract

Alyamani, R. A. S., & Murgatroyd, C. (2018). Epigenetic Programming by Early-Life Stress. *Progress in Molecular Biology and Translational Science* (Vol. 157, pp. 133–150). doi: 10.1016/bs.pmbts.2018.01.004

Anastasiadi, D., Esteve-Codina, A., & Piferrer, F. (2018). Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics & Chromatin*, *11*(1). doi: 10.1186/s13072-018-0205-1

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. doi: 10.1093/bioinformatics/btu638

Audetat, K. A., Galbraith, M. D., Odell, A. T., Lee, T., Pandey, A., Espinosa, J. M., … Taatjes, D. J. (2017). A kinase-independent role for CDK19 in p53 response. *Molecular and Cellular Biology*. *37*(13), e00626-16 Retrieved from http://mcb.asm.org/content/early/2017/04/12/MCB.00626-16.abstract

Balakrishnan, C. N., Mukai, M., Gonser, R. A., Wingfield, J. C., London, S. E., Tuttle, E. M., & Clayton, D. F. (2014). Brain transcriptome sequencing and assembly of three songbird model systems for the study of social behavior. *PeerJ*, *2*, e396–e396. doi: 10.7717/peerj.396

Banerjee, S. B., Arterbery, A. S., Fergus, D. J., & Adkins-Regan, E. (2011). Deprivation of maternal care has long-lasting consequences for the hypothalamic-pituitary-adrenal axis

of zebra finches. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 759–766. doi: 10.1098/rspb.2011.1265

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc B Method*, *57*(1), 289-300.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, *447*(7143), 407–412. doi: 10.1038/nature05915

Bockmühl, Y., Patchev, A. V., Madejska, A., Hoffmann, A., Sousa, J. C., Sousa, N., … Spengler, D. (2015). Methylation at the CpG island shore region upregulates Nr3c1 promoter activity after early-life stress. *Epigenetics*, *10*(3), 247–257. doi: 10.1080/15592294.2015.1017199

Bolton, J. L., Molet, J., Ivy, A., & Baram, T. Z. (2017). New insights into early-life stress and behavioral outcomes. *Current Opinion in Behavioral Sciences*, *14*, 133–139. doi: 10.1016/j.cobeha.2016.12.012

Boyd, R. J., Kelly, T. R., MacDougall-Shackleton, S. A., & MacDougall-Shackleton, E. A. (2018). Alternative reproductive strategies in white-throated sparrows are associated with differences in parasite load following experimental infection. *Biology letters*, *14*(7), 20180194.

Catena, V., Bruno, T., De Nicola, F., Goeman, F., Pallocca, M., Iezzi, S., … Fanciulli, M. (2016). Deptor transcriptionally regulates endoplasmic reticulum homeostasis in multiple myeloma cells. *Oncotarget*, *7*(43), 70546–70558. doi: 10.18632/oncotarget.12060

Cingolani, P., Patel, V., Coon, M., Nguyen, T., Land, S., Ruden, D., & Lu, X. (2012). Using *Drosophila melanogaster* as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Frontiers in Genetics*, *3*, 35–35.

Cottrell, E. C., & Secki, J. R. (2009). Prenatal stress, glucocorticoids and the programming of adult disease. *Frontiers in Behavioral Neuroscience*, *3*,*19*. doi: 10.3389/neuro.08.019.2009

Crean, A. J., & Bonduriansky, R. (2014). What is a paternal effect? *Trends in Ecology & Evolution*, *29*(10), 554–559. doi: 10.1016/j.tree.2014.07.009

Crespi, E. J., Williams, T. D., Jessop, T. S., & Delehanty, B. (2012). Life history and the ecology of stress: how do glucocorticoid hormones influence life-history variation in animals? *Functional Ecology*, *27*(1), 93–106. doi: 10.1111/1365-2435.12009

Desantis, A., Bruno, T., Catena, V., De Nicola, F., Goeman, F., Iezzi, S., … Fanciulli, M. (2015). Che-1-induced inhibition of mTOR pathway enables stress-induced autophagy. *The EMBO Journal*, *34*(9), 1214–1230.

Dhanasiri, A. K. S., Fernandes, J. M. O., & Kiron, V. (2013). Liver Transcriptome Changes in Zebrafish during Acclimation to Transport-Associated Stress. *PLOS ONE*, *8*(6), e65028–e65028.

Diaz-Real, J., Kim, S.-Y., & Velando, A. (2017). Plumage colour and the expression of stress-related genes in gull chicks. *Journal of Avian Biology*, *48*(9), 1216–1225. doi: 10.1111/jav.01460

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21.

Eden, E., Lipson, D., Yogev, S., & Yakhini, Z. (2007). Discovering Motifs in Ranked Lists of DNA Sequences. *PLOS Computational Biology*, *3*(3), e39–e39.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, *10*(1), 48–48. doi: 10.1186/1471-2105-10-48

Finsterwald, C., & Alberini, C. M. (2014). Stress and glucocorticoid receptor-dependent mechanisms in long-term memory: From adaptive responses to psychopathologies. *Neurobiology of Learning and Memory*, *112*, 17–29. doi: 10.1016/j.nlm.2013.09.017

Formica, V. A., Gonser, R. A., Ramsay, S., & Tuttle, E. M. (2004). Spatial dynamics of alternative reproductive strategies: the role of neighbors. *Ecology*, *85*(4), 1125-1136.

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of In Utero and Early-Life Conditions on Adult Health and Disease. *New England Journal of Medicine*, *359*(1), 61–73. doi: 10.1056/NEJMra0708473

Gott, A. J. (2018). Effects of early-life adversity on the adult phenotype in European starlings. Doctoral dissertation, Newcastle University.

Greggor, A. L., Spencer, K. A., Clayton, N. S., & Thornton, A. (2017). Wild jackdaws' reproductive success and their offspring's stress hormones are connected to provisioning rate and brood size, not to parental neophobia. *General and Comparative Endocrinology*, *243*, 70–77. doi: 10.1016/j.ygcen.2016.11.006

Griebel, I. A., Fairhurst, G. D., Marchant, T. A., & Clark, R. G. (2019). Effects of parental and nest-site characteristics on nestling quality in the Tree Swallow (*Tachycineta bicolor*). *Canadian Journal of Zoology*, *97*(1), 63–71. doi: 10.1139/cjz-2018-0109

Griffiths, R., Double, M. C., Orr, K., & Dawson, R. J. G. (1998). A DNA test to sex most birds. *Molecular Ecology*, *7*(8), 1071–1075. doi: 10.1046/j.1365-294x.1998.00389.x

Grunst, A. S., Grunst, M. L., Formica, V. A., Korody, M. L., Betuel, A. M., Barcelo-Serra, M., … Tuttle, E. M. (2018a). Morph-Specific Patterns of Reproductive Senescence: Connections to Discrete Reproductive Strategies. *The American Naturalist*, *191*(6), 744–755. doi: 10.1086/697377

Grunst, A. S., Grunst, M. L., Gonser, R. A., & Tuttle, E. M. (2019). Developmental stress and telomere dynamics in a genetically polymorphic species. *Journal of Evolutionary Biology, 32(2), 134-143.* doi: 10.1111/jeb.13400

Grunst, M. L., Grunst, A. S., Formica, V. A., Korody, M. L., Betuel, A. M., Barcelo-Serra, M., … Tuttle, E. M. (2018b). Actuarial senescence in a dimorphic bird: different rates of ageing in morphs with discrete reproductive strategies. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1892), 20182053. doi: 10.1098/rspb.2018.2053

Heim, C., Newport, D. J., Mletzko, T., Miller, A. H., & Nemeroff, C. B. (2008). The link between childhood trauma and depression: Insights from HPA axis studies in humans. *Psychoneuroendocrinology*, *33*(6), 693–710. doi: 10.1016/j.psyneuen.2008.03.008

Horton, B. M., & Holberton, R. L. (2010). Morph-Specific Variation in Baseline Corticosterone and the Adrenocortical Response in Breeding White-Throated Sparrows (*Zonotrichia albicollis*). *The Auk*, *127*(3), 540–548. doi: 10.1525/auk.2010.09096

Horton, B. M., Moore, I. T., & Maney, D. L. (2014). New insights into the hormonal and behavioural correlates of polymorphism in white-throated sparrows, *Zonotrichia albicollis*. *Animal Behaviour*, *93*, 207–219. doi: 10.1016/j.anbehav.2014.04.015

Horvath, S. (2011). *Weighted network analysis: application in genomics and systems biology*. New York, NY: Springer.

Hu, Z., Chang, Y.-C., Wang, Y., Huang, C.-L., Liu, Y., Tian, F., … DeLisi, C. (2013). VisANT 4.0: Integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Research*, *41*(W1), W225–W231.

Jacquin, L., Blottière, L., Haussy, C., Perret, S., & Gasparini, J. (2012). Prenatal and postnatal parental effects on immunity and growth in 'lactating' pigeons: Prenatal and postnatal parental effects. *Functional Ecology*, *26*(4), 866–875. doi: 10.1111/j.1365-2435.2012.01988.x

Kernbach, M.E., Newhouse, D.J., Miller, J.M., Hall, R.J., Gibbons, J., Oberstaller, J., Selechnik, D., Jiang, R., Unnasch, T.R., Balakrishnan, C.N., & L.B. Martin. Light pollution increases West Nile virus competence of a ubiquitous passerine reservoir species. *In review*.

Kim, S.-H., Kim, H.-J., & Kim, C.-W. (2016). GLCCI1 is a novel component associated with the PI3K signaling pathway in podocyte foot processes. *Experimental & Molecular Medicine*, *48*(5), e233–e233. doi: 10.1038/emm.2016.28

Kinnally, E. L., Feinberg, C., Kim, D., Ferguson, K., Leibel, R., Coplan, J. D., & John Mann, J. (2011). DNA methylation as a risk factor in the effects of early life stress. *Brain, Behavior, and Immunity*, *25*(8), 1548–1553. doi: 10.1016/j.bbi.2011.05.001

Knapton, R. W., Cartar, R. V., & Falls, J. B. (1984). A comparison of breeding ecology and reproductive success between morphs of the white-throated sparrow. *The Wilson Bulletin*, 60-71.

Knapton, R. W., & Falls, J. B. (1983). Differences in parental contribution among pair types in the polymorphic White-throated Sparrow. *Canadian Journal of Zoology*, *61*(6), 1288–1292. doi: 10.1139/z83-173

Kopachena, J. G., & Falls, J. B. (1993). Aggressive Performance as a Behavioral Correlate of Plumage Polymorphism in the White-Throated Sparrow (*Zonotrichia albicollis*). *Behaviour*, *124*(3/4), 249–266.

Krause, E. T., Honarmand, M., Wetzel, J., & Naguib, M. (2009). Early Fasting Is Long Lasting: Differences in Early Nutritional Conditions Reappear under Stressful Conditions in Adult Female Zebra Finches. *PLoS ONE*, *4*(3), e5015. doi: 10.1371/journal.pone.0005015

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

Larsen, D. H., Poinsignon, C., Gudjonsson, T., Dinant, C., Payne, M. R., Hari, F. J., … Lukas, J. (2010). The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. *The Journal of Cell Biology*, *190*(5), 731–740.

Ledón-Rettig, C. C., Richards, C. L., & Martin, L. B. (2013). Epigenetics for behavioral ecologists. *Behavioral Ecology*, *24*(2), 311–324. doi: 10.1093/beheco/ars145

Lewis, C. R., & Olive, M. F. (2014). Early-life stress interactions with the epigenome: potential mechanisms driving vulnerability toward psychiatric illness. *Behavioural Pharmacology*, 1. doi: 10.1097/FBP.0000000000000057

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

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Subgroup, 1000 Genome Project Data Processing. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Lima, S. L. (2009). Predators and the breeding bird: behavioral and reproductive flexibility under the risk of predation. *Biological Reviews*, *84*(3), 485–513. doi: 10.1111/j.1469-185X.2009.00085.x

Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., … Meaney, M. J. (1997). Maternal Care, Hippocampal Glucocorticoid Receptors, and Hypothalamic-Pituitary-Adrenal Responses to Stress. *Science*, *277*(5332), 1659–1662.

Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N., & Ciechanover, A. (2016). The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell Research*, *26*, 869–869.

Lloyd, J. D., & Martin, T. E. (2004). Nest-site preference and maternal effects on offspring growth. *Behavioral Ecology*, *15*(5), 816–823. doi: 10.1093/beheco/arh085

Longmire, J. L., Gee, G. F., Hardekopf, C. L., & Mark, G. A. (1992). Establishing paternity in whooping cranes (*Grus americana*) by DNA analysis. *The Auk*, *109*(3), 522-529.

Louder, M. I. M., Hauber, M. E., & Balakrishnan, C. N. (2018). Early social experience alters transcriptomic responses to species-specific song stimuli in female songbirds. *Behavioural Brain Research*, *347*, 69–76. doi: 10.1016/j.bbr.2018.02.034

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550–550. doi: 10.1186/s13059-014-0550-8

Lowdon, R. F., Jang, H. S., & Wang, T. (2016). Evolution of Epigenetic Regulation in Vertebrate Genomes. *Trends in Genetics*, *32*(5), 269–283. doi: 10.1016/j.tig.2016.03.001

Lowther, J. K. (1961). Polymorphism in the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Canadian Journal of Zoology*, *39*(3), 281–292. doi: 10.1139/z61-031

Lowther, J. K. (1962). *Colour and Behavioural Polymorphism in the White-throated Sparrow, Zonotrichia Albicollis (Gmelin)*. Doctoral dissertation, University of Toronto.

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience*, *10*(6), 434–445. doi: 10.1038/nrn2639

Marco, E. M., Echeverry-Alzate, V., López-Moreno, J. A., Giné, E., Peñasco, S., & Viveros, M. P. (2014). Consequences of early life stress on the expression of endocannabinoid-related genes in the rat brain. *Behavioural Pharmacology*, *25*(5–6), 547–556. doi: 10.1097/FBP.0000000000000068

Marini, F., & Binder, H. (2016). Development of Applications for Interactive and Reproducible Research: a Case Study. *Genomics and Computational Biology*, *3*(1), 39. doi: 10.18547/gcb.2017.vol3.iss1.e39

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

McCoy, C. R., Rana, S., Stringfellow, S. A., Day, J. J., Wyss, J. M., Clinton, S. M., & Kerman, I. A. (2016). Neonatal maternal separation stress elicits lasting DNA methylation changes in the hippocampus of stress-reactive Wistar Kyoto rats. *European Journal of Neuroscience*, *44*(10), 2829–2845. doi: 10.1111/ejn.13404

McEwen, B. S. (2007). Physiology and Neurobiology of Stress and Adaptation: Central Role of the Brain. *Physiological Reviews*, *87*(3), 873–904. doi: 10.1152/physrev.00041.2006

McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonté, B., Szyf, M., … Meaney, M. J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience*, *12*, 342–342.

Meaney, M. J. (2001). Maternal Care, Gene Expression, and the Transmission of Individual Differences in Stress Reactivity Across Generations. *Annual Review of Neuroscience*, *24*(1), 1161–1192. doi: 10.1146/annurev.neuro.24.1.1161

Michopoulos, V., Maney, D. L., Morehouse, C. B., & Thomas, J. W. (2007). A genotyping assay to determine plumage morph in the white-throated sparrow (*Zonotrichia albicollis*). *The Auk*, *124*(4), 1330. doi:10.1642/0004-8038(2007)124[1330:AGATDP]2.0.CO;2

Millstein, R. A., & Holmes, A. (2007). Effects of repeated maternal separation on anxiety- and depression-related phenotypes in different mouse strains. *Neuroscience & Biobehavioral Reviews*, *31*(1), 3–17. doi: 10.1016/j.neubiorev.2006.05.003

Mock, D. W., & Parker, G. A. (1997). *The evolution of sibling rivalry*. Oxford ; New York: Oxford University Press.

Moghadam, H. K., Johnsen, H., Robinson, N., Andersen, Ø., H. Jørgensen, E., Johnsen, H. K., … Tveiten, H. (2017). Impacts of Early Life Stress on the Methylome and Transcriptome of Atlantic Salmon. *Scientific Reports*, *7*(1). doi: 10.1038/s41598-017-05222-2

Monaghan, P. (2014). Organismal stress, telomeres and life histories. *Journal of Experimental Biology*, *217*(1), 57–66. doi: 10.1242/jeb.090043

Mousseau, T. A., & Fox, C. W. (Eds.). (1998). *Maternal effects as adaptations*. New York: Oxford University Press.

Murgatroyd, C., Patchev, A. V., Wu, Y., Micale, V., Bockmühl, Y., Fischer, D., … Spengler, D. (2009). Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nature Neuroscience*, *12*(12), 1559–1566. doi: 10.1038/nn.2436

Noguera, J. C., Kim, S.-Y., & Velando, A. (2017). Family-transmitted stress in a wild bird. *Proceedings of the National Academy of Sciences*, *114*(26), 6794-6799. Retrieved from http://www.pnas.org/content/early/2017/06/06/1706164114.abstract

Nord, A., & Nilsson, J.-Å. (2011). Incubation Temperature Affects Growth and Energy Metabolism in Blue Tit Nestlings. *The American Naturalist*, *178*(5), 639–651. doi: 10.1086/662172

Pértille, F., Brantsæter, M., Nordgreen, J., Coutinho, L. L., Janczak, A. M., Jensen, P., & Guerrero-Bosagna, C. (2017). DNA methylation profiles in red blood cells of adult hens correlate with their rearing conditions. *The Journal of Experimental Biology*, *220*(19), 3579–3587. doi: 10.1242/jeb.157891

Peterson, T. R., Laplante, M., Thoreen, C. C., Sancak, Y., Kang, S. A., Kuehl, W. M., … Sabatini, D. M. (2009). DEPTOR Is an mTOR Inhibitor Frequently Overexpressed in Multiple Myeloma Cells and Required for Their Survival. *Cell*, *137*(5), 873–886. doi: 10.1016/j.cell.2009.03.046

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842.

Quirici, V., Guerrero, C. J., Krause, J. S., Wingfield, J. C., & Vásquez, R. A. (2016). The relationship of telomere length to baseline corticosterone levels in nestlings of an altricial passerine bird in natural populations. *Frontiers in Zoology*, *13*(1). doi: 10.1186/s12983-016-0133-5

R Core Team (2018). R: A language and environment for statistical computing.

Reshetnikov, V. V., Studenikina, A. A., Ryabushkina, J. A., Merkulova, T. I., & Bondar, N. P. (2018). The impact of early-life stress on the expression of HPA-associated genes in the adult murine brain. *Behaviour*, *155*(2–3), 181–203. doi: 10.1163/1568539X-00003482

Romanov, M. N., Tuttle, E. M., Houck, M. L., Modi, W. S., Chemnick, L. G., Korody, M. L., … Ryder, O. A. (2009). The value of avian genomics to the conservation of wildlife. *BMC Genomics*, *10*(Suppl 2), S10. doi: 10.1186/1471-2164-10-S2-S10

Royle, N. J., Smiseth, P. T., & Kölliker, M. (Eds.). (2012). *The evolution of parental care*. Oxford University Press.

Rubenstein, D. R., Skolnik, H., Berrio, A., Champagne, F. A., Phelps, S., & Solomon, J. (2016). Sex-specific fitness effects of unpredictable early life conditions are associated with DNA methylation in the avian glucocorticoid receptor. *Molecular Ecology*, *25*(8), 1714–1728. doi: 10.1111/mec.13483

Schrader, M., Jarrett, B. J. M., & Kilner, R. M. (2018). Parental care and sibling competition independently increase phenotypic variation among burying beetle siblings. *Evolution*, *72(11), 2546-2552*. doi: 10.1111/evo.13607

Schwander, T., Libbrecht, R., & Keller, L. (2014). Supergenes and Complex Phenotypes. *Current Biology*, *24*(7), R288–R294. doi: 10.1016/j.cub.2014.01.056

Shang, F., & Taylor, A. (2011). Ubiquitin–proteasome pathway and cellular responses to oxidative stress. *Free Radical Biology and Medicine*, *51*(1), 5–16. doi: 10.1016/j.freeradbiomed.2011.03.031

Sheldon, E. L., Schrey, A. W., Ragsdale, A. K., & Griffith, S. C. (2018). Brood size influences patterns of DNA methylation in wild Zebra Finches (*Taeniopygia guttata*). *The Auk, 135*(4), 1113–1122. doi: 10.1642/AUK-18-61.1

Silberman, D. M., Acosta, G. B., & Zorrilla Zubilete, M. A. (2016). Long-term effects of early life stress exposure: Role of epigenetic mechanisms. *Pharmacological Research*, *109*, 64–73. doi: 10.1016/j.phrs.2015.12.033

Smith, J.N., Newhouse, D.J., Balakrishnan, C.N., Tuttle, E.M.,& Gonser, R.A. Differential morph gene expression in chromosome three of the white-throated sparrow (*Zonotrichia albicollis*). *In review*.

Spencer, K. A. (2017). Developmental stress and social phenotypes: integrating neuroendocrine, behavioural and evolutionary perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1727). Retrieved from http://rstb.royalsocietypublishing.org/content/372/1727/20160242.abstract

Spencer, K. A., Evans, N. P., & Monaghan, P. (2009). Postnatal Stress in Birds: A Novel Model of Glucocorticoid Programming of the Hypothalamic-Pituitary-Adrenal Axis. *Endocrinology*, *150*(4), 1931–1934. doi: 10.1210/en.2008-1471

Spinney, L. H., Bentley, G. E., & Hau, M. (2006). Endocrine correlates of alternative phenotypes in the white-throated sparrow (*Zonotrichia albicollis*). *Hormones and Behavior*, *50*(5), 762–771. doi: 10.1016/j.yhbeh.2006.06.034

Swett, M. B., & Breuner, C. W. (2009). Plasma Testosterone Correlates with Morph Type across Breeding Substages in Male White-Throated Sparrows. *Physiological and Biochemical Zoology*, *82*(5), 572–579. doi: 10.1086/605392

Szyf, M. (2009). The early life environment and the epigenome. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1790*(9), 878–885. doi: 10.1016/j.bbagen.2009.01.009

Szyf, M., Weaver, I., & Meaney, M. (2007). Maternal care, the epigenome and phenotypic differences in behavior. *Reproductive Toxicology*, *24*(1), 9–19. doi: 10.1016/j.reprotox.2007.05.001

Taylor, S., & Campagna, L. (2016). Avian supergenes. *Science*, *351*(6272), 446–447. doi: 10.1126/science.aae0389

Thomas, J. W., Cáceres, M., Lowman, J. J., Morehouse, C. B., Short, M. E., Baldwin, E. L., … Martin, C. L. (2008). The Chromosomal Polymorphism Linked to Variation in Social Behavior in the White-Throated Sparrow (*Zonotrichia albicollis*) Is a Complex Rearrangement and Suppressor of Recombination. *Genetics*, *179*(3), 1455 LP – 1468.

Thorneycroft, H. B. (1966). Chromosomal Polymorphism in the White-Throated Sparrow, *Zonotrichia albicollis* (Gmelin). *Science*, *154*(3756), 1571-1572.

Thorneycroft, H. B. (1975). A cytogenetic study of the white-throated sparrow, Zonotrichia albicollis (Gmelin). *Evolution*, *29*(4), 611–621. doi:10.1111/j.1558-5646.1975.tb00855.x

Tomalty, K. M. H., Meek, M. H., Stephens, M. R., Rincón, G., Fangue, N. A., May, B. P., & Baerwald, M. R. (2015). Transcriptional Response to Acute Thermal Exposure in Juvenile Chinook Salmon Determined by RNAseq. *G3: Genes|Genomes|Genetics*, *5*(7), 1335 LP – 1349.

Trivers, R.L. (1972). Parental investment and sexual selection. In: *Sexual Selection and the Descent of Man* (B. Campbell, ed), pp. 136–179. Aldine, Chicago, IL.

Turner, J. D., Schote, A. B., Macedo, J. A., Pelascini, L. P. L., & Muller, C. P. (2006). Tissue specific glucocorticoid receptor expression, a role for alternative first exon usage? *Biochemical Pharmacology*, *72*(11), 1529–1537. doi: 10.1016/j.bcp.2006.07.005

Tuttle, E. (2003). Alternative reproductive strategies in the white-throated sparrow: behavioral and genetic evidence. *Behavioral Ecology*, *14*(3), 425–432.

Tuttle, E. M., Grunst, A. S., Grunst, M. L., Korody, M. L., Betuel, A. M., Barcelo-Serra, M., … Gonser, R. A. (2017). Climatically driven changes in population composition and offspring sex-morph ratio in a polymorphic species. *Ecosphere*, *8*(4), e01762–e01762. doi: 10.1002/ecs2.1762

Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., … Balakrishnan, C. N. (2016). Divergence and Functional Degradation of a Sex Chromosome-like Supergene. *Current Biology*, *26*(3), 344–350. doi: 10.1016/j.cub.2015.11.069

van Oers, K., Kohn, G. M., Hinde, C. A., & Naguib, M. (2015). Parental food provisioning is related to nestling stress response in wild great tit nestlings: implications for the development of personality. *Frontiers in Zoology*, *12*(Suppl 1), S10. doi: 10.1186/1742-9994-12-S1-S10

Wade, M. J. (1998). The evolutionary genetics of maternal effects Pp. 5–21 *in* Mousseau T. A & Fox C. W., eds. Maternal effects as adaptations. Oxford Univ. Press, New York.

Wang, Y., Branicky, R., Noë, A., & Hekimi, S. (2018). Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. *The Journal of Cell Biology*, *217*(6), 1915–1928. doi: 10.1083/jcb.201708007

Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., … Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, *7*, 847–847.

Weaver, I. C. G., Meaney, M. J., & Szyf, M. (2006). Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3480 LP – 3485.

Wolf, J. B., & Wade, M. J. (2009). What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1520), 1107–1115. doi: 10.1098/rstb.2008.0238

Zelko, I. N., Mariani, T. J., & Folz, R. J. (2002). Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. *Free Radical Biology and Medicine*, *33*(3), 337–349. doi: 10.1016/S0891-5849(02)00905-X

Zhang, B., & Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. *Statistical Applications in Genetics and Molecular Biology*, *4*. doi: 10.2202/1544-6115.1128

Zhao, H., Zhao, Y. G., Wang, X., Xu, L., Miao, L., Feng, D., … Zhang, H. (2013). Mice deficient in *Epg5*exhibit selective neuronal vulnerability to degeneration. *The Journal of Cell Biology*, *200*(6), 731 LP – 741.

Zinzow-Kramer, W. M., Horton, B. M., McKee, C. D., Michaud, J. M., Tharp, G. K., Thomas, J. W., … Maney, D. L. (2015). Genes located in a chromosomal inversion are correlated

with territorial song in white-throated sparrows. *Genes, Brain and Behavior*, *14*(8), 641–654. doi: 10.1111/gbb.12252

CHAPTER 4. HIGH VARIATION IN THE WHITE-THROATED SPARROW MICROBIOME

Abstract

The host-associated microbiome is essential for host health and plays a crucial role in the development of host physiology and immunity. Microbiome formation is extremely dynamic and is influenced by both environmental factors and host genetics. However, the extent of host genetic control is relatively underexplored in wild species. Here, we assess the role of host genetics in microbiome formation in the white-throated sparrow (*Zonotrichia albicollis*, WTSP), which has two genetic morphs that differ in behavior and physiology. We sequenced the fecal microbiome of nestling and adult WTSPs and found limited support for host genetic influence of microbial composition based on alpha diversity, beta diversity, and clustering analyses. We uncovered extreme variation among WTSP fecal microbiomes, with substantial fluctuation of highly abundant bacterial taxa. For comparison, we also sequenced fecal microbiomes of each remaining *Zonotrichia* species and an outgroup. Like the WTSP, the *Zonotrichia* microbiome is also highly variable. This suggests that some unmeasured environmental factor, rather than host genotype, drives microbiome formation and variation in the WTSP and its relatives.

Introduction

The gut microbiome is crucial for host health and fitness (Cho & Blaser 2012, Gould et al. 2018) and plays a prominent role in shaping host physiology (Pluznick 2014, Jones 2016, Contijoch et al. 2019). Colonization of the host gastrointestinal tract is dependent on both environmental and host-genetic factors (e.g. Goodrich et al. 2014, Rothschild et al. 2018). Environmental influences are particularly important during early life, where the gut microbiome is highly variable and shifts in response to environmental stochasticity and dietary changes (Lozupone et al. 2012, Burns et al. 2016, Dong & Gupta 2019). As organisms mature, their microbiome generally becomes stable (Voreades et al. 2014, Rodriguez et al. 2015), yet it is still largely determined by environmental factors. Host genetics also plays a substantial role by mediating host physiology and behaviors that can influence the abundance of certain microbial taxa (Goodrich et al. 2014). For example, a large survey of mouse genetic lines revealed an effect of certain host quantitative-trait loci on the presence or absence of microbial taxa (Benson et al. 2010). Additionally, there is increasing evidence that a host species' evolutionary history influences its microbiome formation. As hosts diverge, so too do their microbial communities, resulting in species-specific microbiomes (Brooks et al. 2016, Moeller et al. 2016). In this scenario, hosts and their associated microbes may behave as a single entity upon which natural selection can act (i.e. holobiont), which results in parallel changes along the host and microbe phylogenies (Bordenstein & Theis 2015). While much of the groundbreaking work describing the influence of host genetics on microbiome formation has focused in laboratory models (Kostic et al. 2013, Goodrich et al. 2016), there has been a recent surge of investigations into the genetic influence of microbiome formation in wild species (reviewed in Hird 2017, Suzuki 2017, Trevelline et al. 2019). From these studies, the genetic control of microbiome formation appears

to be extremely variable, with moderate (Smith et al. 2015, Sullam et al. 2015, Boodawatta et al. 2018) to little evidence (Hird et al. 2014, Michel et al. 2018, Sun et al. 2018, Grieneisen et al. 2019) of host-genetic control.

These previous studies testing for genetic control of microbiome formation did not investigate species with divergent morphs involving phenotypic differences in multiple systems. To address this, we selected a wild species with permanent genetic morphs having demonstrated pleiotrophic effects on physiology and behavior, the white-throated sparrow (*Zonotrichia albicollis*, WTSP). WTSPs exist in two unique genetic morphs, identifiable by tan and white head stripes, respectively, that result from the presence or absence of a large chromosomal rearrangement on chromosome 2 (Thorneycroft 1966, Thorneycroft 1975). White morphs are nearly always heterozygous for this rearrangement and this rearrangement is not present in tan morphs (Tuttle et al. 2016). The rearrangement causes distinct changes to morph physiology and behavior, notably increasing aggression and circulating steroids in white morph adults (Tuttle et al. 2003, Horton & Holberton 2010, Horton et al. 2014). Additionally, white morph males provision offspring very little and are promiscuous, while tan morph males are highly paternal and defend within nest paternity (Horton et al. 2014). Morphs nearly always mate disassortatively, which produces two stable alternative parental care strategies: biparental (tan male x white female) and female-biased (white male x tan female), which imposes physiological costs in nestlings born into female-biased parental care nests (Newhouse et al. 2018). Morphs also differ in immune function as both adults (Boyd et al. 2018) and nestlings (Newhouse et al. 2018). Thus, WTSPs are a useful model to study how host genotype and corresponding differences in physiology, stress, and immunity influence the microbiome.

Here, we sequenced the fecal microbiome of wild WTSP adults and nestlings of both morphs and nestling rearing environments. In doing so, we tested for the role of host genotype (i.e. morph), as well as age and early life environment in microbiome formation of WTSPs. To place our WTSP results into a broader context, we also sequenced the microbiome of each *Zonotrichia* species and a closely related outgroup to explore trends in microbiome composition across the host phylogeny.

Methods

Sample collection

All animal handling was performed in accordance with the appropriate Institutional Animal Care and Use Committee (IACUC) guidelines. The majority of WTSP samples (n=19) were collected during the 2016 breeding season from a population at Cranberry Lake Biological Station, New York, USA (Indiana State University IACUC#: 562158). We used six adult (five tan, one white) and 13 six-day old nestling (six white, seven tan, six biparental, seven female-biased) samples for this study.

For comparison with our New York WTSP samples, we also sampled each remaining *Zonotrichia* species and an outgroup. Four nestling golden-crowned sparrows (*Zonotrichia atricapilla*, GCSP), three adult GCSP, and one adult white-crowned sparrow (*Zonotrichia leucophrys*, WCS) fecal samples were collected from a breeding population in the Hatcher Pass Management Area, AK, USA in June 2017 (University of Nebraska-Lincoln IACUC #1277). One rufous-collared sparrow (*Zonotrichia capensis*, RCSP) and two dark-eyed junco (*Junco hyemalis*, DEJ) fecal samples were collected from captive individuals at Indiana State University (IACUC #562192) and University of Montana (IACUC #010-16ZCDBS-020916) in February

2017 and September 2016, respectively. Two adult Harris's sparrow (*Zonotrichia querula*, HASP) and two adult WTSP fecal samples were collected in April 2018 during routine banding of migratory birds at the Konza Prairie Biological Station, Kansas, USA (Kansas State University IACUC #3920).

For all wild samples (WTSP, WCS, HASP, GCSP), we collected opportunistic fecal samples during routine measurements for nestlings and during banding of adults. For captive species (RCS & DEJ), we placed birds into a clean cage and placed a sterilized tray at the bottom of the cage to collect fecal droppings. Fecal samples for all species were collected aseptically with a sterile cotton swab, placed into a sterile 1.7 ml microcentrifuge tube, and placed on ice for 1-6 hours until frozen at -20°C until DNA extraction.

DNA extraction, PCR, library preparation, and sequencing

We extracted DNA from fecal samples using the Qiagen PowerSoil Kit (Carlsbad, CA, USA). We followed manufacturer protocol except for the following: we incubated samples in the PowerBead tubes for five minutes at 65°C prior to vortexing, incubated samples for five minutes following addition of the elution buffer to the filter membrane, and eluted purified DNA with 50 µl of elution buffer.

We amplified the V4-V5 region of the bacterial 16s subunit of the ribosomal RNA gene (16s rRNA) with PCR following the Earth Microbiome Project protocol (Caporaso et al. 2012). Each PCR reaction contained 16µl H_2O, 2.5µl 10x Buffer, 2.5µl $MgCl_2$ (25mM), 0.5µl dNTPs (40mM), 0.5µl barcoded 515f forward primer (10µM), 0.5µl barcoded 806r reverse primer (10µM), 0.125µl AmpliTaq Gold (5U/µl), and 2µl DNA template. Each reaction experienced the same thermal cycler conditions of 94°C for 3 minutes, followed by 30 cycles of 94°C for 45

seconds, 50°C for 30 seconds, and 72°C for 90 seconds, then 72°C for 10 minutes and held indefinitely at 4°C. Each sample was amplified in triplicate.

All three PCR reactions per sample were pooled and underwent PCR cleanup using Agecourt AMPure XP magnetic beads (Beckman Coulter, Pasadena, CA, USA). Amplicon DNA concentration was measured with an Invitrogen Qubit 2.0 (Carlsbad, CA, USA). We pooled 5ng DNA from each sample together with a mock sequence sample and sequenced at the Indiana University Center for Genomics and Bioinformatics on one lane of an Illumina MiSeq using 2x150bp reads.

Data analysis

Sequences were analyzed using mothur v1.40.5 (Schloss et al. 2009). Briefly, we merged reads for each sample, deduplicated and denoised reads, and removed chimeras with VSEARCH (Rognes et al. 2016). To classify our sequencing reads into microbial taxa, we aligned our reads to the SILVA v128 database (Quast et al. 2013). We used our known mock sequence to assess error rates and classified sequences into operational taxonomic units (OTUs) based on 99% similarity.

We focused most of our analyses on the effect of morph, age, and rearing environment in shaping the WTSP microbiome. All analyses were performed using Marker Data Profiling in the web tool MicrobiomeAnalyst (Dhariwal et al. 2017). We removed OTUs with less than four counts in 20% of the samples, removed OTUs with low variance (10% inter-quartile range), and rarefied samples to the lowest library size (3,258 reads). We examined alpha diversity by calculating bacterial community richness (observed OTUs, chao1; Chao 1984), community diversity (Shannon 1948), and community evenness (Simpson 1949). To examine beta diversity,

we calculated Bray-Curtis distances between samples (Bray & Curtis 1957). We also performed phylogenetic distance analyses with unweighted and weighted unifrac, which account for presence/absence of bacterial taxa and relative abundance of bacterial taxa, respectively (Lozupone & Knight 2005, Lozupone et al. 2007). We tested for differences with PERMANOVA and visualized with principal coordinates analysis (PCoA). We used a significance cutoff of $p<0.05$ for alpha and beta diversity analyses. We performed differential abundance analyses with DEseq2 (Love et al. 2014) using an FDR adjusted p-value cutoff of < 0.05. Lastly, we compared our WTSP microbiome samples with the remaining *Zonotrichia*. As sample sizes for the remaining species were low, we only performed descriptive analyses of bacterial phyla relative abundances across *Zonotrichia*.

Results

We sequenced 34 samples to an average depth of 27,880 reads per sample (range 3,258-425,790). After filtering, we used 723 OTUs for analysis. We focused our analyses on WTSP samples from our breeding population in New York. We observed substantial microbiome variation in the WTSP based on relative abundance of bacterial phyla and alpha diversity (Figure 4.1, Figure 4.2). Firmicutes and Proteobacteria were the most dominant phyla in our samples. Alpha and beta diversity analyses revealed no differences in morph, age, or nestling rearing environment ($p>0.05$, Table 4.1, Figure 4.2, Figure 4.3). However, comparisons of WTSP morph based on Shannon and Simpson alpha diversity were just above our significance threshold of $p<0.05$ (Table 4.1). Similarly, samples did not cluster by morph, age, or nestling rearing environment (Figure 4.3), nor did they cluster by relatedness (i.e. sibling, parent-offspring).

	Alpha Diversity				Beta Diversity		
	Chao1	**Observed OTUs**	**Shannon**	**Simpson**	**Bray-Curtis Distance**	**Unweighted UniFrac**	**Weighted Unifrac**
Morph	p=0.395 t=-0.873	p=0.145 t=-1.531	p=0.050 t=-2.111	p=0.055 t=-2.086	p=0.201 F=1.284 R^2=0.070	p=0.233 F=1.195 R^2=0.065	p=0.152 F-1.636 R^2=0.087
Age	p=0.641 t=0.483	p=0.822 t=0.231	p=0.998 t=-0.001	p=0.963 t=-0.048	p=0.985 F=0.391 R^2=0.022	p=0.926 F=0.636 R^2=0.036	p=0.939 F=0.147 R^2=0.009
Nest Type	p=0.465 t=0.804	p=0.5793 t=0.565	p=0.557 t=0.607	p=0.463 t=0.808	p=0.759 F=0.780 R^2=0.089	p=0.696 F=0.857 R^2=0.096	p=0.685 F=0.734 R^2=0.084

Table 4.1. Alpha and beta diversity metrics comparing WTSP morph, age, and nest type. T-tests were performed for alpha diversity measurements and PERMANOVA was performed for beta diversity measurements.

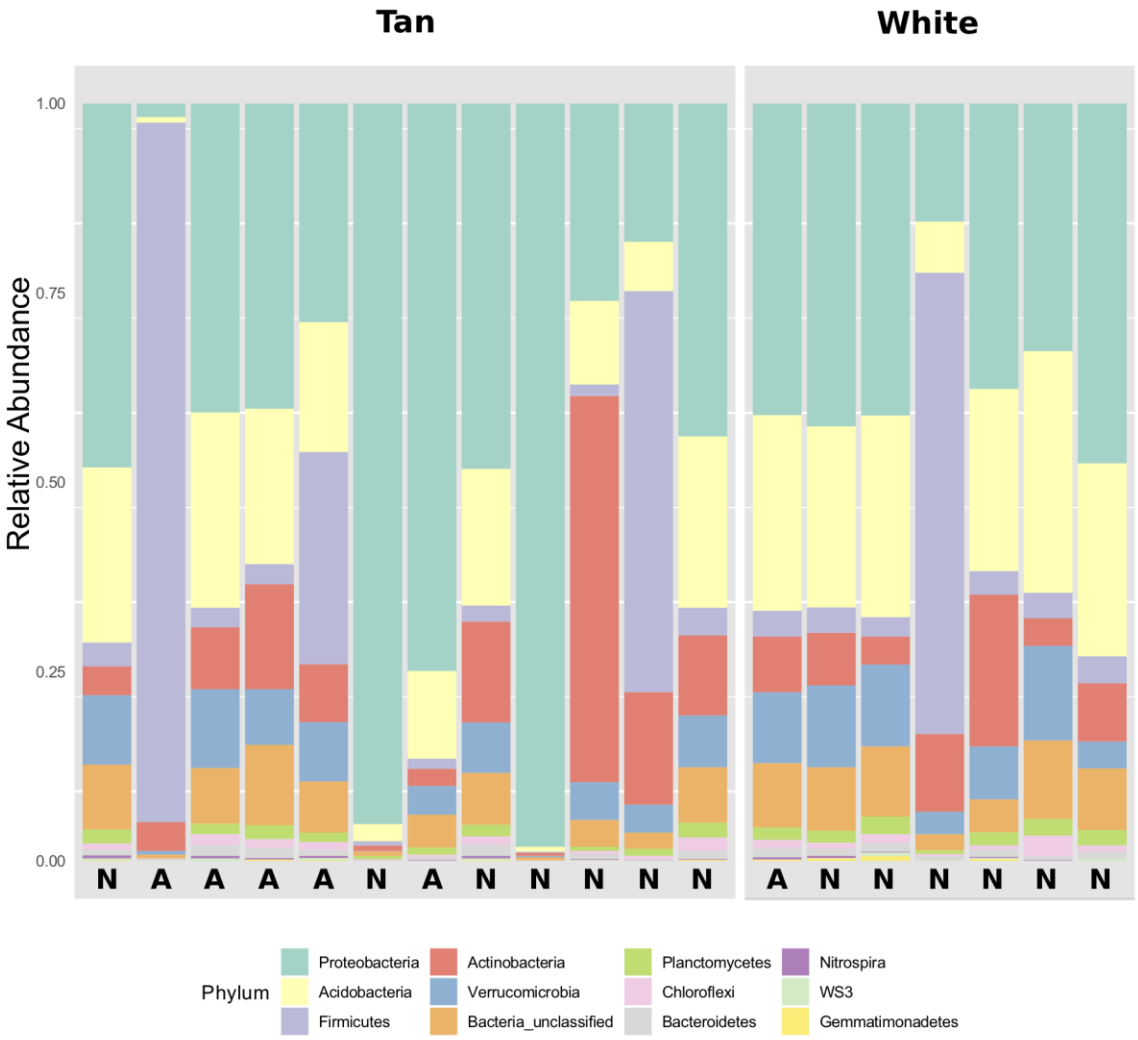


Figure 4.1. Relative abundance of bacterial phyla among WTSP morphs from New York WTSPs. 'N' denotes nestling samples and 'A' denotes adult samples.

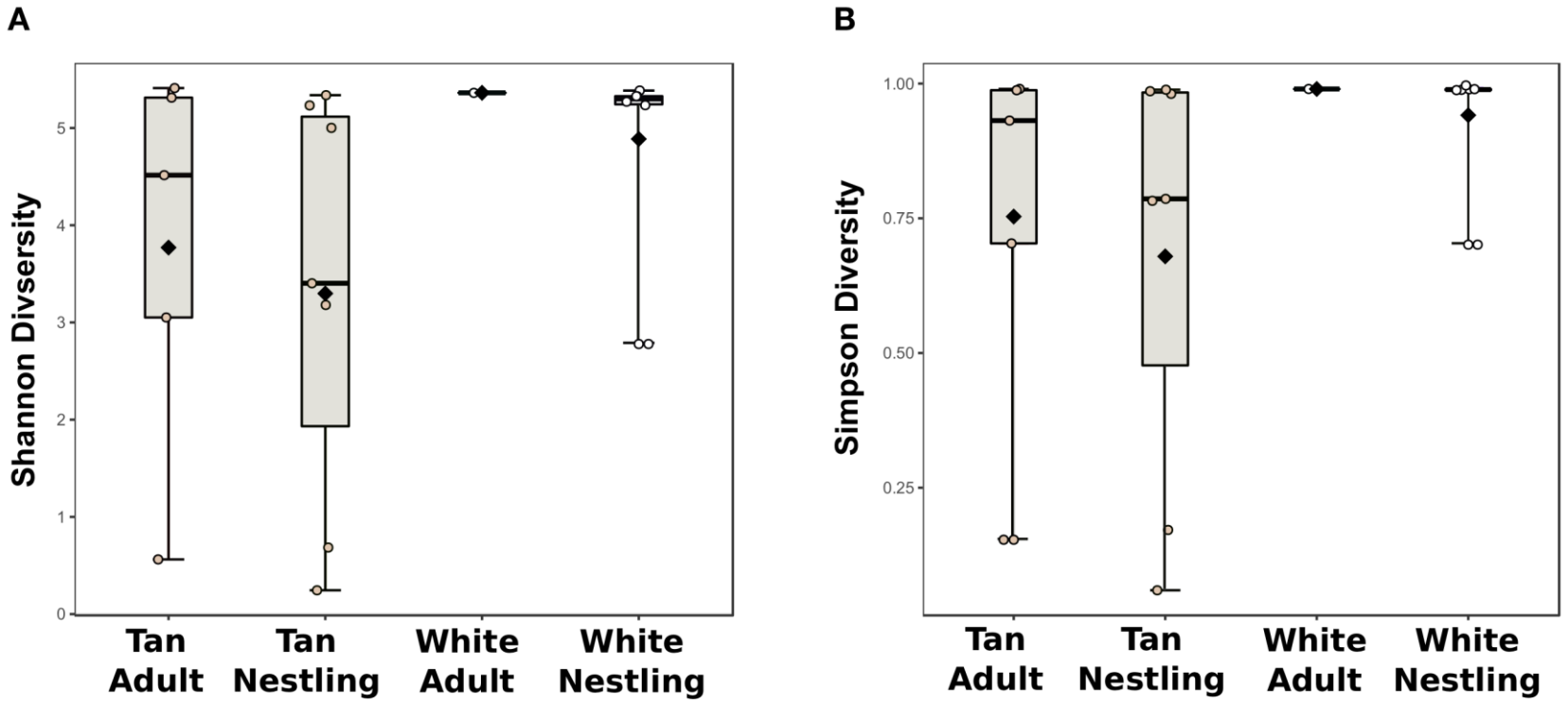


Figure 4.2. (A) Shannon diversity and (B) Simpson diversity indices among New York WTSP morph-age classes.

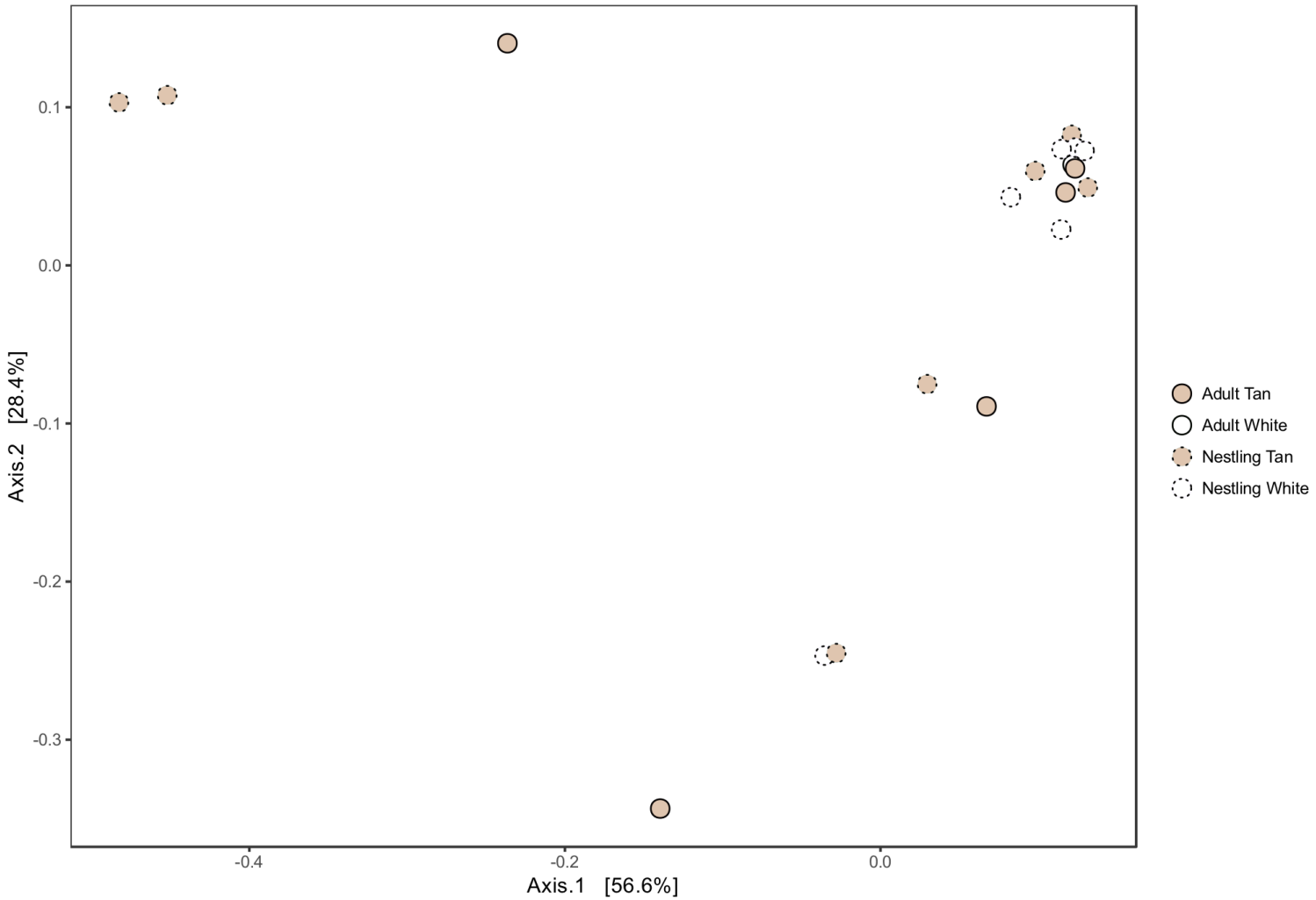


Figure 4.3. PCoA plot of weighted UniFrac values among New York WTSP morph-age classes.

We tested for differential abundance of OTUs with DEseq2 and found *Enterococcus* (FDR = $1.57x10^{-4}$) and unclassified Proteobacteria (FDR=0.047) were significantly increased in nestlings relative to adults. *Rhodococcus* (FDR=0.032) and unclassified Proteobacteria (FDR=$5.69x10^{-7}$) were significantly differentially abundant between the two nestling rearing types, with relatively high levels of each in nestlings from biparental care nests. This appears to have resulted from a few WTSP samples that were dominated by these taxa (Figure 4.1), because no OTUs were significantly differentially abundant after removing these WTSP samples.

Next, we assessed bacterial diversity among all samples in the study. The core microbiome of our samples is dominated by Proteobacteria, and to a lesser extent Acidobacteria, Firmicutes, Actinobacteria, unclassified Bacteria, and Verrucomicrobia (Figure 4.4). Similar to WTSP, we also observed substantial variation among the remaining *Zonotrichia* samples. Relative abundance of bacterial phyla fluctuated among samples and several samples were nearly entirely composed of Firmicutes (*Enterococcus*), Actinobacteria (unclassified Planococcaceae, *Rhodococcus*), or unclassified Proteobacteria (Supplemental Figure 1). Samples did not cluster by host species in hierarchical clustering (Supplemental Figure 2).

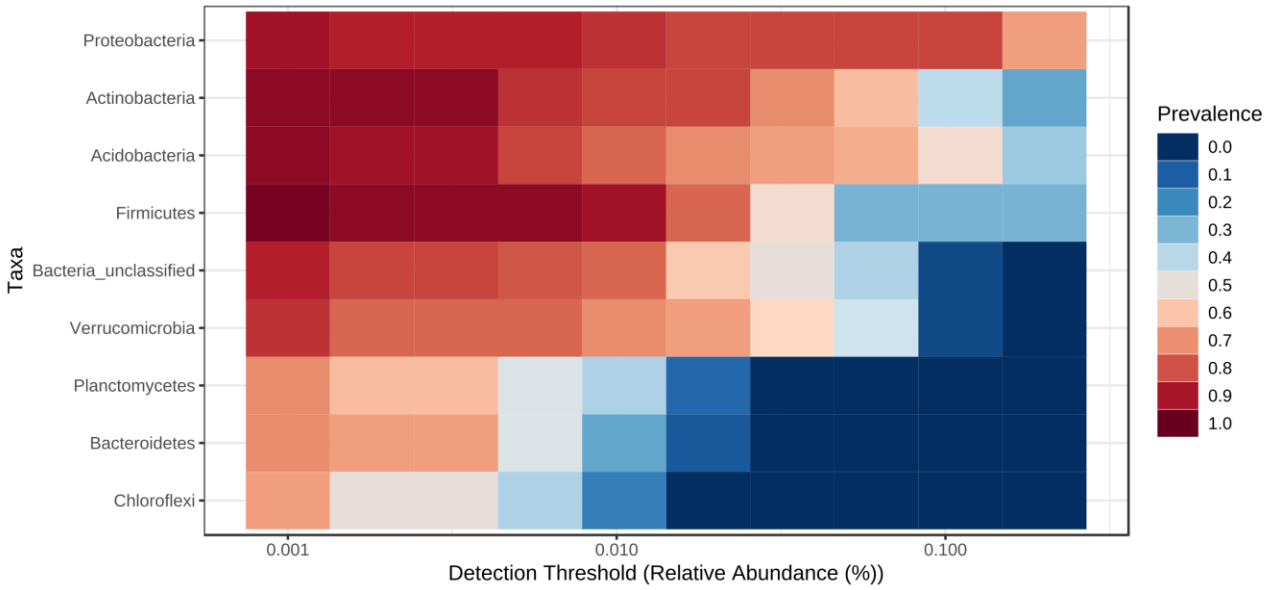


Figure 4.4. The core microbiome of *Zonotrichia* sparrows. Bacterial phyla are sorted by descending sample prevalence and descending relative abundance.

Discussion

We find no evidence of host genetic control in WTSPs, as morphs do not segregate in clustering analyses, do not differ in microbial diversity analyses, and have no OTUs differentially abundant. Interestingly, we observed extreme variation among WTSP samples in relative abundances of microbial taxa and this variation was reflected in our *Zonotrichia* level sampling. This extreme variation likely prevented any differences in alpha or beta diversity measurements in the WTSP microbiome based on age, morph, or nestling rearing environment. Only two OTUs were differentially abundant between nestlings and adults, and only one OTU was differentially abundant between nestling rearing environments.

We were particularly interested in the effect of WTSP morph on microbiome composition. Given the permanent genetic morphs of the WTSP, any differences would reflect a

substantial contribution of host genotype on microbiome formation. Both nestling and adult WTSP morphs appear to differ in the function of the immune system (Boyd et al. 2018, Newhouse et al. 2018). Host immunity and microbes are intimately linked and microbial composition in early life plays an important role in the development of the host immune system (reviewed in Belkaid and Hand 2014). White morph WTSP nestlings maintain elevated transcriptional activity of the innate immune system (Newhouse et al. 2018), and yet this does not seem to substantially alter microbial composition, as we observed no statistically significant differences based on alpha or beta diversity (Table 4.1). However, measurements of Shannon ($p=0.050$) and Simpson ($p=0.055$) diversity between WTSP morphs were near our significance threshold of $p<0.05$. White morph individuals tended to be more OTU-rich (Figure 4.2A) and displayed more community evenness (Figure 4.2B) than tan morph individuals. Low sample sizes of adult white morph WTSPs prevented adult morph comparisons.

Interestingly, we did not observe age-related differences between nestling and adult WTSP microbiomes (Table 4.1). We expected age-related differences, as the vertebrate gut microbiome typically changes from early life to adulthood following dietary transitions (reviewed in Voreades et al. 2014). Previous work in songbirds has revealed age differences (Kreisinger et al. 2017), including a recent study in house sparrows (*Passer domesticus*, Kohl et al. 2018). However, as house sparrows near fledging, their microbiome appeared to stabilize and become adult-like (Kohl et al. 2018). We sampled WTSP nestlings at day six post-hatch, which is near the average fledging date or eight days post-hatch in our population. Nestlings may have already developed a more adult-like microbiome and sampling throughout nestling development will be necessary to tease apart this relationship.

We also did not find any differences in nestling microbiome based on the nest type (i.e. nestling rearing environment). WTSPs exhibit two distinct parental care strategies, female-biased care and biparental care, which occur naturally in equal frequencies. Recent work in WTSP nestlings revealed a distinct transcriptional stress response in nestlings born into female-biased parental care nests, likely results from parental care variation (Newhouse et al. 2018). Stress typically alters microbiome composition resulting from intestinal dysbiosis (Foster et al. 2017, Noguera et al. 2018). Thus, we predicted nestlings in female-biased parental care nests might have a distinct microbiome from their biparental counterparts. The absence of differences between nestlings experiencing different parental care strategies suggests that stress may not sufficient enough to invoke intestinal dysbiosis. Alternatively, the diet among WTSP nestlings may be homogeneous in this population. Further work assessing nestling physiology, diet, and microbiome composition under different rearing environments is needed.

Our negative results could be due in part to the extreme variation we observed in the WTSP microbiome. Previous work showed that the avian microbiome is variable, but typically contains a consistent pattern of one or two dominant phyla across sampling groups (e.g. age, location) (reviewed in Waite & Taylor 2015, Grond et al. 2018). Rather, the relative abundance of several phyla (e.g. Firmicutes, Proteobacteria, Acidobacteria) fluctuated substantially throughout our samples (Figure 4.1). This trend appeared to be random and occurred in nestlings and adults, as well as white and tan morphs. A similar trend has recently been observed in the fecal microbiome of great tits (*Parus major,* Kropáčková et al. 2017a) and in a survey of European passerines (Kropáčková et al. 2017b). However, few studies have investigated the drivers of such variation in avian microbiomes. In chickens, the fecal microbiome is a stochastic mixture of taxa derived from more stable bacterial communities throughout the gastrointestinal

tract (Sekelja et al. 2012). As a result, the fecal microbiome is temporally more variable than other gastrointestinal tissues. The variation we observed here may therefore result from temporal collection of fecal samples and future work should include more gastrointestinal tissues for comparison.

We next performed a survey of the remaining *Zonotrichia* species and an outgroup to explore trends in microbiome variation among the host phylogeny. The *Zonotrichia* core microbiome is also quite variable and is dominated by Proteobacteria, followed by Acidobacteria, Actinobacteria, and Firmicutes (Supplemental Figure 1). This is largely similar to previous work characterizing the avian microbiome (Kohl 2012, Waite & Taylor 2014, Grond et al. 2018). However, our samples have a relatively large proportion of Proteobacteria and Acidobacteria, which are typically found at lower levels in birds (Grond et al. 2018). Previous work describing the cloacal microbiome in *Zonotrichia* sparrows has focused on RCSP (Escallón et al. 2019) and WCS (Phillips et al. 2018). These studies revealed similar high abundances of Proteobacteria, suggesting this phylum might be dominant in *Zonotrichia* sparrows. Samples did not cluster by species (Supplemental Figure 2), revealing that microbiome composition likely results from environmental rather than host genetic factors in these species. Each species was also sampled in a distinct environment, including captivity (RCSP & DEJ), yet samples did not cluster by sampling location, either. Sample sizes for most species were low in this study, limiting interpretation. Further work increasing the sample sizes, standardizing sampling time (e.g. breeding season), and including multiple sampling locations will be needed.

Here, we provided the first detailed analysis of the WTSP microbiome. Despite a strong genetic influence (i.e. morph) on WTSP physiology (Spinney et al. 2006, Swett & Breuner 2009, Horton & Holberton 2010, Horton et al. 2014), this did not appear to influence microbiome

formation. We also provided the first overview into the microbiome of *Zonotrichia* sparrows. In general, we observed substantial microbiome variation in our samples. Future work will be needed to increase sample sizes and extensively control for environment and diet to uncover the drivers of this observed variation.

Acknowledgements

Alice Boyle, Emily Hudson, Dai Shizuka, Lindsay Forrette, Margarida Barcelo-Serra, Rusty Gonser, Maria Stager, and Zac Cheviron provided samples used in this study. Mae Berlow provided guidance on laboratory work. Michael Dozier and Gina Bledsoe provided laboratory assistance.

Data Availability

Sequences will be uploaded to the NCBI Sequence Read Archive.

References

Belkaid, Y., & Hand, T. W. (2014). Role of the Microbiota in Immunity and Inflammation. *Cell*, *157*(1), 121–141. doi: 10.1016/j.cell.2014.03.011

Bodawatta, K. H., Sam, K., Jønsson, K. A., & Poulsen, M. (2018). Comparative Analyses of the Digestive Tract Microbiota of New Guinean Passerine Birds. *Frontiers in Microbiology*, *9*. doi: 10.3389/fmicb.2018.01830

Bray, J. R. and J. T. Curtis. (1957). An ordination of upland forest communities of southern Wisconsin. *Ecological Monographs* 27:325-349.

Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J., & Bordenstein, S. R. (2016). Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLOS Biology*, *14*(11), e2000225. doi: 10.1371/journal.pbio.2000225

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., & Bohannan, B. J. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *The ISME Journal*, *10*(3), 655–664. doi: 10.1038/ismej.2015.142

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., … Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, *6*(8), 1621–1624. doi: 10.1038/ismej.2012.8

Chao, A. (1984). Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, 11:265-270.

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews Genetics*, *13*(4), 260–270. doi: 10.1038/nrg3182

Contijoch, E. J., Britton, G. J., Yang, C., Mogno, I., Li, Z., Ng, R., … Faith, J. J. (2019). Gut microbiota density influences host physiology and is shaped by host and microbial factors. *ELife*, *8*. doi: 10.7554/eLife.40553

Dhariwal, A., Chong, J., Habib, S., King, I. L., Agellon, L. B., & Xia, J. (2017). MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Research*, *45*(W1), W180–W188. doi: 10.1093/nar/gkx295

Dong, T. S., & Gupta, A. (2019). Influence of Early Life, Diet, and the Environment on the Microbiome. *Clinical Gastroenterology and Hepatology*, *17*(2), 231–242. doi: 10.1016/j.cgh.2018.08.067

Escallón, C., Belden, L. K., & Moore, I. T. (2019). The Cloacal Microbiome Changes with the Breeding Season in a Wild Bird. *Integrative Organismal Biology*, *1*(1). doi: 10.1093/iob/oby009

Foster, J. A., Rinaman, L., & Cryan, J. F. (2017). Stress & the gut-brain axis: Regulation by the microbiome. *Neurobiology of Stress*, *7*, 124–136. doi: 10.1016/j.ynstr.2017.03.001

Goodrich, J. K., Davenport, E. R., Waters, J. L., Clark, A. G., & Ley, R. E. (2016). Cross-species comparisons of host genetic associations with the microbiome. *Science*, *352*(6285), 532–535. doi: 10.1126/science.aad9379

Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhman, R., … Ley, R. E. (2014). Human Genetics Shape the Gut Microbiome. *Cell*, *159*(4), 789–799. doi: 10.1016/j.cell.2014.09.053

Gould, A. L., Zhang, V., Lamberti, L., Jones, E. W., Obadia, B., Korasidis, N., … Ludington, W. B. (2018). Microbiome interactions shape host fitness. *Proceedings of the National Academy of Sciences*, *115*(51), E11951–E11960. doi: 10.1073/pnas.1809349115

Grieneisen, L. E., Charpentier, M. J. E., Alberts, S. C., Blekhman, R., Bradburd, G., Tung, J., & Archie, E. A. (2019). Genes, geology and germs: gut microbiota across a primate hybrid zone are explained by site soil properties, not host species. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1901), 20190431. doi: 10.1098/rspb.2019.0431

Hird, S. M. (2017). Evolutionary Biology Needs Wild Microbiomes. *Frontiers in Microbiology*, *8*. doi: 10.3389/fmicb.2017.00725

Hird, S. M., Carstens, B. C., Cardiff, S. W., Dittmann, D. L., & Brumfield, R. T. (2014). Sampling locality is more detectable than taxonomy or ecology in the gut microbiota of the brood-parasitic Brown-headed Cowbird (Molothrus ater). *PeerJ*, *2*, e321. doi: 10.7717/peerj.321

Horton, B. M., & Holberton, R. L. (2010). Morph-Specific Variation in Baseline Corticosterone and the Adrenocortical Response in Breeding White-Throated Sparrows (Zonotrichia albicollis). *The Auk*, *127*(3), 540–548. doi: 10.1525/auk.2010.09096

Horton, B. M., Moore, I. T., & Maney, D. L. (2014). New insights into the hormonal and behavioural correlates of polymorphism in white-throated sparrows, Zonotrichia albicollis. *Animal Behaviour*, *93*, 207–219. doi: 10.1016/j.anbehav.2014.04.015

Jones, R. M. (2016). The Influence of the Gut Microbiota on Host Physiology: In Pursuit of Mechanisms. *The Yale Journal of Biology and Medicine*, *89*(3), 285–297.

Kohl, K. D. (2012). Diversity and function of the avian gut microbiota. *Journal of Comparative Physiology B*, *182*(5), 591–602. doi: 10.1007/s00360-012-0645-z

Kohl, K. D., Brun, A., Caviedes-Vidal, E., & Karasov, W. H. (2019). Age-related changes in the gut microbiota of wild House Sparrow nestlings. *Ibis*, *161*(1), 184–191. doi: 10.1111/ibi.12618

Kostic, A. D., Howitt, M. R., & Garrett, W. S. (2013). Exploring host-microbiota interactions in animal models and humans. *Genes & Development*, *27*(7), 701–718. doi: 10.1101/gad.212522.112

Kreisinger, J., Kropáčková, L., Petrželková, A., Adámková, M., Tomášek, O., Martin, J.-F., … Albrecht, T. (2017). Temporal Stability and the Effect of Transgenerational Transfer on Fecal Microbiota Structure in a Long Distance Migratory Bird. *Frontiers in Microbiology*, *8*. doi: 10.3389/fmicb.2017.00050

Kropáčková, L., Pechmanová, H., Vinkler, M., Svobodová, J., Velová, H., Těšičký, M., … Kreisinger, J. (2017a). Variation between the oral and faecal microbiota in a free-living passerine bird, the great tit (Parus major). *PLOS ONE*, *12*(6), e0179945. doi: 10.1371/journal.pone.0179945

Kropáčková, L., Těšický, M., Albrecht, T., Kubovčiak, J., Čížková, D., Tomášek, O., … Kreisinger, J. (2017b). Codiversification of gastrointestinal microbiota and phylogeny in passerines is not explained by ecological divergence. *Molecular Ecology*, *26*(19), 5292–5304. doi: 10.1111/mec.14144

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, *489*(7415), 220–230. doi: 10.1038/nature11550

Lozupone, C., & Knight, R. (2005). UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235. doi: 10.1128/AEM.71.12.8228-8235.2005

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, *489*(7415), 220–230. doi: 10.1038/nature11550

Michel, A. J., Ward, L. M., Goffredi, S. K., Dawson, K. S., Baldassarre, D. T., Brenner, A., … Chaves, J. A. (2018). The gut of the finch: uniqueness of the gut microbiome of the Galápagos vampire finch. *Microbiome*, *6*(1). doi: 10.1186/s40168-018-0555-8

Moeller, A. H., Caro-Quintero, A., Mjungu, D., Georgiev, A. V., Lonsdorf, E. V., Muller, M. N., … Ochman, H. (2016). Cospeciation of gut microbiota with hominids. *Science*, *353*(6297), 380–382. doi: 10.1126/science.aaf3951

Newhouse, D., Barcelo Serra, M., Tuttle, E., Gonser, R., & Balakrishnan, C. (2018). Parent and offspring genotypes influence gene expression in early life. *BioRxiv*. doi: 10.1101/488684

Noguera, J. C., Aira, M., Pérez-Losada, M., Domínguez, J., & Velando, A. (2018). Glucocorticoids modulate gastrointestinal microbiome in a wild bird. *Royal Society Open Science*, *5*(4), 171743. doi: 10.1098/rsos.171743

Phillips, J. N., Berlow, M., & Derryberry, E. P. (2018). The Effects of Landscape Urbanization on the Gut Microbiome: An Exploration Into the Gut of Urban and Rural White-Crowned Sparrows. *Frontiers in Ecology and Evolution*, *6*. doi: 10.3389/fevo.2018.00148

Pluznick, J. L. (2014). Gut Microbes and Host Physiology: What Happens When You Host Billions of Guests? *Frontiers in Endocrinology*, *5*. doi: 10.3389/fendo.2014.00091

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. doi: 10.1093/nar/gks1219

Rodríguez, J. M., Murphy, K., Stanton, C., Ross, R. P., Kober, O. I., Juge, N., … Collado, M. C. (2015). The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial Ecology in Health & Disease*, *26*(0). doi: 10.3402/mehd.v26.26050

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

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., … Segal, E. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature*, *555*(7695), 210–215. doi: 10.1038/nature25973

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., … Weber, C. F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. doi: 10.1128/AEM.01541-09

Shannon C. A mathematical theory of communication. *Bell Syst. Tech. J*. 1948;27:379–423.

Simpson EH. Measurement of diversity. *Nature*. 1949;163:688.

Smith, C. C., Snowberg, L. K., Gregory Caporaso, J., Knight, R., & Bolnick, D. I. (2015). Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *The ISME Journal*, *9*(11), 2515–2526. doi: 10.1038/ismej.2015.64

Spinney, L. H., Bentley, G. E., & Hau, M. (2006). Endocrine correlates of alternative phenotypes in the white-throated sparrow (Zonotrichia albicollis). *Hormones and Behavior*, *50*(5), 762–771. doi: 10.1016/j.yhbeh.2006.06.034

Sullam, K. E., Rubin, B. E., Dalton, C. M., Kilham, S. S., Flecker, A. S., & Russell, J. A. (2015). Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. *The ISME Journal*, *9*(7), 1508–1522. doi: 10.1038/ismej.2014.231

Sun, C.-H., Liu, H.-Y., Zhang, Y., & Lu, C.-H. (2018). Comparative analysis of the gut microbiota of hornbill and toucan in captivity. *MicrobiologyOpen*, e786. doi: 10.1002/mbo3.786

Suzuki, T. A. (2017). Links between Natural Variation in the Microbiome and Host Fitness in Wild Mammals. *Integrative and Comparative Biology*, *57*(4), 756–769. doi: 10.1093/icb/icx104

Swett, M. B., & Breuner, C. W. (2009). Plasma Testosterone Correlates with Morph Type across Breeding Substages in Male White-Throated Sparrows. *Physiological and Biochemical Zoology*, *82*(5), 572–579. doi: 10.1086/605392

Thorneycroft, H B. (1966). Chromosomal polymorphism in the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Science*, *154*(3756), 1571 LP – 1572.

Thorneycroft, H Bruce. (1975). A cytogenetic study of the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Evolution*, *29*(4), 611–621. doi: 10.1111/j.1558-5646.1975.tb00855.x

Trevelline, B. K., Fontaine, S. S., Hartup, B. K., & Kohl, K. D. (2019). Conservation biology needs a microbial renaissance: a call for the consideration of host-associated microbiota

in wildlife management practices. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1895), 20182448. doi: 10.1098/rspb.2018.2448

Tuttle, E. (2003). Alternative reproductive strategies in the white-throated sparrow: behavioral and genetic evidence. *Behavioral Ecology*, *14*(3), 425–432.

Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., … Balakrishnan, C. N. (2016). Divergence and Functional Degradation of a Sex Chromosome-like Supergene. *Current Biology*, *26*(3), 344–350. doi: 10.1016/j.cub.2015.11.069

Voreades, N., Kozil, A., & Weir, T. L. (2014). Diet and the development of the human intestinal microbiome. *Frontiers in Microbiology*, *5*. doi: 10.3389/fmicb.2014.00494

Waite, D. W., & Taylor, M. W. (2014). Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Frontiers in Microbiology*, *5*. doi: 10.3389/fmicb.2014.00223

Waite, D. W., & Taylor, M. W. (2015). Exploring the avian gut microbiota: current trends and future directions. *Frontiers in Microbiology*, *6*. doi: 10.3389/fmicb.2015.00673

APPENDIX A: Animal Use Declaration

No animal research was performed at East Carolina University. Chapter 1 was performed at the USGS National Wildlife Health Center. Chapter 2 used publicly available data obtained from the NCBI Sequence Read Archive. Chapters 3 & 4 were performed at the Cranberry Lake Biological Station in collaboration with Indiana State University. Additional data for Chapter 4 was generated at the University of Montana, Indiana State University, Kansas State University, and the University of Nebraska-Lincoln.