Mate Choice versus Mate Preference: Inferences about Color-Assortative Mating Differ between Field and Lab Assays of Poison Frog Behavior

Yusan Yang,1,* Simone Blomenkamp,2 Matthew B. Dugas,3 Corinne L. Richards-Zawacki,1,4 and Heike Pröhl5

1. Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15213; 2. Department of Biology, University of Bayreuth, Bayreuth, Germany; 3. School of Biological Sciences, Illinois State University, Normal, Illinois 61790; 4. Smithsonian Tropical Research Institute, Balboa, Ancón, Panama; 5. Institute of Zoology, University of Veterinary Medicine Hannover, Foundation, Hannover, Germany

Submitted September 18, 2018; Accepted December 4, 2018; Electronically published February 15, 2019


Abstract: Codivergence of mating traits and mate preferences can lead to behavioral isolation among lineages in early stages of speciation. However, mate preferences limit gene flow only when expressed as mate choice, and numerous factors might be more important than preferences in nature. In the extremely color polytypic strawberry poison frog (Oophaga pumilio), female mate preferences have codiverged with color in most allopatric populations tested. Whether these lab-assayed preferences predict mating (gene flow) in the wild remains unclear. We observed courting pairs in a natural contact zone between red and blue lineages until oviposition or courtship termination. We found color-assortative mating in a disturbed habitat with high population density but not in a secondary forest with lower density. Our results suggest color-assortative O. pumilio mate choice in the wild but also mating patterns that do not match those predicted by lab-assayed preferences.

Keywords: Oophaga pumilio, color polymorphism, mate choice, behavioral isolation, sexual selection, assortative mating.

Introduction

When male traits and female preferences codiverge among lineages (sympatric or allopatric), any resulting assortative mating may contribute to behavioral isolation, a common early stage in speciation (Panhuis et al. 2001; Ritchie 2007). A typical first step in testing the hypothesis that such divergent preferences play an important role in speciation is to test for assortative mate preferences using laboratory assays, with preferences interpreted as a proxy for what might happen in the wild (Jiang et al. 2013; Scordato et al. 2014). However, unless preferences are absolute (i.e., females always reject the nonpreferred phenotypes), myriad factors in nature can restrict a female’s options (e.g., male-male competition) or alter the relative benefits of expressing a preference as choice (e.g., costs of sampling males; reviewed in Jennions and Petrie 1997; Rosenthal 2017). Because mating, not mate preference, is the ultimate determinant of gene flow among diverged lineages, understanding whether and how mate preferences translate to mate choice in the wild is imperative to understanding the role of mate preferences in the speciation process (Jennions and Petrie 1997; Rosenthal 2017).

The strawberry poison frog (Oophaga pumilio) is one notable example of dramatic phenotypic divergence that may be implicated in the evolution of reproductive isolation (reviewed in Rojas 2016; Dugas 2018). Across most of its range, this small diurnal frog has a red body with blue/black limbs. In and around the Bocas del Toro archipelago of Panama, in contrast, O. pumilio populations are remarkably diverse in both coloration and patterning, with primary body color spanning the visual spectrum (Daly and Myers 1967; Summers et al. 2003). In most populations tested with laboratory preference assays, females prefer to associate and/or court with males of a local color over males from differently colored populations, and color seems to be the trait on which females base this preference (Summers et al. 1999; Reynolds and Fitzpatrick 2007; Maan and Cummings 2008; Richards-Zawacki and Cummings 2011; Yang et al. 2016). There is no evidence of intrinsic reproductive incompatibility between
morphs: experimental crosses among morphs produce viable offspring (Summers et al. 2004) with fertility/fecundity equivalent to that of offspring of intramorph pairs (Dugas and Richards-Zawacki 2015). Divergent preferences have thus been hypothesized as the major potential reproductive barrier among color morphs (Tazzyman and Iwasa 2010).

In one mainland monomorphic population, wild *O. pumilio* court assortatively on the basis of minor variation in red coloration (Gade et al. 2016). Because this suggests that wild individuals attend to color, it provides perhaps the most compelling evidence that color divergence in Bocas del Toro could lead to reproductive isolation among differently colored lineages. However, molecular pedigrees in this species (Richards-Zawacki and Cummings 2011) and an ecologically similar frog (Ursprung et al. 2011), as well as behavioral observations (Meuche et al. 2013), suggest that females are not particularly choosy in the wild. We studied mate choice in a natural contact zone between a red and a blue morph of *O. pumilio* (fig. 1A). Most currently described color morphs of *O. pumilio* occur in allopatry (typically, one morph per island; Summers et al. 2003) with contact unlikely, but this region of the Panamanian mainland (fig. 1A) offers a rare opportunity to compare lab-assayed preferences (Yang et al. 2016) with wild mating patterns. In the laboratory, females associate preferentially with the local color on either side of the contact zone, but all types of females from the polymorphic region prefer to court with red males (Yang et al. 2016). We observed courtship at two sites in the contact zone to test the assortative mating predictions that (1) courting pairs of the same morph would be more likely to mate than pairs consisting of two different color morphs and (2) with color treated in a continuous manner, a pair’s likelihood of mating would increase with color similarity between the male and the female.

**Methods**

**Study Species**

*Oophaga pumilio* is a diurnal, terrestrial frog native to lowland forests on the Caribbean side of Central America (Savage 1968). Within morphs, the sexes do not differ qualitatively in coloration or patterning (Summers et al. 2003; Maan and Cummings 2009). Nearly all currently described color morphs are allopatric (i.e., occur on only one island; Summers et al. 2003), but even isolated, phenotypically distinct populations in this young (1–9 kya) archipelago are relatively undifferentiated at neutral microsatellite loci (Hauswaldt et al. 2011).

Both sexes of *O. pumilio* are polygamous in the wild, but females appear to be the choosier sex and are capable of terminating courtship at any stage (Pröhl and Hödl 1999; Pröhl 2003; Meuche et al. 2013). Males compete for the territories from which they call to attract females, but these territories seem unlikely to offer direct benefits to females (Pröhl and Berke 2001; Dugas 2018). If courtship is successful, a female lays a clutch of ∼5 eggs in the leaf litter. Males tend and hydrate this clutch, but hatching failure is high (>80%: Pröhl and Hödl 1999). When eggs hatch, the female transports tadpoles to water-filled leaf axils and provisions each with unfertilized eggs throughout its development (∼45 days, reviewed by Dugas 2018).

**Study Populations**

We observed courtship and mating at two sites in what appears to be a contact zone between two phenotypically distinct color morphs (red and blue) of *O. pumilio* in the Bocas del Toro region (fig. 1A; Yang et al. 2016, 2018). The two sites we studied appear to differ in habitat structure, frog density, and relative morph frequencies. At site 1 (9°13’15.70"N, 82°13’5.60"W), secondary growth is accompanied by abandoned banana and cacao plots, while at site 2 (9°12’42.48"N, 82°12’53.17"W), these agricultural plantings are absent. We estimated demographic parameters from a 400-m² plot at each site, in which we captured, marked, and recaptured frogs weekly in May–July 2015 (Y. Yang, unpublished data). As is typical in other geographic regions (Donnelly 1989; Pröhl 2002), density is higher in the disturbed habitat, site 1 (∼66.6 frogs/100 m²), than at the relatively undisturbed site 2 (∼28.6 frogs/100 m²; sequential Schnabel estimates; Schnabel 1938). The sex ratio is near 1:1 at both sites, but the relative frequency of morphs appears to be different (site 1: 33% red, 18% blue, 49% intermediate; site 2: 16% red, 52% blue, 32% intermediate).

Male call properties and body size are common predictors of female choice in anurans (Arak 1983) but seem unlikely to mediate morph-assortative mating in the context we studied here. While call properties are associated with mating success (Pröhl and Hödl 1999; Pröhl 2003; Dreher and Pröhl 2014), there are no differences in the advertisement calls among *O. pumilio* color morphs in Panama (Pröhl et al. 2007; M. L. Dye, unpublished data). Body size varies among several *O. pumilio* color morphs (Rudh 2013), but not among red, blue, and intermediate individuals at our focal sites (Y. Yang, unpublished data). Furthermore, male body size is not associated with mating success in *O. pumilio* (Pröhl and Hödl 1999; Gade et al. 2016).

**Behavioral Observations**

When a female *O. pumilio* enters a calling male’s territory, the male attempts to court her by directing softer and slower “courtship calls” (Pröhl 2003) in her direction, approaching and sometimes touching her. The female, if in-
interested, courts back by approaching, staying in proximity to, and sometimes touching the male. As courtship progresses, the male will lead the interested female to several potential oviposition sites within his territory, at one of which the pair may eventually mate (video 1, available online). Successful courtship can take several hours (Pröhl and Hödl 1999), and one partner (most often the female) might terminate the process at any point in the encounter (Limerick 1980; Haase and Pröhl 2002).

Figure 1: A. Map showing the contact zone between red and blue morphs of *Oophaga pumilio* in the Bocas del Toro Archipelago, Panama. Pie charts show the relative red, blue, and intermediate morph frequencies at each location. An ancestral-like morph (red with blue limbs) is found on the northern part of the Aguacate Peninsula (on the mainland) and on the adjacent island of San Cristóbal, while an entirely blue morph occurs farther south on the Aguacate Peninsula. Both morphs, along with phenotypic intermediates, occur in this contact zone. The two study sites (site 1 and site 2) are denoted by larger pie charts. B. Examples of color variation in the contact zone. C. By-eye color categories (color morphs) plotted in a quantitative color space based on photo-derived color scores PC1 and PC2.
We made all observations in May–August 2015, between 0700 and 1300 hours, when *O. pumilio* is most active (Graves 1999; Graves et al. 2005). We located males by listening for advertisement or courtship calls and continued observation if a female was seen courting with the calling male. We observed the courting pairs from more than 1.5 m away (a distance that does not disrupt behavior; Staudt et al. 2010; Gade et al. 2016) until (1) courtship was successful (oviposition occurred) or (2) courtship was unsuccessful, operationally defined as male-female distance of at least 2 m without subsequent approaches by either frog within 10 min. While courting individuals might pause to feed for more than 10 min (H. Pröhl, personal observation), we chose this cutoff to optimize the trade-off between observing such pauses and being able to track and capture both individuals at the end of each observation. We observed 12–160 min of courtship, but this is likely an underestimate of courtship duration, as our observations often started midcourtship.

After each observation, we captured the focal male and female and measured their snout-vent length (SVL) to the nearest 0.01 mm, using dial calipers (except for three females and two males that we were not able to catch). For color quantification, we immediately took photographs with a digital camera (Panasonic DMC-TS5, Kadoma, Osaka, Japan), capturing an image of the dorsum against an 18% gray standard (DGK Color Tools). We also measured the height of the perch where the male was first seen calling to the nearest centimeter, using a flexible meter tape. We marked each individual with a unique toe clip (Funk et al. 2005) before releasing the pair at point of capture. Of the 76 pairs we observed, four males and six females were observed in courtship twice, but none of them was observed engaging with the same potential mate twice.

**Color Quantification**

In this likely contact zone between red and blue *O. pumilio* morphs, there are phenotypically intermediate individuals that likely result from matings between color morphs (fig. 1B). Treating individuals as belonging to discrete color morphs made our results comparable to past laboratory studies of mate preference in this frog (Summers et al. 1999; Reynolds and Fitzpatrick 2007; Maan and Cummings 2008). Quantifying color in a continuous manner, on the other hand, allowed us to test whether or not color similarity (both within and among the morphs) influences mate choice in the contact zone (e.g., whether redder females are more likely to accept redder males). Therefore, we used both by-eye color
categorization and color scores generated from digital photographs for subsequent analyses.

To quantify color in a continuous way, we sampled red (R), green (G), and blue (B) values across five 20 × 20-pixel areas on the frog’s dorsum, using the software ImageJ 1.48v (Schneider et al. 2012). We then standardized these values by taking residuals of mean R, G, or B frog color regressed on R, G, or B scores from the 18% gray standard (a 20 × 20-pixel area) in the same photo. We calculated the mean R, G, and B residuals for each frog and used a principal component analysis (PCA) to reduce the number of color parameters (Stevens et al. 2007; methods followed Dugas et al. 2015).

### Statistical Analyses

For all analyses, data sets from the two sites were treated separately because they differed in biotic and abiotic factors (including habitat structure, color morph frequency, and population density) that are known to influence mate choice (Jennions and Petrie 1997; Rosenthal 2017). We included male and female male body size (SVL) and male perch height as covariates in all the models described below to control for their potential effects on mating success (Pröhl and Hödl 1999).

We first used by-eye color categorizations (red or blue) of courting pair members to test the prediction that pairs of same-morph individuals would be more likely to mate than pairs with individuals of two different color morphs. In this analysis, we excluded pairs containing phenotypically intermediate individuals because nearly all the previous work to which we wanted to compare our results considered discrete, usually allopatric, color morphs. *Oophaga pumilio* have a visual system capable of distinguishing among most color morphs (Siddiqi et al. 2004; Maan and Cummings 2009), but categorical color perception and the corresponding decision threshold (Caves et al. 2018) have not been studied in *O. pumilio*. Using only the individuals at extremes of the spectrum ensured that these are most likely to be perceived as separate color categories. We used a generalized linear model (GLM) with binomial error structure to test the effect of pair type (homomorphic or heteromorphic) on courtship success (yes/no oviposition). An alternative would have been to include effects of male color, female color, and their interaction, but we were not able to run this analysis because of the small and unbalanced sample size (see table 2).

We then tested for assortative mating, using all observations, including pairs that contained intermediate individuals, and continuous color variables (quantitative color scores PC1 and PC2 generated from photographs; see "Color Quantification"). To test the hypothesis that the likelihood of mating (oviposition) increases with color similarity, we calculated the Euclidian distance between the colors of the male and the female in a pair in the PC1-PC2 color space (Endler 1990; fig. 1C) and tested its effect on courtship success (yes/no oviposition), using GLMs with binomial error structure.

We treated each encounter as an independent sample, even though 10 individuals were observed in courtship twice (though always with a unique mate). We were not able to include male and female ID as random effects because our generalized linear mixed model failed to converge as a result of small and unbalanced sample sizes. However, when we reran all analyses without these 10 observations, results were qualitatively similar. These additional analyses and results are included in the archived R files.

We performed all analyses in R 3.2.3 (R Core Development Team 2015). We used the glm function in the stats package (R Core Development Team 2015) to fit the GLMs. We tested the significance of main effects, interaction terms, and covariates, using a likelihood ratio test with the Anova function in the car package, which compares overall model fit with and without a particular effect. Data underlying the “Results” section and figures 1 and 2 are deposited in the Dryad Digital Repository (https://dx.doi.org/10.5061/dryad.fs24q30; Yang et al. 2019).

### Results

#### Color Quantification

As was the case in a previous analysis of this and another *Oophaga pumilio* population (Dugas et al. 2015), the first PC axis (PC1) explained ~60% of the variation, with G and B both loading strongly and negatively; PC2 explained ~36% of the variation, with R loading strongly and positively (table 1). Frogs categorized by eye as blue, red, or

### Table 1: Principal component (PC) analysis for photograph RGB values

<table>
<thead>
<tr>
<th>PC</th>
<th>Eigenvalue</th>
<th>Variance explained (%)</th>
<th>Red</th>
<th>Green</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>1.82</td>
<td>60.7</td>
<td>.029</td>
<td>−.703</td>
<td>−.710</td>
</tr>
<tr>
<td>PC2</td>
<td>1.09</td>
<td>36.2</td>
<td>.954</td>
<td>.232</td>
<td>−.189</td>
</tr>
</tbody>
</table>

Note: Results of a principal components analysis of mean red (B), green (G), and blue (B) color scores from the dorsum of *Oophaga pumilio*. The RGB values were measured in and averaged between five 20 × 20-pixel areas on the dorsum with ImageJ and standardized by taking the residuals from a regression on the gray standard in each photo.
intermediate were significantly different in both PC1 and PC2 and were distinguishable in the quantitative color space based on the two PCs (fig. 1C). Red, intermediate, and blue morphs differ significantly in both PC1 (ANOVA: $F_{2,119} = 22.1, P < .001$) and PC2 (ANOVA: $F_{2,119} = 80.8, P < .001$). The PC1 values were the least negative in red individuals and more negative in intermediate and blue individuals (Tukey post hoc test: red-intermediate: $P < .001$, red-blue: $P < .001$, intermediate-blue: $P = .881$). The PC2 values were most positive in red individuals, lower in intermediate individuals, and lowest in blue individuals (Tukey post hoc test: red-intermediate: $P < .001$, red-blue: $P < .001$, intermediate-blue: $P < .001$).

**Mate Choice Pattern in the Wild**

We observed 32 courting pairs at site 1, 15 of which mated, and 44 courting pairs at site 2, 31 of which mated (details in table 2). The four courting pairs (two at site 1 and two at site 2) with missing body size and photo data (those we were not able to capture) were excluded from all analyses except calculations of overall mating success at the two sites. The overall proportion of interactions that resulted in mating was significantly larger at site 2 than at site 1 ($\chi^2 = 4.31$, df = 1, $P = .038$). All unsuccessful courtships we observed were terminated when the female moved out of proximity to the male.

In the first analysis, using by-eye colors, we found that homomorphic courting pairs were marginally more likely to mate than heteromorphic courting pairs (likelihood ratio test: $LR \chi^2 = 3.51$, df = 1, $P = .061$; fig. 2C) at site 1, but not at site 2 ($LR \chi^2 = 0.002$, df = 1, $P = .966$; fig. 2D). Using photo-based color scores, we found that the probability that a courting pair would mate increased with shorter PC color distance (i.e., more similar male and female colors) at site 1 ($LR \chi^2 = 5.05$, df = 1, $P = .025$; fig. 2A), but not at site 2 ($LR \chi^2 = 1.32$, df = 1, $P = .251$; fig. 2B). The effects of male/female body size and perch height were non-significant in all models (all $LR \chi^2 < 3.11$, all $P > .078$).

**Discussion**

Divergent lab-assayed mate preferences among phenotypic variants or hybridizing species are often interpreted as support for the hypothesis that sexual selection can drive speciation (Jiang et al. 2013; Scordato et al. 2014). Evidence that phenotypically distinct allopatric lineages of *Oophaga pumilio* show such divergent association preferences has been interpreted in a similar way (Summers et al. 1999; Reynolds and Fitzpatrick 2007; Maan and Cummings 2008). In a natural contact zone between red and blue *O. pumilio* morphs, however, support for the hypothesis that assortative preferences have evolved and could contribute to reproductive isolation is mixed at best. In the laboratory, females show assortative preferences for the local color on either side of the contact zone, but red, blue, and phenotypically intermediate females from the polymorphic region all associate preferentially with red males (Yang et al. 2016). In nature, we found evidence of assortative mating, but only at one of two sites we studied. Similar patterns emerged in another polymorphic population of *O. pumilio*. Females of yellow and red morphs both show assortative courtship preferences in the laboratory (Richards-Zawacki and Cummings 2011), but genetic signatures in the wild suggest assortative mating of the more common color morph only (Richards-Zawacki et al. 2012).

Laboratory assays of preference and observations of mating in the wild suggest, at least in two instances of polymorphism in *O. pumilio*, that while color preferences exist, gene flow between morphs may not depend solely, or even primarily, on these preferences. Because selecting and/or rejecting a potential mate often has fitness costs (e.g., time, energy, necessary cognitive ability), the expression of preferences (i.e., choosiness) should be sensitive to social and ecological context (Jennions and Petrie 1997; Rosenthal 2017). In taxonomically diverse animals, the choosiness of the choosy sex tends to increase with population density, perhaps because the cost of searching/sampling for a mate goes down (reviewed in Kokko and Rankin 2006). Consistent with this prediction, *O. pumilio* females at a high-density site (site 1) rejected courting males more frequently than did females at a lower-

<table>
<thead>
<tr>
<th>Table 2: Sample size of each color combination</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Site 1</strong></td>
</tr>
<tr>
<td><strong>Male color</strong></td>
</tr>
<tr>
<td>Blue</td>
</tr>
<tr>
<td>Intermediate</td>
</tr>
<tr>
<td>Red</td>
</tr>
</tbody>
</table>

Note: Number of successful matings (numerator) and total number of observed pairs (denominator) of each male × female color combination in the two populations. Site 1: disturbed habitat, high red frequency, high density; site 2: secondary forest, high blue frequency, low density.
density site (site 2), and individuals mated assortatively by color at the high-density site 1 but not at the low-density site 2. Similarly, differences in color morph frequency can lead to search cost asymmetries between females of more and less frequent phenotypes (as suggested in another polymorphic *O. pumilio* population; Richards-Zawacki et al. 2012), but we were not able to test for asymmetry in assortative mating in this study because of the small and unbalanced sample size. In anurans, mate sampling can take a variety of forms, including acceptance of any territorial/actively calling male (Friedl and Klump 2005; Ursprung et al. 2011), sequential assessment (Ryan 1985), and simultaneous assessment of multiple males (Murphy and Gerhardt 2002; Murphy 2012). While field observation in a low-density population (~6.8 frogs/100 m², compared to ~66.6 and ~28.6 at our two study sites) suggests that female *O. pumilio* do not sample multiple males before mating (Meuche et al. 2013), both theory (Real 1990; Castellano and Cermelli 2011) and empirical evidence from other systems (Backwell and Passmore 1996; Kvarnemo and Forsgren 2000) suggest that mate sampling tactics can vary with search cost and the distribution of male quality. Describing any plasticity in mate sampling tactics of female *O. pumilio* will be an exciting addition to our understanding of how preferences and choice shape reproductive isolation. Finally, differences in the light environment and the structural complexity of the signaling habitat can also influence a female’s ability to distinguish among slight color differences along the red-blue continuum and potentially affect the expression of color preferences (Endler and Théry 1996; Boughman 2001; Maan and Cummings 2009). However, with only two sites, we can-

![Figure 2](image-url)

**Figure 2:** A, B, Likelihood of a courting pair successfully mating in relation to the Euclidean color distance between the male and female at sites 1 (A) and 2 (B), estimated from generalized linear mixed models. The line indicates the estimates and shading the 95% confidence interval of the likelihood function predicted from the generalized linear models. C, D, Proportion of homomorphic and heteromorphic courting pairs that successfully mated at sites 1 (C) and 2 (D).
not distinguish among these various hypotheses and their potential interactions. Testing for a causative link between these social and ecological factors and the degree of color-assortative mating will require further study.

Mate preferences are expected to exert sexual selection only when they are expressed as mate choice (Rosenthal 2017). Phenotypically distinct color morphs of \textit{O. pumilio} express preferences in laboratory assays that do not match mate choice patterns in the wild, and patterns in the wild are inconsistent across sampling localities. Our observation of color-assortative mating at one of the two field sites suggests the potential for limited gene flow among distinct phenotypes, but the pattern and strength of assortative mating may depend on biotic and abiotic factors that are usually removed from standardized laboratory assays. Alongside examples of similar discrepancies in other taxa (Gerhardt 1992; Veen et al. 2001; Brooks 2002; Schunér et al. 2017), these patterns highlight gaps in our knowledge about the extent to which assortative preferences in the laboratory can provide support for the hypothesis that sexual selection can drive speciation.

Acknowledgments

The Smithsonian Tropical Research Institute (STRI) provided logistical support, and we particularly thank Plinio Gondola (Bocas del Toro Research Station). Ling-Wan Chen provided suggestions on statistical analysis. The Panamanian National Authority for the Environment provided research permission (permit SE/A-40-15). This study was supported by a PROMOS scholarship (University of Bayreuth) to S.B., an STRI short-term fellowship to Y.Y., and a grant from the National Science Foundation (1146370) to C.L.R.-Z. This work was done in compliance with Institutional Animal Care and Use Committee protocols at Tulane University (0382 and 0382R) and STRI (2014-0515-2017).

Literature Cited


Natural History Editor: Edmund D. Brodie III

Top left, Oophaga pumilio blue morph. Top right, O. pumilio red morph. Bottom, red O. pumilio male courting a blue female. Photo credit: Yusan Yang.