Author for correspondence:
Marcella D. Baiz
e-mail: mbaiz@buffalo.edu

THE ROYAL SOCIETY PUBLISHING

Association between the gut microbiome and carotenoid plumage phenotype in an avian hybrid zone

Marcella D. Baiz[1], Andrew W. Wood[2] and David P. L. Toews[2]

[1]Department of Biological Sciences, University at Buffalo, Buffalo, NY 14260, USA
[2]Department of Biology, Pennylvania State University, University Park, PA 16802, USA

MDB, 0000-0002-1629-6737

Vertebrates host complex microbiomes that impact their physiology. In many taxa, including colourful wood-warblers, gut microbiome similarity decreases with evolutionary distance. This may suggest that as host populations diverge, so do their microbiomes, because of either tight coevolutionary dynamics, or differential environmental influences, or both. Hybridization is common in wood-warblers, but the effects of evolutionary divergence on the microbiome during secondary contact are unclear. Here, we analyse gut microbiomes in two geographically disjunct hybrid zones between blue-winged warblers (*Vermivora cyanoptera*) and golden-winged warblers (*Vermivora chrysoptera*). We performed 16S faecal metabarcoding to identify species-specific bacteria and test the hypothesis that host admixture is associated with gut microbiome disruption. Species identity explained a small amount of variation between microbiomes in only one hybrid zone. Co-occurrence of species-specific bacteria was rare for admixed individuals, yet microbiome richness was similar among admixed and parental individuals. Unexpectedly, we found several bacteria that were more abundant among admixed individuals with a broader deposition of carotenoid-based plumage pigments. These bacteria are predicted to encode carotenoid biosynthesis genes, suggesting birds may take advantage of pigments produced by their gut microbiomes. Thus, host admixture may facilitate beneficial symbiotic interactions which contribute to plumage ornaments that function in sexual selection.

1. Introduction

Vertebrates occupy environments teeming with microscopic organisms. Moreover, individual animals themselves serve as an ecological niche occupied by diverse microbial communities that vary in diversity across body sites and over time [1]. Microbiome variation between individuals can also be great; however, variation between conspecifics is often less than variation between heterospecific individuals [2,3]. In many host taxa, the extent of gut microbiome differentiation is positively associated with evolutionary distance, an observation called phylosymbiosis [4]. This may suggest that as host populations diverge, so do their microbiomes, either as a result of tight coevolutionary dynamics with their hosts, or reflecting differential environmental influences in allopatry, or both.

A question that has been under-explored is, what happens to these divergent microbiomes when host populations come into secondary contact? During allopatric divergence, differentiation of host-associated microbiomes may lead to species-specificity. If hybridization occurs upon secondary contact, hybrid microbiomes could include influence from both parental taxa. Because bacteria can form intimate associations with their host and carry out important physiological functions, hybrid offspring may carry deleterious combinations of species-specific microbes or suffer incompatible interactions involving divergent host alleles. This framework is an extension of the Dobzhansky–Muller model of hybrid incompatibilities [5,6], which accounts for differentiation between parental species' microbiomes [7,8].
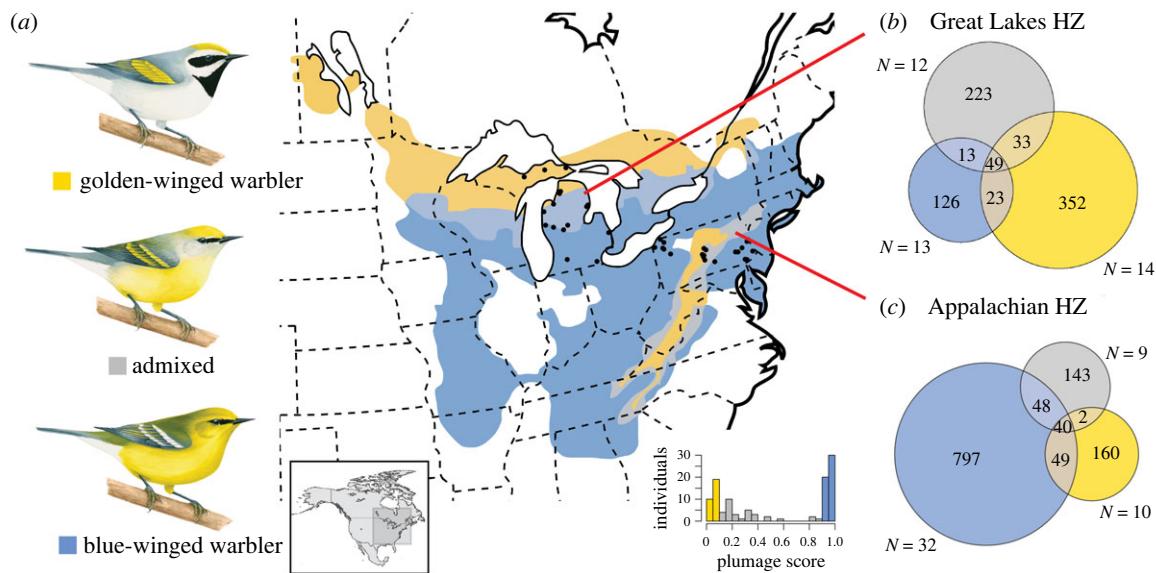
Figure 1. Summary of study system and sample data. (*a*) Parental plumage phenotypes, and an intermediate admixed (Brewster's) phenotype, Illustrated by Liz Clayton Fuller. The map shows the breeding range of the study species and our sampling localities. The histogram depicts the distribution of individuals included in this study based on our plumage score of admixture. (*b,c*) Number of amplicon sequence variants (ASVs) represented in individuals scored for plumage characters in the rarefied dataset years 2021–2022 for Great Lakes individuals (*b*), and Appalachian individuals (*c*). Sample sizes are indicated beside each group in (*b,c*). HZ, hybrid zone.

However, hybridization is not always deleterious. Beneficial mutations can spread between species via adaptive introgression, and admixture creates novel recombinants that can be favoured by selection [9]. Hybrid hosts may provide an intermediate environment (e.g. physiology, morphology, diet) to their symbionts that is favourable to a wider diversity of microbes than either parental species is independently able to accommodate, increasing the number of potentially beneficial interactions. Similarly, admixed genomes may confer parental alleles that allow hybrids to take advantage of beneficial microbial products that are only usable by one parental species. The frequency at which hybrids are expected to face beneficial, deleterious, or neutral symbiotic interactions in natural systems is unclear.

In a previous analysis of 15 wood-warbler species, we found that gut microbiomes exhibit phylosymbiosis [10]. Hybridization is common in the warbler family and several sister taxa form natural hybrid zones [11]. In hybrid zones, parental and hybrid individuals coexist, so environmental influences on their microbiomes should be similar. Further, parental alleles are shuffled among admixed individuals, which may allow divergent host phenotypes to be associated with certain features of the microbiome. Here, we present an analysis of gut microbiome variation in the blue-winged warbler × golden-winged warbler (*Vermivora cyanoptera* × *Vermivora chrysoptera*) hybrid zone system.

Non-recombining, mitochondrial DNA suggests that blue-winged and golden-winged warblers are sister taxa that diverged at 1.1 Ma [12]. They are the only two extant members of their genus, which diverged from its sister clade at approximately 6 Ma [13]. By contrast, nuclear genome-wide divergence is extremely low, consistent with a long history of gene flow [14,15]. In fact, divergence is limited to six small genomic regions, most of which contain genes known to be involved in plumage traits [15]. The parental plumages primarily differ in two pigment classes: the presence/absence of a melanin-based (black) facial patterning and the extent of carotenoid-based (yellow) body coloration (figure 1*a*). Birds synthesize melanin pigments via the evolutionarily conserved melanogenesis pathways, whereas they are unable to synthesize carotenoid pigments but are known to deposit dietary carotenoids into feather tissue [16]. Blue-winged warblers are very yellow overall (i.e. their feathers have many visible carotenoids), with blue–grey wings bearing thin white wing bars, and a black eyeline. Golden-winged warblers are grey and white overall, with a striking black throat patch and black eye mask. In golden-winged warblers, yellow plumage is restricted to the crown and wing bars. Hybrid individuals exhibit a continuum of variation along these two trait axes [17].

Both species are long-distance migrants. In the non-breeding season, blue-winged warblers occur in Mexico, Central America and the Caribbean, whereas golden-winged warblers occur primarily in Central America and northern South America. Their breeding ranges overlap in two disjunct regions in eastern North America. In the Great Lakes region, interbreeding occurs where the species' distributions overlap at their parapatric range boundary, while in the Appalachian region the hybrid zone is more mosaic as the range of the golden-winged Warbler is enveloped within the range of the blue-winged warbler (figure 1*a*). Breeding habitat preferences of the parental species are very similar [18], but differences may occur in some areas [19]. Notably, habitat differences are likely more divergent on disparate non-breeding grounds [20]. The warblers are insectivores that tend to associate with early successional, shrubby patches containing few tall trees. In contact zones, parental and admixed individuals can be found nesting in the same patch. Thus, this system serves as a useful model for understanding species-specific influences that shape avian microbiomes during early stages of speciation.

In this study, we quantify gut microbiome variation across two regions of hybridization in *Vermivora* warblers. Our aims are to (i) identify species-specific gut bacteria, (ii) quantify potential differences in gut microbiome diversity between individuals in the Great Lakes and the Appalachian regions, and (iii) test the hypotheses that admixed individuals exhibit disrupted microbiome diversity (i.e. divergent, or outside the typical range of parental microbiome diversity) and that microbiome variation is correlated with host admixture. To quantify gut microbiome diversity, we performed 16S amplicon sequencing of faecal samples. We also predicted metagenome features for admixed individuals to generate hypotheses about the function of bacteria associated with host traits.

2. Methods

(a) Sampling

We captured birds in mist-nets during May–July each year between 2019 and 2022 (figure 1a). In 2019 and 2020, we only sampled birds in the Appalachian hybrid zone. In 2021 and 2022, we sampled birds in both hybrid zones (electronic supplementary material, table S1). We held birds in a paper bag for up to 10 min before collecting whole faeces by scraping them directly into a sample tube containing 0.5% SDS lysis buffer [21]. We froze samples at −20°C within two weeks of collection until DNA extraction was performed. In total, we analysed 158 samples from 144 individuals.

(b) DNA extraction and sequencing

We processed and sequenced faecal samples following the procedures described in Baiz *et al.* [10]. Briefly, we extracted DNA using an SPRI bead DNA extraction method. Before carrying out amplifications, we randomized samples to avoid within-plate batch effects during polymerase chain reaction (PCR). Next, we prepared 16S amplicon libraries using a two-step multiplex dual-index approach. In the first step, we used universal PCR 515F/806R primers [22] to target the V4 region of the bacterial 16S rRNA gene and amplified each sample in triplicate. In the second step, we appended dual P5 and P7 Illumina indexes to each library via PCR so that samples could be pooled and multiplexed. See Baiz *et al.* [10] for reaction conditions and details of our protocol.

We included two types of negative controls in our sequencing to ensure minimal contamination. First, negative extraction controls ($N = 4$) were processed the same way as faecal samples, except we used sample buffer (not containing faeces) in the DNA extraction. Second, we sequenced one negative control from each PCR ($N = 5$).

We quantified DNA using a Qubit 4.0 fluorometer (Invitrogen), and normalized library concentrations by pooling an equal amount of DNA from each PCR. Samples were then sequenced using the 600-cycle kit on a single lane of an Illumina MiSeq instrument with 250 × 250 paired-end sequencing at the Penn State Genomics Core Facility.

(c) Sequence data processing

We imported sequence data in *QIIME 2* v.2020.8 [23] using *qiime dada2 denoise-paired*, and assigned taxonomic classification to amplicon sequence variants (ASVs) using the SILVA database v.138 SSURef NR99 [24]. We removed ASVs assigned as mitochondria, chloroplast, eukaryote, or that could not be assigned. We also removed possible contaminant ASVs using the *decontam* package v.1.16.0 [25], by using the prevalence method (threshold = 0.5) to detect ASVs more prevalent in negative controls than in faecal samples. We then used *qiime phylogeny align-to-tree-mafft-fasttree* to generate a phylogenetic tree from our ASV sequences.

We next rarefied our dataset to standardize sequencing depth across our samples (range 39–88 461 reads per sample, mean = 24 100 reads per sample). We used *vegan::rarecurve* v.2.6-4 [26] to visualize the minimum sequencing depth necessary to capture the majority of ASVs sequenced in each sample (electronic supplementary material, figure S1). Based on this analysis, we used *phyloseq::rarefy_even_depth* v.1.40.0 [27] to rarefy our dataset such that all samples contained 3076 reads. We removed 25 samples that fell below this threshold and ten additional samples from nine recaptured individuals, such that each individual is only represented by one sample in the final dataset. After filtering and rarefaction, our dataset contained 4564 ASVs and 123 samples.

(d) Host admixture

We determined the extent of host admixture by visually assessing the plumage phenotype. Previous analysis of 132 individuals has shown that plumage phenotype is nearly perfectly correlated with genome-wide admixture ($R^2 = 0.94$) and thus is a very strong indicator of hybrid ancestry in this system [15]. We scored nine plumage characters that vary between the parental species (electronic supplementary material, table S2). Plumage scores vary from 0 to a maximum sum of 32. We calculated plumage scores as $1 - $ (character sum/32), such that golden-winged warbler individuals received a score of 0 and blue-winged warbler individuals received a score of 1. In some tests, we used the plumage score of the six carotenoid characters (i.e. 'carotenoid score'), which ranges from 5 to 20, where blue-winged warblers have low carotenoid scores (more yellow) and golden-winged warblers have high carotenoid scores (less yellow).

We could not score all nine plumage characters for four individuals because either they had begun to moult feathers in their greater secondary coverts, so we could not score wing bar width ($N = 2$ individuals, 8 of 9 characters scored), or our scoring was limited because certain characters were not captured in photographs taken of the bird during banding by colleagues who kindly provided us samples ($N = 2$ individuals, 5 and 7 of 9 characters could be scored). Because the majority of characters could be scored for these individuals, we included them in our analysis and calculated plumage scores by dividing the sum of scored characters by the maximum possible sum of the scored characters. For our categorical analyses involving admixture, we considered 'admixed' individuals to have a plumage score between 0.1 and 0.9 ($N = 32$), blue-winged warblers to have a plumage score greater than 0.9 ($N = 50$), and golden-winged warblers to have a plumage score less than 0.1 ($N = 29$). There were 12 individuals for which plumage could not be given a numerical score because samples were kindly provided by colleagues who did not capture photographs, but noted no obvious signs of admixture (golden-winged warbler: $N = 1$, blue-winged warbler $N = 11$). These categorical analyses included a total of 111 individuals with numerical plumage scores.

For some tests, we performed categorical analyses based solely on the dominant plumage phenotype (regardless of the numerical plumage score), where we consider admixed individuals to only include the 'named' hybrid classes. Our dataset contains samples from intermediate hybrids with discrete plumage characteristic of the 'Brewster's' hybrid type, thought by some to include the F1 phenotype [28], though this has not been confirmed. We categorized hybrids as 'Brewster's' ($N = 12$) if they were mostly grey in colour, but lacked the black facial patches characteristic of golden-winged warblers. Otherwise, individuals were categorized as either parental species based on their dominant plumage phenotype (blue-winged warbler: $N = 64$, golden-winged warbler: $N = 47$). These analyses included all 123 individuals in the rarefied dataset, and represent a crude yet conservative analysis of gut microbiome variation across plumage phenotypes.

(e) Hybrid zone analyses

For analyses where we analysed contact zones separately, we partitioned samples by region to account for potential fine-scale gut microbiome variation due to geography (electronic supplementary material, figure S2). We partitioned each contact zone into three regions

based on localities where we encountered obvious intermediate hybrids (i.e. 'Brewster's' type). In the Great Lakes hybrid zone, we encountered only phenotypic blue-winged warblers south of latitude 43.4° N. Sampling sites north of this region, but within the lower peninsula, are where we encountered all Brewster's type individuals. The remaining sites were dominated by individuals with the golden-winged warbler phenotype. In the Appalachian hybrid zone, we encountered only individuals with blue-winged plumage at sites west of longitude −78° W and east of longitude −77.1° W. Between these coordinates we encountered Brewster's-type hybrids.

(f) Gut microbiome diversity

We calculated two different measures of gut microbiome diversity: beta diversity quantifies dissimilarity between-individual microbiomes, and alpha diversity quantifies within-individual microbiome diversity (i.e. richness). For beta diversity, we calculated four different metrics using the *phyloseq::distance*. Bray–Curtis distance measures differences in community composition weighted by ASV abundances, whereas Jaccard is based only on ASV presence/absence. Unweighted UniFrac measures the phylogenetic distance between communities based on ASV presence/absence, whereas weighted UniFrac is similar but weights branch lengths by ASV abundance. We used permutational multivariate analysis of variance (PERMANOVA) to test for significant effects of different variables on beta diversity using *vegan::adonis2*, and set the 'by' parameter to 'margin' to account for effects of multiple variables. A homogeneous dispersion among groups is an assumption for PERMANOVA tests, so we calculated multivariate homogeneity of group dispersions for significant variables using *vegan::betadisper* and used permutation tests to test for deviations from homogeneity using *vegan::permutest*.

We calculated three alpha diversity metrics, using *phyloseq::estimate_richness* to calculate the Chao1 and Shannon index and *btools::estimate_pd* v.0.0.1 [29] to calculate Faith's phylogenetic diversity (PD). Chao1 quantifies ASV richness, while Shannon index accounts for evenness of the community (the equity in ASV abundances). PD is a measure of the sum of branch lengths in the phylogeny representing the ASV community assemblage. We fitted linear models (LMs) using the *lm* function in R and ran type II analysis-of-variance (ANOVA) tests to assess the significance of different variables on alpha diversity using *car::Anova* v.3.1-2 [30].

In our PERMANOVA and LM analyses, we tested for additive effects of sampling year (2019–2022), species (based on numerical plumage score: blue-winged/admixed/golden-winged), hybrid zone locality (Great Lakes/Appalachia) or sampling region (in separate analysis of each hybrid zone). We alternatively ran each model with hybrid zone (or region) nested within sampling year to account for the substructure in our study design. Results of our nested models were very similar to the additive models and are reported in electronic supplementary material, tables S3 and S4.

(g) Identifying amplicon sequence variants associated with host traits

We used microbiome multivariable associations with linear models—*MaAsLin2* v.1.10.0 [31]—to identify ASVs associated with host features, including species identity and plumage scores. We also included sampling year and hybrid zone as fixed effects. We used rarefied ASV counts as input in our models, and used blue-winged warblers as the reference group for tests involving categorical fixed effects of the host. Significance was determined by applying the default Benjamini–Hochberg correction for multiple hypothesis testing using a false discovery rate (FDR) threshold of $Q < 0.25$.

We verified significant associations by repeating analyses using ASV relative abundances instead of rarefied counts as input (electronic supplementary material, table S5). We calculated relative abundances using the pre-rarefied data by dividing the ASV count by the total number of reads for that sample. Our results were generally consistent, so we only report results from analyses using rarefied ASV counts in the main text.

(h) Microbiome function

We used *PICRUSt2* v.2.0.3_b [32] to predict the functional significance of microbiome features associated with plumage scores. *PICRUSt2* places 16S rRNA gene sequences in a phylogeny containing reference genomes from the Integrated Microbial Genomes database, uses the tree to predict gene-family copy numbers for each ASV, and produces a predicted metagenome for each sample. Mean nearest sequenced taxon index value for our samples was 0.040 ± 0.04 s.d., suggesting the available reference genomes allowed high accuracy of metagenome predictions [33]. We outputted the predicted abundance of *Kyoto encyclopedia of genes and genomes* (KEGG) orthologues, and of KEGG enzymes for each sample, and then used the *MaAsLin2* approach outlined above to identify orthologues and enzymes for which abundance is associated with plumage traits. We used the KEGG mapper reconstruct tool (https://www.genome.jp/kegg/mapper/reconstruct.html) to link KEGG pathways to orthologues and enzymes that had significant associations with plumage traits.

3. Results

(a) General patterns of diversity

We detected 4564 ASVs in our rarefied dataset, classified in 30 phyla. Proteobacteria represented 36% of the total reads, followed by Actinobacteriota (17% of reads), then Planctomycetota (11% of reads). Remaining phyla each represented less than 7% of the total reads (electronic supplementary material, figure S3).

We were able to assess host ancestry by scoring plumage for 111 individuals, among which we delimited 4300 ASVs. Only 234 of these ASVs (5.4%) were shared among blue-winged, golden-winged and admixed individuals. Proteobacteria and Actinobacteria represented 44 and 24% of the shared ASVs, and other phyla represented less than 6% (electronic supplementary material, table S6). When examining samples from 2021 to 2022 (the years we sampled both hybrid zones), golden-winged individuals carried the most private ASVs (not shared between groups) in the Great Lakes, whereas in Appalachia, blue-winged individuals carried the most private ASVs (figure 1b). The latter was likely influenced by uneven sample sizes (figure 1); however, the increase in the number of private blue-winged ASVs is disproportionate to the increase in sample size for that group, suggesting the pattern is biologically meaningful.

In the full dataset, admixed individuals tended to exhibit higher mean alpha diversity (electronic supplementary material, figure S4), though linear models revealed an effect of sampling year (LM: Chao1 $F = 50.40$, $p < 0.001$; Shannon $F = 30.11$, $p < 0.001$; PD $F = 47.29$, $p < 0.001$), but no effect of admixture or hybrid zone locality (electronic supplementary material, table S3). This pattern is due to unusually high alpha diversity for samples collected in 2020 (electronic supplementary material, figure S5). Similarly, year explained a significant amount of variation between microbiomes (PERMANOVA: Bray–Curtis $R^2 = 0.061$, $p = 0.001$; Jaccard $R^2 = 0.045$, $p = 0.001$; Unifrac $R^2 = 0.068$, $p = 0.001$), as did hybrid zone locality when examining Jaccard distance ($R^2 = 0.013$, $p = 0.001$). Permutation tests indicated that beta dispersion was not homogeneous among samples collected in different years (Bray–Curtis and Jaccard, $p < 0.01$), or in different hybrid zone localities (Jaccard, $p = 0.001$). Inspection of dispersion plots revealed tight clustering among samples collected in 2019–2020 from the Appalachian hybrid zone, and 2020 samples clustered mostly separate from samples collected from both hybrid zones in 2021–2022 (electronic supplementary material, figure S6).

(b) Species-specific microbiomes

To identify features of species-specific microbiomes, we examined the subset of individuals with plumage scores inconsistent with admixture (i.e. parental individuals). We identified a single ASV strongly associated with host species (model 1, table 1). This ASV (Proteobacteria: genus *Rickettsia*) was more abundant in golden-winged than blue-winged warblers. It was also present in eight out of 32 (25%) admixed individuals, but was only present in a single blue-winged warbler sample at low abundance (count = 8).

(c) Effects of hybrid zone locality

To further examine potential differences in patterns of microbiome diversity between hybrid zones, we excluded samples from years 2019–2020 because we only collected samples from both localities in 2021–2022.

We first examined the subset of parental individuals separately in each hybrid zone to ask whether variation in parental microbiomes differs between hybrid zones. We found that species and sampling region both explained a significant amount of variation between parental microbiomes in the Great Lakes hybrid zone when looking at UniFrac distance (PERMANOVA: species $R^2 = 0.054$, $p = 0.045$; sampling region $R^2 = 0.106$, $p = 0.025$; figure 2a), and permutation tests were consistent with homogeneous beta dispersions for each variable ($p > 0.05$).

In the Appalachian hybrid zone, however, species did not explain a significant amount of variation between microbiomes using any distance metric (figure 2b). Sampling region and year were the only variables that were significant, and this was when looking at Jaccard distance (PERMANOVA: region $R^2 = 0.055$, $p = 0.016$; year $R^2 = 0.030$, $p = 0.014$). Beta dispersion was non-homogeneous for samples collected in different years ($F = 5.914$, $p = 0.023$), and clustering appeared tighter for samples collected in 2021. Similarly, when examining alpha diversity of parental individuals we found sampling region to be the only variable to show a significant effect on Chao1 diversity in the Great Lakes hybrid zone only (LM: $F = 3.75$, $p = 0.040$). This effect seems to be driven by increasing Chao1 values for golden-winged warblers in the north as we found a positive relationship between latitude and Chao1 values (LM: $F = 5.48$, $p = 0.028$), while accounting for sampling year and host species in the model (figure 2c).

Because we found some evidence for an effect of hybrid zone locality on beta diversity, we next examined blue-winged, admixed, and golden-winged individuals separately to determine if the effect size of hybrid zone locality differed among groups. When using Jaccard distance, hybrid zone locality explained 6.2% of the variation between golden-winged warbler microbiomes ($p = 0.002$), 5.5% of the variation between microbiomes of admixed individuals ($p = 0.040$), and 2.6% of the variation between blue-winged warbler microbiomes ($p = 0.041$). We did not observe a significant effect of hybrid zone locality on alpha diversity in any group (electronic supplementary material, table S3). Sampling year again explained a small amount of variation in some tests for alpha and beta diversity (electronic supplementary material, table S3). Beta dispersion among samples was homogeneous when grouped by each variable, except in the case of sampling year for admixed individuals when using weighted UniFrac ($F = 18.554$, $p = 0.003$)—samples collected in 2021 clustered more tightly than 2022.

(d) Relationship between admixture and gut microbiome diversity

We ran several *MaAsLin2* analyses to examine the relationship between admixture phenotype and ASV abundances, including hybrid zone locality and sampling year as fixed effects. We first ran a model where we grouped individuals by dominant plumage phenotype to identify ASVs possibly associated with each discrete host phenotype (figure 3). This analysis revealed an ASV in the genus *Jatrophihabitans* (phylum Actinobacteriota), associated with the 'Brewster's' phenotype (model 2, table 1). This ASV was represented in 19% of the admixed individuals (including non-'Brewster's'), but only 8 and 7% of blue-winged, and golden-winged individuals, respectively, and it occurred in samples from both hybrid zones and in samples collected in years 2020, 2021 and 2022. This model also uncovered another Actinobacteriota ASV in the genus *Kineococcus* with elevated counts for individuals with the blue-winged warbler phenotype, and we again detected the *Rickettsia* ASV associated with the golden-winged warbler phenotype (model 2, table 1). Only one of the 32 admixed individuals carried both the blue-winged warbler-associated ASV and the golden-winged warbler ASV at low counts (1 and 15 reads, respectively). This individual had a plumage score of 0.25 and was sampled in the Appalachian hybrid zone.

To assess the robustness of our results to the way we coded phenotype, we ran an alternative model that included the continuous numerical plumage score instead of the dominant plumage phenotype. In this model, we identified one ASV negatively associated with plumage score (model 3, table 1). This was the same *Rickettsia* ASV associated with golden-winged warblers.

Next, we retained only admixed individuals to minimize species-effects and more closely examine a model fitted on individuals with recombinant plumage characters. In this model, plumage score was associated with four ASVs, each in the phylum Proteobacteria (genera *Sphingomonas*, *Pseudomonas*, *Allorhizobium–Neorhizobium–Pararhizobium–Rhizobium*, *Methylobacterium–Methylorubrum*).
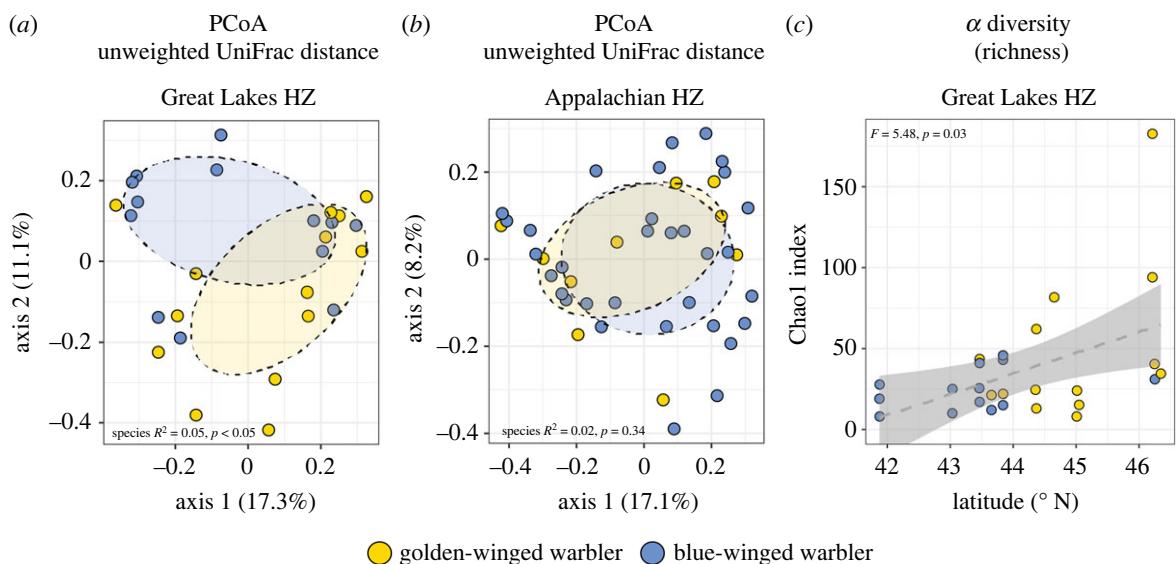
Figure 2. Gut microbiome diversity of parental individuals in each hybrid zone. (*a*,*b*) Principal coordinate analysis (PCoA) of unweighted UniFrac distance between microbiomes in the Great Lakes (*a*), and Appalachia (*b*). Ellipses are drawn at 50% confidence level. (*c*) Positive association between gut microbiome richness (Chao1 index) and latitude in the Great Lakes hybrid zone.

Table 1. Summary of ASVs significantly associated with plumage phenotypes (FDR < 0.25). 'variable' is the fixed effect associated with ASV abundance. ASV IDs are SILVA identifiers. Association denotes the direction and level (for categorical variables) of the association between ASV abundance and the fixed effect variable. 'coef.' is the *MaAsLin2* model coefficient and FDR is the false discovery rate.

model	variable, dataset	ASV ID	taxonomy	association	coef.	FDR	*p*-value
1	species, parental individuals	ca0365837eefeb9bf94c58dd59c0e8c3	*Rickettsia* sp.	+golden-winged warbler	0.810	0.069	0.002
2	dominant plumage phenotype, full dataset	dd070aa4a7da33679f1038d652140369	*Jatrophihabitans* sp.	+Brewster's	1.837	0.032	0.002
		b1dde0ed7ab1f036845b5643cfaa699e	*Kineococcus* sp.	—golden-winged warbler	−1.130	0.154	0.012
		ca0365837eefeb9bf94c58dd59c0e8c3	*Rickettsia* sp.	+golden-winged warbler	0.712	0.013	<0.001
3	plumage score, full dataset	ca0365837eefeb9bf94c58dd59c0e8c3	*Rickettsia* sp.	—	−0.325	0.046	0.003
4	plumage score, admixed individuals	b5c8ed4496d79b916859402cdf11f0a7	*Sphingomonas* sp.	+	0.804	0.093	0.003
		5aec9bd35889489e4a05c78a82358060	*Pseudomonas* sp.	+	0.875	0.133	0.005
		fc6622c636a5210293fb2873fc4761d9	*Allorhizobium–Neorhizobium–Pararhizobium–Rhizobium* sp.	+	0.987	0.210	0.011
		04ecfad5772d2e09a84a0f5ef460536c	*Methylobacterium–Methylorubrum* sp.	+	1.241	0.237	0.013
5	carotenoid score, admixed individuals	b5c8ed4496d79b916859402cdf11f0a7	*Sphingomonas* sp.	—	−0.811	0.057	0.002
6	carotenoid score, full dataset	ca0365837eefeb9bf94c58dd59c0e8c3	*Rickettsia* sp.	+	0.261	0.230	0.023
		3ab43fcc04514ef20f0fe8689a8c5ea7	family: Microbacteriaceae	—	−0.346	0.230	0.023
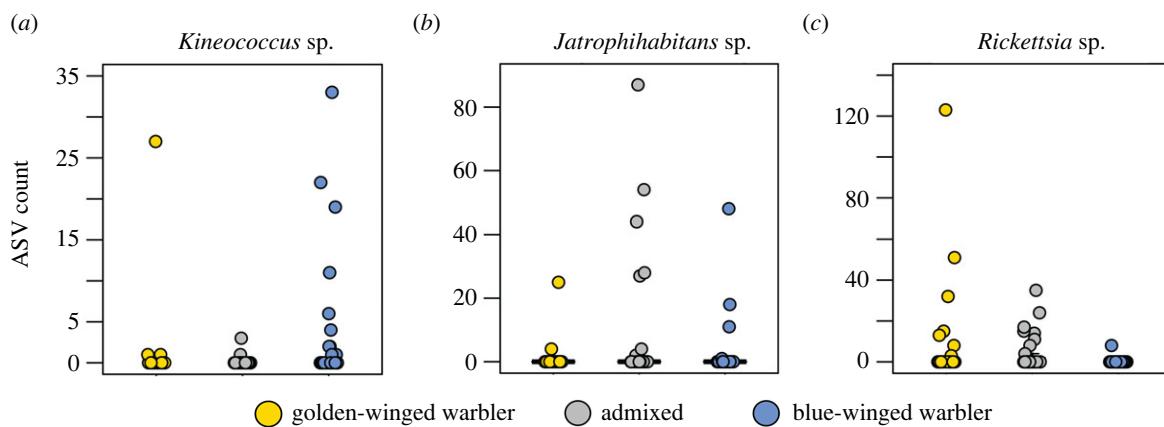
Figure 3. Counts of ASVs associated with host phenotype. We identified one ASV associated with each discrete phenotype: blue-winged warbler (*a*), Brewster's (*b*) and golden-winged warbler (*c*). Individuals are grouped by ancestry according to plumage score. Summary statistics for associations shown in each panel are provided in table 1 (model 2).

Abundance of each ASV was positively associated with plumage score, so higher ASV counts were observed in admixed individuals with more blue-winged warbler-like characters (model 4, table 1). Blue-winged warblers are very yellow in colour, as a result of carotenoid pigments in their plumage. Six of the nine plumage characters in our plumage score are carotenoid-based characters, suggesting abundances of these ASVs may be closely tied to carotenoid deposition in plumage. When we used plumage score of only the six carotenoid characters (i.e. 'carotenoid score') in our model, the *Sphingomonas* ASV remained significant (model 5, table 1). When including parental individuals in the model with carotenoid score, we detected an additional ASV in the family Microbacteriaceae (phylum Actinobacteriota) associated with more-carotenoid-rich (blue-winged-like) plumage, as well as the *Rickettsia* ASV reported above, which was again associated with less-carotenoid-rich (golden-winged-like) plumage (model 6, table 1).

(e) Predicted function of admixed microbiomes

Our metagenome prediction analysis yielded 7579 KEGG orthologs and 2314 KEGG enzymes in our dataset. The majority of these features were represented in predicted metagenomes of admixed individuals (6986 orthologues, 92%; 2158 enzymes, 93%).

In a model including only admixed individuals, we identified 132 KEGG orthologues for which abundance is associated with plumage score (electronic supplementary material, table S7), including four orthologues in the 'carotenoid biosynthesis' pathway (*crtW*: beta-carotene ketolase (CrtW type) coef. = 1.442, FDR = 0.229, $p = 0.010$; *crtC*: carotenoid 1,2-hydatase coef. = 1.690, FDR = 0.108, $p = 0.003$; *crtD*: 1-hydroxycarotenoid 3,4-desaturase coef. = 1.690, FDR = 0.108, $p = 0.003$; *crtF*: demethylspheroidene *O*-methyltransferase coef. = 1.690, FDR = 0.108, $p = 0.003$). Other orthologues significantly associated with plumage score could be linked to various metabolic pathways as well as pathways under 'environmental information processing', 'cellular processes', and 'genetic information processing' (electronic supplementary material, table S8). When using plumage carotenoid score in the model, we identified 23 significant orthologues (electronic supplementary material, table S9), 19 of which overlapped with the results from the model using plumage score, including one orthologue in the carotenoid biosynthesis pathway (*crtW* coef. = −1.412, FDR = 0.224, $p = 0.008$). Each of these orthologues in the carotenoid biosynthesis pathway exhibited an association with plumage score such that admixed individuals with more blue-winged warbler-like characters (broader carotenoid plumage) exhibited higher ASV abundances.

We identified 43 KEGG enzymes with abundance positively associated with plumage score in admixed individuals (electronic supplementary material, table S10), including three enzymes in the 'carotenoid biosynthesis' pathway, which correspond to the KEGG orthologues identified above (*crtC* coef. = 1.656, FDR = 0.121, $p = 0.003$; *crtD* coef. = 1.656, FDR = 0.121, $p = 0.003$; *crtF* coef. = 1.648, FDR = 0.128, $p = 0.003$). Using plumage carotenoid score in the model yielded five significant enzymes, none of which was in the carotenoid biosynthesis pathway (electronic supplementary material, table S11).

4. Discussion

Our analyses support a limited effect of host species on gut microbiome diversity in this hybrid zone system, although we identified one ASV associated with each parental species and one ASV associated with admixed individuals. Patterns in microbiome diversity differed between the two disjunct hybrid zones we studied. Species identity explained a small amount of variation between microbiomes in the Great Lakes, but not in the Appalachian hybrid zone. Hybrid zone locality explained more variation among golden-winged warbler than blue-winged warbler microbiomes. We found that admixed individuals generally did not host both species-specific ASVs, and that abundance of certain ASVs is correlated with the degree of admixture. Metacommunity predictions suggest that admixed individuals with broader deposition of carotenoid pigments in their plumage host gut bacteria that encode genes in the carotenoid biosynthesis pathway.

(a) Limited species-effects on the *Vermivora* gut microbiome

In a previous analysis of 15 warbler species—not including *Vermivora*—we found that host species identity explained more variation between gut microbiomes than sampling year, locality, or diet diversity [10]. Further, gut microbiome similarity was

correlated with evolutionary relatedness between host species, suggesting evolutionary history may play a significant role in shaping gut microbiome structure in this clade. Thus, we may expect the gut microbiome to be associated with host genetics, as has been shown in other taxa [34]. In the current study, however, effects of sampling year and locality generally explained more variation between gut microbiomes of parental species.

We hypothesize the limited species-effect we observed likely relates to the unique evolutionary history of blue-winged and golden-winged warblers. The parental species diverged relatively recently, at approximately 1 Ma [12], such that they have not evolved strong reproductive isolating barriers [14], and their genomes are highly homogenized, likely due to a long history of hybridization [15]. If the warbler gut microbiome is under host genetic control, perhaps limited genome-wide divergence constrains structuring of the gut microbiome between these parental species. Alternatively, environmental effects could overpower species-effects on gut microbiome structure. Differences in habitat preference are not readily apparent as both species can be found in early successional shrublands, and occur in the same patch within hybrid zones. This may mean hybrid zone individuals nest in the same vegetation [18] and consume the same prey. Thus, limited habitat divergence could explain the broadly homogenized microbiomes we observed for parental taxa in sympatry.

Despite limited microbiome structuring by species, we did observe some host-effects. Consistent with our findings among a larger sampling of warbler species [10], approximately 5% of the variation between microbiomes of parental individuals in the Great Lakes region was explained by species when using UniFrac distance. This suggests that in the Great Lakes region, parental species host somewhat taxonomically distinct gut microbiomes.

We also discovered one ASV associated with each parental phenotype. Abundance of the *Rickettsia* ASV exhibited a significant correlation with plumage score among admixed individuals, consistent with additive genetic effects. This may suggest that the *Rickettsia* ASV exhibits a biased affinity for hosts having more golden-winged warbler ancestry. *Rickettsia* are obligate intracellular bacteria known to be associated with invertebrates that, when transmitted to vertebrates, can act as pathogens [35]. Therefore, it is possible that golden-winged warblers preferentially consume a type of invertebrate prey that hosts this ASV, and this dietary preference is genetic, or that golden-winged warblers are preferentially associated with some invertebrate parasite (e.g. tick) transmitting this ASV to its host.

The blue-winged warbler plumage phenotype was associated with an ASV in the genus *Kineococcus*. Cultured species in this genus grow in orange-colored colonies and encode genes in the carotenoid biosynthesis pathway [36]. *Kineococcus* has been reported in the caecum of native chickens [37] and thus might be a common avian gut symbiont. We only detected this *Kineococcus* ASV at low counts (less than 4 reads) in two of the admixed individuals, suggesting high affinity for non-admixed blue-winged warblers.

(b) Hybrid zone locality drives differences in microbiome diversity

We observed different patterns in *Vermivora* gut microbiome diversity between the two disjunct hybrid zones we studied. Greater differentiation among parental microbiomes in the Great Lakes could be explained by differences in carry-over effects from non-breeding sites [20], within-species population structure, hybrid zone structure and geography, or a combination of these factors. Blue-winged warblers in both hybrid zones and golden-winged warblers that breed in the Great Lakes migrate to Central America after breeding, whereas golden-winged warblers that breed in Appalachia migrate to South America [20]. As a result, in the non-breeding season Appalachian golden-winged warblers are exposed to environments in South America that have recently experienced deforestation at higher rates than birds on non-breeding grounds in Central America [20]. Thus, carry-over effects from a wider geographical range of non-breeding sites may disproportionately drive microbiome differentiation between breeding birds in different contact zones for golden-winged warblers. This difference in migratory connectivity is associated with genetic differentiation [38], so it is also possible that within-species population genetic structure is greater for golden-winged warblers, which could directly drive gut microbiome structure if the microbiome is under genetic control. However, this difference in migratory behaviour and population structure cannot explain our finding that host species explains some gut microbiome variation in the Great Lakes—where both species migrate to Central America—but not in Appalachia, where parental species migrate to disparate non-breeding grounds (figure 2).

Ecogeographic differences between contact zones may also help explain our findings. In the Great Lakes, blue-winged and golden-winged warblers have a parapatric distribution, and individuals with intermediate plumage occur at the centre of the contact zone. Thus, parental individuals in the north and south are allopatric. This contact zone lies at the boundary of the northern forest and eastern temperate forest ecoregions. However, in the Appalachian hybrid zone, the range of golden-winged warblers is enveloped within the distribution of blue-winged warblers and parental individuals occur more sympatrically. The individuals we sampled here occur in a single ecoregion—the eastern temperate forest. Therefore, although nesting habitat preferences are highly similar among the parental species, Great Lakes individuals—especially golden-winged individuals in the north—may experience some different environment-specific influences on the gut microbiome compared with Appalachian individuals. This is supported by our findings of a greater influence of hybrid zone locality on gut microbiome similarity among golden-winged warblers than among blue-winged warblers (electronic supplementary material, table S3).

We acknowledge that we sampled across a relatively narrow transect in each hybrid zone, and that broader sampling could refine or refute our interpretations. Nevertheless, our findings highlight the potential nuance of environmental influences on the gut microbiome, which vary across the landscape within this system. This suggests that for hybridizing taxa, gut microbiome variation is dependent on the geographical context of secondary contact and is possibly shaped by carry-over effects from non-breeding grounds.

(c) Features of admixed microbiomes

Previous vertebrate gut microbiome studies of hybrid individuals have focused on mammals and fish. In these studies, shifts in hybrid gut microbiome composition relative to parental individuals are common, including effects dependent on the direction of

the cross, and elevated alpha diversity for F1 hybrids [8]. Elevated alpha diversity in hybrid microbiomes might be expected if parental microbiomes are characterized by distinct features, if genetic ancestry underlies microbiome composition, and as a result hybrid microbiomes encompass a mixture of species-specific microbiome features (as hosts do for plumage traits). Our results are generally inconsistent with this, as we found no significant differences in alpha diversity among groups. One of the two species-specific ASVs we identified (genus *Kineococcus*) was only represented in two admixed individuals, and at low counts. Further, one of these individuals was the only admixed individual we examined that carried both species-specific ASVs. This pattern of limited occurrence of species-specific ASVs in admixed microbiomes—and limited co-occurrence of these ASVs—would be expected for bacteria that are incompatible on the heterospecific genetic background, and/or incompatible with heterospecific microbes in the shared gut environment [7].

Similar to parental individuals, a large proportion of the ASVs in admixed individuals were private. One ASV in the genus *Jatrophihabitans* exhibited higher abundance among admixed individuals. Collectively, these results suggest that although admixed *Vermivora* microbiomes are not highly distinct from parental taxa, there are some discernible differences. *Jatrophihabitans* species have been isolated from soil [39], and plant tissue [40]. Thus, it is possible that warblers acquire this ASV by gleaning invertebrates from vegetation. However, the mechanism by which admixed individuals attain elevated abundances is not clear. *Jatrophihabitans* taxa have also been detected in other animal gut microbiome datasets [41], but to our knowledge the function of this genus in the avian gut has not been studied.

Our analyses do not point to any overt signs of dysbiosis (i.e. unhealthy, disrupted microbiome composition) for admixed individuals. We did not detect any ASVs with elevated abundance in admixed microbiomes that are known to cause disease in birds, and microbiome richness did not differ significantly between admixed and parental individuals. Although we are not feasibly able to measure health or fitness for individuals in our current dataset, we observed several individuals, including admixed males that returned to breed in multiple years, indicating good health. As an anecdote, we sampled one 'Brewster's' male in the Great Lakes that had been banded by another researcher in 2015—banding records therefore indicate that this individual was likely 8 years old at the time of our sampling. This longevity of admixed males may not be uncommon, as one historical account documented a Brewster's male that returned to breed for 6 years at a site in the Appalachian hybrid zone [42]. It is important to note that hybrid inviability or hybrid sterility is expected to disproportionately affect female warblers because, in birds, females are heterogametic [43], though hybrid inviability/sterility has not been studied in *Vermivora*. Our dataset only included seven females, none of which exhibited admixed plumage. Thus, it may not be unexpected that admixed microbiomes do not show overt dysbiosis in our dataset. It will be the target of future research to study the relationship between the microbiome and the fertility and viability of early generation hybrids of both sexes.

(d) Gut microbiomes vary with the degree of admixture

We found some support for our hypothesis that microbiome variation is associated with the degree of host admixture. This finding is consistent with results from a limited number of hybrid zone studies [44,45] (but see [46]). We identified several ASVs for which abundance correlates with host ancestry based on plumage characters. One of these ASVs (genus *Rickettsia*) had higher abundance in golden-winged than in blue-winged warblers, consistent with additive effects of host ancestry underlying ASV abundance.

Genome-wide differentiation between golden-winged and blue-winged warblers is restricted to six regions, four of which are upstream of genes known to be associated with plumage traits [15]. This constrains the possible number of divergent alleles that may underlie microbiome composition in an ancestry-dependent manner. It may also suggest that plumage genes dictate the proliferation of certain taxa in the *Vermivora* gut microbiome.

Research on model organisms has uncovered a clear link between host genotypes and the microbiome [34]. Many previous studies link host genes that function in immunity, growth and development to certain features of the microbiome. In birds, research on the genetic basis of microbiome assembly has been limited to poultry and, to our knowledge, ours is the first study on wild birds to link microbiome features to host traits that vary with genetic ancestry.

More research is necessary to identify host genes that potentially underlie microbiome structure in warblers. Although phylosymbiosis is not as strong in birds as in non-flying mammals [47], it has become clear that microbiome differentiation is expected to follow host divergence in some avian clades [2,10,48,49]. The current study highlights the utility of hybrid zones for uncovering the genetic basis of avian microbiome assembly in a clade exhibiting phylosymbiosis [10].

(e) Metagenome predictions link carotenoid biosynthesis pathway to plumage variation

Our results uncovered an unexpected and exciting potential link between the gut microbiome and carotenoid plumage coloration. Like other vertebrates, birds cannot endogenously produce carotenoid pigments, and it is known that they derive plumage carotenoids from their diet [16,50]. Warblers are insectivores and almost exclusively consume arthropods. A straightforward mechanism for carotenoid uptake is via the consumption of herbivorous arthropods as a source of plant-derived carotenoids. Bacteria-derived carotenoids have been overlooked as a source of plumage pigments, but our results and other recent work [51] suggest the gut microbiome could be directly linked to avian coloration.

The abundance of several ASVs was elevated in admixed individuals that were more yellow overall. Further, analysis of metagenome predictions uncovered correlations between plumage scores and predicted abundances of KEGG orthologues and enzymes in the carotenoid biosynthesis pathway. This suggests that the *Vermivora* gut microbiome contains bacteria that encode carotenoid-processing genes, and that gut microbiome function may contribute to plumage colour variation in this system.

Notably, each of the ASVs correlated with yellower plumage scores are related to taxa known to produce carotenoid pigments (*Rhizobium* [52], Microbacteriacae [53], *Sphingomonas* [54], *Pseudomonas* [55], *Methylorubrum* [56]). In culture, colonies of certain taxa in the family Microbacteriaceae and genera *Sphingomonas* and *Pseudomonas* are yellow in colour. Further, pigments produced by

Sphingomonas include zeaxanthin [57], which is an isomer of lutein—the dominant plumage carotenoid of two warbler species closely related to *Vermivora* [50].

Limited genome-wide divergence between parental species includes a region upstream of the beta-carotene oxygenase 2 (*BCO2*) gene [15], which encodes an enzyme that cleaves carotenoids like lutein into shorter products that do not get incorporated in feathers. Thus, it is possible that *Vermivora* genotypes at *BCO2* regulate the processing of bacterial-derived carotenoids, which would provide a potential link between the gut microbiome, host genetics and plumage phenotype. We have observed interannual variation in carotenoid coloration for admixed individuals that we recaptured in multiple years [17], suggesting a possible environmental influence such as interannual variation in the proliferation of carotenoid-producing gut bacteria, and/or variation in diet.

Evidence for use of bacterial-derived carotenoids in host coloration has been reported in molluscs [58], sponges [59] and whiteflies [60]. To our knowledge, our results represent the first evidence in vertebrate hosts. Carotenoid-pigmented plumage plays an important role in mate selection and is an honest indicator of male quality [61]. In golden-winged warblers, plumage ornamentation is associated with territory quality and may influence male–male competition [62]; however, the role of plumage colour variation in the reproductive success of admixed individuals has not been studied.

Our results add to recent studies that draw links between microbiomes and sexual selection [63]; however, additional research is necessary to confirm the gut microbiome as a source of plumage carotenoids for warbler hosts. One caveat to our study is that we sampled birds during the breeding season after they moulted into their breeding plumage, which they began to do at the end of the previous breeding season [64]. Thus, our plumage scores reflect the deposition of plumage carotenoids that occurred the prior year and our interpretations presume the bacteria we identified in our samples were present at similar abundances during that time. This would be consistent with the minor influence of sampling year on gut microbiome composition that we observed here (excluding 2020 samples), and the limited annual effects observed in previous studies [10,65]. We note that our finding of unusually high alpha diversity for samples collected in 2020 mirrors results in our recent analysis of gut microbiome diversity among 15 other warbler species [10], a pattern that possibly resulted from batch effects during sample processing.

Our interpretations are also limited by our metabarcoding approach. Our ASV counts may be biased because PCR amplification was used [66]. We attempted to address this by analysing ASV relative abundances in addition to rarefied counts, and results of both approaches were generally consistent (electronic supplementary material, table S5). Additionally, our functional interpretations are based on metagenome predictions because 16S metabarcoding only captures a single region of the bacterial genome. It will be desirable to use a more precise approach to elucidate the possible link between gut microbiome function and plumage coloration.

5. Concluding remarks

To our knowledge, this is the first study to analyse admixed gut microbiomes in an avian hybrid zone. We found differences in the factors that explained microbiome variation in each disjunct hybrid zone, and uncovered several ASVs for which abundance is correlated with host ancestry. Our results suggest that both the geographical context of admixture and host genotype are important in shaping gut microbiome differentiation between these young taxa. Our analysis of metagenome predictions supports a link between gut microbiome function and plumage colour variation. This may suggest that birds take advantage of carotenoids produced by their symbiotic gut bacteria, which would implicate gut microbiome function in sexual selection.

References

1. Lloyd-Price J et al. 2017 Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**, 61–66. (doi:10.1038/nature23889)

2. Trevelline BK, Sosa J, Hartup BK, Kohl KD. 2020 A bird's-eye view of phylosymbiosis: weak signatures of phylosymbiosis among all 15 species of cranes. *Proc. R. Soc. B* **287**, 20192988. (doi:10.1098/rspb.2019.2988)

3. Bodawatta KH, Koane B, Maiah G, Sam K, Poulsen M, Jønsson KA. 2021 Species-specific but not phylosymbiotic gut microbiomes of New Guinean passerine birds are shaped by diet and flight-associated gut modifications. *Proc. R. Soc. B* **288**, 20210446. (doi:10.1098/rspb.2021.0446)

4. Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR. 2016 Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biol.* **14**, e2000225. (doi:10.1371/journal.pbio.2000225)

5. Muller H. 1942 Isolating mechanisms, evolution, and temperature. *Biol. Symp.* **6**, 71–125.

6. Dobzhansky T. 1936 Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* **21**, 113–135. (doi:10.1093/genetics/21.2.113)

7. Brucker RM, Bordenstein SR. 2012 Speciation by symbiosis. *Trends Ecol. Evol.* **27**, 443–451. (doi:10.1016/j.tree.2012.03.011)

8. Miller AK, Westlake CS, Cross KL, Leigh BA, Bordenstein SR. 2021 The microbiome impacts host hybridization and speciation. *PLoS Biol.* **19**, e3001417. (doi:10.1371/journal.pbio.3001417)

9. Marques DA, Meier JI, Seehausen O. 2019 A combinatorial view on speciation and adaptive radiation. *Trends Ecol. Evol.* **34**, 531–544. (doi:10.1016/j.tree.2019.02.008)

10. Baiz MD, Benavides CA, Miller ET, Wood AW, Toews DPL. 2023 Gut microbiome composition better reflects host phylogeny than diet diversity in breeding wood-warblers. *Mol. Ecol.* **32**, 518–536. (doi:10.1111/mec.16762)

11. Williamson JL, Gyllenhaal EF, Oliver KD, Brady SS, Johnson AB, Michelsohn MJ, Andersen MJ. 2021 Predictable outcomes of warbler hybridization: synthesis and an exceptional yellow × black-throated blue barbler (*Setophaga petechia* × *S. caerulescens*) pairing. *Wilson J. Ornithol.* **133**, 82–102. (doi:10.1676/21-00031)

12. Wood AW, Szpiech ZA, Lovette IJ, Smith BT, Toews DPL. 2023 Genomes of the extinct Bachman's warbler show high divergence and no evidence of admixture with other extant *Vermivora* warblers. *Curr. Biol.* **33**, 2823–2829. (doi:10.1016/j.cub.2023.05.058)

13. Barker FK, Burns KJ, Klicka J, Lanyon SM, Lovette IJ. 2015 New insights into New World biogeography: an integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. *Auk* **132**, 333–348. (doi:10.1642/AUK-14-110.1)

14. Vallender R, Friesen VL, Robertson RJ. 2007 Paternity and performance of golden-winged warblers (*Vermivora chrysoptera*) and golden-winged X blue-winged warbler (*V. pinus*) hybrids at the leading edge of a hybrid zone. *Behav. Ecol. Sociobiol.* **61**, 1797–1807. (doi:10.1007/s00265-007-0413-3)

15. Toews DPL, Taylor SA, Vallender R, Brelsford A, Butcher BG, Messer PW, Lovette IJ. 2016 Plumage genes and little else distinguish the genomes of hybridizing warblers. *Curr. Biol.* **26**, 2313–2318. (doi:10.1016/j.cub.2016.06.034)

16. Toews DPL, Hofmeister NR, Taylor SA. 2017 The evolution and genetics of carotenoid processing in animals. *Trends Genet.* **33**, 171–182. (doi:10.1016/j.tig.2017.01.002)

17. Baiz MD, Kramer GR, Streby HM, Taylor SA, Lovette IJ, Toews DPL. 2020 Genomic and plumage variation in *Vermivora* hybrids. *Auk* **137**, ukaa027. (doi:10.1093/auk/ukaa027)

18. Will TC. 1986 The behavioral ecology of species replacement: blue-winged and golden-winged warblers in Michigan. PhD thesis, University of Michigan, Ann Arbor, MI. See https://www.proquest.com/docview/303489548?pq-origsite=gscholar&fromopenview=true.

19. Confer JL, Barnes KW, Alvey EC. 2010 Golden- and blue-winged warblers: distribution, nesting success, and genetic differences in two habitats. *Wilson J. Ornithol.* **122**, 273–278. (doi:10.1676/09-136.1)

20. Kramer GR *et al.* 2018 Population trends in *Vermivora* warblers are linked to strong migratory connectivity. *Proc. Natl Acad. Sci. USA* **115**, E3192–E3200. (doi:10.1073/pnas.1718985115)

21. White P, Densmore L. 1992 Mitochondrial DNA isolation. In *Molecular genetic analysis of populations: a practical approach* (ed. AR Hoelzel), pp. 29–58. Oxford, UK: IRL Press at Oxford University Press.

22. Caporaso JG *et al.* 2012 Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624. (doi:10.1038/ismej.2012.8)

23. Bolyen E *et al.* 2019 Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857. (doi:10.1038/s41587-019-0209-9)

24. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013 The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596. (doi:10.1093/nar/gks1219)

25. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018 Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226. (doi:10.1186/s40168-018-0605-2)

26. Oksanen J *et al.* 2022 *vegan: Community ecology package*. See https://CRAN.R-project.org/package=vegan.

27. McMurdie PJ, Holmes S. 2013 phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217. (doi:10.1371/journal.pone.0061217)

28. Parkes KC. 1951 The genetics of the golden-winged X blue-winged warbler complex. *Wilson Bull.* **63**, 5–15.

29. Battaglia T. 2022 *btools: A suite of R functions for all types of microbial diversity analyses*. See https://github.com/twbattaglia/btools.

30. Fox J, Weisberg S. 2019 *An R companion to applied regression*, 3rd edn. Thousand Oaks, CA: Sage. See https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

31. Mallick H *et al.* 2021 Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, e1009442. (doi:10.1371/journal.pcbi.1009442)

32. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020 PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**, 685–688. (doi:10.1038/s41587-020-0548-6)

33. Langille MGI *et al.* 2013 Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821. (doi:10.1038/nbt.2676)

34. Ryu EP, Davenport ER. 2022 Host genetic determinants of the microbiome across animals: from *Caenorhabditis elegans* to cattle. *Annu. Rev. Anim. Biosci.* **10**, 203–226. (doi:10.1146/annurev-animal-020420-032054)

35. Perlman SJ, Hunter MS, Zchori-Fein E. 2006 The emerging diversity of *Rickettsia*. *Proc. R. Soc. B* **273**, 2097–2106. (doi:10.1098/rspb.2006.3541)

36. Bagwell CE *et al.* 2008 Survival in nuclear waste, extreme resistance, and potential applications gleaned from the genome sequence of *Kineococcus radiotolerans* SRS30216. *PLoS ONE* **3**, e3878. (doi:10.1371/journal.pone.0003878)

37. Saxena S, Saxena VK, Tomar S, Sapcota D, Gonmei G. 2016 Characterisation of caecum and crop microbiota of Indian indigenous chicken targeting multiple hypervariable regions within 16S rRNA gene. *Br. Poult. Sci.* **57**, 381–389. (doi:10.1080/00071668.2016.1161728)

38. Toews DPL, Taylor SA, Streby HM, Kramer GR, Lovette IJ. 2019 Selection on *VPS13A* linked to migration in a songbird. *Proc. Natl Acad. Sci. USA* **116**, 18 272–18 274. (doi:10.1073/pnas.1909186116)

39. Kim S-J, Moon J-Y, Lim J-M, Hamada M, Ahn J-H, Weon H-Y, Suzuki K, Kwon S-W. 2015 *Jatrophihabitans soli* sp. nov., isolated from soil. *Int. J. Syst. Evol. Microbiol.* **65**, 1759–1763. (doi:10.1099/ijs.0.000173)

40. Madhaiyan M, Hu CJ, Kim S-J, Weon H-Y, Kwon S-W, Ji L. 2013 *Jatrophihabitans endophyticus* gen. nov., sp. nov., an endophytic actinobacterium isolated from a surface-sterilized stem of *Jatropha curcas* L. *Int. J. Syst. Evol. Microbiol.* **63**, 1241–1248. (doi:10.1099/ijs.0.039685-0)

41. Dumonteil E, Pronovost H, Bierman EF, Sanford A, Majeau A, Moore R, Herrera C. 2020 Interactions among *Triatoma sanguisuga* blood feeding sources, gut microbiota and *Trypanosoma cruzi* diversity in southern Louisiana. *Mol. Ecol.* **29**, 3747–3761. (doi:10.1111/mec.15582)

42. Carter TD. 1944 Six years with a Brewster's warbler. *Auk* **61**, 48–61. (doi:10.2307/4079596)

43. Price TD, Bouvier MM. 2002 The evolution of F1 postzygotic incompatibilities in birds. *Evolution* **56**, 2083–2089. (doi:10.1111/j.0014-3820.2002.tb00133.x)

44. Lin D, Lacey EA, Bach BH, Bi K, Conroy CJ, Suvorov A, Bowie RCK. 2020 Gut microbial diversity across a contact zone for California voles: implications for lineage divergence of hosts and mitonuclear mismatch in the assembly of the mammalian gut microbiome. *Mol. Ecol.* **29**, 1873–1889. (doi:10.1111/mec.15443)

45. Nielsen DP, Harrison JG, Byer NW, Faske TM, Parchman TL, Simison WB, Matocq MD. 2023 The gut microbiome reflects ancestry despite dietary shifts across a hybrid zone. *Ecol. Lett.* **26**, 63–75. (doi:10.1111/ele.14135)

46. Grieneisen LE, Charpentier MJE, Alberts SC, Blekhman R, Bradburd G, Tung J, Archie EA. 2019 Genes, geology and germs: gut microbiota across a primate hybrid zone are explained by site soil properties, not host species. *Proc. R. Soc. B* **286**, 20190431. (doi:10.1098/rspb.2019.0431)

47. Song SJ *et al.* 2020 Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *mBio* **11**, 10.1128/mbio.02901-19. (doi:10.1128/mbio.02901-19)

48. Kropáčková L *et al.* 2017 Codiversification of gastrointestinal microbiota and phylogeny in passerines is not explained by ecological divergence. *Mol. Ecol.* **26**, 5292–5304. (doi:10.1111/mec.14144)

49. Loo WT, García-Loor J, Dudaniec RY, Kleindorfer S, Cavanaugh CM. 2019 Host phylogeny, diet, and habitat differentiate the gut microbiomes of Darwin's finches on Santa Cruz Island. *Scient. Rep.* **9**, 18781. (doi:10.1038/s41598-019-54869-6)

50. McGraw KJ, Beebee MD, Hill GE, Parker RS. 2003 Lutein-based plumage coloration in songbirds is a consequence of selective pigment incorporation into feathers. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **135**, 689–696. (doi:10.1016/S1096-4959(03)00164-7)

51. Marques Silva S, Xavier R, Gomes ACR, Beltrão P, Cardoso GC, Trigo S. 2023 Phenotypic associations of common waxbill gut and feather microbiome diversity in a shared environment. *Biol. J. Linn. Soc.* **141**, blad082. (doi:10.1093/biolinnean/blad082)

52. Evans WR, Fleischman DE, Calvert HE, Pyati PV, Alter GM, Rao NS. 1990 Bacteriochlorophyll and photosynthetic reaction centers in *Rhizobium* strain BTAi 1. *Appl. Environ. Microbiol.* **56**, 3445–3449. (doi:10.1128/aem.56.11.3445-3449.1990)

53. Xie F, Niu S, Lin X, Pei S, Jiang L, Tian Y, Zhang G. 2021 Description of *Microbacterium luteum* sp. nov., *Microbacterium cremeum* sp. nov., and *Microbacterium atlanticum* sp. nov., three novel C_{50} carotenoid producing bacteria. *J. Microbiol.* **59**, 886–897. (doi:10.1007/s12275-021-1186-5)

54. Zhu L, Wu X, Li O, Qian C, Gao H. 2012 Cloning and characterization of genes involved in nostoxanthin biosynthesis of *Sphingomonas elodea* ATCC 31461. *PLoS ONE* **7**, e35099. (doi:10.1371/journal.pone.0035099)

55. Fukaya Y, Takemura M, Koyanagi T, Maoka T, Shindo K, Misawa N. 2018 Structural and functional analysis of the carotenoid biosynthesis genes of a *Pseudomonas* strain isolated from the excrement of autumn darter. *Biosci. Biotechnol. Biochem.* **82**, 1043–1052. (doi:10.1080/09168451.2017.1398069)

56. Mo X-H, Zhang H, Wang T-M, Zhang C, Zhang C, Xing X-H, Yang S. 2020 Establishment of CRISPR interference in *Methylorubrum extorquens* and application of rapidly mining a new phytoene desaturase involved in carotenoid biosynthesis. *Appl. Microbiol. Biotechnol.* **104**, 4515–4532. (doi:10.1007/s00253-020-10543-w)

57. Asker D, Beppu T, Ueda K. 2007 *Sphingomonas astaxanthinifaciens* sp. nov., a novel astaxanthin-producing bacterium of the family Sphingomonadaceae isolated from Misasa, Tottori, Japan. *FEMS Microbiol. Lett.* **273**, 140–148. (doi:10.1111/j.1574-6968.2007.00760.x)

58. Liu H, Tan KS, Zhang X, Zhang H, Cheng D, Ting Y, Li S, Ma H, Zheng H. 2020 Comparison of gut microbiota between golden and brown noble scallop *Chlamys nobilis* and its association with carotenoids. *Front. Microbiol.* **11**, 480323. (doi:10.3389/fmicb.2020.00036)

59. Miki W, Otaki N, Yokoyama A, Izumida H, Shimidzu N. 1994 Okadaxanthin, a novel C_{50}-carotenoid from a bacterium, *Pseudomonas* sp. KK10206C associated with marine sponge, *Halichondria okadai*. *Experientia* **50**, 684–686. (doi:10.1007/BF01952874)

60. Sloan DB, Moran NA. 2012 Endosymbiotic bacteria as a source of carotenoids in whiteflies. *Biol. Lett.* **8**, 986–989. (doi:10.1098/rsbl.2012.0664)

61. Hill GE. 1991 Plumage coloration is a sexually selected indicator of male quality. *Nature* **350**, 337–339. (doi:10.1038/350337a0)

62. Jones JA, Tisdale AC, Bakermans MH, Larkin JL, Smalling CG, Siefferman L. 2017 Multiple plumage ornaments as signals of intrasexual communication in golden-winged warblers. *Ethology* **123**, 145–156. (doi:10.1111/eth.12581)

63. Rowe M, Veerus L, Trosvik P, Buckling A, Pizzari T. 2020 The reproductive microbiome: an emerging driver of sexual selection, sexual conflict, mating systems, and reproductive isolation. *Trends Ecol. Evol.* **35**, 220–234. (doi:10.1016/j.tree.2019.11.004)

64. Pyle P. 1997 *Identification guide to North American birds, part I. Point Reyes station*. Bolinas, CA: Slate Creek Press.

65. Escallón C, Belden LK, Moore IT. 2019 The cloacal microbiome changes with the breeding season in a wild bird. *Integr. Organ. Biol.* **1**, oby009. (doi:10.1093/iob/oby009)

66. James ARM, Geber MA, Toews DPL. 2022 Molecular assays of pollen use consistently reflect pollinator visitation patterns in a system of flowering plants. *Mol. Ecol. Resour.* **22**, 361–374. (doi:10.1111/1755-0998.13468)

67. Baiz MD, Wood AW, Toews DPL. 2024 Association between the gut microbiome and carotenoid plumage phenotype in an avian hybrid zone. GitHub repository. (https://github.com/baizm/vermivora_hz_microbiomes)

68. Baiz MD, Wood AW, Toews DPL. 2024 Association between the gut microbiome and carotenoid plumage phenotype in an avian hybrid zone. Figshare. (doi:10.6084/m9.figshare.c.7159234)