

ORIGINAL ARTICLE

Allele-specific *cis*-regulatory methylation of the gene for vasoactive intestinal peptide in white-throated sparrows

Mackenzie R. Prichard  | Kathleen E. Grogan  | Jennifer R. Merritt  |
Jessica Root  | Donna L. Maney 

Department of Psychology, Emory University,
Atlanta, Georgia, USA

Correspondence

Mackenzie R. Prichard, Department of
Psychology, Emory University, 36 Eagle Row,
Atlanta, Georgia, USA.
Email: mackenzie.prichard@emory.edu

Present address

Kathleen E. Grogan, Departments of
Anthropology and Biology, University of
Cincinnati, Cincinnati, Ohio, USA.

Jennifer R. Merritt, Zuckerman Mind Brain
Behavior Institute and Department of Ecology,
Evolution and Environmental Biology,
Columbia University, New York, New
York, USA.

Jessica Root, Department of Pharmacology
and Chemical Biology, Emory University,
Atlanta, Georgia, USA.

Funding information

National Institutes of Health, Grant/Award
Numbers: F31MH114509, R01MH082833

Abstract

White-throated sparrows (*Zonotrichia albicollis*) offer a unique opportunity to connect genotype with behavioral phenotype. In this species, a rearrangement of the second chromosome is linked with territorial aggression; birds with a copy of this “supergene” rearrangement are more aggressive than those without it. The supergene has captured the gene *VIP*, which encodes vasoactive intestinal peptide, a neuromodulator that drives aggression in other songbirds. In white-throated sparrows, *VIP* expression is higher in the anterior hypothalamus of birds with the supergene than those without it, and expression of *VIP* in this region predicts the level of territorial aggression regardless of genotype. Here, we aimed to identify epigenetic mechanisms that could contribute to differential expression of *VIP* both in breeding adults, which exhibit morph differences in territorial aggression, and in nestlings, before territorial behavior develops. We extracted and bisulfite-converted DNA from samples of the hypothalamus in wild-caught adults and nestlings and used high-throughput sequencing to measure DNA methylation of a region upstream of the *VIP* start site. We found that the allele inside the supergene was less methylated than the alternative allele in both adults and nestlings. The differential methylation was attributed primarily to CpG sites that were shared between the alleles, not to polymorphic sites, which suggests that epigenetic regulation is occurring independently of the genetic differentiation within the supergene. This work represents an initial step toward understanding how epigenetic differentiation inside chromosomal inversions leads to the development of alternative behavioral phenotypes.

KEYWORDS

alternative phenotypes, behavioral polymorphism, bisulfite sequencing, chromosomal inversions, epigenetic regulation, hypothalamus, supergene

1 | INTRODUCTION

The white-throated sparrow (*Zonotrichia albicollis*) is a newly emerging model of the evolution of supergenes and their effects on social

behavior.¹ The species occurs in two plumage morphs (Figure 1), white-striped (WS) and tan-striped (TS), each of which make up about 50% of any population. WS birds of both sexes are heterozygous for a supergene, referred to as ZAL2^m, which comprises a series of nested

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Genes, Brain and Behavior published by International Behavioural and Neural Genetics Society and John Wiley & Sons Ltd.



FIGURE 1 The white-striped (WS) and tan-striped (TS) morphs in white-throated sparrows. The WS morph (on right, pictured perched on a branch) is distinguished from the TS morph (inset, left) by the bold black and white striping pattern on the crown and the larger white throat patch. Main photo credit M. R. Prichard, left photo credit B. M. Horton

inversions. TS birds are homozygous for the standard arrangement, ZAL2. The supergene is linked to well-documented variation in social behaviors such as aggression and parental provisioning.^{2–4} WS birds of both sexes defend territories more aggressively than TS birds, whereas TS birds, particularly males, provision nestlings more frequently than do WS birds.^{2,4}

As a result of an unusual disassortative mating system in which WS-WS breeding pairs are highly uncommon,⁵ ZAL2^m is in a near-constant state of heterozygosity. Recombination between the ZAL2 and ZAL2^m arrangements is therefore suppressed, leading to linkage disequilibrium and genetic divergence inside the supergene.^{8–10} This divergence has affected a number of genes known to mediate both aggression and parental behavior. To date, however, only one gene has been shown definitively to contribute causally to the behavioral phenotype. *ESR1*, which encodes estrogen receptor alpha, is over-expressed in the ventromedial arcopallium in WS birds relative to TS birds; Merritt et al.¹¹ showed experimentally that this over-expression is necessary for the morph difference in aggression. We believe that the behavioral polymorphism in this species is highly unlikely to be caused by a single gene, however, and other candidates must be considered.

A near neighbor of *ESR1*, the gene that encodes vasoactive intestinal peptide (*VIP*), is such a candidate. This gene is both highly evolutionarily conserved and associated with social behavior across vertebrates, including many birds (*Gallus gallus*,¹²; *Cyanistes coeruleus*,¹³; *Spizella pusilla*, *Melospiza melodia*^{14,15}). The peptide is expressed in many brain regions thought to be related to social behavior, including those that make up a social behavior network.^{16,17} One of these regions is the anterior hypothalamus (AH), which contains a large population of *VIP*-expressing neurons in songbirds.¹⁸ In field and song sparrows, *VIP* immunoreactivity in the AH correlates with aggressive behaviors.^{14,15} Further, Goodson et al.¹⁹ showed a causal effect of *VIP* expression on aggressive behavior; knockdown of *VIP*

expression in AH reduced aggression in violet-eared waxbills (*Uraeginthus granatina*) and zebra finches (*Taeniopygia guttata*). *VIP* knockdown did not affect other behaviors, such as affiliative or anxiety-like behaviors,¹⁹ providing evidence that *VIP* expression in the AH is causal for aggressive behavior specifically.

This evidence from other species strongly suggests that *VIP* could represent a significant contributor to the behavioral polymorphism in white-throated sparrows. Further supporting this view, we previously showed that *VIP* is expressed at higher levels in the AH of WS than TS birds.²⁰ Further, *VIP* expression in the AH predicted rates of aggressive singing in both morphs and sexes.²⁰ We hypothesize that these morph differences in expression are driven by genetic variation between the ZAL2 and ZAL2^m alleles of this gene. Although the coding regions do not differ between the ZAL2 allele of *VIP* (*VIP*²) and the ZAL2^m allele (*VIP*^{2m}),⁸ there are many single nucleotide polymorphisms (SNPs) and insertion/deletion polymorphisms upstream of the transcription start site (TSS)⁸ (Figure S1). Thus, the genetic variation in this putative cis-regulatory element (hereafter called the *VIP* CRE) could cause the observed morph differences in *VIP* expression.

Differential regulation of the *VIP*² and *VIP*^{2m} alleles could be caused not only by genetic variation, but also by epigenetic variation in the same regulatory region. Here, we focus specifically on variation in the methylation of CG dinucleotides (CpGs). DNA methylation affects transcription factor binding²¹ and is vital for tissue-specific regulation of gene expression.^{22,23} DNA methylation in CREs is most commonly, but not always, associated with reduced expression of the associated gene.^{24,25} Using whole-genome bisulfite sequencing in white-throated sparrows, Sun et al.²⁶ found that ZAL2^m was, overall, less methylated than ZAL2. They reported further that, throughout the rearranged region, the degree of differential methylation between alleles was negatively associated with the degree of differential expression of those alleles. Therefore, we expected that the known higher *VIP* expression in WS birds²⁰ might be explained by relative hypomethylation of the *VIP*^{2m} CRE which could result in overall higher levels of expression of *VIP* in WS birds.

In this study, we used next-generation sequencing to answer three central questions about the methylation of *VIP* in samples of white-throated sparrow hypothalamus. First, are the two alleles of the *VIP* CRE differentially methylated? Because *VIP* expression in the AH is higher in WS than TS birds,²⁰ we predicted that the *VIP*^{2m} CRE would be less methylated than the *VIP*² CRE in DNA extracted from this region. Such differences in DNA methylation could result from epigenetic variation alone, such that the same CpGs are methylated to different degrees on *VIP*² and *VIP*^{2m}. This expectation is supported by the findings of Sun et al.,⁸ who reported hypomethylation across the ZAL2^m rearrangement even when considering only CpG sites that were not disrupted by polymorphisms. Alternatively, because fixed differences in the *VIP* CRE create/disrupt CpGs, differential methylation could also be driven by genetic variation (Figure S1). Thus, we were interested in the extent to which DNA methylation at shared versus polymorphic sites contributes to potential differential methylation of the region.

Second, does methylation of this CRE predict the expression of *VIP* in AH? To answer this question, we tested whether the degree of allele-specific methylation of the *VIP* CRE, either regionally or at specific individual CpGs, predicts the degree of expression of the corresponding allele in the same hypothalamic tissue samples.

Third, because the morph difference in territorial aggression occurs only in breeding adults, we also tested our hypotheses in nestlings collected at post-hatch day seven, long before territorial behavior develops. This approach allowed us to ask whether, at such an early stage, methylation of this CRE already differs by allele or predicts expression of *VIP*. It also allowed us to look at how overall patterns of epigenetic regulation change in this region between an early stage of development and adulthood.

2 | MATERIALS AND METHODS

2.1 | Tissues and extractions

All procedures involving animals were approved by the Emory University Institutional Animal Care and Use Committee (IACUC), adhered to guidelines set forth by the National Institutes of Health Guide for the Care and Use of Laboratory Animals, and abided by all federal, state and local laws. Adult white-throated sparrows of both sexes and morphs ($n = 9$ TS males, 4 WS males, 5 TS females and 8 WS females) were collected in the Hemlock Stream Forest in Argyle, Maine in May and June 2013. We used song playback to lure adults into mist nets. Average time to capture was $5.96 \text{ min} \pm 0.89 \text{ SEM}$ and there was no effect of morph or sex on latency to capture (effect of morph, $p = 0.307$; effect of sex, $p = 0.654$). Nestlings ($n = 3$ TS males, 4 WS males, 5 TS females, 7 WS females) were collected from six nests (1–4 nestlings per nest) at the same field site in 2010 and 2011. The nestlings were collected on post-hatch day seven, approximately 2 days before the natural fledging age in this species.²⁷ Adults and nestlings were anesthetized with isoflurane and then euthanized by rapid decapitation within 10 min of capture and all tissue was collected within 15 min. Brains and livers were flash-frozen on powdered dry ice in the field and then shipped on dry ice to Emory University, where they were stored at -80°C . Sex was determined by visual inspection of gonads. Sex and morph were later confirmed by performing PCR on DNA extracted from the liver samples.^{9,28}

The samples of hypothalamus used in this study were microdissected as previously described.²⁹ Briefly, adult and nestling brains were sectioned on a cryostat at $300 \mu\text{m}$. Four punches of the hypothalamus from each brain were taken using a 1.0 mm punch tool.³⁰ Punches were centered on the midline to include both hemispheres. Two punches, a dorsal punch just below the anterior commissure and a ventral one just below the dorsal one (Figure S2), were taken in each of two consecutive sections. The adult punches were originally collected as part of an RNAseq study by Zinzow-Kramer et al.²⁹; the extracted RNA was also used in studies by Grogan et al.³¹ and Merritt et al.¹¹ to measure the expression of steroid-related genes. For those

studies, all four punches had been pooled for RNA/DNA extraction. These punches included the AH as well as other regions of the hypothalamus. Although some of these other regions likely contain a small number of *VIP* cells, the AH is by far the predominant *VIP* cell population in the sampled area in white-throated sparrows (B. M. Horton, unpublished). The DNA was bisulfite-converted and sequenced as described below. In nestlings, because of the much higher yield of RNA extracted,³¹ two separate pools of tissue were made for each nestling: one from the two dorsal punches and one from the two ventral punches. These dorsal and ventral samples were kept separate throughout RNA/DNA extraction and DNA bisulfite sequencing. So that the area sampled would be comparable to that sampled in adults, the data from nestlings were averaged across the dorsal and ventral samples for the statistical analysis. Because of this difference in the way the adult and nestling samples were processed, we did not make quantitative comparisons across age in our study. For further details on how these tissue samples were collected and processed, see Grogan et al.³¹; Merritt et al.¹¹; and Zinzow-Kramer et al.²⁹

DNA and RNA were extracted from the punches using a Qiagen Allprep DNA/RNA micro kit (Qiagen; Valencia, CA, USA) with modifications described by Zinzow-Kramer et al.,²⁹ which allows total RNA and genomic DNA extraction from the same samples. We used the DNA for bisulfite-converted sequencing to measure DNA methylation (described in the next section) and the RNA for the RNA sequencing (see “RNA Sequencing” below).

2.2 | Measurements of DNA methylation

2.2.1 | Bisulfite conversion and polymerase chain reaction

Two hundred nanograms of genomic DNA extracted from the punches of hypothalamus described above were bisulfite-converted and prepared for sequencing with a Zymo EZ DNA Methylation Lightning Kit. Because the sequences immediately upstream of TSS's typically contain important regulatory elements,²³ we targeted that region of the *VIP* gene. Four primer pairs were designed using the Zymo Bisulfite Primer Seeker software (Zymo; Irvine, CA, USA) and included adapter sequences (Table S1). Polymerase chain reaction (PCR) product lengths were 305–423 base pairs (bp) and spanned 28 SNPs and 74 CpG sites (Figure S1). Overall, these amplicons spanned 1282 bp upstream of the *VIP* TSS (NW_005081596.1:2258182–2,259,462). We refer to this region as the *VIP* CRE.

Each $25 \mu\text{l}$ PCR included $0.5 \mu\text{l}$ JumpStart Taq DNA polymerase, $2.5 \mu\text{l}$ JumpStart buffer containing 40 mM MgCl_2 , 5 nM of each primer and 8 ng of bisulfite-converted DNA. Cycling conditions were 95°C for 60 s, then 40 cycles of 95°C for 30 s, 55°C for 30 s and 70°C for 60 s, followed by a single final extension phase at 70°C for 5 min. Ten μl of each PCR product was visualized on a 1% agarose gel. Any PCRs that did not show a band were rerun, modifying the concentration of MgCl_2 until a band of the expected size was visible.

2.2.2 | Targeted bisulfite sequencing

Five microliters of each PCR product from all four primer pairs were pooled for each sample of the hypothalamus, for a total of 20 μ l per bird. Before sequencing, these samples were indexed with the 16S Metagenomic Sequencing Library kit (Illumina; San Diego, CA, USA). At the Emory Integrated Genomics Core, samples were run on the Agilent bioanalyzer to confirm indexing and DNA quality. The samples were then pooled into a single lane and sequenced on an Illumina MiSeq in 300 bp paired-end reads.

2.2.3 | Analysis of DNA methylation

We obtained an average of 305,000 reads per sample ($SD \pm 174,288$; $IQR = 164,233$). We assessed read quality with FastQC³² and trimmed the adapter sequences using Cutadapt.³³ An average of 7000 reads per sample were filtered out based on a Phred score cut-off of 30. We then aligned these reads to a bisulfite-converted genome with Bowtie2.³⁴ We created the bisulfite-converted genome with Bismark_genome_preparation³⁵ from a TS ZAL2 scaffold (NCBI accession: NW_005018596.1). The scaffold was N-masked at all SNPs except those disrupting CpG sites; the polymorphic CpG sites within the sequenced region were manually added to the reference genome to extract methylation status from all possible CpG sites following Merritt et al.¹¹ Reads with methylation at three or more non-CpG cytosines indicate possible incomplete bisulfite conversion or overestimated non-CG methylation and were filtered out (mean 3.6% per sample). We then identified the ZAL2 from the ZAL2^m reads using SNPsplit by assigning reads to one of the two alleles based on the SNPs that did not affect CpGs.³⁶ We extracted methylation counts using the bismark_methylation_extractor³⁵ and then used bismark2-bedGraph to convert the output to bedGraph format, which we used for data analysis in R (version 4.2.1).

2.3 | RNA sequencing

To test whether the degree of methylation predicted expression of the *VIP* alleles, we used previously existing RNAseq data (see References 11,29). RNAseq data were not available for some of the females and nestlings for which we had bisulfite sequencing data; the final sample sizes for the RNAseq portion of the study were as follows: adult males $n = 9$ TS, 4 WS; adult females $n = 3$ TS, 7 WS; nestling males $n = 2$ TS, 4 WS; nestling females $n = 5$ TS, 6 WS.

For the adult males, the mRNA sequencing was performed at the Genomics Core of the Emory National Primate Research Center (Atlanta, GA). RNA libraries were constructed according to the manufacturer's protocol using the Illumina TruSeq RNA Sample Preparation Kit v2 (Illumina Inc., San Diego, CA, USA) and sequenced on an Illumina HiSeq 1000 in 100 bp paired end-read reactions. For more detail, see Zinzow-Kramer et al.²⁹

For the adult females and nestlings, RNA library preparation and sequencing was performed by the Duke Center for Genomic and Computational Biology Sequencing and Genomic Technologies Core in Durham, NC. Libraries were constructed using a Kapa Stranded mRNA-Seq Kit (Kapa Biosystems, Boston, MA, USA) following the manufacturer's instructions, then sequenced on an Illumina HiSeq 4000 in 150 bp paired-end reactions. For more detail, see Merritt et al.¹¹

2.4 | Statistical analysis

All statistics were conducted in RStudio (RStudio Team 2022, R version 4.2.1). We conducted all analyses in adults and nestlings separately because of slight methodological differences in tissue collection and processing in the two age groups (see above), which prevented direct comparisons. All methylation data were \log_{10} transformed prior to statistical analyses. All statistical modeling for nestlings (see below) included a random effect of "nest" to account for sibling relatedness. Models that included both alleles from WS birds (i.e., both VIP^2 and VIP^{2m}) included a random effect of individual. For Benjamini-Hochberg adjustments of the alpha levels, the false discovery rate (FDR) was set at 5%.³⁷ Note that all effect sizes are also presented in the supplementary tables.

2.4.1 | Allelic differences in DNA methylation

Bisulfite treatment converts unmethylated cytosines to thymines. Therefore, we calculated percent methylation at each CpG as the number of cytosine reads divided by the total number of cytosine and thymine reads at that site. This calculation estimates the percentage of methylated CpGs out of the total number of CpGs. At polymorphic CpG sites, that is, sites present on only one allele due to genetic variation, the allele without the site was assigned zero percent methylation. Percent methylation was then averaged across the CRE to obtain a regional measure of percent methylation for each allele (VIP^2 or VIP^{2m}) for each bird. In all statistical analyses, the decimal proportion of methylated CpG sites was \log_{10} transformed to address right-skewed distribution. Because of the high number of zeros in the data set (both naturally occurring and artificially placed as described above), each proportion was increased by a small constant equal to half of the smallest value of the data set (+0.000219) prior to log transformation. Because the ZAL2 allele could be differentially methylated in TS versus WS birds, we needed to define three alleles in our data set: VIP^2 CRE in TS birds (herein TS- VIP^2), the VIP^2 CRE in WS birds (WS- VIP^2) and the VIP^{2m} CRE in WS birds (WS- VIP^{2m}).

To test for an effect of allele on percent methylation, we used ANOVAs within each age group using the stats (v4.2.1, R Core Team³⁸) and nlme packages (v3.1-157).^{39,40} Sex was initially included in these analyses to test for sex differences; however, we had no a priori reason to expect sex differences and sex was not part of our

hypotheses. Therefore, sex was dropped from the models upon confirming there were no main effects of sex or interactions between sex and allele. The ANOVAs included a random effect of individual to control for repeated measures (methylation of $WS-VIP^2$ and $WS-VIP^{2m}$) in WS birds. We calculated the eta-squared (η^2) effect size for each of these ANOVAs with the *effectsize* package.⁴¹ To determine which alleles were differentially methylated, we conducted *post-hoc* ANOVAs to compare DNA methylation between $TS-VIP^2$ and $WS-VIP^2$ and between $WS-VIP^2$ and $WS-VIP^{2m}$. To determine the effect of polymorphic CpGs on allelic differences in methylation of the *VIP* CRE, we ran these analyses in parallel, with the polymorphic CpGs included and excluded.

To explore how differential methylation manifests across the epigenetic landscape, we conducted ANOVAs as above, using the same three levels ($TS-VIP^2$, $WS-VIP^2$, $WS-VIP^{2m}$), to test for an effect of allele on percent methylation of each CpG. As above, we followed up significant effects of allele with *post-hoc* ANOVAs to test for a difference between $TS-VIP^2$ and $WS-VIP^2$ and between $WS-VIP^2$ and $WS-VIP^{2m}$. When comparing $WS-VIP^2$ and $WS-VIP^{2m}$, we included a random effect of individual to account for comparison within-subjects. Because our primary goal was to show patterns across the epigenetic landscape rather than to positively identify all individual CpGs that were differentially methylated, we performed these *post-hoc* tests on the basis of the significance of main effects before Benjamini-Hochberg adjustment. For interested readers, significant effects that survived the adjustment, as well as effect sizes for all comparisons, are presented in the supplementary tables. For the polymorphic CpGs, we tested whether methylation of the allele with the CpG differed from the assumed zero methylation of the alternative allele by using a one-sample, one-tailed *t*-test. We also tested for differences between the morphs in the methylation of individual CpGs on VIP^2 with a student's *t*-test. Effect sizes (Cohen's *d*) were calculated using the *lsr* package (v0.06⁴²).

To investigate whether our positive findings could be related to the degree of methylation at any particular site, we performed bootstrapping analyses (iteration $n = 1000$) using the *boot* package (v1.3–28).^{43,44} For each shared CpG, we calculated the probability of finding an allelic difference in methylation equal to or larger than the difference we actually observed. We then tested, using regression, whether this probability was predicted by the degree of methylation at that site. This analysis was carried out for both the $TS-VIP^2$ versus $WS-VIP^2$ and the $WS-VIP^2$ versus $WS-VIP^{2m}$ pairwise comparisons.

2.4.2 | Regressions of DNA methylation on expression

To test whether methylation of the *VIP* CRE predicted *VIP* expression, we ran linear regressions of \log_2 transformed *VIP* read counts normalized to library size on methylation of the *VIP* CRE.¹¹ Each of these regressions was performed separately for each allele ($TS-VIP^2$; $WS-VIP^2$; $WS-VIP^{2m}$), using allele-specific reads and percent methylation

of that allele within each age group. Because the RNA sequencing was carried out in separate runs for the male and female adults, and to control for sex in these models to the extent possible, we included sex as a fixed-effect in all regressions for both adults and nestlings. To test whether methylation of any individual CpG sites predicted expression, we ran separate analyses for each CpG. These analyses consisted of linear regressions of the percent methylation of each CpG on the expression of the corresponding allele for a total of three sets of regressions (one per allele) at every CpG site in each age group.

3 | RESULTS

3.1 | Adults

We first investigated the methylation of the entire *VIP* CRE, considered as a whole. This region contains a total of 74 CpG sites. Of those sites, 66 are present on both the VIP^2 and VIP^{2m} alleles and eight are polymorphic; that is, they are present on only one allele because of genetic differentiation. Of the polymorphic sites, three are present only on VIP^2 and five only on VIP^{2m} . For our analyses, the methylation of polymorphic CpG sites was assumed to be zero on the allele without the site. We found that the level of DNA methylation varied across the region, with several clear peaks and troughs in similar locations in all three alleles (Figure 2). The average percent methylation was low across the three alleles ($TS-VIP^2$: 7.10% [SD = 12.3%]; $WS-VIP^2$: 7.52% [SD = 14.1%]; $WS-VIP^{2m}$: 6.27% [SD = 13.4%]). We found no main effect of sex on the methylation of this region, either with polymorphic CpG sites included ($p = 0.612$, $F_{[1,24]} = 0.265$, $\eta^2 = 0.011$) or excluded ($p = 0.606$, $F_{[1,24]} = 0.273$, $\eta^2 = 0.011$), and sex was removed from the model. ANOVA showed a significant main effect of allele on DNA methylation ($p = 0.009$, $F_{[2,10]} = 7.737$, $\eta^2 = 0.607$; Table S2). *Post-hoc* *t*-tests showed that the region was less methylated on $WS-VIP^2$ than on $WS-VIP^{2m}$ ($p = 0.003$, $t = -3.878$, $d = 0.901$; Figure 3A; Table S2). To explore the possibility that the differential methylation between the WS alleles could be explained by differentiation of the genetic sequence itself, we repeated the analysis with the polymorphic sites excluded. The main effect of allele then became slightly larger ($p = 0.002$, $F_{[2,10]} = 13.245$, $\eta^2 = 0.726$; Table S2) and the difference in percent methylation between $WS-VIP^2$ and $WS-VIP^{2m}$ persisted ($p = <0.001$, $t = -5.081$, $d = 1.082$; Figure 3A; Table S2). Thus, the differential methylation is unlikely to be explained by the loss or gain of CpG sites due to genetic differentiation. There was no difference in average percent methylation between $TS-VIP^2$ and $WS-VIP^2$ (Table S2); that is, we could not detect epigenetic differentiation between the TS and WS versions of the VIP^2 CRE.

We next tested for differential methylation at individual CpGs. Initial ANOVAs showed that the main effect of allele on the percent methylation was significant at alpha of 0.05 for eight of the 66 CpG sites present on both alleles ("shared" sites) (Table S3). We then

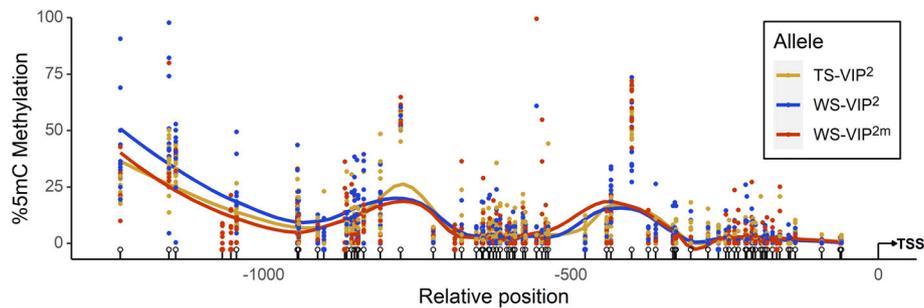


FIGURE 2 The epigenetic landscape of the *VIP* CRE in adults, showing similar patterns of methylation in the three alleles. The X-axis indicates the position of each CpG relative to the transcription start site (TSS, indicated by an arrow on the bottom right). Open lollipops indicate CpG sites shared by both alleles; blue and red lollipops indicate polymorphic sites on *VIP*² and *VIP*^{2m}, respectively. Each dot on the graph represents the percent methylation (on the Y-axis) of that CpG on one allele in one individual. The solid lines depict a LOESS smooth function indicating how the degree of methylation changes across the region. The TS-*VIP*², WS-*VIP*² and WS-*VIP*^{2m} alleles are indicated in tan, blue and red, respectively

proceeded to make pairwise comparisons between the TS-*VIP*² and WS-*VIP*² alleles and between the WS-*VIP*² and WS-*VIP*^{2m} alleles for each of these eight CpGs (Figure 3B, Table S3). These post-hoc ANOVAs indicated that, after corrections for multiple comparisons, none of the sites were differentially methylated between TS-*VIP*² and WS-*VIP*². Four of the sites were, however, differentially methylated between WS-*VIP*² and WS-*VIP*^{2m}. Of these four sites, three were significantly more methylated on WS-*VIP*² and one (2259083) was more methylated on WS-*VIP*^{2m}. Our bootstrapping analyses showed that for the shared CpG sites, the probability of finding a difference equal to or larger than the ones we detected was not predicted by the degree of methylation (TS-*VIP*² vs. WS-*VIP*² alleles, $R^2 = 0.002$, $F_{(1,64)} = 0.097$, $p = 0.757$; WS-*VIP*² vs. WS-*VIP*^{2m} alleles, $R^2 = 0.001$, $F_{(1,64)} = 0.656$, $p = 0.421$).

To assess the potential contribution of genetic differentiation to differential methylation of the CRE, we tested for differential methylation at each polymorphic CpG site. This part of our investigation was important because we did not want to assume that any particular polymorphic site was significantly methylated above zero. In WS birds, all eight of the polymorphic sites—three on *VIP*² and five on *VIP*^{2m}—were significantly different from the assumed zero methylation of the alternative allele. None of the three *VIP*²-specific CpG sites showed a morph difference in percent methylation (Figure 3C, Table S4).

Lastly, we investigated whether methylation of the *VIP* CRE predicted expression of *VIP* mRNA. For this analysis, we used RNAseq data from the same tissue punches (Figure S2).^{10,11,31} The percent methylation of the region, averaged across the CRE for each *VIP* allele, did not predict the expression of that allele (Table S5). Allelic *VIP* expression was, however, significantly predicted by methylation of some of the individual CpGs. We noted eight individual CpGs that predicted expression of the corresponding allele, all of which were located on TS-*VIP*² (Figure 4; Table S6). None of these CpGs were polymorphic, nor were they differentially methylated (compare Table S3 vs. S6). There were no individual CpGs in the other alleles that significantly predicted expression after Benjamini-Hochberg adjustment (Tables S7 and S8).

3.2 | Nestlings

The epigenetic landscape changes considerably over the course of development.^{45–47} Therefore, we also looked at the methylation of the *VIP* CRE in nestlings. The pattern of DNA methylation across the region in nestlings at seven days post-hatch largely resembled that of adults and was similar across the three alleles, with several noticeable peaks (Figure S3). Considering the region as a whole, the average percent methylation was low in nestlings, as it was in adults (TS-*VIP*²: 5.30% [SD = 10.8%]; WS-*VIP*²: 5.93% [SD = 12.5%]; WS-*VIP*^{2m}: 5.66% [SD = 12.6%]). As was the case for adults, we found no sex differences in the methylation of this region either with polymorphic CpG sites included ($p = 0.095$, $F_{[1,24]} = 3.278$, $\eta^2 = 0.215$) or excluded ($p = 0.098$, $F_{[1,24]} = 3.220$, $\eta^2 = 0.211$). Average percent methylation differed significantly among the alleles only when polymorphic CpGs were excluded ($p = 0.041$, $F_{[2,9]} = 4.651$, $\eta^2 = 0.508$; Table S9), consistent with the pattern we observed in adults. Post-hoc t-tests showed that WS-*VIP*^{2m} was less methylated than WS-*VIP*² at shared sites ($p = 0.016$; $t = -2.907$, $d = 0.534$; Table S9; Figure S4A).

As in adults, we next tested for differential methylation of individual CpGs. Initial ANOVAs that included all three alleles showed a main effect of allele for six out of the 66 shared CpGs (Table S10). Among these six, post-hoc analyses within WS birds showed that four sites were more methylated on WS-*VIP*² than on WS-*VIP*^{2m} and two were more methylated on WS-*VIP*^{2m} than on WS-*VIP*² (Figure S4B). None of the CpG sites in nestlings were differentially methylated between the TS- and WS-*VIP*² alleles (Table S10-11; Figure S4B,C); that is, differences in methylation occurred exclusively between the two WS alleles, WS-*VIP*² and WS-*VIP*^{2m}. Our bootstrapping analyses showed that for the shared CpG sites, the probability of finding a difference equal to or larger than the ones we detected was not predicted by the degree of methylation (TS-*VIP*² vs. WS-*VIP*² alleles, $R^2 = 0.003$, $F_{(1,64)} = 0.194$, $p = 0.661$; WS-*VIP*² vs. WS-*VIP*^{2m} alleles, $R^2 = 0.042$, $F_{(1,64)} = 2.801$, $p = 0.099$). All of the polymorphic CpGs in nestlings were significantly more methylated than the assumed zero methylation of the alternative allele (Figure S4C,D; Table S11).

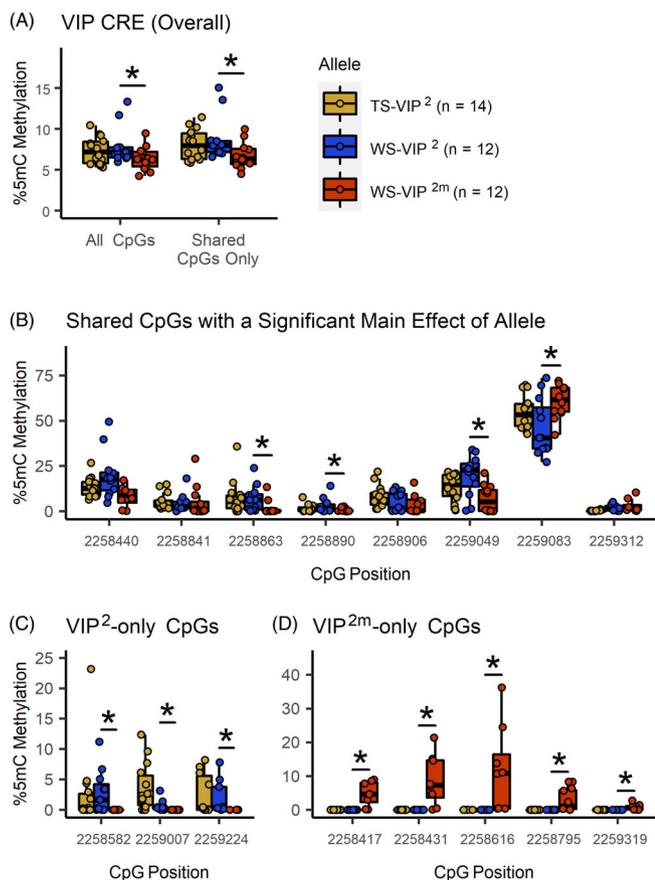


FIGURE 3 Allelic differences in percent methylation (5mC% methylation) of the *VIP* CRE in adults. Panel A shows the average percent methylation of the region as a whole with polymorphic sites included (left) and excluded (right; see Table S2). Panel B shows the percent methylation of shared CpGs at which there was a significant main effect of allele. Panels C and D show the percent methylation for polymorphic CpGs on *VIP²* and *VIP^{2m}*, respectively. For this analysis, the percent methylation was set at zero for the allele without the site (see Methods). For all polymorphic sites, methylation was higher than zero (the presumed methylation of that location on the other allele). In Panels B–D, the positions of each CpG are shown on the X-axis, in order from furthest to closest to the transcription start site from left to right. Significant differences between alleles after correction for multiple comparisons are denoted by asterisks (see also Tables S3 and S4)

The region-wide methylation of this CRE did not predict allelic-specific expression in nestlings (Table S5). There were also no individual CpGs that predicted allelic expression after correction for multiple comparisons (Figure S5; Tables S6–S8).

3.3 | Qualitative comparison between adults and nestlings

To explore the potential for regulatory hot spots that are conserved between nestlings and adults, we next looked for similarities between the age groups in our pattern of results. As a caveat, all comparisons

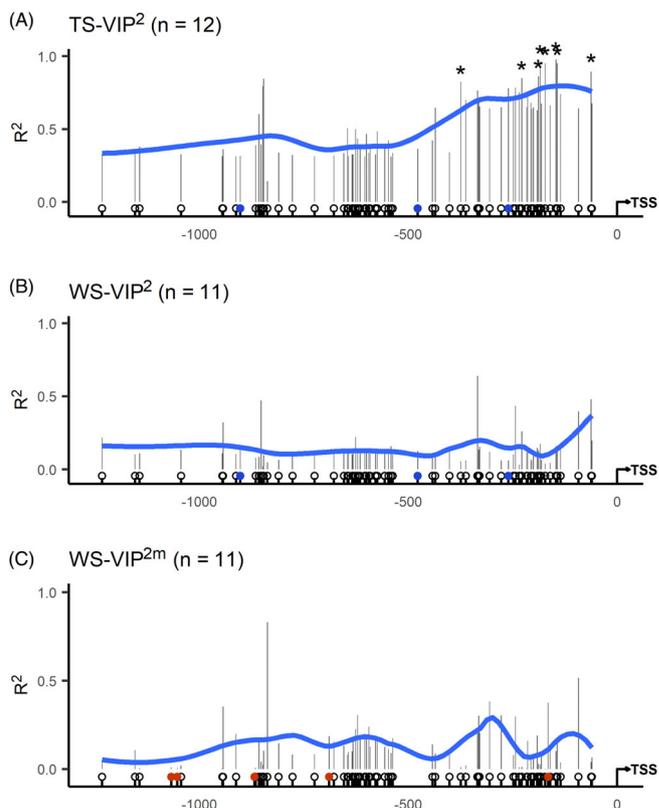


FIGURE 4 Relationships between percent methylation of individual CpGs and expression of the corresponding allele in the same tissue samples. R^2 values are plotted for *TS-VIP²* (A), *WS-VIP²* (B) and *WS-VIP^{2m}* (C) as a function of position along the *VIP* CRE. In each panel, the magnitude of R^2 for each regression is indicated by a vertical black line at the position of that CpG relative to the transcription start site (TSS, indicated by an arrow on the bottom right). The blue line in each panel represents a LOESS smooth function over the R^2 values and shows how the relationship between methylation and expression changes across the region. Open lollipops indicate CpG sites shared by the *VIP²* and *VIP^{2m}* alleles; blue and red lollipops indicate polymorphic sites on *VIP²* and *VIP^{2m}*, respectively. * indicates a p -value < alpha after Benjamini-Hochberg correction (see also Tables S6–S8)

across age group in this study were necessarily qualitative in nature. Because the tissue samples were processed slightly differently in adults and nestlings (see methods), we could not test for interactions between age and DNA methylation. Thus, we could not assess quantitatively the degree to which the differential methylation of any particular CpG may differ between adults and nestlings.

As noted above, the DNA methylation landscape appeared to be quite similar between the age groups across the sequenced region, with peaks and troughs at the same locations (Figures 2, S3). At the level of individual CpG sites, as one might expect, the greatest similarities were found among the polymorphic sites (Figure 5). Methylation of all eight such sites was significantly different from zero in both adults and nestlings. We found fewer similarities at shared sites. Of the shared CpG sites that were significantly differentially methylated between the WS alleles in the post-hoc tests, two were detected only

in adults and four only in nestlings. Only two were detected in both age groups (2259049, 2259083); one of these was more methylated on *WS-VIP*² and one was more methylated on *WS-VIP*^{2m}. The site that was more methylated on *TS-VIP*² than *WS-VIP*² (2259312) was detected only in adults. When considering the shared and polymorphic sites together, we noted that differentially methylated sites seemed to cluster at the upstream end of the CRE, furthest from the TSS, in both ages (Figure 5).

All eight of the CpG sites that predicted expression of any allele, at any age, were located on *TS-VIP*² in adults only (Figure 6). We did not observe notable similarities between adults and nestlings in the patterns of predictive sites, either overall or within either of the *WS* alleles. The predictive CpG sites on the *TS-VIP*² allele in adults were concentrated near the TSS; this clustering was not observed in nestlings (Figure 6).

4 | DISCUSSION

In this study, we showed that epigenetic regulation of the *VIP* gene is a potential modulator of allele-specific *VIP* expression in white-throated sparrows. Because expression of *VIP* is already known to be causal for aggression in at least two other species of songbird,¹⁹ this epigenetic regulation may contribute to the well-known alternative behavioral phenotypes in this species (see References 2,20). Focusing on a likely *cis*-regulatory region encompassing 1.2 kb upstream of the *VIP* start site, we showed that the supergene allele of *VIP*, which is present only in individuals of the more aggressive *WS* morph,^{5,48,49} is less methylated than the standard allele in DNA extracted from hypothalamus (Figure 3A). Notably, this differential DNA methylation was evident even in nestlings, only seven days after hatching (Figure S4A).

4.1 | Differential methylation of the standard versus supergene alleles

The region we analysed contained 74 CpG sites, creating diverse opportunities for allele-specific DNA methylation. Of those sites, eight were polymorphic, meaning they exist on only one of the two alleles. Genetic variation at CpG sites, particularly in regulatory elements such as the one analysed here, is thought to alter epigenetic landscapes in an allele-specific fashion.^{26,50,51} In a previous study,¹¹ we showed evidence of this phenomenon for *ESR1*, a gene that is only ~500 kb upstream from *VIP* in the white-throated sparrow genome (NCBI reference sequence: NW_005081596.1). Allele-specific methylation of an upstream *cis*-regulatory region on *ESR1* was explained by polymorphic sites; in other words, when we limited our analysis to non-polymorphic sites, we could not detect allele-specific methylation of the region. For *ESR1*, therefore, genetic differentiation inside the supergene potentially disrupts CpG sites, reducing opportunities for methylation on *ZAL2*^m.

Even though *ESR1* and *VIP* may be co-regulated and even coadapted,¹ the findings we report here for *VIP* are dissimilar to our previous findings for *ESR1* in that the lower methylation of the supergene allele of *VIP* could not be attributed solely to polymorphic CpGs. When considering only non-polymorphic CpG sites, that is, those shared between the two alleles, we still detected lower methylation of the supergene allele of *VIP* (Figure 3A, S4A). In adults and nestlings, the effect of allele increased when we considered only shared sites. Thus, a nontrivial amount of the epigenetic differentiation between the two alleles is occurring at shared CpG sites. This methylation seemed biased toward the standard allele; in adults, three out of the four shared sites that were differentially methylated between the supergene and standard allele in *WS* birds were more methylated on the standard allele (Figure 3B); the same was the case for four out of six sites in nestlings (Figure S4B). We noted, however, that such sites constituted only about 10% of the total shared sites, suggesting that allele-specific methylation of the region as a whole is driven either by a large effect of a few sites or collectively by sites that are themselves not significantly differentially methylated. It is also possible that the methylation of the polymorphic CpGs on *WS-VIP*^{2m}, which is relatively high (Figure 3D), could compensate for the differential methylation at shared sites, making the overall methylation of the two alleles more similar. Such a mechanism could be consistent with the idea that degeneration of the *ZAL2*^m chromosome drives dosage compensation, potentially via opportunities for methylation.⁸

Epigenetic differentiation of non-polymorphic CpG sites was also found by Okhovat et al.⁵⁰ In prairie voles (*Microtus ochrogaster*), allelic differences in methylation of an enhancer in the gene encoding the vasopressin receptor *V1a* (*avpr1a*) were highly significant even when polymorphic CpG sites were excluded from the analysis. These results suggest that genetic variation can influence local methylation of nearby, non-polymorphic CpG sites. Using the data from this study, we could not detect evidence of such simply by testing for correlations between methylation and proximity to SNPs. The relationship between genetic variation and local epigenetic modifications of non-polymorphic sites is likely to be much more complex.

In a genome-wide study of DNA methylation in white-throated sparrows, Sun et al.²⁶ assigned genes on *ZAL2/2*^m into three categories: more methylated on *ZAL2*, more methylated on *ZAL2*^m, or “extremely hypomethylated” on *ZAL2*^m. Although we found here that methylation of *WS-VIP*^{2m} was significantly lower than *WS-VIP*², the difference would not place this region of the *VIP* gene into the “extremely hypomethylated” category, which is based on the size of the difference between *ZAL2* and *ZAL2*^m.²⁶ We detected differential DNA methylation primarily at the upstream end of the CRE (Figure 5), where methylation was generally higher than at the downstream end, close to the TSS; the likelihood of our findings did not depend, however, on the level of methylation (see bootstrapping analyses). In mammals, areas surrounding TSSs are generally less methylated than other areas of the genome (as cited in Reference 23). Work in human cell lines has shown that expression of the *VIP* gene can be regulated by far upstream elements.⁵² Methylation of such elements, for

example enhancers, can significantly affect gene transcription generally.^{51,52} Thus, upstream sequences of the *VIP* gene should be analysed in future studies on white-throated sparrows.

4.2 | Relationships between DNA methylation and allele-specific *VIP* expression

Although most of the differential DNA methylation we detected was located on the upstream end of our sequence, where the DNA tended to be more methylated, the regions that most strongly predicted *VIP* mRNA expression were located in the hypomethylated region immediately upstream of the start site. This hypomethylated region seemed to be enriched with CpGs that predicted the expression of the standard allele, particularly in adults (Figures 4, 6, S5). Notably, these predictive sites were shared, rather than polymorphic, and did not seem to cluster with polymorphic sites. This finding stands in contrast with the neighboring *ESR1* gene, for which all predictive sites were either polymorphic or inside a cluster containing at least one polymorphic site.¹¹ For the *VIP* CRE, however, as was the case for differential methylation between alleles (Figure 3A, S4A), the important regulatory variation seems to stem from epigenetic differentiation at non-polymorphic sites.

All of the CpGs that significantly predicted allele-specific *VIP* expression were located on the standard allele in TS birds, *TS-VIP²*. This result is consistent with the possibility that local trans-regulatory elements vary both according to morph and developmental stage. That is, the availability and mechanisms of action of trans factors might themselves depend on morph and age, thus altering the exact region of the CRE that is most important for *VIP²* expression (see Reference 46). Also notable was our finding that for most of the predictive CpGs, methylation was positively correlated with *VIP* expression. In general, this relationship is expected to be negative; in their genome-wide study of DNA methylation in white-throated sparrows, Sun et al.²⁶ found evidence consistent with the idea that DNA methylation of cis-regulatory regions, particularly those immediately upstream of TSS's, dampens expression. Other authors have found

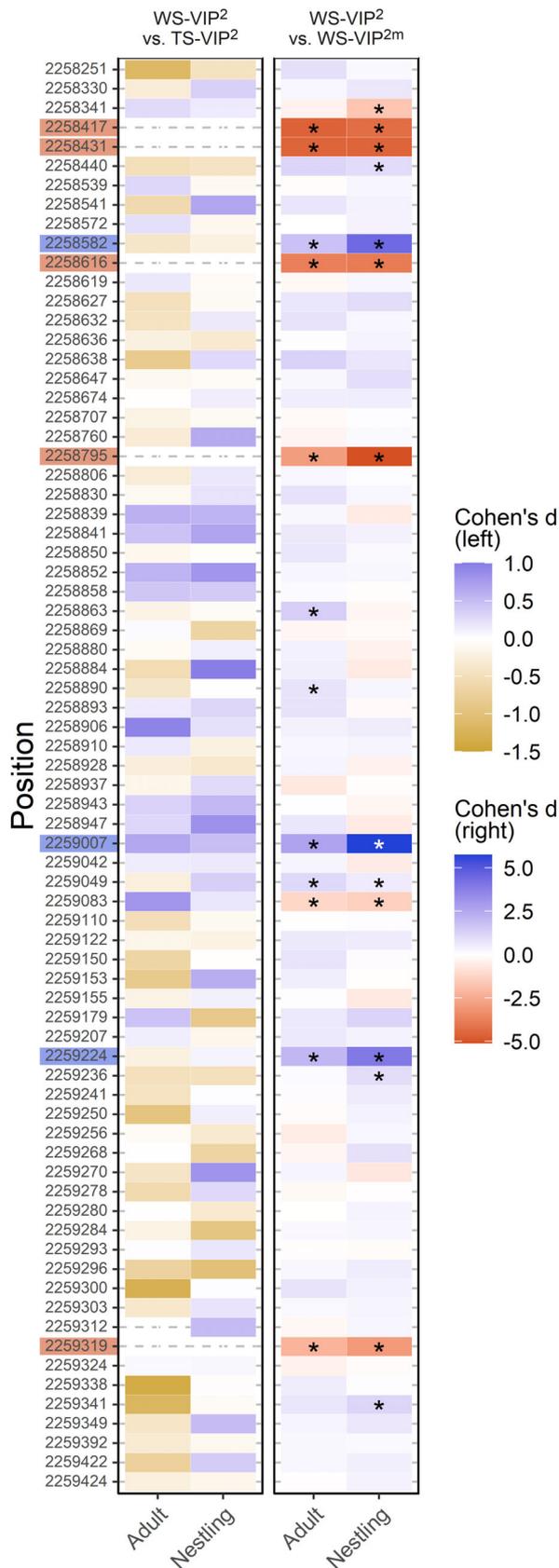
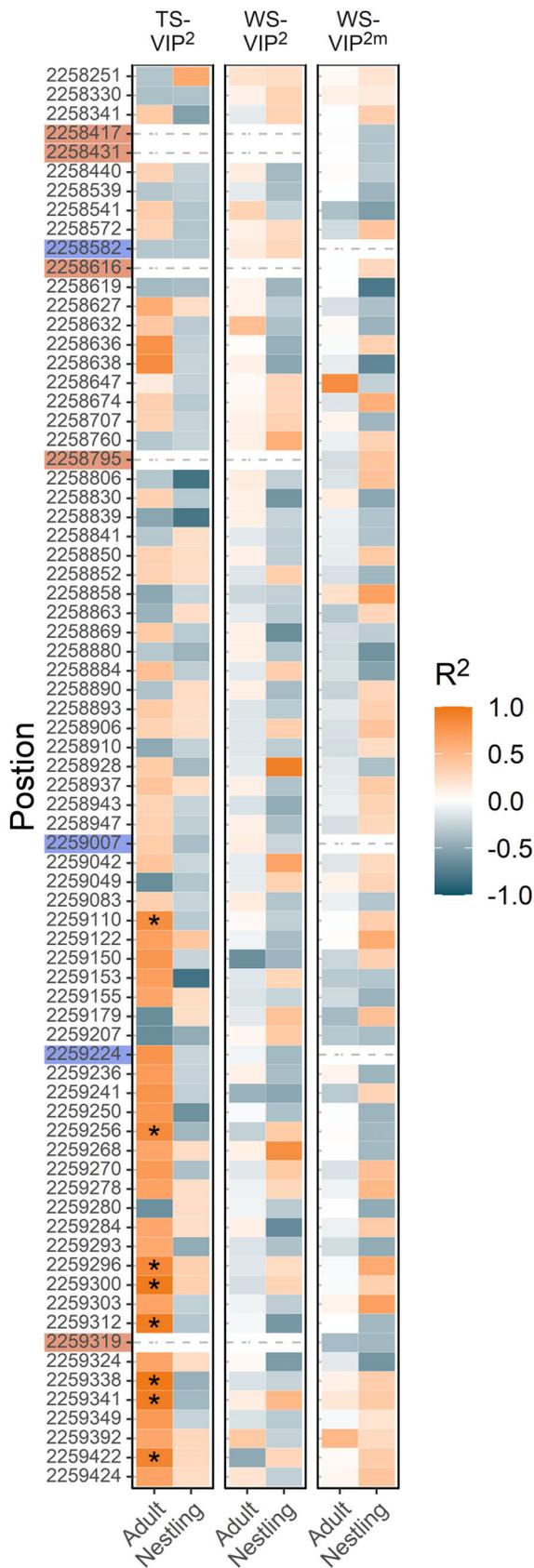


FIGURE 5 Effects of allele on the percent methylation of each CpG in the *VIP* CRE. Comparisons between *TS-VIP²* and *WS-VIP²* are indicated in the columns on the left, with tan and blue indicating greater methylation of *VIP²* in TS birds or WS birds, respectively. Comparisons between *WS-VIP²* and *WS-VIP^{2m}* are shown in the columns on the right, with blue and red indicating greater methylation of *VIP²* or *VIP^{2m}*, respectively. Darker colors indicate a larger effect size (Cohen's *d*) with positive values corresponding to higher methylation of *WS-VIP²*. CpGs are listed in order according to their proximity to the transcription start site, with the closest to the start site at the bottom and furthest upstream at the top. Locations of CpGs are shown by the numbers on the left, with highlighting to indicate allele-specific CpGs; blue indicates *VIP²* and red *VIP^{2m}*. Because the first two columns compare sites on *VIP²*, there are no data for *VIP^{2m}*-specific CpGs (gray dashed lines). For sample sizes, see Figure 3 and S4. *Significant effect of allele ($p < \text{Benjamini-Hochberg corrected alpha}$) in the pairwise post-hoc comparison indicated at the top of each panel. For complete statistical results see Table S3 and S4 (adults) and Table S10 and S11 (nestlings)

positive relationships, however, between methylation and expression, for example in ants (*Camponotus floridanus* and *Harpegnathos saltator*⁵³) and termites (*Zootermopsis nevadensis*).⁵⁴



4.3 | Limitations of this study

Our over-arching hypothesis is that epigenetic differentiation of *ZAL2*^m inside the supergene causes differential expression of *VIP*, resulting in behavioral differences between birds with and without the supergene.^{1,20} We cannot rule out the possibility, however, that behavioral differences themselves drive differential methylation of *VIP*. Epigenetic marks on DNA, in particular DNA methylation, are widely understood to be profoundly influenced by experience.^{55–58} This epigenetic regulation is likely to alter social behavior.⁵⁹ Indeed, in trees wallows (*Tachycineta bicolor*), methylation of the gene encoding the *VIP* receptor has been shown to be affected by engaging in territorial behavior over a period of just two days.⁶⁰ TS and WS white-throated sparrows differ in the degree to which they engage in territorial aggression,² which could at least partly explain the differential methylation we detected in this study. Although the standard and supergene alleles were differentially methylated even in nestlings, which do not yet defend breeding territories, other behavioral differences in nestlings could contribute to DNA methylation. We believe, however, that if differential methylation of the *VIP* CRE is driven primarily by morph differences in behavior, the effect would likely have manifested as morph-specific methylation of the standard allele, which is present in both morphs. Of the 69 CpG sites on the standard allele, none were differentially methylated between the morphs in adults or in nestlings (Figures 3, 5, S4). Therefore, it does not seem that methylation of this CRE is significantly influenced by morph-specific factors such as hormonal milieu, gene expression, or behavior. Our evidence is more consistent with a model wherein the morph-specific environment interacts with WS-*VIP*^{2m}-specific sequences to produce allele-specific methylation in WS birds.

We could not test formally for differences between the age groups (nestlings vs. adults) due to slight differences in the methods by which the tissue samples and data were processed. Levels of DNA methylation appeared to be lower in nestlings than in adults, which is consistent with reported levels of methylation in this species genome-wide.²⁶ We also noted that although the differentially methylated sites were similarly located in both age groups (Figure 5), they were not identical; further, sites that predicted allele-specific expression were found only in adults (Figure 6). These discrepancies, which may result from developmental, experiential, or seasonal processes, did not cause major

FIGURE 6 Relationships between methylation of individual CpGs and allele-specific expression of *VIP*. The colors indicate the effect size R², with orange denoting a positive relationship and teal a negative relationship between expression of the *VIP* allele and the percent methylation of the CpG site on that allele. Locations of CpGs are shown by the numbers on the left, with highlighting to indicate allele-specific CpGs; blue indicates *VIP*² and red *VIP*^{2m}. CpGs are listed in order according to their proximity from the transcription start site, from closest at the bottom to furthest upstream at the top. Dashed lines indicate polymorphic CpGs not present on that allele. For sample sizes see Figures 4 and S5. **p* < Benjamini-Hochberg-corrected alpha level based on multiple linear regression; see Tables S6–S8 for the complete results

differences in the overall genetic landscape of the region analysed in this study (Figure 2, S3), nor did they prevent us from detecting region-wide differential methylation at both ages (Figure 3A, S4A). Because transcription factors interact with stretches of DNA that span as many as 30 bp,⁶¹ methylation of neighboring, yet distinct sites within the same consensus sequence may impact expression similarly.^{58,62}

We measured DNA methylation of the *VIP* CRE in a heterogeneous sample of hypothalamic tissue in wild animals behaving in their natural habitat. These approaches may have affected our results by increasing variation. Our hypothalamic samples, although they contained a major *VIP* cell population, also contained other types of cells that do not express *VIP* at all. The latter cells may suppress *VIP* expression through modifications of chromatin accessibility and increased DNA methylation, which should occur equally on both alleles. Thus, this heterogeneity could have introduced irrelevant variation. Subsequent studies should isolate *VIP* cells more specifically in order to further explore these relationships. Further, as reported by Bentz et al.,⁶⁰ engaging in social behaviors can affect methylation of *VIP*-related genes in songbirds. Okhovat et al.⁵⁰ showed that in prairie voles, methylation of an enhancer in the *avpr1a* gene predicted expression more strongly in lab-reared than wild-caught animals, suggesting that environmental variation could introduce noise into the data. Thus, the effects we report here may have been stronger if we could have controlled that variation more tightly. We believe, however, that studying free-living animals offers powerful advantages such as increased ecological validity and the ability to choose model organisms that are best-suited to the research questions. The white-throated sparrow continues to provide rich opportunities for understanding gene-behavior relationships.

ACKNOWLEDGMENTS

The authors thank Brent Horton, Dan Sun, Soojin Yi and Wendy Zinzow-Kramer for technical assistance. We also thank Sherryl Goodman and Aubrey Kelly for feedback on an earlier draft of this manuscript. The authors report no conflict of interest.

FUNDING INFORMATION

National Institutes of Health R01MH082833 to Donna L. Maney and F31MH114509 to Jennifer R. Merritt.

DATA AVAILABILITY STATEMENT

Data Availability: The data and code that support the findings of this study are available at https://github.com/mpricha/WTSP_VIPMethylation. The raw sequencing data are available at Dryad <https://doi.org/10.5061/dryad.547d7wm9s>. RNA sequencing is available on NCBI: <https://www.ncbi.nlm.nih.gov/proxy/library.emory.edu/bioproject/657006>.

ORCID

Mackenzie R. Prichard  <https://orcid.org/0000-0003-3218-9082>
Kathleen E. Grogan  <https://orcid.org/0000-0002-0112-1766>
Jennifer R. Merritt  <https://orcid.org/0000-0002-1189-0454>

Jessica Root  <https://orcid.org/0000-0002-9150-9579>

Donna L. Maney  <https://orcid.org/0000-0002-1006-2358>

REFERENCES

- Maney DL, Merritt JR, Prichard MR, Horton BM, Yi SV. Inside the supergene of the bird with four sexes. *Horm Behav.* 2020;126:104850.
- Horton BM, Moore IT, Maney DL. New insights into the hormonal and behavioural correlates of polymorphism in white-throated sparrows, *Zonotrichia albicollis*. *Anim Behav.* 2014;93:207-219.
- Kopachena JG, Falls JB. Aggressive performance as a behavioral correlate of plumage polymorphism in the white-throated sparrow (*Zonotrichia albicollis*). *Behaviour.* 1993a;124(3/4):249-266.
- Kopachena JG, Falls JB. Re-evaluation of morph-specific variations in parental behavior of the white-throated sparrow. *Wilson Bull.* 1993b;105:48-59.
- Thornycroft HB. A cytogenetic study of the white-throated sparrow, *Zonotrichia albicollis* (GMELIN). *J Evol.* 1975;29(4):611-621.
- Davis JK, Mittel LB, Lowman JJ, et al. Haplotype-based genomic sequencing of a chromosomal polymorphism in the white-throated sparrow (*Zonotrichia albicollis*). *J Hered.* 2011;102(4):380-390.
- Huynh LY, Maney DL, Thomas JW. Chromosome-wide linkage disequilibrium caused by an inversion polymorphism in the white-throated sparrow (*Zonotrichia albicollis*). *Heredity.* 2011;106(4):537-546.
- Sun D, Huh I, Zinzow-Kramer WM, Maney DL, Yi SV. Rapid regulatory evolution of a nonrecombining autosome linked to divergent behavioral phenotypes. *Proc Natl Acad Sci.* 2018;115(11):2794-2799.
- Thomas JW, Caceres M, Lowman JJ, et al. The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (*Zonotrichia albicollis*) is a complex rearrangement and suppressor of recombination. *Genetics.* 2008;179(3):1455-1468.
- Tuttle EM, Bergland AO, Korody ML, et al. Divergence and functional degradation of a sex chromosome-like supergene. *Curr Biol.* 2016;26(3):344-350.
- Merritt JR, Grogan KE, Zinzow-Kramer WM, et al. A supergene-linked estrogen receptor drives alternative phenotypes in a polymorphic songbird. *Proc Natl Acad Sci.* 2020;117:21673-21680.
- Kuenzel WJ, McCune SK, Talbot RT, Sharp PJ, Hill JM. Sites of gene expression for vasoactive intestinal polypeptide throughout the brain of the chick (*Gallus domesticus*). *J Comp Neurol.* 1997;381(1):101-118.
- Montagnese CM, Székely T, Csillag A, Zachar G. Distribution of vasotocin- and vasoactive intestinal peptide-like immunoreactivity in the brain of blue tit (*Cyanistes coeruleus*). *Front Neuroanat.* 2015;9:90.
- Goodson JL. Territorial aggression and dawn song are modulated by septal vasotocin and vasoactive intestinal polypeptide in male field sparrows (*Spizella pusilla*). *Horm Behav.* 1998;34(1):67-77.
- Goodson JL, Wilson LC, Schrock SE. To flock or fight: neurochemical signatures of divergent life histories in sparrows. *Proc Natl Acad Sci.* 2012;109(Supplement 1):10685-10692.
- Kingsbury MA. New perspectives on vasoactive intestinal polypeptide as a widespread modulator of social behavior. *Curr Opin Behav Sci.* 2015;6:139-147.
- Kingsbury MA, Wilson LC. The role of *VIP* in social behavior: neural hotspots for the modulation of affiliation, aggression, and parental care. *J Integr Comp Biol.* 2016;56(6):1238-1249.
- Goodson JL. The vertebrate social behavior network: evolutionary themes and variations. *Horm Behav.* 2005;48(1):11-22.
- Goodson JL, Kelly AM, Kingsbury MA, Thompson RR. An aggression-specific cell type in the anterior hypothalamus of finches. *J Proc Natl Acad Sci.* 2012;109(34):13847-13852.
- Horton BM, Michael CM, Prichard MR, Maney DL. Vasoactive intestinal peptide as a mediator of the effects of a supergene on social behaviour. *Proc Royal Soc B: Biol Sci.* 2020;287(1924):20200196.

21. Heberle E, Bardet AF. Sensitivity of transcription factors to DNA methylation. *Essays Biochem.* 2019;63(6):727-741.
22. Mendizabal I, Yi SV. Whole-genome bisulfite sequencing maps from multiple human tissues reveal novel CpG islands associated with tissue-specific regulation. *Hum Mol Genet.* 2016;25(1):69-82.
23. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet.* 2008;9(6):465-476.
24. Eden S, Cedar H. Role of DNA methylation in the regulation of transcription. *Curr Opin Genet Dev.* 1994;4(2):255-259.
25. Mamrut S, Harony H, Sood R, et al. DNA methylation of specific CpG sites in the promoter region regulates the transcription of the mouse oxytocin receptor. *PLOS One.* 2013;8(2):e56869.
26. Sun D, Layman TS, Jeong H, et al. Genome-wide variation in DNA methylation linked to developmental stage and chromosomal suppression of recombination in white-throated sparrows. *Mol Ecol.* 2021;30(14):3453-3467.
27. Falls JB, Kopachena JG. White-throated sparrow (*Zonotrichia albicollis*). In: Poole AF, ed. *White-Throated Sparrow (Zonotrichia Albicollis), Version 1.0. In Birds of the World.* Cornell Lab of Ornithology; 2020.
28. Griffiths R, Double MC, Orr K, Dawson RJG. A DNA test to sex most birds. *Mol Ecol.* 1998;7(8):1071-1075.
29. Zinzow-Kramer WM, Horton BM, McKee CD, et al. Genes located in a chromosomal inversion are correlated with territorial song in white-throated sparrows. *Genes Brain Behav.* 2015;14(8):641-654.
30. Palkovits M. Microdissection of individual brain nuclei and areas. In: Boulton AA, Baker GB, eds. *General Neurochemical Techniques.* Humana Press; 1985:1-17.
31. Grogan KE, Horton BM, Hu Y, Maney DL. A chromosomal inversion predicts the expression of sex steroid-related genes in a species with alternative behavioral phenotypes. *Mol Cell Endocrinol.* 2019;495:110517.
32. Andrews S. FastQC: a quality control tool for high throughput sequence data [online]; 2010. Available Online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
33. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal.* 2011;17(1):10-12.
34. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012;9(4):357-359.
35. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. *Bioinformatics.* 2011;27(11):1571-1572.
36. Krueger F, Andrews SR. SNPsplit: allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res.* 2016;5:1479.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Methodol.* 1995;57(1):289-300.
38. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022. <https://www.r-project.org/>
39. Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS.* Springer; 2000.
40. Pinheiro J, Bates D, R Core Team. nlme: Linear and nonlinear mixed effects models. R package version 3. 2022. 1-157.
41. Ben-Shachar M, Lüdtke D, Makowski D. Effectsize: estimation of effect size indices and standardized parameters. *J Open Source Softw.* 2020;5(56):2815.
42. Navarro DJ. Learning statistics with R: a tutorial for psychology students and other beginners. (version 0.6) University of new South Wales. Sydney, Australia; 2015.
43. Canty A, Ripley B. Boot: bootstrap R (S-Plus) functions. R Package Version 1.3-28; 2021.
44. Davison AC, Hinkley DV. *Bootstrap Methods and their Applications.* Cambridge University Press; 1997.
45. Rubenstein DR, Skolnik H, Berrio A, Champagne FA, Phelps S, Solomon J. Sex-specific fitness effects of unpredictable early life conditions are associated with DNA methylation in the avian glucocorticoid receptor. *Mol Ecol.* 2016;25(8):1714-1728.
46. Sinha S, Jones BM, Traniello IM, et al. Behavior-related gene regulatory networks: a new level of organization in the brain. *Proc Natl Acad Sci.* 2020;117(38):23270-23279.
47. Szyf M, McGowan P, Meaney MJ. The social environment and the epigenome. *Environ Mol Mutagen.* 2008;49(1):46-60.
48. Horton BM, Hu Y, Martin CL, et al. Behavioral characterization of a white-throated sparrow homozygous for the ZAL2^m chromosomal rearrangement. *Behav Genet.* 2013;43(1):60-70.
49. Thornycroft HB. Chromosomal polymorphism in the white-throated sparrow, *Zonotrichia albicollis* (GMELIN). *Science.* 1966;154(3756):1571-1572.
50. Okhovat M, Berrio A, Wallace G, Ophir AG, Phelps SM. Sexual fidelity trade-offs promote regulatory variation in the prairie vole brain. *Science.* 2015;350(6266):1371-1374.
51. Okhovat M, Maguire SM, Phelps SM. Methylation of *avpr1a* in the cortex of wild prairie voles: effects of CpG position and polymorphism. *R Soc Open Sci.* 2017;4(1):160646.
52. Liu D, Krajniak K, Chun D, et al. VIP gene transcription is regulated by far upstream enhancer and repressor elements. *Biochem Biophys Res Commun.* 2001;284(1):211-218.
53. Bonasio R, Li Q, Lian J, et al. Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Curr Biol.* 2012;22(19):1755-1764.
54. Glastad KM, Gokhale K, Liebig J, Goodisman MAD. The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. *Sci Rep.* 2016;6(1):37110.
55. Aristizabal MJ, Anreiter I, Halldorsdottir T, et al. Biological embedding of experience: a primer on epigenetics. *Proc Natl Acad Sci.* 2020;117(38):23261-23269.
56. Day JJ, Sweatt JD. DNA methylation and memory formation. *Nat Neurosci.* 2010;13(11):1319-1323.
57. Herb BR, Shook MS, Fields CJ, Robinson GE. Defense against territorial intrusion is associated with DNA methylation changes in the honey bee brain. *BMC Genomics.* 2018;19(1):216.
58. Weaver ICG, Cervoni N, Champagne FA, et al. Epigenetic programming by maternal behavior. *Nat Neurosci.* 2004;7(8):847-854.
59. Seebacher F, Krause J. Epigenetics of social behaviour. *Trends Ecol Evol.* 2019;34(9):818-830.
60. Bentz AB, George EM, Wolf SE, et al. Experimental competition induces immediate and lasting effects on the neurogenome in free-living female birds. *Proc Natl Acad Sci.* 2021;118(13):e2016154118.
61. Stewart AJ, Hannehalli S, Plotkin JB. Why transcription factor binding sites are ten nucleotides long. *Genetics.* 2012;192(3):973-985.
62. Keshet I, Yisraeli J, Cedar H. Effect of regional DNA methylation on gene expression. *Proc Natl Acad Sci.* 1985;82(9):2560-2564.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Prichard MR, Grogan KE, Merritt JR, Root J, Maney DL. Allele-specific *cis*-regulatory methylation of the gene for vasoactive intestinal peptide in white-throated sparrows. *Genes, Brain and Behavior.* 2022;21(8):e12831. doi:10.1111/gbb.12831