Journal of Heredity, 2023, **114**, 207–218
https://doi.org/10.1093/jhered/esad006
Advance access publication 20 February 2023
Original Article

American Genetic Association

OXFORD

Original Article

Epigenetic potential: Promoter CpG content positively covaries with lifespan and is dependent on gene function among vertebrates

Elizabeth L. Sheldon[1], Aaron W. Schrey[2], M. Ellesse Lauer[2], Lynn B. Martin[1]

[1]College of Public Health, University of South Florida, Global Health and Infectious Disease Research Center and Center for Genomics, Tampa, FL, United States,
[2]Department of Biology, Georgia Southern University, Savannah, GA, United States

Address correspondence to E.L. Sheldon at the address above, or e-mail: elsheldon@usf.edu.

Corresponding Editor: Bridgett vonHoldt

Abstract

Variation in DNA methylation is associated with many ecological and life history traits, including niche breadth and lifespan. In vertebrates, DNA methylation occurs almost exclusively at "CpG" dinucleotides. Yet, how variation in the CpG content of the genome impacts organismal ecology has been largely overlooked. Here, we explore associations between promoter CpG content, lifespan and niche breadth among 60, amniote vertebrate species. The CpG content of 16 functionally relevant gene promoters was strongly, positively associated with lifespan in mammals and reptiles, but was not related to niche breadth. Possibly, by providing more substrate for CpG methylation to occur, high promoter CpG content extends the time taken for deleterious, age-related errors in CpG methylation patterns to accumulate, thereby extending lifespan. The association between CpG content and lifespan was driven by gene promoters with intermediate CpG enrichment—those known to be predisposed to regulation by methylation. Our findings provide novel support for the idea that high CpG content has been selected for in long-lived species to preserve the capacity for gene expression regulation by CpG methylation. Intriguingly, promoter CpG content was also dependent on gene function in our study; immune genes had on average 20% less CpG sites than metabolic- and stress-related genes.

Key words: DNA methylation, epigenetic age, epigenetic potential, niche breadth, phenotypic plasticity, vertebrates

Background

A key aim of ecological epigenetics is to relate patterns of epigenetic diversity to the evolutionary forces impacting organismal ecology and life history (Rey *et al.* 2020). Interspecific variation in 2 fundamental traits: niche breadth (the diversity of resources used or environments tolerated by an individual, population, or species) and lifespan (the length of time an individual lives), has recently been associated with epigenetic variation in DNA methylation (Herrera *et al.* 2012; Sexton *et al.* 2017; Horvath and Raj 2018). Yet, the proximal mechanism by which DNA methylation might influence these traits remains unclear.

The incorporation of methylation to the genome does not alter the DNA sequence, but when methylation falls within gene regulatory regions (e.g. gene promoters) it can affect phenotypic plasticity via its effects on gene expression (Boyes and Bird 1991; Bird 2002; Weaver *et al.* 2004; Klose and Bird 2006; Zhi *et al.* 2013; Lemire *et al.* 2015). Patterns of DNA methylation can be modified in response to various external cues, contributing to the ability of individuals to respond flexibly to environmental variation within their lives (Weaver *et al.* 2004; Pigliucci 2006; Lea *et al.* 2016; Leung *et al.* 2016; Sheldon et al. 2018, 2020; Pfennig 2021). Such environmentally induced phenotypic plasticity is thought to be intrinsically linked with the evolution of species-level niche breadth (niche expansion or contraction), as it largely underscores the capacity of organisms to perform flexibly across multiple environments (Via and Lande 1985; Lynch and Gabriel 1987; Gomulkiewicz and Kirkpatrick 1992; Gavrilets and Scheiner 1993; Gilchrist 1995; Pigliucci 2006; Herrera *et al.* 2012; Sexton *et al.* 2017) (however, it is recognized that the expression of a fixed phenotype that functions in diverse conditions may also foster a broad niche; Holt 2009; Fraebel *et al.* 2020).

In vertebrates, DNA methylation occurs almost exclusively at CpG dinucleotides (adjacent cytosine and guanine linked by a phosphate) (Suzuki and Bird 2008) (other sequence contexts have been detected but at much lower frequencies (Ramsahoye *et al.* 2000; Lister *et al.* 2009). Sequence variants that increase or decrease CpG content are thus extremely relevant for increasing or decreasing the potential for phenotypic plasticity via CpG methylation (Feinberg and Irizarry 2010; Kilvitis *et al.* 2017). The CpG content of cis-regulatory regions (e.g. gene promoters) has consequently been referred to as "epigenetic potential" (Kilvitis *et al.* 2017); where high epigenetic potential (more CpG sites) represents more opportunities for

Received January 12, 2023; Accepted February 7, 2023

the addition and/or removal of methylation in response to external cues, thus potentially more opportunities to regulate environmentally induced phenotypic plasticity (Auld *et al.* 2010; Botero *et al.* 2015; Kilvitis *et al.* 2017).

Positive selection for high epigenetic potential has been detected in several animal populations thought to be reliant on phenotypic plasticity for survival and reproduction (Peng *et al.* 2011; Hanson *et al.* 2022; Smith *et al.* 2022). For example, around 40% of adaptive genetic polymorphisms in hypoxia-related genes were found to modify promoter CpG content in highland human populations relying on gene expression flexibility to regulate oxygen homeostasis (Peng *et al.* 2011; Julien et al. 2017). Further, across a globally expanding population of songbirds, invasive birds moving into novel environments were found to have higher promoter CpG content in immune-related genes compared with native birds facing purportedly less environmental fluctuations (Hanson et al. 2020, 2022).

Recent experimental work has also indicated that high epigenetic potential is associated with greater inducibility and reversibility (i.e. flexibility) of gene expression (Hanson *et al.* 2021). Epigenetic potential has also been associated with fitness consequences; specifically, high CpG content in an important immune gene (Toll-like receptor 4) is positively correlated with the capacity of house sparrows (*Passer domesticus*) to resist a bacterial infection (Sheldon *et al.* In review). Because CpG content is a genetic motif, it is also an inherently heritable trait that could be subject to natural selection if it is associated with fitness effects (Hanson *et al.* 2022). CpG content could thus represent a proximal mechanism whereby the potential for epigenetic regulation of phenotypic plasticity is linked to ecological processes (e.g. niche breadth expansion; Sexton *et al.* 2017). Yet, to date, this idea has not been explored. The first aim of this study was to explore whether the promoter CpG content of selected, DNA methylation regulated genes is positively associated with interspecific niche breadth variation among vertebrates; potentially indicative of its positive selection during niche expansion.

In addition to the regulation of life-long, environmentally induced phenotypic plasticity, variation in CpG methylation also constitutes an important component of developmental plasticity and the coordination of stable cell lineage differentiation (West-Eberhard 2005). Errors in the maintenance of critical CpG methylation patterns are associated with reduced cellular function, destabilized cell fates, and various age-related disease states (Yang *et al.* 2023). In vertebrates, stereotypical errors (losses and gains of CpG methylation across the genome) tend to occur with age, and the accumulation of these errors can be used to predict chronological and biological age with unprecedented accuracy (Horvath and Raj 2018; Parrott and Bertucci 2019). By increasing the genomic substrate available for methylation to occur, it has recently been proposed that high CpG content may extend the time taken for deleterious, age-related errors in CpG methylation patterns to accumulate (Bertucci and Parrott 2020). In other words, CpG content may help explain interspecific variation in lifespan, as increased CpG content may be protective against the age-related erosion of the epigenetic landscape (Bertucci and Parrott 2020). In line with this, 2 recent studies have found positive correlations between genome-wide promoter CpG content and lifespan in mammals (McLain and Faulk 2018) and other vertebrate species (Mayne *et al.* 2019).

It is important to note that gene promoters with very high (CpG islands) or very low CpG content do not seem to contribute to variation in transcriptional activity (Weber *et al.* 2007). For example, CpG island promoters are generally unmethylated even when inactive, and promoters with low CpG content often remain active despite being methylated (possibly because the recruitment of methyl-CpG-binding domain proteins to methylated CpGs is not sufficient at low densities for transcriptional repression; Weber *et al.* 2007). Promoters with "intermediate" CpG enrichment are therefore most significant for transcriptional regulation via methylation (Weber *et al.* 2007). Accordingly, the second aim of our study was to explore whether potential associations between promoter CpG content and lifespan/niche breadth largely occur in a subset of promoters with intermediate opposed to high or low levels of CpG enrichment (Feinberg and Irizarry 2010). If associations between lifespan/niche breadth and CpG content are driven by promoters with intermediate CpG enrichment, this would bolster the premise that CpG content may play a functional role in niche-breadth/lifespan evolution among vertebrates.

In our study, we explore associations between promoter CpG content and lifespan/niche breadth among 60, amniote vertebrate species (see Methods section for why we focus on amniotes) that vary widely in niche breadth and lifespan. We predict that promoter CpG content will be positively associated with niche breadth and/or lifespan by increasing the potential for the epigenetic regulation of phenotypic plasticity and/or the maintenance of the epigenetic landscape. We further predict that if such trends are associated with functional epigenetic regulation of gene expression, they would be driven by promoters with intermediate CpG enrichment. Instead of a genome-wide analysis, here, we focus on 16, well-studied, homologous genes; genes whose expression are known to be regulated by promoter CpG methylation and whose function is likely to impact organismal fitness across multiple environments (Supplementary Table 1). Specifically, we focus on genes associated with immune, metabolic, and stress response regulation. We focus on this subset of functionally relevant genes for 2 main reasons. Firstly, our hypotheses predict that promoter CpG content may be a target for selection as it holds fitness consequences relating to the potential for epigenetic gene expression regulation. We do not, therefore, aim to identify correlations between lifespan/niche breadth and promoter CpG content in a subset of functionally irrelevant loci that are not directly regulated by promoter DNA methylation (e.g. genes in epistasis) (while correlations with such genes may be indicative of other important interactions between the genome, epigenome, and organismal ecology/life history, detangling these processes is beyond the scope of this study; Mayne *et al.* 2019). Secondly, intraspecific studies linking CpG content with organismal ecology have to date, focused on specific, functional genes (e.g. *ESPA1* [Yang *et al.* 2023] and *TLR-4* [Sheldon *et al.* In review]), here, we aim to explore whether similar trends may occur interspecifically to warrant a more general concept of "epigenetic potential."

Methods

Species and gene selection

We focus on amniotes as all types of DNA sequences (i.e. introns, exons, promoters, enhancers, etc.) are subject to methylation in this phylum, while non-vertebrates

are characterized with mosaic-like methylation patterns interspersed by unmethylated domains (Aliaga *et al.* 2019). Among vertebrates, mammalian-like isochores (i.e. regions of DNA with a high degree of uniformity in guanine and cytosine [GCs]) have been reported in sauropsids (birds and reptiles) but not teleosts and lissamphibians (Bernardi and Bernardi 1990; Wicker 2004). Compositional heterogeneity in GCs is thus likely to have evolved in the amniote ancestor (Romiguier *et al.* 2010). Consequently, to avoid capturing effects of GC content related to large phylogenetic distances (i.e. across clades), only birds, mammals, and reptiles were included in our analyses. Amphibians and fish were also excluded because these Classes were insufficiently represented in the NCBI database (see below) (Database NCBI 2016), and they tend to occupy distinct habitats from birds, reptiles, and mammals, making direct niche breadth comparisons problematic.

The specific selection of species for our study involved 4 criteria: 1) species needed to be represented in the National Center for Biotechnology Information (NCBI) (Database NCBI 2016) GenBank nucleic acid sequence database; 2) species needed to share gene homology with the other species included in our study; 3) species needed to be distantly related to others in the same Class (such that the taxonomic diversity of each Class was well represented by different Orders); and 4) species needed to possess unique life history traits, geographic distributions and habitats relative to others in the set. To meet criteria 1) and 2), we used the NCBI system "HomoloGene" to detect homologous (including paralogs and orthologs) gene sequences automatically. To gauge gene homology, we chose 3 conserved genes (*TLR3*, *BDNF*, and *MC4R*) and listed the species for which homologs were consistently present. Eighty-eight species of birds, 58 species of reptile, and 106 species of mammal had these gene homologs in the NCBI system. From these species, Orders were unequally represented, with many being represented by only 1 species and few by >30 species. As such, we included only 1 species from each available Order in our study. This approach enabled us to include 20 Mammalia, 20 Aves, and 20 Reptilia species for our study (see Supplementary Fig. 2 for all species involved in the study). For orders with >1 species, or for Classes that were not represented by 20 Orders, we chose the species that allowed us to best meet criteria 3) and 4). We conducted a power analysis using the R package SIMR in to assess whether our sample size (N = 60 species) was large enough to test detect statistically meaningful effects (Green and MacLeod 2015).

Rather than the whole genome, we focus on 16, specific, well-studied homologous genes in our study, as we sought to compare variation in CpG content in genes that have a direct relationship with DNA methylation regulated phenotypic plasticity and influence fitness (Supplementary Table 1). The specific selection of genes for our study involved 3 criteria: 1) genes needed to be homologous among the 60 species, 2) genes needed to be directly associated with functions relevant to coping with environmental variation (we focused on genes involved in energy metabolism, immunity, and stress), and 3) gene expression needed to be regulated by promoter CpG methylation. We used the "summary" and resource materials listed in the "bibliography" section of NCBI's gene search feature to identify whether a gene met the above criteria (Database NCBI 2016). We identified 16 genes from these criteria: metabolism-related genes: *cpt1A*, *FASN*, *MC4R*, *MDH1*, *PDK1*, immune-related genes: *NLRX1*, *RIG-1*, *TLR3*, *TLR4*, and *TLR5*, and stress-related genes: *BDNF*, *HSP60*, *HSP90A1*, *HSPA5*, *HSPA9*, and *HSPH1* (Supplementary Table 1).

Quantification of promoter CpG content

Gene promoters were identified in NCBI as –499 to 100 bp regions flanking the transcription start site (TSS) of each gene in NCBI. This region (500 bp upstream of the TSS) is widely recognized as the genomic region containing the gene promoter (Weber *et al.* 2007; Hanson *et al.* 2022). The number of CpG's (and CpA's—used here as a "control") in each sequence was quantified. Although Illumina sequencing does not tend to sequence well through GC-rich regions (Kim *et al.* 2021; Rhie *et al.* 2021), genome coverage graphs and missing sequence indicators (highlighted by black bars on the chromosome in the NCBI database) were visually examined before collecting the data, and in the rare instances where >5% of bases of sequence were missing in the 500 bp upstream of the TSS, the sequence was excluded from our analyses.

Life history and lifespan estimations

Average lifespan for each species was collected from "AnAge" (de Magalhães and Costa 2009). Body mass and reproductive capacity data (see equation below) were collected from "Animal Diversity" (Myers *et al.* 2022). We used "reproductive capacity" in our models, as plasticity is often associated with a species pace of life (Mazalov *et al.* 1996; Kokko and Sutherland 2001; Dammhahn *et al.* 2018; Ratikainen and Kokko 2019). For birds and reptiles, reproductive capacity was calculated using the equation below, a similar equation was used for mammals using gestation length, interbirth interval and litter size in exchange of variables related to egg laying:

$$\text{Reproductive capacity} = \frac{(\text{lifespan} - \text{age at}}{[(\text{incubation length} + \text{interlay}}$$

Niche breadth estimations

To calculate a proxy of niche breadth we estimated and summed proxies of each species global, geographic distribution and habitat use (values toward 1 being having a broad niche breadth and values toward 0 being having a narrow niche breadth). To estimate the geographic distribution of a species, species ranges were visualized on a global map from the IUCN database (Supplementary Fig. 1). A Mercator projection grid was then manually cast over each range map and the number of squares the range spanned were counted (Supplementary Fig. 1). This total was then divided by the total number of squares on the Mercator projection (175 squares) grid to calculate a proxy of geographic distribution. To estimate habitat use, another putative driver of plasticity, eleven habitat categories (e.g. urban, rainforest, desert, tundra, wetland, polar, etc.) were identified from the main IUCN habitat classes. The total categories occupied by each species was then discerned from IUCN databases, and divided by the total number (eleven) of habitat Classes to obtain a proxy for habitat use. "Geographic distribution" and "habitat usage" were moderately correlated (R = 0.49). When both terms were included in the phylogenetic generalized

least squares (PGLS) model, neither variable showed statistically significant associations with CpG content ($P > 0.89$ for all vertebrate classes)—equivalent to trends detected with "niche breadth" considered as a singular term. We therefore did not include each variable as a separate fixed effects in our analyses to reduce the risk of multicollinearity (associations between independent variables that reduces the reliability of statistical inference).

Statistical analyses

To test whether CpG content was associated with species traits, we first compared promoter CpG content with promoter CpA content among species. We avoided comparisons between CpG and GpG as the likelihood of the occurrence of these dinucleotides are conditional; they are composed of the same bases, so the presence of 1 dinucleotide (CpG or GpC) increases the probability of the other pro/preceding it by a factor of 4. We then asked whether CpG content was random with respect to phylogeny by testing for phylogenetic signals with Pagel's lambda (λ) (Pagel 1999) using the R package GEIGER (Harmon *et al.* 2008). The influence of phylogeny increases with lambda from 0 (no phylogenetic signal) to 1 (strong phylogenetic signal). To determine whether lambda was significantly different from zero, we used a likelihood ratio test in "R" (Harmon *et al.* 2008).

We then used a PGLS model to investigate the associations between lifespan and niche breadth and, CpG content while controlling for potential phylogenetic effects. Data were analyzed using the "ape" and "nlme" packages in R. An Ornstein–Uhlenbeck model of evolution was assumed in all models, which is a modification of the Brownian motion model that assumes the amount of evolutionary change along a branch decreases exponentially. Phylogenetic trees were created using timetree.org (Kumar *et al.* 2017). Species variables included as fixed effects in our PGLS model were: niche breadth, reproductive capacity, and a residual measure of lifespan that controlled for body mass (lifespan and body mass were positively correlated in birds ($r = 0.501$), reptiles ($r = 0.718$), and mammals ($r = 0.884$)).

Association between CpG content among Phyla

Visual inspection of Fig. 3 showed that bird species are characterized by either high or low (but not intermediate) CpG content across all 16 genes, we therefore categorized each vertebrate species into either a "high" or "low" group depending on whether average CpG content across 16 genes in that species was above or below the mean across all genes for all species. We next asked whether variation in CpG content could be used to group species into high or low categories, and what genes were most important in distinguishing these categories. We did this by conducting 3 linear discriminant function analyses for each vertebrate class with the "R" package "MASS." These analyses predicted the probability of belonging to high or low categories based on the CpG content of either immune-, metabolic-, or stress-related genes.

Association between CpG content among gene function

We then conducted a linear mixed model to determine whether CpG content differed significantly among metabolic-, stress-, and immune-related genes. Gene function (i.e. immune, stress, or metabolism) was included as a fixed effect, and Vertebrate Class was included as a random effect.

Association between promoter CpG content and promoter CpG enrichment

Promoters with an intermediate CpG enrichment are more likely to be involved in transcriptional regulation (possibly because active transcriptional repression is not sufficient at low methylation cytosine densities, and because CpG islands tend to be unmethylated). Therefore, we classified gene promoters into 3 categories to distinguish promoters with high (CpG islands), intermediate, and low CpG enrichment. We calculated a CpG ratio using the following formula from Weber *et al.* (2007):

$$\text{CpG ratio} = \frac{(\text{No. of CpGs} \times \text{No. of base pairs})}{(\text{No. of } C's \times \text{No. of } G's)}$$

The 3 categories of promoters were then determined as follows: HCPs (high-CpG promoters) contain a 500-bp area with CpG ratio above 0.75 and GC content above 55%; LCPs (low-CpG promoters) do not contain a 500-bp area with a CpG ratio above 0.48; and ICPs (intermediate CpG promoters) are neither HCPs nor LCPs (Weber *et al.* 2007). A linear mixed model was conducted to explore the relationship between CpG content and lifespan for each CpG promoter category. Lifespan was included as a fixed effect and gene and Vertebrate Class were included as random effects.

Results

Promoter CpG content was minimally influenced by phylogenetic effects

Average CpG content across 16 genes was not associated with phylogeny (Pagel's lambda results did not differ significantly from zero; Table 1) (Pagel 1999). Average CpG content ranged from 6.9 to 56.4 CpGs across all species in this study, and this variation was most extensive in Aves (note, the more variable red-blue gradient across this class compared with the other clades; Fig. 1). Average CpG content in gene promoters was also lower and more variable than CpA content in Aves, Mammalia, and Reptilia species (Table 2), indicating that equivalently sized genetic motifs differed depending on their potential for methylation. Our power analysis suggested that a

Table 1. Interspecific variation in promoter CpG content for 3 Classes of vertebrates.

Class	Promoter CpG content		
	Pagel's lambda	Log-likelihood	*P*
All Classes	0.014	−235.364	0.819
Reptilia	<0.001	−59.551	0.999
Aves	<0.001	−87.119	0.999
Mammalia	1.457	−64.687	0.135

Pagel's lambda (Pagel 1999) indicates weak phylogenetic effects on CpG content across all Classes (note, lambda values >1 for Mammalia indicates greater evolution at the "root" of the tree that decreases toward the tips, although this effect was nonsignificant) (Wang and Clarke 2014).

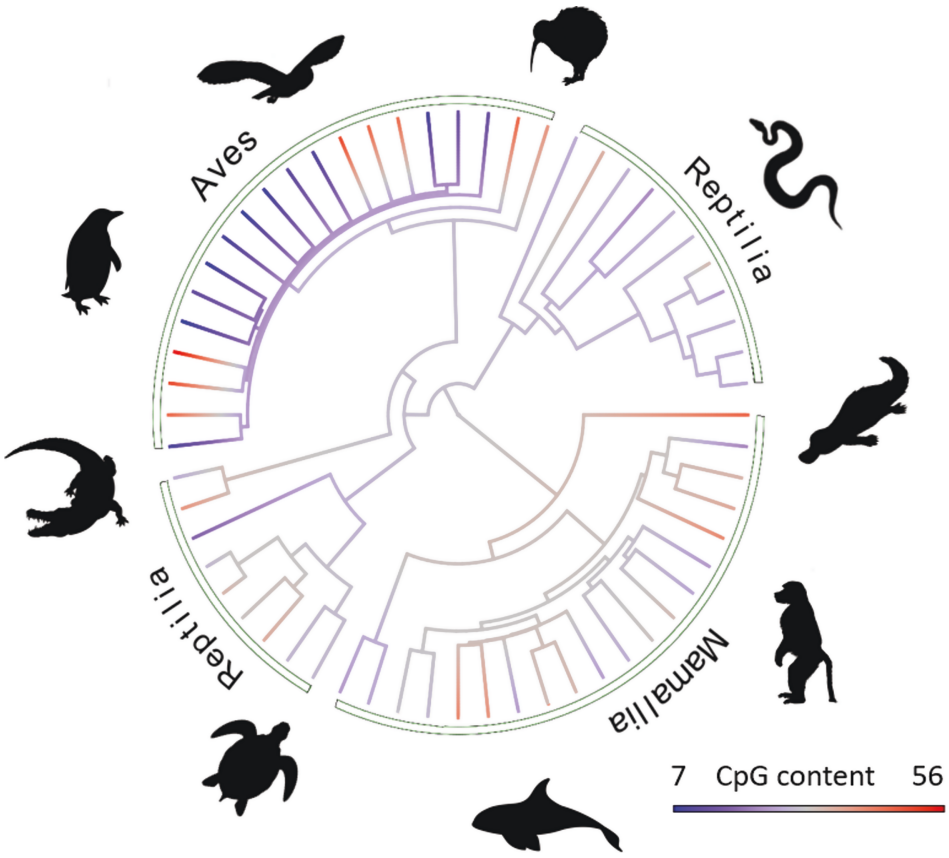

Fig. 1. Average CpG content across 16 genes for 3 vertebrate Classes: Reptilia, Mammalia, and Aves. Species silhouettes are positioned relative to their associated branch location on the phylogenetic tree. The color shading of branches denotes the level of average CpG content for each species, irrespective of genes (from blue: low CpG content, to red: high CpG content, see inset color bar for effect size).

Table 2. CpG content in all 16 genes among 3 vertebrate Classes.

Class	CpG content (SD)	Range	CpA content (SD)	Range
Aves	27.87 (29.48)	1 to 108	35.41 (8.91)	12 to 65
Mammalia	32.69 (23.65)	1 to 104	35.42 (9.08)	18 to 77
Reptilia	29.76 (19.87)	1 to 79	36.85 (8.54)	9 to 59

CpA content has also been included to compare variation with a dinucleotide with no potential for methylation, and hence no capacity to regulate phenotypic plasticity.

sample size of $N = 30$ is required to detect an effect of our fixed factors on telomere length change with 83% power and 0.05 significance level, thus our sample size of $N = 60$ was sufficient.

CpG content was positively associated with lifespan across mammals and reptiles

In mammals, lifespan was the only life history trait associated with CpG content (lifespan: estimate = 0.508706 [Std Error = 0.107], $t = 3.069394$, $P = 0.008$) (Fig. 2a). Niche breadth (estimate = 5.771 [Std Error = 4.861], $t = 1.187$, $P = 0.252$), and reproductive capacity (estimate = 0.078 [Std Error = 0.148], $t = 0.529$, $P = 0.604$) were not associated with CpG content among mammals.

In reptiles, lifespan was also the only life history trait associated with CpG content (lifespan estimate = 0.299 [Std Error = 0.051], $t = 5.78$, $P < 0.001$) (Fig. 2b). Reproductive capacity (estimate = −0.002 [Std Error = 0.002], $t = -1.073$, $P = 0.299$) and niche breadth (estimate = 10.697 [Std Error = 3.787], $t = 2.824$, $P = 0.072$) were not associated with CpG content.

CpG content was not associated with avian lifespan (estimate = 0.136 [Std Error = 0.388], $t = 0.352$, $P = 0.729$), niche breadth (estimate = 56.58 [Std Error = 47.881], $t = 1.181$, $P = 0.255$), or reproductive capacity (estimate = −0.197 [Std Error = 0.215], $t = -0.917$, $P = 0.373$). Because CpG content was distinctly distributed in birds (see below), we assessed whether this distribution was affected by genome length. However, no such associations were detected (estimate = <0.001 [Std Error = <0.001], $t = 0.068$, $P = 0.945$).

CpG content is uniquely distributed into high and low clusters in birds

Twelve bird species (blue dots in Fig. 3) had low average CpG content (below the Class average) and low variation in CpG content, but 8 bird species (red dots in Fig. 3) had high CpG content (above the Class average) and high variation in CpG content among genes. These patterns were largely driven by differences among stress and metabolic genes; for immune genes, all species of birds had relatively low CpG content and variation (Figs. 3 and 4). This dichotomy was not present in mammals and reptiles (note, the red and blue dots in Fig. 3 overlap for mammals and reptiles). Indeed, a linear discriminant analysis found that CpG content in metabolic- and stress-related genes could assign species of birds into high

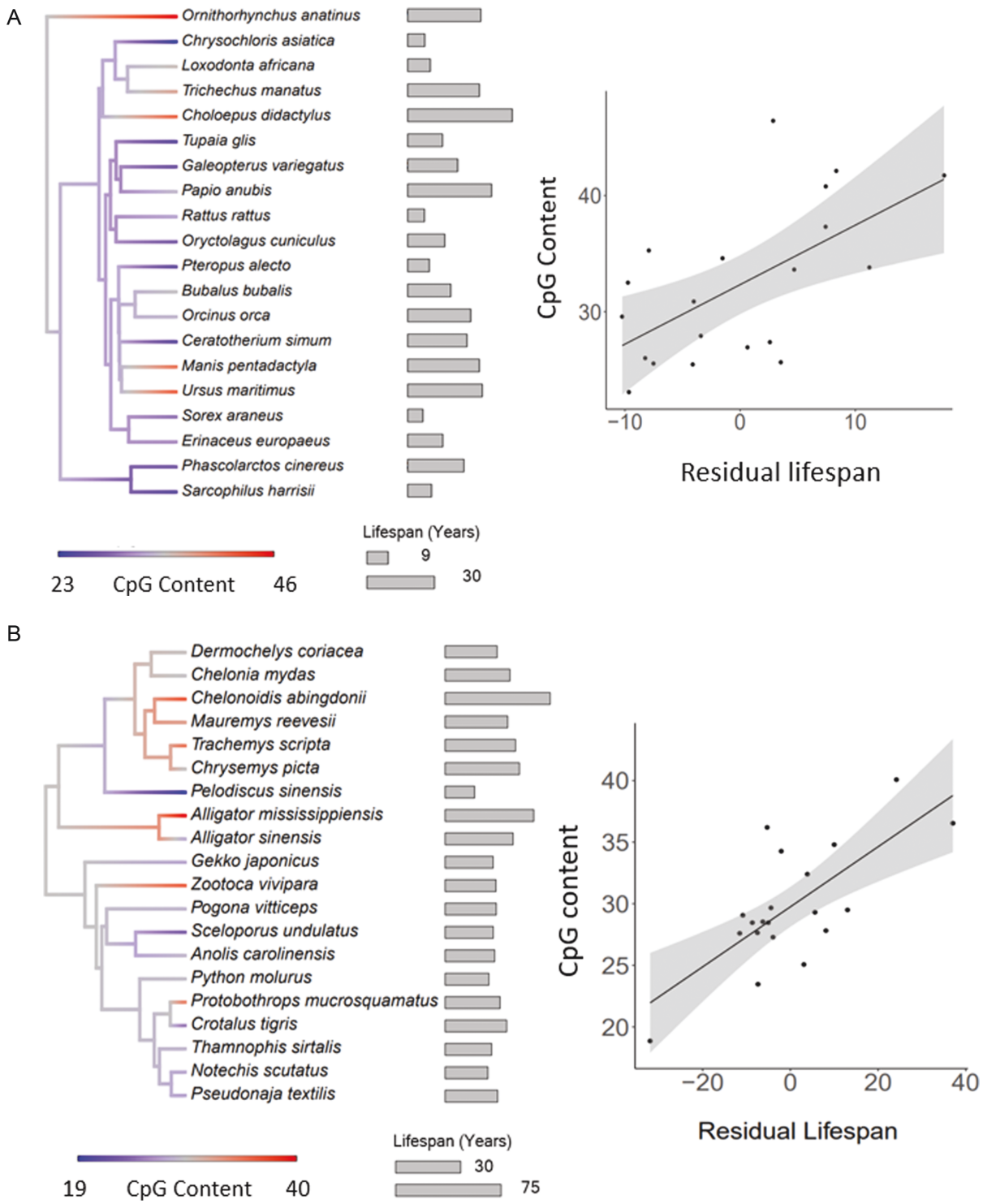

Fig. 2. A phylogenetic tree with the evolution of CpG content mapped as a continuous trait (see inset color bar for effect sizes) for a) each mammal and b) each reptile species. Each species' average lifespan (years) is plotted next to its names as a bar chart. The positive relationship between CpG content and lifespan is also illustrated as a scatterplot (with a regression line and 95% confidence intervals) with lifespan illustrated as a residual value in relation to body mass

and low categories with 100% accuracy. This same approach applied to mammals and reptiles had only 57% and 85% accuracy, respectively (Supplementary Table 3). CpG content in immune genes performed poorly in assigning species of any class to high or low CpG content categories (Supplementary Table 3).

Variation in CpG content was contingent on gene function

CpG content in metabolic- and stress-related genes were greater and more variable than those of immune-related genes among all classes (immune vs. metabolism: estimate = 18.209 [Std Error = 1.937], t = 9.399, $P < 0.001$; immune vs. stress:

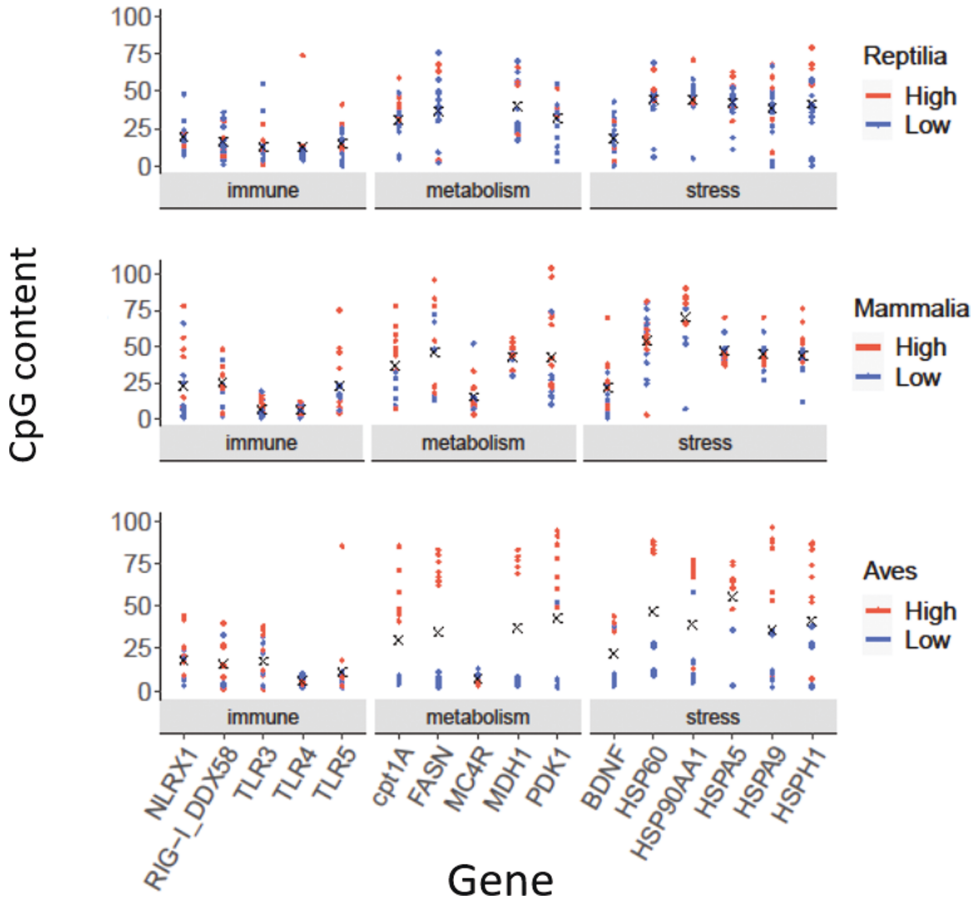

Fig. 3. CpG content in each gene promoter for each species (red dots = species with average CpG content in the upper half of the range, blue dots = species with average CpG content in the lower half of the range). Black "X"'s represents average CpG densities for each gene. Note: *MCR4* was not available in reptiles due to poor sequence coverage of this gene.

estimate = 22.683 [Std Error = 1.791], $t = 12.662$, $P < 0.001$) (Fig. 4). 17.39% of variation in CpG content was driven by gene categories, whereas only 1.538% of variation was driven by Vertebrate Class.

Correlations between lifespan and CpG content was driven by promoters with intermediate CpG enrichment

The association between CpG content and lifespan was primarily driven by promoters with intermediate CpG enrichment (low, intermediate, and high CpG enrichment categories were calculated following Weber *et al.* (2007), and the equation in the Methods section of this study) (Fig. 5). Across promoters with low and high CpG enrichment, lifespan was not associated with CpG content (estimate = 0.095 [Std Error = 0.062], $t = 1.523$, $P = 0.129$ and estimate = 0.019 [Std Error = 0.053], $t = 0.365$, $P = 0.716$, respectively). Across promoters with intermediate CpG enrichment, lifespan was positively associated with CpG content across all vertebrate groups (estimate = 0.136 [Std Error = 0.063], $t = 2.157$, $P = 0.032$).

Discussion

In our study, we tested whether promoter CpG content in a subset of functional genes was associated with niche breadth and/or lifespan among 60 amniote vertebrate species. We found significant interspecific variation in CpG content that was especially pronounced and uniquely distributed among birds. When controlling for phylogenetic relatedness among species, CpG content was positively associated with lifespan in mammals and reptiles, but not birds. Importantly, this trend was driven by promoters with intermediate CpG enrichment that are predisposed to de novo methylation and transcriptional regulation (Weber *et al.* 2007). Our findings support the hypothesis that high CpG content has been selected for in long-lived species, potentially to preserve the capacity for gene expression regulation by CpG methylation (however further work is necessary before causal relationships between CpG content, methylation status and gene expression can be confirmed; Sheldon *et al.* In review). We did not find conclusive evidence that CpG content was related to niche breadth (i.e. geographic distribution and habitat utilization). However, we did find that CpG content differed among types of genes: immune genes had much lower and less variable CpG content than metabolic- and stress-related genes.

Promoter CpG content, niche breadth, and phenotypic plasticity

In line with the concept of epigenetic potential (Kilvitis *et al.* 2017; Hanson *et al.* 2021), we predicted that CpG content would be positively correlated with species-level niche breadth. However, we did not detect such a relationship in our study. Environmentally induced and stochastic epigenetic

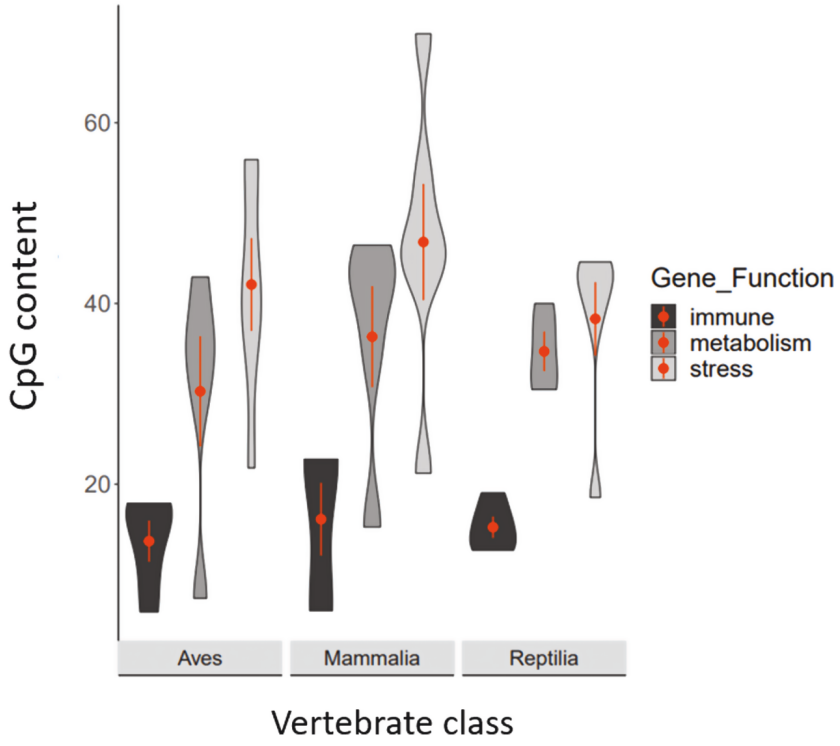

Fig. 4. A violin plot illustrating variation in CpG content across the 3 gene categories for each vertebrate Class. Orange dots represent the median values with error bars spanning the first to third quartiles.

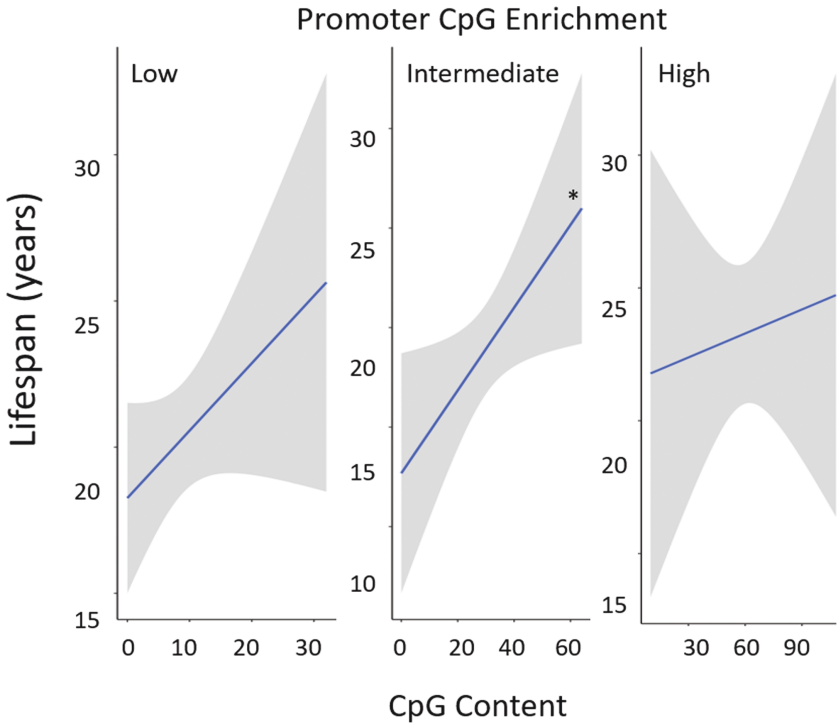

Fig. 5. The relationship between CpG content and lifespan across all vertebrates for gene promoters with low (n = 229), intermediate (n = 244)*, and high (n = 331) CpG content groups (regression line and 95% confidence intervals are illustrated). * indicates a significant relationship.

modifications are 2 fundamentally different mechanisms enabling phenotypic plasticity via variation in DNA methylation. Environmentally induced epigenetic variation involves the perception of an environmental signal, and the subsequent epigenetic silencing or activation of specific genes resulting in an environment-specific phenotype (Rando and Verstrepen 2007; Chinnusamy and Zhu 2009; Beldade *et al.* 2011; Verhoeven and Preite 2014). This form of plasticity is generally advantageous for dealing with predictable environmental changes (Lachmann and Jablonka 1996; DeWitt *et al.* 1998;

Reed *et al.* 2011; Scheiner and Holt 2012). Stochastic epigenetic changes, by contrast, have been proposed to be among the potential mechanisms underlying diversified bet-hedging strategies, advantageous to organisms needing to cope with unpredictable environments (Piggot 2010; Casadesús and Low 2013; Herman *et al.* 2014; Vogt 2015).

The proxies of niche breadth used in this study; geographic distribution and number of habitats utilized, capture the extent to which a species has adapted its phenotype to different environments, but do not necessarily reflect the extent to which an *individual* can cope with environmental unpredictability (Grantham *et al.* 2016; Leung *et al.* 2016). Perhaps we did not detect an association between interspecific niche breadth and CpG content because CpG content as we measured it here is more relevant for stochastic epigenetic modifications than environmentally mediated plasticity. This perspective may explain why intraspecific studies have previously detected higher CpG content in individuals experiencing unpredictable conditions (e.g. range expanding populations of birds) than those experiencing more predictable conditions (e.g. birds at an established range core) (Hanson *et al.* 2022). CpG content might yet be an important capacitor of phenotypic plasticity, but its effects might be subtler than we initially expected. Ultimately, more research is needed to disentangle exactly what aspects of plasticity CpG content may accommodate.

It is also worth noting that the *number* of CpG sites in gene promoters is not the only form of CpG content involved in the mediation of plasticity. For example, the *position* of the CpG site in the gene promoter (rather than the CpG content per se) may constitute a mechanism through which single-nucleotide polymorphisms could affect plasticity via epigenetics (Hellman and Chess 2007; Dayeh *et al.* 2013). Moreover, the specific location of promoters of all genes in all species is apt to vary, but we had no way to account for this in our study. Weak associations between niche breadth and CpG content may also have been missed due to sample size ($n = 16$ genes in our study). A complimentary study including more functional genes may detect more subtle associations between CpG content and niche breadth.

Promoter CpG content is positively correlated with lifespan

In our study, lifespan was significantly, positively correlated with lifespan among mammals and reptiles. Similar trends have also been detected in 2 other studies: in mammals (Mayne *et al.* 2019) and other vertebrate species (McLain and Faulk 2018). It has recently been suggested that high CpG content could extend lifespan by delaying the time taken for deleterious losses of epigenetic information (i.e. DNA methylation profiles associated with transcriptional regulation) to accumulate (Mills *et al.* 1999; Oberdoerffer *et al.* 2008; Bertucci and Parrott 2020). The present study supports this reasoning by showing that the correlation between CpG content and lifespan was driven by promoters with intermediate CpG enrichment—known to be targets for de novo methylation and transcriptional activity via methylation (Weber *et al.* 2007). Alternatively, or additionally, high CpG content could maintain the potential for methylation state alterations in response to external cues (plasticity), which may also delay biological aging (as the mismatch between the environment and a non-plastic individual's phenotype is likely to increase as time passes (Ratikainen and Kokko 2019). It remains unclear why the relationship between lifespan and CpG content was not detected among birds in our study. Although, similarly, Mayne *et al.* (2019) found that promoter CpG-lifespan estimates were considerably less accurate in birds compared with mammals and reptiles.

CpG dinucleotides are generally depleted from the amniotic genome (apart from at CpG islands) (Belle *et al.* 2004), which is thought to reflect inherent mutability of methylated cytosines to thymine (a methylated cytosine is only 1 hydrolytic deamination away from mutation to a thymine; Coulondre *et al.* 1978). Because CpG sites are exceptionally prone to mutations (Coulondre *et al.* 1978), CpG content could reflect the extent of historical CpG-related mutations experienced by a lineage, which might have little to do with selection for lifespan and more to do with the control of transposons, or speciation (Pértille *et al.* 2019). However, repair pathways do exist to selectively remove thymine from a T:G mismatch in the context of CpG dinucleotides (Hendrich *et al.* 1999). Consequently, while CpG content may be depleted by cytosine methylation, CpG content may also reflect differential selection pressures during the evolution of life history. This has been demonstrated in corals (*Platygyra daedalea*); thermally stressed corals showed *higher* CpG content in the promoter of a thermotolerance gene than unstressed corals even though the thermally stressed corals showed *higher* levels of DNA methylation (Smith *et al.* 2022).

Dichotomy of promoter CpG content among birds

Just under half (8) of the bird species in our study had high average CpG content (~45 CpGs per promoter), and just over half (12) had low CpG content (~15 CpGs); a trend driven by the CpG content of metabolic- and stress-related genes (all birds had an average of 15 CpGs in immune genes, see below). The average CpG content of each promoter among birds was either lower or higher than the average CpG content of each promoter among mammalian and reptilian species (i.e. the average promoter CpG content for birds fell at the extremes of the rest of the data set). This pattern is difficult to interpret, and it was unexpected. We speculated that distinctions between avian versus mammalian and reptilian average CpG content may be driven by genome size. Among amniotes, birds have much smaller genomes than mammals and reptiles; mammal and reptilian genomes range between 1.0 and 8.2 Gb while bird genomes are 0.91 to 1.3 Gb (Jarvis *et al.* 2014). Smaller genomes may have less regulatory variation, thus high CpG content may constitute an additional mechanism for regulating gene expression not afforded by other regulatory processes available to large genomes. Yet, among the bird species we studied, genome size was unrelated to CpG content, so this possibility seems unlikely. Moreover, birds had both higher *and* lower CpG content than mammals and reptiles. Thus, "pan-avian" genomic traits would be unlikely to lend insight into the high/low distribution of CpG content among birds.

We consequently considered whether avian lifestyles corresponded with high vs. low CpG content. A study by Jarvis *et al.* (2014) showed that waterbirds (Aequornithia) had lower "GC" content than land birds (Telluraves). Whereas this metric (GC) is not necessarily related to the epigenetically relevant CpG content we measured, the 5 waterbirds in our study did have lower CpG content on average (CpG

= 16.34, SD = 15) than the 15 land birds (CpG = 27, SD= 20). Further, water/land bird categorizations (in addition to other categorizations e.g. predatory vs. non-predatory, vocal vs. non-vocal learning, etc.) did not map to CpG content variation. Future work, with greater sample sizes across more avian life history strategies is clearly necessary to foster additional investigation.

Promoter CpG content and gene function

Immune genes had approximately 20% lower and less variable CpG content than metabolic- and stress-related genes, a trend that was apparent across all 3 vertebrate Classes. The immune genes included in our analyses were not an arbitrary choice, but have a comparable function immunologically, being involved in the recognition and distinction of pathogens from normal cells. Other immune genes with distinct functions (i.e. cytokines as signaling molecules or antimicrobial peptides as direct antagonists of pathogens) might not have revealed the same variation in CpG content. Nevertheless, for these particular immune genes, our results may suggest that their signaling pathways are strictly and finely regulated by DNA methylation among taxa. Indeed, aberrant signaling of these defense response immune genes is likely to alter immune homeostasis with potential implications for immunopathology conditions (e.g. autoimmune responses) (Klasing 2004; Martin *et al.* 2017). That said, the CpG content of 1 "immune gene"; *TLR-4*, has been shown to range from 6 to 10 CpG sites in 1 species—the house sparrow (*P. domesticus*), indicating moderate intraspecific variation in immune gene CpG content within 1 species. Together, these observations could suggest that within the seemingly narrow range of immune gene CpG content among species, variation at the individual level is still prominent, and contingent on the selective pressures at play. Higher and more variable CpG content in metabolic- and stress-related genes, could suggest that the regulation by DNA methylation is less constrained (Zagouri *et al.* 2012; Zabkiewicz *et al.* 2014). It is, however, premature to make any particular conclusions about patterns of CpG content among genes; much larger studies with many more genes of various functions are critical.

Conclusions

Our study, and related ones (McLain and Faulk 2018; Mayne *et al.* 2019), detect positive associations between CpG content and longevity at the species level, supporting a role for CpG content in mediating interspecific variation in lifespan (Nabholz *et al.* 2011; Bertucci and Parrott 2020). Our study bolsters and expands upon this avenue of research by showing that correlations between CpG content and lifespan were driven by transcriptionally relevant promoters—those with intermediate CpG enrichment (Weber *et al.* 2007). Our study suggests that promoter CpG content is not associated with interspecific differences in niche breadth, possibly because they are masked by the stochastic genetic divergences involved in speciation. Further work, assessing associations between ecological/life history traits and genes on a promoter-by-promoter level may be useful to detect whether specific gene features (e.g. gene function, nucleotide composition, position on the genome, etc.) are more likely to be subject to selection on promoter CpG content. Two findings from our study also remain somewhat inconclusive: 1) that across 60 vertebrate species, birds had either particularly low or high CpG content, and 2) that among the 16 genes analyzed, immune gene promoters tended to have much fewer and less variable CpG content than metabolic- or stress-related genes. Altogether, while CpG content varies interspecifically and intergenically in some intelligible ways, it is apparent that there is much more to learn about this form of genomic variation.

Supplementary material

Supplementary material is available at *Journal of Heredity* online.

Funding

This work was supported by a National Science Foundation grant: NSF-IOS 2027040.

Acknowledgments

We would like to acknowledge the contributions of Mark Ravinet and Erich Jarvis, whose suggestions and insight were valuable to the interpretation of our data.

Data availability

Data for this study can be found in the supplementary materials and the publicly available repository: Sheldon, Elizabeth (2022): Selection for promoter CpG content_DATA.csv. Figshare. Dataset. https://doi.org/10.6084/m9.figshare.20714404.v1.

References

Aliaga B, Bulla I, Mouahid G, Duval D, Grunau C. Universality of the DNA methylation codes in Eukaryotes. *Sci Rep*. 2019;9:1–11.

Auld J, Agrawal A, Relyea R. Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proc R Soc Lond B Biol Sci*. 2010;277:503–511.

Beldade P, Mateus AR, Keller RA. Evolution and molecular mechanisms of adaptive developmental plasticity. *Mol Ecol*. 2011;20(7):1347–1363.

Belle EM, Duret L, Galtier N, Eyre-Walker A. The decline of isochores in mammals: an assessment of the GC content variation along the mammalian phylogeny. *J Mol Evol*. 2004;58:653–660. doi:10.1007/s00239-004-2587-x

Bernardi G, Bernardi G. Compositional patterns in the nuclear genome of cold-blooded vertebrates. *J Mol Evol*. 1990;31(4):265–281.

Bertucci EM, Parrott BB. Is CpG density the link between epigenetic aging and lifespan? *Trends Genet*. 2020;36(10):725–727.

Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 2002;16:6–21.

Botero C, Weissing F, Wright J, Rubenstein D. Evolutionary tipping points in the capacity to adapt to environmental change. *Proc Natl Acad Sci USA*. 2015;112:184–189.

Boyes J, Bird A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*. 1991;64(6):1123–1134.

Casadesús J, Low DA. Programmed heterogeneity: epigenetic mechanisms in bacteria. *J Biol Chem*. 2013;288(20):13929–13935.

Chinnusamy V, Zhu JK. RNA-directed DNA methylation and demethylation in plants. *Sci China C Life Sci*. 2009;52(4):331–343.

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature*. 1978;274(5673):775–780.

Dammhahn M, Dingemanse NJ, Niemelä PT, Réale D. Pace-of-life syndromes: a framework for the adaptive integration of behaviour, physiology and life history. *Behav Ecol Sociobiol*. 2018;72:62.

Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2016;44(Database issue):D7–D19.

Dayeh TA, Olsson AH, Volkov P, Almgren P, Rönn T, Ling C. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia*. 2013;56(5):1036–1046.

de Magalhães JP, Costa J. A database of vertebrate longevity records and their relation to other life-history traits. *J Evol Biol*. 2009;22(8):1770–1774.

DeWitt TJ, Sih A, Wilson DS. Costs and limits of phenotypic plasticity. *Trends Ecol Evol*. 1998;13(2):77–81.

Feinberg AP, Irizarry RA. Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc Natl Acad Sci USA*. 2010;107(suppl 1):1757–1764.

Fraebel DT, Gowda K, Mani M, Kuehn S. Evolution of Generalists by Phenotypic Plasticity. *iScience*. 2020;23(11):101678. PMID: 33163936; PMCID: PMC7600391.

Gavrilets S, Scheiner SM. The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *J Evol Biol*. 1993;6:31–48.

Gilchrist GW. Specialists and generalists in changing environments. I. Fitness landscapes of thermal sensitivity. *Am Nat*. 1995;146(2):252–270.

Gomulkiewicz R, Kirkpatrick M. Quantitative genetics and the evolution of reaction norms. *Evolution*. 1992;46:390–411.

Grantham ME, Antonio CJ, O'Neil BR, Zhan YX, Brisson JA. A case for a joint strategy of diversified bet hedging and plasticity in the pea aphid wing polyphenism. *Biol Lett*. 2016;12(10):20160654.

Green P, MacLeod CJ. SIMR: an R package for power analysis of generalised linear mixed models by simulation. *Methods Ecol Evol*. 2015;7(4):493–498. doi:10/1111/2041-210X.12504

Hanson HE, Koussayer B, Kilvitis HJ, Schrey AW, Maddox JD, Martin LB. Epigenetic potential in native and introduced populations of house sparrows (*Passer domesticus*). *Integr Comp Biol*. 2020:60(6):1458–1468.

Hanson HE, Zimmer C, Koussayer B, Schrey AW, Maddox JD, Martin LB. Epigenetic potential affects immune gene expression in house sparrows. *J Exp Biol*. 2021;224(6):jeb238451.

Hanson HE, Wang C, Schrey AW, Liebl AL, Ravinet M, Jiang RHY, Martin LB. Epigenetic potential and DNA methylation in an ongoing house sparrow (Passer domesticus) Range Expansion. *Am Nat*. 2022;200(5):662–674. doi:10.1086/720950. Epub 2022 Sep 21. PMID: 36260844.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. GEIGER: investigating evolutionary radiations. *Bioinformatics*. 2008;24:129–131.

Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. *Science*. 2007;315(5815):1141–1143.

Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*. 1999;401:301–304.

Herman JJ, Spencer HG, Donohue K, Sultan SE. How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. *Evolution*. 2014;68(3):632–643.

Herrera CM, Pozo MI, Bazaga P. Jack of all nectars, master of most: DNA methylation and the epigenetic basis of niche width in a flower-living yeast. *Mol Ecol*. 2012;21(11):2602–2616.

Holt RD. Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proc Natl Acad Sci USA*. 2009;106:19659–19665.

Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*. 2018;19(6):371–384.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*. 2014;346(6215):1320–1331.

Kilvitis HJ, Hanson H, Schrey AW, Martin LB. Epigenetic potential as a mechanism of phenotypic plasticity in vertebrate range expansions. *Integr Comp Biol*. 2017;57(2):385–395.

Kim M, Xi H, Park S, Yun Y, Park J. Genome-wide comparative analyses of GATA transcription factors among seven *Populus* genomes. *Sci Rep*. 2021;11:16578.

Klasing KC. The costs of immunity. *Dong Wu Xue Bao*. 2004;50(6):961–969.

Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci*. 2006;31:89–97.

Kokko H, Sutherland WJ. Ecological traps in changing environments: ecological and evolutionary consequences of a behaviorally mediated Allee effect. *Evol Ecol Res*. 2001;3:537–551.

Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol Biol Evol*. 2017;34(7):1812–1819.

Lachmann M, Jablonka E. The inheritance of phenotypes: an adaptation to fluctuating environments. *J Theor Biol*. 1996;181:1–9.

Lea AJ, Altmann J, Alberts SC, Tung J. Resource base influences genome-wide DNA methylation levels in wild baboons (*Papio cynocephalus*). *Mol Ecol*. 2016;25(8):1681–1696.

Lemire M, Zaidi SHE, Ban M, Ge B, Aïssi D, Germain M, Kassam I, Wang M, Zanke BW, Gagnon F, et al. Long-range epigenetic regulation is conferred by genetic variation located at thousands of independent loci. *Nat Commun*. 2015;6:6326.

Leung C, Breton S, Angers B. Facing environmental predictability with different sources of epigenetic variation. *Ecol Evol*. 2016;6(15):5234–5245.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315–322.

Lynch M, Gabriel W. Environmental tolerance. *Am Nat*. 1987;129:283–303.

Martin LB, Kilvitis HJ, Brace AJ, Cooper L, Haussmann MF, Mutati A, Fasanello V, O'Brien S, Ardia DR. Costs of immunity and their role in the range expansion of the house sparrow in Kenya. *J Exp Biol*. 2017;220(12):2228–2235.

Mayne B, Berry O, Davies C, Farley J, Jarman S. A genomic predictor of lifespan in vertebrates. *Sci Rep*. 2019;9:17866.

Mazalov V, Perrin N, Dombrovsky Y. Adaptive search and information updating in sequential mate choice. *Am Nat*. 1996;148:123–137.

McLain AT, Faulk C. The evolution of CpG density and lifespan in conserved primate and mammalian promoters. *Aging (Milano)*. 2018;10(4):561–572.

Mills KD, Sinclair DA, Guarente L. MEC1-dependent redistribution of the Sir3 silencing protein from telomeres to DNA double-strand breaks. *Cell*. 1999;97(5):609–620.

Myers P, et al. The Animal Diversity Web (online). 2022. https://animaldiversity.org

Nabholz B, Künstner A, Wang R, Jarvis ED, Ellegren H. Dynamic evolution of base composition: causes and consequences in avian phylogenomics. *Mol Biol Evol*. 2011;28(8):2197–2210.

Oberdoerffer P, Michan S, McVay M, Mostoslavsky R, Vann J, Park S-K, Hartlerode A, Stegmuller J, Hafner A, Loerch P, et al. SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. *Cell*. 2008;135(5):907–918.

Pagel M. Inferring the historical patterns of biological evolution. *Nature*. 1999;401(6756):877–884.

Parrott BB, Bertucci EM. Epigenetic aging clocks in ecology and evolution. *Trends Ecol Evol*. 2019;34(9):767–770.

Peng Y, Yang Z, Zhang H, Cui C, Qi X, Luo X, Tao X, Wu T, Chen H, Shi H, et al. Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas. *Mol Biol Evol*. 2011;28:1075–1081. doi:10.1093/molbev/msq290

Pértille F, Da Silva VH, Johansson AM, Lindström T, Wright D, Coutinho LL, Jensen P, Guerrero-Bosagna C. Mutation dynamics of CpG dinucleotides during a recent event of vertebrate

diversification. *Epigenetics*. 2019;14(7):685–707. doi:10.1080/15592294.2019.1609868. Epub 2019 May 9. PMID: 31070073; PMCID: PMC6557589.
Pfennig DW. *Phenotypic plasticity & evolution: causes, consequences, controversies*. Boca Raton (FL): CRC Press/Taylor & Francis; 2021.
Piggot P. Epigenetic switching: bacteria hedge bets about staying or moving. *Curr Biol*. 2010;20(11):R480–R482.
Pigliucci M. Evolution of phenotypic plasticity: where are we going now? *Trends Ecol Evol*. 2006;20(9):481–486.
Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci USA*. 2000;97(10):5237–5242.
Rando OJ, Verstrepen KJ. Timescales of genetic and epigenetic inheritance. *Cell*. 2007;128(4):655–668.
Ratikainen II, Kokko H. The coevolution of lifespan and reversible plasticity. *Nat Commun*. 2019;10:538.
Reed TE, Schindler DE, Waples RS. Interacting effects of phenotypic plasticity and evolution on population persistence in a changing climate. *Conserv Biol*. 2011;25:56–63.
Rey O, Eizaguirre C, Angers B, Balazar-soares M, Sagonas K, Prunier JG, Blanchet S. Linking epigenetics and biological conservation: towards a conservation epigenetics perspective. *Funct Ecol*. 2020;34:414–427. https://doi.org/10.1111/1365-2435.13429
Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021;592(7856):737–746.
Romiguier J, Ranwez V, Douzery EJ, Galtier N. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Res*. 2010;20:1001–1009.
Scheiner SM, Holt RD. The genetics of phenotypic plasticity. X. Variation versus uncertainty. *Ecol Evol*. 2012;2(4):751–767.
Sexton JP, Montiel J, Shay JE, Stephens MR, Slatyer RA. Evolution of ecological niche breadth. *Annu Rev Ecol Evol Syst*. 2017;48(1):183–206.
Sheldon EL, Schrey AW, Ragsdale AK, Griffith SC. Brood size influences patterns of DNA methylation in wild Zebra Finches (Taeniopygia guttata). *The Auk*. 2018;135(4):1113–1122. https://doi.org/10.1642/AUK-18-61.1
Sheldon EL, Schrey AW, Hurley LL, Griffith SC. Dynamic changes in DNA methylation during postnatal development in zebra finches *Taeniopygia guttata* exposed to different temperatures. *J Avian Biol*. 2020;51(5):e02294.
Sheldon EL, Zimmer C, Hanson H, Koussayer B, Schrey A, Reese D, Wiglet P, Wedley A, Martin LB. In review. High epigenetic potential in the Toll-like receptor 4 promoter is associated with resistance to a Salmonella enterica infection; implications for invasion dynamics.
Smith EG, Hazzouri KM, Choi JY, Delaney P, Al-Kharafi M, Howells EJ, Aranda M, Burt JA. Signatures of selection underpinning rapid coral adaptation to the world's warmest reefs. *Sci Adv*. 2022;8(2):eabl7287. doi:10.1126/sciadv.abl8287. Epub 2022 Jan 12. PMID: 35020424.
Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet*. 2008;9(6):466–467.
Verhoeven KJ, Preite V. Epigenetic variation in asexually reproducing organisms. *Evolution*. 2014;68(3):644–655.
Via S, Lande R. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution*. 1985;39:505–522.
Vogt G. Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. *J Biosci*. 2015;40:159–204.
Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ. Epigenetic programming by maternal behavior. *Nat Neurosci*. 2004;7(8):847–854.
Weber M, Hellmann I, Stadler M, Ramos L, Pääbo S, Rebhan M, Schübeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*. 2007;39(4):457–466. doi:10.1038/ng1990. Epub 2007 Mar 4. PMID: 17334365.
West-Eberhard MJ. Developmental plasticity and the origin of species differences. *Proc Natl Acad Sci USA*. 2005;102(suppl 1):6543–6549.
Wicker T. International Chicken Genome Sequencing Consortium. *Nature*. 2004;432(7018):695–716.
Yang JH, Hayano M, Griffin PT, Amorim JA, Bonkowski MS, Apostolides JK, Salfati EL, Blanchette M, Munding EM, Bhakta M, et al. Loss of epigenetic information as a cause of mammalian aging. *Cell*. 2023;186(2):305–326.e27. doi:10.1016/j.cell.2022.12.027. Epub 2023 Jan 12. PMID: 36638792.
Zabkiewicz J, Pearn L, Hills RK, Morgan RG, Tonks A, Burnett AK, Darley RL. The PDK1 master kinase is over-expressed in acute myeloid leukemia and promotes PKC-mediated survival of leukemic blasts. *Haematologica*. 2014;99(5):858–864.
Zagouri F, Bournakis E, Koutsoukos K, Papadimitriou CA. Heat shock protein 90 (hsp90) expression and breast cancer. *Pharmaceuticals*. 2012;5(9):1008–1020.
Zhi D, Aslibekyan S, Irvin MR, Claas SA, Borecki IB, Ordovas JM, Absher DM, Arnett DK. SNPs located at CpG sites modulate genome-epigenome interaction. *Epigenetics*. 2013;8(8):802–806.