



# Genome-wide epigenetic isolation by environment in a widespread *Anolis* lizard

Guinevere O. U. Wogan<sup>1</sup> | Michael L. Yuan<sup>1</sup> | D. Luke Mahler<sup>2</sup> | Ian J. Wang<sup>1</sup>

<sup>1</sup>Department of Environmental Science, Policy, and Management, College of Natural Resources, University of California, Berkeley, CA, USA

<sup>2</sup>Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada

## Correspondence

Ian J. Wang, Department of Environmental Science, Policy, and Management, College of Natural Resources, University of California, Berkeley, CA, USA.  
Email: ianwang@berkeley.edu

## Funding information

Society for the Study of Amphibians and Reptiles; NSF Dimensions of Biodiversity, Grant/Award Number: DEB-1542534

## Abstract

Epigenetic changes can provide a pathway for organisms to respond to local environmental conditions by influencing gene expression. However, we still know little about the spatial distribution of epigenetic variation in natural systems, how it relates to the distribution of genetic variation and the environmental structure of the landscape, and the processes that generate and maintain it. Studies examining spatial patterns of genetic and epigenetic variation can provide valuable insights into how ecological and population processes contribute to epigenetic divergence across heterogeneous landscapes. Here, we perform a comparative analysis of spatial genetic and epigenetic variation based on 8,459 single nucleotide polymorphisms (SNPs) and 8,580 single methylation variants (SMVs) from eight populations of the Puerto Rican crested anole, *Anolis cristatellus*, an abundant lizard in the adaptive radiations of anoles on the Greater Antilles that occupies a diverse range of habitats. Using generalized dissimilarity modelling and multiple matrix regression, we found that genome-wide epigenetic differentiation is strongly correlated with environmental divergence, even after controlling for the underlying genetic structure. We also detected significant associations between key environmental variables and 96 SMVs, including 42 located in promoter regions or gene bodies. Our results suggest an environmental basis for population-level epigenetic differentiation in this system and contribute to better understanding how environmental gradients structure epigenetic variation in nature.

## KEYWORDS

anole, DNA methylation, isolation by distance, landscape epigenetics, landscape genetics, population structure

## 1 | INTRODUCTION

Mounting evidence suggests that epigenetic changes may play important roles in organismal responses to spatial and temporal environmental variation, yet we still know little about how epigenetic variation is distributed within and among populations in natural systems (Foust et al., 2016; Herrera, Medrano, & Bazaga, 2016; Hu & Barrett, 2017; Schmid et al., 2018; Verhoeven, vonHoldt, & Sork, 2016). DNA methylation, the most commonly studied epigenetic modification, is closely linked to the environment in two important

ways. First, environmental stimuli can induce intragenerational epigenetic changes, and environmentally induced methylation can vary under different environmental conditions, both genome-wide and at specific loci (Downen et al., 2012; Metzger & Schulte, 2017; Putnam, Davidson, & Gates, 2016). Second, DNA methylation can affect gene expression and therefore contribute to phenotypic plasticity across different environments (Artemov et al., 2017; Greenspoon & Spencer, 2018; Jaenisch & Bird, 2003). These two mechanisms can also be related—for instance, evidence suggests that environmentally induced DNA methylation may underlie a portion of normal

phenotypic plasticity and physiological responses to different environments (Artemov et al., 2017; Downen et al., 2012; Jaenisch & Bird, 2003; Rubenstein et al., 2016). Hence, epigenetic variation should be closely tied to environmental variation, and a variety of spatial processes can shape epigenetic profiles.

However, patterns of epigenetic variation should also, at least to some extent, reflect underlying patterns of genetic variation (Dubin et al., 2015; Foust et al., 2016; Herrera et al., 2016). Because of epigenetic inheritance, the same processes that contribute to generating spatial patterns of genetic structure can act on epigenetic variation as well—epigenetic alleles (“epialleles”) should move between populations with gene flow (Herrera et al., 2016; Herrera, Medrano, & Bazaga, 2017), and random drift in isolation could reduce epigenetic variance within populations (Richards, Bossdorf, & Verhoeven, 2010; Trucchi et al., 2016). However, epigenetic inheritance is imperfect, meaning that epigenetic variation is “reset” to some degree between generations (Angers, Castonguay, & Massicotte, 2010; Jablonka & Raz, 2009; Schmitz et al., 2013), and rates of change in epigenetic markers (“epimutations”) are expected to be higher than genetic mutation rates (Angers et al., 2010; Jiang, Mithani, et al., 2014; Trucchi et al., 2016). So, ostensibly, population processes can lead to patterns of epigenetic isolation by distance (IBD), and differences between genetic IBD and epigenetic IBD can be attributed to processes acting on epigenetic, but not genetic, variation (Herrera et al., 2016).

Theoretically, then, various ecological and evolutionary processes can also generate patterns of epigenetic isolation by environment, the correlation of epigenetic and environmental distances (Herrera et al., 2017). As with genetic isolation by environment (IBE; Sexton, Hangartner, & Hoffmann, 2014; Wang & Bradburd, 2014), epigenetic IBE could result from biased dispersal (Wang & Bradburd, 2014) or selection against migrants from divergent environments (Herrera et al., 2017; Sexton et al., 2014), if either is influenced by phenotypes affected by epigenetic variation. In both cases, reduced dispersal between different environments would lead to neutral divergence in epigenetic profiles through epigenetic drift (Richards et al., 2010; Trucchi et al., 2016). Additionally, because epigenetic changes can be environmentally induced (Angers et al., 2010; Radford et al., 2014; Verhoeven et al., 2016), epigenetic profiles may diverge between environments if different conditions contribute to differentially induced effects (Downen et al., 2012; Dubin et al., 2015; Herman, Spencer, Donohue, & Sultan, 2014). These differences could be heritable or re-induced every generation, and epigenetic IBE could result under this scenario with no differences in fitness. Indeed, a few studies have now demonstrated local adaptation related to epigenetic variation (e.g., Alakärppä et al., 2018; Dubin et al., 2015; He et al., 2018; Platt, Gugger, Pellegrini, & Sork, 2015) and close associations between epigenetic variants and environmental gradients in a variety of natural systems (e.g., Foust et al., 2016; Gugger, Fitz-Gibbon, Pellegrini, & Sork, 2016; Keller, Lasky, & Yi, 2016). Others, however, have found overall conservation of epigenetic profiles between populations inhabiting very different conditions, suggesting molecular or developmental constraints on

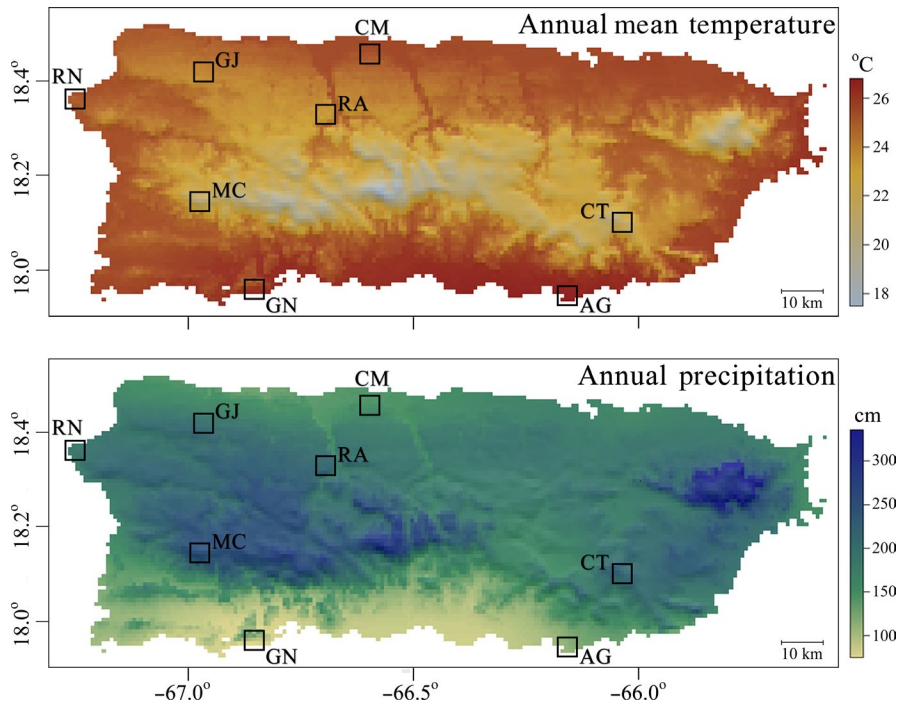
epigenetic variation (e.g., Trucchi et al., 2016). So, when and where patterns of epigenetic IBE emerge and what environmental factors are involved remain open questions, and studies comparing spatial patterns of genetic and epigenetic variation across heterogeneous landscapes are powerful for quantifying the signatures of epigenetic responses to environmental variation (Foust et al., 2016; Gugger et al., 2016; Herrera et al., 2016; Schmitz et al., 2013).

Here, we examine spatial genetic and epigenetic variation in the Puerto Rican crested anole (*Anolis cristatellus*), from eight populations distributed across a wide range of habitats, to identify the geographical and environmental factors contributing to population divergence. This species exhibits genetic IBD and IBE (Wang, Glor, & Losos, 2013), environmentally structured morphological variation (Gunderson, Siegel, & Leal, 2011; Leal & Fleishman, 2004; Winchell, Reynolds, Prado-Irwin, Puente-Rolón, & Revell, 2016), and phenotypic plasticity in response to local climate conditions (Kolbe, VanMiddlesworth, Losin, Dappen, & Losos, 2012; Perry, Dmi'el, & Lazell, 2000), making it an excellent system in which to investigate the drivers of spatial epigenetic variation. Specifically, we test whether epigenetic IBD and IBE are present even when controlling for any underlying genetic structure, under the hypothesis that epigenetic-specific characteristics—such as the responsiveness of DNA methylation to the environment, which could result in increased IBE, and its higher “epimutation” rate, which could decrease IBD—could lead to deviations in patterns of spatial genetic and epigenetic variation. We evaluate both genome-wide and locus-specific variation to quantify the effects of geographical and environmental variables acting broadly on the whole genomic background as well as differentially across different regions of the genome.

## 2 | MATERIALS AND METHODS

### 2.1 | Study system and sample collection

The Puerto Rican crested anole, *Anolis cristatellus*, is a medium-sized (50–80 mm snout to vent length) arboreal lizard with an island-wide distribution across Puerto Rico. Classified as a trunk-ground ecomorph in the adaptive radiations of anoles on the Greater Antilles (Losos, 2009; Mahler, Ingram, Revell, & Losos, 2013; Williams, 1983), *A. cristatellus* occupies broad surfaces, typically tree trunks, close to the ground and has a stocky build, large head and relatively long limbs (Henderson & Powell, 2009). Puerto Rico contains a diverse range of habitats, from montane tropical wet forests to lower montane semideciduous woodlands to coastal shrublands and dry forests, and *A. cristatellus* is found in nearly all environments on the island, including disturbed and anthropogenic habitats (Henderson & Powell, 2009; Winchell et al., 2016). Previous research on *A. cristatellus* has found evidence of genetic IBD and IBE in mitochondrial DNA (mtDNA) (Wang et al., 2013) and shifts in ecologically important morphological, behavioural and physiological traits between different habitats (Gunderson et al., 2011; Leal & Fleishman, 2004; Otero, Huey, & Gorman, 2015; Winchell et al., 2016), suggesting a



**FIGURE 1** Sampling localities for eight populations of *Anolis cristatellus* on maps of annual mean temperature (top) and total annual precipitation (bottom), superimposed on shaded relief maps, from Puerto Rico [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

variety of processes may drive population divergence between environments in this system.

We collected 79 adult male *A. cristatellus* during a 12-day period in June 2017 from eight localities from across Puerto Rico (Figure 1). These localities include seven protected forests and an area of relatively intact natural forest along the northwestern coast: Aguirre State Forest (AG), Cambalache State Forest (CM), Carite State Forest (CT), Guajataca State Forest (GJ), Guanica State Forest (GN), Maricao State Forest (MC), Rio Abajo State Forest (RA) and the Rincon peninsula (RN). We collected 10 specimens at each locality except for Rincon, from which we collected nine. These sites span a gradient from mesic to xeric habitat, from the montane wet forests in the central mountains (Carite and Maricao) to the coastal dry forests along the southern coast (Aguirre and Guanica). We recorded the GPS coordinates where each specimen was collected using handheld GPS devices. From each specimen, we removed liver tissue and preserved it in RNAlater for genetic and epigenetic analysis.

## 2.2 | Epigenetic sequencing

We extracted whole genomic DNA from liver samples using the GeneJET Genomic DNA Purification Kit (Thermo Scientific). Because methylation profiles are highly tissue-specific (Lokk et al., 2014; Schilling & Rehli, 2007; Vanyushin, Mazin, Vasilyev, & Belozersky, 1973), liver was chosen because of its physiological importance in thermoregulation through energy storage and release (Rui, 2014) and because changes in gene expression in liver tissue have been linked to climate adaptation in anoles (Campbell-Staton et al., 2017). We characterized genome-wide CpG methylation profiles using reduced representation bisulfite sequencing (RRBS; Meissner

et al., 2005), following the protocol of Boyle et al. (2012). RRBS relies on enzymatic digestion using *MspI* to target fragments that begin and/or end with CpG sites, thus enriching for potentially methylated genomic regions. Like other methods that rely on bisulfite conversion, RRBS uses sodium bisulfite to convert unmethylated cytosines to uracils. These converted uracils are then transcribed as thymines following PCR amplification. Site-specific methylation levels can then be inferred using the ratio of cytosine reads to the total number of reads mapped to a reference genome.

In brief, we spiked 180 ng of DNA from each sample with 0.1 ng of unmethylated lambda phage DNA to determine sample-specific bisulfite conversion efficiencies. Samples were digested at 37°C overnight using the restriction enzyme *MspI* (New England Biosystems), followed by end repair and A-tailing. We then ligated NEXTFlex Bisulfite-Seq Barcodes (BIOO Scientific), which are fully methylated and thus resist bisulfite conversion, to the libraries and then pooled 17–19 libraries based on their unique barcodes. Pooled libraries were bisulfite-converted using the EpiTect Fast Bisulfite Conversion Kit (Qiagen). We then PCR-amplified our final libraries for 16 cycles using PfuTurbo Cx Hotstart Taq (Agilent Technologies), which overcomes uracil stalling that would otherwise occur due to bisulfite conversion. We performed single-end sequencing (1 × 100) for each pool on nine lanes of the Illumina HiSeq 2500 at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

## 2.3 | Sequence alignment and variant calling

We assessed read quality using FASTQC (Andrews, 2010). To generate methylation profiles, we used informatics approaches designed

specifically for use with bisulfite sequencing. We used BS-SEEKER2 (Guo et al., 2013) to build an indexed genome from the soft-masked *Anolis carolinensis* genome (AnoCar2; Alföldi et al., 2011) and then used end-to-end mapping in the BS-SEEKER2 integrated version of BOWTIE2 (Langmead & Salzberg, 2012) to individually align reads for each sample to the AnoCar2 genome. We called methylation levels for each cytosine at single base resolution; methylation levels are measured as the frequency of methylation for a given site, which corresponds to the proportion of cells with methylated bases at that locus. Using the METHYLKIT R package (Akalin et al., 2012), we then identified single methylation variants (SMVs) as sites that had greater than a 10% range in percent methylation among samples and were not missing from more than 10% of individuals, with at least 10× coverage for a minimum of five individuals per population, following parameters from similar studies (e.g., Gugger et al., 2016). To ensure that C/T genetic polymorphisms were not misinterpreted as variation in methylation (because unmethylated cytosines are read as thymines in RRBS), we identified loci that were potentially C/T polymorphic as any for which we found an A read on the strand opposite the CpG site and conservatively removed all of them from our SMV data set.

To generate a single nucleotide polymorphism (SNP) data set for analyses of genetic variation, we analysed the read data using a customized RADseq pipeline adapted from previous work (Bi et al., 2012; Singhal, 2013). First, we assessed read quality using FASTQC (Andrews, 2010), filtered each sample for the MspI cut site, allowing a maximum 1-bp mismatch, then cleaned the data by removing polymerase chain reaction (PCR) duplicates and low complexity reads with TRIMMOMATIC (Bolger, Lohse, & Usadel, 2014), and trimmed the adapters using SKEWER (Jiang, Lei, Ding, & Zhu, 2014). We removed reads matching contaminants using BOWTIE2 (Langmead & Salzberg, 2012). We used CD-HIT (Li & Godzik, 2006) to cluster reads with a 0.96 similarity index and used a generalized vertebrate repeat masking scheme with REPEATMASKER (Chen, 2004). We then created a pseudoreference genome across the samples and mapped clusters to it using BOWTIE2 (Langmead & Salzberg, 2012). We performed indel-realignment with GATK (McKenna et al., 2010) and then identified variants in SAMTOOLS (Li et al., 2009), using AnoCar2 as a reference genome. We set parameters such that we retained the mapping quality, used a minimum base quality score of 20, counted orphans, retained read depth, calculated base quality on the fly, computed genotype likelihoods to a bcf file, and used a faidx-indexed reference file. We piped this to BCFTOOLS for variant calling, used the consensus caller with *p*-value threshold set to .1, and retained one SNP per locus (Li et al., 2009).

We used SNP-CLEANER to filter variant sites (<https://github.com/tplinderth/>) using an even coverage filter, which controls for exceptionally high or low read depth in any sample, and an exact test of Hardy-Weinberg equilibrium (HWE) with a minimum *p*-value of .0001 (Fumagalli, Vieira, Linderth, & Nielsen, 2014). For the variant sites that passed all filters, we then performed Bayesian SNP and individual genotype calling in ANGSD (Korneliussen, Albrechtsen, &

Nielsen, 2014), using a folded site frequency spectrum and AnoCar2 as a reference genome. We set a *p*-value of  $1 \times 10^{-6}$  for calling SNPs and a 0.95 posterior probability threshold for calling genotypes. Because unmethylated cytosines are read as thymines in RRBS, we removed all C/T polymorphic loci to ensure that our SNP data set contained only true genetic polymorphisms.

We used a pseudoreference genome for mapping clusters and AnoCar2 for calling SNPs because, for nonmodel organisms, mapping to a divergent reference genome can introduce mapping biases that can impact downstream analyses. For example, reads may be filtered as low quality and removed when in fact they may simply represent a region of genome divergence between the reference and the focal taxon, thus reducing inferred genomic differentiation (Sarver et al., 2017). Pseudoreference genomes have been demonstrated to reduce alignment error and downstream variant detection, as the reads used to generate the assembly are aligned back onto themselves (Bi et al., 2012; Singhal, 2013).

## 2.4 | GIS data layers

We downloaded 19 bioclimatic data layers with 30 arc-sec resolution (~1 km) from the WorldClim2 database (<https://www.worldclim.org>; Fick & Hijmans, 2017). These data layers represent annual trends (e.g., annual mean temperature), seasonality (e.g., precipitation seasonality), and extreme or limiting environmental factors (e.g., precipitation of driest month) for a set of ecologically relevant temperature and precipitation variables. We also downloaded a set of data layers for the enhanced vegetation index (EVI) and the normalized difference vegetation index (NDVI) at 1-km spatial resolution from the MODIS database (<https://modis.gsfc.nasa.gov>). These data layers represent two commonly used vegetation metrics that provide consistent spatial and temporal comparisons of vegetation canopy greenness and structure, derived from red, near-infrared and blue wavebands collected by remote sensing satellites. NDVI measures the density of greenness over a land surface, and EVI is a similar metric that uses additional wavelengths of reflected light to minimize canopy-soil variations and improve sensitivity over dense vegetation conditions. These data layers are available at 16-day and monthly temporal resolution. Because vegetation changes seasonally, we chose monthly vegetation layers corresponding to the peak of the dry season and peak of the wet season, based on average monthly precipitation in Puerto Rico. Hence, we included four vegetation layers in our data set: dry season EVI, wet season EVI, dry season NDVI and wet season NDVI.

Because bioclimatic and vegetation variables often covary, we reduced the collinearity in our set of 23 environmental layers by performing principal components analysis (PCA) for rasters in the RSTOOLBOX R package (Leutner, Horning, Schwald-Willmann, & Hijmans, 2019). The raster PCA produces a set of principal component rasters (PC rasters) that represent the orthogonal transformations of the original data layers. We analysed the scree plot of the

total variance in the data explained by each PC to determine the number of PC rasters to retain for downstream analyses.

## 2.5 | SNP and SMV outlier detection

To identify SNPs and SMVs with allele or methylation frequencies that are statistical outliers after correcting for population structure, we performed a genome scan using the `PCADAPT` R package (Luu, Bazin, & Blum, 2017). The `PCADAPT` approach uses PCA to characterize population structure and then calculates the Mahalanobis distance for each SNP or SMV in principal component space to identify candidate loci that are outliers with respect to the genomic background. Simulation data suggest that `PCADAPT` is robust to hierarchical population structure and outperforms competing methods in cases of admixture and population divergence (Luu et al., 2017). SNP loci identified as outliers by this method are potentially under selection, and SMV loci are potentially influenced by different forces than those operating on the whole epigenomic background. For both tests, we identified statistically significant outliers after performing a false discovery rate (FDR) correction with a cutoff of 1%. To evaluate whether different regions of the genome exhibit different levels of epigenetic divergence, we constructed Manhattan plots with the log of the  $p$ -value for each SMV, based on the `PCADAPT` analysis, at its position in the genome using a custom R script.

For downstream analyses, we partitioned our SNP and SMV data into outlier and nonoutlier data sets. The nonoutlier data sets were used to calculate genetic and epigenetic structure, and we used the outlier SMV data set to test for locus-specific associations between epigenetic and environmental variation. Methylation that falls outside of promoter regions or gene bodies is not known to affect gene expression (Jones, 2012; Lou et al., 2014; Zemach, McDaniel, Silva, & Zilberman, 2010), and therefore environmental associations with SMVs in intergenic regions are more likely to be spurious or transient. So, we characterized the genomic context of each outlier SMV identified by `PCADAPT` by comparing its genomic location to the annotated *A. carolinensis* genome using the `GENOMATION` R package (Akalın, Franke, Vlahoviček, Mason, & Schübeler, 2015). Because not all gene elements will be conserved between *A. carolinensis* and *A. cristatellus*, particularly promoter regions, the putative genomic context for each SMV should be interpreted with some caution, but this information can help to identify any major patterns between environmental associations and genomic context. We also identified the gene and putative function associated with each of the outlier SMVs that were located in promoter or gene body (exon and intron) regions, based on the *A. carolinensis* genome, using the Ensembl genome browser (Zerbino et al., 2018).

## 2.6 | Genetic and epigenetic structure

To characterize population genetic structure, we calculated pairwise  $F_{ST}$  values between each of our study localities from our set

of putatively neutral SNPs using the `HIERFSTAT` R package (Goudet & Jombart, 2015). To retrieve an analogous metric for population epigenetic structure, we calculated pairwise  $\Phi_{ST}$  values from our set of nonoutlier SMVs using hierarchical analysis of molecular variance (AMOVA) (Foust et al., 2016) in the `PEGAS` R package (Paradis, 2010). We also investigated whether the genomic context of SMV sites affected epigenetic differentiation by characterizing the genomic context for each nonoutlier SMV, as we did for the outlier SMVs above, and then calculating population pairwise  $\Phi_{ST}$  values for exon, intron, promoter and intergenic sites. We compared the resulting  $\Phi_{ST}$  matrices to each other and to the matrix for all nonoutlier SMVs using Mantel tests in the `VEGAN` R package (Oksanen, Kindt, Legendre, Ohara, & Stevens, 2007).

To examine spatial genetic structure, we generated a genetic covariance matrix and performed a genetic PCA using the `ADEGENET` R package (Jombart, 2008) and performed analysis of population structure and admixture in `STRUCTURE` (Pritchard, Stephens, & Donnelly, 2000). We ran an admixture model with a 10,000-step burn-in followed by a 35,000-step Markov chain Monte Carlo (MCMC) procedure and performed 10 replicates of each  $k$  value from  $k = 2$  to  $k = 10$ . We checked for convergence between chains and between runs and used `STRUCTURE HARVESTER` (Earl & vonHoldt, 2012) and `CLUMPP` (Jakobsson & Rosenberg, 2007) to aggregate replicates for each  $k$  value and `DISTRUCT` (Rosenberg, 2004) to visualize the individual admixture probabilities.

## 2.7 | Isolation by distance and environment

To quantify the environmental and geographical drivers of spatial variation in genetic and epigenetic composition, we used generalized dissimilarity modelling (GDM), a form of nonlinear matrix regression analysis (Ferrier, Manion, Elith, & Richardson, 2007; Fitzpatrick & Keller, 2015), and multiple matrix regression with randomization (MMRR) analysis (Wang, 2013). We used these complementary analyses to test the null hypothesis that spatial epigenetic structure should simply reflect spatial genetic structure against alternative hypotheses that there is significant epigenetic IBD or IBE even after controlling for any underlying genetic structure.

In GDM, predictor variables are first transformed using a series of I-spline basis functions, and then models are fit using maximum-likelihood estimation (Ferrier et al., 2007). Variables are standardized so their resulting coefficients can be compared, and GDM is robust to collinearity among predictor variables. In a GDM, the coefficient for each variable describes the proportion of compositional turnover explained by that variable and is determined by the maximum height of its I-spline (Ferrier et al., 2007; Fitzpatrick & Keller, 2015). The slope of the I-spline indicates the rate of compositional turnover and how this rate varies at any point along the gradient concerned, while holding all other variables constant (Landesman, Nelson, & Fitzpatrick, 2014). Hence, GDM is a powerful method for distinguishing between the effects of environmental dissimilarity (IBE) and geographical distance (IBD) on genetic and epigenetic structure



because it accounts for nonstationary rates of turnover along environmental gradients (Warren, Cardillo, Rosauer, & Bolnick, 2014).

We performed the GDM analysis using the `GDM` R package (Manion, Lisk, Ferrier, Nieto-Lugilde, & Fitzpatrick, 2016). GDM fits a generalized linear model of the form:

$$-\ln(1 - d_{ij}) = \alpha_0 + \sum_{p=1}^n |f_p(x_{pi}) - f_p(x_{pj})|$$

where  $d$  is any distance between localities  $i$  and  $j$ ,  $\alpha_0$  is the intercept,  $p$  is the number of covariates, and  $f_p(x)$  are the I-spline-transformed predictor variables (Fitzpatrick & Keller, 2015). For each GDM, we used the raster PCs and pairwise geographical distances as predictor variables. For the raster PCs, we extracted values from each raster for the coordinates of each locality. For the geographical distances, we calculated Euclidean distances between localities and minimum topographic distances that account for the overland distance imposed by topographic relief, using the `topoDistance` R package (Wang, 2019). We fitted models with either Euclidean distance or topographic distance separately, keeping all other variables the same, to compare their effects. For the GDM analysis of genetic variation we used the matrix of pairwise  $F_{ST}$  values as the response variable, and for the analysis of epigenetic variation we used the matrix of pairwise  $\Phi_{ST}$  values. Finally, because epigenetic variation can depend on genetic variation, we also fit a GDM with  $\Phi_{ST}$  as the response variable and with  $F_{ST}$  as a predictor, in addition to the environmental (PC) variables and geographical distances.

To identify the optimal model for each GDM analysis and to evaluate the statistical significance of the model and each predictor variable, we used a backward elimination and variable permutation model selection procedure implemented in the `GDM` package (Ferrier et al., 2007; Manion et al., 2016). This procedure begins with the full model containing all predictor variables (geographical distance and environmental PCs 1–5, plus  $F_{ST}$  for the epigenetic analysis) then iteratively removes the variable with the lowest coefficient and recalculates the model fit and the significance values. Under the permutation procedure, model significance is evaluated by permuting all of the predictor variables, refitting the model under each permutation to generate a null distribution of deviance-explained values (model fit scores), and then comparing the model for the original data to the distribution derived from the permutations. The significance of each predictor variable is evaluated by permuting each variable individually to generate a null distribution of the change in deviance explained for the model and then comparing the contribution of each variable to the model against the null distribution. The final result is a model that retains only the most important predictor variables, with significance values for the whole model and for each predictor (Manion et al., 2016). We used 999 permutations in each of the permutation tests.

After fitting the GDM to these data, we visualized spatial patterns of genetic and epigenetic turnover (changes in genetic or epigenetic composition) by projecting the model onto the PC rasters. This process generates a new raster with a colour value assigned to each cell based on its predicted genetic or epigenetic composition.

Greater differences in the colours between cells indicate greater predicted genetic or epigenetic differences. Finally, to better understand the shape of any nonlinear relationships between the response and predictor variables, we plotted the I-spline basis functions for each predictor variable in the final models.

We performed the MMRR analysis using the “MMRR” function in R (Wang, 2013). MMRR performs multiple linear regression on distance matrices using permutations of the response variable to assess significance because of the nonindependence of distance matrix data. As with the GDM analyses, we applied a backward elimination procedure to perform variable selection, starting with the full set of predictor variables and then removing the variable with the lowest coefficient and refitting the model until only statistically significant variables ( $p < .05$ ) remained. All tests of significance were performed with 999 permutations, and we performed MMRR analyses on epigenetic structure with and without  $F_{ST}$  as a predictor variable to match our GDM analyses. Following the GDM and MMRR analyses, we plotted  $F_{ST}$ ,  $\Phi_{ST}$  and residual  $\Phi_{ST}$  (after partialling out  $F_{ST}$ ) against geographical distance and environmental distance to illustrate the patterns of IBD and IBE in our data.

## 2.8 | Genetic- and epigenetic-environmental associations

To quantify associations between SNPs or SMVs and environmental factors, we used generalized linear mixed models (GLMMs) implemented in the `spAMM` (spatial mixed models) package in R (Rousset & Ferdy, 2014). Significant associations could result from divergent selection between environments or direct influences of the environment on methylation levels at a locus (Gugger et al., 2016), and linear mixed models have been used to identify gene-environment associations in both genetic and epigenetic data (Gugger et al., 2016; Lobréaux & Melodelima, 2015; Yoder et al., 2014). GLMMs adjust for the correlation structure among populations using a kinship matrix estimated from the data (Lobréaux & Melodelima, 2015; Rousset & Ferdy, 2014; Sul & Eskin, 2013). For the SMVs, we used a kinship matrix based on the correlations in methylation profiles between populations, constructed in `METHYLKIT` (Akalin et al., 2012), because kinship matrices based on epigenetic variation, rather than genetic variation, have been shown to better capture pairwise relatedness and control for the rate of false positives when testing for epigenetic-environmental associations (Gugger et al., 2016).

We then performed a GLMM analysis for each of the outlier SMVs to test for environmental associations. For these analyses, we considered three environmental variables, with low to moderate pairwise correlations ( $r = .051-.423$ ), that could act as environmental stressors or forces of selection on *Anolis* lizards: (a) maximum temperature of the warmest month ( $T_{max}$ ; BIO5), which was highly correlated with minimum temperature of the coldest month ( $T_{min}$ ; BIO6,  $r = .976$ ), either of which could represent thermal stress; (b) temperature annual range ( $T_{range}$ ; BIO7), which was highly correlated with isothermality (BIO3,  $r = .954$ ), both of which are measures of temperature stability;

and (c) precipitation seasonality (BIO15), which was correlated with dry season precipitation (BIO17,  $r = .845$ ) and vegetation (dry season EVI,  $r = .644$ ) and represents habitat stability. Experiments have shown that *A. cristatellus* and other anoles respond to variation in thermal environment, even over short periods (Hertz, 1992; Kolbe et al., 2012; Muñoz et al., 2014), and exhibit physiological and behavioural divergence in response to humidity gradients and changes in habitat structure (Gorman & Hillman, 1977; Gunderson et al., 2011; Hertz & Huey, 1981; Leal & Fleishman, 2004).

In GLMM, the SNP or SMV data comprise the response variable, environmental factors are coded as fixed effects, and a correlation matrix (typically population structure, relatedness or geographical distance) is a covariate with a random effect (Lobréaux & Melodelima, 2015). For each SMV, we used the population mean methylation frequency as the response variable, the epigenetic kinship matrix as a covariate, and maximum temperature ( $T_{max}$ ), temperature range ( $T_{range}$ ) and precipitation seasonality as predictor variables. For each SNP, we used the major allele frequency for each population as the response variable, the  $F_{ST}$  matrix as a covariate, and  $T_{max}$ ,  $T_{range}$  and precipitation seasonality as predictor variables. For each analysis, we tested a set of hierarchical models that included the three environmental predictors on their own, in pairs, and all together, resulting in a set of six nested GLMMs plus a null model that included only the covariate matrix. We compared models based on their Akaike information criterion (AIC) scores and calculated the statistical significance of the best model compared to the null model based on their log likelihoods (Rousset & Ferdy, 2014). We then adjusted  $p$ -values to  $q$ -values using the FDR method, implemented in the `QVALUE` R package (Storey, Bass, Dabney, Robinson, & Warnes, 2019), to correct for multiple tests.

## 3 | RESULTS

### 3.1 | Sequence alignment and variant calling

Our sequencing of 79 *Anolis cristatellus* yielded 9,642,016 cleaned and filtered reads, totalling 911,701,107 bp. Bisulfite conversion rates were consistently high (>98.23%) for all samples, and mapping efficiency ranged from 37.93% to 42.15% for each sample (Table S1). For our genetic data set, after all filtering and quality control steps, we recovered 8,459 SNP loci that were present in at least 85% of the 79 samples in our data set, with no individual missing more than 10% data. The average coverage for these SNPs was  $11.4 \pm 38.5$  reads per locus per individual. For our epigenetic data set, our sequencing efforts resulted in low coverage of CpG sites for two samples from Guajataca (GJ), so we removed these from our data set. For the remaining 77 samples, after applying a 10% missing data threshold and requiring that SMV sites be represented in at least five samples per population with at least  $10\times$  coverage, we recovered 53,852 SMVs. After applying a cut-off of at least a 10% range in methylation frequencies among samples we retained 9,772 SMVs, and after filtering for potential

C/T polymorphisms our final epigenetic data set included 8,580 SMVs with average coverage of  $111.7 \pm 264.7$  reads per locus per individual. Mean methylation frequencies across all loci for each population varied little, ranging from 34.7% to 36.1%.

### 3.2 | GIS data layers

Each of the 23 environmental data layers in our data set showed substantial spatial variation across Puerto Rico. The PCA on these raster layers resulted in five layers retained for downstream analyses, based on the scree plot, that cumulatively explained 94.27% of the variance in the original variables (Table S2). Loadings for each variable indicated that raster PC1 and raster PC2 were primarily described by temperature variables, with PC1 reflecting annual mean, minimum and maximum temperatures and PC2 reflecting temperature range and seasonality. Raster PC3 was weighted heavily by the vegetation variables (NDVI and EVI) and wet season precipitation. The highest loadings for PC4 were temperature seasonality and precipitation seasonality, and PC5 was composed largely of variables representing bioclimatic extremes, including mean temperature of the driest quarter and precipitation of the warmest quarter (Table S2).

### 3.3 | SNP and SMV outlier detection

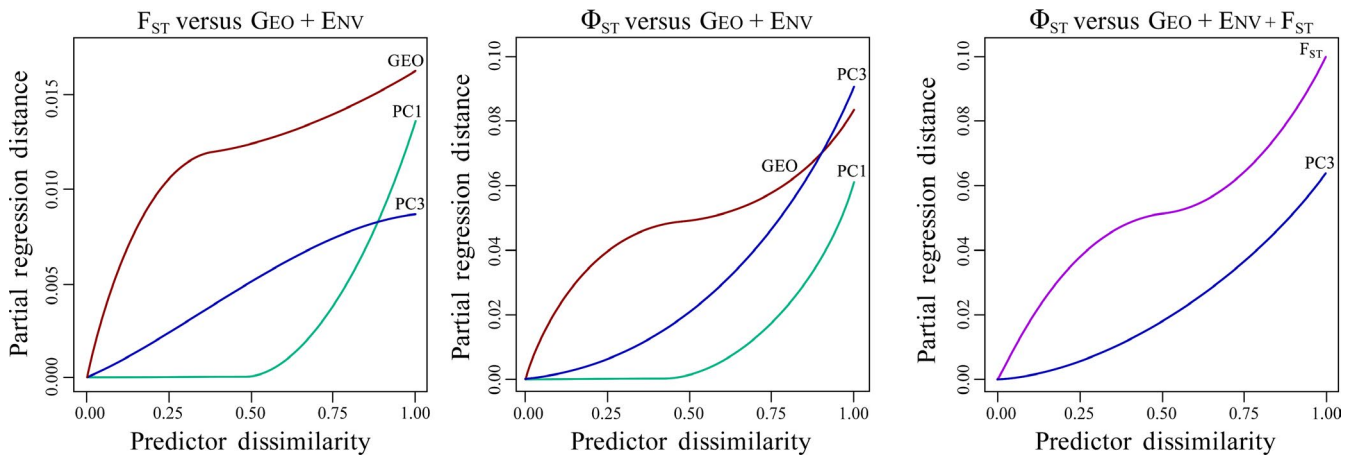
Based on the `PCADAPT` analysis, we identified 24 SNPs and 177 SMVs that were significant outliers after FDR correction for multiple testing. The range in methylation frequencies between populations for each of these loci averaged 72.8%, suggesting there are many highly differentiated SMVs across the genome. The Manhattan plot of epigenetic variation showed that highly divergent loci were dispersed throughout the genome and did not appear to cluster in any regions (Figure S1). Gene annotation, based on the *Anolis carolinensis* genome, assigned 72 of the outlier SMVs to promoter or gene body regions: six in exons, 56 in introns and 10 in promoters (Table S3). These genes had a wide range of putative functions, many associated with molecular and cellular processes (Table S3).

### 3.4 | Genetic and epigenetic structure

Our estimates of genetic structure revealed low to moderate levels of genetic differentiation between populations, with pairwise  $F_{ST}$  values ranging from 0.038 to 0.071 and mean  $F_{ST} = 0.054 \pm 0.010$  SD (Table 1). We also found moderate to high levels of epigenetic differentiation, with pairwise  $\Phi_{ST}$  values ranging from 0.049 to 0.255 and mean  $\Phi_{ST} = 0.128 \pm 0.053$  SD (Table 1). All pairwise values of  $F_{ST}$  and  $\Phi_{ST}$  were significantly different from zero after correction for multiple tests ( $p < .05$ ). The  $\Phi_{ST}$  matrices constructed from SMVs with different genomic contexts (introns, exons,

**TABLE 1** Pairwise estimates of genetic  $F_{ST}$  (below diagonal) and epigenetic  $\Phi_{ST}$  (above diagonal) for eight populations of *Anolis cristatellus*; all values are statistically significant after correction for multiple tests

	AG	CM	CT	GJ	GN	MC	RA	RN
AG	—	0.193	0.087	0.097	0.255	0.105	0.174	0.200
CM	0.065	—	0.085	0.175	0.107	0.099	0.059	0.119
CT	0.048	0.045	—	0.099	0.173	0.049	0.086	0.116
GJ	0.068	0.046	0.052	—	0.226	0.061	0.165	0.174
GN	0.051	0.065	0.053	0.067	—	0.144	0.149	0.121
MC	0.052	0.047	0.042	0.045	0.040	—	0.084	0.082
RA	0.070	0.038	0.051	0.048	0.070	0.050	—	0.106
RN	0.071	0.052	0.056	0.039	0.068	0.049	0.054	—



**FIGURE 2** GDM-fitted I-splines for each predictor variable significantly associated with genetic ( $F_{ST}$ ) and epigenetic distances ( $\Phi_{ST}$ ), based on 8,459 SNPs and 8,580 SMVs. Included are the results of three GDM analyses: genetic distance versus geographical (Geo.) and environmental (Env.) distances (left), epigenetic distance versus geographical and environmental distances (middle), and epigenetic distance versus versus geographical, environmental and genetic distances (right). The maximum height of each I-spline indicates the total amount of genetic or epigenetic turnover associated with that variable, while holding all other variables constant, and the shape of each curve indicates how the rate of compositional turnover varies with predictor dissimilarity [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

promoters and intergenic regions) were all significantly correlated with each other and with the  $\Phi_{ST}$  matrix for all nonoutlier SMVs ( $p \leq .022$  in all cases).

Analyses of population structure identified two major genetic clusters ( $k = 2$ ) that corresponded to the four northernmost populations (CM, GJ, RA and RN) and the two southernmost populations (AG and GN), with the two populations in the central mountains (CT and MC) showing putative admixture between the northern and southern clusters (Figure S2). The genetic PCA suggested well-differentiated populations, overall, with MC and CT midway between the northern and southern populations along the first genetic principal component (gPC1) and populations divided east–west along gPC2 (Figure S3). The epigenetic PCA suggested more population overlap in epigenetic space than genetic space but also differentiated the same population groups (Figure S4).

### 3.5 | Isolation by distance and environment

Generalized dissimilarity modeling analyses found evidence of significant IBD and IBE for both genetic and epigenetic differentiation

(Figure 2; Table 2). In all cases, topographic distances were a slightly better fit than Euclidean distances. The best model for genetic structure explained a very high proportion of the deviance in the data (deviance explained = 0.765,  $p = .001$ ; Table 2) and included topographic distance, environmental PC1 (temperature) and environmental PC3 (vegetation). The strongest individual effect came from topographic distance ( $\beta_{IBD} = 0.016$ ,  $p = .002$ ), followed by PC1 ( $\beta_{PC1} = 0.013$ ,  $p = .045$ ) and PC3 ( $\beta_{PC3} = 0.008$ ,  $p = .040$ ). However, the combined effects of PC1 and PC3 indicate a slightly stronger signal of IBE ( $\beta_{IBE} = 0.021$ ) than IBD, overall.

For GDM analyses of epigenetic structure, when we included only topographic distance and the environmental PCs as predictor variables the best model found significant effects of topographic distance ( $\beta_{IBD} = 0.083$ ,  $p = .004$ ), PC1 ( $\beta_{PC1} = 0.061$ ,  $p = .048$ ) and PC3 ( $\beta_{PC3} = 0.090$ ,  $p = .002$ ) on epigenetic differentiation (Table 2; Figure 2). Together, the regression coefficients for environmental PC1 (temperature) and PC3 (vegetation) indicate a strong signal of IBE ( $\beta_{IBE} = 0.151$ ). This model was a significant fit to the data and explained a large portion of the variance in the data (deviance explained = 0.684,  $p = .002$ ). When we also included  $F_{ST}$  as a predictor, the best model found that  $F_{ST}$  had the largest effect on epigenetic



**TABLE 2** Results of GDM and MMRR analyses for (a) genetic differentiation ( $F_{ST}$ ) versus geographical distance and environmental dissimilarity, (b) epigenetic differentiation ( $\phi_{ST}$ ) versus geographical distance and environmental dissimilarity, and (c) epigenetic differentiation versus geographical distance, environmental dissimilarity, and genetic differentiation

<b>(a) Genetic versus Geo. + Env.</b>				
	<b>GDM</b>		<b>MMRR</b>	
	Dev. Expl. = 0.765, $p = .001$		$r^2 = .381, p = .028$	
Variable	Coefficient	$p$ -value	Coefficient	$p$ -value
Geo. Dist.	0.016	.002	0.396	.044
PC1	0.013	.045	0.223	.162
PC3	0.008	.040	0.377	.036
<b>(b) Epigenetic versus Geo. + Env.</b>				
	<b>GDM</b>		<b>MMRR</b>	
	Dev. Expl. = 0.684, $p = .002$		$r^2 = .448, p = .022$	
Variable	Coefficient	$p$ -value	Coefficient	$p$ -value
Geo. Dist.	0.083	.004	0.233	.192
PC1	0.061	.048	0.073	.629
PC3	0.090	.002	0.591	.006
<b>(c) Epigenetic versus Geo. + Env. + Gen.</b>				
	<b>GDM</b>		<b>MMRR</b>	
	Dev. Expl. = 0.710, $p < .001$		$r^2 = .756, p = .001$	
Variable	Coefficient	$p$ -value	Coefficient	$p$ -value
PC3	0.063	.007	0.305	.013
$F_{ST}$	0.100	.001	0.613	.001

Note: GDM results include the total deviance explained (Dev. Expl.) and significance ( $p$ ) for the fitted model. MMRR results include the model fit ( $r^2$ ) and significance ( $p$ ). For each model, the significance value of each predictor variable based on 999 permutations is included next to its regression coefficient. Variables with values in italics were not retained in the best MMRR model after backward elimination but are shown here for comparison with the GDM results.

differentiation ( $\beta_{F_{ST}} = 0.100, p = .001$ ) and that PC3 also had a significant effect ( $\beta_{PC3} = 0.063, p = .007$ ; Table 2; Figure 2). Again, this model explained a large proportion of the variance in the data and was a significant fit (deviance explained = 0.710,  $p < .001$ ). Projections of the GDM models onto the environmental data layers showed similar spatial patterns of turnover in genetic and epigenetic composition (Figure 3).

Results from MMRR were highly concordant with the GDM results (Table 2). For analyses of genetic structure, the MMRR results indicated significant IBD and IBE in a model that was a significant fit to the data ( $r^2 = .381, p = .028$ ). In these results, geographical distance ( $\beta_{IBD} = 0.396, p = .044$ ) and PC3 ( $\beta_{PC3} = 0.377, p = .036$ ) had significant effects on genetic differentiation, but PC1 did not ( $\beta_{PC1} = 0.223, p = .162$ ). For analyses of epigenetic structure versus geographical distance and environmental dissimilarity, the MMRR

results indicated a strong effect of PC3 on epigenetic differentiation ( $\beta_{PC3} = 0.591, p = .006$ ) but did not find significant effects of geographical distance ( $\beta_{IBD} = 0.233, p = .192$ ) or PC1 ( $\beta_{PC1} = 0.073, p = .629$ ). For analyses of epigenetic structure that also included  $F_{ST}$  as a predictor, MMRR found strong effects of both genetic differentiation ( $\beta_{F_{ST}} = 0.613, p = .001$ ) and PC3 ( $\beta_{PC3} = 0.305, p = .013$ ) on epigenetic differentiation. The model including  $F_{ST}$  was a better fit to the data ( $r^2 = .756, p = .001$ ) than the one including only geographical and environmental variables ( $r^2 = .448, p = .022$ ; Table 2).

Plots of genetic  $F_{ST}$  and epigenetic  $\phi_{ST}$  against geographical distance and environmental distance show positive correlations, indicating genetic and epigenetic IBD and IBE (Figure 4). When the effects of  $F_{ST}$  are controlled for, the plot of residual  $\phi_{ST}$  (after partialling out  $F_{ST}$ ) shows that the correlation with geographical distance (IBD) flattens out but the positive correlation with environmental distance remains (Figure 4).

### 3.6 | Genetic- and epigenetic-environmental associations

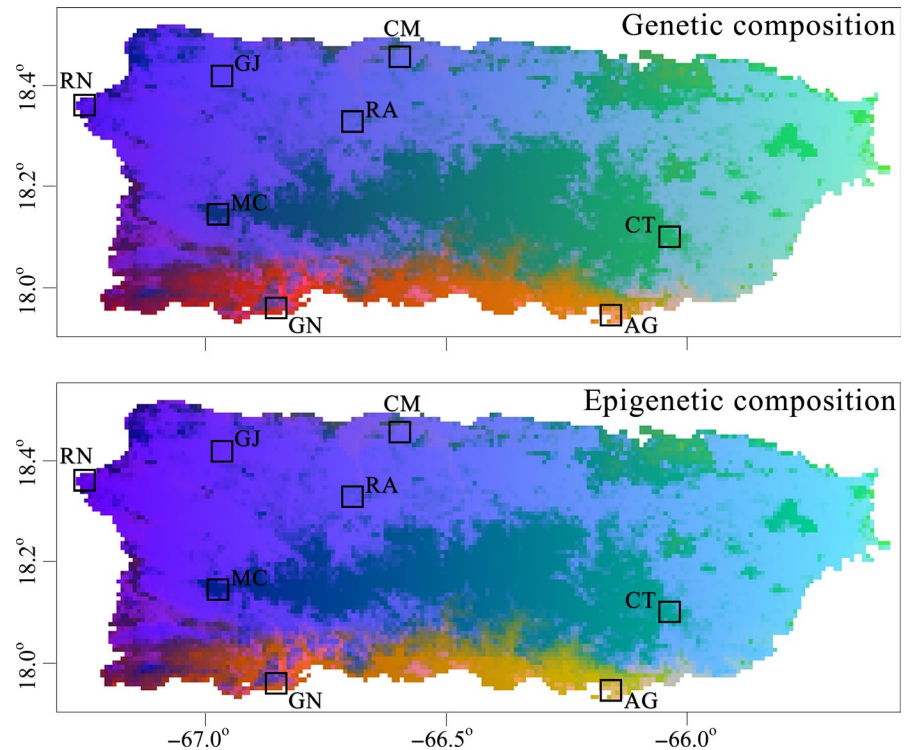
We detected significant associations with environmental variables for 96 out of the 177 outlier SMVs after FDR correction ( $q < 0.05$ ): four in promoters, five in exons, 33 in introns and 54 in intergenic regions (Table 3; Table S3). For all but one of the 96 SMVs, GLMM analysis indicated a significant correlation with precipitation seasonality. Of these 95 SMVs, 14 were also significantly associated with maximum temperature of the warmest month ( $T_{max}$ ) and 25 with temperature annual range ( $T_{range}$ ). The single SMV not associated with precipitation seasonality was significantly associated with  $T_{max}$  and  $T_{range}$  (Table S3). We did not see any systematic differences between SMVs in genic and intergenic regions with respect to their environmental associations (Table 3). SMVs in exons did have a high rate of significant environmental association (83% compared to 40%–59% for introns, promoters and intergenic regions), but our data set includes only six SMVs in exons, preventing meaningful statistical comparisons.

For our genetic data set, we found significant environmental associations with allele frequencies in only two of the 24 outlier SNPs. Both were significantly associated with  $T_{max}$  and  $T_{range}$  but not with precipitation seasonality ( $\beta_{T_{max}} = 0.006, \beta_{T_{range}} = 0.014, p = .012$ ; and  $\beta_{T_{max}} = 0.005, \beta_{T_{range}} = 0.017, p = .012$ ).

## 4 | DISCUSSION

Although recent studies have shown that epigenetic variation may play an important role in population-level responses to environmental heterogeneity (Alakärppä et al., 2018; Platt et al., 2015; Richards, Schrey, & Pigliucci, 2012; Xie et al., 2015), so far little evidence exists for spatial patterns of epigenetic differentiation in natural populations that are independent from genetic structure (Dubin et al., 2015; Foust et al., 2016; Herrera et al., 2016; Hu & Barrett, 2017). In

**FIGURE 3** Maps of spatial turnover in projected genetic (a) and epigenetic (b) composition, based on GDM analysis. Fitted generalized dissimilarity models are used to transform the underlying geographical and environmental variables, resulting in projections of biological composition in which increasingly different colours indicate greater genetic or epigenetic dissimilarity [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



particular, empirical demonstration of epigenetic IBE has been limited to only a few plant systems (Foust et al., 2016; Herrera et al., 2017). By performing landscape genetic and landscape epigenetic analyses for a species inhabiting a diverse range of habitats, we were able to quantify not only genome-wide and locus-specific patterns of variation but also the specific environmental variables associated with epigenetic divergence. Our results provide clear evidence that geographical and environmental factors play important roles in structuring spatial genetic and epigenetic variation in a natural vertebrate system.

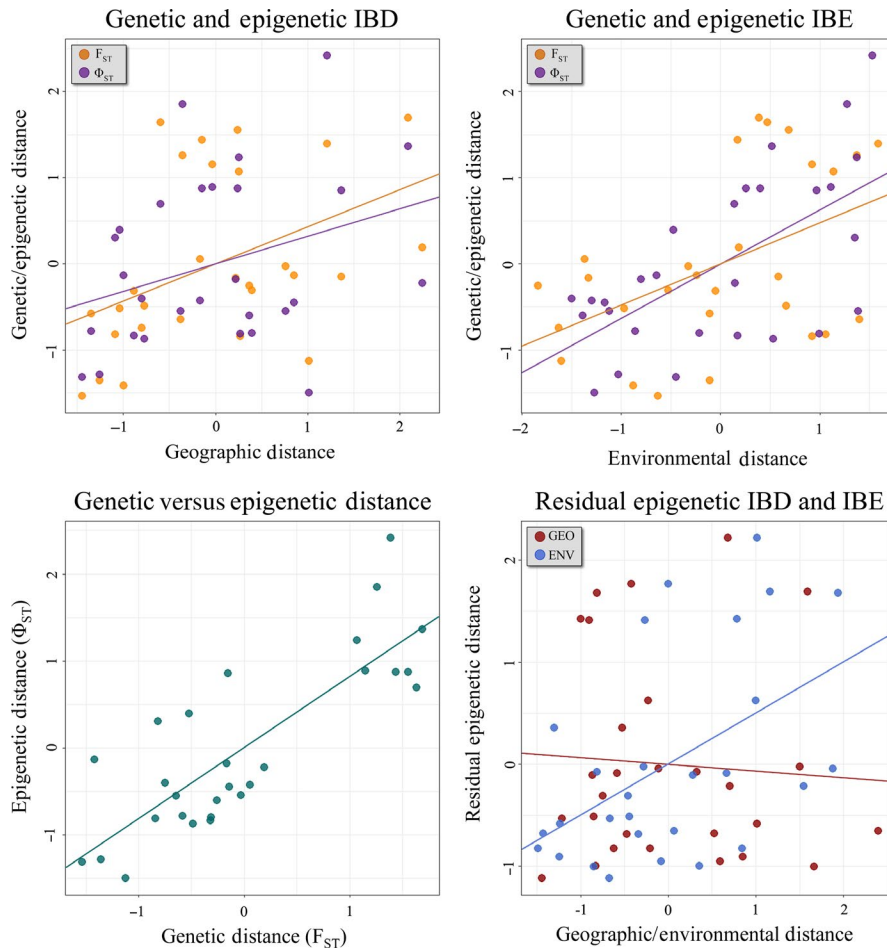
#### 4.1 | Spatial genetic and epigenetic structure

We found that spatial genetic variation does, indeed, explain a large amount of epigenetic variation in *Anolis cristatellus* (Table 2). Projections of spatial turnover in genetic and epigenetic composition show obvious similarities (Figure 3), and plots of genetic and epigenetic IBD and IBE indicate similar slopes (Figure 4), suggesting that genetic and epigenetic variation may have shared responses to geographical and environmental factors (Herrera et al., 2017) or that epigenetic variation could depend on genetic variation. However, we also found strong evidence that epigenetic structure in this system is not only attributable to the pattern of genetic variation. Our GDM and MMRR analyses revealed significant epigenetic IBE, even after accounting for the effects of genetic structure (Table 2). In each analysis, environmental PC3 (vegetation) has about half to two-thirds as large an effect as genetic  $F_{ST}$  on epigenetic structure (GDM:  $\beta_{PC3} = 0.063$ ,  $\beta_{F_{ST}} = 0.100$ ; MMRR:  $\beta_{PC3} = 0.305$ ,  $\beta_{F_{ST}} = 0.613$ ), suggesting that about one-third of the

variation in epigenetic distances ( $\phi_{ST}$ ) explained by the models is due to IBE (Table 2). Additionally, after accounting for genetic structure, signals of epigenetic IBD and effects of PC1 (temperature) disappear, probably because  $F_{ST}$  captures variation in PC1 and geographical distance (Table 2; Figures 2 and 4). Hence, we do not find evidence for epigenetic IBD independent of genetic structure but do find clear evidence for a strong pattern of genome-wide epigenetic IBE.

Our analyses of genetic variation, by contrast, recovered significant patterns of both geographical and environmental structure. The results of our STRUCTURE and genetic PCA analyses revealed major geographical subdivision of our sampled populations between the northern and southern parts of Puerto Rico (Figure S2 and gPC1 in Figure S3) and some structure reflecting their east-west orientation as well (gPC2 in Figure S3). These results are broadly consistent with previous phylogeographic analyses of mtDNA data in *A. cristatellus* (Kolbe et al., 2007; Revell, Harmon, Langerhans, & Kolbe, 2007), which identified deep splits between southwestern Puerto Rico and the rest of the island, with additional substructure between south-east, northeast and northwest populations. GDM and MMRR each found significant IBD and IBE (Table 2), consistent with a previous analysis of mtDNA variation in *A. cristatellus* (Wang et al., 2013). They differed slightly in identifying which environmental variables contributed to IBE (Table 2), but these differences are minor and probably result from MMRR fitting linear relationships between untransformed variables but GDM modelling nonlinear relationships (Figure 2; Ferrier et al., 2007; Fitzpatrick & Keller, 2015; Wang, 2013).

A variety of selective and nonselective processes can generate genetic IBE, including natural and sexual selection against migrants



**FIGURE 4** Plots of (a) genetic and epigenetic IBD ( $F_{ST}$  and  $\Phi_{ST}$  vs. geographical distance), (b) genetic and epigenetic IBE ( $F_{ST}$  and  $\Phi_{ST}$  vs. geographical distance), (c) genetic distance versus epigenetic distance, and (d) epigenetic IBD and IBE after controlling for genetic structure (residual  $\Phi_{ST}$  vs. geographical distance and environmental distance) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 3** SMV outliers by genetic element, including the total number of outliers, the number of outliers with significant environmental associations (Sig. env. assoc.), and the number of those SMVs associated with maximum temperature of the warmest month ( $T_{max}$ ), temperature annual range ( $T_{range}$ ) and precipitation seasonality (Precip.)

Element	Total outliers	Sig. env. assoc.	$T_{max}$	$T_{range}$	Precip.
Exon	6	5	0	0	5
Intron	56	33	11	3	33
Promoter	10	4	0	1	4
Intergenic	105	54	15	11	53
Total	177	96	26	15	95

from divergent environments, differential fitness of admixed offspring and biased dispersal (Wang & Bradburd, 2014). Because a substantial amount of spatial variation in genome-wide methylation is explained by genetic structure in *A. cristatellus*, these same processes may drive epigenetic IBE, too, by reducing dispersal between different environments and allowing epigenetic divergence through random drift (Richards et al., 2010; Trucchi et al., 2016). Experimental evidence has shown that *A. cristatellus* from different environments diverge physiologically, including in temperature regulation and water loss (Gunderson et al., 2011; Hertz, 1992; Huey, 1983), and morphologically in traits associated with locomotor performance (Winchell et al., 2016), both of which could lead to natural selection against migrants from different environments. Other studies have

shown that habitat structure leads to divergence in anole dewlaps—the longitudinal flap of skin that extends below the lower jaw and is used in intraspecific signalling—and even the timing of reproductive cycles (Leal & Fleishman, 2004; Otero et al., 2015). Environmental PC1 is an axis that primarily describes temperature variation, and PC3 primarily comprises vegetation layers. So, it seems reasonable that genetic and epigenetic IBE could result from any of these processes in this system, although experiments linking gene flow or genetic differentiation to divergent natural or sexual selection are needed to provide confirmation.

Additionally, for epigenetic variation, two other mechanisms could potentially generate IBE as well. First, because epigenetic changes can be environmentally induced, differential methylation

could occur across environmental gradients of the factors that induce methylation (Angers et al., 2010; Downen et al., 2012; Jaenisch & Bird, 2003), potentially leading to divergence between environments across the epigenome. Second, the effect of methylation on gene expression means that epigenetic variation can be linked to phenotypic plasticity (Artemov et al., 2017; Baerwald et al., 2016; Dimond & Roberts, 2016), potentially leading to epigenetic divergence between environments for gene body methylation affecting transcription if it results from a plastic response to the local environment or contributes to differential survival or biased dispersal. Differences between genetic and epigenetic IBE indicate differences in how the genome and methylome respond to environmental variation and are probably attributable to processes such as these that act on epigenetic but not genetic variation. Our results emphasize that genetic structure cannot entirely explain variation in genome-wide methylation across a landscape, and further studies linking variation in methylation to phenotypic changes could provide insights into the effects of epigenetic variation on population divergence and plasticity in different environments (Verhoeven et al., 2016).

## 4.2 | Climate-associated SMVs

Our analyses of individual SMVs found 96 outliers strongly associated with key climatic variables (Table S3)—about half of these were in promoter or gene body regions, suggesting potential involvement in adaptive or plastic responses to local environments, but many fell in intergenic regions, where methylation has no known functional effect. Several of the outlier SMVs had methylation frequencies significantly correlated with either maximum temperature of the warmest month ( $T_{\max}$ , 26 SMVs) or temperature annual range ( $T_{\text{range}}$ , 15 SMVs; Table 3; Table S3). These loci did not overlap with genes previously predicted to have roles in thermal adaptation in nonavian reptiles (Wollenberg Valero et al., 2014), but genes outside of these predicted networks can be involved in responses to climate variation in *Anolis* lizards (Campbell-Staton et al., 2017). Several studies have proposed that methylation may provide a mechanism for rapid responses to climate change (Dimond & Roberts, 2016; Jeremias et al., 2018; Metzger & Schulte, 2017; Platt et al., 2015), and further studies could help to understand whether methylation is involved in thermal adaptation or acclimatization in small ectotherms, such as *A. cristatellus*, which are particularly vulnerable to climate change (Metzger & Schulte, 2017; Sinervo et al., 2010).

All but one of the 96 environmentally associated outlier SMVs, including all 42 SMVs in promoters or gene bodies, were significantly linked to precipitation seasonality, indicating that DNA methylation may play a role in mediating responses to environmental variability in rainfall and its effects on habitat structure. Several of these SMVs were located in genes belonging to biological pathways that could be important for facilitating responses to environmental conditions (Table S3), including ATPase activity (ATP9A), Rho-GTPase activity (ARHGAP15, CHN1 and PAK1), cell-cell adhesion (DSC2), energy metabolism (PPF1BP1) and transmembrane signalling

(GRIK4, SLC38A7, SLC6A7, TENM2, and XPR1). Epigenetic divergence in genes involved in Rho-GTPase and signal transduction pathways has been identified in a variety of systems and associated with environmental stress (Baerwald et al., 2016; Hu et al., 2019; Hu, Pérez-Jvostov, Blondel, & Barrett, 2018; Uren Webster et al., 2018). Moreover, one sodium symporter gene, SLC6A7, is in the family of gamma-aminobutyric (GABA) transporters that includes two genes, SLC6A1 and SLC6A8, found to be differentially expressed in *Anolis carolinensis* populations responding to severe winter storms (Campbell-Staton et al., 2017). However, whether these genes are functionally relevant in different environments and whether methylation at these loci alters gene expression in *A. cristatellus* remain unknown. Nevertheless, strong relationships between outlier SMVs and climate variables and genome-wide evidence of IBE suggest that methylation is at least influenced by the environment if not a driver of local adaptation or phenotypic plasticity (Gugger et al., 2016; Keller et al., 2016; Wilschut, Oplaat, Snoek, Kirschner, & Verhoeven, 2016).

## 5 | CONCLUSIONS

Isolation by environment is an important pattern of spatial genetic variation in a wide range of systems (Nanninga, Saenz-Agudelo, Manica, & Berumen, 2014; Sexton et al., 2014; Shafer & Wolf, 2013; Zhang, Wang, Comes, Peng, & Qiu, 2016), and studies of genetic IBE have provided valuable insights into the modes of divergence involved in early biological diversification (Sexton et al., 2014; Wang & Bradburd, 2014). Genetic IBE appears to be fairly commonplace in the *Anolis* radiations of the Greater and Lesser Antilles (Thorpe, Barlow, Malhotra, & Surget-Groba, 2015; Thorpe, Surget-Groba, & Johansson, 2008; Wang et al., 2013), and our results suggest that, in *A. cristatellus*, epigenetic IBE occurs independently of spatial genetic structure as well. Diversification in anoles is closely tied to ecological variation (Glor, Kolbe, Powell, Larson, & Losos, 2003; Mahler, Revell, Glor, & Losos, 2010), and anoles are known to exhibit rapid genetic and phenotypic responses to changes in their biotic and abiotic environments (Campbell-Staton et al., 2017; Stuart et al., 2014). Our results showing strong relationships between epigenetic and environmental spatial structure in *A. cristatellus* raise the question of whether epigenetic variation contributes to adaptation and persistence across diverse habitats in this and other systems. Future research characterizing the processes underlying epigenetic IBE will play an important role in advancing our understanding of the links between epigenetic variation and the environment.

## ACKNOWLEDGEMENTS

We thank J. Boyko, L. Gray and B. White for their hard work in the field, A. Lea for advice on library preparation, T. Vilgalys for help filtering SMV data, and R. Bell, W. Brinkerhoff, J. Frederick, N. Graham, D. Hart, M. McElroy, K. O'Connell, I. Prates, R. Schott and E. Westeen for helpful comments on drafts of the manuscript. We also thank the Departamento de Recursos Naturales y Ambientales (DRNA) in





Puerto Rico for granting scientific collecting permits for this project (#2017-IC-018) and the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley for support with DNA sequencing. This study was supported by a Society for the Study of Amphibians and Reptiles Grant-in-Herpetology to M.L.Y. and an NSF Dimensions of Biodiversity grant (DEB-1542534) to I.J.W.

## AUTHOR CONTRIBUTIONS

G.O.U.W. and M.L.Y. designed the study with input from D.L.M. and I.J.W. G.O.U.W. and I.J.W. performed the field work, and G.O.U.W. and M.L.Y. collected the molecular data. G.O.U.W. and I.J.W. analysed the data with help from M.L.Y. G.O.U.W. and I.J.W. wrote the paper, with contributions from M.L.Y. and D.L.M.

## ORCID

Guinevere O. U. Wogan  <https://orcid.org/0000-0003-1877-0591>  
Ian J. Wang  <https://orcid.org/0000-0003-2554-9414>

## DATA AVAILABILITY STATEMENT

All DNA sequences can be downloaded from the NCBI BioProject ID PRJNA579702. The SNP and SMV data sets and environmental PC rasters are available from the Dryad Digital Repository (<https://doi.org/10.6078/D1X97D>).

## REFERENCES

- Akalin, A., Franke, V., Vlahoviček, K., Mason, C. E., & Schübeler, D. (2015). genomation: A toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics*, 31(7), 1127–1129. <https://doi.org/10.1093/bioinformatics/btu775>
- Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13(10), R87. <https://doi.org/10.1186/gb-2012-13-10-r87>
- Alakärppä, E., Salo, H. M., Valledor, L., Cañal, M. J., Häggman, H., & Vuosku, J. (2018). Natural variation of DNA methylation and gene expression may determine local adaptations of Scots pine populations. *Journal of Experimental Botany*, 69(21), 5293–5305. <https://doi.org/10.1093/jxb/ery292>
- Álföldi, J., Di Palma, F., Grabherr, M., Williams, C., Kong, L., Mauceli, E., ... Lindblad-Toh, K. (2011). The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*, 477(7366), 587–591. <https://doi.org/10.1038/nature10390>
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data* (Version 0.11.8). Retrieved from <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Angers, B., Castonguay, E., & Massicotte, R. (2010). Environmentally induced phenotypes and DNA methylation: How to deal with unpredictable conditions until the next generation and after. *Molecular Ecology*, 19(7), 1283–1295. <https://doi.org/10.1111/j.1365-294X.2010.04580.x>
- Artemov, A. V., Mague, N. S., Rastorguev, S. M., Zhenilo, S., Mazur, A. M., Tsygankova, S. V., ... Prokhortchouk, E. B. (2017). Genome-wide DNA methylation profiling reveals epigenetic adaptation of stickleback to marine and freshwater conditions. *Molecular Biology and Evolution*, 34(9), 2203–2213. <https://doi.org/10.1093/molbev/msx156>
- Baerwald, M. R., Meek, M. H., Stephens, M. R., Nagarajan, R. P., Goodbla, A. M., Tomalty, K. M. H., ... Nichols, K. M. (2016). Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Molecular Ecology*, 25(8), 1785–1800. <https://doi.org/10.1111/mec.13231>
- Bi, K., Vanderpool, D., Singhal, S., Linderroth, T., Moritz, C., & Good, J. M. (2012). Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, 13(1), 403. <https://doi.org/10.1186/1471-2164-13-403>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 1–7. <https://doi.org/10.1093/bioinformatics/btu170>
- Boyle, P., Clement, K., Gu, H., Smith, Z. D., Ziller, M., Fostel, J. L., ... Meissner, A. (2012). Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biology*, 13(10), R92. <https://doi.org/10.1186/gb-2012-13-10-r92>
- Campbell-Staton, S. C., Cheviron, Z. A., Rochette, N., Catchen, J., Losos, J. B., & Edwards, S. V. (2017). Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. *Science*, 357(6350), 495–498. <https://doi.org/10.1126/science.aam5512>
- Chen, N. (2004). Using RepeatMasker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics*, 5(1), 4.10.1–4.10.14. <https://doi.org/10.1002/0471250953.bi0410s05>
- Dimond, J. L., & Roberts, S. B. (2016). Germline DNA methylation in reef corals: Patterns and potential roles in response to environmental change. *Molecular Ecology*, 25(8), 1895–1904. <https://doi.org/10.1111/mec.13414>
- Downen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Downen, J. M., Nery, J. R., ... Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences of the United States of America*, 109(32), E2183–E2191. <https://doi.org/10.1073/pnas.1209329109>
- Dubin, M. J., Zhang, P., Meng, D., Remigereau, M.-S., Osborne, E. J., Paolo Casale, F., ... Nordborg, M. (2015). DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *eLife*, 4, e05255. <https://doi.org/10.7554/eLife.05255>
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, 13(3), 252–264. <https://doi.org/10.1111/j.1472-4642.2007.00341.x>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, 18(1), 1–16. <https://doi.org/10.1111/ele.12376>
- Foust, C. M., Preite, V., Schrey, A. W., Alvarez, M., Robertson, M. H., Verhoeven, K. J. F., & Richards, C. L. (2016). Genetic and epigenetic differences associated with environmental gradients in replicate populations of two salt marsh perennials. *Molecular Ecology*, 25(8), 1639–1652. <https://doi.org/10.1111/mec.13522>
- Fumagalli, M., Vieira, F. G., Linderroth, T., & Nielsen, R. (2014). ngsTools: Methods for population genetics analyses from next-generation sequencing data. *Bioinformatics*, 30(10), 1486–1487. <https://doi.org/10.1093/bioinformatics/btu041>
- Glor, R. E., Kolbe, J. J., Powell, R., Larson, A., & Losos, J. B. (2003). Phylogenetic analysis of ecological and morphological diversification in Hispaniolan trunk-ground anoles (*Anolis cybotes* group). *Evolution*, 57(10), 2383–2397. <https://doi.org/10.1111/j.0014-3820.2003.tb00250.x>
- Gorman, G. C., & Hillman, S. (1977). Physiological basis for climatic niche partitioning in two species of Puerto Rican *Anolis* (Reptilia, Lacertilia,



- Iguanidae. *Journal of Herpetology*, 11(3), 337–340. <https://doi.org/10.2307/1563246>
- Goudet, J., & Jombart, T. (2015). *hierfstat: Estimation and tests of hierarchical F-statistics (Version 0.04-22) [R]*.
- Greenspoon, P. B., & Spencer, H. G. (2018). The evolution of epigenetically mediated adaptive transgenerational plasticity in a subdivided population. *Evolution*, 72(12), 2773–2780. <https://doi.org/10.1111/evo.13619>
- Gugger, P. F., Fitz-Gibbon, S., Pellegrini, M., & Sork, V. L. (2016). Species-wide patterns of DNA methylation variation in *Quercus lobata* and their association with climate gradients. *Molecular Ecology*, 25(8), 1665–1680. <https://doi.org/10.1111/mec.13563>
- Gunderson, A. R., Siegel, J., & Leal, M. (2011). Tests of the contribution of acclimation to geographic variation in water loss rates of the West Indian lizard *Anolis cristatellus*. *Journal of Comparative Physiology B*, 181(7), 965–972. <https://doi.org/10.1007/s00360-011-0576-0>
- Guo, W., Fiziev, P., Yan, W., Cokus, S., Sun, X., Zhang, M. Q., ... Pellegrini, M. (2013). BS-Seeker2: A versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics*, 14(1), 774. <https://doi.org/10.1186/1471-2164-14-774>
- He, L. I., Wu, W., Zinta, G., Yang, L., Wang, D., Liu, R., ... Zhu, J.-K. (2018). A naturally occurring epiallele associates with leaf senescence and local climate adaptation in *Arabidopsis* accessions. *Nature Communications*, 9(1), 460. <https://doi.org/10.1038/s41467-018-02839-3>
- Henderson, R. W., & Powell, R. (2009). *Natural history of West Indian reptiles and amphibians*. Gainesville, FL: University Press of Florida.
- Herman, J. J., Spencer, H. G., Donohue, K., & Sultan, S. E. (2014). How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. *Evolution*, 68(3), 632–643. <https://doi.org/10.1111/evo.12324>
- Herrera, C. M., Medrano, M., & Bazaga, P. (2016). Comparative spatial genetics and epigenetics of plant populations: Heuristic value and a proof of concept. *Molecular Ecology*, 25(8), 1653–1664. <https://doi.org/10.1111/mec.13576>
- Herrera, C. M., Medrano, M., & Bazaga, P. (2017). Comparative epigenetic and genetic spatial structure of the perennial herb *Helleborus foetidus*: Isolation by environment, isolation by distance, and functional trait divergence. *American Journal of Botany*, 104(8), 1195–1204. <https://doi.org/10.3732/ajb.1700162>
- Hertz, P. E. (1992). Temperature regulation in Puerto Rican *Anolis* lizards: A field test using null hypotheses. *Ecology*, 73(4), 1405–1417. <https://doi.org/10.2307/1940686>
- Hertz, P. E., & Huey, R. B. (1981). Compensation for altitudinal changes in the thermal environment by some *Anolis* lizards on Hispaniola. *Ecology*, 62(3), 515–521. <https://doi.org/10.2307/1937714>
- Hu, J., Askary, A. M., Thurman, T. J., Spiller, D. A., Palmer, T. M., Pringle, R. M., & Barrett, R. D. H. (2019). The epigenetic signature of colonizing new environments in *Anolis* lizards. *Molecular Biology and Evolution*, 36(10), 2165–2170. <https://doi.org/10.1093/molbev/msz133>
- Hu, J., & Barrett, R. D. H. (2017). Epigenetics in natural animal populations. *Journal of Evolutionary Biology*, 30(9), 1612–1632. <https://doi.org/10.1111/jeb.13130>
- Hu, J., Pérez-Jvostov, F., Blondel, L., & Barrett, R. D. H. (2018). Genome-wide DNA methylation signatures of infection status in Trinidadian guppies (*Poecilia reticulata*). *Molecular Ecology*, 27(15), 3087–3102. <https://doi.org/10.1111/mec.14771>
- Huey, R. B. (1983). Natural variation in body temperature and physiological performance in a lizard (*Anolis cristatellus*). In A. G. J. Rhodin, & K. Miyata (Eds.), *Advances in herpetology and evolutionary biology: Essays in honor of Ernest E. Williams* (pp. 484–490). Cambridge, MA: Museum of Comparative Zoology.
- Jablonka, E., & Raz, G. (2009). Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology*, 84(2), 131–176. <https://doi.org/10.1086/598822>
- Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nature Genetics*, 33(3s), 245–254. <https://doi.org/10.1038/ng1089>
- Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jeremias, G., Barbosa, J., Marques, S. M., Asselman, J., Gonçalves, F. J. M., & Pereira, J. L. (2018). Synthesizing the role of epigenetics in the response and adaptation of species to climate change in freshwater ecosystems. *Molecular Ecology*, 27(13), 2790–2806. <https://doi.org/10.1111/mec.14727>
- Jiang, C., Mithani, A., Belfield, E. J., Mott, R., Hurst, L. D., & Harberd, N. P. (2014). Environmentally responsive genome-wide accumulation of de novo *Arabidopsis thaliana* mutations and epimutations. *Genome Research*, 24(11), 1821–1829. <https://doi.org/10.1101/gr.177659.114>
- Jiang, H., Lei, R., Ding, S.-W., & Zhu, S. (2014). Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*, 15(1), 182. <https://doi.org/10.1186/1471-2105-15-182>
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13(7), 484–492. <https://doi.org/10.1038/nrg3230>
- Keller, T. E., Lasky, J. R., & Yi, S. V. (2016). The multivariate association between genomewide DNA methylation and climate across the range of *Arabidopsis thaliana*. *Molecular Ecology*, 25(8), 1823–1837. <https://doi.org/10.1111/mec.13573>
- Kolbe, J. J., Glor, R. E., Schettino, L. R., Lara, A. C., Larson, A., & Losos, J. B. (2007). Multiple sources, admixture, and genetic variation in introduced *Anolis* lizard populations. *Conservation Biology*, 21(6), 1612–1625. <https://doi.org/10.1111/j.1523-1739.2007.00826.x>
- Kolbe, J. J., VanMiddlesworth, P. S., Losin, N., Dappen, N., & Losos, J. B. (2012). Climatic niche shift predicts thermal trait response in one but not both introductions of the Puerto Rican lizard *Anolis cristatellus* to Miami, Florida, USA. *Ecology and Evolution*, 2(7), 1503–1516. <https://doi.org/10.1002/ece3.263>
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, 15(1), 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Landesman, W. J., Nelson, D. M., & Fitzpatrick, M. C. (2014). Soil properties and tree species drive  $\beta$  diversity of soil bacterial communities. *Soil Biology and Biochemistry*, 76, 201–209. <https://doi.org/10.1016/j.soilbio.2014.05.025>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Leal, M., & Fleishman, L. J. (2004). Differences in visual signal design and detectability between allopatric populations of *Anolis* lizards. *The American Naturalist*, 163(1), 26–39. <https://doi.org/10.1086/an.2004.163.issue-1>
- Leutner, B., Horning, N., Schwald-Willmann, J., & Hijmans, R. J. (2019). *RStoolbox: tools for remote sensing data analysis* (Version 0.2.6). Retrieved from <https://cran.rproject.org/web/packages/RStoolbox/index.html>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>

- Lobréaux, S., & Melodelima, C. (2015). Detection of genomic loci associated with environmental variables using generalized linear mixed models. *Genomics*, 105(2), 69–75. <https://doi.org/10.1016/j.ygeno.2014.12.001>
- Lokk, K., Modhukur, V., Rajashekar, B., Märtens, K., Mägi, R., Kolde, R., ... Tõnisson, N. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biology*, 15(4), 3248. <https://doi.org/10.1186/gb-2014-15-4-r54>
- Losos, J. B. (2009). *Lizards in an evolutionary tree: Ecology and adaptive radiation of anoles*. Berkeley, CA: University of California Press.
- Lou, S., Lee, H.-M., Qin, H., Li, J.-W., Gao, Z., Liu, X., & Yip, K. Y. (2014). Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome Biology*, 15(7), 408. <https://doi.org/10.1186/s13059-014-0408-0>
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>
- Mahler, D. L., Ingram, T., Revell, L. J., & Losos, J. B. (2013). Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, 341(6143), 292–295. <https://doi.org/10.1126/science.1232392>
- Mahler, D. L., Revell, L. J., Glor, R. E., & Losos, J. B. (2010). Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution*, 64(9), 2731–2745. <https://doi.org/10.1111/j.1558-5646.2010.01026.x>
- Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., & Fitzpatrick, M. C. (2016). *gdm: Functions for generalized dissimilarity modeling (version 1.2.3)*. Retrieved from <https://cran.r-project.org/web/packages/gdm/index.html>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S., & Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Research*, 33(18), 5868–5877. <https://doi.org/10.1093/nar/gki901>
- Metzger, D. C. H., & Schulte, P. M. (2017). Persistent and plastic effects of temperature on DNA methylation across the genome of threespine stickleback (*Gasterosteus aculeatus*). *Proceedings of the Royal Society B: Biological Sciences*, 284(1864), 20171667. <https://doi.org/10.1098/rspb.2017.1667>
- Muñoz, M. M., Stimola, M. A., Algar, A. C., Conover, A., Rodriguez, A. J., Landestoy, M. A., ... Losos, J. B. (2014). Evolutionary stasis and lability in thermal physiology in a group of tropical lizards. *Proceedings of the Royal Society B: Biological Sciences*, 281(1778), 20132433. <https://doi.org/10.1098/rspb.2013.2433>
- Nanninga, G. B., Saenz-Agudelo, P., Manica, A., & Berumen, M. L. (2014). Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, 23(3), 591–602. <https://doi.org/10.1111/mec.12623>
- Oksanen, J., Kindt, R., Legendre, P., Ohara, B., & Stevens, M. H. H. (2007). *Vegan: Community ecology package. R package version 1.8.8*. Retrieved from <http://r-forge.r-project.org/projects/vegan>
- Otero, L. M., Huey, R. B., & Gorman, G. C. (2015). A few meters matter: Local habitats drive reproductive cycles in a tropical lizard. *The American Naturalist*, 186(3), E72–E80. <https://doi.org/10.1086/682359>
- Paradis, E. (2010). pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26(3), 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Perry, G., & Dmi'el, R., & Lazell, J. (2000). Evaporative water loss in insular populations of *Anolis cristatellus* (Reptilia: Sauria) in the British Virgin Islands. III. Response to the end of drought and a common garden experiment. *Biotropica*, 32(4a), 722–728. <https://doi.org/10.1111/j.1744-7429.2000.tb00520.x>
- Platt, A., Gugger, P. F., Pellegrini, M., & Sork, V. L. (2015). Genome-wide signature of local adaptation linked to variable CpG methylation in oak populations. *Molecular Ecology*, 24(15), 3823–3830. <https://doi.org/10.1111/mec.13230>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- Putnam, H. M., Davidson, J. M., & Gates, R. D. (2016). Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible corals. *Evolutionary Applications*, 9(9), 1165–1178. <https://doi.org/10.1111/eva.12408>
- Radford, E. J., Ito, M., Shi, H., Corish, J. A., Yamazawa, K., Isganaitis, E., ... Ferguson-Smith, A. C. (2014). *In utero* undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science*, 345(6198), 1255903. <https://doi.org/10.1126/science.1255903>
- Revell, L. J., Harmon, L. J., Langerhans, R. B., & Kolbe, J. J. (2007). A phylogenetic approach to determining the importance of constraint on phenotypic evolution in the neotropical lizard *Anolis cristatellus*. *Evolutionary Ecology Research*, 9, 261–282.
- Richards, C. L., Bossdorf, O., & Verhoeven, K. J. F. (2010). Understanding natural epigenetic variation. *New Phytologist*, 187(3), 562–564. <https://doi.org/10.1111/j.1469-8137.2010.03369.x>
- Richards, C. L., Schrey, A. W., & Pigliucci, M. (2012). Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. *Ecology Letters*, 15(9), 1016–1025. <https://doi.org/10.1111/j.1461-0248.2012.01824.x>
- Rosenberg, N. A. (2004). dstruct: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4(1), 137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Rousset, F., & Ferdy, J.-B. (2014). Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography*, 37(8), 781–790. <https://doi.org/10.1111/ecog.00566>
- Rubenstein, D. R., Skolnik, H., Berrio, A., Champagne, F. A., Phelps, S., & Solomon, J. (2016). Sex-specific fitness effects of unpredictable early life conditions are associated with DNA methylation in the avian glucocorticoid receptor. *Molecular Ecology*, 25(8), 1714–1728. <https://doi.org/10.1111/mec.13483>
- Rui, L. (2014). Energy metabolism in the liver. *Comprehensive Physiology*, 4(1), 177–197. <https://doi.org/10.1002/cphy.c130024>
- Sarver, B. A. J., Keeble, S., Cosart, T., Tucker, P. K., Dean, M. D., & Good, J. M. (2017). Phylogenomic insights into mouse evolution using a pseudoreference approach. *Genome Biology and Evolution*, 9(3), 726–739. <https://doi.org/10.1093/gbe/evx034>
- Schilling, E., & Rehli, M. (2007). Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics*, 90(3), 314–323. <https://doi.org/10.1016/j.ygeno.2007.04.011>
- Schmid, M. W., Heichinger, C., Schmid, D. C., Guthörl, D., Gagliardini, V., Bruggmann, R., ... Grossniklaus, U. (2018). Contribution of epigenetic variation to adaptation in *Arabidopsis*. *Nature Communications*, 9(1), 4446. <https://doi.org/10.1038/s41467-018-06932-5>
- Schmitz, R. J., Schultz, M. D., Urlich, M. A., Nery, J. R., Pelizzola, M., Libiger, O., ... Ecker, J. R. (2013). Patterns of population epigenomic diversity. *Nature*, 495(7440), 193–198. <https://doi.org/10.1038/nature11968>
- Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, 68(1), 1–15. <https://doi.org/10.1111/evo.12258>
- Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters*, 16(7), 940–950. <https://doi.org/10.1111/ele.12120>

- Sinervo, B., Mendez-de-la-Cruz, F., Miles, D. B., Heulin, B., Bastiaans, E., Villagran-Santa Cruz, M., ... Sites, J. W. (2010). Erosion of lizard diversity by climate change and altered thermal niches. *Science*, 328(5980), 894–899. <https://doi.org/10.1126/science.1184695>
- Singhal, S. (2013). De novo transcriptomic analyses for non-model organisms: An evaluation of methods across a multi-species data set. *Molecular Ecology Resources*, 13(3), 403–416. <https://doi.org/10.1111/1755-0998.12077>
- Storey, J. D., Bass, A. J., Dabney, A., Robinson, D., & Warnes, G. (2019). *qvalue: Q-value estimation for false discovery rate control. (Version 2.14.1) [R package]*.
- Stuart, Y. E., Campbell, T. S., Hohenlohe, P. A., Reynolds, R. G., Revell, L. J., & Losos, J. B. (2014). Rapid evolution of a native species following invasion by a congener. *Science*, 346(6208), 463–466. <https://doi.org/10.1126/science.1257008>
- Sul, J. H., & Eskin, E. (2013). Mixed models can correct for population structure for genomic regions under selection. *Nature Reviews Genetics*, 14(4), 300. <https://doi.org/10.1038/nrg2813-c1>
- Thorpe, R. S., Barlow, A., Malhotra, A., & Surget-Groba, Y. (2015). Widespread parallel population adaptation to climate variation across a radiation: Implications for adaptation to climate change. *Molecular Ecology*, 24(5), 1019–1030. <https://doi.org/10.1111/mec.13093>
- Thorpe, R. S., Surget-Groba, Y., & Johansson, H. (2008). The relative importance of ecology and geographic isolation for speciation in anoles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1506), 3071–3081. <https://doi.org/10.1098/rstb.2008.0077>
- Trucchi, E., Mazzarella, A. B., Gilfillan, G. D., Lorenzo, M. T., Schönswetter, P., & Paun, O. (2016). BsRADseq: Screening DNA methylation in natural populations of non-model species. *Molecular Ecology*, 25(8), 1697–1713. <https://doi.org/10.1111/mec.13550>
- Uren Webster, T. M., Rodriguez-Barreto, D., Martin, S. A. M., Oosterhout, C. V., Orozco-terWengel, P., Cable, J., ... Consuegra, S. (2018). Contrasting effects of acute and chronic stress on the transcriptome, epigenome, and immune response of Atlantic salmon. *Epigenetics*, 13(12), 1191–1207. <https://doi.org/10.1080/15592294.2018.1554520>
- Vanyushin, B. F., Mazin, A. L., Vasilyev, V. K., & Belozersky, A. N. (1973). The content of 5-methylcytosine in animal DNA: The species and tissue specificity. *Biochimica Et Biophysica Acta (BBA) – Nucleic Acids and Protein Synthesis*, 299(3), 397–403. [https://doi.org/10.1016/0005-2787\(73\)90264-5](https://doi.org/10.1016/0005-2787(73)90264-5)
- Verhoeven, K. J. F., vonHoldt, B. M., & Sork, V. L. (2016). Epigenetics in ecology and evolution: What we know and what we need to know. *Molecular Ecology*, 25(8), 1631–1638. <https://doi.org/10.1111/mec.13617>
- Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, 16, 175–182. <https://doi.org/10.1111/evo.12134>
- Wang, I. J. (2019). Topographic path analysis for modeling dispersal and functional connectivity: Calculating topographic distances using the topoDistance R package. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041-210X.13317>
- Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, 23(23), 5649–5662. <https://doi.org/10.1111/mec.12938>
- Wang, I. J., Glor, R. E., & Losos, J. B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology Letters*, 16(2), 175–182. <https://doi.org/10.1111/ele.12025>
- Warren, D. L., Cardillo, M., Rosauer, D. F., & Bolnick, D. I. (2014). Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology & Evolution*, 29(10), 572–580. <https://doi.org/10.1016/j.tree.2014.08.003>
- Williams, E. E. (1983). Ecomorphs, faunas, island size, and diverse end points in island radiations of *Anolis*. In R. B. Huey, E. R. Pianka, & T. W. Schoener (Eds.), *Lizard ecology: Studies of a model organism* (pp. 326–370). Cambridge, MA: Harvard University Press.
- Wilschut, R. A., Oplaat, C., Snoek, L. B., Kirschner, J., & Verhoeven, K. J. F. (2016). Natural epigenetic variation contributes to heritable flowering divergence in a widespread asexual dandelion lineage. *Molecular Ecology*, 25(8), 1759–1768. <https://doi.org/10.1111/mec.13502>
- Winchell, K. M., Reynolds, R. G., Prado-Irwin, S. R., Puente-Rolón, A. R., & Revell, L. J. (2016). Phenotypic shifts in urban areas in the tropical lizard *Anolis cristatellus*. *Evolution*, 70(5), 1009–1022. <https://doi.org/10.1111/evo.12925>
- Wollenberg Valero, K. C., Pathak, R., Prajapati, I., Bankston, S., Thompson, A., Usher, J., & Isokpehi, R. D. (2014). A candidate multimodal functional genetic network for thermal adaptation. *PeerJ*, 2, e578. <https://doi.org/10.7717/peerj.578>
- Xie, H. J., Li, H., Liu, D., Dai, W. M., He, J. Y., Lin, S., ... Qiang, S. (2015). ICE1 demethylation drives the range expansion of a plant invader through cold tolerance divergence. *Molecular Ecology*, 24(4), 835–850. <https://doi.org/10.1111/mec.13067>
- Yoder, J. B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N. D., & Tiffin, P. (2014). Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics*, 196(4), 1263–1275. <https://doi.org/10.1534/genetics.113.159319>
- Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, 328(5980), 916–919. <https://doi.org/10.1126/science.1186366>
- Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., ... Flicek, P. (2018). Ensembl 2018. *Nucleic Acids Research*, 46(D1), D754–D761. <https://doi.org/10.1093/nar/gkx1098>
- Zhang, Y.-H., Wang, I. J., Comes, H. P., Peng, H., & Qiu, Y.-X. (2016). Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure in a Tertiary relict species, *Emmenopterys henryi* (Rubiaceae). *Scientific Reports*, 6, 24041. <https://doi.org/10.1038/srep24041>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Wogan GOU, Yuan ML, Mahler DL, Wang IJ. Genome-wide epigenetic isolation by environment in a widespread *Anolis* lizard. *Mol Ecol*. 2020;29:40–55. <https://doi.org/10.1111/mec.15301>