Genomic variation in the black-throated green warbler (Setophaga virens) suggests divergence in a disjunct Atlantic Coastal Plain population (S. v. waynei)

John Carpenter[1], Alexander Worm[2], Than Boves[2], Andrew Wood[3], Joseph Poston[4], and David Toews[3]

[1]North Carolina Wildlife Resources Commission
[2]Arkansas State University
[3]Pennsylvania State University
[4]Catawba College

April 16, 2024

Abstract

New World wood warblers (family: Parulidae) can exhibit strong phenotypic differences among species, particularly in song and plumage. However, within-species variation in these warblers—often designated as subspecies—is much more subtle and has led to significant debate over the origin, maintenance, and conservation status of populations that differ. A species that exhibits controversial subspecific status is the black-throated green warbler (Setophaga virens), a Neotropical-Nearctic migrant that breeds throughout eastern and boreal North America with several isolated populations at the margins of its range. In particular, uncertainty has lingered over the status of S. v. waynei, a disjunct population along the southeast Atlantic Coastal Plain of the United States that differs morphologically and ecologically from the nominate subspecies. Despite its unique circumstances, the subspecific status of S. v. waynei remains questionable in the absence of any population-wide genomic analyses. Here, we employ whole-genome resequencing to estimate the genetic distinctiveness among samples collected across the entirety of S. virens breeding range, including from putative S. v. waynei. Despite detecting low global differentiation (FST = 0.027) across the entire species, we observed discrete genetic clustering among S. v. waynei. Principal components analysis of genome-wide differences shows the main axis of variation separates S. v. waynei from all other S. v. virens samples. We also found that S. v. waynei is most similar to another isolated population from the Piedmont of North Carolina and detected evidence of a historical north-to-south geographic dispersal among the entire complex. Combined with previously documented ecological and morphological distinctness, our results support that S. v. waynei be considered a distinct and recognized subspecies worthy of targeted conservation efforts.

Genomic variation in the black-throated green warbler (*Setophaga virens*) suggests divergence in a disjunct Atlantic Coastal Plain population (*S. v. waynei*)

Running title: Divergence in the Black-throated Green Warbler

John P. Carpenter[1]*, Alexander J. Worm[1,2], Than J. Boves[2], Andrew W. Wood[3], Joseph P. Poston[4], David P.L. Toews[3]*

[1]Wildlife Diversity Program, North Carolina Wildlife Resources Commission, Raleigh, NC 27699, USA

[2]Department of Biological Sciences, Arkansas State University, Jonesboro, Arkansas 72401, USA

[3]Department of Biology, Penn State University, University Park, PA 16802, USA

[4]Department of Biology, Catawba College, Salisbury, NC 28144, USA

*Corresponding authors

John P. Carpenter:*john.carpenter@ncwildlife.org*

David P.L. Toews: *toews@psu.edu*

Abstract (about 250 words)

New World wood warblers (family: Parulidae) can exhibit strong phenotypic differences among species, particularly in song and plumage. However, within-species variation in these warblers—often designated as subspecies—is much more subtle and has led to significant debate over the origin, maintenance, and conservation status of populations that differ. A species that exhibits controversial subspecific status is the black-throated green warbler (*Setophaga virens*), a Neotropical-Nearctic migrant that breeds throughout eastern and boreal North America with several isolated populations at the margins of its range. In particular, uncertainty has lingered over the status of*S. v. waynei* , a disjunct population along the southeast Atlantic Coastal Plain of the United States that differs morphologically and ecologically from the nominate subspecies. Despite its unique circumstances, the subspecific status of *S. v. wayne* i remains questionable in the absence of any population-wide genomic analyses. Here, we employ whole-genome resequencing to estimate the genetic distinctiveness among samples collected across the entirety of *S. virens* breeding range, including from putative *S. v. waynei* . Despite detecting low global differentiation ($F_{ST} = 0.027$) across the entire species, we observed discrete genetic clustering among *S. v. waynei* . Principal components analysis of genome-wide differences shows the main axis of variation separates*S. v. waynei* from all other *S. v. virens* samples. We also found that *S. v. waynei* is most similar to another isolated population from the Piedmont of North Carolina and detected evidence of a historical north-to-south geographic dispersal among the entire complex. Combined with previously documented ecological and morphological distinctness, our results support that *S. v. waynei*be considered a distinct and recognized subspecies worthy of targeted conservation efforts.

KEYWORDS

Setophaga virens waynei , whole-genome resequencing, principal component analysis, phylogeny, next-generation sequencing, population genomics, Parulidae

INTRODUCTION

Ecotypic variation within a species often corresponds to selection pressures that give rise to local adaptation, and studies of local adaptation are important for understanding how selection may lead to evolution in natural systems. These kinds of population-level ecotypic differences have been commonly described in freshwater fishes (stream and lake stickleback, *Gasterosteus aculeatus* ; Hendry et al., 2002), coastal and inland annuals (*Mimulus guttatus* ; Lowry et al., 2008), and island finches (large- and small-billed *Geospiza fortis* ; Hendry et al., 2009). Among vagile, widely distributed, continental birds, however, local adaptation at a broad scale is less commonly described (but see Alcaide et al., 2014). Specifically, in avian taxa, geographic variants—described variably as populations, races, or subspecies depending on the magnitude of differentiation—are often expressed through plumage or song, with a less obvious connection to their environments (Toews et al., 2016; Toews, Taylor, et al., 2016). These phenotypic differences can be further obscured by overlap with respect to ecological affinities and have presented significant debate to their contribution to biodiversity and conservation science (Mayr, 1982; Zink, 2004).

Wood warblers of North, Central, and South America (Parulidae) are particularly notable among bird species for lacking strong ecological disparities: the classic study of niche partitioning by MacArthur (1958) was in fact initiated because “ecologists studying them have concluded that any differences in the species' requirements must be quite obscure”. Indeed, although MacArthur observed fine-scale habitat partitioning, it is still generally appreciated that some of the most distinct differences among warbler species occur within plumage and song (Kent et al., 2021; Price et al., 2000). Moreover, any morphological and ecogeographic

differences within most species are not discrete, but clinal in nature, such as the gradual increase in body size with increasing latitudes (Jones et al., 2003; Youtz et al., 2020).

One exception to the general pattern—where there appears to be discrete, non-clinal morphometric variation within a warbler species—occurs within the black-throated green warbler (*Setophaga virens*). This Neotropical-Nearctic migrant is a common breeder in upland, mixed-hardwood and evergreen forests of the eastern USA and much of Canada (Morse & Poole, 2020). Its occurrence is much rarer in several disjunct breeding locations scattered across the southern periphery of its breeding range (Haggerty, 2009; Mumford et. al., 1984; Rodewald, 1997), including the south Atlantic Coastal Plain where one formerly recognized subspecies, *S. virens waynei* , was first described (Bangs, 1918). Colloquially referred to as Wayne's warbler, *S. v. waynei* is uniquely associated with swamp and non-riverine wetland habitats (Sprunt, Jr., 1953; Watts et al., 2011; Worm & Carpenter, 2021) and separated by more than 400 km (and up to 1200 m elevation) from the closest conspecifics in the southern Appalachian Mountains. While plumage differences from the nominate form may be subtle, *S. v. waynei* is reported to be smaller, especially in bill dimensions, even when compared to individuals from the same latitude (Morse & Poole, 2020). Despite these reported differences in morphology and habitat use—as well as continued recognition of*S. v. waynei* as a Species of Greatest Conservation Need in three U.S. state Wildlife Action Plans (North Carolina Wildlife Resources Commission, 2015; South Carolina Department of Natural Resources, 2014; Virginia Department of Game and Inland Fisheries, 2015)—this population is now considered "non-diagnosable" (Morse & Poole, 2020), leaving the status of *S. v. waynei* as a meaningful subspecies in doubt.

Genetic diagnosability is important for conservation to appropriately delineate management units and prioritize limited resources; however, differences between *S. virens* and *S. v. waynei* as reflected in the genome remain unknown. This uncertainty may delay critical conservation action for *S. v. waynei,* which has experienced recent declines (Watts et al., 2011) likely related to loss and degradation of its breeding habitat (Davis et al., 1997; Richardson, 1983).

To address this taxonomic debate and fill a critical knowledge gap, we apply whole-genome resequencing to quantify, for the first time, genetic difference across the *S. virens* complex, including samples obtained from putative *S. v. waynei* individuals. Our goals are to determine: 1) whether *S. v. waynei* cluster independently from other *S. virens* populations, 2) if there are genetic differences among the groups, including the extent and distribution of the divergence (i.e., is differentiation clustered into regions of the genome), and 3) approximately when and where *S. v. waynei*separated from the nominate form. This information is essential to better understand the unique evolutionary history of the species (Bermingham et al., 1992) and to increase our ability to conserve and manage a putative subspecies of concern within the complex.

MATERIALS AND METHODS

Study sites and data collection

We selected ten populations from across *S. virens* breeding distribution to capture range-wide genomic diversity (Figure 1). During their breeding season (April-June), we used mist nets with audio playback of conspecific song to capture *S. v. virens* males in Arkansas, Indiana, New York, central North Carolina, and Tennessee, as well as putative *S. v. waynei* males in coastal North Carolina (Table 1). For all individuals, we collected [?]50 µL of blood from the brachial vein, which was stored in either Queen's Lysis buffer or ethanol (Seutin et al., 1991). In addition, two *S. v. virens*tissue samples from New York were donated by The Cornell University Museum of Vertebrates.

Genetic Sample Preparation

Our final dataset (n = 29) included five *S. v. virens*samples from New York that were sequenced in a previous study (Baiz, Wood, Brelsford, Lovette, & Toews, 2021). These samples were prepared with an identical method but sequenced separately from the new samples included in the current study (n = 24). For phylogenetic analyses we also included a single golden-cheeked warbler (*Setophaga chrysoparia*) as an outgroup, also published in Baiz et al., (2021). Genomic DNA was extracted from blood using the

Qiagen DNeasy Blood and Tissue kit following the nucleated blood spin-column protocol with these minor modifications: we used 75 µL of blood and the 56°C incubation was performed overnight. All samples were normalized to 2ng/µL and DNA was sheared with a Covaris S220 to a target insert size of 350bp, according to the TruSeq Nano protocol. We then prepared sequencing libraries using the Illumina TruSeq Nano DNA kit and submitted them to the Pennsylvania State University Genomics Core Facility for sequencing (150nt, paired-end) on a single NextSeq High Output lane.

Whole-Genome Resequencing Analysis

We removed adapter sequences and quality trimmed reads using AdapterRemoval v2.1.7 (Schubert et al., 2016). We then aligned reads to a new chromosome-level assembly of a yellow-rumped warbler (*Setophaga coronata*) (Baiz et al., 2021) using Bowtie2 (Langmead & Salzberg, 2012). PCR duplicates were marked with Picard tools (Broad Institute, 2021).

We analyzed the resultant assemblies using the ANGSD bioinformatics pipeline (Korneliussen et al., 2014), the most appropriate method for low-coverage data. To measure the extent of differentiation, we calculated the global estimation of F_{ST} between the main range *S. v. virens* and *S. v. waynei* . We then generated a windowed estimate of F_{ST} in 10kb windows across the genome, comparing *S. v. virens* and *S. v. waynei* , thus quantifying whether different parts of the genome are more divergent than other regions of the genome between the groups.

To identify genes within the divergent regions, we used the annotation information associated with the *S. coronata* genome (Baiz et al., 2021), which used SNP gene predictions, trained on the Zebra Finch (*Taeniopygia guttata*) genome, within the MAKER annotation pipeline. We focused on coding sequences where MAKER had either transcript or protein matches with the Zebra Finch annotation. We first identified genes within the two regions (see below) that were bounded by 10Kb F_{ST} windows >0.15. We also focused on two genes intersecting the two highestF_{ST} windows.

We next performed Principal Components Analysis (PCA) to determine the genome-wide signal of clustering. We used PCAngsd (Meisner & Albrechtsen, 2018) to generate a covariance matrix from genome-wide genotype likelihoods, and R 4.0.5 (R Core Team, 2021) to calculate and plot eigenvalues. Lastly, we constructed a bootstrapped phylogeny to estimate relationships of the various populations and to test roughly when *S. v. waynei* separated from the main group. Genotype likelihoods generated by ANGSD for 31,241,801 sites were analyzed using ngsDist v1.0.8 (Vieira et al., 2016) and run with 100 bootstrap replicates. We produced trees from the resultant distance matrices using FastME v2.1.6.2 (Lefort et al., 2015) and combined support values on the main tree using RAxML v8.2.12 (Stamatakis, 2014).

RESULTS

We generated genomic data for 9 individuals breeding in coastal North Carolina (putative *S. v. waynei)* , and 20 individuals from the nominate subspecies, *S. v. virens* , which were distributed across their breeding range. Genome-wide F_{ST} indicated low (0.027) overall differentiation across the *S. virens* complex. However, principal component analysis revealed three highly clustered groups, one of which consisted of only *S. v. waynei* (Figure 2). Samples from coastal North Carolina (*S. v. waynei)* were distinct along PC1, which explained 8% of the total variance in the data, and strongly separated from *S. v. virens* . Samples of *S. v. virens* from the Uwharrie National Forest in central North Carolina separated along PC2 (4.1% variance) with the remaining group composed of *S. v. virens* samples from Arkansas, Tennessee, Indiana, and New York (Figure 2).

Differentiation was not homogenous across the genome (Figure 3A). We identified two regions of the genome, one on chromosome 6 (Figure 3B) and one on the sex (Z) chromosome (Figure 3C) that showed elevated differentiation above the background. On chromosome 6, the elevated region spanned 16,950,000bp and 17,040,000bp (a length of 80,000bp) and had six windows with F_{ST} above 0.15 (the maximum windowed F_{ST} in this region was 0.28). The region on the Z chromosome, between 11,910,000bp and 12,150,000bp (a length of 240,000bp) contained 10 windows above F_{ST} 0.15 (the maximum windowed F_{ST} in this region

was 0.32). The chromosome 6 region included five annotated genes that were bounded by the region of elevated divergence; the Z chromosome region had four annotated genes, and one gene of unknown function (Figure 4).

Phylogenetic relatedness among the samples clustered most samples by geographic locality (Figure 5). The initial split within the species occurred between samples from New York and other samples from southern populations. Samples from putative *S. v. waynei* formed a monophyletic group, which was sister to birds breeding in the Uwharrie National Forest of central North Carolina.

DISCUSSION

In concordance with previously described habitat (Bangs, 1918) and morphological (Sprunt, Jr., 1953) differences, we provide strong evidence that black-throated green warblers breeding along the southern Atlantic Coastal Plain of the United States are genomically distinct from the nominate subspecies found throughout much of the forested portions of eastern North America. As such, this isolated population is worthy of classification as a true subspecies and thus deserves conservation and management efforts required to sustain its persistence.

Patterns of Genomic Divergence

Differentiation between *S. v. virens* and *S. v. waynei* was most pronounced in two regions of the genome, one along chromosome 6 and the other on the Z chromosome (Figure 3). Baiz et al., (2021) compared patterns of differentiation between nine sister species across the*Setophaga* genus. Notably, these regions on chromosome 6 and the Z are also commonly differentiated between several other warbler pairs. However, in those instances, regions of divergence are much larger, on the scale of 1-2 Mb, whereas in the present case the divergent regions are an order of magnitude smaller. We suggest this is likely because all pairs in Baiz et al., (2021) are between fully reproductively isolated species, each sharing a common ancestor much older than that between*S. v. virens* and *S. v. waynei* . Thus, divergence between those pairs have had more time to accrue greater differences in their genomes. Additionally, this study adds to the number of examples in birds where the Z chromosome is highlighted as highly divergent between related taxa, beyond expectations of a reduced effective population size (Ellegren, 2011; Oyler-McCance et al., 2015; Ruegg et al., 2014).

Because the regions of divergence between *S. v. virens* and*S. v. waynei* are small, we can identify individual genes within them that were potentially the target of divergent selection between the subspecies. Although the two regions—on chromosome 6 and the Z chromosome—include multiple genes, we focused on two that directly intersected the two most divergent 10Kb windows. In the chromosome 6 region, the gene that intersects the most divergent window is broad substrate specificity ATP-binding cassette transporter ABCG2 isoform X1 (*ABCG2;* NCBI reference XP_030131844.2). *ABCG2* is involved in protein transport, extra- and intra-cellularly (Ma et al., 2020), and recently implicated in egg coloration differences in mallard ducks (*Anas platyrhynchos;* Liu et al., 2021). The most divergent window within the entire genome, on the Z chromosome, falls within the gene Metallo-Beta-Lactamase Domain Containing 2 (*MBLAC2*). The MBL superfamily of enzymes has a diverse phenotypic role but has been associated with hydrolase activity (Malgapo et al., 2021). While we do not know the function of these genes in warblers, they are both involved in metabolic processes and cellular transport and present a possible connection to the distinct habitat (and therefore diet) these warblers inhabit. Both are good candidates for further study of the implications of gene functionality in birds.

More broadly, the overall limited genomic differentiation suggests these warblers are in the early stages of divergence. Previous studies of closely related avian taxa have commonly identified shared divergence peaks between independent groups (Delmore et al., 2015; Toews, Campagna, et al., 2016). These have generally been understood as genomic regions of reduced recombination, which amplify the influence of divergence via linked selection. However, given the large size and many genes that fall within these regions in previous studies, pinpointing putative targets of selection has been challenging. In the *S. v. virens* complex, however, the small size of these divergent regions has allowed us to identify *ABCG2* and *MBLAC2* as two possible targets, and it will be important to understand whether these specific genes have contributed to common

divergence patterns in other species pairs.

Population Structure

Although the Atlantic coast population was most distinct in our dataset, the nested structure of *S. v. waynei* within the clades in the south suggests a relatively recent origin of this group. Our results also suggest that *S. v. waynei* are most closely related to other southern populations (Figure 5), especially to *S. v. virens*breeding <150 km away in the Uwharrie National Forest. This pattern is consistent with post-glacial geographic dispersal from north to south. However, the Uwharrie population does not appear to be directly intermediate between the coastal populations and the inland group—as would be expected if it was intergrades of the two subspecies—but instead shows evidence of independent evolution (e.g., divergence along PC2 in Figure 2).

Despite substantial geographic isolation from the contiguous *S. v. virens* breeding distribution (e.g., the Arkansas population is separated by >900 km), the fact that the other isolated populations did not cluster or provide as clear a signal of isolation-by-distance as *S. v. waynei,* could be explained in several ways. First, it is possible that these locations have been only recently colonized, thus not providing enough time for genetic differentiation to occur. This is likely the case in Arkansas, where*S. v. virens* was only first documented breeding in the mid-1990s (Rodewald, 1997). Additionally, unlike the ecological conditions of the southern Atlantic Coastal Plain, the habitats in these other disjunct areas are not as clearly different from the rest of *S. v. virens'*breeding range—at least south of the Boreal Forest—so selective pressures may not have promoted rapid differentiation. Finally, despite rather extensive geographic separation, it is possible that isolation in these areas is incomplete, with gene flow occurring occasionally when individuals forgo the rest of their migratory journey (natal or breeding dispersal from the contiguous breeding distribution) and choose to breed at these locations. In support of this possibility, most birds sampled in Arkansas were young male (second-year or first-time breeding) birds—the age at which dispersal distance for many birds is greatest (Greenwood & Harvey, 1982).

From a conservation perspective, the genetic diagnosability of *S. v. waynei* supports its subspecific taxonomic status, and likely the validity of a unique evolutionary history as a population that exhibits fixed genetic differences. Thus, it would also be appropriate to manage*S. v. waynei* as an independent management unit for conservation actions to maintain the unique diversity that this subspecies exhibits. Our use of low-coverage sequencing and genotype likelihoods (within the ANGSD pipeline) makes it challenging to estimate per-SNP F_{ST} values. However, using successively smaller genomic windows (Figure S1) we show that many dozens of SNPs within the Z chromosome region have F_{ST} values >0.7, and within the chromosome 6 region at least one SNP with F_{ST}>0.7. Thus, these nearly fixed genetic differences are consistent with *S. v. waynei* having a unique evolutionary history worthy of conservation even though genome-wide divergence is low overall.

Acknowledgements

Funding provided to the N.C. Wildlife Resources Commission through the Pittman-Robertson Federal Aid Wildlife Restoration Grant, J.P. Poston from the Department of Environment and Sustainability at Catawba College, and T.J. Boves from the College of Sciences and Mathematics at Arkansas State University. D.P.L.T. was supported by Pennsylvania State University and startup funds from the Eberly College of Science and the Huck Institutes of the Life Sciences. We thank several partner agencies for permitting access to and data collection on their properties, including the U.S. Fish and Wildlife Service, U.S. Forest Service, N.C. Division of Parks and Recreation, N.C. Forest Service, and The Nature Conservancy. We also thank Alix E. Matthews, Emily R. Donahue, Macayla Upright, Jacqueline Poston, Scott Poston, John Poston, Grace Poston, and Benjamin Nickley for their help collecting field samples. All birds were captured and handled under the auspices of U.S. Geological Survey federal banding permits #23877 and #06557, as well as IACUC protocols approved by Arkansas State University (Protocol #FY17-18-319).

DATA ACCESSIBILITY

All sequencing and genome information will be uploaded to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under Bioproject #PRJNA630247.

AUTHOR CONTRIBUTIONS

J.P.C., A.J.W., D.P.L.T., T.J.B., and J.P.P. conceived the study and performed field work. D.P.L.T. and A.W.W. processed genetic samples and conducted statistical analyses; J.P.C., A.J.W., D.P.L.T., T.J.B., A.W.W., and J.P.P. wrote and reviewed the manuscript.

REFERENCES

Alcaide, M., Scordato, E. S. C., Price, T. D., & Irwin, D. E. (2014). Genomic divergence in a ring species complex. *Nature* ,*511* (7507), 83–85. https://doi.org/10.1038/nature13285

Baiz, M. D., Wood, A. W., Brelsford, A., Lovette, I. J., & Toews, D. P. L. (2021). Pigmentation Genes Show Evidence of Repeated Divergence and Multiple Bouts of Introgression in *Setophaga* Warblers.*Current Biology* , *31* (3), 643–649. https://doi.org/10.1016/j.cub.2020.10.094

Bangs, O. (1918). A new race of the black-throated green wood warbler.*Proceedings of the New England Zoological Club* , *6* , 93–94.

Bermingham, E., Rohwer, S., Freeman, S., & Wood, C. (1992). Vicariance biogeography in the Pleistocene and speciation in North American wood warblers: a test of Mengel's model. *Proceedings of the National Academy of Sciences* , *89* (14), 6624–6628.

Broad Institute. (2021). Picard: A set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data and formats such as SAM/BAM/CRAM and VCF. Retrieved from http://broadinstitute.github.io/picard/

Davis, K. N., Henderson, B., & Daniels, S. (1997). Inventory of Atlantic white cedar remnant stands in North Carolina. *U.S. Air Force BPA No. F31610-95-AV026* .

Delmore, K. E., Hübner, S., Kane, N. C., Schuster, R., Andrew, R. L., Câmara, F., Guigó, R., & Irwin, D.E. (2015). Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation.*Molecular Ecology* , *24* (8), 1873–1888. https://doi.org/10.1111/mec.13150

Ellegren, H. (2011). Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nature Reviews Genetics* , *12* (3), 157–166. https://doi.org/10.1038/nrg2948

Greenwood, P. J., & Harvey, P. H. (1982). The natal and breeding dispersal of birds. *Annual Review of Ecology and Systematics* ,*13* (1), 1–21. https://doi.org/10.1146/annurev.es.13.110182.000245

Haggerty, T. M. (2009). *Alabama Breeding Bird Atlas 2000-2006* . Alabama Ornithological Society. Retrieved from https://conservationgis.alabama.gov/ALBBA/BBA

Hendry, A. P., Huber, S. K., De Leon, L. F., Herrel, A., & Podos, J. (2009). Disruptive selection in a bimodal population of Darwin's finches. *Proceedings of the Royal Society B: Biological Sciences* ,*276* (1657), 753–759. https://doi.org/10.1098/rspb.2008.1321

Hendry, A. P., Taylor, E. B., & McPhail, J. D. (2002). Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution* , *56* (6), 1199–1216. https://doi.org/10.1111/j.0014-3820.2002.tb01432.x

Jones, J., Gibb, C. E., Millard, S. C., Barg, J. J., Girvan, K., Veit, M. L., ... Robertson, R. J. (2003). Multiple Selection Pressures Generate Adherence to Bergmann's Rule in a Neotropical Migratory Songbird. *Journal of Biogeography* , *32* (10), 1827–1833. https://doi.org/10.1111/j.1365-2699.2005.01320.x

Kent, C. M., Huh, K. M., Hunter, S. C., Judson, K., Powell, L. L., & Sherry, T. W. (2021). High resource overlap and small dietary differences are widespread in food-limited warbler (Parulidae) communities. *Ibis* . https://doi.org/10.1111/ibi.13006

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* ,*15* (1), 1–13. https://doi.org/10.1186/s12859-014-0356-4

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* , *9* (4), 357–359. https://doi.org/10.1038/nmeth.1923

Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution* , *32* (10), 2798–2800. https://doi.org/10.1093/molbev/msv150

Liu, H., Hu, J., Guo, Z., Fan, W., Xu, Y., Liang, S., Liu, D., Zhang, Y., Xie, M., Tang, J., Huang, W., Zhang, Q., Xi, Y., Li, Y., Wang, L., Ma, S., Jiang, Y., Feng, Y., Wu, Y., Cao, J., Zhou, Z. & Hou, S. (2021). A single nucleotide polymorphism variant located in the cis-regulatory region of the ABCG2 gene is associated with mallard egg colour. *Molecular Ecology* , 30: 1477-1491. https://doi.org/10.1111/mec.15785

Lowry, D. B., Rockwood, R. C., & Willis, J. H. (2008). Ecological reproductive isolation of coast and inland races of *Mimulus guttatus* . *Evolution: International Journal of Organic Evolution* ,*62* (9), 2196–2214. https://doi.org/10.1111/j.1558-5646.2008.00457.x

Ma, S., Liu, H., Sun, W., Mustafa, A., Xi, Y., Pu, F., Li, Y., Han, C., Bai, L., & Hua, H. (2020). Molecular evolution of the ATP-binding cassette subfamily G member 2 gene subfamily and its paralogs in birds.*BMC Evolutionary Biology* , *20* (85). https://doi.org/10.1186/s12862-020-01654-z

MacArthur, R. H. (1958). Population Ecology of Some Warblers of Northeastern Coniferous Forests. *Ecology* , *39* (4), 599–619. https://doi.org/10.2307/1931600

Malgapo M.I.P., Safadi J.M., & Linder, M.E. (2021). Metallo-β-lactamase domain-containing protein 2 is S-palmitoylated and exhibits acyl-CoA hydrolase activity. *Journal of Biological Chemistry* ,*296* :100106. https://doi.org/10.1074/jbc.RA120.015701

Mayr, E. (1982). Of What Use Are Subspecies? *The Auk* ,*99* (3), 593–595. Retrieved from https://www.jstor.org/stable/4085949

Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics* ,*210* (2), 719–731. https://doi.org/10.1534/genetics.118.301336

Morse, D. H., & Poole, A. F. (2020). Black-throated Green Warbler (Setophaga virens), version 2.0 (A. F. Poole, Ed.). Retrieved August 18, 2021, from Birds of North America website: https://doi.org/10.2173/bna.55

Mumford, R. E., Keller, C. E., & Zimmerman, W. (1984). *The birds of Indiana* . Bloomington, IN: Indiana University Press.

North Carolina Wildlife Resources Commission. (2015). *North Carolina Wildlife Action Plan* . Raleigh, NC.

Oyler-McCance, S. J., Cornman, R. S., Jones, K. L., & Fike, J. A. (2015). Z chromosome divergence, polymorphism and relative effective population size in a genus of lekking birds. *Heredity* ,*115* (5), 452–459. https://doi.org/10.1038/hdy.2015.46

Price, T., Lovette, I. J., Bermingham, E., Gibbs, H. L., & Richman, A. D. (2000). The imprint of history on communities of North American and Asian warblers. *American Naturalist* , *156* (4), 354–367. https://doi.org/10.1086/303397

R Core Team. (2021). *R A language and environment for statistical computing* . Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Richardson, C. J. (1983). Pocosins: Vanishing Wastelands or Valuable Wetlands? *BioScience* , *33* (10), 626–633. https://doi.org/10.2307/1309491

Rodewald, P. G. (1997). Two new Breeding Species of Wood-Warblers (Parulinae) in Arkansas. *The Southwestern Naturalist* ,*42* (1), 106–107.

Ruegg, K., Anderson, E. C., Boone, J., Pouls, J., & Smith, T. B. (2014). A role for migration-linked genes and genomic islands in divergence of a songbird. *Molecular Ecology* , *23* (19), 4757–4769. https://doi.org/10.1111/mec.12842

Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes* , *9* (1), 1–7. https://doi.org/10.1186/s13104-016-1900-2

Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* , *69* (1), 82–90. https://doi.org/10.1139/z91-013

South Carolina Department of Natural Resources. (2014). *South Carolina's State Wildlife Action Plan (SWAP) 2015* . Columbia, SC.

Sprunt, Jr., A. (1953). Wayne's Black-throated Green Warbler. In A. C. Bent (Ed.), *Bulletin 203 of the United States National Museum*(pp. 308–316). Washington, D.C.: U.S. Government Printing Office.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* ,*30* (9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Toews, D. P. L., Brelsford, A., Grossen, C., Milá, B., & Irwin, D. E. (2016). Genomic variation across the Yellow-rumped Warbler species complex. *The Auk* , *133* (4), 698–717. 10.1642/auk-16-61.1

Toews, D. P. L., Campagna, L., Taylor, S. A., Balakrishnan, C. N., Baldassarre, D. T., Deane-Coe, P. E., ... Winger, B. M. (2016). Genomic approaches to understanding population divergence and speciation in birds. *The Auk* , *133* (1), 13–30. 10.1642/auk-15-51.1

Toews, D. P. L., Taylor, S. A., Vallender, R., Brelsford, A., Butcher, B. G., Messer, P. W., & Lovette, I. J. (2016). Plumage Genes and Little Else Distinguish the Genomes of Hybridizing Warblers. *Current Biology* , *26* (17), 2313–2318. https://doi.org/10.1016/j.cub.2016.06.034

Vieira, F. G., Lassalle, F., Korneliussen, T. S., & Fumagalli, M. (2016). Improving the estimation of genetic distances from Next-Generation Sequencing data. *Biological Journal of the Linnean Society* , *117* (1), 139–149. https://doi.org/10.1111/bij.12511

Virginia Department of Game and Inland Fisheries. (2015). *Virginia's 2015 Wildlife Action Plan* . Henrico, VA.

Watts, B. D., Paxton, B. J., & Smith, F. M. (2011). Status and Habitat Use of the Wayne's Black-throated Green Warbler in the Northern Portion of the South Atlantic Coastal Plain. *Southeastern Naturalist* ,*10* (2), 333–344. https://doi.org/10.1656/058.010.0212

Worm, A. J., & Carpenter, J. P. (2021). Description of Black-throated Green Warbler (Wayne's) (*Setophaga virens waynei*) Nests in Coastal North Carolina. *The Chat* , *85* (2), 41–44.

Youtz, J., Miller, K. D., Bowers, E. K., Rogers, S. L., Bulluck, L. P., Johnson, M., ... Ames, E. M. (2020). Bergmann's rule is followed at multiple stages of postembryonic development in a long-distance migratory songbird. *Ecology and Evolution* , *10* (19), 10672–10686. https://doi.org/10.1002/ece3.6721

Zink, R. M. (2004). The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society B: Biological Sciences* , *271* (1539), 561–564. https://doi.org/10.1098/rspb.2003.2617

Tables and Figures (with captions)

TABLE 1 Study sites and dates of sampling. All captured black-throated green warblers (*S. virens*) were male, except where indicated. NWR = National Wildlife Refuge, WMA = Wildlife Management Area. [a] Samples sequenced previously (Baiz et al., 2021), [b] Frozen sample, [c] Female

State, Location	State, Location	Date	ID	Latitude	Longitude
Arkansas					
Ozark National Forest	Ozark National Forest	25 April 2020	290005180	35.726	-93.810
Ozark National Forest	Ozark National Forest	25 April 2020	290005185	35.726	-93.809
Ozark National Forest	Ozark National Forest	26 April 2020	290005186	35.727	-93.809
Ozark National Forest	Ozark National Forest	13 June 2020	9	35.669	-93.894
Indiana	Indiana				
Hoosier National Forest	Hoosier National Forest	27 May 2020	290005282	39.033	-86.342
Hoosier National Forest	Hoosier National Forest	27 May 2020	290005284	39.036	-86.312
Hoosier National Forest	Hoosier National Forest	28 May 2020	290005286	39.032	-86.342
Hoosier National Forest	Hoosier National Forest	28 May 2020	290005285	39.032	-86.344
New York	New York				
Adirondack Park[a]	Adirondack Park[a]	18 May 2013	CUMV4068	43.934	-74.837
Adirondack Park[a]	Adirondack Park[a]	7 June 2018	284029390	43.716	-74.754
Adirondack Park[a]	Adirondack Park[a]	21 June 2018	284029440	43.736	-74.512
Adirondack Park[ab]	Adirondack Park[ab]	7 June 2018	284029389	43.714	-74.756
Lake Placid[abc]	Lake Placid[abc]	5 June 2015	CUMV4914	44.248	-73.985
North Carolina	North Carolina				
Alligator River NWR	Alligator River NWR	26 May 2020	201010926	35.780	-75.821
Alligator River NWR	Alligator River NWR	26 May 2020	201010923	35.768	-75.798
Bladen Lakes State Forest	Bladen Lakes State Forest	20 March 2020	201010941	34.748	-78.503
Bladen Lakes State Forest	Bladen Lakes State Forest	22 April 2020	201010934	34.747	-78.502
Bladen Lakes State Forest	Bladen Lakes State Forest	22 April 2020	201010933	34.712	-78.524
Bladen Lakes State Forest	Bladen Lakes State Forest	1 May 2020	201010931	34.725	-78.544
Croatan National Forest	Croatan National Forest	4 May 2020	215063530	34.770	-76.971

Croatan National Forest	Croatan National Forest	11 May 2020	215063531	34.770	-76.970
Green Swamp Preserve	Green Swamp Preserve	17 April 2020	201010939	34.103	-78.307
Uwharrie National Forest	Uwharrie National Forest	1 June 2018	264043090	35.408	-80.045
Uwharrie National Forest	Uwharrie National Forest	6 June 2019	264042908	35.408	-80.041
Uwharrie National Forest	Uwharrie National Forest	10 May 2020	264042920	35.414	-80.035
Tennessee					
Royal Blue WMA	Royal Blue WMA	13 May 2020	290005238	36.314	-84.244
Royal Blue WMA	Royal Blue WMA	14 May 2020	23	36.315	-84.259
Royal Blue WMA	Royal Blue WMA	14 May 2020	290005271	36.316	-84.258
Royal Blue WMA	Royal Blue WMA	15 May 2020	290005272	36.356	-84.282

FIGURE 1 Breeding range (darker shaded area) of black-throated green warbler (*Setophaga virens virens*) and putative subspecies, Wayne's warbler (*S. v. waynei*). Study sites are indicated with black symbols where *S. v. virens* were sampled and yellow symbols where *S. v. waynei* were sampled. NF = National Forest, WMA = Wildlife Management Area, SF = State Forest.

FIGURE 2 Principal components analysis (PCA) differentiating populations across the *Setophaga virens* North American breeding range. All populations are *S. v. virens* except for those individuals from North Carolina (NC) identified as *S. v. waynei* .

FIGURE 3 Windowed F_{ST} estimates (10kb) comparing the genomes of *Setophaga virens virens and S. v. waynei*(A), including two regions—one on chromosome 6 (B) and one on the Z (sex) chromosome (C)—with elevated differentiation above the background.

FIGURE 4 Detail view of regions of elevated divergence comparing the genomes of *Setophaga virens virens and S. v. waynei.* Grey points illustrate F_{ST} estimates for 10kb windows. Variable colors indicate different annotated genes in these regions on (a) chromosome 6 (five annotated genes) and (b) the Z chromosome (four annotated genes and one gene of unknown function).

FIGURE 5 Bootstrap phylogeny from ngsDist based on genome-wide genotype likelihoods arcoss all *Setophaga virens* samples and one outgroup (*S. chrysoparia*).

FIGURE S1 Detail view of narrowed windowedF_{ST} estimates (2 bp) comparing the genomes of*Setophaga virens virens and S. v. waynei* for two regions (A) on chromosome 6 (B) the Z chromosome with elevated differentiation above the background. The red line indicates F_{ST} of 0.7.