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

Raphaël Scherrer,[[1]](#footnote-2) 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,  
University of Groningen, Groningen, The Netherlands  
 Current address: Department of Biology, Washington University, St. Louis, MO, USA  
 Current address: Department of Biology, Center for Computational and Integrative Biology,  
Rutgers University–Camden, Camden, NJ, USA

Animal signals evolve in an ecological context. Locally adapting animal sexual signals can be especially important for initiating or reinforcing reproductive isolation during the early stages of speciation. Previous studies have demonstrated that dewlap color in *Anolis* lizards can be highly variable between populations in relation to both biotic and abiotic adaptive drivers at relatively large geographical scales. Here, we investigated differentiation of dewlap coloration among habitat types at a small spatial scale, within multiple islands of the West Indies, as this may give new insights into the local scale at which adaptation is possible. We explored variation in dewlap coloration in the most widespread species of anole, *Anolis sagrei*, across three characteristic habitats (name them?) spanning the Bahamas and the Cayman Islands. Using reflectance spectrometry paired with supervised machine learning, we found significant differences in spectral properties of the dewlap between habitats within small islands, sometimes over very short distances. Passive divergence in dewlap phenotype associated with isolation-by-distance did not explain our results. On the other hand, these habitat-specific dewlap differences varied in magnitude and direction across islands, and thus our primary test for adaptation – parallel responses across islands – was not supported. We suggest that neutral processes or selection could be involved in several ways, including sexual selection. Our results shed new light on the scale at which signal color polymorphism can be maintained in the presence of gene flow, and the relative role of local adaptation and other processes in driving these patterns.

**Keywords** — reflectance, adaptation, sexual signal, machine learning, polymorphism

# 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, competition for mates, species recognition, aposematism, and cooperation (Bradbury & Vehrencamp, 2011). A primary evolutionary factor shaping communication signals is the sensory system and behavior of recipients (the sensory drive hypothesis; Endler & McLellan (1988; Endler, 1992, 1998)). Over the past decades, scientists have established that signals evolve in an ecological context and are dependent on environmental conditions (Endler, 1992, 1993b; a). Just as different habitats may favor different combinations of ecomorphological traits to maximize performance and fitness (Arnold, 1983), they may also shape different forms of a signal, so as to maximize its transmission and detection (e.g. Seehausen (1997)), or reduce its detection by unintended recipients such as predators (Endler, 1984, 1990, 1991; Halfwerk *et al.*, 2014). 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 (Felsenstein, 1976; García-Ramos & Kirkpatrick, 1997; Dieckmann & Doebeli, 1999; Lenormand, 2002; Hendry *et al.*, 2007a). This genetic homogenization has been confirmed empirically in systems such as stick insects (Nosil & Crespi, 2004) and stickleback (Hendry *et al.*, 2007b). Yet, examples of microgeographic adaptation, i.e. adaptation at smaller scales than the range of dispersal, exist, highlighting the potential of some organisms to respond to selection in the face of gene flow (see Richardson *et al.* (2014) and references therein). Examples include small scale adaptation in fragmented areas in Australian fruit flies (Willi & Hoffmann, 2012) and local adaptation to predation pressure in North American salamanders (Richardson & Urban, 2013). 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 (Richardson *et al.*, 2014).Lizards of the neotropical genus *Anolis* are an excellent group for studying the eco-evolutionary dynamics of local adaptation and natural selection (Losos, 2009). 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 (Sigmund, 1983; Driessens *et al.*, 2014, 2015) and territorial displays (Losos, 1985; Macedonia & Stamps, 1994; Macedonia *et al.*, 2013) as well as in predator deterrence (Leal & Rodríguez-Robles, 1995, 1997; Leal & Rodriguez-Robles, 1997). Dewlap characteristics vary widely among the approximately species of the genus (Nicholson *et al.*, 2007). Interspecific variation in dewlap coloration is implicated in species recognition (Williams, 1969; Rand & Williams, 1970; Williams & Rand, 1977; Losos, 1985; Macedonia & Stamps, 1994; Fleishman, 2000; Macedonia *et al.*, 2013), and this function could have had a role in initiating or reinforcing reproductive isolation during speciation (Lambert *et al.*, 2013; Geneva *et al.*, 2015; Ng *et al.*, 2017).  
Within species, studies have shown a link between variation in dewlap coloration and differences in habitats or climatic conditions (Macedonia, 2001; Leal & Fleishman, 2002, 2004; Thorpe & Stenson, 2002; Thorpe, 2002; Vanhooydonck *et al.*, 2009; Ng *et al.*, 2012, 2013, 2016; Driessens *et al.*, 2017). Some studies suggest that those differences may be adaptive and that dewlaps may have evolved to maximize detectability given local light conditions (Fleishman & Persons, 2001; Leal & Fleishman, 2002, 2004). Although this claim is further supported by recent findings that dewlap colors are perceived differently under different levels of shading (Fleishman *et al.*, 2020), other studies found conflicting patterns of between-habitat variation that did not support the sensory drive hypothesis (Fleishman *et al.*, 2009; Ng *et al.*, 2012; Macedonia *et al.*, 2014).  
Previous studies investigating variation in anole dewlaps compared populations at relatively large geographical scales, e.g. between islands (Vanhooydonck *et al.*, 2009; Driessens *et al.*, 2017) or within large islands such as Puerto Rico (Leal & Fleishman, 2004) or Hispaniola (Ng *et al.*, 2012, 2016). These large scales and marine barriers should reduce gene flow (Ng & Glor, 2011; Lambert *et al.*, 2013; Richardson *et al.*, 2014; Ng *et al.*, 2017). That said, examples do exist of divergence in dewlap coloration at smaller scales or between populations with high degrees of gene flow (Thorpe & Stenson, 2002; Thorpe, 2002; Stapley *et al.*, 2011; Ng *et al.*, 2016).  
The species *Anolis sagrei* is widespread across islands of the West Indies (Reynolds *et al.*, 2020). It has been the subject of numerous studies concerning local adaptation (Losos *et al.*, 1994, 1997, 2001; Kolbe *et al.*, 2012), biological invasion (Kolbe *et al.*, 2008), and sexual selection (Tokarz, 2002, 2006; Tokarz *et al.*, 2005; Driessens *et al.*, 2014, 2015; Steffen & Guyer, 2014) among many other topics. Between-island variation in the mainly orange-red color of its dewlap was shown to be better explained by climatic variables (Driessens *et al.*, 2017) than by proxies for biotic factors such as sexual selection or predation pressure (Vanhooydonck *et al.*, 2009; Baeckens *et al.*, 2018). 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 differentiation despite gene flow.  
Here, we analyzed the color characteristics of *A. sagrei* dewlaps within nine islands in the Bahamas and Cayman Islands. These island systems presently, if not historically, comprise relatively small islands, with no major geographic barriers within islands limiting dispersal for this promiscuous species (Kamath & Losos, 2018). These islands all share three characteristic native West Indian small-island 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 (Howard, 1950; Schoener, 1968). The Cayman Islands and the Bahamas have been colonized independently by *A. sagrei* from Cuba (van de Schoot (2016) unpublished thesis; Reynolds *et al.* (2020)), such that these archipelagos constitute an ideal suite of natural replicates to explore within-island dewlap diversity across multiple islands.  
Our sampling design included sites in close proximity; the median distance between two sites within an island was km. Combining reflectance spectrometry and supervised machine learning, 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, compared 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 (Losos, 2011).

# Methods

## Data collection

We sampled 466 male *A. sagrei* from seven islands in the Bahamas Archipelago – Abaco, North Andros, South Andros, South Bimini, Eleuthera, Long Island, and Ragged Island – and two in the Cayman Islands – Cayman Brac and Little Cayman (Figure [3](#fig:maps)A). These islands were chosen to span the breadth of the West Indian range of *A. sagrei*, because they have highly similar habitat types, and because the *A. sagrei* on each island group are derived from ancient and distinct colonization events from Cuba (i.e. relatively evolutionarily independent, Reynolds *et al.* (2020)). Three habitats were sampled on each island based on characterizations by Howard (1950) and Schoener (1968). Each habitat is clearly distinguishable by its dominant vegetation type — xeric beach 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 [[tab:counts]](#tab: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 [3](#fig:maps)B), and % of all pairwise distances within islands were less than km. Additionally, there are no major barriers to dispersal (such as mountains or grassland) on any of the islands that we sampled.

## Reflectance measurements

We measured reflectance between 300nm and 700nm wavelength, a range from ultraviolet to red that encompasses the colors visible to most lizards and vertebrates in general (Lazareva *et al.*, 2012). 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 (Endler, 1990). The device was regularly standardized with a Spectralon white standard (Labsphere, North Sutton, NH, USA). Reflectance was measured at the center of the dewlap. Reflectance curves were smoothed using the R package pavo (Maia *et al.*, 2013) as well as with custom R functions, down to one reflectance value at each nanometer in wavelength from 300 to 700nm.

## Analysis

We tested for detectable differences in dewlap coloration between populations from different habitats across islands by following an analytic pipeline in several steps. First, we used multivariate analyses of variance to assess the relative contributions of islands, habitats and habitat-by-island interactions on the partitioning of variation in color space. Second, and provided that habitat-by-island interactions were found, we investigated habitat-differences in dewlap color for each island separately using machine learning classification. Third, for each island where multivariate differences were detected using our machine learning pipeline, we used univariate analyses of variance to decompose the signal among the different dimensions of color space. Fourth, for each significant between-habitat variation found in univariate analyses, we used post-hoc tests to determine which habitats were responsible for the differences. Last, to get insights into the spatial scale of phenotypic variation, for each significant contrast between two habitats detected along a given dimension on a given island, we performed multiple pairwise Wilcoxon tests to assess differences in dewlap coloration among the sites involved in that significant contrast, and recorded the geographical distance between sites that were found significant. In parallel, we tested a possible effect of isolation-by-distance, an alternative cause of phenotypic divergence between localities based on diffusion approximation and dispersal distance, irrespective of habitat-types. We did so using a permutation test to assess the significance of the correlation between geographical distances and phenotypic distances among sites within each island.  
All analyses in this study were performed in R 3.6.1 (R Core Team, 2019) and the source code can be found at https://github.com/rscherrer/dewlap, presently private.

### Dimensionality reduction

Because neighboring wavelengths are highly collinear and redundant in reflectance, we reduced the dimensionality of the data using principal component analysis (PCA), as per Cuthill *et al.* (1999) and Leal & Fleishman (2002). We performed PCA on data from all islands combined, as well as 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 [[tab:pcavariances]](#tab:pcavariances)). PCs need not represent the same wavelengths across islands because they are fitted on different datasets. Nevertheless, PC1 was highly collinear with brightness for all islands (Figure [12](#fig:brightness)), while the other PCs captured the chromatic variation (i.e. irrespective of brightness) in dewlap color.

### Among-island variance partitioning

We performed a two-way nonparametric multivariate analysis of variance (PERMANOVA, Anderson (2001), R package vegan, Oksanen *et al.* (2019)) to identify differences in coloration between islands, habitats and habitats within islands, using principal components fitted on data from all islands together. We used a nonparametric test because although no multivariate outliers were detected based on the Mahalanobis distance, the assumption of multivariate normality was violated in several habitats on several islands (Henze-Zirkler’s test, Henze & Zirkler (1990), R package MVN, Korkmaz *et al.* (2014), , Table [[tab:multinorm]](#tab:multinorm)).

### Within-island machine learning

We performed a machine learning classification analysis on the first four principal components within each island separately, using random forests (Breiman, 2001). Random forests are a versatile, intuitive, and powerful algorithm commonly used in machine learning, using decision trees to predict the labels of particular observations based on their multivariate coordinates. These coordinates, or variables, are passed through a series of successive decision nodes, each examining a given variable of any given observation (James *et al.*, 2013). The prediction for each observation is an aggregate over a large number of decision trees, each tree being trained on a subset of observations sampled with replacement from the dataset, and each tree being allowed to examine only a subset of the variables. This allows the random forest to overcome the individual errors of all trees in the predictions it makes.  
To detect differences in dewlap coloration between habitats, we measured the success of random forests in reassigning individual lizards to their correct habitat of origin, based solely on their principal component scores. In machine learning, this so-called cross-validation procedure is typically done in two steps (James *et al.*, 2013). First, a random forest is trained in recognizing features of dewlap coloration most associated with the different habitats, by being presented with multiple observations, making predictions about them, and updating its own decision rules based on whether the prediction deviates from the truth. Then, once trained, the patterns that the random forest has learned to recognize are tested by presenting new, previously unseen observations to the random forest, and measuring the proportion of correct predictions. This proportion, or success score, can then be statistically assessed against random guessing using a binomial test.  
The cross-validation procedure requires that the data be split into a training set and a testing set. To remove any bias due to the set that is being sampled for training, it is common practice to use k-fold cross-validation (James *et al.*, 2013), where the data are split into random bins and independent machines are trained, each taking one of the bins as a testing set and the rest for training, and where classification success is measured by summing all correct classifications from the machines.  
Here, we used a k-fold cross-validation procedure with , where each training set consisted of 80% of the data and the machine was tested on the remaining 20%. Each training set was conditioned on containing at least five lizards from each of the three habitats. We also down-sampled the training set to the sample size of the least represented habitat, to ensure that the different habitats were equally represented. To further remove any bias due to the specific random split into the different bins, we replicated each k-fold cross-validation five times. We then averaged the five resulting confusion matrices across replicates, where each confusion matrix shows the number of lizards from each habitat reassigned into each habitat. For each island, we then used the average proportion of correctly reassigned lizards (i.e. the proportion of observations on the diagonal of the average confusion matrix) as an estimate of classification success. This score was tested against random guessing by comparing it to a binomial distribution with number of trials being the number of lizards on that island and success probability , representing the rate of successful classification by chance when three habitats are involved.  
We used the machine learning fitting functions in the R package rminer (Cortez, 2020), which calls random forest routines from the randomForest package (Liaw & Wiener (2002), implementation from the original random forest algorithm (Breiman, 2001)). For each random forest, we optimized the number of trees in the forest and the number of variables examined by each tree using the grid hyperparameter search procedure implemented in rminer, to choose between two numbers of trees (500 or 1,000) and four numbers of principal components examined per tree (1 to 4), using rminer’s ordered holdout validation method with of the data used for training.  
We validated the results of our analysis by using two other widely used machine learning classification methods: linear discriminant analysis and support vector machines (Cristianini & Shawe-Taylor, 2000; James *et al.*, 2013), both accessible in rminer (Cortez, 2020).  
To know which wavelengths were most used to assign data points to each habitat, we trained another set of random forests, this time directly on reflectance data (taken every 5nm from 300 to 700nm) instead of principal components. We recorded the relative importance of each wavelength for each habitat, as measured by the mean decrease in accuracy during wavelength permutation, implemented in the randomForest package (Liaw & Wiener, 2002).

### Univariate analyses

For each island where significant differences in dewlap coloration were detected between habitats, we used multiple univariate analyses of variance (ANOVA) to identify possible principal components underlying the observed differences. We constructed our ANOVA models in two steps, as per Zuur (2009). 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, Pinheiro & Bates (2000; Pinheiro *et al.*, 2020)). 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, Bartoń (2019)), 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 evaluated the normality of the standardized residuals (residuals divided by their standard error, which can differ among habitats in a GLS model) of each fitted ANOVA model using Shapiro-Wilk’s test, with P-values adjusted for multiple testing using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995). In cases where significant deviations from normality were detected (, Table [[tab:normality]](#tab:normality)) we performed Kruskal-Wallis’s nonparametric test to back up the ANOVA results.  
To know which habitat-populations were different from which in dewlap coloration, we performed different post-hoc multiple comparison tests (all implemented in the PMCMRplus package, (Pohlert, 2020)), depending on which assumptions were met. In cases where normality and homoscedasticity were met (i.e. OLS-ANOVA was the best fit), we used Tukey’s honest significant difference test. When normality was met but not homoscedasticity (i.e. GLS-ANOVA was the best fit), we used Dunnett’s T3 test. Finally, whenever we used Kruskal-Wallis’s test because the ANOVA residuals were not normally distributed, we used Nemenyi’s test for post-hoc comparisons.

### 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 [[tab:sites]](#tab: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 (Gittleman & Kot, 1990). 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 in the R package geosphere; Hijmans (2019)) 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 tests (Hollander *et al.*, 2013) 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 variance. The P-values were adjusted using a Benjamini-Hochberg correction for multiple testing (Benjamini & Hochberg, 1995).

# 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 (beach scrub, primary coppice and mangroves). We found that most of the variation in coloration was partitioned between islands (two-way PERMANOVA, , , explained variance %). Nonetheless, we did find evidence for differences in dewlap coloration between habitat-types, and those were mostly island-specific (habitat-by-island interaction term, , , %), with a significant portion of the variation explained by an habitat effect across all islands, but this effect was relatively small (, , %).  
We subsequently tested for differences in dewlap coloration between habitat-populations within each island, using within-island principal component scores (to maximize the variation captured for each island, see Methods). Our within-island random forest classification analyses revealed detectable differences in dewlap coloration on eight out of the nine islands in our sample: Abaco, Bimini, Cayman Brac, Eleuthera, Little Cayman, Long Island, North Andros and South Andros. The accuracy of random forest classification exceeded random expectation more often than expected by chance for all these islands (Table [[tab:randomforests]](#tab:randomforests)). Accuracy was as high as 73% for Cayman Brac. We obtained similar results using other machine learning approaches such as support vector machines (Table [[tab:ksvms]](#tab:ksvms)) and linear discriminant analysis (Table [[tab:ldas]](#tab:ldas)), except that these methods did not detect significant differences on Eleuthera and North Andros. We did not find evidence of spatial autocorrelation in dewlap coloration between the sites within islands, except for Eleuthera (Table [[tab:autocorrelation]](#tab:autocorrelation)). We now describe the specific differences detected on each island.  
On Abaco, dewlaps from the mangrove habitat were the best discriminated, while dewlaps from the beach scrub habitat were often mistaken for dewlaps from the coppice habitat (Fig. [2](#fig:Abaco)D). Importance analysis revealed that beach scrub and mangrove lizards mostly differed in reflectance in the ultraviolet (UV) end of the spectrum (below 400nm, Fig. [[fig:Abaco\_supplement]](#fig:Abaco_supplement)F), where mangrove dewlaps had higher UV reflectance relative to beach scrub lizards, and coppice lizards had an intermediate UV reflectance between the two other habitats (Fig. [2](#fig:Abaco)B). Consistent with this, our analyses of variance detected significantly lower PC2 scores in mangrove lizards than in the two other habitats (Fig. [2](#fig:Abaco)E, Table [[tab:anova]](#tab:anova)), representing a higher UV-reflectance relative to red (Fig. [2](#fig:Abaco)C). Beach scrub lizards also scored lower on PC3 (Fig. [2](#fig:Abaco)E, Table [[tab:anova]](#tab:anova)), indicating less curvature of the reflectance profile and relatively higher reflectance at intermediate wavelengths (blue-to-yellow) than at the ends of the range (Fig. [2](#fig:Abaco)C). Differences were detected between sites both at large ( 100km) and short ( 1km) distances (Fig. [[fig:Abaco\_supplement]](#fig:Abaco_supplement)G).  
On Bimini, the random forests mostly correctly classified lizards from the coppice and mangrove habitats while often misclassifying lizards from the beach scrub habitat (Fig. [4](#fig:Bimini)D). Relatively flat importance profiles for beach scrub lizards suggested that brightness was used instead of a particular wavelength to identify some of the beach scrub dewlaps (Fig. [4](#fig:Bimini)F). Indeed, some beach scrub dewlaps were substantially brighter than the rest (Fig. [4](#fig:Bimini)B, C), a pattern that was captured by our analysis of variance along PC1 (i.e. brightness, Fig. [4](#fig:Bimini)C, E, Table [[tab:anova]](#tab:anova)). The random forests also used UV reflectance to discriminate between coppice and mangrove dewlaps (Fig. [4](#fig:Bimini)F), which could reflect the significant difference we detected along PC3 between these two habitats (Fig. [4](#fig:Bimini)C, E, Table [[tab:anova]](#tab:anova)). Beach scrub lizards were characterized by elevated red reflectance relative to UV (as represented by PC2, (Fig. [4](#fig:Bimini)C, E)), and beach scrub and mangrove lizards were characterized by a more even distribution of the reflectance along the spectrum (as represented by PC3, (Fig. [4](#fig:Bimini)C, E)), in contrast to coppice lizards which harbored a stronger curvature at intermediate wavelengths (Fig. [4](#fig:Bimini)C, E). On this island, the beach scrub and coppice habitats were separated by a few hundred meters, making this contrast the smallest geographical scale at which differences in coloration were found in our study (Fig. [4](#fig:Bimini)G).  
On Cayman Brac, all three habitats could be well discriminated against each other (Fig. [5](#fig:CaymanBrac)D), with UV reflectance appearing to be an important variable differentiating beach scrub and mangrove dewlaps (Fig. [5](#fig:CaymanBrac)F). In contrast, coppice dewlaps had a relatively flat importance profile, suggesting that brightness made them more distinct rather than any particular wavelength (Fig. [5](#fig:CaymanBrac)F). Consistent with this, coppice dewlaps were significantly different from all other dewlaps along PC1 (Fig. [5](#fig:CaymanBrac)E, Table [[tab:anova]](#tab:anova)). At a distance between 2 and 3km (Fig. [5](#fig:CaymanBrac)G), dewlaps in the beach scrub habitat reflected more red light (as represented by PC2, Fig. [5](#fig:CaymanBrac)C, E) and more UV (as represented by PC3, Fig. [5](#fig:CaymanBrac)C, E) than in the mangrove habitat. Coppice lizards were also characterized by a higher UV reflectance than mangrove dewlaps (PC3, Fig. [5](#fig:CaymanBrac)C, E, Table [[tab:anova]](#tab:anova)).  
Eleuthera was the only island where we detected significant spatial autocorrelation (Table [[tab:autocorrelation]](#tab:autocorrelation)), that is, sites that were closer geographically tended to have populations of lizards with more similar dewlap colors. Although random forests detected between-habitat differences in dewlap color, other approaches did not (Tables [[tab:ldas]](#tab:ldas) and [[tab:ksvms]](#tab:ksvms)), suggesting that the differences may be small. Consistent with this, the only significant univariate difference detected was for PC2 between beach scrub and mangrove lizards, where beach scrub lizards had higher levels of red reflectance and mangrove lizards higher levels of UV reflectance (Fig. [6](#fig:Eleuthera)C, E, Table [[tab:anova]](#tab:anova)).  
Little Cayman was characterized by a better discrimination of mangrove lizards from the rest than between beach scrub and coppice lizards (Fig. [7](#fig:LittleCayman)D). Mangrove dewlaps were most distinct with respect to their reflectance in short wavelengths (Fig. [7](#fig:LittleCayman)F), with significantly lower UV reflectance (as represented by PC2, Fig. [7](#fig:LittleCayman)C, E, Table [[tab:anova]](#tab:anova)). Beach scrub lizards were characterized by brighter dewlaps than coppice lizards (PC1), and also more convex curves, i.e., slightly higher UV and red reflectance (as represented by higher PC3 scores), than lizards from the other two habitats (Fig. [7](#fig:LittleCayman)C, E, Table [[tab:anova]](#tab:anova)).  
On Long Island the three habitats were relatively well discriminated (Fig. [8](#fig:LongIsland)D). Importance profiles indicated that short wavelengths were used to discriminate between beach scrub and mangrove lizards (Fig. [8](#fig:LongIsland)F). Beach scrub lizards had more curved reflectance profiles than in either of the two other habitats, with higher levels of UV and red reflectance relative to intermediate wavelengths (PC3, Fig. [8](#fig:LongIsland)C, E, Table [[tab:anova]](#tab:anova)). Beach scrub lizards also differed from mangrove lizards along PC4 (Fig. [8](#fig:LongIsland)E), which represented a rather small portion of the variance not already explained by the first three principal components, and is therefore difficult to interpret (Fig. [8](#fig:LongIsland)C). Coppice lizards were significantly darker than mangrove and beach scrub lizards (PC1, Fig. [8](#fig:LongIsland)C, E, Table [[tab:anova]](#tab:anova)).  
On North Andros, although the random forest classification was significant (, Table [[tab:randomforests]](#tab:randomforests)) and the average confusion matrix indicated that lizards from beach scrub were particularly well predicted (Fig. [9](#fig:NorthAndros)D), no significant univariate differences were detected along any of the four PCs (Fig. [9](#fig:NorthAndros)E, Table [[tab:anova]](#tab:anova)). Importance analysis of full-spectrum random forests showed higher importance scores near the UV-end of the spectrum in discriminating beach scrub dewlaps from the rest (Fig. [9](#fig:NorthAndros)F). Besides, reflectance curves of beach scrub dewlaps appeared more similar to each other in the UV range than dewlaps from other habitats (Fig. [9](#fig:NorthAndros)B), suggesting that the machines may have used this low within-habitat variance, as opposed to between-habitat differences in means, to correctly classify beach scrub lizards. A small sample size on this islands may also have contributed to a lack of power in detecting univariate differences using analyses of variance (Table [[tab:counts]](#tab:counts)).  
On South Andros beach scrub and coppice dewlaps could be discriminated better against each other than with mangrove dewlaps (Fig. [11](#fig:SouthAndros)D), with importance profiles supporting UV-reflectance as a predictor of coppice lizards (Fig. [11](#fig:SouthAndros)F). Coppice lizards had more curved reflectance profiles than beach scrub lizards (PC3), and lizards from both habitats differed along PC4, which is again more difficult to interpret (Fig. [11](#fig:SouthAndros)C, E, Table [[tab:anova]](#tab:anova)). Beach scrub lizards also differed from mangrove lizards in PC4 (Fig. [11](#fig:SouthAndros)C, E, Table [[tab:anova]](#tab:anova)).  
Classification success was not significantly better than expected by chance on Ragged Island (Table [[tab:randomforests]](#tab:randomforests)) where nearly no habitat could be differentiated from any other based on reflectance.

# Discussion

Two main insights follow from our results. First, excluding North Andros where the follow-up univariate analyses were not significant, we detected significant differences in dewlap coloration between habitats within seven out of the nine islands investigated, suggesting a putatively high potential for local differentiation of dewlap coloration in *Anolis sagrei*. Second, we found differences in coloration along different dimensions of color space, and in different directions, on different islands.  
Detectable differences in dewlap color between populations are surprising, as habitats were often in close geographical proximity to each other (as close as a few hundred meters on Bimini and most of the time within ten kilometers). Indeed, given that (1) the populations were continuously distributed between the habitats, (2) different habitat-populations were not monophyletic with respect to mitochondrial haplotypes (van de Schoot (2016) unpubl.), and (3) *A. sagrei* has been shown to be a highly mobile species (Kamath & Losos, 2018), we would have expected more homogeneous distributions of color phenotypes within islands due to extensive gene flow, with fewer differences between populations, especially those in close proximity.  
Several scenarios could account for these findings. One explanation is an adaptive one: populations living in different habitats could be phenotypically adapted to their local environmental conditions. Given that the brightly colored dewlap of *A. sagrei* is used as a communication signal, its color may be a target for selection if the transmission or perception of the signal differs from one habitat to another, for example because of differences in ambient light, according to the sensory-drive hypothesis (Endler & McLellan, 1988; Endler, 1992, 1998). The sensory-drive hypothesis has been tested multiple times for dewlap coloration in *Anolis* lizards, with mixed results. Some authors found support for it (Leal & Fleishman, 2002, 2004), while others did find differences in dewlap coloration between habitats, but those were inconsistent with the sensory-drive hypothesis (Fleishman *et al.*, 2009; Ng *et al.*, 2012).  
If our results were an example of sensory drive, we would have expected to see consistent differences between habitat-populations across islands (a pattern that would have been a compelling indicator of adaptation, Losos (2011)). This is because environmental conditions that may be relevant to color signal detectability such as light, temperature, moisture and vegetation, are consistent within the three main and clearly distinct habitat-types found across the sampled islands, i.e. beach scrub, primary coppice and mangroves (Howard, 1950; Schoener, 1968). Moreover, the patterns of divergence expected under a sensory drive scenario should be consistent with increased detectability given the local light conditions, such as the high contrasts with background vegetation found in the UV-range by Leal & Fleishman (2002) and Leal & Fleishman (2004).  
Instead, we found differences in the way dewlap color differs between habitats across islands. While short-wavelengths (UV reflectance) were often involved in color differences, they were not involved on all islands where significant differences were detected. On some islands, other or additional variables differed, such as brightness, red reflectance or the reflectance at the ends of the spectrum visible to *Anolis* lizards (UV and red, Lazareva *et al.* (2012)) relative to intermediate wavelengths (blue-to-yellow). Similar portions of the spectrum were sometimes involved in opposite directions on different islands, such as on Abaco and Cayman Brac, where mangrove lizards had a higher UV-reflectance than beach scrub lizards on the former, but a lower UV-reflectance on the latter. Under a sensory-drive scenario, we would have expected the same variables to be consistently divergent between habitats, or at least in a consistent direction.  
Consistent patterns across islands would have been a good clue for a sensory-drive explanation, but in particular consistent differences between habitats that are most different in their local conditions regarding the ecological function of the dewlap, such as ambient light. For example, if ambient light is an important factor shaping dewlap coloration, we would expect mangrove and beach scrub lizards, both inhabiting areas with high light penetration, to harbor more similar dewlaps, and to differ significantly from lizards from the coppice habitat, where irradiance is low. Overall, the observed heterogeneity of divergence patterns across islands provides no support to a sensory-drive explanation.  
Phenotypic plasticity could be another cause for dewlap color variation between habitats, where different environmental conditions produce different phenotypes in different habitats. Indeed, the yellow, orange and red colors in anoline dewlaps are produced by pterins and carotenoids (Ortiz, 1962; Ortiz *et al.*, 1962; Ortiz & Williams-Ashman, 1963; Ortiz & Maldonado, 1966; Macedonia *et al.*, 2000; Steffen & McGraw, 2007, 2009). Animals can synthesize pterins from nucleotides, but lack the ability to synthesize carotenoids (Goodwin, 1984; Hill *et al.*, 2002; Hill & McGraw, 2006). Different food qualities across sites within islands could therefore potentially cause detectable differences in coloration. Alternatively, more subtle effects on dewlap color could arise from developmental plasticity and depend, e.g., on differences in egg-rearing conditions. However, more data are needed to test these hypotheses, and although some work has shown plastic responses of dewlap color in response to parasites in *A. sagrei* (Cook *et al.*, 2013), we find it unlikely to account for the widespread habitat differences we found. Besides, studies testing the effect of carotenoid deprivation (Steffen *et al.*, 2010; Ng *et al.*, 2013) and heritability (Cox *et al.*, 2017) of dewlap coloration in *A. sagrei* and another species with a carotenoid-based dewlap, *A. distichus*, found little support for phenotypic and developmental plasticity in dewlap coloration.  
Genetic drift is another process that can account for differences in phenotype between localities, especially in small populations. One way this could proceed is through isolation-by-distance, where more distant populations accumulate more differences through time because of the reduced effect of gene flow at larger geographical scales relative to the dispersal range of the species (Rousset, 2004). Here, we only found a significant correlation between phenotypic and geographical distances on Eleuthera to support this scenario. On all the other islands, in contrast, populations from closer sites were not phenotypically more similar, which argues against isolation-by-distance. That said, there were often few sampling sites per island in our study, whose locations were not uniformly chosen within the islands, and so the true extent of isolation-by-distance may be difficult to test. Other, less trivial forms of drift may be at play than isolation-by-distance, but nevertheless, we did find significant differences in color phenotype at relatively small spatial scales, sometimes in neighboring habitats, on islands where gene flow is probably highly pervasive, as suggested by high rates of encounter between males and females (Kamath & Losos, 2018), making the divergence of habitat-populations by drift in relative genetic isolation an unlikely scenario.  
A number of alternative explanations remain. First, there could be sexual selection for different dewlap colors in different locations. Indeed, although the sensory-drive hypothesis may include a sexual selection aspect, e.g., if the optimal male phenotype in a given habitat is a function of female perception, sexual preferences may also be arbitrary and independent of the habitat (Andersson, 1994), and so could differ across habitats and islands. Although one previous study has found no link between dewlap coloration and body size dimorphism in *A. sagrei* (a proxy for the intensity of sexual selection) in an among-island comparison (Baeckens *et al.*, 2018), within-island data are lacking to test the plausibility of this scenario.  
Alternatively, selective pressures may be different in similar habitats from one island to another, because of other environmental variables not accounted for by the habitat-type classification we used (Howard, 1950; Schoener, 1968). The islands we sampled exhibit variation in some climatic variables, but also in densities of predators and anole congeners, which have all been shown to correlate with variation in *A. sagrei* mean dewlap coloration among islands (Vanhooydonck *et al.*, 2009; Baeckens *et al.*, 2018).  
Finally, different island-populations could also respond differently to similar selective pressures, resulting in various between-habitat divergence patterns across islands. Several factors could account for this. For example, the founder populations of each island, which colonized the islands independently (van de Schoot, unpublished thesis; Driessens *et al.* (2017; Reynolds *et al.*, 2020)), could have exhibited different dewlap colors at the time of colonization, as may be suggested by the larger differences we observed between than within islands. In turn, different initial phenotypes could have led to different ways in which populations would have diverged between habitats. Moreover, the different founding populations may have also consisted in different subsets of the standing genetic variation of their Cuban ancestor due to potential bottlenecks (Reynolds *et al.*, 2020), which may have constrained the way they would later respond to the local selective pressures of their respective islands.  
Altogether, our results show that dewlap color of *A. sagrei* commonly varies between habitat types, even in close geographical proximity, within islands of the West Indies, and that coloration differs in different ways from one island to another. We discussed several non-mutually exclusive mechanisms that could explain these observations, but more data are needed to thoroughly test these hypotheses. Nevertheless, heterogeneous patterns of divergence across islands do not support an adaptive sensory-drive scenario, and we propose that within-island dewlap color variation may be underlain by a more subtle mosaic of factors.

## Acknowledgements

Collection permission was granted by the Bahamas Environment, Science and Technology Commission, the Bahamas National Trust, the Bahamas Ministry of Agriculture, 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, Jason Fredette, Quyhn Quach, Wendy Jesse, Inbar Maayan, Alexis Harrison and Melissa Kemp for assistance in the field and Pratik Gupte, Max Lambert and James Stroud for helpful discussions. Funding for this work was provided by the Templeton Foundation (to JBL), NSF DEB #1927194 (to JBL and AJG), NSF DEB #1500761 (to AJG), NSF DBI #1609284 (to CMD), and a Harvard Museum of Comparative Zoology Putnam Expedition Grant (to RGR).

1. Corresponding author: r.scherrer@rug.nl [↑](#footnote-ref-2)