Dewlap color variation in *Anolis sagrei* is maintained between habitats within islands of the West Indies

Raphaël Scherrer [[1]](#footnote-21), Colin M. Donihue,  
R. Graham Reynolds, Jonathan B. Losos and Anthony J. Geneva  
 Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology  
Harvard University, Cambridge, MA, USA  
 Department of Biology, University of North Carolina Asheville, Asheville, NC, USA  
 Current address: Groningen Institute for Evolutionary Life Sciences,  
Groningen, The Netherlands  
 Current address: Department of Biology, Washington University, St. Louis, MO, USA

Animal signals evolve in an ecological context. Moreover, locally adapting animal sexual signals can be especially important for initiating or reinforcing reproductive isolation during the early stages of speciation. Dewlap color in *Anolis* lizards can be highly variable between populations in relation to both biotic and abiotic adaptive drivers, albeit at relatively large geographical scales. Here, we investigated local adaptation of the dewlap across habitat-types at a small spatial scale, as this may give an indication of how conditions for the early stages of speciation may be met. We explored variation in dewlap coloration in one widespread species, *Anolis sagrei*, across three characteristic habitats spanning the Bahamas and the Cayman Islands. Using reflectance spectrometry as well as supervised machine learning, we found some consistent differences in spectral properties of the dewlap between habitats within small islands. Passive divergence in dewlap phenotype associated with isolation-by-distance did not explain our results. Instead, the observed patterns in dewlap coloration are more consistent with an adaptive explanation in these *A. sagrei* populations, as one would otherwise expect differences within islands to be erased by gene flow at such small geographical scales. Although these habitat-specific dewlap differences vary in magnitude and direction across islands, and islands themselves differ substantially, we found a suite of consistent archipelago-wide differences between habitat types, suggesting parallel responses to similar selective pressures. While at present, populations from these different habitats probably experience too much gene flow to follow distinct evolutionary lineages, should additional barriers arise between habitat-specific populations, the observed disruptive selection on dewlap coloration may facilitate ecological speciation.

**Keywords** — *Anolis*, reflectance, local adaptation, sexual signal, supervised machine learning

# Introduction

The staggering diversity of animal communication signals has long been of interest to evolutionary biologists. Animals use chemical, mechanical, electromagnetic, and visual signals to communicate in a wide variety of contexts, including, for example, competition for mates, species recognition, aposematism, and cooperation . A primary evolutionary factor shaping communication signals is the sensory system and behavior of their recipient(s) (the sensory drive hypothesis; ). Over the past decades, scientists have established that signals evolve in an ecological context and are dependent on environmental conditions . Just as different habitats may favor different combinations of eco-morphological traits to maximize performance and fitness , they may also shape different forms of a signal, so as to maximize its transmission and detection (e.g. ), or reduce its detection by unintended recipients such as predators . This selective pressure may drive the local adaptation of communication signals.  
One potential barrier to the maintenance of localized signal divergence is the homogenizing effect of gene flow. Population genetics theory suggests that gene flow may counteract local adaptation between localities and prevent divergence altogether, especially at small spatial scales, because of the inflow of maladapted alleles or because of the breaking of linkage between coevolving loci . This has been confirmed empirically in systems such as stick-insects and sticklebacks . Yet, examples of microgeographic adaptation, i.e. adaptation at smaller scales than the range of dispersal, exist, highlighting a high potential of some organisms to respond to selection in the face of gene flow (see and references therein). Examples include small scale adaptation in fragmented areas in Australian fruit flies , or local adaptation to predation pressure in North American salamanders . Therefore, despite evidence that local adaptation may be particularly difficult at small spatial scales where gene flow tends to cause adjoining populations to remain genetically homogeneous, the potential adaptive response of species traits, in particular communication signals, to localized differences in habitats remains relatively unknown .  
Lizards of the neotropical genus *Anolis* are a model system for studying the eco-evolutionary dynamics of local adaptation and natural selection . A particularly conspicuous trait of anoles is their dewlap; an extensible flap of skin that is typically sexually dimorphic and used as a communication signal in courtship , competition as well as in predator deterrence . Dewlap characteristics vary widely among the approximately species of the genus . Interspecific variation in dewlap coloration is implicated in species recognition , and possibly involved in speciation .  
Within species, studies have shown a link between variation in dewlap coloration and differences in habitats or climatic conditions . Some studies suggest that those differences may be adaptive, and that dewlaps may have evolved to maximize detectability given local light conditions . Other studies testing this hypothesis, however, found no pattern .  
Previous studies investigating variation in anole dewlaps compared populations at relatively large geographical scales, e.g. between islands or within large islands such as Puerto Rico or Hispaniola . These large scales should reduce gene flow . That said, examples do exist of divergence in dewlap coloration at smaller scales or between populations with high degrees of gene flow .  
The species *Anolis sagrei* is widespread across islands of the West Indies . It is a model organism in studies of local adaptation , biological invasion and sexual selection . Between-island variation in the mainly orange-red color of its dewlap was shown to be better explained by climatic variables than biotic factors such as sexual selection or predation pressure . How intra-island differences in habitat may contribute to the diversity of dewlap coloration, however, remains unexplored, and may reveal new insights into the scale of local adaptation despite gene flow.  
The island bank systems of the Bahamas and Cayman Islands comprise relatively small islands, with no major geographic barriers within islands limiting dispersal for this promiscuous species . These islands all share three characteristic native West Indian habitat-types – beach scrub bush, closed-canopy primary coppice forest, and mangrove forest – that are often spatially intermingled. These habitats contrast in environmental parameters including vegetation community, light irradiance, humidity and temperature . Each of these islands has been colonized independently by *A. sagrei* (, van de Schoot et al. unpubl.), such that these archipelagos constitute an ideal suite of natural replicates to explore within-island dewlap diversity across multiple islands.  
Here, we analyzed the color characteristics *A. sagrei* dewlaps within nine islands in the Bahamas and Cayman Islands, combining reflectance spectrometry and supervised machine learning. Our sampling design included sites in close proximity (the median distance between two sites within an island was km). We tested for divergence in dewlap phenotype between habitats within islands and between islands across the range of A. sagrei. We predicted that if light conditions in the environment indeed drive color evolution, dewlaps should be most similar between beach scrub and mangrove forest, which both have high levels of light irradiance, contrary to the darker, closed-canopy coppice forest. Similar, if detectability is maximized given the local conditions, we expected darker and more contrasting dewlaps in high irradiance habitats. Finally, if habitat characteristics are strong determinants of dewlap color variation, similar patterns should be observed across multiple islands . We found strong support for fine-scale, within-island differences in coloration between lizards inhabiting the three habitat-types in several color space dimensions, suggesting a potentially strong effect of divergent selection. However, the divergence patterns we observed did not match our *a priori* predictions and were inconsistent between islands. We found no evidence of isolation-by-distance as an explanation for the observed differences. Our results are nevertheless consistent with small-scale adaptive maintenance of signal polymorphism despite presumed considerable opportunity for gene flow.

# Methods

## Data collection

We sampled 466 lizards from seven islands in the Bahamas Archipelago – Abaco, North Andros, South Andros, South Bimini, Eleuthera, Long Island, Ragged Island – and two in the Cayman Islands – Cayman Brac and Little Cayman (Figure [fig:map]). These islands and island banks were chosen to span the West Indian range of *Anolis sagrei*. Three habitats were sampled on each island based on characterizations by and . Each habitat is clearly distinguishable by their dominant vegetation type — xeric coastal scrub (open, relatively dry habitat consisting of low vegetation or isolated trees), primary coppice forest (closed-canopy forest) and mangrove forest (wet coastal habitat with trees growing in brackish water and high light penetration). Sample sizes are given in Table [suptab:counts]. Our sampling design enabled us to test for differences between habitats at a coarse and fine geographical scale. The median distance between two localities within an island was km, with some islands being sampled at smaller or larger scales (Figure [supfig:map], Table [suptab:sites]). % of all pairwise distances within islands were below km. Additionally, there are no major barriers to dispersal (such as mountains) on any of the islands that we sampled.

## Reflectance measurements

We measured reflectance between 300 and 700nm wavelength, a range that encompasses the colors visible to most lizards and vertebrates in general . Measurements were taken with an Ocean Optics USB4000 spectrometer, a pulsed Xenon light source (PX-2, Ocean Optics, Largo, FL, USA) and a reflectance probe protected by a black anodized aluminum sheath. Measurements were taken with a 45-degree inclination to prevent specular reflection . The device was regularly standardized with a Spectralon white standard (Labsphere, North Sutton, NH, USA). Reflectance was measured at the center of the dewlap.

## Analysis

All analyses in this study were performed in R 3.6.1 .

### Dimensionality reduction

Reflectance curves were smoothened using the R package pavo as well as custom R functions, down to one reflectance value at each nanometer in wavelength from 300 to 700nm. Because neighboring wavelengths are highly collinear in reflectance, we reduced the dimensionality of the data using principal component analysis (PCA), as per and . We performed PCA on each island separately and systematically retained the first four principal components (PC), which together always explained more than of the variance across islands (Table [suptab:pcavariances]). PC1 explained between and % of the variance across islands; PC2 explained –%; PC3 –% and PC4 –%. The first four PCs explained similar proportions of variance when calculated for all islands together (Table [suptab:pcavariances]). PCs need not represent the same wavelengths across islands because they are fitted on different datasets. Nevertheless, PC1 was very collinear with brightness for all islands (Figure [supfig:brightness], Table [suptab:brightness]). PC2 correlated highly with the red and ultraviolet ends of the spectrum, which were inversely correlated with each other (Fig. [fig:anova]A). Higher PCs corresponded to various combinations of wavelengths. Because PC1 correlated uniformly with all wavelengths across the spectrum we considered PC2 onwards to capture the chromatic dimensions of color space, i.e. the relative contributions of the wavelengths regardless of brightness.

### Pooled analyses

In addition to within-island PCA, we performed a PCA on pooled data from the whole archipelago. The first four principal components explained 91.3% of the variance (Table [suptab:pcavariances]). Again PC1 strongly correlated with brightness (Fig. [supfig:brightness\_pooled], Table [suptab:brightness]). PC2 was positively correlated to short wavelengths (ultraviolet to blue) and negatively correlated to long wavelengths (green to red, Fig. [supfig:pooled]B). PC3 was strongly negatively correlated with UV reflectance and positively correlated with blue-green. PC4 was made of a mosaic of wavelengths, correlating positively with blue and red but negatively with ultraviolet and yellow.  
We used this dataset to partition the variance in dewlap coloration among islands, habitats and habitats within islands, using a two-way multivariate analysis of variance (MANOVA) with an interaction term. However, because the assumptions of parametric MANOVA were violated for all islands but Ragged Island (multivariate normality, Henze-Zirkler’s test, , R package MVN, , Table [suptab:multinorm]; and homogeneity of covariance matrices, Box’s M-test, , R package heplots, , Table [suptab:covariance]), we used a semi-parametric MANOVA instead (R package MANOVA.RM, ), with P-values calculated from a bootstrap procedure with 1,000 iterations. We calculated the proportion of variance explained by islands, habitats and the habitat-by-island interaction using partial effect sizes on a MANOVA-approximation of the analysis (R package heplots, ).

### Machine learning

Because of the aforementioned violations of the MANOVA assumptions, and to reduce the chances of false discovery, we conducted multivariate group comparisons using support vector machines (SVMs), a model-free, powerful nonparametric supervised machine learning technique.  
Machine learning for group comparison has become more popular in ecology and evolution in the recent years (e.g. ). In particular, SVMs are designed to find the best possible nonlinear boundaries between labelled groups of points in multidimensional spaces, without assumptions about the distribution of the data . This makes them well suited to field biological data, which often violate the assumptions of classical linear modeling and can be, as in the case of coloration, inherently highly multivariate . First, a machine is trained to recognize differences between groups within a subset of the data called the training set. Significance of differences is then assessed by testing the accuracy of that fitted machine in predicting the group-labels of data points that were not included in the training, called a testing set, based solely on their multivariate coordinates. This cross-validation procedure results in a proportion of correctly classified points, or generalization accuracy score, which can be compared to that expected under random guessing using a binomial test.  
In this study, we performed SVM classifications on each island separately. We used a standard five-fold cross-validation procedure, where the data were randomly split into five bins of approximately equal sizes. Each bin was in turn taken as the testing set while the rest was used as a training set, thus resulting in five trained machines per cross-validation. We replicated this procedure 100 times for each island to account for stochastic outcomes. We performed binomial tests to evaluate the significance of deviations in observed mean generalization accuracy per island to null expectations under random guessing. Each training data set was downsampled to the size of its least represented habitat to ensure balanced training samples. We ensured that each habitat was represented by at least five data points in the training set.  
All classification analyses were repeated using the more classical linear discriminant analysis (LDA), a supervised machine learning technique finding linear boundaries that maximize the differences between groups, albeit assuming multivariate normality and homogeneity of covariance matrices . We used the R package rminer for SVMs, and MASS for LDAs. We used rminer’s default heuristic search option to automatically tune the Gaussian kernel parameter and the complexity parameter for the SVMs.  
The same procedure was repeated on principal components from the whole archipelago (see Pooled analyses) to evaluate the significance of archipelago-wide differences in dewlap coloration across habitats.  
All machine learning classifications performed on principal components were also repeated on the original reflectance datasets reduced to 50-nm spaced wavelengths from 300 to 700nm.  
We conducted one-dimensional sensitivity analyses using rminer to determine the relative importance of the different input variables during classification where significant differences were detected, both on machines trained on principal components and machines trained on non-transformed reflectance at various wavelengths. In parallel, we conducted univariate analyses of variance to independently test the importance of different variables in between-habitat variation, on islands where the machines detected significant differences based on binomial tests (next section).

### Univariate analyses

For each island where significant differences in multivariate dewlap coloration were detected between habitats, we used multiple univariate analyses of variance (ANOVA) to identify which variables were responsible for the observed differences. We constructed our ANOVA models in two steps, as per . In a first step, we accounted for heterogeneity of variances across groups by systematically comparing the goodness-of-fit of an ANOVA model estimated with ordinary least squares (OLS) with that of a model estimated with generalized least squares (GLS), which allowed one estimate of residual variance per habitat (using the R package nlme, ). Both models were fitted with restricted maximum likelihood (REML). Goodness-of-fit was estimated using Akaike’s Information Criterion corrected for small sample sizes (AICc, R package MuMIn, ), and the estimation method yielding the lowest AICc was retained. In a second step, we re-fitted the retained model with maximum likelihood (ML) to test for the effect of habitat-type using likelihood ratio tests (LRT) between a model including a habitat-term and a null model lacking the habitat-term.  
We tested the assumptions of the parametric ANOVA for each island included in the univariate analyses. For all islands where deviations from multivariate normality were detected in at least one habitat (Table [suptab:multinorm]), we assessed univariate normality for each principal component (Shapiro-Wilk’s test, Table [suptab:normality]). For skewed PCs that deviated significantly from normality, we repeated the analysis using a nonparametric Kruskal-Wallis tests . We found no multivariate outliers based on the Mahalanobis distance (package MVN, ). We used the cases of better fit of the GLS model relative to the OLS model as evidence for heterogeneity of variances, which were then accounted for by the GLS approach (Table [tab:anova]).  
Significant post hoc contrasts were assessed using Tukey’s Honest Significant Difference (HSD) test whenever the assumptions of normality and homogeneity of variances was met , Dunnett’s T3 method when only homogeneity of variances was violated but not normality , and Nemenyi’s test when normality was violated . All post hoc tests were performed with the R package PMCMRplus .  
We used the same procedure to investigate which variables, if any, were involved in archipelago-wide multivariate differences between habitats detected in our two-way MANOVA design (see Pooled analyses). However, in the first step or our model comparison procedure, we added mixed-effect equivalents of our OLS and GLS models, this time with island as a random effect. The resulting four models were compared and the best fitting variance structure was retained as explained above.

### Spatial autocorrelation

We tested for within-island spatial autocorrelation between the geographical distances among sampling sites and their euclidean distances in multivariate color space (mean PC1 to PC4 per site, Table [suptab:sites]), regardless of habitat-type. Because often only a few sites were sampled per island, we could not get meaningful results from tests that use sites as units of observation, such as Moran’s I test . Instead, we designed a permutation test where we randomly reshuffled individual lizards across sites within islands 1,000 times each, and systematically recalculated Pearson’s correlation coefficient between geographic distances (computed as geodesic distances, R package geosphere, ) and phenotypic distances. We used the resulting null distributions of correlation coefficients to assess the significance of the observed spatial autocorrelation for each island.

### Site differences

In this study, we were interested in the minimum spatial scale at which significant differences between habitats could be detected within islands. We performed multiple pairwise nonparametric Wilcoxon-Mann-Whitney’s tests to compare dewlap coloration between sites with different habitat-types, for each pair of habitats and each variable where significant differences were detected with our analyses of variances. The P-values were adjusted using a Benjamini-Hochberg correction for multiple testing .

# Results

We tested for variation in *A sagrei* dewlap coloration between populations living in three characteristic habitat types across nine islands that span the West Indian range of the brown anole (Fig [fig:map], [supfig:map]). We found that most of the variation in coloration partitioned between islands (two-way semi-parametric MANOVA, modified ANOVA-type statistic (MATS) = 2009.6, P < 0.001, Fig. [supfig:reflectance], explained variance %, MANOVA approximation). Nonetheless, we did find evidence for differences in dewlap coloration between habitat-types, and those were mostly island-specific (habitat-by-island interaction term, MATS = 384.4, P < 0.001, explained variance %), leaving a small but significant portion of the variation explained by an archipelago-wide habitat effect (MATS = 42.5, P = 0.001, %).  
The small archipelago-wide effect of habitat-type was detected for PC1, PC2 and PC3 (mixed-effect ANOVA with island as a random effect, Table [suptab:anova-pooled]), but this effect was too small for post hoc tests to find which habitats differed. Archipelago-wide differences in dewlap coloration between habitats were also detected by SVMs trained on pooled data regardless of island identity, both for PCA data and reflectance scores (Fig. [supfig:classif-svm-pca-pooled], [supfig:classif-svm-refl-pooled]). This pattern seemed to be driven by mangrove lizards being correctly reassigned more often than predicted by chance. Sensitivity analyses on these machines suggest a relatively small role of long wavelengths (red reflectance) in driving this pattern (Fig. [supfig:importance-svm-refl-pooled]), but did not reveal strong differences between the PCs in relative importance (Fig. [supfig:importance-svm-pca-pooled]). Archipelago-wide differences were not detected by LDA classifiers at all (Fig. [supfig:classif-lda-pca-pooled], [supfig:classif-lda-refl-pooled]). So the difference could be in something else than means.  
Within islands, SVM classifiers correctly assigned individuals to their habitat of origin based solely upon dewlap coloration on five islands: Abaco, Bimini, Cayman Brac, Little Cayman, and Long island (Fig. [fig:classif-svm-pca]). An LDA approach yielded similar success rates (Fig. [supfig:classif-lda-pca]), suggesting robust differences between these populations. Of the five islands, Little Cayman was the best discriminated with a mean SVM generalization success of 73.4% (Table [suptab:classif-svm-pca]). The results of the classification analyses on PCA data were very similar to results from SVMs and LDAs trained on reflectance values at 50nm-spaced wavelengths from 300 to 700nm (Fig. [supfig:classif-svm-refl] and [supfig:classif-lda-refl]).  
Differentiation in dewlap coloration occurred in multiple dimensions of color space. Moreover, the differences in dewlaps between habitats were not always consistent between islands, thus, we will discuss the habitat-specific variation in dewlap coloration for each island where significant differences were detected in turn (Fig. [fig:anova], Tables [tab:anova], [suptab:kruskal]). Figure [fig:anova]A provides a key to map principal component scores to the underlying wavelengths.  
On Abaco, dewlaps did not differ in PC1, which represents brightness. Mangrove lizards had significantly lower PC2 scores, corresponding to higher ultraviolet reflectance and lower red reflectance. Coastal beach scrub lizards had lower scores on PC3, corresponding to lower ultraviolet reflectance and higher blue reflectance.  
On Bimini, coastal beach scrub lizards had significantly brighter dewlaps than lizards from mangroves (PC1), but mangrove lizards had higher PC2 scores than beach scrub lizards, indicating higher violet and blue reflectance, and lower red reflectance. Lizards from primary coppice had higher PC3 scores overall, which correlated very positively with ultraviolet reflectance.  
On Cayman Brac, coppice-lizard dewlaps were significantly less bright than lizards from the other habitats. Coastal beach scrub lizards had dewlaps that scored low on PC2, corresponding to lower violet-blue and more red, while the mangrove lizards exhibited the opposite: relatively higher levels of violet-blue and less red. In PC3 space we found that dewlaps from lizards in the coastal habitat had high ultraviolet reflectance, coppice lizards had intermediate levels, and mangrove lizards had relatively low levels.  
On Little Cayman, the dewlaps of coppice lizards were significantly darker (PC1) than coastal-lizards. Mangrove lizards had less ultraviolet and redder dewlaps (PC2). The dewlaps of the coastal beach scrub lizards had higher levels of red and ultraviolet reflectance and less blue reflectance than the dewlaps of the other habitat-populations (PC3).  
On Long Island, lizards from the coppice habitat had darker dewlaps than lizards from the other habitats (PC1). Coastal lizards had relatively more ultraviolet and less blue-green reflectance in their dewlaps (PC3). These coastal-habitat lizards also scored lower on PC4, corresponding to slightly more violet and green-yellow dewlaps, and less blue dewlaps, than the mangrove lizards on the island.  
Sensitivity analyses on classifiers suggested an overall higher relative importance for PC2 and PC3 in determining between-group differences on Abaco, both in SVM and LDA classifiers (Fig. [supfig:importance-svm-pca], [supfig:importance-lda-pca]), consistent with our ANOVA results (Fig. [fig:anova]B). There was no strong signal of differences in relative importance among principal components on the other islands. Sensitivity analyses of SVMs trained on reflectance scores rather than principal components revealed, however, a consistently higher importance of ultraviolet reflectance in between-group differences on all islands (Fig. [supfig:importance-svm-refl]). This pattern was not recovered for LDAs trained on reflectance scores (Fig. [supfig:importance-lda-refl]).  
We did not find significant spatial autocorrelation between the sampling sites on the islands where we detected a significant habitat effect. We did, however, detect a significant positive signal of autocorrelation on Eleuthera (Table [suptab:autocor]), suggesting possible color differentiation through isolation-by-distance on this island.  
In contrast, differences in dewlap coloration between habitats were often detected in close geographical proximity. For example, Bimini, Cayman Brac and Little Cayman were among the smallest islands in our study (Fig. [supfig:map]). Besides, for pairs of habitats where significant differences in dewlap coloration were detected along some principal components, comparisons of the actual sampling sites indicate that the most detectable differences involved sites that were mostly 5-10km apart. Our most extreme case of significant differences occurred for PC3 between the beach scrub site and a coppice site, separated from each other by a a few hundreds of meters at most on Bimini (multiple pairwise Wilcoxon-Mann-Whitney tests, Fig. [supfig:distances]).  
Patterns of differentiation were inconsistent across the five most significant islands. Contrasts in principal components between habitats, calculated on pooled data from the whole archipelago, were not similar, for any component, among islands (Fig. [supfig:contrasts]; MANOVA, Pillai’s trace = 0.354, , P = 0.36). No pattern of variation was shared by all five significant islands, along any dimension. Some patterns did seem more common however, such as darker dewlaps among coppice lizards (Cayman Brac, Little Cayman, and Long Island, Fig. [fig:anova]) or the intermediate position of coppice lizards in chromatic color space (Cayman Brac and Long Island). In other cases, patterns of differentiation were reversed from one island to another, with more ultraviolet reflecting dewlaps in mangroves than in coastal habitat on Abaco and Cayman Brac, but the opposite on Little Cayman and Long Island. Overall, it seemed that patterns of heterogeneity of variance were often driven by higher variances in coloration within beach scrub lizards (Fig. [fig:anova], Table [tab:anova]). Yet other patterns were idiosyncratic, such as the combination of higher red and ultraviolet reflectance in coastal lizards on Little Cayman, where the rule seemed to be a negative correlation between ultraviolet and red reflectance across every other island.

# Discussion

#### Dewlap coloration differs between habitat-types

We found that male dewlap coloration in *A. sagrei* significantly varied between habitat-types (beach scrub bush, primary coppice forest and mangrove forest) on five islands of the West Indies: Abaco, Bimini, Cayman Brac, Little Cayman and Long Island. However, the habitat-specific variation in dewlaps was not consistent between these islands. Although those results are consistent with adaptation at a very local scale, other evolutionary drivers could be at work, including phenotypic plasticity, random drift, or historical contingency, including multiple colonization events. We reject this last explanation because all of the island populations in this study are strictly monophyletic, implying a single colonization event per island (van de Schoot, unpublished thesis; ). While random drift cannot be completely ruled out, we see little evidence for a role of phenotypic isolation-by-distance (spatial autocorrelation) in explaining the differences we report. We did detect a significant signal of isolation-by-distance on Eleuthera, but we did not find differences in dewlap coloration between habitats on this island.

#### Divergence is maintained in the face of gene flow

We detected differences between habitats at relatively small spatial scales, most of the time between sites 5-10km apart, and sometimes a few hundred meters away. This, together with the lack of geographic barriers within islands and the high mobility of *A. sagrei* individuals , imply ample opportunity for extensive gene flow between our study populations. In agreement with this, populations from different habitats were found to not be monophyletic based on <insert marker name here> (van de Schoot et al. unpublished thesis). The maintenance of phenotypic differences at such a small spatial scales, in the face of gene flow, may support an adaptive explanation .  
The persistence of dewlap color divergence despite gene flow has been documented in multiple species of anoles, sometimes in relation to environmental conditions. and found divergent dewlap coloration in the face of gene flow between subspecies of *A. distichus* across Hispaniola, and proposed this as a mechanism of reproductive isolation in the early stages of speciation . found that dewlap color polymorphism was maintained in the absence of genetic structure between populations of *A. apletophallus* from Panama. found that divergence in dewlap coloration matched habitat-type better than mitochondrial lineage in *A. roquet* on Martinique, and a convergent pattern was found in *A. trinitatis* on the featureless island of St Vincent . Finally, divergence in body coloration, but not dewlap coloration, was also reported in *A. conspersus* on another small island, Grand Cayman . Altogether, our results align with previous studies suggesting a generally high adaptive potential of dewlap coloration in anoles over small distances, which could apply to *A. sagrei*.

#### Dewlap coloration could be locally adapted

Presently, we do not know the adaptive drivers of dewlap color divergence observed in this study. Some degree of parallel evolution is usually a good indicator for an adaptive process , and convergent patterns of dewlap color evolution in similar environments across islands and species have been documented in other species . Here, however, we found little consistency in between-habitat divergence across the nine islands studied; except maybe for dewlap brightness being lower in the primary coppice habitat on three islands, and the primary coppice habitat being intermediate in coloration between mangrove and beach scrub on two islands. Those patterns could reflect adaptation to components of the habitat , but it is not clear what those components of the habitat might be. Although those patterns could reflect adaptation to components of the habitat , it is not clear what those components might be. Alternatively, neutral processes could produce patterns similar to local adaptation, such that this possibility cannot be completely ruled out (full subsection for this?).  
Previous studies have suggested that dewlap coloration may have evolved to be maximally detectable under local light conditions imposed by the environment, primarily through UV contrast (i.e. UV-brighter dewlaps in UV-dark, mesic habitats and UV-darker dewlaps in UV-bright, xeric habitats), in *A. cristatellus* and *A. cooki* from Puerto Rico . Although UV reflectance was commonly involved in between-habitat divergence, we found no such patterns in *A. sagrei*, where instead, we found the darkest dewlaps in the darkest, mesic habitat – primary coppice forest – on three islands, and dewlaps often differed the most between beach scrub and mangrove forest, two xeric habitats with similar, high irradiance levels . The inconsistent and idiosyncratic patterns we observed suggest that dewlap color variation between habitats cannot be predicted by habitat identity alone. Studies of Jamaican and Hispaniolan anoles similarly found between-habitat differences in dewlap coloration but no evidence for higher detectability . Habitats on different islands may also differ in other aspects than light conditions such as densities of predators or other anole species, which have been shown to affect among-island dewlap diversity . In particular, recently showed that dewlaps with spotted patterns occurred more often in *A. sagrei* on islands with more coexisting species of anoles. Therefore, dewlap coloration in may be under the influence of a mosaic of interacting local environmental factors, which need not be the same across islands.

#### A role of phenotypic plasticity is unlikely

Phenotypic plasticity could explain differences in coloration between habitat populations because of diet differences. The yellow, orange and red coloration in anoline dewlaps are produced by pterins and carotenoids . Animals lack the ability to synthesize carotenoids, and those must therefore be found in the diet, while pterins are synthesized from nucleotides . However, experimental manipulation of dietary carotenoid content showed no effect on dewlap coloration in *A. sagrei* nor in *A. distichus* , which also has an orange-based dewlap. This makes a plastic response to differences in diet across habitats unlikely. Developmental plasticity (e.g. due to differences egg rearing conditions) is also unlikely because dewlap coloration develops at sexual maturity in anoles . The differences we observed could therefore be heritable. This hypothesis is further supported by , who found a high degree of heritability of dewlap coloration in *A. sagrei*. Although most studies used one or two-generation common garden experiments and thus could not rule out transgenerational plastic effects , dewlap coloration generally seems to not be a plastic trait. This further reinforces an adaptive explanation, where dewlap color could be under differential natural and/or sexual selection in these different habitats.

#### Sexual selection could be at play

Substantial levels of promiscuity in *A. sagrei* suggest ample opportunity for female mate choice and sexual selection . A number of studies have suggested that characteristics of the dewlap are indicators of male quality and may therefore act as a cue in mate choice, according to the “good genes” model of sexual selection . found lower orange reflectance in dewlaps with heavily parasitized *A. brevirostris*, suggesting a trade-off in carotenoid use between the immune response and pigment deposition. found that lower UV and orange-red reflectance predict contest-winning success between males. further found that more yellow and red dewlaps (relative to UV) predict better body condition, and that higher yellow and UV reflectance at the margin of the dewlap predict higher hematocrit (the concentration of red blood cells), indicating a better health.  
Other aspects of the dewlap may be important for sexual selection. For example, dewlap size in was found to correlate with bite force and sexual size dimorphism (a proxy for sexual selection; ). Even though display frequency increases in the presence of females , no link was found between dewlap display frequency and mating success or individual quality . Our report of inconsistent divergence between islands, however, is at odds with the good-genes model of sexual selection, which would predict the evolution of dewlap coloration in the same direction across the archipelago. This is in agreement with , who found no link between the average island-dewlap coloration and sexual size dimorphism in *A. sagrei* across the West Indies. If sexual selection is at play in our system, it more likely involves divergent female preferences that are not linked to male quality but rather to components of the environment, or may well be arbitrary (Fisherian sexual selection; ).

#### Implications in the context of speciation

Local adaptation can be a precursor to ecological speciation, a process that may have given rise to the adaptive radiation of *Anolis* lizards . Ecologically-mediated divergence of a sexual signal may be a potent path to the evolution of reproductive isolation through divergent sexual selection . Evidence suggests that dewlap coloration could take this role in anoles , or at least that it is frequently involved in species recognition . Although this signal is not detected at the phylogenetic scale of the whole genus , sexual signals are often evolutionarily very labile , and the anole dewlap in particular is capable of rapid macroevolution; for example, *A. conspersus* on Grand Cayman evolved a UV-blue dewlap from an ancestral orange dewlap in 2 to 3 million years . We present evidence of multiple cases of potentially adaptive maintenance of habitat-associated dewlap divergence over small geographical scale in *A. sagrei* across the West Indies. While these populations do not appear to be in the process of speciation, our results strongly suggest that the anoline dewlap has enough micro-scale, local adaptive potential to participate in the build-up of reproductive isolation, should it be recruited for assortative mating.

## Acknowledgements

Collection permission was granted by the Bahamas Environment, Science and Technology Commission, the Bahamas National Trust, and the Cayman Islands Department of the Environment. The authors thank Sofia Prado-Irwin, Pavitra Muralidhar, Nicholas Herrmann, Richard E. Glor, Alberto R. Puente-Rolón, Kevin Aviles-Rodriguez, Kristin Winchell, and Melissa Kemp for assistance in the field. Funding for this work was provided by the Templeton Foundation (to JBL), NSF DEB #1500761 (to AJG), NSF DBI #1609284 (to CMD), and a Harvard Museum of Comparative Zoology Putnam Grant (to RGR).

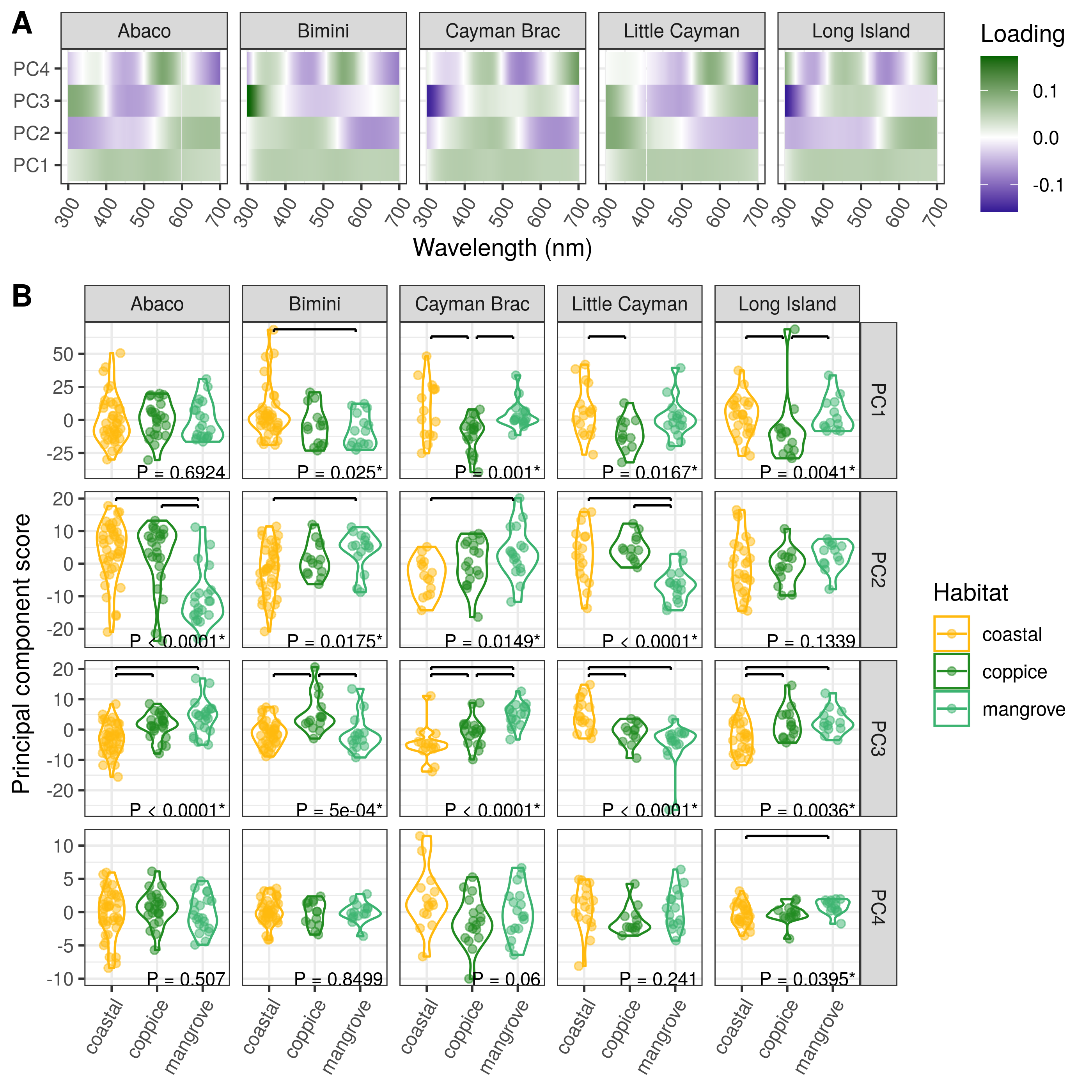
# Figures

![Map of the West Indies with sampled islands highlighted in black.](data:application/pdf;base64,)

Map of the West Indies with sampled islands highlighted in black.

SVM classification accuracy across islands based on principal component data. Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The dashed line is the density of a corresponding null binomial distribution, which would be expected under random guessing (testing sets with 20% of the observations for each island and success probability of ). Inset plots show the corresponding average confusion matrices and represent the proportion of lizards from each habitat (columns) reassigned in each other habitat (rows), with an interpretation guide in the right panel. Binomial test P-values indicate deviations of the mean classification accuracy to the null distribution. \*, P < 0.05.

SVM classification accuracy across islands based on principal component data. Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The dashed line is the density of a corresponding null binomial distribution, which would be expected under random guessing (testing sets with 20% of the observations for each island and success probability of ). Inset plots show the corresponding average confusion matrices and represent the proportion of lizards from each habitat (columns) reassigned in each other habitat (rows), with an interpretation guide in the right panel. Binomial test P-values indicate deviations of the mean classification accuracy to the null distribution. \*, P < 0.05.



Dewlap color variation between habitat-types across the most significant islands. (A) Mapping of reflectance at various wavelengths onto the principal components (loadings from the PCA rotation matrix). (B) Distribution of PC scores between habitats along the first four PCs on each island where significant between-habitat differences were detected using SVMs. P-values are reported for univariate ANOVA (or Kruskal-Wallis tests when applicable, see Methods). Post hoc significant differences at a 0.05 error rate are indicated with horizontal bars. \*, P < 0.05.

# Tables

[tab:anova]

# Supplementary Figures

![Map of the sampling sites and corresponding habitats across nine islands of the West Indies.](data:application/pdf;base64,)

Map of the sampling sites and corresponding habitats across nine islands of the West Indies.

Correlation between dewlap brightness (as measured by the mean reflectance from 300 to 700nm in wavelength) and PC1 score for each island. Pearson’s squared correlation coefficients are reported. \*, P < 0.05.

Correlation between dewlap brightness (as measured by the mean reflectance from 300 to 700nm in wavelength) and PC1 score for each island. Pearson’s squared correlation coefficients are reported. \*, P < 0.05.

Correlation between dewlap brightness (as measured by the mean reflectance from 300 to 700nm in wavelength) and PC1 score across the whole archipelago. Pearson’s squared correlation coefficient is reported. \*, P < 0.05.

Correlation between dewlap brightness (as measured by the mean reflectance from 300 to 700nm in wavelength) and PC1 score across the whole archipelago. Pearson’s squared correlation coefficient is reported. \*, P < 0.05.

5-95th percentile range of lizard dewlap reflectance values (in % of incoming light) across wavelengths for each island and each habitat.

5-95th percentile range of lizard dewlap reflectance values (in % of incoming light) across wavelengths for each island and each habitat.

Archipelago-wide SVM classification accuracy based on principal component data. Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Archipelago-wide SVM classification accuracy based on principal component data. Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Archipelago-wide SVM classification accuracy based on reflectance data at 50nm-intervals in wavelength (see Methods). Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Archipelago-wide SVM classification accuracy based on reflectance data at 50nm-intervals in wavelength (see Methods). Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Sensitivity analyses of the different input variables in the archipelago-wide SVM classification on principal component data (Figure [supfig:classif-svm-pca-pooled]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide SVM classification on principal component data (Figure [supfig:classif-svm-pca-pooled]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide SVM classification on reflectance data at 50nm-intervals in wavelength (Figure [supfig:classif-svm-refl-pooled]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide SVM classification on reflectance data at 50nm-intervals in wavelength (Figure [supfig:classif-svm-refl-pooled]), with relative importance computed for every machine.

Archipelago-wide LDA classification accuracy based on principal component data. Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Archipelago-wide LDA classification accuracy based on principal component data. Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Archipelago-wide LDA classification accuracy based on reflectance data at 50nm-intervals in wavelength (see Methods). Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

Archipelago-wide LDA classification accuracy based on reflectance data at 50nm-intervals in wavelength (see Methods). Machines were trained on individual dewlaps regardless of island identity. The histogram shows the accuracy distribution over 100 replicates for each five cross-validation bins. The legend is the same as in Figure [fig:classif-svm-pca].

LDA classification accuracy across islands based on principal component data. Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The legend is the same as in Figure [fig:classif-svm-pca].

LDA classification accuracy across islands based on principal component data. Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The legend is the same as in Figure [fig:classif-svm-pca].

SVM classification accuracy across islands based on reflectance data at 50nm-intervals in wavelength (see Methods). Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The legend is the same as in Figure [fig:classif-svm-pca].

SVM classification accuracy across islands based on reflectance data at 50nm-intervals in wavelength (see Methods). Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The legend is the same as in Figure [fig:classif-svm-pca].

LDA classification accuracy across islands based on reflectance data at 50nm-intervals in wavelength (see Methods). Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The legend is the same as in Figure [fig:classif-svm-pca].

LDA classification accuracy across islands based on reflectance data at 50nm-intervals in wavelength (see Methods). Histograms show accuracy distributions over 100 replicates for each five cross-validation bins per island. The legend is the same as in Figure [fig:classif-svm-pca].

Sensitivity analyses of the different input variables in the within-island SVM classification on principal component data (Figure [supfig:classif-svm-pca]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the within-island SVM classification on principal component data (Figure [supfig:classif-svm-pca]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the within-island LDA classification on principal component data (Figure [supfig:classif-lda-pca]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the within-island LDA classification on principal component data (Figure [supfig:classif-lda-pca]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide SVM classification on reflectance at 50nm-intervals in wavelength (Figure [supfig:classif-svm-refl]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide SVM classification on reflectance at 50nm-intervals in wavelength (Figure [supfig:classif-svm-refl]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide LDA classification on reflectance at 50nm-intervals in wavelength (Figure [supfig:classif-lda-refl]), with relative importance computed for every machine.

Sensitivity analyses of the different input variables in the archipelago-wide LDA classification on reflectance at 50nm-intervals in wavelength (Figure [supfig:classif-lda-refl]), with relative importance computed for every machine.

Spatial scale of between-habitat differences in dewlap coloration. For each variable and each pair of habitats where significant differences were detected (Figure [fig:anova]), we performed multiple post hoc pairwise comparisons between the sites involved (Figure [supfig:map], Table [suptab:sites]), using nonparametric Wilcoxon-Mann-Whitney tests. Here we report, for each pair of habitats, the distances between sites that significantly differed in dewlap coloration at an error rate of 0.05 (P-values corrected with the Benjamini-Hochberg procedure for multiple testing).

Spatial scale of between-habitat differences in dewlap coloration. For each variable and each pair of habitats where significant differences were detected (Figure [fig:anova]), we performed multiple post hoc pairwise comparisons between the sites involved (Figure [supfig:map], Table [suptab:sites]), using nonparametric Wilcoxon-Mann-Whitney tests. Here we report, for each pair of habitats, the distances between sites that significantly differed in dewlap coloration at an error rate of 0.05 (P-values corrected with the Benjamini-Hochberg procedure for multiple testing).

Test of parallel divergence between islands. Differences in habitat-means, or contrasts, are shown for two pairs of habitats for each principal component on each island, rescaled so the standard deviation of the means along each principal component is one. The contrasts represent the patterns of between-habtiat variation on each island, for a given principal component. The absence of clustering of islands by variable indicates that islands differ in their between-habitat divergence patterns. This is confirmed by a non-significant MANOVA test of the between versus within-variable variance in contrasts.

Test of parallel divergence between islands. Differences in habitat-means, or contrasts, are shown for two pairs of habitats for each principal component on each island, rescaled so the standard deviation of the means along each principal component is one. The contrasts represent the patterns of between-habtiat variation on each island, for a given principal component. The absence of clustering of islands by variable indicates that islands differ in their between-habitat divergence patterns. This is confirmed by a non-significant MANOVA test of the between versus within-variable variance in contrasts.

# Supplementary Tables

[suptab:counts]

[suptab:sites]

[suptab:pcavariances]

[suptab:brightness]

[suptab:multinorm]

[suptab:covariance]

[suptab:normality]

[suptab:anova-pooled]

[suptab:classif-svm-pca]

[suptab:kruskal]

[suptab:autocor]

1. Corresponding author: r.scherrer@rug.nl [↑](#footnote-ref-21)