How does variation in a mating trait evolve? Insights from studies of color signals and their perception in a highly polymorphic poison frog

by

Kimberly Anne Smith Howell

Bachelor of Arts, Luther College, 2015

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kimberly Anne Smith Howell

It was defended on

December 1, 2022

and approved by

Dr. Jessica Stephenson, Assistant Professor,
Department of Biological Sciences, University of Pittsburgh

Dr. Jon P. Boyle, Professor,
Department of Biological Sciences, University of Pittsburgh

Dr. Miler T. Lee, Assistant Professor,
Department of Biological Sciences, University of Pittsburgh

Dr. Lauren O'Connell, Assistant Professor,
Department of Biology, Stanford University

Dissertation Advisor: Dr. Corinne L. Richards-Zawacki, Professor,
Department of Biological Sciences, University of Pittsburgh

How does variation in a mating trait evolve? Insights from studies of color signals and their perception in a highly polymorphic poison frog

Kimberly Anne Smith Howell, Ph.D.

University of Pittsburgh, 2022

Phenotypic polymorphism is commonly thought of as a step towards speciation in animals. Sexual selection can contribute to the development and maintenance of phenotypic polymorphism if mate preferences diverge along with the trait and reduces mating between the different variants. Color is a common sexually selected trait and divergence in color-based sexually selected traits has been found. The strawberry poison frog, *Oophaga pumilio* is highly variable in color across the Bocas del Toro Archipelago in Panama and color appears to be under natural and sexual selection, though the forces that drove divergence are unclear. My dissertation focused on understanding the divergence of color and color perception in this species. First, I tested for variation in expression and sequence differences in the light-absorbing protein opsin, to test whether differences might contribute to known color-based behavioral biases. Overall, I did not find convincing evidence that differences in color vision have evolved among distinctly colored populations, contrary to findings in aquatic systems. Next, I used a controlled breeding study to test for differences in the inheritance patterns and heritability of dorsal coloration between three color morphs to better understand how this trait has evolved and how it may further evolve. I found evidence for simple Mendelian dominance in one cross and additive genetic variance in two other crosses, suggesting the genetic architecture of color differs between morphs. I also found evidence heritability differs between morphs suggesting different populations might have different responses to selection. Finally, I compared patterns of trait variation among populations across an

island with patterns of variation in neutral loci, testing for isolation by adaptation (IBA). I found a pattern of IBA for body size, suggesting this trait may under selection, but not for dorsal coloration or mating call. Overall, my research contributes new insights into the mechanisms by which coloration has diverged in this species, which is fast becoming a model species for the study of phenotypic divergence and speciation. More broadly, my work contributes to our understanding of animal evolution by providing an additional case study of how intraspecific variation in sexually selected traits can evolve.

Table of Contents

List of Tables

List of Figures

Preface

First and foremost, I need to thank my wonderful advisor, Dr. Cori Richards-Zawacki, for giving me a lab home in my first year and letting me run with a project that was totally new to the lab and a bit outside her area of expertise. Thanks for the support and encouragement through all of the ups and downs and in-betweens. Thanks to all of my labmates over the years for helping to create a supportive and fun environment to work in including, Dr. Layla Freeborn, Dr. Yu-san Yang, Dr. Sebastian Echeverri, Dr. Verónica Saenz, Dr. Tiffany Betras, Dr. Michel Ohmer, Dr. Karie Altman, Dr. Brandon Hoenig, Marco Gonzalez. Special thanks to Layla for doing the field work and processing for my eye samples, increasing her own long field days. I also want to thank my wonderful committee members, Dr. Jessica Stephenson, Dr. Jon Boyle, Dr. Miler Lee, and Dr. Lauren O'Connell for their help and encouragement over the years. Thanks especially to Jon for all the hours spent troubleshooting my opsin expression experiment, and to Jess for joining in my committee in my last semester of graduate school and for providing really good feedback on an earlier draft. Thanks also to the many undergrads who helped make this research possible including Jill Feinstein, Mikaela Baratka, Tabea Schneider, Avery Miklos, Sarah, Lofland, Ankit Chadha, and Risha Musuku. Mikaela Baratka deserves special appreciation for staring at frog videos for two years for a project that proved extremely frustrating and ultimately did not succeed.

Thanks also to my support network outside of science. To my dear friend Zoe, thanks for all of the support and laughter from our (semi) monthly hours long phone calls. They helped me to take a step back from the graduate school life and helped ground me. Thanks also to Katie, I'm really glad we both ended up on the Ph.D. journey together, albeit in different disciplines and universities. It was so nice getting to chat with someone outside of my program who knew what

graduate school was like and I really looked forward to our chats. Thanks to my family for loving and encouraging me, even if you didn't understand what I was doing all of the time. Thanks especially to my parents and my brother and future sister-in-law for all of their support throughout my graduate school experience.

1.0 Introduction

The evolution of trait polymorphism is commonly posited as a step along the way to speciation (rev. in McKinnon *et al.* 2010). Polymorphisms in a wide variety of traits used in mate selection have been noted in many animal species. For example, polymorphisms in coloration occur in birds (Kokko *et al.* 2014), fish (Sandkam *et al.* 2015, Maan *et al.* 2008), and frogs (Myers and Daly 1976, Twomey et al. 2014) and in body size in fish (Oliveira et al. 2001), and damselflies (Tsubaki 2003). The evolution of color polymorphism is especially intriguing because color serves multiple functions in animals, including camouflage (Zimova *et al.* 2018), aposematic warning coloration (butterflies, Finkbeiner *et al.* 2014, coral snakes Banci *et al.* 2020), mimicry (Pfennig and Mullen 2010, Kunte 2009), and as a signal used in mate choice (Brooks and Endler 2001). Because animal coloration can be used in many contexts, understanding the mechanisms that generate color differences within species can help us to understand the roles that alternative evolutionary forces, like natural selection, sexual selection, and genetic drift, play in shaping these important signals. My work focuses on understanding how sexual selection generate divergence in animal color signals. It is well known that sexual selection can aid in the divergence of sexual trait if the individuals bearing divergent sexual traits mate assortatively (rev. in Bolnick and Fitzpatrick 2007). However, it is less obvious how these trait polymorphisms arise to begin with. My work focuses on understanding the drivers of mating trait divergence in a polymorphic poison frog, *Oophaga pumilio.*

1.1 Color Perception

Sexual selection as an agent of evolutionary change relies on members of the showier sex, typically males, successfully signaling their fitness to members of the choosier sex, typically females. The ability of the chooser to perceive differences in the signals of potential mates will influence which mate is chosen, as better perceptual ability gives the chooser more information to use (rev. in Kokko *et al.* 2002 and in Ryan and Cummings 2013). Males with a non-optimal signal with regard to female perception may have their signal lost against background noise, resulting in fewer or no mates for the male (Endler 1992). Therefore, in most animals, where females are the choosier sex, the evolution of variation in male sexual signals is thought to be constrained by the perceptual abilities of the female sensory system. Animals generally have better perceptual abilities in certain portion(s) of their total sensory range. For example, a female may be able to see colors across the range from 300 – 700nm in wavelength but can detect and/or discriminate colors from 500-600nm most easily. This means that males are more successful in attracting the female's attention and interest if the value of his sexual signal falls within the range that the female can perceive best, because that maximizes the efficiency with which the sexual signal is transmitted (Endler 1992). If signal perception varies among females, this has been proposed to result in males "specializing" on different ranges of the sexual signal, producing variation in the signal driven by variation in the female sensory system (Endler and Basolo 1998). The second chapter of my dissertation focuses on understanding the role that variation in the *Oophaga pumilio* visual system may have played in divergence in color-based mating signal divergence.

1.2 Genetics of Coloration

My study species, *O. pumilio*, exhibits a large range of color phenotypes which are thought to have arisen rapidly. So, what permitted these color phenotypes to diverge so quickly? Gaining a complete understanding of the evolution of a quantitative trait, as color often is, requires knowledge of the trait's genetic architecture. This includes knowing how many genes are involved in producing the trait, their location in the genome, the dominance patterns among alleles, and any epistatic or pleiotropic interactions (rev. in Mackay 2001). Dominance effects can shield recessive alleles from selection, particularly at low frequencies, and selection on genes with epistatic effects can strengthen selection on additive genetic variation, often resulting in reduced genetic variation in a trait (rev. in Hansen 2006). If multiple genes affecting a trait are tightly linked, they will tend to be inherited together, and thus selection will act on the two loci at once (rev. in Nosil *et al.* 2009). Thus, some knowledge of the genetic architecture of a trait is critical to understanding how the trait has and might continue to evolve. Thus, in my third chapter I use a controlled breeding experiment to test hypotheses about the genetic architecture of color in three distinctly colored *O. pumilio* populations.

1.3 Genotypic and Phenotypic Structure

The patterns that arise within and among populations in genes that are evolving neutrally (i.e., by mutation and genetic drift) can be used as a backdrop on which to make inferences about the evolutionary processes that are shaping phenotypic variation. Comparing the patterns of neutral genetic variation with that of diversity in phenotypic traits putatively under selection is one

approach that can allow us to make inferences about the evolutionary forces generating trait variation and polymorphisms – the raw material on which selection acts. However, few studies have tested for concordance between patterns of neutral genetic variation and variation in sexually selected traits in animals. In my fourth chapter, I take this approach, comparing genetic variation in neutral loci to patterns of variation in traits putatively under sexual selection, like color, mating call, and body size, to better understand the role of sexual selection vs. genetic drift in generating the striking phenotypic variation seen among *O. pumilio* populations.

1.4 Study System and Chapter Summary

My dissertation contributes to our understanding of the evolution of divergence in animal coloration and its perception using a highly polymorphic poison frog species as a case study. The strawberry dart frog, *Oophaga pumilio*, is a species that has rapidly diverged in color across the Bocas del Toro archipelago of Panama (Summers *et al*. 2003). On most islands in the archipelago there is just a single *O. pumilio* morph that typically differs from populations on nearby islands in color (Myers and Daly 1976, Wang and Summers *et al.* 2010, Rudh *et al.* 2010). This pattern alone might suggest that genetic drift played a large role in generating the diversity of color. However, the island chain developed quite recently, in the past 1000-5000 years (Anderson and Handley 2002), suggesting that drift alone could not likely have produced the observed color diversity. Color in *O. pumilio* is used both as an aposematic signal of frog toxicity (Saporito et al. 2007, Hegna *et al.* 2010), as a sexual signal in female mate choice (Reynolds and Fitzpatrick 2007, Maan and Cummings 2008, Richards-Zawacki and Cummings 2010, Maan and Cummings 2012), and by males when evaluating rival males (Yang *et al* 2016). However, while both natural selection

and sexual selection appear to be operating on color, the relative roles that the various evolutionary forces have played in generating the diversity in color remain unclear.

My research used several approaches to investigate the evolution of coloration in this frog. Sexually selected signals are selected to be maximally detectable by the intended receiver (Endler 1992). Therefore, the direction of selection on color evolution depends in part on the visual capabilities of the frogs. Differences between *O. pumilio* morphs in their color vision may contribute to the previously observed differential color-based mate preference and male-male aggression biases. I investigated whether several aspects of frog color vision differed between morphs to test the hypothesis that differences in color vision to contribute to the frogs' known color-based behavioral biases. I tested for convergent evolution in the expression of opsins, the light sensing proteins in the eye, between different morphs of the same color, using green and red morphs. I also compared expression between morphs from two polymorphic populatoins with those of their nearest monomorphic neighbors. In addition to opsin expression differences, I sequenced the opsin genes from different morphs to test for variation that might contribute to differences in their color vision.

I also investigated the genetic architecture and relative contribution of genes and environment to the production of color, to estimate whether the potential for selection to shape coloration differs between the morphs. *Oophaga pumilio* uses both genetically based and dietarily derived pigments to produce color, and the suite of pigments used to produce color varies by morph (Freeborn 2020). I used a captive breeding study to investigate the inheritance patterns and clarify the contribution of additive genetic variation to the color phenotype. I used controlled breeding of frogs from three morphs, including F1 crosses between morphs, as well as backcrosses between the F1 and parental morph,s to investigate the inheritance patterns of color. I also used pure morph

breeding in the three morphs to test for differences between them in narrow-sense heritability, a measure of the amount of phenotypic variance attributable to additive genetic variance (vs. environmental and other genetic effects).

Finally, I used landscape genetics to examine the evolutionary forces acting on color, as well as two other traits that could be evolving under sexual selection in *O. pumilio*. I compared neutral genetic variation with variation in phenotypic traits across populations from an island with a polymorphic region in the north, and variation in shades of red throughout the rest of the island. This design allows me to compare patterns of divergence in traits with that of neutral divergence, thus testing whether drift or selection is contributing more to divergence.

In summary, my dissertation takes a variety of approaches to improve our understanding of the evolution of color divergence in *O. pumilio.* My findings contribute to our understanding of the evolution of color in animals, a trait known to exhibit intraspecific variation in a variety of taxa (Sandkam *et al*. 2015, Ng *et al*. 2012, Stapley *et al*. 2011, Williams et al 2012, rev. in Gray and McKinnon 2007).

2.0 Do differences in color vision exist between color morphs of *O. pumilio* that could contribute to observed behavioral color biases?

2.1 Summary

Animals send and receive signals for a variety of purposes with signals that are selected for maximal detection by the intended receiver. Selection on signals used for mate choice acts on the signal to be easy to detect by the receiving choosier sex's sensory systems. For color-based sexual signals, the male's signal should be tuned to the female's visual system, which has been observed in a few taxa for which color is a sexual signal, but studies have focused more on aquatic systems. Here, I test for differences in color vision differences in the terrestrial polymorphic poison frog, *Oophaga pumilio*. Females of this frog species generally prefer to court with males of their own color morph and I hypothesize that this bias may be due in part to variation in the expression or sequence of the opsin proteins among the different morphs. I made several comparisons to test my hypothesis. First, I compared populations of three red and three green morphs to test for convergence of opsin expression and sequence by color. I also compared the morphs of two polymorphic regions with their nearby monomorphic populations to test for better discrimination in the polymorphic frogs compared to their monomorphic neighbors. I found no significant differences in opsin expression between morphs for the red and green comparison nor for one of the two polymorphic populations (DBP) and significantly higher RH1 expression for the polymorphic morphs of the second polymorphic region (CEM) than in their neighboring monomorphic populations. I also found a few mutations in opsin sequence that were only found in two individuals of one of the polymorphic populations (DBP) and several others that showed

no morph patterns. Other mutations found include a nonsense mutation and a mutation at a residue important for opsin function. Overall, I did not find convincing evidence that differences in color vision have evolved among distinctly colored populations, contrary to findings in aquatic systems, though I did substitutions, suggesting that the raw material for selection on opsin sequence is present for future evolution.

2.2 Introduction

The ability to send and receive signals is critical to animal communication. Successful signaling often results in direct fitness benefits, for instance a prairie dog heeding a warning call given by a sentry in the colony and avoiding a predator (Kiriazis and Slobodchikoff 2006), or a predator avoiding eating a toxic prey item by detecting its bright aposematic signal (Mappes *et al.* 2005). Successfully obtaining a mate also often relies on successful signaling to the opposite sex. For example, male elephants signal their sexual receptivity with loud stereotyped calls, termed rumbles, as well as odor cues to attract females (rev. in Vidya and Sukumar 2005). Additionally, access to mates can rely on showy sexual signals to either attract a mate such as the male peacock's long, showy tail (Loyau *et al.* 2005) or the superb lyrebird's complex call (Dalziell and Welbergen 2016). For a signal to be effective, three things must happen; the signaler must produce the signal, the signal must travel to the receiver, and the signal must be perceivable by the receiver and stand out from background noise in the environment (Endler 1992). That is, the superb lyrebird must vocalize and create the call, the call must stand out from the other noise in the forest and then the call must be perceivable by ears of a nearby female. Thus, the relevant sensory system of the receiver is an important component of signal evolution, and selection favors signals that are tuned

to match the receiver's sensory system capabilities. Variation in male signal perceivability by females can thus act as a selective force shaping male sexual traits.

Sexual selection as an agent of evolutionary change relies on members of one, typically males, successfully signaling their fitness to members of the choosier sex, typically females. The ability of the chooser to perceive differences in the signals of potential mates will influence which mate is chosen, as better perceptual ability gives the chooser more information to use (rev. in Kokko *et al.* 2002 and in Ryan and Cummings 2013). Males with a non-optimal signal with regard to female perception may have their signal lost against background noise, resulting in fewer or no matings for the male (Endler 1992). Therefore, in most animals where females are the choosier sex, the evolution of variation in male sexual signals is thought to be constrained by the perceptual abilities of the female sensory system. Animals generally have better perceptual abilities in some portion(s) of their total sensory range. For example, a female may be able to see colors across the range from 300 – 700nm in wavelength but can detect and/or discriminate colors from 500-600nm most easily. This means that males are more successful in attracting the female's attention and interest if the value of his sexual signal falls within the range that the female can perceive best because that maximizes the efficiency with which the sexual signal is transmitted (Endler 1992). If signal perception varies among females, this has been proposed to result in males "specializing" on different ranges of the sexual signal, producing variation in the signal driven by variation in the female sensory system (Endler and Basolo 1998). This has been seen in sticklebacks (Boughman 2001) and in African cichlid fish (Maan *et al.* 2006), where variation in the light environment has produced females with a higher sensitivity to the more abundant wavelengths of light available in their habitats and a corresponding shift in male coloration to exploit this peak in female sensitivity. Likewise, sexual signals have evolved differences in response to female perception using other

sensory modalities. For example, urban birds call at higher frequencies than non-urban birds, so their call can be better heard by females over the loud, lower frequency urban noise (Patricelli and Blickley 2006) Thus, the sensory capabilities of the females are important for their perception of the male signal..

Vision in vertebrates is made possible by light sensing photoreceptors (rods and cones) in the retina of the eye, which possess the photopigment opsin. Color vision is possible when multiple cone types, each containing an opsin protein that absorbs light at a different wavelength, are present in the retina of an organism's eye (rev. in Cronin *et al.* 2014). These opsin proteins absorb incoming light, which produces a conformational change in the chromophore, a vitamin-A based signal ligand bound to the opsin protein (rev. in Bloch 2016). This reaction sets off the phototransduction cascade which, among other events, results in the disassociation of the opsin and the chromophore (known as photoreceptor bleaching) and causes a signal to be sent through the nervous system to the brain (Fain *et al.* 2010). The relative signal strengths from the different cone types in the same section of the retina are compared by the brain to determine the color of an object (rev. in Gegenfurtner and Kiper 2003, Cronin *et al.* 2014). This complex sequence of events allows for numerous ways that color vision can be modified by evolution.

Changes to the opsin proteins are one mechanism of visual adaptation in vertebrates. The amino acid sequence of the opsin protein, particularly at certain key sites, is important in determining the range of wavelengths that the opsin absorbs (Yokoyama and Radlwimmer 1998, rev in Yokoyama 2008, rev. in Bowmaker 2008). Thus, changes in the amino acid sequence can shift the wavelength of maximum absorption (λ_{max}) of the protein and over time, evolution can select for those changes that improve the color vision in areas of the color spectrum most relevant to the animal, where selection favors better sensitivity and/or discrimination (rev. in Price 2017).

For example, in species where sexual selection acts on a color-based trait, selection for opsin mutations that allow better color discrimination of the sexual signal would be expected. This type of population-specific variation in opsin amino acid sequence has been found in *Pundamilia* cichlid fishes (rev. in Maan and Seehausen 201) and between guppy species (Sandkam *et al.* 2015b), both of which have females that exhibit differential color-based mate preferences.

Changes to the expression of the opsin mRNA transcripts is another way that vertebrate visual systems can evolve to more effectively detect and discriminate important signals. Higher opsin expression is generally thought to indicate a higher density of the opsin protein in the retina, either due to higher density of photoreceptors, or due to increased density of opsin protein per photoreceptor (rev. in Price 2017). A higher density of the visual pigment in the retina increases the number of photons captured by the photoreceptor (rev. in Cronin *et al.* 2014). Since each photoreceptor has a unique opsin protein, expression of the opsin genes can be used as a measure of the animal's visual sensitivity across different regions of the visible spectrum and daily cycling of opsin expression has been shown to correlate with behavioral sensitivity to that opsin's color range in zebrafish (Li *et al.* 2005). Thus, variation in the expression of the different opsin proteins can be used to infer differences in color vision among individuals or populations. Studies in both the African cichlids (Smith *et al.* 2011) and guppies (Sandkam *et al.* 2015a, Sandkam *et al.* 2015b) have found intraspecific opsin expression differences associated with differences in mate choice, suggesting a role for color vision in the divergence of sexually selected traits via sexual selection. Sexual selection could thus drive the process of speciation by producing the phenotypic variation necessary for divergence if mate preference diverges along with the trait and reduces mating between the different variants (Ritchie 2007, Servedio and Boughman 2017). This could generate behavioral reproductive isolation if there is variation in the sexual signal (typically male

ornamentation) that co-diverges with female preference for this trait (Coyne and Orr 2004, Panhuis *et al.* 2001).

A few studies have shown convincing evidence for color vision differences associated with differential female preferences, creating disruptive selection that ultimately drives speciation. However, these have focused on aquatic animals, taking advantage of characteristic shifts in the wavelengths of available light with water depth and water clarity (rev. in Carleton 2009). However, to understand more fully the role that intraspecific differences in sensory systems play in speciation we need to study how this mechanism works in a wider range of taxa. The optical properties of air and water vary in how they affect attenuation of light of various wavelengths. The characteristics of light available for signaling/detection can be affected by both the water quality/depth in aquatic systems (rev. in Cronin *et al* 2014) and, for example, the density of the tree canopy in a terrestrial environment (Endler 1993). Because the light environments of terrestrial and aquatic ecosystems differ, the properties of the eyes of animals in each type of habitat have evolved in different ways. Thus, it is important to understand the potential influence of color vision on divergence in sexually selected traits and preferences for them in terrestrial systems as well as in aquatic ones. However, most work in terrestrial systems has tested for phylogenetic patterns in opsin presence among bird species (Bloch 2015, Bloch *et al.* 2015). The few studies that have tested for intraspecific differences in color vision associated with sexual selection in terrestrial animals have not found a clear link between color vision and sexual selection (lizard, Yewers *et al.* 2015; damselflies, Bybee *et al*. 2012 Huang *et al.* 2014).

Oophaga pumilio is a terrestrial poison dart frog that shows extreme color polymorphism across the Bocas del Toro archipelago in northwestern Panama (Myers and Daly 1976, Wang and Summers *et al.* 2010, Rudh *et al.* 2010). This archipelago formed recently (est. 1000 – 5200 years

ago, Anderson and Handley 2002), and recent work suggests that sexual selection played a role in shaping the evolution of color variation in this species. Female *O. pumilio* tend to prefer males of the local color (Reynolds and Fitzpatrick 2007, Maan and Cummings 2008, Richards-Zawacki and Cummings 2010, Maan and Cummings 2012) and territorial males are more aggressive toward encroaching males of the local color (Yang *et al* 2016) though these behavioral biases appear to be stronger in monomorphic populations than in the few polymorphic ones (Yang *et al*. 2016, Richards-Zawacki *et al.* 2012). Thus, it appears that both sexes have the capability to distinguish differences between color morphs and differences in color vision could have contributed to the evolution of these differential color-based biases. Recent evidence suggests that female mate preference and male aggression biases are learned as tadpoles through imprinting on the mother's color (Yang *et al.* 2019b). Thus, the ability to accurately distinguish colors appears to be important from an early age in this species.

However, our knowledge of the visual system of this (and most other) frogs is far from complete. What we do know comes about *O. pumilio* from microspectrophotometry data from a single captive individual. These data suggest that *O. pumilio* have trichromatic color vision owing to three cone types being present in the retina, a red cone with wavelength of maximum absorption, λ_{max} of 561nm, a green cone with a λ_{max} of 489nm and a blue cone with λ_{max} of 466nm (Siddiqi *et al.* 2004). While the typical red (LWS) and blue (SWS1) cone opsins are present in the *O. pumilio* genome, it appears that the typical green cone opsin, RH2, has been lost from Dendrobatid frogs, including *O. pumilio* (Wan *et al.* in review). However, based on the microspectrophotometry (MSP) data from Siddiqui *et al.* (2004), the pigment in the green cones of *O. pumilio* has a similar MSP absorbance spectrum to the pigment in the rod, suggesting that the green cones are using the same opsin protein as the rods (RH1).

Here I test the hypothesis that two aspects of color vision, opsin sequence and opsin mRNA expression, differ between morphs of *O. pumilio* in ways that putatively improve discrimination of the local frog color. I tested for convergence in color vision characteristics between populations containing morphs of a similar color, taking advantage of the existence of multiple morphs of *O. pumilio* that appear to have converged on red or green coloration. If color vision has been important in shaping mate preference, the red morphs should have more similarity to each other in their opsin sequence and expression than they do to the green populations and vice versa. Additionally, I examine these same color vision components in two polymorphic populations, where frogs must be able to accurately discriminate between the different colors present in order to show color associated behavioral biases. Therefore, I predicted that frogs in the polymorphic populations would have visual systems that are better able to discriminate between their two or more local colors than neighboring monomorphic populations would. The polymorphic nature of this species offers a unique chance to test for differences in the visual system between different color morphs of a species that is putatively diverging due to sexual selection. Additionally, this study improves our understanding of intraspecific variation in the visual system of a terrestrial species. My findings offer a test of the idea that divergence in visual systems can contribute to the early stages of speciation.

2.3 Methods

To test for differences in opsin sequence and expression between *O. pumilio* morphs, I collected five frogs from each of 10 monomorphic populations as well as five per morph from two polymorphic populations. Three of these monomorphic populations contained green frogs and

another three contained red frogs; these were used for the red/green comparisons (Figure 2-1). Red is the ancestral color of this species though seems to have evolved more than once (Freeborn 2020) and the three green populations converged on green (Figure 2-2), suggesting that green evolved multiple times. The remaining four monomorphic populations used were the closest (geographically, and often genetically (Figures 2-1 and 2-2)) monomorphic populations bearing a similar coloration to one of the polymorphic morphs. We used these monomorphic/polymorphic population sets to test for differences in color vision between frogs that encounter a wider variety of conspecific coloration in their local population (polymorphic) and their monomorphic neighbors (Figure 2-1). That is, for the Dolphin Bay polymorphic population (which contains red, blue, and intermediately colored brown frogs), I used a nearby red monomorphic (SCBH) and a blue monomorphic (NRA) population to compare with the red and blue frogs from the polymorphic area (the intermediate brown phenotype is not found elsewhere) Likewise, for the cemetery polymorphic population, which contains red and yellow-green frogs, I used nearby yellow-green (ICN) and red (LL) monomorphic populations as comparisons (Figure 2-1).

Frogs were hand captured in the field and housed them in plastic terraria for transport to the Smithsonian Tropical Research Institute's Bocas del Toro field station. Frogs were euthanized by rapid decapitation followed by double pithing, after which the eyes were quickly dissected out, then punctured them with a dissecting pin (to allow penetration of RNAlater). Eyes were then placed in RNAlater (Ambion), and stored them overnight at 4^{o} C before moving them to -20^{o} C. The samples were shipped on dry ice to the University of Pittsburgh where they were again stored at -20 ^{o}C.

I extracted total RNA from the retina of one eye per animal and reverse transcribed the mRNA into cDNA. To do this, immediately prior to extraction, I dissected out the retina of one

haphazardly chosen eye and homogenized it in Trizol (Invitrogen) with a mini pestle in a 1.5mL tube. I extracted total RNA using the DirectZol RNA mini-prep kit (Zymo) following the manufacturer's instructions and including the optional 15min on-column DNase incubation. The resulting RNA was then quantified using an Epoch Microplate Spectrophotometer (Biotek) and 250ng of total RNA was converted into cDNA using the SuperScript IV VILO kit (Invitrogen), which includes both oligo(dt)18 and random hexamers to prime the cDNA synthesis reaction. Additionally, I ran a no reverse transcriptase reaction (NRT) that had all the same components as the reverse transcribed (RT) samples, except for the reverse transcriptase. I used these samples as controls to confirm the lack of contamination in the sample during qPCR reactions. cDNA samples were diluted 1:10 in molecular grade water prior to qPCR to overcome PCR inhibition.

2.3.1 Sequencing

To test for opsin sequence differences among morphs, I designed primers to span the entirety of each *O. pumilio* opsin gene transcript. Since the opsin transcripts are longer than a typical Sanger sequence (transcript lengths: LWS-1223, RH1-1406, SWS1-1309), I designed two overlapping primer sets (Table 2-2), each spanning one half of the gene transcript, from the transcripts of each *O. pumilio* opsin gene (provided by L. O'Connell) using Primer3 (Untergasser *et al*. 2012, Koressaar and Remm 2007). Initially I used GoTaqGreen (Promega) for PCR reactions. However due concerns about the high error rate of Taq polymerase, I switched to a hot-start, high-fidelity polymerase Q5 (New England Biolabs, Ipswich, MA) part way through the study. To account for this switch, I redid the sequencing with Q5 for any sample that had a change in the GoTaqGreen data to confirm the SNP was real and not a result of the high error rate of the

Taq. PCR reactions were done in 25 μL total volumes starting with an initial denaturation at 95°C for 90s, followed by 40-45 cycles of denaturation at 95°C for 30s, annealing at 63-67°C for 30s, and extension at 72°C for 40s with a final extension of 72°C for 7 min to ensure enough product for sequencing from the cDNA samples. I visualized PCR products on a 0.75% 50mL agarose gel and cut out bands of the expected size. I then cleaned-up the samples using the Monarch DNA Gel Extraction Kit (New England Biolabs, Ipswich, MA). Samples were Sanger Sequenced on an ABI 3730xl at Azenta Life Sciences (Chelmsford, MA).

Sequence ends were trimmed and trace files visually examined for quality. I manually checked any base pairs that were called as "N" for possible heterozygosity using Geneious v 9.1.8 (www.geneious.com). Following sequence quality checks, I aligned the sequence halves for each gene with the full mRNA transcript (provided by L. O'Connell) in MEGA v 7.026 (Kumar *et al.* 2016), manually combined sequences to create contigs, and translated contigs using ExPASy (web version, Swiss Institute of Bioinformatics). In cases where the two sequence parts failed to overlap, I inserted the appropriate number of Xs to fill the gap. mRNA and protein sequences were aligned in MEGA v 7.026 (Kumar *et al.* 2016) and examined for any differences between sequences. For every SNP and amino acid change, I manually examined the trace file at the relevant base pair(s) to confirm the SNP.

2.3.2 Gene Expression Assay

I developed qPCR primers for the *O. pumilio* opsin proteins. Since the *O. pumilio* genome assembly is not complete (Rogers *et al*. 2018), I used unpublished opsin transcripts provided to us by L. O'Connell. I then aligned these opsin transcripts to the scaffolds of the *O. pumilio* genome

assembly (Rogers et al. 2018) in Geneious v 9.1.8 (https:www.geneious.com) to find the exon-exon boundaries within the transcript. I used primer3 (Untergasser *et al.* 2012, Koressaar and Remm 2007) to design primer sets that spanned these exon-exon boundaries, ideally on the 3' ends, in the *O. pumilio* transcripts, to ensure that I only amplified cDNA in my qPCR assay. To design primers for my reference gene, I first found the transcript from the annotated transcriptome using a BLAST search in GenBank (Altschul *et al.* 1990) and completed the steps outlined above to find primer sets where one primer spanned the exon-exon boundary. I used a rod specific phosphodiesterase (PDE6g) that is involved in the rod phototransduction cascade as a references gene as it appears to have consistent expression between individuals. I calculated the primer efficiency for all primer sets by running a qPCR assay (see below) on a five step 1/10 dilution series featuring a cocktail of templates in triplicate. I then calculated primer efficiency (following Rasmussen 2001). Primer sets for each gene were chosen such that efficiency was 90-110%.

After choosing primers, I ran qPCR on a QuantStudio 3 (Applied Biosystems) on the cDNA samples to measure gene expression of the opsins relative to my reference gene. I ran 20uL reactions with 10uL of SYBR Select Master Mix (Applied Biosystems), 250nM of each primer (1uL each of 5uM primer), 2uL of template, and added water to 20uL. The qPCR protocol was as follows: 2min at 50°C to activate the uracil-DNA glycosylase in the master mix, followed by an initial denaturation/activation of AmpliTaq polymerase stage at 95°C for 2 min. This was then followed by 40 cycles of 95ºC for 15s, 55°C for 15s and 72°C for 30s. After the amplification stage I ran a melt curve analysis to determine the number of products per well. This consisted of 95°C for 15s, 55°C for 1min and then ramping up to 95°C at a rate of 0.15°C/s. Samples were run in triplicate with all genes from a given sample run on the same plate. If one of the opsin genes required a redo of the qPCR, the reference gene was rerun as well and used for analysis of that

gene to avoid differences between plate set-up and qPCR run adding to variation within samples. A redo of the qPCR for a particular sample was deemed necessary if the reference gene Ct value was outside of its normal range of values as this gene is expressed consistently across samples. That is, the qPCR was rerun if the reference gene had a Ct value greater than one standard deviation above that of the reference gene (i.e., if Ct > 28), indicating that the reference gene values were off as this gene is consistently expressed across individuals. A redo was also done if the standard deviation of the triplicates for any gene was above 0.5. For each qPCR run, I included for each sample on the plate a no-reverse transcription control well for all four genes (i.e., four no RT controls per sample). These controls included all qPCR components but the NRT sample was used for the template instead of the RT. Each run also had a no template control well for each gene to ensure there were no contaminants introduced during qPCR setup. I calculated the relative quantity of transcripts for each sample using the delta Ct method (Pfaffl 2001), calculated as the Ct value of the reference gene (PDE6g) minus the Ct value of the opsin gene. This delta Ct value was used for all downstream analyses.

2.3.3 Statistical Analysis

All analyses were run in R v. 3.6.2 (R Core Team 2021) except where noted. I tested model assumptions by 1) looking for linearity of the data by checking residual vs fitted value plots, 2) checking normality in the residual by checking the Q-Q plots of the residuals and 3) checking for outliers using residuals vs leverage plots. In all cases I was able to conclude that model assumptions were met. I used the *lm* function in package "*lme4*" (Bates *et al.* 2015) to specify linear models to test for differences in opsin expression between monomorphic populations in the

red and green frog comparison. I used separate linear models for each gene with main effects of frog color (red vs. green), population, sex (male vs. female), and a color by population interaction term. I used Tukey tests in GraphPad Prism v 9.3.1 for Windows (GraphPad Software, San Diego, CA) for post-hoc comparisons of expression for each population pair of the red and green populations (GLB, HOSP, OLP, NIPA, PUNT, and CAN in Figure 2-1). I tested each of the two polymorphic populations for differences in opsin expression among its constituent color morphs in separate linear models. Each morph was included as a separate population in these models, as were the nearest (geographically) monomorphic populations that shared the same coloration. These linear models included color, population, sex, and a color by population interaction term as main effects. I conducted post-hoc comparisons of expression for each population set (for the Dolphin Bay population comparison, DPB (three morphs), SCBH, RA; for the Cemetery polymorphic population comparison, CEM (two morphs), ICN, LL, Figure 2-1) using Tukey tests.

I also compared opsin expression profiles for all three genes together by using the delta Ct values from all three opsin proteins. If opsin expression differ among populations or morphs, morphs of the same color should have similar overall opsin expression profiles. To test this, I ran a principal components analysis (PCA) on the expression values from all three opsin genes. Separate PCA's were run for each comparison (green/red populations, cemetery population and neighboring monomorphic populations, and Dolphin Bay population and neighboring monomorphic populations) using the *prcomp* function in R. I then visually examined graphs of PC1 and PC2 for clustering by population and/or morph. If opsin expression profiles are similar for frogs of the same color, red frogs should cluster together in the region of PC space where the red opsin (LWS) loads, and this should be also true with green frogs for the green opsin (RH1) and blue frogs (SWS1).

2.4 Results

2.4.1 Sequencing

I sequenced LWS, RH1, and SWS1 for 49 individuals from 12 populations across Bocas del Toro. For most populations I sequenced two individuals but for a few populations (NIPA, NRA) I sequenced three individuals because one individual was amplified using GoTaqGreen (Promega) before I made the shift to using high fidelity Q5 polymerase (New England Biolabs). All but two of the SNPs I detected in this study found were in samples that were sequenced with Q5, and one of the two SNPs found using GoTaqGreen was also found in two other individuals that were sequenced with Q5, suggesting that the inclusion of the GoTaqGreen samples did not influence the conclusions of my study. My initial sequencing pass of two individuals per morph resulted in a nonsense mutation in one of the polymorphic populations. I therefore decided to increase my sequencing sampling to include all of the individuals included in this study for both polymorphic sites (5 per morph, totals - CEM (10), DBP (15)). I was able to sequence all three opsin genes for most individuals. However, I was unable to obtain sequences for part of RH1 and part of SWS1 for one individual from the GLB population. Additionally, I was not able to obtain overlapping sequences from my two primer sets per gene for all individuals. This necessitated filling the gap with Xs (for the missing base pairs) in 13 sequences from 11 individuals and was more common in RH1 and SWS1 (5 and 6 samples, respectively) than in LWS (2 samples). The gaps ranged from 3-172 bp long with a mean gap size of 58.3bp. Gaps were longest in the SWS1 sequences (mean = 90.4bp) compared to LWS (mean = 62.5bp) and RH1 (mean = 19.3bp).

I found several SNPs resulting in changes to the amino acid sequence. LWS and RH1 each had four amino acid residues with missense mutations, one of which was shared by individuals from different populations (two in LWS, three in RH1). One of the missense mutations in LWS was found in the K296 site (bovine rhodopsin numbering) which is critical for chromophore binding. Another missense mutation in LWS was shared by two frogs from the polymorphic DBP that were different color morphs. SWS1 had 11 amino acid residues with missense mutations and one with a nonsense mutation (residue 108, residue 111 in bovine rhodopsin (Hargrave *et al.* 1983)). One of residues had a missense mutation that was shared by two frogs of different morphs from the polymorphic DBP population. Most frogs with a SNP were heterozygous for the wild type sequence at that base pair, but this was not the case for the K296 mutation or the nonsense mutation (Figure 2-3). Additionally, both the nonsense mutation and the K296 mutation were found in frogs sequenced using Q5.

2.4.2 Gene Expression Assay

RNA extractions were successful for all 74 retina samples. Most samples resulted in extractions that permitted us to use 250 ng of RNA as template in the cDNA reaction. Some samples produced nearly 250 ng of RNA, and for these I investigated the potential impact of this lower template amount on my results using my qPCR data. The average Ct values for samples that contained < 250 ng of RNA were generally less than one Ct value higher than the average of samples with the full 250 ng (range of Ct difference between groups 0.32-1.029). However, this difference in starting template values was controlled for via the reference gene, as concentrations of opsins and reference gene would have been equally reduced.

There were no significant differences in expression between red and green frogs for any of the opsins (Figure 2-4; linear model: $F_{4,24} \leq 1.88$, $p \geq 0.15$). Comparing among the overall opsin means for the red/green frog comparison, RH1 generally had the highest expression (mean Ct = 26.2 ± 1.30 SD, mean delta Ct = -0.19 ± 0.84 SD), followed by LWS (mean Ct = 28.9 ± 1.34 SD, mean delta Ct = -2.665 ± 0.68 SD), and SWS1 had the lowest expression (mean Ct value = 30.1 ± 1.13 SD, mean delta Ct = -4.053 ± 0.84 SD). A similar pattern was seen in the comparison of opsin expression in Dolphin Bay polymorphic zone frogs (red and blue morphs) with their nearest monomorphic neighbors. Here again, no significant difference in expression was observed between morphs or populations for any of the opsin genes (Figure 2-5; linear model: $F_{4,20} \leq 1.2$, $p \geq 0.33$) and RH1 had the highest expression (mean Ct =26.02 ± 1.62, mean delta Ct = 0.802 ± 0.62), followed by LWS (mean Ct =29.145 ±, 1.79 mean delta Ct =-2.683 ± 0.60) then SWS1 (mean Ct = 30.157 ± 1.51, mean delta Ct = -3.453 ± 0.76). For the Cemetery red/yellow polymorphic zone, I found no significant differences between morphs or populations in expression for LWS (linear model: $F_{4,15} = 1.9$, $p = 0.16$). However, I found marginally non-significant differences in expression in SWS1 (linear model: $F_{4,15} = 3.019$, $p = 0.052$). Additionally, for SWS1 the monomorphic population ICN was marginally non significantly lower than two populations, cemetery red ($p = 0.08$), and LL ($p = 0.085$). However, I found significant differences in expression for RH1 among samples (linear model: $F_{4,15} \geq 4.09$, $p \leq 0.02$) with a significant color by population interaction ($p = 0.034$). Both Cemetery morphs had greater mean RH1 expression than either of the nearby monomorphic morphs, though the polymorphic morphs were only significantly higher than the green ICN morph ($p \leq 0.01$). Similar to the other comparisons, RH1 had the highest expression (mean Ct = 25.613 ± 1.03 SD, mean delta Ct = 0.139 ± 0.66), followed by LWS (mean

Ct = 28.427 ± 1.11 SD, mean delta Ct = -2.614 ± 0.70), and SWS1 (mean Ct = 29.723 ± 1.00, mean delta Ct = -4.066 ± 0.93).

The PCA analysis including expression for all three opsins showed significant eigenvalues for PC1 (1.495) and PC2 (1.001), which explained 49.848 and 33.367 percent of the variation, respectively. RH1 and SWS1 both aligned strongly and negatively with PC1 (Table 2-4), while LWS aligned strongly and positively with PC2 (Table 2-4). When PC1 and PC2 were graphed together, frogs did not generally cluster together by population, nor did they cluster together by dorsal color category (i.e., red, green, etc.) (Figures 2-6-8).

2.5 Discussion

Overall, my hypothesis that opsin expression and opsin sequence differ between morphs was largely unsupported by the data. There were no mutations found in residues at any of the key sites for spectral tuning that have been identified in other taxa (Yokoyama and Radlwimmer 1998), suggesting that the missense mutations I found may not impact the λ_{max} values of the opsin proteins. Additionally, most of the mutations in the sequences that I found were only found in one individual and almost all of the mutations that were found in multiple individuals were not found in individuals from the same population, nor did they typically occur in frogs of the same color, suggesting the mutations did not coevolve with color. It should be noted, however, that my sampling scheme of only five frogs per morph likely missed some of the variation present in opsin gene sequences. Deeper sampling across a smaller number of populations might reveal spatial patterns in some of the mutations found in this study.

However, I did still find some interesting results from the opsin sequencing as a couple mutations found in this study do likely have a major impact on color vision for individual frogs. The nonsense mutation found the SWS1 opsin in a red frog from the polymorphic cemetery population occurs early in the protein and likely has a major impact on that frog's color vision. Since the frog is homozygous for the mutation, this suggests that this frog has dichromatic color vision with red and green cones, and therefore sees quite differently than its trichromatic conspecifics. Since this frog was from the polymorphic cemetery population, this frog likely has a reduced ability to distinguish between the red and yellow-green morphs, as these variants would appear as different shades of red to it (Jefferson and Harvey 2006). Another mutation that is also likely to have a major impact on color vision is the mutation of the lysine residue at position 311 in the O. *pumilio* LWS transcript (296 by bovine rhodopsin numbering) as it is a mutation in the site of chromophore binding (rev. in Pepe 1999). This mutation was also found in the homozygous state, so there is no wild-type protein to rescue the mutant phenotype. Since the mutation results in a change in the amino acid residue (Table 2-3), the chromophore binding is likely to be disrupted and result in limited functionality of the LWS opsin for that frog. That suggests that the frog has dichromatic color vision using green and blue cones and, since the *O. pumilio* green cones have low absorption of wavelengths of light longer than 575nm (Siddiqui *et al.* 2004), this frog likely has difficulty with detecting photons of light in the orange and red ranges of the color spectrum. Since this frog is from the Dolphin Bay polymorphic population, which has red, blue, and intermediate (brownish) colored frogs (Yang *et al.* 2019a), this suggests that that frog likely has difficulty discriminating between red frogs and intermediate frogs on the red end of the spectrum. Additionally, mutations in this site are linked with retinitis pigmentosa, a group of diseases

characterized by retina degradation in humans (Robinson *et al.* 1992), suggesting that this mutation might have wider ranging effects on the frog than just impacting color vision.

It is hard to infer what impact the other mutations I found would have on the frog's color vision as they are not near any of the known key sites based on the previously studied (non-amphibian) species. Several of the mutations are in sites that are invariable across a wide range of species (see Table 2-3), suggesting that these are important sites for the protein and that a mutation in the amino acid residue could impact functionality by changing the λ_{max} or by disrupting protein folding. Additionally, since many of the mutations found are in the trans-membrane domains, it is possible that they could cause minor disruption to the protein structure, which could change interactions between residues and potentially have impacts on protein λ_{max}, particularly since many changes result in a change of charge or polarity at that residue (Table 2-3). Additionally, my finding of more mutations in the SWS1 compared to LWS and RH1 is consistent with the findings of a survey of several frog clades, including the poison frog family Dendrobatidae, which found more amino acid residues under positive selection in SWS1 than in LWS and RH1 (Wan *et al.* in review). This suggests stronger selection on SWS1, perhaps due to the apparent loss of SWS2 in Dendrobatid frogs (Wan *et al.* in review), and the shifting of the SWS1 opsin from having a λ_{max} in the UV range to having a λ_{max} the blue range in *O. pumilio*. However, none of these sites under positive selection in the broader amphibian opsin study were found in sites known to be key for spectral tuning in other, previously studied taxa (Wan *et al.* in review), nor were any of the mutations that I found in my study. This suggests the potential for amphibian key sites to be different than those found in other non-amphibian taxa, though a recent survey of sequence variation in frog species of Texas did find variation in several key sites known from other species (Schott *et al* 2022). Perhaps some of the known key sites are identical for amphibians, while other

key sites are unique to amphibians. Studies that specifically test for key sites that results in shifts in λ_{max} in amphibian opsins are necessary to better understand opsin evolution in this group.

Opsin expression was fairly consistent across individuals and across morphs as in all frogs; RH1 had the highest expression and SWS1 had the lowest expression, with LWS expression falling between the reference gene (PDE6g) and SWS1. Comparisons of opsin gene expression for red and green frogs from multiple populations also revealed no consistent expression differences by color. This comparison between red and green frogs showed no differences between the frog color groups and there were no consistent patterns among populations within color categories, suggesting that opsin expression hasn't evolved in concert with frog color. While I did not see differences in opsin expression levels between morphs in my red/green comparison, I did see some differences in expression when I compared expression between frogs from polymorphic and monomorphic populations. Here I found that, RH1 expression differed significantly between morphs, and SWS1 was marginally non-significantly different between morphs in one polymorphic population. Both morphs in the polymorphic Cemetery population had greater expression of RH1 than frogs from the monomorphic yellow-green population. The greater expression of the RH1 opsin for the Cemetery morphs suggests that opsin expression is upregulated in these frogs, which would allow for better color discrimination ability in the yellow to orange-red range. Increased opsin expression has been linked to increased opsin density in the photoreceptors and to increased sensitivity to light (rev in Price 2017). Therefore, it is conceivable that this increase in expression for the RH1 is due to selection favoring frogs being able to better distinguish between the red and yellow frogs present in the Cemetery polymorphic population. An increase in the RH1 expression would increase the overlap in wavelengths absorbed by the green and red cones, thus allowing for better color discrimination in the overlap region which covers the

green, yellow, and orange part of the color spectrum (Siddiqui *et al.* 2004, Cronin et al. 2014). However, because RH1 appears to be used in both the rods and green cones (Wan *et al.* in review), this increase in expression of RH1 for the cemetery morphs could also be explained by increased rod usage. Most studies that have tested for changes in rhodopsin (RH1) mRNA transcript levels throughout the day have found that numbers of mRNA transcripts increase throughout the day and decrease overnight (Hartman et al 2001, Kamphuis et al 2005, Yu et al. 2007, but see McGinnis et al. 1992). Therefore, in these systems it appears that rod expression is highest during daylight, and therefore my measure of increased RH1 expression in cemetery morphs of *O. pumilio* might be due to increased rod expression in that population. However, lighting conditions are not expected to differ much among *O. pumilio* populations (Yeager 2015) and thus an increase in RH1 for rod usage in just this population seems unlikely. However, since I can't distinguish between expression of RH1 used for rods vs cones, changes in expression due to increased rod usage cannot be ruled out. Additionally, all samples were taken at approximately the same time of day and therefore differences associated with circadian variation of opsin expression are unlikely to explain the observed differences.

While I did find some differences in opsin sequences and expression levels in *O. pumilio*, most of the differences that I can infer would impact color vision were present in the polymorphic Cemetery and Dolphin Bay populations and not in the monomorphic populations. One potential reason that only the Cemetery morphs differed in expression compared to nearby monomorphic populations is because the morphs are closer in color (yellow vs. red) in the Cemetery populations than the Dolphin Bay morphs (blue vs. red). Discriminating between red and blue is much easier as the two are further apart on the color spectrum. Thus, changes in opsin expression to improve color discrimination ability might not have been selected for in the Dolphin Bay frogs. However,

Dolphin Bay did have several residues with mutations shared by two frogs, which was not seen in any other population. While this could potentially be due to evolution of spectral tuning to the different color morphs present in the polymorphic population, it more likely represents a sampling bias, as 15 individuals (5 red, 5 blue, 5 intermediate) from this population were sampled, compared to only 10 (5 red, 5 yellow) in the Cemetery population, and two in every other population in this study. Since most of the mutations found were in samples from the polymorphic populations, this suggests that sampling a larger number of individuals from each population would likely yield more mutations and a better test for differences in the frequencies of mutations among populations.

In summary, differences in color vision due to expression or sequence differences do not appear to have been a driving force behind the development of different color-based behavioral biases seen in different color morphs of O. pumilio as neither opsin expression nor opsin sequence showed consistent differences amongst morphs of the same color. However, variation in both opsin sequence and expression are present in *O. pumilio* and appear to be distinct in the polymorphic populations. Thus, it is possible for color vision to evolve via differences in opsin expression and/or sequence in this species as variation exists, and is not particularly rare, in both. This may be especially true for the polymorphic populations, as we found a large number of expression and sequence differences in these populations and selection on vision in cases where variation exists within the population may be especially strong. However, as sample sizes were uneven between polymorphic and monomorphic populations and were generally small overall, deeper sampling within populations needs to be done before solid conclusions can be drawn. Further work should thus investigate whether these changes in opsin expression and sequence result in any meaningful change in the color vision of the frog. Overall, I found little evidence for differences spectral tuning among morphs, contrasting with results from aquatic systems (Smith *et al.* 2011, Sandkam *et al.*

2015a, Sandkam *et al.* 2015b), and suggesting that properties of water as a medium for vision might produce stronger selection on the visual systems compared to air. However, *O. pumilio*'s hypothesized recent radiation into distinct color morphs, which is hypothesized to have occurred as the islands formed (~1000 – 5200 ya, Anderson and Handley 2002), is more recent than the cichlid radiation (300,000 ya, rev. in Seehausen 2006). This suggests the possibility that *O. pumilio* morphs haven't had enough time to diverge in their opsin expression and sequence. Additionally, it is possible that while opsin expression and sequence do not vary between morphs, that other aspects of color vision do. For instance, variation in chromophore usage between morphs could be used in spectral tuning, as frogs have been found to use two different chromophore forms and the form of the chromorphore that is used affects $\lambda_{max.}$. However, what little we know about chromophore usage in amphibians comes primarily from differences found between tadpoles and adult frogs in a single species and this pattern has not been further investigated in frogs to my knowledge (Liebman and Entine 1968). Additionally, *O. pumilio* photoreceptors contain an oil droplet, which could be used to tune color vision via pigmentation in the oil droplet (Cronin *et al* 2014). Limited work on captive *O. pumilio* have found clear oil droplets in the cones (Hailman 1976, Siddiqui *et al.* 2004), suggesting a lack of tuning, but as these studies used a limited number of captive frogs that were not fed their typical wild diet, it is possible there might be differences in wild frogs that have yet to be discovered. While this study adds to our knowledge of color vision differences in sexually selected terrestrial species, there is still much more to investigate to better understand color vision's role in shaping color-based biases and their role sexual selection for terrestrial animals.

2.6 Figures and Tables

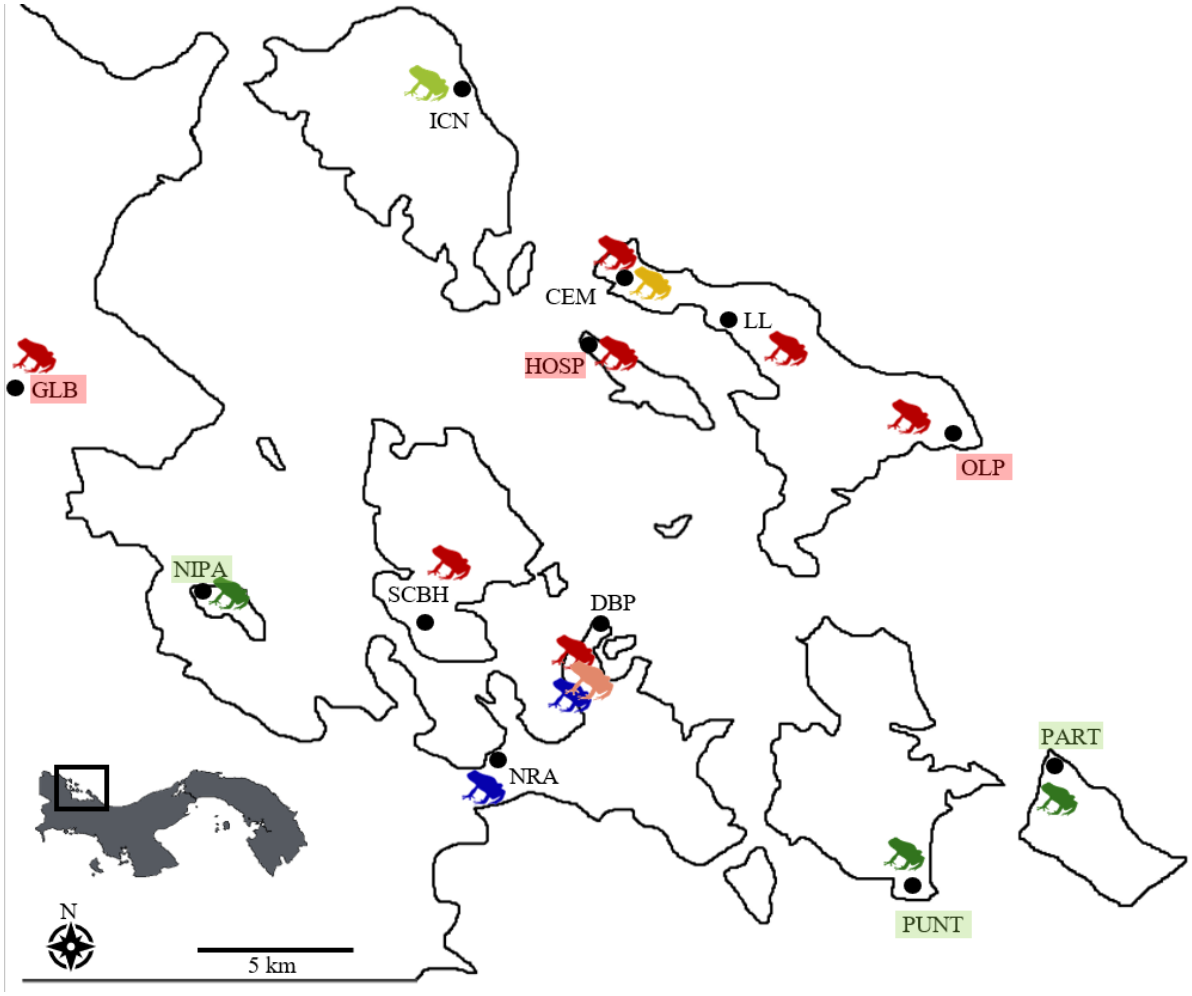


Figure 2-1 Map of Collection Locations

Map of *O. pumilio* sampling locations (black dots) in Bocas del Toro, Panama. The color of the frog icon indicates the color morph(s) present at that location. CEM and DBP are both polymorphic populations. Population names of locations that were used in the comparison of red and green frogs are highlighted in red and green, respectively.

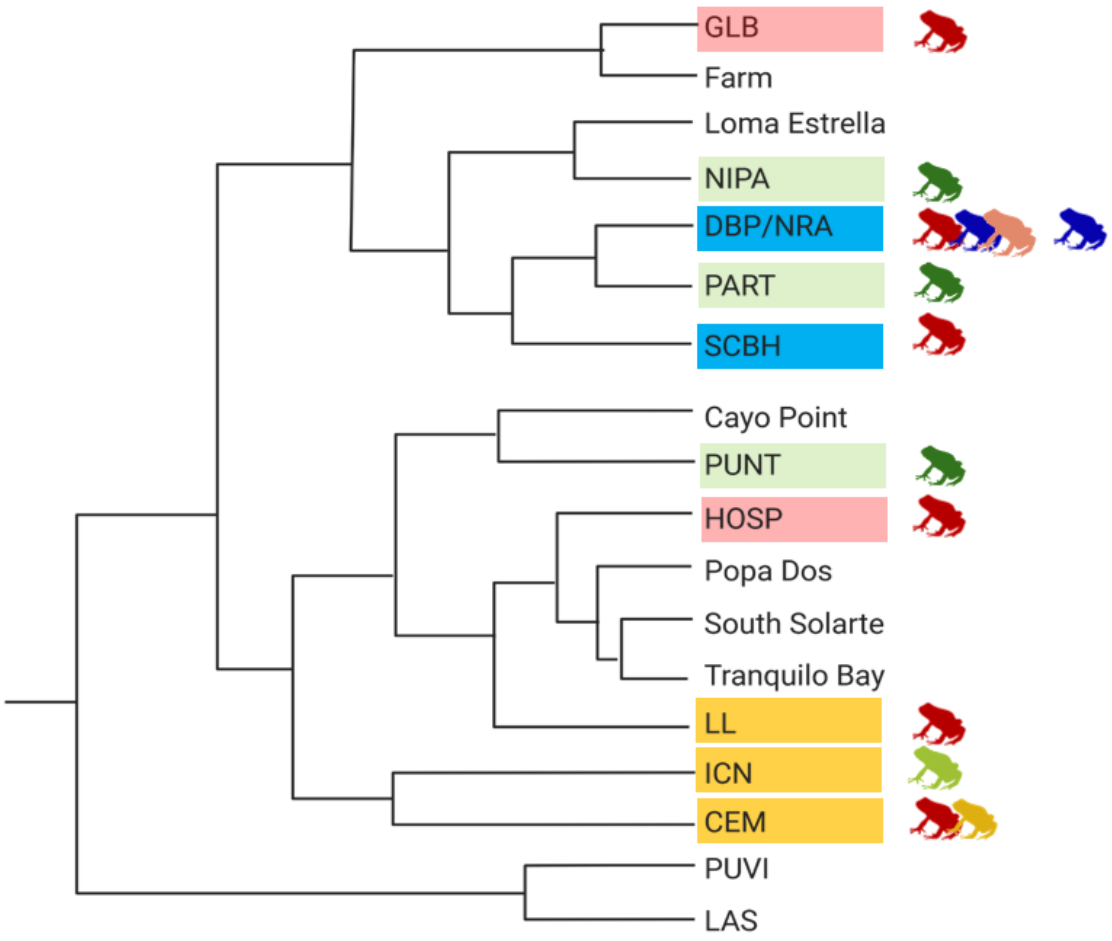


Figure 2-2 Phylogeny of Morphs

Phylogenetic relationships among the sampled *O. pumilio* morphs. The populations used in this study are denoted by frog icons with the color morph(s) present at that location. Because they were so geographically proximate, frogs from monomorphic blue and polymorphic (red and blue) localities were combined when the tree was generated. Populations included in the green/red comparison are highlighted in red or green, respectively. Populations included in the DBP polymorphic/monomorphic comparison are highlighted in blue. Populations included in the CEM polymorphic/monomorphic comparison are highlighted in orange. Phylogeny reproduced from Freeborn (2020) and created with BioRender.com.

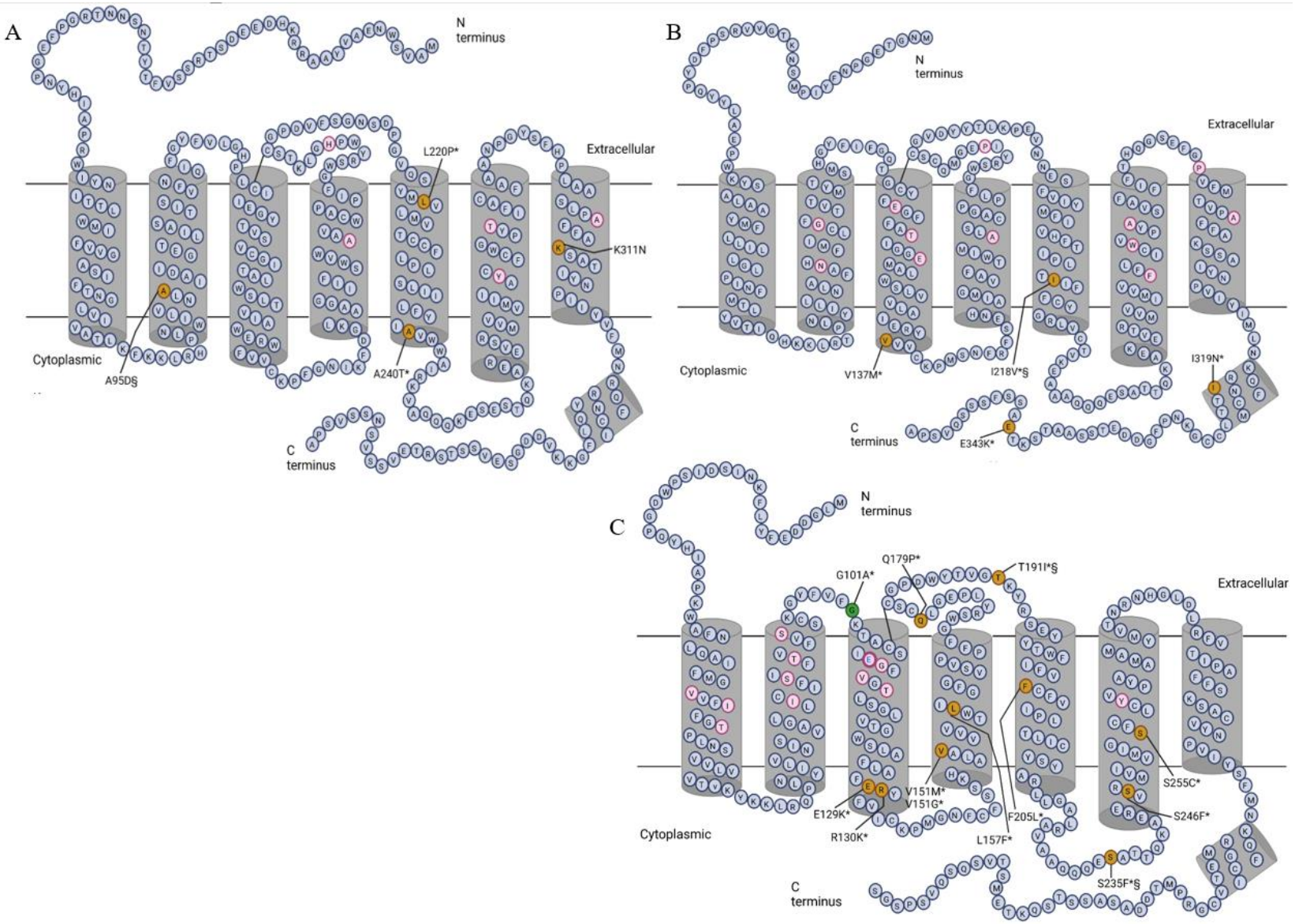


Figure 2-3 Sequencing Results

Diagrams of *O. pumilio* opsin proteins. A) is LWS, B) is RH1, C) is SWS1. Pink residues indicate sites previously identified as key sites for spectral tuning in other taxa (Yokoyama and Radlwimmer 1998). Orange residues are where a missense mutation was found in this study (with the mutation listed in the text next to the residue and numbering based on *O. pumilio* residue numbers). Blue text for the amino acid residue and a thicker border indicates a nonsense mutation. The § symbol indicates that the mutation was found in multiple frogs in this study. The * indicates that the frog(s) was/were heterozygous for the mutation. The one mutation found only in GoTaqGreen is marked with a green circle. Figures were created with BioRender.com.

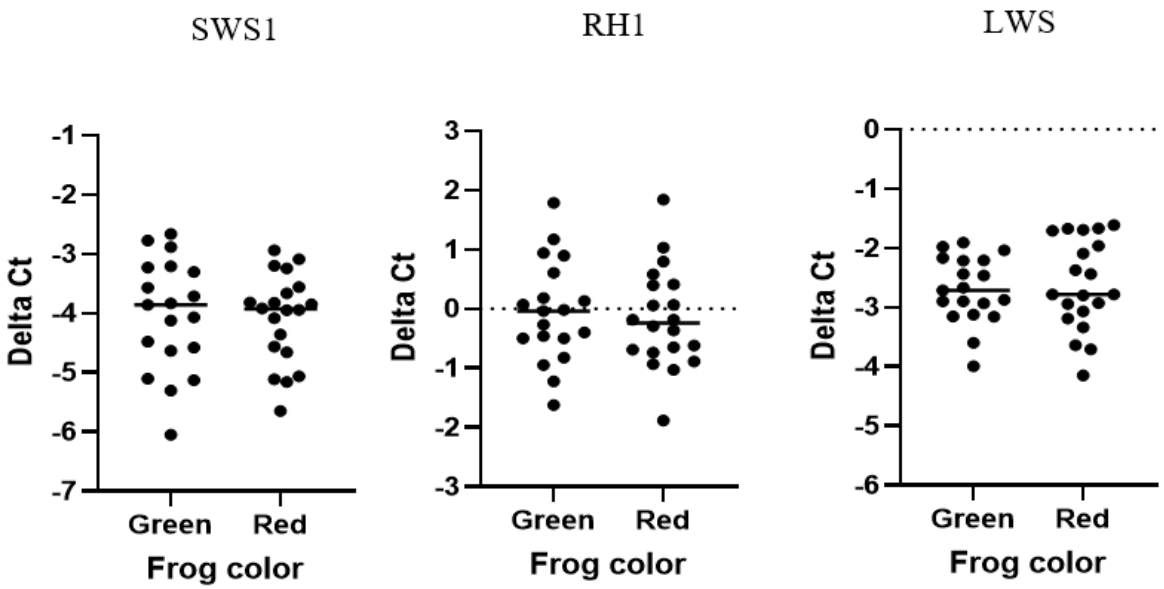


Figure 2-4 Opsin Expression Results-Red/Green Comparison

Opsin expression relative to PDE6g reference gene by frog color for the populations included in the red/green frog color comparison. Delta Ct was calculated as the reference gene Ct – opsin Ct. The dots indicate the delta Ct value of a single individual and the bar is the mean of the group. The dotted line is the x-axis for graphs where 0 is present and is where Ct values of the reference gene and opsin gene are equal. There were no significant differences in expression between green and red frogs for any opsin gene (Linear model, $F_{4,24} \leq 1.88$, $p \geq 0.15$.

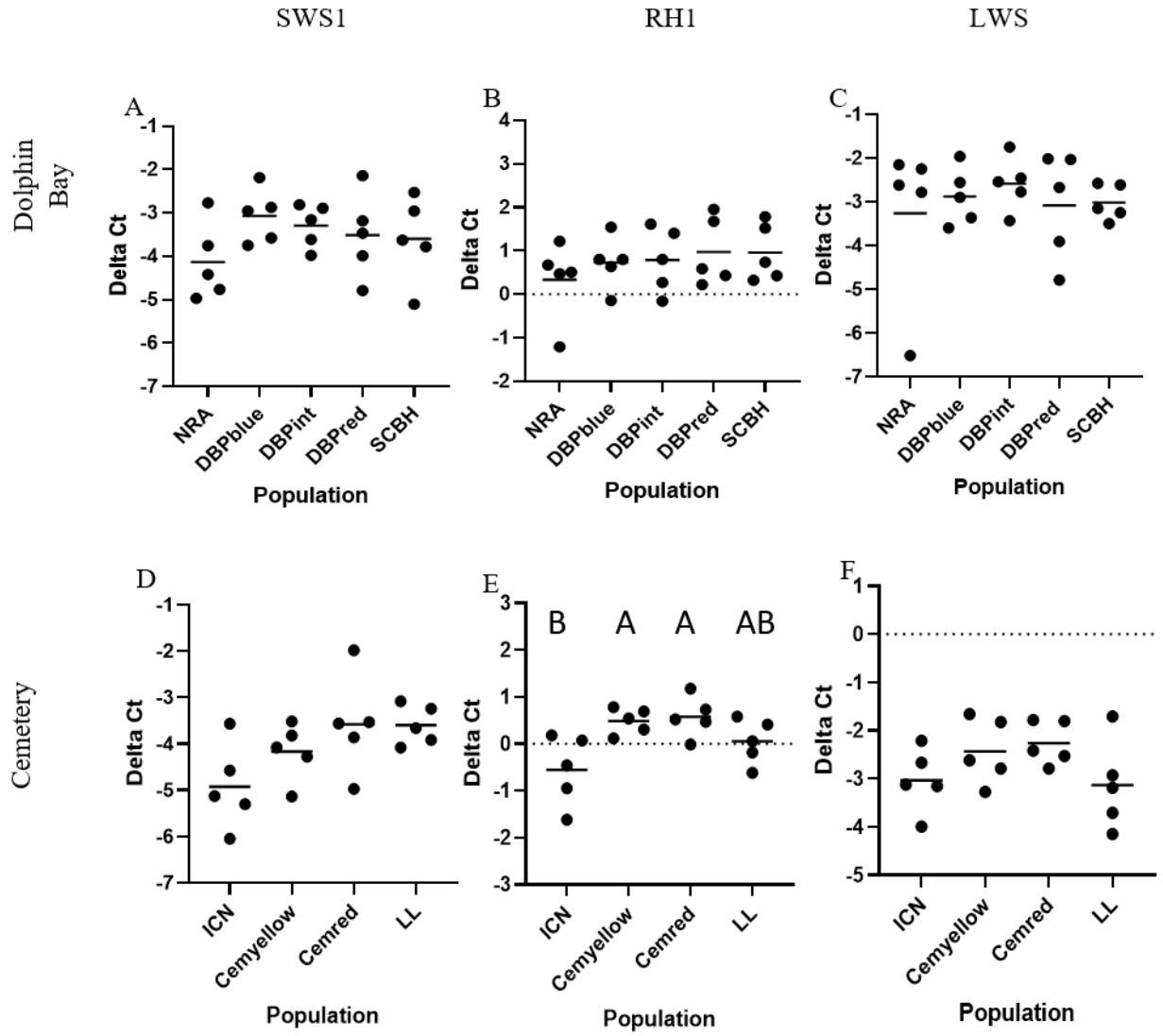


Figure 2-5 Opsin Expression Results- Polymorphic Populations

Opsin expression by morph for the polymorphic populations and their nearest neighbors. Panels A-C show the data for the Dolphin Bay polymorphic morphs (DBPblue, DBPint, DBPred) and nearby monomorphic populations (NRA (blue) and SCBH (red)) for each opsin (SWS1, RH1, LWS) and panels D-F show the data for the Cemetery polymorphic population (Cemyellow, Cemred) and the nearby monomorphic populations (ICN (green) and LL (red)). Letters at the top

of panel E indicate groups with significant differences in opsin expression. Delta Ct was calculated as the reference gene Ct – opsin Ct. The dots indicate the delta Ct value of a single individual and the bar is the mean of the group. The dotted line is the x-axis for graphs where 0 is present and is where Ct values of the reference gene and opsin gene are equal.

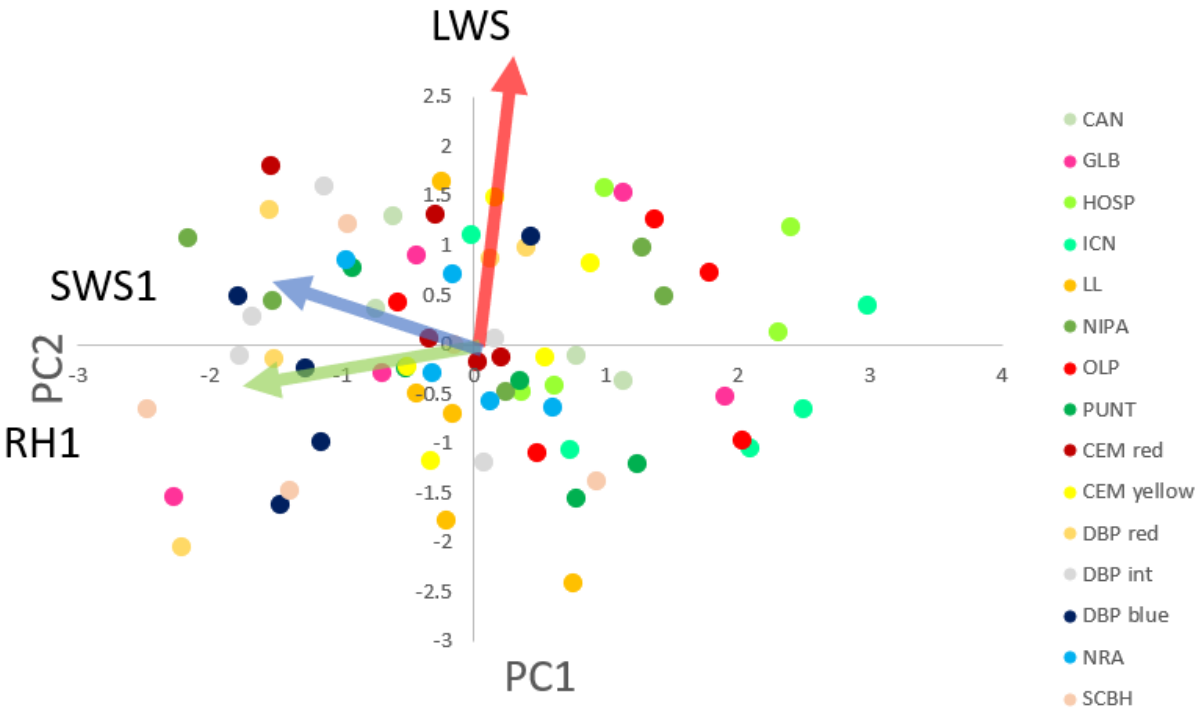


Figure 2-6 Opsin Expression Profile for Complete Data Set

Principal compents 1 and 2 for overall opsin expression profiles by population/polymorphic morphs. Each population/polymorphic morph is represented by a different colored dot that is in the same color category as the frog color (i.e. a red frog is represented by a shade of red, but this does not indicate the shade of red of the frog). Arrows indicate the loadings of the three opsin genes.

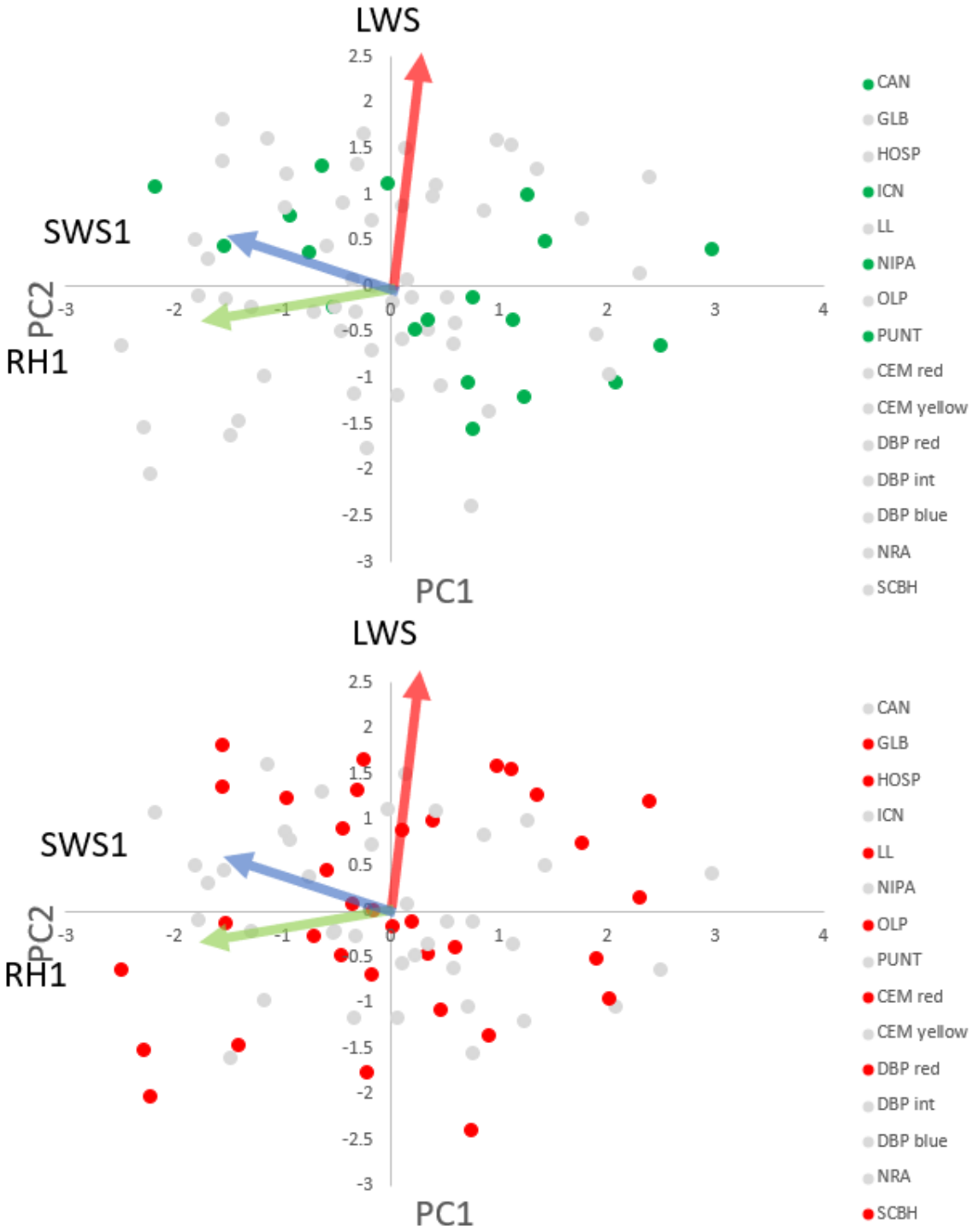


Figure 2-7 Opsin Expression Profile for Green and Red frogs

Principal compents 1 and 2 for overall opsin expression profiles in the green (A) and red (B) for the populations included in the green/red comparison. Arrows indicate the loadings of the three opsin genes.

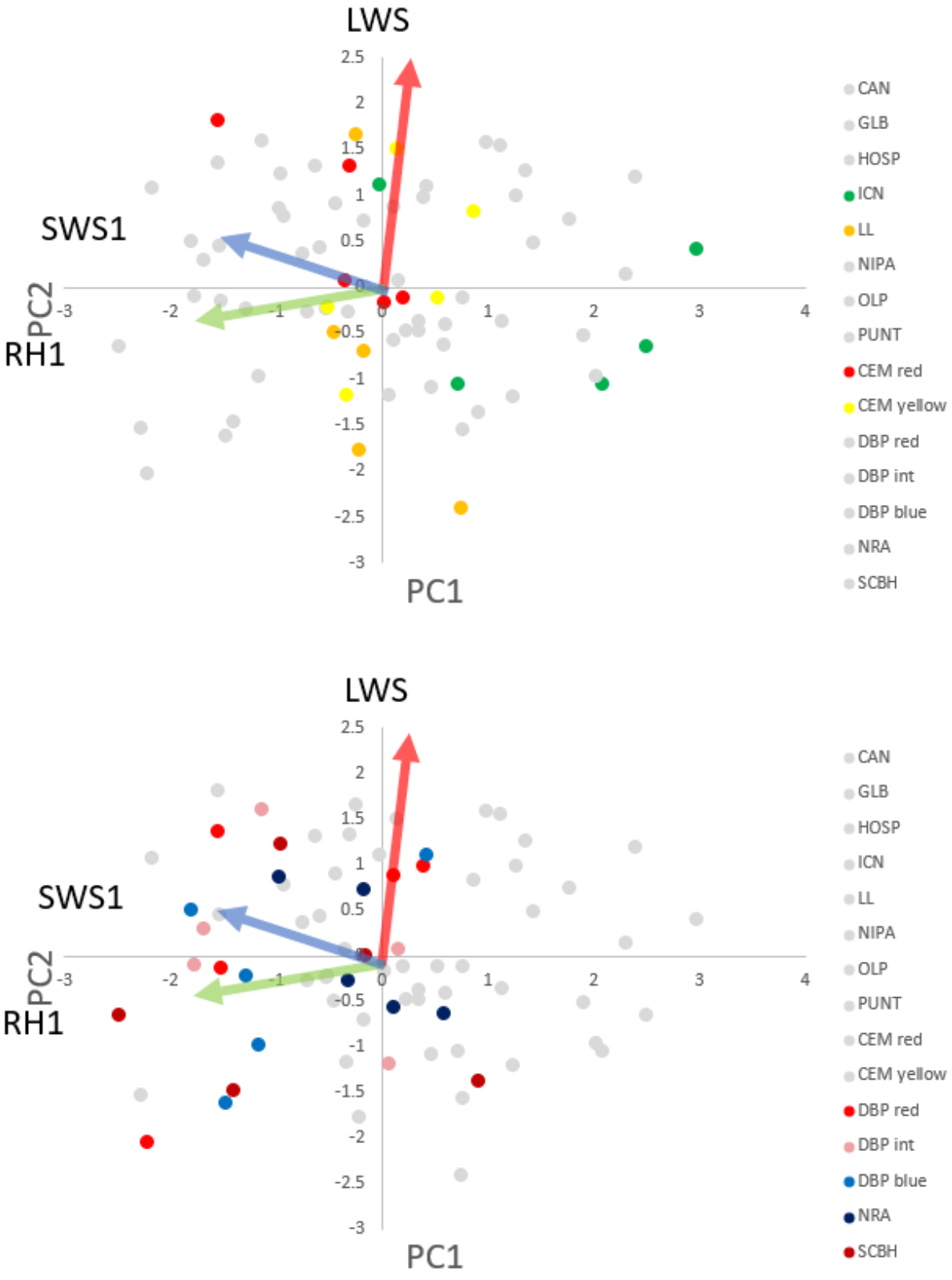


Figure 2-8 Opsin Expression Profile for CEM and DBP Polymorphic Populations

Principal compents 1 and 2 for overall opsin expression profiles for the poylmorphic and monomorphic compulations included in the Cemetery (A) and Dolphin Bay (B) comparisons. Arrows indicate the loadings of the three opsin genes.

Table 2-1 qPCR Primers

O. pumilio opsin primers used for qPCR reactions in this study.

Gene	Primer	Primer sequence	Primer length	Primer start (5'end)	Product size	Location of exon-exon boundary spanned
LWS	LWS-forward	5’AGATTTTTGGCTACTTCGTC	20	394	76	462
	LWS-reverse	5’GTAATGCCACAAACTGAAAC	20	469		
RH1	RH1-forward	5’CTCCACTCTTTGGATGGTC	19	536	80	556
	RH1-reverse	5’CGGCTTCAGGGTATAGTAGTC	21	615		
SWS1	SWS1-forward	5’CATGGGAAACTTTTGCTTCA	20	455	116	559
	SWS1-reverse	5’CTGGCAAATACCTGCTCCA	19	570		
PDE6g	PDE6g-forward	5’GCAGACCAGGCAGTTCAAG	19	316	103	403
	PDE6g-reverse	5’GATGACGGTGATGTCTGTTCC	21	418		

Table 2-2 Sequencing Primers

O. pumilio opsin primer sets used in sequencing.

Gene	Primer	Primer sequence	Primer length	Primer start (5'end)	Product size
LWS 1st half	LWS-forward	5’CCAAAGCTAGATCAAAGAAATAGGA	25	3	678
	LWS-reverse	5’ACTGAACACATCCGGACCAC	20	680	
LWS 2nd half	LWS-forward	5’GGAAAATTGGCTGCTGGTG	19	549	651
	LWS-reverse	5’TGTGCAGACCATAGGAAGAAG	21	1199	
RH1 1st half	RH1-forward	5’CCTTTTAAGAGCCGCCACTA	20	9	701
	RH1-reverse	5’ACACCAGTCGGCCATAGCA	19	709	
RH1 2nd half	RH1-forward	5’CTCCACTCTTTGGATGGTC	19	536	815
	RH1-reverse	5’CTACGGAGCAGTGTCGCATC	20	1350	
SWS1 1st half	SWS1-forward	5’AGTCCGACGATCGTGAAAAC	20	11	569
	SWS1-reverse	5’GTAAACCCTCTGGCAAATACCT	22	579	
SWS1 2nd half	SWS1-forward	5’CACCTGGCTCATTGGTTTC	19	506	729
	SWS1-reverse	5’TTGGGCGAGGTAGAACATTG	20	1234	

Table 2-3 Summary of Mutations

Summary of amino acid residue mutations found. Mutations are given with the numbering according to *O. pumilio* opsin proteins. Sample names are the same abbreviations used for sampling locations in Fig. 1. Amino acid properties are coded as: N = nonpolar, P = polar, + = positive charge, - = negative charge. Additional variants, if any, found at these sites in other species are listed, along with the species used and their GenBank accession numbers.

Gene	Mutation	Number of frogs with mutation	Sample found in	Amino acid properties of change	Amino acid variants at residue	Other species used	GenBank Accession Numbers
LWS	A95D	2	DBP08 DBP11	P→ X	invariable	*Nanorana parkeri, Ambystoma tigrinum, Cynops pyrrhogaster, Xenopus laevis, X. tropicalis, Rana catesbeiana*, human red, human green	XP_018416216.1 AAC96070.1 BAB55453.1 NP_001096331.1 NP_001084114.1 PIO33359.1 NP_064445.2 NP_001041646.1
	L220P	1	CAN04	N →N	M, I, L		
	A240T	1	CEM10	N→ P	H, Q		
	K311N	1	DBP04	+ → P	invariable		
RH1	V137M	1	ICN03	N → N	L, I	*Xenopus laevis, X. tropicalis, Nanorana parkeri, Ornithorhynchus anatinus, Crocodylus porosus, Tachyglossus aculeatus, Gavialis*	NP_001080517.1 NP_001090803.1 XP_018410729.1 NP_001121099.1 XP_019392098.1 AFO70161.1 XP_019360201.1 XP_005426698.1
	I218V	3	CAN05 GLB04 LL02	N → N	I, V		
	I319N	1	DBP08	N → P	I, L		

	E343K	1	CEM09	- → +	invariable	*gangeticus, Geospiza fortis*	
SWS1	G101A	1	CEM04	N → N	invariable	*Xenopus laevis, X. tropicalis, Rana catesbeiana, Nanorana parkeri, Haliaeetus leucocephalus, Picoides pubescens, Mus pahari, Condylura cristata, Caunus lupis familiaris, Uta stansburiana*	NP_001079121.1 NP_001119548.1 PIO33958.1 XP_018416245.1 XP_010567082.1 XP_009898521.1 XP_021045511.1 XP_004677073.1 XP_539386.2 AAZ79909.1
	E108Stop	1	CEM03	- → stop	E, D		
	E129K	1	CEM03	- → +	invariable		
	R130K	1	CEM03	+ → +	invariable		
	V151 M/G	2	CEM04 DBP10	N → N	invariable		
	L157F	1	ICN02	N → N	invariable		
	Q179P	1	DBP10	P → N	invariable		
	T191I	2	DBP11 DPB15	P → N	invariable		
	F205L	1	NIPA02	N → N	invariable		
	S235F	1	NIPA03	P → N	invariable		
	S255C	1	DBP15	P → P	invariable		

Table 2-4 PCA Summary

PCA coordinates for each opsin gene.

	PC1	PC2	PC3
LWS	0.132226	0.979358	0.152885
RH1	-0.88142	-0.0875526	0.464154
SWS1	-0.8583375	0.240775	-0.45308

3.0 Estimating the heritability and inheritance patterns of color in a polymorphic frog

3.1 Summary

Understanding the evolution of a trait requires understanding the genetic architecture including the number of genes involved and their interaction. The narrow-sense heritability (h^2) can be used to estimate the contribution of additive genetic variation to phenotype and as an estimate of the potential response to selection. In this study, I test for differences in heritability and inheritance of color in three morphs of a color polymorphic frog, *Oophaga pumilio*. Color in this species is hypothesized to be under both natural and sexual selection and has both genetic and environmentally based components. Additionally, the proportion of phenotype produced by the different pigment types differs among morphs, suggesting heritability may differ among morphs. I found the red morph to have the highest heritability and the green morph the lowest, though confidence intervals overlapped for all estimates. This suggests that different populations might have different responses to selection on coloration. I also found evidence for a simple dominant/recessive relationship controlling dorsal color between one pair of morphs where green is dominant to blue, whereas the other two pairs showed evidence for an additive genetic relationship. This study provides evidence for the potential for further evolution of differentiation in color, a complex trait used in a variety of contexts, between morphs of the same species and contributes to our understanding of color inheritance in this species.

3.2 Introduction

Color in animals has evolved to serve multiple functions. In many species, color has evolved for camouflage, allowing the animal to blend in with its surroundings and hide from predators or prey. Animals that live in the arctic such as snowshoe hares, arctic foxes, and rock ptarmigans have evolved the ability to change their coloration seasonally to match the predominant color of their environment; white when snow is present and brown when it is not (rev. in Zimova *et al.* 2018). At the other extreme, bright coloration has evolved as a warning signal of toxicity, alerting potential predators of an animal's unpalatability. Such aposematic signals are designed to be highly contrasting with the background environment and conspicuous to potential predators. Bright colors are often paired with black elements to enhance visibility of the aposematic signal, such as the red or yellow bands on the black wings of *Heliconius* butterflies (Finkbeiner *et al.* 2014) and the black bands accompanying red and yellow bands on coral snakes (Banci *et al.* 2020). Similar signals have also evolved in some palatable species to mimic the warning coloration of a sympatric toxic species and take advantage of the fact that predators have learned to avoid that signal. For example, there are a number of non-venomous snakes that have a similar banding pattern to venomous coral snakes (rev. in Pfennig and Mullen 2010). Similarly, several species of palatable swallowtail butterflies have evolved to mimic the wing color pattern of toxic sympatric butterfly species (rev. in Kunte 2009). However, color is not only used to avoid predators. Color can serve as a sexual signal that is used by animals in attracting and choosing mates as well. A good example is the orange coloration of male guppies, which is attractive to female guppies (Brooks and Endler 2001). Because animal coloration can be used in many contexts, understanding the mechanisms that generate color differences within species can help us to understand the role that selection plays in shaping animal color signals.

Gaining a complete understanding of the evolution of a quantitative trait such as color requires knowledge of the trait's genetic architecture. This includes knowing how many genes are involved in producing the trait, their location in the genome, the dominance patterns among alleles, and any epistatic or pleiotropic interactions (rev. in Mackay 2001). Dominance effects can shield recessive alleles from selection, particularly at low frequencies, and selection on genes with epistatic effects can strengthen selection on additive genetic variation, often resulting in reduced genetic variation in a trait (rev. in Hansen 2006). If multiple genes affecting a trait are tightly linked, they will tend to be inherited together, and thus selection will act on the two loci at once (rev. in Nosil *et al.* 2009). Thus, some knowledge of the genetic architecture of a trait is critical to understanding how the trait has and might continue to evolve. Studies often attempt to characterize the genetic architecture of a trait by partitioning the total variation in phenotype into multiple components, namely variation due to additive genetic variance (V_A), dominance effects (V_D) and epistatic effects (V_I) and the rest is attributable to variation due to environment (V_E) (rev. in Huang and Mackay 2016). Approaches typically used to determine the genetic architecture of complex traits like color in model organisms include mutagenesis studies and QTL mapping (rev. in Mackay 2001). However, for non-model organisms, the genetic tools needed to quantify and pinpoint loci contributing to a trait and determine gene interactions are often lacking. In these cases, controlled breeding studies offer an alternative approach for testing hypotheses about the evolution of animal color signals.

Controlled breeding studies provide a way to estimate color heritability, or how much of the variability in a color phenotype is due to genetic variation, and test hypotheses about different modes of inheritance (Visscher *et al*. 2008). The narrow-sense heritability (h^2) of a trait, is calculated as the amount of additive genetic variance (V_A) in the trait divided by the total

phenotypic variance (V_P) (de Villemereuil 2012). V_A can be estimated using measurements of trait values associated with a pedigree and requires no prior knowledge of the genes underlying the trait, which makes it a tractable approach non-model species. Because heritability is a population-level estimate that depends on the amount of additive genetic variation for a trait in a population, it can differ both within and among species (Visscher *et al*. 2008). In the case of color, for example, different populations might use different pigment and/or cell types to produce color. Reliance on certain pigment types, like carotenoids which are dietarily derived (rev. in Weaver *et al.* 2018), may result in lower heritability as a greater proportion of the variation in phenotype is attributable to non-genetic causes. Since selection requires heritable variation in a trait for evolution to occur, the response to selection can vary among populations if trait variants have different heritabilities. Selection can produce larger evolutionary changes in a trait per generation for populations in which that trait has high heritability. For example, if the orange color, mentioned above, that females prefer in male guppies has high heritability and differences between males are therefore primarily due to genotypic variation, females acting on their preference for orange can drive the evolution of male coloration toward more orange. In contrast, if heritability of the trait is low and diet is largely responsible for variation in orange coloration in males then the selection by females of the most orange males would not produce much change in the overall trait mean for the population. Thus, the heritability of a trait is indicative of how effective selection on a trait can be shaping the evolution of a trait.

Oophaga pumilio is a terrestrial poison frog species that is found along the Caribbean coast of Central America from southern Nicaragua to north-western Panama. Throughout most of its range, it is a red frog with blue to black limbs (Wang and Summers 2010). However, in the Bocas del Toro archipelago of north-western Panama, this species exhibits a diversity of coloration that

spans the visible spectrum, with most color morphs found on separate islands (Myers and Daly 1976, Wang and Summers 2010, Rudh *et al.* 2010). Some morphs display bright aposematic coloration, while others (despite being toxic) have cryptic coloration (Rudh *et al.* 2010). Recent work has focused on understanding the evolutionary forces acting on color in this species as a case study for understanding the evolution of divergence in animal color signals. Evidence from clay model studies suggests that color functions as an aposematic signal in the brightly colored morphs (Saporito *et al.* 2007, Hegna 2012, Richards-Zawacki *et al.* 2013). However, evidence suggests that natural selection is not currently a strong driver of divergence in color (Yeager 2015). There is also strong evidence that sexual selection is also acting on color. Female *O. pumilio* generally prefer to court with males of the local color (Reynolds and Fitzpatrick 2007, Maan and Cummings 2008, Richards-Zawacki and Cummings 2010, Maan and Cummings 2012) and territorial males are more aggressive to encroaching males that are the local color (Yang *et al.* 2016). Moreover, morphs in two polymorphic locations show color-based biases in both female choice (Yang *et al.* 2016, Richards-Zawacki *et al.* 2012) and male aggression (Yang *et al.* 2018), suggesting that ongoing sexual selection on color is occurring.

Although recent work has shed light on the evolutionary forces acting on color in *O. pumilio*, we still know very little about the genetic architecture of color in this species. Evidence from one pair of color morphs found in sympatry on the northern tip of Bastimentos Island suggests a simple Mendelian pattern of inheritance (with red being dominant to yellow, Richards-Zawacki *et al.* 2012). However, in another polymorphic population, the presence of red, blue, and a diversity of intermediately colored (brownish) individuals (Yang *et al.* 2019) suggests a more complex genetic architecture. Furthermore, intermediate coloration seen in F1 hybrids generated through captive breeding of several color morphs (Dugas and Richards-Zawacki 2015) suggests that color

often shows incomplete dominance or more complex inheritance in this species. Differences in coloration among *O. pumilio* populations can be attributed to variation in the contents and structure of the dermal chromatophore unit, which is comprised of structural elements such as iridiphores, chromatophores that house pigments the animal itself produces (e.g., melanophores), and chromatophores that house dietarily derived carotenoids (Freeborn 2020). Attempts to date to elucidate the genetic basis of color variation in this species have been hindered its large and incompletely- assembled genome, which is made up in large part by repetitive elements (Rogers *et al.* 2018). Using a genetic linkage map approach, several color-associated SNPs have been identified. However, due to the aforementioned issues with the frog's genome, these SNPs could not be mapped with confidence to particular genes or gene regions (Freeborn 2020).

Given the limited genomic resources available for *O. pumilio*, I felt that a breeding study, resulting in color phenotype measurements from frogs with a known pedigree, held the most promise for improving our understanding of the genetic architecture of color and potential for adaptive evolution in this species. To this end, I set up a controlled breeding experiment using three color morphs; one with aposematic (red-orange) coloration and two with cryptic (green, blue) coloration, each of which differ in which pigments and cell types are combined used to produce color (Freeborn 2020). Within this study design I tested several hypotheses about dorsal color inheritance in *O. pumilio*, focusing on the coloration of the dorsal surface because of its relevance in the contexts of both natural and sexual selection (Myers and Daly 1976, Wang and Summers 2010, Rudh *et al.* 2010, Reynolds and Fitzpatrick 2007, Maan and Cummings 2008, Richards-Zawacki and Cummings 2010, Maan and Cummings 2012, Yang *et al.* 2016, Yang *et al*. 2018). First, I tested the alternative hypotheses that color shows (1) simple Mendelian genetics, (2) incomplete dominance, or (3) is produced by more additive genetic variation. I predicted that color

would follow an additive model of inheritance due to the fact that in a smaller study, F1 crosses between purebred individuals of the three color morphs appeared to be intermediate in color between the two parental morphs (Dugas and Richards-Zawacki 2015). Next, I hypothesized that heritability of dorsal color differs among morphs and predicted that heritability would be different for red (aposematic) frogs compared to green and blue (cryptic) frogs as these color groups have likely been shaped by different forms of natural selection. Alternatively, red and green phenotypes might have lower heritability than blue because red and green coloration is imparted in large part by carotenoids, which are accumulated and/or synthesized from precursors that come from the frogs' natural diets. Previous work has shown that carotenoids comprise a larger proportion of the dermal chromatophore unit in red and green compared to blue *O. pumilio* populations (Freeborn 2020). Thus, there might be a stronger effect of environment on phenotype for red and green frogs than for blue frogs whose color is produced in large part by genetically derived pigments (e.g., melanins) and/or structural elements (e.g., iridophores) in the dermis. The results of this study will contribute to our understanding of how effectively selection can act on color variation in a species that is vast becoming a model for understanding how animal color variation evolves.

3.3 Methods

3.3.1 Breeding Design

Frogs were bred in captivity in a colony kept first at the Smithsonian Tropical Research Station's Bocas del Toro Research Station in Panama, then at Tulane University, and later at the University of Pittsburgh. This colony was founded by frogs collected from the wild in 2008. I used

frogs from several different color morphs of *O. pumilio* originating from the islands of Bastimentos (monomorphic red population) and Popa (monomorphic green population) as well as from the Aguacate peninsula (monomorphic blue population) of mainland Panama. Frogs were kept in breeding pairs consisting of one male and one female in 37 x 22 x25 cm plastic terraria with four water-filled PVC tubes (one in each corner) for tadpole deposition and a plant for egg deposition. Frogs were fed springtails and fruit flies. The fruit flies were gut loaded with carotenoids, which has been found to improve reproductive success (Dugas and Richards-Zawacki 2013). All other animal husbandry details for the colony can be found in Dugas and Richards-Zawacki (2015). To study heritability and inheritance of color, I set up several different types of crosses. To estimate heritability of dorsal color in each of our color morphs, I used breeding pairs that consisted of male and female frogs that were the same morph. I also examined the inheritance patterns of color using breeding pairs consisting of males and females of different color morphs. The offspring of these crosses were also paired to one of the parental phenotypes to create backcrosses. Due to differences in the number of animals of each sex available for the different morphs, not all crosses and backcrosses were reciprocal.

3.3.2 Color Quantification

I took dorsal and ventral photographs of each adult frog that founded the breeding colony and of each offspring once the frog had reached at least 90 days post-metamorphosis (the time point by which all frogs in a pilot study had reached their adult coloration). Frogs were initially photographed on graph paper but from 2017 on were photographed on a standard (18%) grey card alongside a color standard with a digital camera mounted on a camera stand directly above the frog.

I analyzed frog color from photographs by extracting red (R), green (G), and blue (B) values in the RGB color system from a sample of pixels using ImageJ version 1.50i (Schindelin *et al.* 2012). I measured the average RGB values for five samples along the dorsal surface of each frog's body: one on the head, one in the shoulder region, one on each side of the body, and one near the venter (Figure 3-3). Photographs analyzed prior to 2017 were measured using samples that were 20 x 20 pixels in size. Photographs analyzed later used samples that were scaled in size to the size of the frog such that the length of one side of the sampling area box was 1/3 of the distance between the frog's eyes. This was done to ensure that each sampling area was covering a similar proportion of the frog's dorsal skin surface. I cropped each sample square as a separate file and extracted the average RGB values from it using a custom macro (see supplement) and the RGBMeasure tool implemented in ImageJ (supplement). The result was mean R, G, and B values (which range from 0 to 256 where a value of 0, 0, 0 is black and a value of 256, 256, 256 is white) for each the sampled body areas. I quality checked the data using histograms to identify outliers. Outliers were only removed if they were in a histogram bin by themselves and two or more bins away from the next bin that contained data. I averaged RGB values from each of the five sampling regions to obtain a single set of dorsal mean R, G, and B values for each frog. To compare color scores among frogs I used proportional color scores, which I calculated by dividing each color mean (R, G, or B) by the sum of all three (e.g., proportional R score =Mean R/(Mean R + Mean G + Mean B), and analogously for G and B).

3.3.3 Statistical Analysis

3.3.3.1 Mendelian Genetics

To test the hypothesis of simple Mendelian dominant/recessive inheritance for dorsal color, I plotted color PC1 vs. PC2 for each cross using the *ggplot2* package in R (Wickham 2016) and included a 95% confidence ellipse. F1 hybrids were considered to have a different phenotype from the parental morphs if there was no overlap between the 95% confidence ellipses of the F1 offspring and the ellipses for either of the parental phenotypes. To test for incomplete dominance inheritance patterns, I plotted PC1 vs. PC2 for the parental morphs and backcrosses. To test for the expected 1:1 ratio between the parental phenotype and the hybrid phenotype, with both directions of the backcross treated separately, under incomplete dominance, I tallied the individuals from each backcross whose color phenotype fell within vs. outside of the parental 95% confidence ellipses. Points that were on the line of the ellipse were considered in the ellipse. I used these counts in χ^2 tests with the null hypothesis being twice as many offspring with non-parental phenotypes than offspring with phenotypes falling within either parental ellipse for all three cross types.

3.3.3.2 Sex Linkage

I tested the role of parental sex in determining frog color (e.g., whether the phenotype differs when the mother is red and the father is blue or vice versa) using reciprocal crosses between two color morphs. To do this, I ran a univariate animal model (i.e., a linear model that uses pedigree relationships as a random effect, see Wilson *et al.* 2010) in the *MCMCglmm* package in R (Hadfield 2010), with color PC1 as the response variable and including color data from reciprocal backcrosses from the Popa x Bastimentos cross. These backcrosses were the offspring of F1 hybrid

individuals backcrossed to a Popa individual (i.e., both F1 hybrid males x Popa females and vice versa). The parental individuals of the backcrosses were included in the model to account for relatedness among parents (e.g., the presence of siblings) as the animal model requires any animal listed as a parent to be included as an entry in the pedigree. However, color data for the parental morphs was not included in the model. The mother's phenotype was included as a fixed effect and the pedigree, included as a random effect, was fit to a gaussian distribution. I used relatively uninformative priors fit with an inverse gamma distribution and ran the model for 500,000 iterations after a 50,000 burn-in period with a thinning parameter of 20 (de Villemereuil 2012). I confirmed convergence of the model by checking for consistency across the traces created using the function "plot", broken down by fixed effects (Sol) and random effects (VCV). I verified a lack of autocorrelation for both fixed and random effects using the function "autocorr".

3.3.3.3 Sexual Dimorphism

I tested for color differences between the sexes and whether the extent of such color dimorphism differed among morphs using a univariate animal model with color PC1 as the response variable using *MCMCglmm*. I included dorsal color data for all of the pure morph individuals for which the sex was known. Sex, and the interaction between sex and morph were included as fixed effects. I fit the model to a gaussian distribution and used relatively uninformative priors based on an inverse gamma distribution. I ran the model for 500,000 iterations after a 50,000 burn-in period with a thinning parameter of 20. I checked to confirm convergence of the model using the "plot" function for both fixed and random effects verified a lack of autocorrelation using the "autocorr" function for fixed and random effects.

3.3.3.4 Color Heritability

To test for differences in color heritability (narrow sense) between color morphs, I ran a multivariate animal model using *MCMCglmm*. The model included the values of dorsal color PC1 for each of the three morphs as separate response variables and included the pedigree information for each morph as random effects. I used relatively uninformative priors with an inverse gamma distribution that specified equal variance explained by the residuals and fixed effect. I ran the model with a gaussian distribution for 100,000 iterations after a 10,000 burn-in period with a thinning parameter of 10 (de Villemereuil 2012). I checked for convergence of the model using the "plot" function for both fixed and random effects and verified a lack of autocorrelation using the "autocorr" function for fixed and random effects. Heritability was calculated for each morph (each response variable) as the variance component estimate from the pedigree information (random effect) divided by the total variance for each response variable and 95% confidence intervals for the estimate were calculated using the function "HPDinterval" (de Villemereuil 2012).

3.4 Results

3.4.1 PCA

The PCA resulted in eigenvalues > 1 for PC1 (1.79) and PC2 (1.21), which together explained 100% of the variation in color (Table 3-1). The proportion of R was positively associated with PC1 (0.97), while the proportion of B was negatively associated with PC1 (-0.89). The proportion of G was negatively associated with PC2 (-0.98) (Table 3-2). However, since frogs

varied in PC1 value by color, where blue frogs had negative values, green frogs had negative but close to zero values, and red frogs had positive values, only PC1 was used for subsequent analyses.

3.4.2 Mendelian Genetics

In two of my three cross types (Bastimentos x Popa and Bastimentos x Aguacate), the F1 hybrids were clearly distinct from the parental phenotypes in color PC space (Figure 3-4 A and E). However, in the cross between Popa and Aguacate, the F1 frogs had considerable phenotypic overlap with the Popa frogs (Figure 3-4 C). This suggests that I can rule out a simple dominant/recessive relationship for the Bastimentos x Popa and Bastimentos x Aguacate crosses. However, for the Popa x Aguacate cross few if any individuals have phenotypes that fall outside of the parental morph ellipses. In this case, I cannot rule out simple Mendelian dominance of green over blue color, as this ratio appears consistent with the expected 4:0 ratio of the parental colors, though the sample size of F1 individuals is too small to test this statistically.

The backcrosses for each cross type had ellipses that overlapped with the backcrossed parent's ellipse, but the extent of overlap depended on the cross. In the Bastimentos x Popa cross, the Popa backcrosses had much more overlap with the pure Popa frogs than the Bastimentos backcrosses had with the pure Bastimentos frogs (Figure 3-4 B). In the Popa and Aguacate cross, there was considerable overlap between backcrossed and parental phenotype for all of the backcross types (Figure 3-4 D). The Aguacate x Bastimentos cross had the smallest number of backcrossed individuals (seven), three of which had phenotypes that fell within the PC space of pure Bastimentos frogs and four of which fell out in between the PC spaces of the pure morph clusters (Figure 3-4 F). My χ^2 goodness-of-fit tests using the backcross data revealed that

Bastimentos x Popa backcrosses were significantly different from the expected 1:1 ratio for incomplete dominance for both direction of the backcross, with more backcrossed individuals falling in the Popa phenotype space than expected for the Popa backcrosses (χ^2_1 =6.38, n =106, p = 0.012), and more falling outside of the parental ellipse than expected for the Bastimentos backcrosses (χ^2_1 = 16.67, p < 0.001). A similar pattern was found for the Popa x Aguacate cross, with more backcrossed frogs falling into the Popa phenotype space and fewer falling in the intermediate phenotype space than expected for the Popa backcrosses (χ^2_1 = 26.13, n = 30, p < 0.001) and more frogs falling outside of the parental ellipse for the Aguacate backcrosses than in the Aguacate ellipse (χ^2_1 = 7.81, n= 37, p = 0.005). I did not have enough samples (n = 7 backcross frogs) to conduct a statistical test for the Bastimentos and Aguacate cross.

3.4.3 Sex Linkage

There were 72 backcrossed frogs with a Popa frog as the mother, representing 7 sets of full siblings, and 49 backcrossed frogs with Popa as the father, representing 6 sets of full siblings, in the Bastimentos x Popa dataset. I found no significant difference in color between offspring resulting from the two backcross types (animal model, fixed effect of Mother's phenotype: p = 0.409) suggesting that the direction of the cross does not affect color.

3.4.4 Sexual Dimorphism

The main effect of sex was not significant (animal model, p = 0.67) in the model comparing color and sex, suggesting that across all morphs there is no consistent difference in color between male and female frogs. However, when we tested for an interaction between morph and sex, we

found a marginally significant result for the Aguacate morph (animal model, $p = 0.063$), suggesting that the difference in color between males and females was marginally smaller in Aguacate blue frogs than it was in Bastimentos red frogs. Bastimentos females had higher PC1 scores than Bastimentos males, while Aguacate females had lower PC1 scores than Aguacate males, though the difference between sexes was much less pronounced for Aguacate (Figure 3-6). The interaction term for the Popa and Aguacate comparison was not significant (animal model, $p = 0.20$).

3.4.5 Heritability

The multivariate animal model returned moderate heritability estimates for all three morhps. The Popa (green) frogs had the lowest heritability estimate (0.359) and Bastimentos (red) had the highest heritability estimate (0.537) with the Aguacate (blue) frogs falling in between (0.42) (Figure 3-5). Although I did not test for statistical differences between these estimates, the 95% confidence intervals for all three estimates overlap.

3.5 Discussion

My goal was to test hypotheses about the genetic architecture of color in the polymorphic frog *O. pumilio* using a controlled breeding experiment. I was able to reject the null hypothesis of simple Mendelian dominance/recessive inheritance for dorsal coloration in two of the three cross types tested: Bastimentos x Popa, and Bastimentos x Aguacate. F1 individuals from these two cross types had color phenotypes that were distinct (Figure 4 A and E) from both parental morphs. In the cross that didn't result in a distinct phenotype for the F1 frogs, Popa x Aguacate, the colors

of the parental morphs are fairly close to each other on the color spectrum (green vs. blue), so there may simple have been less available color space between them for the F1 phenotype to occupy (Figure 3-4 C). I also had a small sample size (n = 10) of F1 frogs for this cross type, which may also have limited my ability to detect phenotype differences between the F1 crosses and the parental morphs. However, the ratios of the F1 frogs falling within the phenotype space of the two parental morphs for the Popa x Aguacate are consistent with the 4:0 (dominant:recessive) expected ratio for simple Mendelian genetics with Popa (green) being dominant to Aguacate (blue). This finding is further supported by the backcross data which also produced a similar pattern and was significantly different from incomplete dominance expectations but not for simple Mendelian expectations. This would be the second apparent case of simple Mendelian inheritance in this species. The first report was in a pair of red and yellow morphs from a polymorphic population on the northern tip of Bastimentos Island (Richards-Zawacki *et al.* 2012). Additionally, given that the Popa backcrosses fell more in Popa color space for both crosses but more outside of parental color space for Aguacate and Bastimentos suggests that Popa (green) alleles may be dominant to alleles from both other (blue and red) morphs, though this effect is less obvious when Popa is crossed with Bastimentos due to the additive nature of color inheritance for this cross. My finding of what looks to be simple Mendelian dominance of green over blue in one cross type and additive genetic control of color in two other cross types along with the previous finding of simple Mendelian genetics between the two other morphs suggests that the genetic mechanisms that underlie the color differences among populations and morphs vary in their architecture.

I also tested for sex linkage of color inheritance and sexual dimorphism in color. I found no significant effect of cross direction on color, suggesting that color is not sex-linked in the morphs I examined. However, I was only able to test this in a smaller set of crosses for which I

had sufficient numbers of offspring in both cross directions. This included one set of parental backcrosses for the Bastimentos x Popa cross, with a Bastimentos male and a Popa female. Therefore, it is possible that sex-linkage occurs in the other cross types that I was unable test with the available data. In particular, the different directions of the F1s from the Aguacate and Popa cross, while extremely limited, suggest that there may be a difference between these two directions of the F1 cross. Perhaps there are some sex-linked genes involved in color determination for some of the *O. pumilio* morphs, though my data do not permit an adequate test of this for all of the morphs investigated here. There was also no overall effect of sex on color, suggesting that across all three morphs are no consistent differences in color between the sexes. However, I did find a marginally significant interaction between morph and sex for the comparison between Aguacate and Bastimentos. This result suggests that male Bastimentos frogs are on average a bit redder than female Bastimentos frogs and that male Aguacate frogs are on average a bit bluer than female Aguacate frogs (Figure 3-6). This could indicate a degree of sexual dimorphism in *O. pumilio*, potentially due to sexual selection acting on male color, though this result should be interpreted with caution as this effect was not statistically significant.

The narrow sense heritability estimates for dorsal coloration I found in *O. pumilio* are moderate, ranging from 0.359 to 0.537, and comparable to heritability estimates of obtained from other studies on of wild populations. For example, the heritability of sexually selected traits in birds, estimated using a parent-offspring regression, range in narrow-sense heritability values from 0.44 to 0.584 (Hegyi *et al.* 2002, Potti and Canal 2011) and the heritability of several variables relating to egg-shell color in pied flycatchers obtained from a study using the animal model range in value from 0.15 to 0.54 (Morales *et al.* 2010). The heritability of melanistic coloration in Australian magpies, in contrast, is very high ($h^2 = 0.92$ from an animal model, Dobson *et al.* 2019).

These heritability estimates are also similar to heritabilities found for a wider range of morphological traits and study species (h^2 range ~0.25 to 0.85, rev. in Visscher *et al.* 2015). This suggests that the moderate amount of heritability in color that I obtained may be common for complex naturally and/or sexually selected traits, as is often the case for color.

I found that the red (Bastimentos) *O. pumilio* morph had the greatest estimated heritability of the three morphs I tested while the green morph (Popa) had the lowest. While the differences were not great, the fact that the largest difference was between red and green was unexpected given what we know about how these colors are produced in the skin. In *O. pumilio*, red and green dorsal coloration results from similarly structured dermal chromatophore units. In both red and green frogs, a large variety of carotenoid pigments, which are either obtained intact or assembled from precursors in the frog's diet, play an important role in determining the animal's color. In contrast, for blue frogs, color comes mainly from melanin pigments the frogs produce and reflecting platelets, which are a form of structural color (Freeborn 2020). I had predicted that red and green frogs would have the lowest heritability due to both colors being strongly influenced by an animal's diet. However, carotenoid-based red coloration is typically produced by conversion of yellow carotenoid pigments ingested by an animal (Weaver *et al.* 2018). This could mean that red coloration, even though it is dependent upon animals obtaining dietary precursors, requires a greater involvement of genes and hence has a greater additive genetic variance component than green coloration does. The intermediate position of the blue morph is surprising as this color is mostly produced by melanins and reflecting platelets (Freeborn 2020), which presumably are both under genetic control. However, the observed differences in heritability could also have been driven by the different types of selection acting on the colors. Green and blue are considered to be cryptic colors (Rudh *et al.* 2010) and red is thought to be aposematic in *O. pumilio* (Saporito *et al.*

2007). This would suggest that there is the potential for faster evolution in the aposematic, rather than the cryptic colors, which could have been selected for by predator learning. However, given that the confidence intervals obtained in this study overlapped for each of the heritability estimates and the fact that I was unable to test for differences statistically, it is also possible that the variation in heritability I detected among morphs is not biologically relevant.

Overall, I found evidence for differences in the genetic architecture of color between three color morphs of *O. pumilio*. I also found a moderate amount of heritability in color suggesting that a moderate amount of phenotypic variation is explainable by additive genetic variation as opposed to other sources such as dominance effects, epistasis, or the environment. Since the heritability estimates that I found are similar to those found for sexually selected traits in birds (Hegyi *et al.* 2002, Potti and Canal 2011), this suggests that a moderate amount of heritability is common for sexually selected traits and adds support to the commonly held thought that color is under selection in *O. pumilio*. Because the spectacular variation in color we see in this species is heritable, this allows selection to act on these phenotypes. To the extent that the heritability differences we found among morphs are biologically meaningful, selection may be able more finely tune color in certain morphs compared to others.

My finding of moderate heritability of color also suggests that natural and sexual selection could have shaped the variation we see in color today, as has been long hypothesized. However, the frogs in this study were from a captive population and reared in a consistent environment. Future studies using wild animals with a known pedigree would be a good addition to our understanding because they may provide a more accurate measure of the contribution that some of these non-genetic factors make to color variation in the wild. Nevertheless, the results of this study contribute to our understanding of the genetic architecture of this trait, and its evolutionary

potential, as well as add to our understanding of the evolution of complex traits more broadly. This study highlights the complexity of the genetic architecture of coloration, and that different colors of the same species can have different inheritance patterns within a single species. It also suggests that different colors of the same species can have different heritability of color and thus differ in their response to selection. My findings highlight that a controlled breeding approach can provide useful information on the inheritance of complex traits in species that lack the genetic tools for QTL mapping. Additionally, this study provides evidence for the potential for further evolution of differentiation in color, a complex trait used in a variety of contexts, between morphs of the same species. Overall, this study contributes to our understanding of the how intraspecific divergence in animal color signals may evolve.

3.6 Figures and Tables

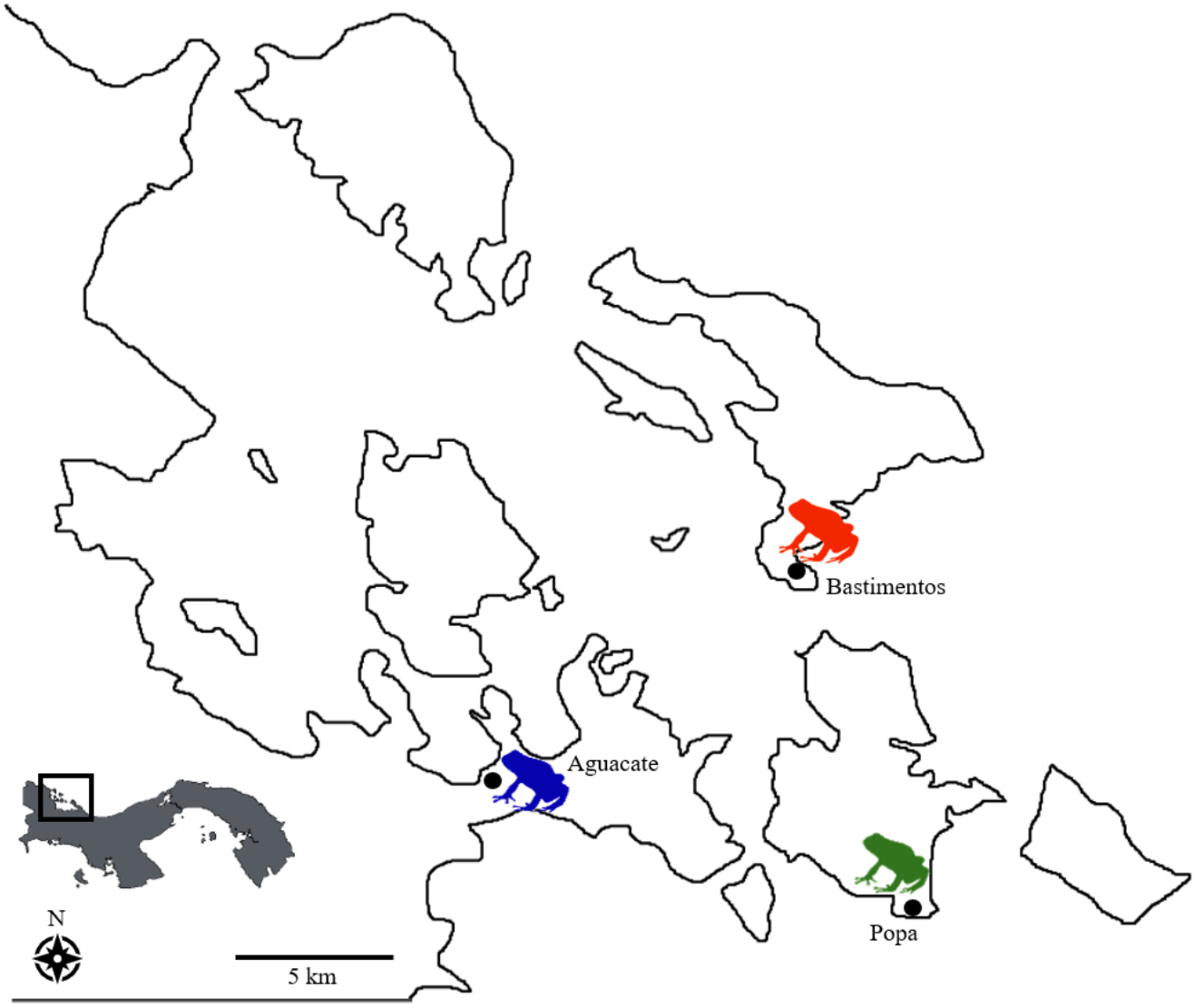


Figure 3-1 Map of Morphs Used

Location of populations from which founding frogs from the colony were collected in Bocas del Toro, Panama. The map of Panama in the bottom left corner indicates the region of focus. The color of the frog icon represents the approximate dorsal color of frogs in the population.

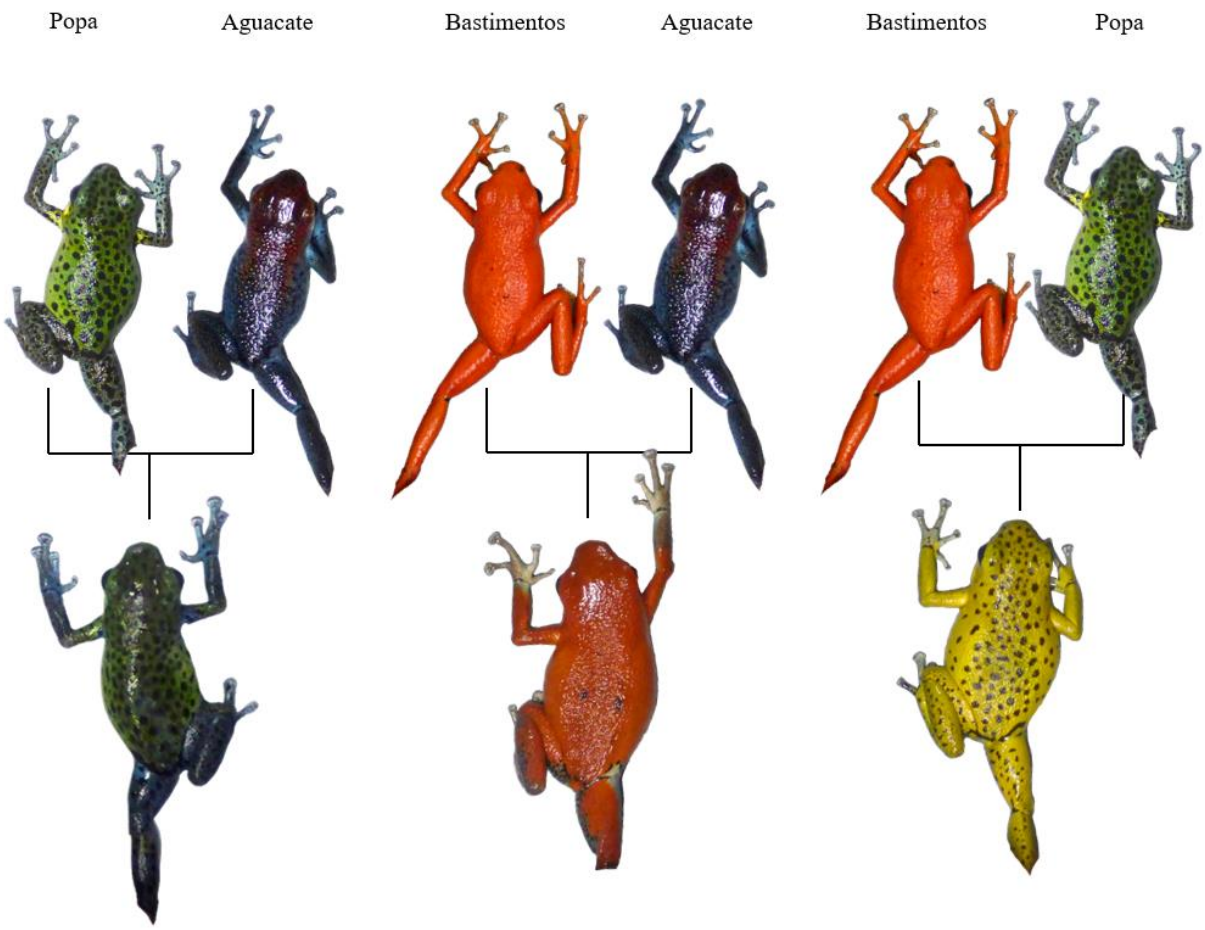


Figure 3-2 F1 Hybrid Phenotypes

Color of F1 individuals produced from all crosses between Bastimentos, Popa, and Aguacate frogs.

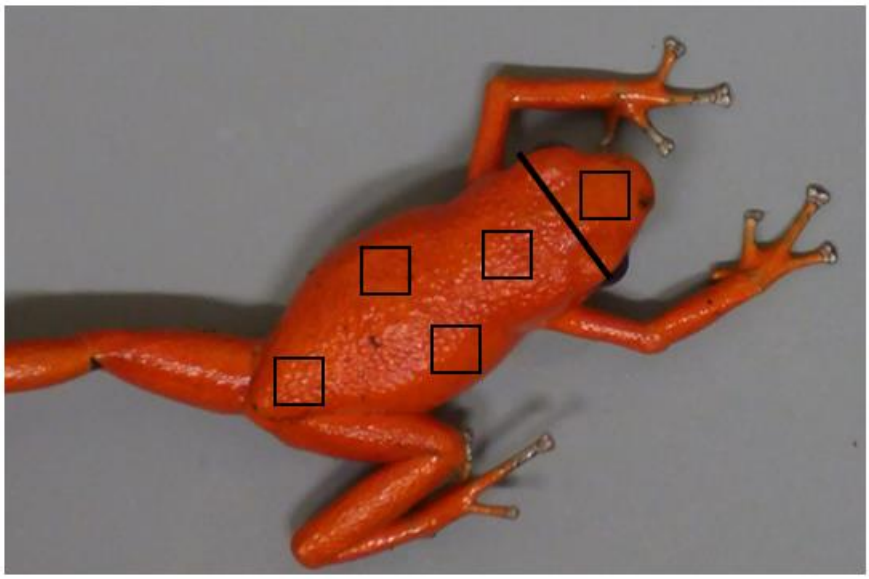
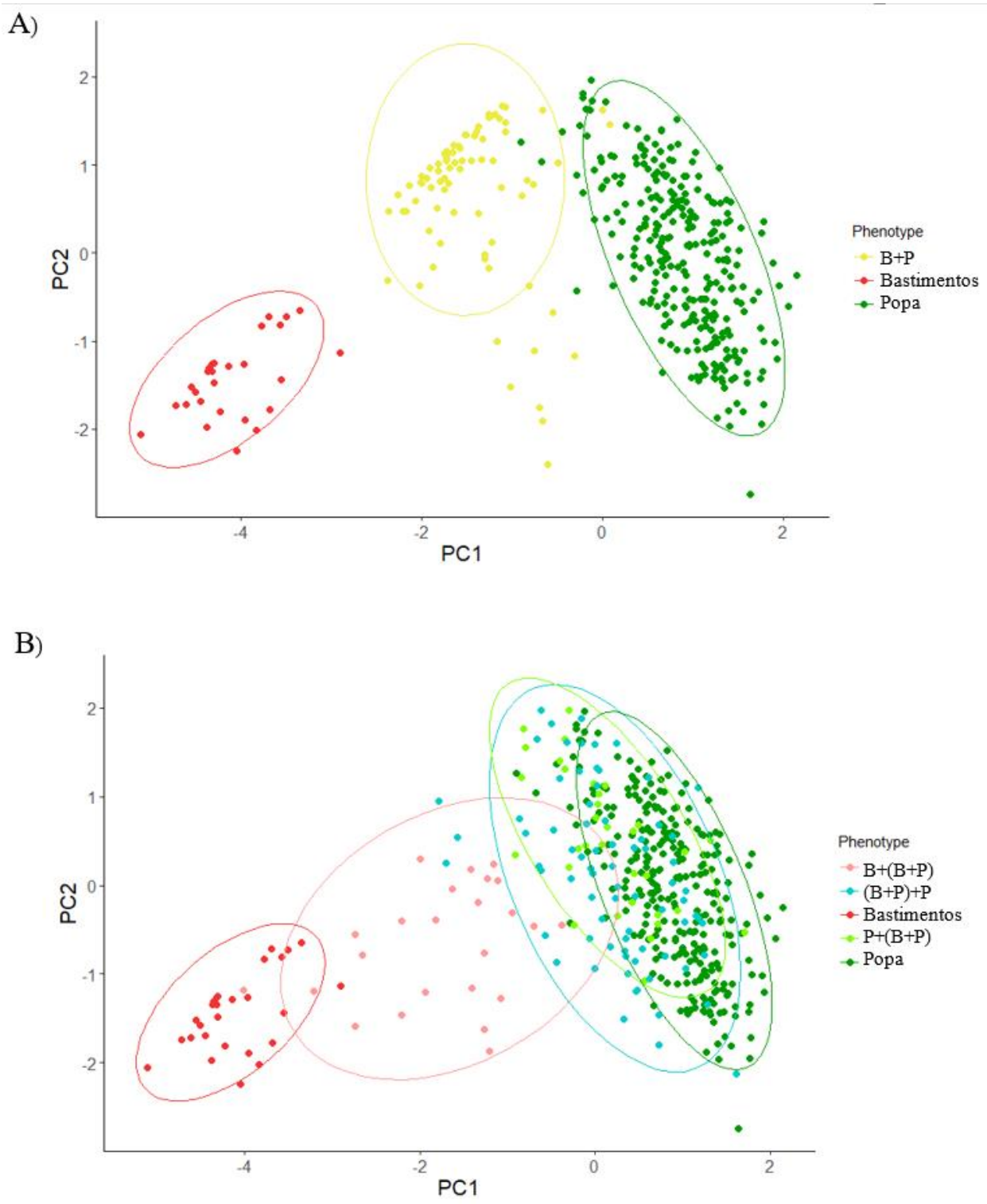
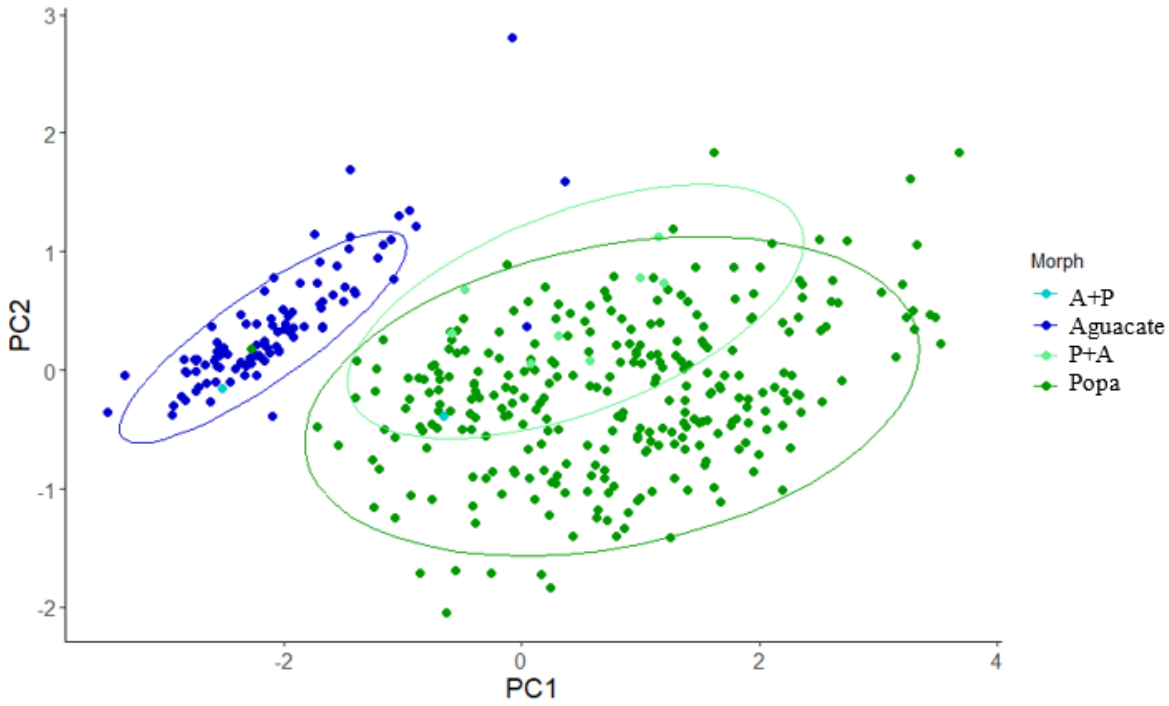


Figure 3-3 Color Sampling Locations

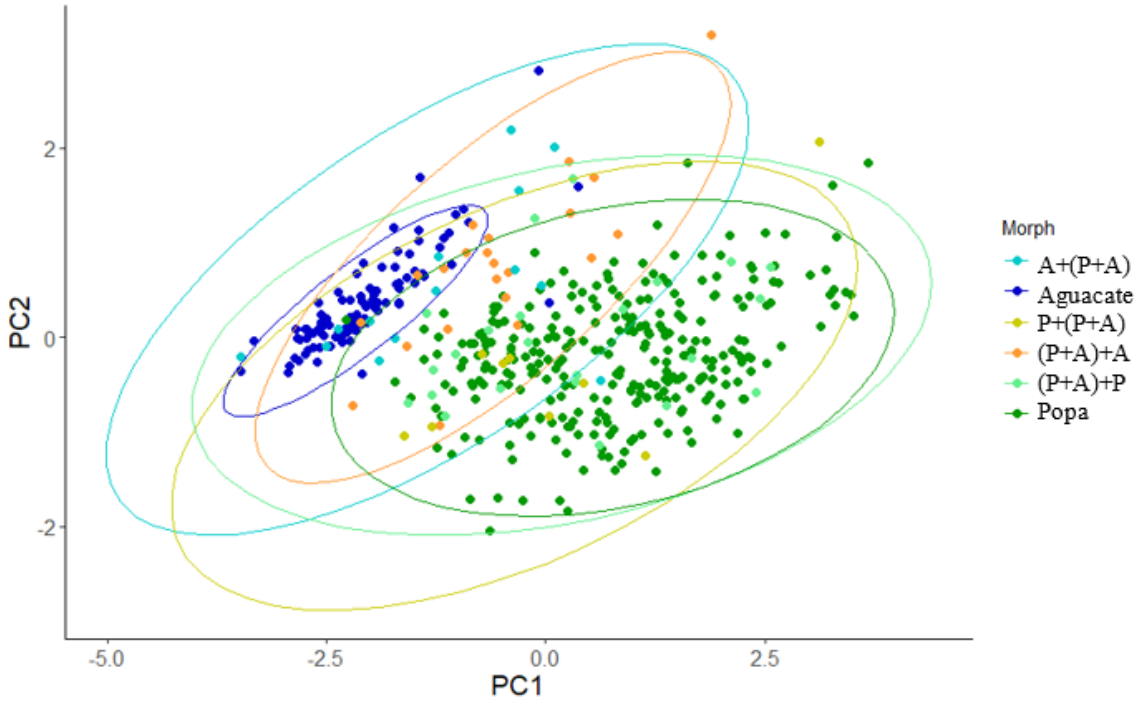
The approximate locations of color sampling on each frog. The line used to determine the size of the box is indicated between the eyes.


A)
PC2
PC1
2
1
0
-1
-2
-4
-2
0
2
Phenotype
B+P
Bastimentos
Popa
B)
PC2
PC1
2
1
0
-1
-2
-4
-2
0
2
Phenotype
B+(B+P)
(B+P)+P
Bastimentos
P+(B+P)
Popa

C)


PC2
PC1
3
2
1
0
-1
-2
-2
0
2
4
Morph
A+P
Aguacate
P+A
Popa

D)


PC2
PC1
2
0
-2
-5.0
-2.5
0.0
2.5
Morph
A+(P+A)
Aguacate
P+(P+A)
(P+A)+A
(P+A)+P
Popa

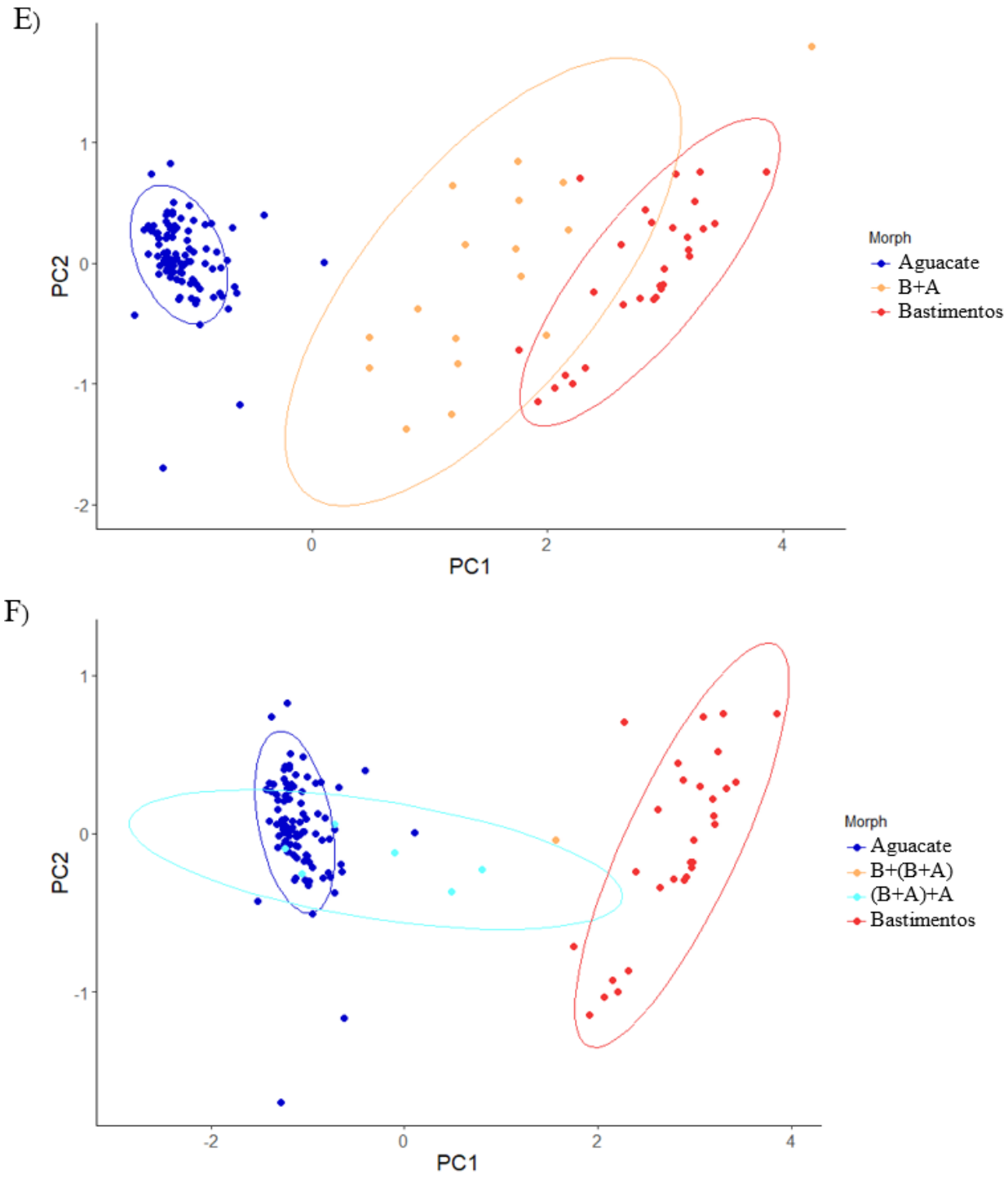


Figure 3-4 Color Phenotypes for Each Cross Type

Plots of dorsal color PC1 vs. dorsal color PC2 for different *O. pumilio* morphs and cross types. Panels (A) and (B) show data from the Bastimentos (red frogs) and Popa (green frogs)

crosses, (C) and (D) show data from the Aguacate (blue frogs) and Popa crosses, and (E) and (F) show data from the Aguacate and Bastimentos crosses. Panels (A), (C), and (E) all show data from the two parental morphs and the F1 hybrid frogs, whereas (B), (D), and (F) all show data from the two parental morphs and the backcrosses. Hybrids are listed as father then mother in cross names, with morph locations abbreviated to just the first letter of the population name (e.g., B = Bastimentos, P = Popa, A = Aguacate). Parentheses in the backcross names indicate that the parent was a hybrid (i.e., B+(B+P) is a Bastimentos male crossed with a female that is an F1 cross between a Bastimentos male and a Popa female). Ellipses represent 95% confidence intervals.

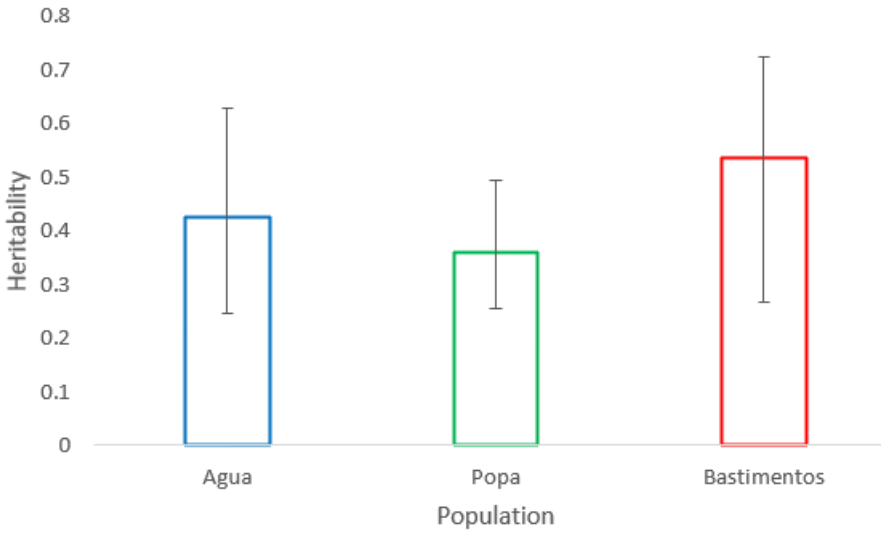


Figure 3-5 HeritabilityData

Heritability estimates of color for three morphs of *O. pumilio* obtained from a trivariate MCMCglmm animal model with color of each morph as a separate response variable. Error bars represent 95% confidence intervals. The color of the bar indicates the color of the frog.

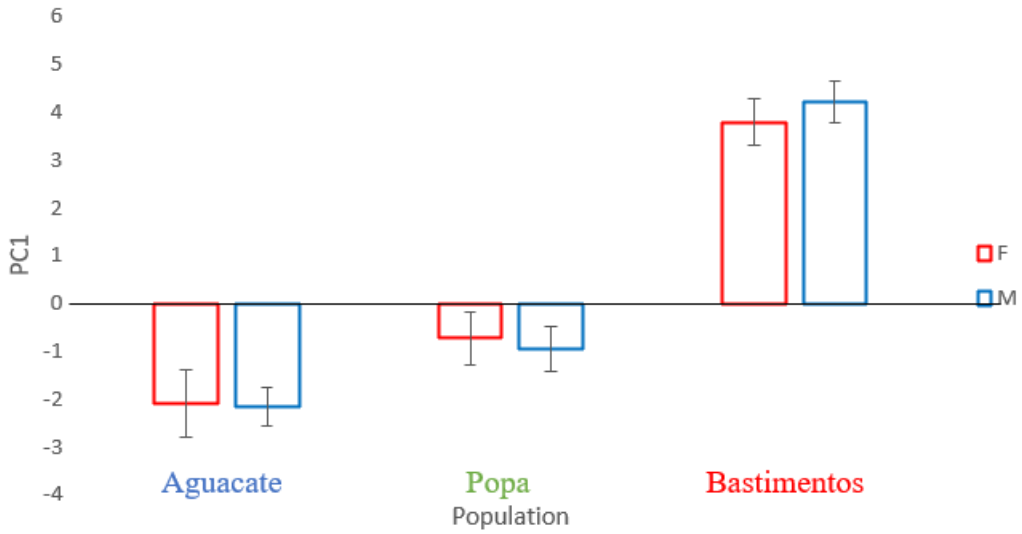


Figure 3-6 PC1 by Population and Sex

Mean (±SD) PC1 values for males and females by morph.

Table 3-1 Summary of Color PCA Eigenvalues

For each PC, the value of the eigenvalue, its variance, and the cumulative variance in color explained are given

	Eigenvalue	Variance	Variance percent
PC1	1.790809	5.97E+00	59.69364
PC2	1.209191	4.03E+01	100
PC3	1.81E-26	3.06E-29	100

Table 3-2 Summary of Variable Loadings from Color PCA

	PC1	PC2	PC3
RAVG/Total	0.975934	0.218065	-6.27E-16
GAVG/Total	-0.1883854	-0.9820952	-4.09E-16
BAVG/Total	-0.8960315	0.44399	-5.97E-16

4.0 Contrasting patterns of genetic and phenotypic variation in insular populations of the strawberry poison frog, *Oophaga pumilio*

4.1 Summary

Polymorphism is thought to be a step towards speciation but how this trait variation is produced remains unclear. Comparison of variation in neutral genetic loci with that of traits putatively under selection can be used to make inferences about the evolutionary processes that generate phenotypic variation. For natural selection, neutral loci are expected to be associated with selected traits in a way that is not structured by geographic distance, a pattern termed isolation by adaptation (IBA), though it is unclear if sexual selection produces a similar pattern. Here I test for IBA in three phenotypic traits putatively under selection in the strawberry dart frog *Oophaga pumilio* This species exhibits extreme phenotypic variation in coloration, particularly in the Bocas del Toro archipelago of Panama, which appears to have arisen recently and rapidly. Since coloration is used both as an aposematic signal and in mate choice, this trait has likely diverged in response to strong selection. Body size and male advertisement calls, however, are not known to be evolving via sexual or natural selection in *O. pumilio*, despite being traits commonly used for mate choice in other species. In this study I tested for concordant differences between neutral genetic variation and color, advertisement calls, and body size, hypothesizing a pattern of IBA across the island of Bastimentos, which has documented variation in color. I found a clear pattern of isolation by distance in the neutral genetic markers and that genetic distance was significantly or nearly significantly associated with differences in body size in males and females, respectively, suggesting that body size is evolving neutrally. I also found that call distance was not significantly

associated with genetic distance. Differences in color among sampling sites were pronounced, but were not significantly associated with genetic distances in neutral loci as expected for IBA. Since color is known to be under sexual selection in this species, this suggests that the expected patterns of IBA may not hold for traits evolving under sexual selection.

4.2 Introduction

The evolution of trait polymorphism is commonly posited as a step along the way to speciation (rev. in McKinnon *et al.* 2010). Polymorphisms in a wide variety of traits used in mate selection have been noted in many species. For example, polymorphisms in coloration occur in birds (Kokko *et al.* 2014), fish (Sandkam *et al.* 2015, Maan *et al.* 2008), and frogs (Myers and Daly 1976, Twomey et al. 2014), polymorphisms in mating call occur in frogs (Howard and Young 1998) and birds (rev. in Lemon 1975, Slabbekoorn and Smith 2002), and polymorphisms in body size occur in two fish: Arctic charr (rev. in Hindar and Jonsson 1993), and swordtails (Ryan et al. 1992). Sexual selection can aid in the divergence of a polymorphic sexual trait if the individuals bearing divergent sexual traits mate assortatively (rev. in Bolnick and Fitzpatrick 2007). However, it is less obvious how polymorphisms arise to begin with. Comparing the patterns of neutral genetic variation with that of diversity in phenotypic traits putatively under selection is one approach that can allow us to make inferences about the evolutionary forces generating trait variation and polymorphisms – the raw material on which selection acts. However, few studies have tested for concordance between patterns of neutral genetic variation and variation in sexually selected traits in animals.

The patterns that arise within and among populations in genes that are evolving neutrally (i.e., by mutation and genetic drift) can be used as a backdrop on which to make inferences about the evolutionary processes that are shaping phenotypic variation. For example, a pattern of isolation by distance (IBD), where differences in neutral genetic loci between individuals or populations increase with geographic distance, is the expected result when genetic drift is the main evolutionary process at work (rev. in Nosil *et al.* 2009, Orsini *et al.* 2013). This pattern is strongest when individuals have low dispersal (i.e., gene flow among populations is minimal). In this case genetic structure results from different alleles drifting to fixation in different populations that have limited connectivity via gene flow (rev. in Nosil *et al.* 2009, Orsini *et al.* 2013).

In contrast, isolation by adaptation (IBA) produces a pattern where the magnitudes of differences in phenotypes under selection are not associated with geographic distance, with limited divergence in neutral loci, but where the magnitude of trait differs. This pattern is expected to result when adaptation to different biotic and/or abiotic conditions in a heterogenous landscape is the main evolutionary process at work. Under this scenario, migrants among differently adapted populations are selected against due to an unfavorable match with the abiotic environment, with heterospecifics (e.g., predators, competitors, parasites, or prey), or both, thus allowing neutral loci to diverge via drift alongside loci under selection in separate demes (rev. in Nosil *et al.* 2009, Orsini *et al.* 2013, Wang and Bradburd 2014). However, the resulting pattern in phenotypic variation produced depends on the strength of selection. Strong selection that acts against migrants to prevent gene flow would result in a correlated divergence in neutral alleles and the selected traits that is not associated with geographic distance and thus would produce a pattern of IBA. Weak selection, however, would still allow for geneflow between color morphs so should produce a pattern of IBD. While the theory on which expectations for IBA was developed focused almost

exclusively on adaptation in response to differential natural selection, differential sexual selection based on variation in phenotypic traits used in mate choice or in male-male competition has been hypothesized to produce similar patterns if immigrants are selected against (rev. in Wang and Bradburd 2014). Though, to my knowledge, no empirical test for IBA under sexual selection has been done.

The existence of both polymorphic and polytypic populations of the strawberry dart frog (*Oophaga pumilio*) in the Bocas del Toro region of Panama make this species an ideal one for research on the evolutionary processes that drive phenotypic divergence. *Oophaga pumilio* is a small forest-dwelling frog with a range that extends along the Caribbean side of Central America from central Nicaragua to western Panama. Throughout most of its range, the frog is a bright red color with blue or black limbs. However, in the Bocas del Toro archipelago of western Panama, numerous color morphs that span the visible spectrum have been described (Myers and Daly 1976, Wang and Summers *et al.* 2010, Rudh *et al.* 2010). This color variation appears to have evolved recently and rapidly as the archipelago, which was formed by sea level rise, took on its present form only during the past 1000-5000 years (Anderson and Handley 2002).

In *O. pumilio*, coloration is used as an aposematic signal, warning of toxicity to would-be predators (Saporito *et al.* 2007, Maan and Cummings 2012). However, color is also a target of sexual selection in these frogs, where females use color as a cue to choose among potential mates (e.g., Summers *et al.* 1999, Reynolds and Fitzpatrick 2007, Maan and Cummings 2008,

Richards-Zawacki and Cummings 2010, Maan and Cummings 2012) and males use color to choose how they interact with rival males (Yang *et al.* 2018). Specifically, female *O. pumilio* from the Bocas del Toro region of Panama generally prefer to court with males bearing a similar color phenotype to their own (Summers *et al.* 1999, Maan and Cummings 2008, Reynolds and

Fitzpatrick 2007, Richards-Zawacki and Cummings 2010, Yang *et al.* 2016). Similarly, males have been found to be more aggressive in their territorial interactions toward other males bearing the local coloration (Yang et al. 2018). Taken together, these studies suggest that the frog's striking color variation may have arisen, at least in part, due to sexual selection.

Several prior studies have looked for associations between coloration and genetic structure in *O. pumilio* across the Bocas del Toro archipelago, resulting in varied and at times, conflicting, inferences about the roles of selection and genetic drift (Rudh *et al.* 2007, Wang and Summers 2010, Brown *et al.* 2010, Gehara *et al.* 2013). However, several studies using both modelling and genetic approaches have reported results consistent with divergent selection contributing to the observed variation in coloration (Brown *et al.* 2010, Gehara *et al.* 2013). A likely contributor to the conflicting results among these studies may have been that each one sampled a different subset of populations/color morphs and small numbers of frogs per population, and most sampled from just one collection location per island/morph. By their nature these sampling schemes preclude investigations of the effects of ongoing gene flow and/or selection on patterns of phenotypic variation. In hopes of yielding new insights into the evolutionary forces shaping phenotypic divergence in this species, I designed the present study with a focus on characterizing patterns of variation across finer geographic scales, and across populations connected by contiguous habitat.

While color is important for the later stages of mate choice, courtship in *O. pumilio* begins with females approaching the calls made by territorial males (Pröhl and Hödl 1999). This suggests that along with coloration, the properties of male advertisement calls may be shaped by sexual selection. Call divergence driven by sexual selection is common in frogs (Ryan and Rand 1990, Boul *et al.* 2007, Moreno-Gómez *et al.* 2015). However, prior work on *O. pumilio* populations from across the species' range found call divergence to be uncorrelated with neutral genetic

divergence (i.e., no pattern of IBA), suggesting that calls may not be diverging due to sexual selection in this species (Pröhl *et al.* 2007). However, tests for correlations between call divergence and genetic divergence in the polymorphic Bocas del Toro region have thus far only been made among frogs from different islands (i.e., among populations with no contemporary gene flow). The present study, with its finer scale sampling across more contiguous habitat, is therefore better suited to test for patterns that allow us to distinguish the action of genetic drift from selection in shaping male calls.

Body size is also a trait frequently used in sexual selection (Darwin's finches, Price 1938, dung flies, Banckenhorn *et al.* 2000, Javan sparrows, Hasegawa *et al*. 2011, two-spotted gobies, Borg *et al.* 2005) that has been documented to be subject to female choice in a toad species (*B. quercicus*, Wilbur et al. 1978), though this is not the case for all frog species tested (*Hyla arborea*, Friedl and Klump 2001; *Bufo americanus, B. terrestris* x *americanus, B. terrestris*, Wilbur *et al.* 1978) and it has not yet been tested for its role in female choice in *O. pumilio*. However, body size in amphibians can also be indirectly under selection via selection on male advertisement calls. Body size in frogs correlates with call frequency, where bigger frogs call at lower frequencies Thus, selection for call frequency indirectly selects for male body size. Female preference for lower frequency calls, has been seen in the treefrog *Hyla chrysoscelis* (Morris and Yoon 1989), highlighting that indirect selection on body size has been found in at least one frog species. Larger body size can also be selected for as advantageous during male combat and evidence suggests that larger males are more likely to win territorial disputes in *O. pumilio* (Clause 2017).

I hypothesized that genetic structure among *O. pumilio* populations, even from a single island, would be strong since these frogs are thought to move little throughout their lives. Both sexes have relatively small home ranges; males defend territories that have been estimated to range

in size from 5.0 to 34.85 m^2 and females generally occupy somewhat larger (3.9-93.8 m^2) home ranges that overlap with the territories of several males (Meuche *et al.* 2011). The occurrence and extent of juvenile dispersal in *O. pumilio* remains unknown. Juveniles of other terrestrial frog species have been found to disperse up to ~ 500m from where they were born, but more often that distance is 200m or less (Driscoll 1997, Osawa and Katsuno 2001) and several frog species have been found to have strong genetic structure across similarly small geographic distances (Monsen and Blouin 2003, Knopp and Merilä 2009, Pan *et al*. 2019, Ferreira *et al.* 2020, but see Furman *et al.* 2016). Additionally, since the islands were connected in the past (Anderson and Handley 2002), neighboring islands should show this past connectivity in the genetic structure of neutral loci.

My focal island, Bastimentos, is one of the larger islands in the Bocas del Toro archipelago and is different from many of the islands these frogs inhabit in that the populations of *O. pumilio* that call it home are variable in color. Across Bastimentos, the frogs exhibit a wide range of dorsal coloration from reds to oranges, yellows, and even light greens, with variation in black dorsal spotting pattern and ventral and limb color occurring across the island as well (Figure 4-1). Across most of the island, color variants are found in allopatry. However, populations at the northwestern tip of Bastimentos are polymorphic (sites 4, 5, 6 in Figure 4-1), and contain frogs ranging in dorsal color from red to yellow-green, all with black spots or patterning on the dorsal side (Richards-Zawacki and Cummings 2011). Thus, Bastimentos provides a unique opportunity to examine the association between neutral genetic variation and variation in coloration, a trait that is known to be important in the context of mate choice and reproduction even among individuals within the polymorphic populations (Richards-Zawacki and Cummings 2011, Richards-Zawacki *et al.* 2012). While variation in calls hasn't been tested across Bastimentos, there is considerable variation in call parameters between islands (Pröhl 2007), suggesting the capacity for differential selection on

calls. Body size variation across Bastimentos has not been reported to my knowledge, frogs on Bastimentos are on average larger (20 mm ±0.45) than other islands including nearby Solarte (17 mm ±0.35) (Galeano and Harms 2015), suggesting there may be finer scale variation across an island that has yet to be tested for.

In this study, I compare variation in phenotypic traits putatively under sexual selection with genetic variation in neutral microsatellite loci for *O. pumilio* individuals sampled from localities that span the island of Bastimentos (~ 1-2.5 km apart). By testing frog sampling locations across a single island, I eliminate aquatic barriers as a confounding variable allowing us to more directly test for associations with neutral genetic structure and phenotypic traits. I also sampled populations from three neighboring islands that, due to their recent (~1000 – 5200 years ago, Anderson and Handley 2002) connectivity with Bastimentos, may be important to understanding the regional context for genetic and phenotypic variation. Because the Bocas del Toro archipelago formed so recently, drift is unlikely to be the sole cause of the observed variation in *O. pumilio* coloration. Thus, I hypothesize that color will show a pattern of IBA, that is, differences in color variation will not associate with differences in geographic distance but will associate with differences in genetics. For male calls, given that variation in call parameters does not correlate with genetic distance in frogs across the species' range (Pröhl *et al.* 2007), I hypothesize that this pattern would persist across a finer geographic scale and will show a pattern of IBA. And finally, since body size is also related to male call frequency and body size is used in mate choice in other frog species, I hypothesize that body size will show IBA. My study contributes to our understanding of sexual selection as a potential source of adaptation for patterns of IBA by examining genetic and phenotypic variation in a polymorphic species. Since sexual selection preferences aren't always acted on in mate choice and allow for some variability in male traits, it is possible that sexually

selected traits might vary continuously with distance in a pattern of IBD, especially given that the traits studied are continuous variables that change gradually across the island. My study uses *O. pumilio* as a test to see if the predictions of IBA vs. IBD that were developed with natural selection in mind can be used to identify cases where sexual selection may be shaping variation instead.

4.3 Methods

4.3.1 Sample Collection

From 13 July 2007 to 30 July 2007 *O. pumilio* were captured at 16 locations (spaced ~ 1-2.5 km apart) on the island of Bastimentos (Figure 4-1). Later on (between 18 May 2008 and 20 July 2008), frogs were captured from four locations on the three neighboring islands, Colon, Solarte, and Popa. photographs of the frogs' dorsal and ventral sides were taken for color analysis. Photographs were taken in the field on an 18% grey card standard background using a Sony Cybershot camera under ambient light. The frogs' snout-vent length (SVL) was also measured using dial calipers. tissue samples were collected by clipping off a single toe pad with a sterile razor blade before releasing each frog at its point of capture. Toe clips were placed in a tissue preservative (salt-saturated DMSO and EDTA solution) and stored at room temperature prior to genomic DNA extraction. For this, a Qiagen DNeasy Blood and Tissue kit was used, following the protocol for animal tissue.

4.3.2 Genetic Analysis

To characterize genetic structure across the island of Bastimentos, I genotyped 8-12 individuals from each sampling location at eight different polymorphic microsatellite loci (Oop_C3, Oop_B8, Oop_H5, Oop_B9, Oop_F1 Oop_E3, Oop_D4, Oop_C11, Hauswaldt *et al.* 2009). The forward primer of each pair was labeled with a fluorescent dye (6-FAM, VIC, HEX, or NED). I amplified each microsatellite region for fragment size determination using PCR. The PCR mixtures contained 4.8μL of GoTaq Green Master Mix (Promega, contains Taq polymerase, buffer, dNTPs, and $MgCl_2$), 0.6μL of each primer, 4.8μL of molecular grade water, and 1.2μL of DNA for a total reaction volume of 12μL. PCR cycling conditions were as follows: initial denaturation at 95˚C for 5 min, 35 cycles of denaturation at 95˚ for 30 s, 55˚C for 90s, and 72˚C for 45s, followed by a final extension period at 60˚C for 30 min (Hauswaldt *et al.* 2009).

I determined microsatellite fragment size using an Applied Biosystems 3730 DNA analyzer. Microsatellite fragment size was scored for each individual and locus using the GeneScan 500 Liz ladder (Applied Biosystems) as a reference in Geneious version 9.2.1 (http://www.geneious.com, Kearse *et al.* 2012). I called alleles based on the bins generated by Geneious, which were based on the locus repeat unit size (4 bp) and spanned the full range of allele sizes that I observed for that locus, which sometimes extended beyond those described in Hauswaldt *et al.* (2009). I repeated the PCR and fragment analysis steps for samples with no amplification as well as those that produced ambiguous results in the initial fragment analysis (i.e. faint peaks or seemingly more than two peaks of a similar height). If the second attempt produced peaks that could be called with confidence, then those allele sizes were recorded. Otherwise, that locus was treated as missing data for that individual. To minimize potential errors in my microsatellite scoring, I ran my data through several quality control tests. First, I checked for

scoring errors due to stutter and large allele dropout using *MicroChecker* v. 2.2.3 (Van Oosterhout *et al.* 2004). I then tested each microsatellite locus for deviation from Hardy-Weinberg equilibrium, for linkage disequilibrium, and for the presence of null alleles using the allele frequency analysis tool in *Cervus* v3.0.7 (Kalinowski *et al.* 2007).

To estimate the genetic distance between sampling locations, I calculated pairwise F_{st} values between all locations using the web version of *GenePop* v. 4.7.5 (Raymond and Rousset 1995, Rousset 2008). To infer the appropriate number of genetic clusters represented by the samples and to assign individuals membership within these genetic clusters I used *Structure* (version 2.3.4, Pritchard *et al.* 2000). I ran the program with a burn-in of 10,000 iterations and a run of 100,000 iterations using an admixture model and did not use collection locations as prior information in determining cluster assignment (i.e., the LOCPRIOR model was not used). I ran the simulation for values of K ranging from K=1 to K=10 with 10 runs per K value. The *Structure* output was visualized using *Clumpak* and the appropriate number of genetic clusters was computed using the Evanno method (Evanno *et al.* 2005) as implemented in *Clumpak* (Kopelman *et al*. 2015).

4.3.3 Color Analysis

I used *ImageJ* v. 1.50i (Schneider *et al.* 2012) and photographs of each frog to characterize color variation. In *ImageJ*, I extracted color information from six 20 x 20 pixel squares on each of the dorsal and ventral images of each frog (Appendix Figure 1). The squares were positioned to minimize the area of the square that contained black (melanistic) spots, which some frogs have on their dorsa, as I was interested in characterizing variation in background coloration. In male frogs, I avoided measuring color of the dark throat patch, which only males possess on their ventral sides, by placing the measurement area just ventral to the throat patch. I used the RGB Measure tool in

ImageJ to calculate the mean value of all pixels within each measurement square for each of the three color channels in the RGB color model (R, G, and B corresponding to Red, Green and Blue, each of which range in value from 0 to 255). The values from all six squares from a frog's body surface were averaged together to get a mean frog score per surface, which was then used to get sampling location mean dorsal and ventral scores for R, G, and B.

4.3.4 Call Collection

I recorded advertisement calls made by male *O. pumilio* at several of the sampling locations on Bastimentos and Solarte (Figure 4-1) during four time periods: 10 March 2008 to 26 July 2008, 10-24 October 2008, 14 December 2016, and 5-9 March 2017. I recorded 10 calling males per location using a Marantz Professional Solid State Recorder (PMD660) at a sampling rate of 44.1 kHZ and a Sennheiser shotgun (MKH70) microphone positioned ~1m in front of the calling male. I recorded calls between the hours of 0700 and 1700, when the frogs are most active (Pröhl 1997). I recorded up to 10 call bouts, where a call bout is defined as a single call sequence, per male but included in the dataset all males for which I had at least five recorded call bouts. When possible, I also recorded air temperature when the frog was calling using a handheld weather station (Kestrel 2000).

4.3.5 Call Analysis

Oophaga pumilio calls consist of a series of nearly identical notes, which are each made up of a series of pulses. Parameters of the recorded male advertisement calls were analyzed in *Raven* v. 1.5 (Bioacoustics Research Program 2014) at a sampling rate of 24,000 Hz. For each call,

I determined the call duration (time from start to end of call bout), the note rate (number of notes/call duration), and the dominant frequency (frequency in Hz with the maximum energy) of the call. I also measured the pulse rate for 10 individual notes in the middle of each call bout and calculated duty cycle for each call bout as the note rate multiplied by the average note length of the call bout (Appendix Figure 2).

4.3.6 Statistical Analysis

All analyses were done in R version 3.3.1 (R Core Team 2013) except where otherwise noted. To test the hypothesis of isolation by distance for sampling locations on Bastimentos, I used a Mantel test (Mantel 1967) to compare pairwise Fst/(1-Fst) (Rousset 1997) values to distances between sampling locations. I used Haversine distances, which take into account the curvature of the earth. I calculated these using the coordinates I measured at the time of sampling at each location using a handheld GPS unit (Garmin eTrex 10, WGS 1984 reference) and using the 'distHaversine' function in the R package *geosphere*.

For color, I did separate analyses for dorsal and ventral surfaces using the average R, G, and B values for each frog. To test for differences among sampling sites, I compared the square-root transformed mean R, G, and B values for individual frogs in separate linear models using sampling location as a fixed effect. I used post-hoc Tukey HSD. I also calculated the Mahalanobis distance between sampling location means of each color channel (R, G, and B) to find the average color value for each location. I calculated differences in collection locations for each color channel separately using a linear model with collection location as a fixed factor and did pairwise post-hoc comparisons correcting for multiple comparisons using a Bonferroni correction. The Mahalanobis distances between sampling sites were then compared pairwise with the corresponding pairwise

Fst/(1-Fst) values using Mantel tests. Additionally, I ran partial Mantel tests comparing pairwise distances for each variable (dorsal and ventral color, call, and body size) with Fst/(1-Fst) while controlling for geographic distance (Smouse *et al.* 1986).

To test for associations between call variation and genetic variation, I first calculated the average Mahalanobis distance between call parameters for each collection location. I then used a Mantel test to compare pairwise distance between call parameters and genetic distance.

To test for associations between body size variation and genetic variation, I used a Mantel test to compare the absolute value of the difference in mean SVL between sampling locations with pairwise Fst/(1-Fst) values. I then used a Mantel test to compare pairwise distance in body size with genetic distance.

4.4 Results

I scored alleles for 198 individuals from 20 different sampling locations (8 to 12 animals per location, mode of 10) at eight microsatellite loci (Table 4-1). Only one locus (Oop_E3) had a large amount of unscorable samples (Table 4-1), mostly from the sampling locations in the southern portion of Bastimentos. I ran the structure analysis without the E3 locus and found a similar pattern, so I left it in to provide more data for the structure analysis. Each locus was highly polymorphic with the number of alleles per locus ranging from 17 to 33. Tests of Hardy-Weinberg equilibrium (HWE) revealed that samples from many locations were significantly out of Hardy-Weinberg equilibrium for several loci after Bonferroni correction (Appendix Table 4) likely due to small sample sizes per collection location. Across all samples, only loci E3 and B8 were not significantly out of HWE. There was no significant linkage disequilibrium for any of the eight

microsatellite loci. The proportion of null alleles was generally low, but two loci had proportions of null alleles around 0.3 (C3 and B9, Appendix Table 4). There was also no evidence of scoring error due to stutter or large allele dropout, though there was too much missing data to run tests for several loci, though these loci were not consistent across sampling locations.

Overall, the genetic distance between sampling locations was low (pairwise Fst ranged from 0 to 0.0892, Table 4-2). However, there was a clear association between genetic distance (Fst) and geographic distance among sampling locations on Bastimentos Island (Mantel test, $p < 0.001$, 9,999 permutations) and this pattern persisted when all sampling locations, including those on Popa, Colon, and Solarte islands, were included (Mantel test, $p = 0.047$, 9,999 permutations). *Structure* analysis suggested there were two genetic clusters on Bastimentos (Figure 4-2). Sampling locations on the northwest part of Bastimentos generally belonged to one cluster, while the sampling locations in the eastern portion of the island belonged to the other. Sampling locations along the southwestern coast showed mixed assignment to the two genetic clusters while sampling locations from nearby islands mostly matched the cluster assignment of the nearest sampling location on Bastimentos (Figure 4-3).

Dorsal coloration differed significantly among sampling sites for all three color channels (R, G, and B: linear models, $F_{1, 472} \geq 19.62$, $p < 0.001$) as did ventral coloration for all three color channels (linear model: $F_{1, 471} \geq 12.2$, $p < 0.001$). Post-hoc tests for dorsal color resulted in many significant comparisons post-correction, but the patterns of non-significant collection sites varied for each color channel (Appendix Table 1). For the R channel, sites 12-15, located in the middle of the wide section of the island generally were not significantly different from each other, but most other collection sites were significantly different from one or both neighboring sites as well as significantly different from the majority of collection sites. For the G channel, sites 10-19,

located in the entire southern half of Bastimentos were generally all non-significantly different from each other. For the B channel, the three polymorphic collection sites (4, 5, 6) were significantly different from all other populations except for the comparison between sites 4 and 5. Sites 7-10, located just outside the polymorphic region in the narrow part of the northern island, were not significantly different from each other, and the populations towards the south generally formed a cluster of nonsignificant comparisons (typically 12, 13, 14, 15, and 17 with 18 and 19 sometimes included). Post-hoc tests for the ventral side also resulted in many significant comparisons. For the R channel, site 16 was significantly different from every other collection site, and sites 5, 6, 8, 9, and 19 were significantly different from most other sites (1-2 nonsignificant comparisons). For the G channel, site 11 was significantly different from every other collection site, and sites 8, 9, and 19 were significantly different from most other sites (1-2 nonsignificant comparisons). For the B channel, sites 4, 11, 15, and 19 were significantly different from most other sites (1-2 nonsignificant comparisons). Pairwise Mahalanobis distances, which encompass all three color channels, were not significantly correlated with Fst/(1-Fst) for either dorsal (Mantel test, $p = 0.168$, 99,999 permutations) or ventral (Mantel test, $p = 0.100$, 99,999 permutations) sides. Partial Mantel tests accounting for the effect of geographic distance on correlations between phenotype and genetic distances resulted in similar or greater p-values than the corresponding simple Mantel tests for both dorsal and ventral (partial-Mantel test: $p = 0.99$ dorsal, $p = 0.39$ ventral, 99,999 permutations).

Distance in call variation was calculated using pairwise Mahalanobis distances, which describe variation in all the measured call metrics together, were not significantly correlated with Fst/(1-Fst) (Mantel test, $p = 0.101$, 99,999 permutations). The mean dominant frequency of the calls I recorded was 5.01 kHz ± 0.331 SD, which is a bit higher than has been previously reported

for this species (4.45 kHz ± 0.41, Pröhl 2007). A partial mantel test comparing call distance and genetic distance while controlling for geographic distance produced similar results to the full Mantel test (partial Mantel test, p= 0.095, 99,999 permutations). Pairwise Mahalanobis distances between collection sites were generally pretty low (under 3.0), but collection site 15 had consistently higher Mahalanobis distances between each of the other sites. The call parameters of pulse rate, duty cycle, dominant frequency and note rate were all significantly different across sites ($F_{8, 43} \geq 3.86$, $p < 0.001$).

To compare variation in the microsatellite markers with that of body size, I analyzed males and females separately, since females are larger than males (Chaves-Acuña et al. 2020) and selection may be acting differently on body size for the different sexes (Zhang *et al.* 2012, Nali *et al.* 2014). In male *O. pumilio,* body size differences were significantly associated with pairwise Fst values (Mantel test, p = 0.0315, 99,999 permutations), but female body size differences were marginally non-significantly associated with pairwise Fst values (Mantel test, p = 0.0597, 99,999 permutations). This pattern persisted when controlling for geographic distance (partial Mantel test, p = 0.046 males, p = 0.074 females, 99,999 permutations).

4.5 Discussion

In this study, I tested for a pattern of isolation by adaptation (IBA) in three traits putatively evolving under sexual selection in a highly color polymorphic poison frog: coloration, male calls, and body size. Because fine-scale genetic structure is common in frogs (Monsen and Blouin 2003, Knopp and Merilä 2009, Pan *et al*. 2019, Ferreira *et al.* 2020) and adults of this species are known to occupy small home ranges (Pröhl and Berke 2001) I hypothesized that I would find a pattern of

isolation by distance (IBD) in neutral genetic loci across the island of Bastimentos, which I indeed found. Since Bastimentos had prior connectivity with neighboring islands (Anderson and Handley 2002), I predicted that this pattern of IBD would extend from Bastimentos to neighboring islands, which I also found to be true. Furthermore, neighboring sampling locations on Bastimentos tended to cluster together in my Structure analyses and frogs from locations at both the northern and southern tips of the island were assigned to the same genetic clusters as sampling locations on the adjacent islands (Figure 4-3), which would have been neighboring populations prior to island separation. Since Bastimentos was previously connected to the mainland via connections with neighboring islands on both its north and south ends, the apparent genetic break near the center of Bastimentos could have resulted from frogs having colonized the island from both ends.

The range of pairwise Fst values obtained in this study (Fst = 0-0.0892) suggest that in *O. pumilio,* color diverged across Bastimentos despite a high level of gene flow throughout the island. The Fst estimates obtained for *O. pumilio* are similar to Fst values previously obtained for frogs in studies of frog populations with continuous land connectivity (e.g., *Rana sylvatica* Fst =0.006-0.048, Newman and Squire 2001). Higher Fst estimates that extend well beyond the values found in this study have been found in frog species for which a mountain range extends likely causes a barrier to gene flow and allows for more genetic differentiation (*R. luteiventris* Fst = -0.009-0.52, Funk *et al.* 2005; *R. cascadae* Fst = 0.01-0.52, Monsen and Blouin 2004). There doesn't appear to be any barrier that could cause the genetic break seen in this study as frogs on Bastimentos and across the archipelago typically occupy forest habitat with similar light characteristics and plant communities (Summers *et al.* 2003). Additionally, while there is some variation in elevation on Bastimentos (range: 0 to 62 m asl), differences are minimal and do not associate with the genetic break. Additionally, habitat is fairly homogenous across the island and is nearly contiguous due to

a minimal human presence on the island that is mostly restricted to the coast. The low amount of genetic differentiation obtained across the sampling locations in this study suggests either current or recent gene flow among the sampling locations on Bastimentos.

I did not find the hypothesized pattern of IBA in color, as differences in color, a trait putatively evolving under sexual selection among *O. pumilio* populations, were not correlated with genetic distances among neutrally evolving loci. I also found little evidence for color variation across the genetic break found in the Structure analysis (Supp. Fig. 3). This suggests that genetic drift cannot explain the differences in color we see, nor can they be explained by genetic demes that are isolated due to selection as is expected under IBA. Therefore, the results of my study suggest that coloration is evolving faster than would be expected by drift alone, since Fst values between sampling locations that are quite different in color are often quite low. For example, previous studies have found the Fst values between red and "non-red", which encompasses orange, yellow, and light green, frogs in the polymorphic area of northwest Bastimentos to be near zero (Fst = 0.0003 from sampling location 4, Richards-Zawacki *et al*. 2012). However, frogs in these populations had the greatest differences in dorsal coloration. This color polymorphism appears to be stable (Richards-Zawacki *et al*. 2013), despite ongoing gene flow between the color morphs (Richards-Zawacki *et al*. 2012). The stability of color polymorphisms in this species have been proposed to result from maternal imprinting, which generates color-associated biases in both female mate choice and male-male competition (Yang *et al*. 2019). At the other extreme, I found several population pairs (i.e. collection site pairs 7 and 16, 7 and 19, 8 and 16, 15 and 18) that had low dorsal coloration distances (Mahalanabis distance range of 0.03-0.9) but higher genetic distance (Fst range 0.075-0.089) among them indicating that color can be similar between sampling locations connected by lower levels of gene flow. This taken together with the findings

from the polymorphic region suggests that the expectation under IBA might not work when sexual selection is at play. Additionally, color differences between collection sites differed between dorsal and ventral sides. Site 6 had the highest average pairwise Mahalanabis distance in dorsal coloration, while site 19 had the highest average distance for the ventral side. This suggests differences in how selection is acting on dorsal and ventral sides.

My color results suggest that sexual selection has the potential to disrupt the patterns expected under IBA. While divergence in sexually selected traits can be driven by ecological differences among populations (e.g., differences in light environments: Maan and Seehausen 2006, Maan and Seehausen 2010; the presence/absence of predators: Endler and Houde 1995, Sandkam *et al.* 2015; availability of dietary components that contribute to sexually selected traits: Craig and Foote 2001, Ballentine 2006), this need not be the case. My case study with *O. pumilio* demonstrates this well as the largest differences I found in coloration, the trait for which I have the greatest empirical support for sexual selection from previous studies, were found among individuals from the same sampling localities (i.e., in polymorphic populations). While aposematism theory would predict strong selection against immigrants bearing non-local coloration (Briolat *et al.* 2019), and thus the potential for patterns of IBA to arise Previous work both in the polymorphic region of Bastimentos (Richards-Zawacki et al. 2013) and more broadly in *O. pumilio* (Dreher *et al.* 2015), has shown that predators are equally likely to attack red, orange, yellow, and green morphs of *O. pumilio.* Thus, natural selection against non-local color morphs appears to be weak at best, allowing gene flow to prevent the buildup of associations between color and neutral genetic loci (i.e., the pattern predicted under IBA).

The lack of a pattern of IBA for color in my study could have been driven by the particular form of learning that generates biases in mating behaviors in *O. pumilio*. Experimental evidence

suggests female mate preferences and male biases in territorial behavior are learned at the tadpole stage via imprinting on the mother's color (Yang *et al.* 2019). Population genetic models have shown that this form of learning can quickly build associations between novel colors and behavioral biases for them and also allow new phenotypes to persist by providing negative frequency dependent selection. While this mechanism is presumably good at generating stable polymorphisms (Yang et al. 2019), frogs don't always find a mate of their preferred color, thus producing gene flow between color morphs that breaks down any associations between neutral and color-associated loci (Richards-Zawacki and Cummings 2011, Richards-Zawacki et al. 2012, Yang et al. 2016).

The relationship between variation in male advertisement calls, another trait putatively under sexual selection, and genetic variation was more complex. While all of the call parameters tested except for call duration were significantly different across collection sites, overlapping confidence intervals on most of them and small samples sizes suggest that this finding should be interpreted with caution. Note rate is particularly interesting as many 95% confidence intervals do not overlap and the variation in note rate was split across the genetic break found in the microsatellite loci, where northern sites in the blue genetic cluster had lower note rates than those in the southern orange genetic cluster. Since microsatellites are presumed neutral loci, this suggests that at least some call parameters are evolving neutrally and are thus not under sexual selection. Duty cycle and pulse rate were more variable for the two polymorphic locations (sites 5 and 6) then they were for other sites tested, suggesting that there may be differences between morphs in these call parameters and that there may be some sexual selection for these traits. However, I do not have morph data for the frogs tested so I cannot check for a difference between morphs in these

call parameters. Additionally, I did not take body size measurements of the frogs recorded for call analysis so I am unable to test for an association between body size and call parameters.

Because body size can be indirectly selected for along with call variation, I predicted that variation in body size would show a pattern of IBA. While I found that male body size matched this prediction and differences were significantly associated with genetic distances in the microsatellite loci used, female body size was marginally non-significantly associated with genetic distance. This suggests that there may be selection for larger males. Females may prefer larger males, or larger males may be more likely to be able to obtain and hold a territory and therefore more likely to obtain mates. For females, the marginal non-significance could simply be caused by lower statistical power as for many sampling locations I had larger sample sizes of males (n = 13 to 23) compared to females (n = 7 to 16). It could also suggest weaker selection on females compared to males for larger body size.

Overall, I did not find a pattern of IBA in color, a trait known to be under sexual selection in this species, and the only trait that I found a pattern consistent with IBA was male body size, a trait not known to be under sexual selection in this species but which has been shown to be influenced by sexual selection in another frog species (Wilbur et al. 1978, Howard and Kluge 1985). This suggests that the expectations for IBA do not seem to work well in distinguishing trait evolution due to sexual selection from trait evolution due to drift. Despite the lack of and IBA pattern, sexual selection does appear to be important in shaping coloration in *O. pumilio* and while several studies have shown that coloration is an important trait for female mate choice in lab assays (Maan and Cummings 2008, Reynolds and Fitzpatrick 2007, Richards-Zawacki and Cummings 2010, Yang *et al.* 2016), few studies have examined the possibility in more realistic settings (but see Richards-Zawacki *et al.* 2012, Yang *et al.* 2018). Given the prior evidence for sexual selection

on color and the lack of pattern of IBA, my results suggest that the IBA framework developed for natural selection is not useful for distinguishing between drift and sexual selection. Future work should investigate whether sexual selection produces consistent patterns of trait variation and neutral structure.

4.6 Figures and Tables

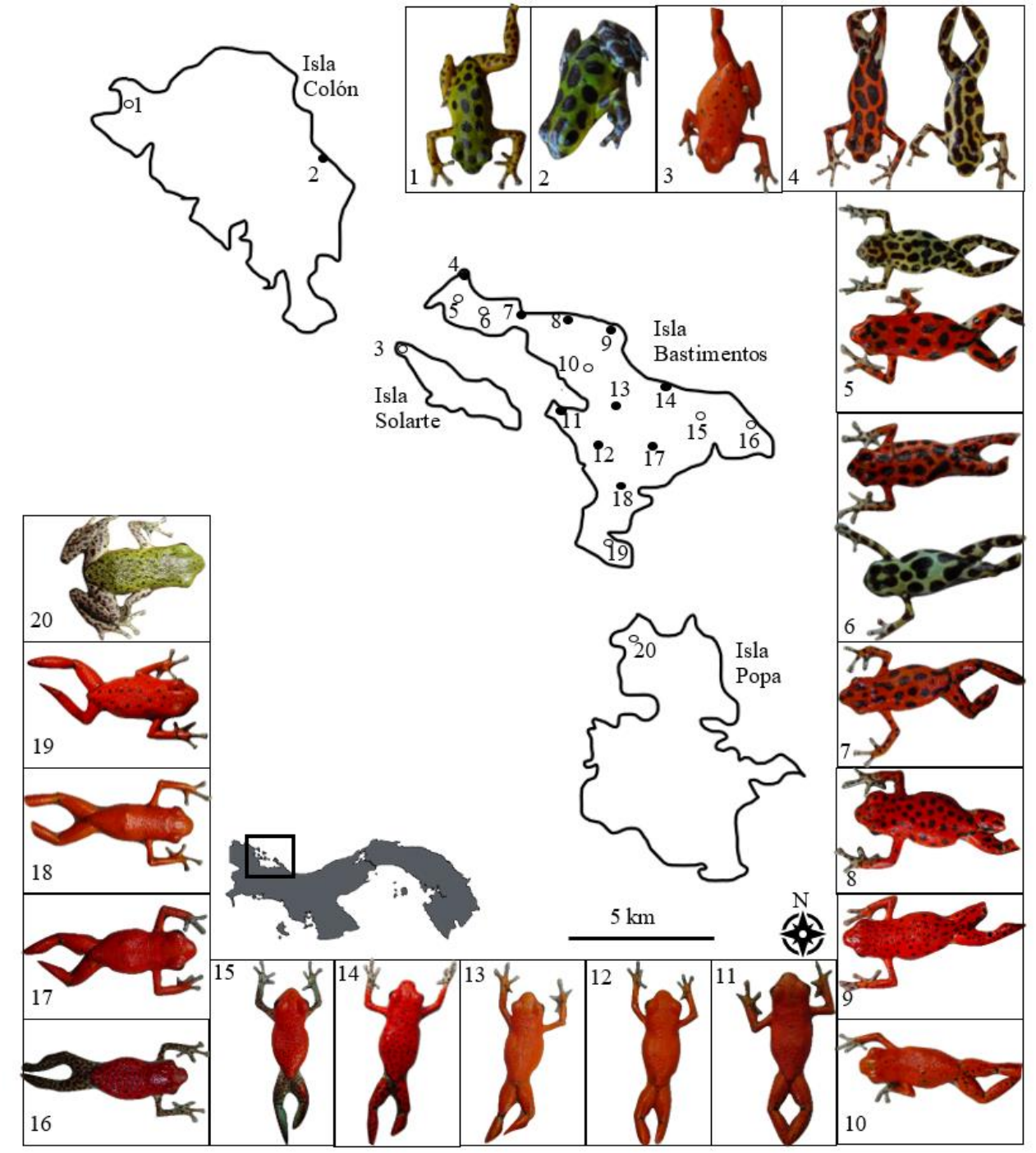


Figure 4-1 Collection Locations and Color Variation Across Bastimentos

Map of collection sites on Isla Bastimentos and neighboring islands in Bocas del Toro, Panama. Frog images show the variation of dorsal color across Bastimentos and adjacent islands. Sampling locations 4-6 are polymorphic. Sampling locations with an open circle (sampling locations 1, 3, 5, 6,10, 15, 16, 19, 20) are where we have collected and analyzed call data.

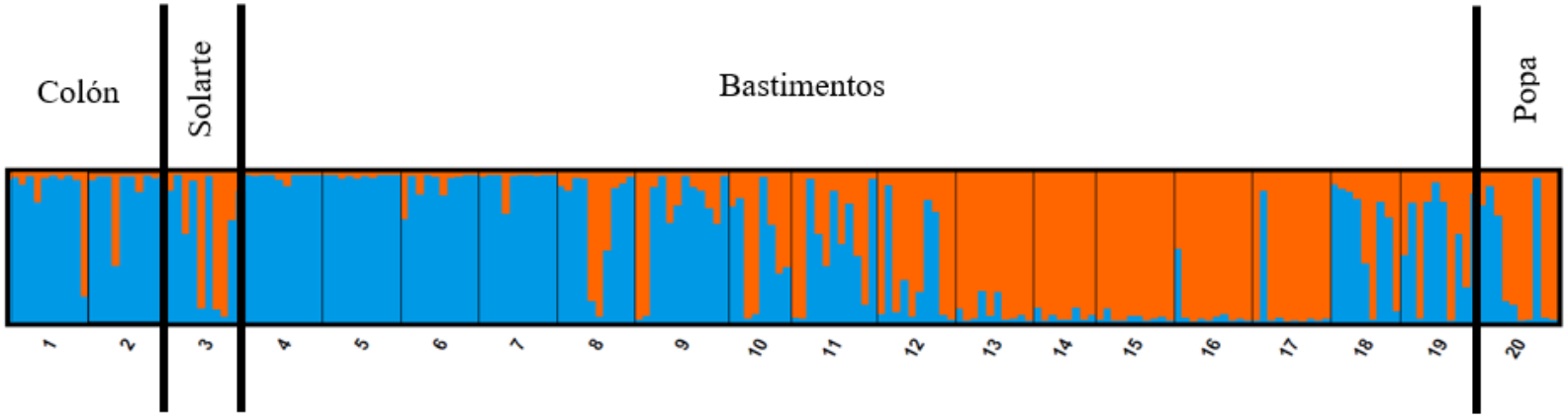


Figure 4-2 Structure Plot

Structure output for K=2 genetic clusters across Isla Bastimentos. Each vertical bar represents one individual's assignment to each genetic cluster. Numbers at the bottom correspond to the sampling locations in Figure 4-1.

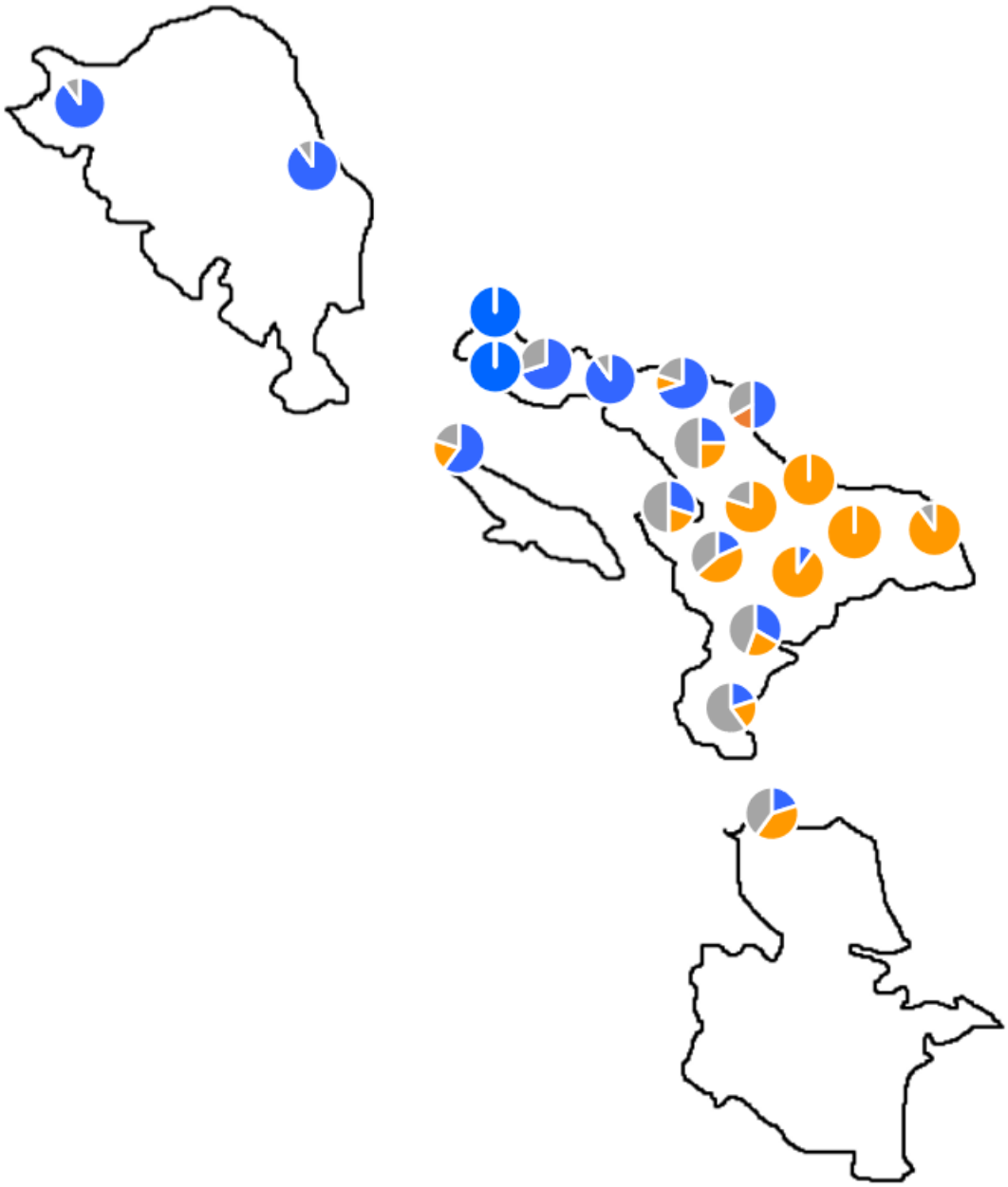


Figure 4-3 Results of Structure Output Visualized on Islands

Each pie chart represents the proportion of individuals sampled at that location that were clearly assigned to one of the two genetic clusters assigned by Structure for K = 2 (blue or orange) or didn't clearly (less than 80% of bar on structure plot) belong in either cluster (grey).

Table 4-1 Summary of Microsatellites

Summary data of microsatellite data across loci at the level of sampling locality. Variables included are the mean, minimum, and maximum for the 20 sampled sites for observed heterozygosity (Hobs), expected heterozygosity (Hexp), number of alleles (N-alleles), and number of individuals with data for that locus (Nind) and the proportion of null alleles (null)

	Hobs				Hexp			N-alleles			Nind			null		
Locus	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean
C11	0.64	0.17	1	0.89	0.80	0.94	8.45	4	11	8.95	6	12	0.06	-0.08	0.23	0.06
E3	0.72	0.00	1	0.76	0.00	1.00	6.75	0	12	5.9	0	10	0.01	-0.07	0.12	0.01
F1	0.82	0.67	1	0.88	0.77	0.97	8.55	6	14	9.15	6	11	0.00	-0.07	0.11	0.001
H5	0.73	0.40	1	44.31	0.80	0.87	9.00	5	14	8.05	4	12	0.01	-0.14	0.11	0.01
C3	0.46	0.00	0.9	0.80	0.00	0.92	6.45	0	11	7.95	0	12	0.27	-0.03	0.78	0.27
D4	0.54	0.00	0.9	0.80	0.00	0.92	7.10	0	12	8.45	0	12	0.18	-0.10	0.74	0.18
B8	0.79	0.25	1	0.90	0.68	1.00	8.70	4	12	8.35	4	12	0.00	-0.08	0.05	0.002
B9	0.47	0.00	0.9	0.87	0.78	0.94	7.45	4	11	8.45	6	10	0.22	-0.01	0.63	0.22

Table 4-2 Pairwise Fst values

Values for pairwise Fs between sampling locations on Bastimentos and the neighboring islands of Bocas del Toro, Panama. Populations 1-2 are from Colon, 3 is from Solarte, 4-19 are from Bastimentos, and 20 is from Popa. (Table is split between pages)

	1	2	3	4	5	6	7	8	9	10
1	0									
2	0.0157	0								
3	0.0490	0.0099	0							
4	0.0000	0.0114	0.0484	0						
5	0.0198	0.0262	0.0524	0.0100	0					
6	0.0109	0.0016	0.0510	0.0223	0.0076	0				
7	0.0121	0.0029	0.0147	0.0006	0.0185	0.0145	0			
8	0.0173	0.0387	0.0787	0.0188	0.0545	0.0361	0.0239	0		
9	0.0235	0.0032	0.0310	0.0268	0.0354	0.0199	0.0203	0.0223	0	
10	0.0298	0.0232	0.0422	0.0400	0.0565	0.0241	0.0164	0.0247	0.0084	0
11	0.0230	0.0051	0.0326	0.0269	0.0290	0.0307	0.0235	0.0330	0.0000	0.0228
12	0.0351	0.0226	0.0547	0.0411	0.0490	0.0212	0.0499	0.0258	0.0209	0.000
13	0.0410	0.0367	0.0598	0.0623	0.0623	0.0318	0.0568	0.0293	0.0234	0.0179
14	0.0419	0.0485	0.0370	0.0449	0.0483	0.0311	0.0511	0.0450	0.0141	0.0204
15	0.0532	0.0556	0.0692	0.0636	0.0765	0.0660	0.0816	0.0580	0.0422	0.0489
16	0.0659	0.0687	0.0737	0.0800	0.0835	0.0533	0.0774	0.0815	0.0290	0.0294
17	0.0567	0.0589	0.0486	0.0706	0.0697	0.0540	0.0721	0.0455	0.0151	0.0476
18	0.0331	0.0014	0.0405	0.0234	0.0211	0.0122	0.0256	0.0281	0.0143	0.0000
19	0.0766	0.0323	0.0892	0.0824	0.0787	0.0422	0.0704	0.0795	0.0400	0.0464
20	0.0441	0.0232	0.0458	0.0198	0.0293	0.0274	0.0363	0.0123	0.0030	0.0195

11	12	13	14	15	16	17	18	19	20
0									
0.0224	0								
0.0141	0.0179	0							
0.0053	0.0261	0.0237	0						
0.0400	0.0192	0.0144	0.0253	0					
0.0369	0.0167	0.0209	0.0423	0.0243	0				
0.0188	0.0246	0.0081	0.0000	0.0056	0.0302	0			
0.0025	0.0025	0.0253	0.0132	0.0663	0.0355	0.0497	0		
0.0532	0.0470	0.0460	0.0655	0.0703	0.0555	0.0611	0.0492	0	
0.0053	0.0098	0.0163	0.0158	0.0139	0.0309	0.0000	0.0000	0.0476	0

5.0 Conclusion

My dissertation aimed to improve our understanding of the evolution of divergence of color signals and their perception using studies of the polymorphic poison frog *O. pumilio*. I tested for differences in opsin expression and sequence and I found evidence for upregulated opsin expression in polymorphic red and yellow frogs that might contribute to a better ability of frogs in that population to discriminate the two local morphs. I also found a handful of opsin mutations, two of which likely have a major effect on color vision, one nonsense mutation early in the SWS1 protein of a cemetery (polymorphic red/yellow population) frog, and one a mutation in the chromophore binding residue of the LWS of a Dolphin Bay (polymorphic red/blue population) frog, although the effects of the other mutations I found remain unclear. Overall, I found little evidence for spectral tuning between morphs, contrasting with results from aquatic systems (Smith *et al.* 2011, Sandkam *et al.* 2015a, Sandkam *et al.* 2015b), which suggests that properties of water as a medium for vision might produce stronger selection on the visual systems compared to air. However, the hypothesized timing of *O. pumilio*'s divergence into distinct color morphs is much more recent than that of the cichlid radiation, suggesting that perhaps *O. pumilio* morphs simply haven't had enough time to diverge in their opsin expression and sequence. Future investigations into the color vision of this system should aim to sequence a larger number of frogs per population in order to test for differences in mutation frequencies among populations and morphs. Additionally, testing whether these differences in the frog's visual system affect the frog's color vision, particularly its ability to detect differences between morphs, should be investigated, perhaps via behavioral assays color discrimination ability or identification of key sites for λ_{max} tuning in O. pumilio opsins. My attempts, early in my PhD, at behaviorally testing for differences

between morphs in their ability to discriminate between colors via the optomotor response failed to produce a consistent response in this species. Therefore, other types of assays should be attempted to try and understand the color discrimination ability of *O. pumilio* and whether there are differences between the populations and morphs. Evidence suggests that *O. pumilio* learn their color-based biases, at least in part, via imprinting on the mother's color as a tadpole (Yang *et al.* 2019). This suggests that tadpoles can see in color, however, we do not know what the tadpole's color vision is like, how early it develops, and when it achieves an adult-like form. Understanding how tadpole color vision develops will improve our understanding of how the observed color-based biases develop as well.

I also found evidence that heritability might differ slightly between three color morphs of *O. pumilio* and that the genetic architectures underlying the dorsal coloration of the morphs differ. I found evidence for differences in the genetic architecture of color between three color morphs of *O. pumilio*. I also found a moderate amount of heritability in color in all three morphs, suggesting that a moderate amount of phenotypic variation is explainable by additive genetic variation as opposed to other sources of variation such as dominance effects, epistasis, or the environment. Since the heritability estimates that I found are similar to those found for sexually selected traits in other species (Hegyi *et al.* 2002, Potti and Canal 2011), this suggests that a moderate amount of heritability is common for sexually selected traits and adds support to the commonly held thought that color is evolving under selection in *O. pumilio*. This study also highlights the complexity of the genetic architecture of coloration, and that different colors of the same species can have different inheritance patterns. It also suggests that different colors of the same species may have different narrow sense heritability values for color and thus differ in their response to selection.

My findings highlight that a controlled breeding approach can provide useful information on the inheritance of complex traits in species that lack the genetic tools for QTL mapping. Additionally, this study provides evidence for the potential for further evolution of a complex trait, and thus the potential for further evolution of differentiation in color between morphs of the same species. Since I found differences between crosses in the inheritance pattern of color, further investigations using controlled breeding experiments with additional morphs will help to elucidate whether different morphs of the same color are inherited similarly (i.e., have simple vs. more additive genetic inheritence). Additionally, an artificial selection experiment could potentially be used to investigate differences in the evolvability of each morph. A better understanding of the number of genes involved in producing color would also help improve our understanding of the genetic basis of color and its inheritance patterns. This could allow for QTL mapping and mutagenesis studies to determine loci involved in color and their affect (rev. in Mackay 2001). However, attempts at mapping color-associated SNPs in *O. pumilio* thus far haven't been able to pinpoint genes associated with color, likely due to the current incompletely assembled genome of this species (Freeborn 2020). These other genetic techniques could become a fruitful avenue for exploring the evolution of coloration in this species if new genome assembly techniques improve our understanding of the frog's genome.

Finally, I examined variation in putatively sexually selected traits and their association with genetic distance in neutral loci and geographic distance. Overall, I did not find a pattern of isolation by adaptation (IBA) in color, a trait known to be under sexual selection in this species, and the only trait in which I did find a pattern consistent with IBA was male body size, a trait not known to be under sexual selection in this species but which has been documented to be evolving under sexual selection another frog species (Wilbur et al. 1978). This suggests that the expectations for

IBA do not seem to work well in distinguishing trait evolution due to sexual selection from trait evolution due to drift. Given the prior evidence for sexual selection on color and the lack of an IBA pattern in color, my results suggest that the IBA framework, which was developed for natural selection may not be an accurate or useful approach for distinguishing between drift and sexual selection. Similar studies on other sexually selected species can help to determine if sexually selected traits produce different patterns of variation than natural selection, or if the patterns found in *O. pumilio* are unusual and the causes for the aberrant pattens should be further investigated.

Overall, my work contributed to our understanding of divergence in animal coloration by providing a detailed case study from a terrestrial, color-polymorphic species. While I didn't find convincing evidence to suggest that opsin expression and/or sequence variation have contributed divergence in coloration, I did find evidence for variation in these variables among indivduals, and especially in polymorphic population. This suggests that the raw material for future divergence in color vision exists among these newly-diverged color morphs. Furthermore, my study adds to our understanding of how the frogs perceive the sexually selected signals that they use in mate choice. My work also provided insight into the genetic architecture of color and highlighted that within a single species, different colors can vary in their patterns of inheritance and heritability. This provides a first step toward understanding and predicting how the response to selection might differ among populations bearing divergent phenotypes. Finally, I showed that comparisons between patterns of phenotypic variation in traits putatively under sexual selection and neutral genetic variation cannot always be expected to follow the pattern of isolation by adaptation, which was developed for the study of traits under natural selection. Taken together, my research highlights that despite limited genetic tools, insights into the evolution of intraspecific divergence in a sexually selected trait can be made for non-model organisms. Furthermore, my work adds to

our understanding of how sexually selected traits diverge in terrestrial species, which have received much less frequently studied than aquatic species to date.

Appendix A Supplemental Material for Chapter 2

Appendix A.1 Linear Model Outputs for Red-Green Comparison

Appendix A.1.1 LWS Opsin Output

```
Call:
lm(formula = LWSdct ~ Color + Population + Color:Population +
    Sex, data = Redgreenanalysis)

Residuals:
    Min      1Q  Median      3Q     Max
-1.0023 -0.4034  0.0102  0.5077  0.9124

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      -2.46152    0.39358  -6.254 1.83e-06 ***
Color            -0.27565    0.62066  -0.444    0.661
Population       -0.07032    0.08169  -0.861    0.398
Sex               0.19971    0.28169   0.709    0.485
Color:Population  0.12233    0.17356   0.705    0.488
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6074 on 24 degrees of freedom
Multiple R-squared:  0.05855,    Adjusted R-squared: -0.09835
F-statistic: 0.3732 on 4 and 24 DF,  p-value: 0.8254
```

Appendix A.1.2 RH1 Opsin Output

```
Call:
lm(formula = RH1dct ~ Color + Population + Color:Population +
    Sex, data = Redgreenanalysis)

Residuals:
```

```
     Min       1Q   Median       3Q      Max
-1.53631 -0.70867 -0.03946  0.64773  1.78440

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       0.003913   0.570852   0.007    0.995
Color             1.262920   0.900209   1.403    0.173
Population        0.125816   0.118485   1.062    0.299
Sex              -0.506550   0.408562  -1.240    0.227
Color:Population -0.494895   0.251733  -1.966    0.061
.
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1

Residual standard error: 0.8809 on 24 degrees of freedom
Multiple R-squared:   0.1775,      Adjusted R-squared:
0.04045
F-statistic: 1.295 on 4 and 24 DF,  p-value: 0.2999
```

Appendix A.1.3 SWS1 Opsin Output

```
Call:
lm(formula = SWS1dct ~ Color + Population +
Color:Population +
    Sex, data = Redgreenanalysis)

Residuals:
     Min       1Q   Median       3Q      Max
-1.42219 -0.61089  0.01092  0.51842  1.17866

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      -2.90544    0.48443  -5.998 3.43e-06
***
Color            -0.41068    0.76393  -0.538    0.596
Population       -0.12879    0.10055  -1.281    0.212
Sex              -0.42160    0.34671  -1.216    0.236
Color:Population -0.08155    0.21362  -0.382    0.706
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1

Residual standard error: 0.7476 on 24 degrees of
freedom
Multiple R-squared:  0.2386,     Adjusted R-squared:
0.1117
F-statistic:  1.88 on 4 and 24 DF,  p-value: 0.1467
```

Appendix A.2 Linear Model Outputs for CEM

Polymorphic/Monomorphic Comparison

Appendix A.2.1 LWS Opsin Output

```
Call:
       lm(formula = LWSdct ~ Color + +Population +
Color:Population +
           Sex, data = Cemanalysis)

       Residuals:
           Min      1Q  Median      3Q     Max
       -1.1643 -0.3846  0.1059  0.3832  1.0263

       Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
       (Intercept)      -1.12203    1.30191  -0.862    0.402
       Color            -0.45523    0.90695  -0.502    0.623
       Population       -0.37221    0.61743  -0.603    0.556
       Sex              -0.50227    0.38292  -1.312    0.209
       Color:Population -0.01303    0.49779  -0.026    0.979

       Residual standard error: 0.685 on 15 degrees of
freedom
       Multiple R-squared:  0.3388,     Adjusted R-squared:
0.1625
       F-statistic: 1.921 on 4 and 15 DF,  p-value: 0.1591
```

Appendix A.2.2 RH1 Opsin Output

```
       Call:
       lm(formula = RH1dct ~ Color + Population +
Color:Population +
           Sex, data = Cemanalysis)
```

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.94975 -0.31113  0.06131  0.29724  0.85268

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -0.2492     0.9734  -0.256   0.8014
Color              0.9481     0.6781   1.398   0.1824
Population         0.5723     0.4617   1.240   0.2341
Sex                0.2887     0.2863   1.008   0.3293
Color:Population  -0.8654     0.3722  -2.325   0.0345
*
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1

Residual standard error: 0.5122 on 15 degrees of
freedom
Multiple R-squared:  0.5216,     Adjusted R-squared:
0.3941
F-statistic: 4.089 on 4 and 15 DF,  p-value: 0.01948
```

Appendix A.2.3 SWS1 Opsin Output

```
Call:
lm(formula = SWS1dct ~ Color + Population +
Color:Population +
    Sex, data = Cemanalysis)

Residuals:
     Min       1Q   Median       3Q      Max
-1.37169 -0.35527  0.03732  0.37891  1.22752

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2.8776     1.4713  -1.956   0.0694 .
Color             -0.3883     1.0250  -0.379   0.7101
Population         0.6213     0.6978   0.890   0.3873
Sex               -0.6217     0.4327  -1.437   0.1714
Color:Population  -0.5664     0.5626  -1.007   0.3300
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1

Residual standard error: 0.7741 on 15 degrees of freedom
Multiple R-squared:  0.446,      Adjusted R-squared:
0.2983
F-statistic: 3.019 on 4 and 15 DF,  p-value: 0.05177
```

Appendix A.3 Linear Model Outputs for DBP

Polymorphic/Monomorphic Comparison

Appendix A.3.1 LWS Opsin Output

```
Call:
lm(formula = LWSdct ~ Color + Population + Color:Population +
    Sex, data = DBPanalysis)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7910 -0.3812  0.1271  0.7676  1.4094

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -6.1922     3.2275  -1.919   0.0694 .
Color              2.4532     1.9007   1.291   0.2115
Population         0.5548     0.6258   0.886   0.3859
Sex                0.7668     0.4912   1.561   0.1342
Color:Population  -0.5106     0.4001  -1.276   0.2165
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.03 on 20 degrees of freedom
Multiple R-squared:  0.1364,    Adjusted R-squared: -0.03638
F-statistic: 0.7894 on 4 and 20 DF,  p-value: 0.5457
```

Appendix A.3.2 RH1 Opsin Output

```
Call:
lm(formula = RH1dct ~ Color + Population + Color:Population +
    Sex, data = DBPanalysis)
```

```
Residuals:
     Min       1Q   Median       3Q      Max
-2.90263 -0.32877  0.09507  0.65828  1.68352

Coefficients:
                  Estimate   Std.   Error   t   value
Pr(>|t|)
(Intercept)        3.5804     3.2963   1.086    0.290
Color             -0.7642     1.9412  -0.394    0.698
Population        -0.5123     0.6392  -0.801    0.432
Sex               -0.3799     0.5017  -0.757    0.458
Color:Population   0.1223     0.4086   0.299    0.768

Residual  standard  error:  1.052  on  20  degrees  of
freedom
Multiple R-squared:  0.1031,    Adjusted R-squared:
-0.07628
F-statistic: 0.5748 on 4 and 20 DF,  p-value: 0.6841
```

Appendix A.3.3 SWS1 Opsin Output

```
Call:
lm(formula = SWS1dct ~ Color + Population + Color:Population +
    Sex, data = DBPanalysis)

Residuals:
     Min       1Q   Median       3Q      Max
-1.54888 -0.50947  0.04075  0.42065  1.34570

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -5.9588     2.5567  -2.331   0.0304 *
Color              2.4766     1.5057   1.645   0.1156
Population         0.5213     0.4958   1.051   0.3056
Sex                0.1143     0.3891   0.294   0.7720
Color:Population  -0.5520     0.3169  -1.742   0.0969 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8162 on 20 degrees of freedom
Multiple R-squared:  0.1965,    Adjusted R-squared:  0.0358
F-statistic: 1.223 on 4 and 20 DF,  p-value: 0.3326
```

Appendix B Supplemental Material for Chapter 3

Appendix B.1 RGB Measure ImageJ Batch Color Analysis Macro

```
// This macro batch processes a folder of images,
// measuring the RGB values separately and with the
// option of specifying an ROI for all of the processed
// images. The optional ROIs should have the same name
// as the corresponding image and an ".roi" extension.
// Press Esc to abort.

  requires("1.33n");
  dir = getDirectory("Choose a Directory ");
  list = getFileList(dir);
  run("Set Measurements...",
     "  mean display redirect=None decimal=3");
  roi = "";
  start = getTime();
  titles = newArray(list.length);
  run("Clear Results");
  setBatchMode(true); // runs up to 20 times faster
  j = 0;
  for (i=0; i<list.length; i++) {
     path = dir+list[i];
     if (endsWith(path, ".roi"))
        roi = path;
     else {
        open(path);
        title = getTitle();
        titles[j++] = title;
        //print(i+"  "+title);
        run("RGB Split");
        measure(""+title+" (red)", roi);
        measure(""+title+" (green)", roi);
        measure(""+title+" (blue)", roi);
     }
  }
  reformatResults(titles);
  //print((getTime()-start)/1000);

  function measure(title, roi) {
     selectImage(title);
     if (roi!="") open(roi);
     run("Measure");
     close();
  }

  function reformatResults(titles) {
     n = nResults/3;
     reds = newArray(n);
     greens = newArray(n);
     blues = newArray(n);
     for (i=0; i<n; i++) {
```

```
        reds[i] = getResult("Mean", i*3);
        greens[i] = getResult("Mean", i*3+1);
        blues[i] = getResult("Mean", i*3+2);
      }
      run("Clear Results");
      for (i=0; i<n; i++) {
        setResult("Label", i, titles[i]);
        setResult("Red", i, reds[i]);
        setResult("Green", i, greens[i]);
        setResult("Blue", i, blues[i]);
      }
       updateResults()
}
```

Appendix B.2 Model Outputs

Appendix B.2.1 Sex Linkage Linear Model Output

```
Iterations = 50001:499981
 Thinning interval  = 20
 Sample size  = 22500

 DIC: 256.7213

 G-structure:  ~animal

       post.mean l-95% CI u-95% CI eff.samp
animal    0.1241  0.01577   0.3104     6048

 R-structure:  ~units

      post.mean l-95% CI u-95% CI eff.samp
units    0.4177   0.2607   0.5865    10379

 Location effects: PC1 ~ Mother_Phenotype

                      post.mean  l-95% CI  u-95% CI  eff.samp
pMCMC
(Intercept)             -0.07347 -0.36646  0.21603     21519
0.616
Mother_PhenotypePopa   -0.15510 -0.55030   0.22133     22500
0.409
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
1
```

Appendix B.2.2 Sexual Dimorphism Linear Model Output

```
Iterations = 50001:499981
 Thinning interval  = 20
 Sample size  = 22500

 DIC: 478.7447

 G-structure:  ~animal

       post.mean  l-95% CI u-95% CI eff.samp
animal   0.04757 0.0009254  0.09329     3523

 R-structure:  ~units

      post.mean l-95% CI u-95% CI eff.samp
units    0.2327   0.1813   0.2858     6132

 Location effects: PC1 ~ Sex + morph + Sex * morph

                  post.mean  l-95%  CI  u-95%  CI  eff.samp
pMCMC
(Intercept)     -2.08727 -2.24326 -1.92674    22500 <4e-
05 ***
SexM               -0.05052 -0.28518  0.18956     24745
0.6730
morphBast        5.88381  5.43267  6.30959    23618 <4e-
05 ***
morphPopa        1.43012  1.21096  1.65685    21474 <4e-
05 ***
SexM:morphBast    0.48121 -0.02362  0.98652      25025
0.0632 .
SexM:morphPopa   -0.18220 -0.46236  0.09523      24026
0.1973
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1
```

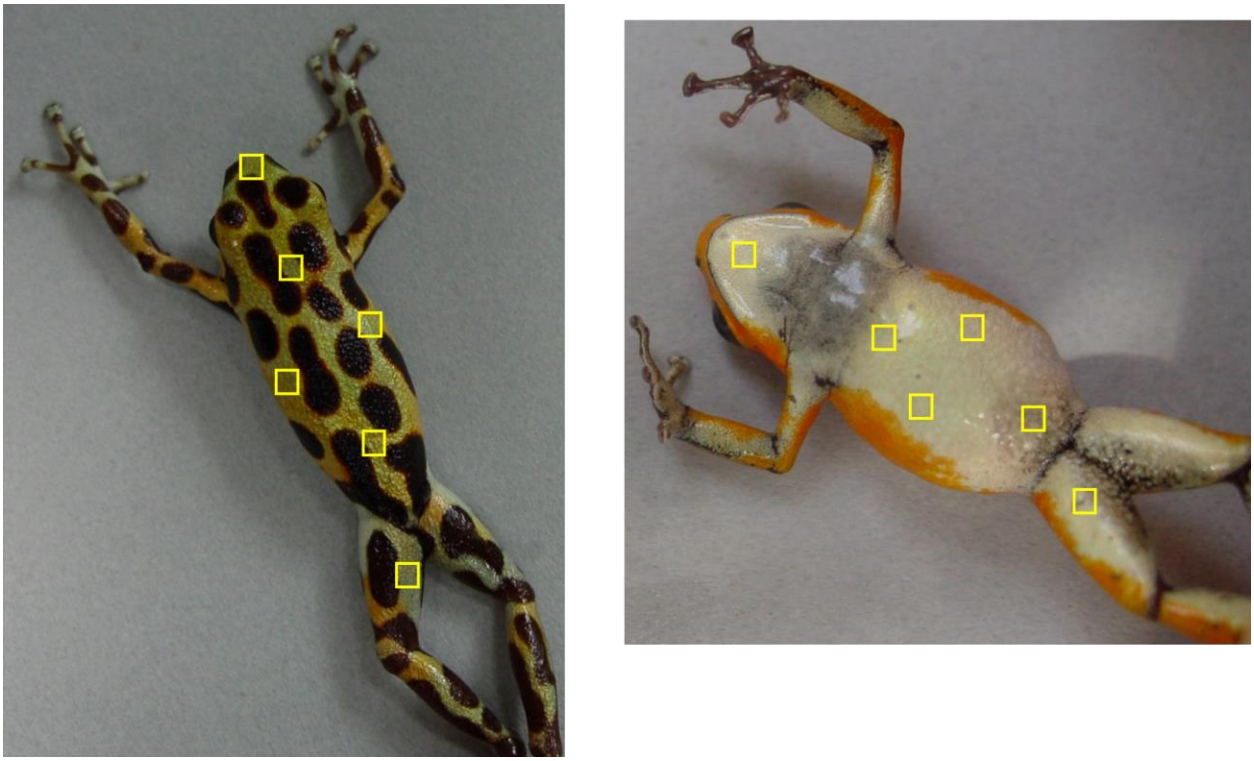
Appendix B.2.3 Color Heritability Animal Model Output

```
Iterations = 10001:99991
 Thinning interval  = 10
 Sample size  = 9000

 DIC: 560.0002
```

```
        G-structure:  ~us(trait):animal

                                     post.mean l-95% CI u-95% CI
eff.samp
       traitPC1_A:traitPC1_A.animal  0.2151768  0.10213  0.34585
1481.1
       traitPC1_P:traitPC1_A.animal  0.0019206 -0.05892  0.06569
1052.2
       traitPC1_B:traitPC1_A.animal -0.0002565 -0.11917  0.12399
927.4
       traitPC1_A:traitPC1_P.animal  0.0019206 -0.05892  0.06569
1052.2
       traitPC1_P:traitPC1_P.animal  0.1220977  0.07327  0.17491
3297.4
       traitPC1_B:traitPC1_P.animal  0.0012137 -0.07261  0.07519
996.1
       traitPC1_A:traitPC1_B.animal -0.0002565 -0.11917  0.12399
927.4
       traitPC1_P:traitPC1_B.animal  0.0012137 -0.07261  0.07519
996.1
       traitPC1_B:traitPC1_B.animal  0.2936366  0.10976  0.53226
484.5

        R-structure:  ~us(trait):units

                                    post.mean  l-95% CI  u-95% CI
eff.samp
       traitPC1_A:traitPC1_A.units   0.285378    0.1755    0.4055
1981.7
       traitPC1_P:traitPC1_A.units   0.002541   -0.1006    0.1130
588.3
       traitPC1_B:traitPC1_A.units -0.004298   -0.1532    0.1417
609.0
       traitPC1_A:traitPC1_P.units   0.002541   -0.1006    0.1130
588.3
       traitPC1_P:traitPC1_P.units   0.201583    0.1595    0.2462
7484.3
       traitPC1_B:traitPC1_P.units -0.003471   -0.1155    0.1051
775.0
       traitPC1_A:traitPC1_B.units -0.004298   -0.1532    0.1417
609.0
       traitPC1_P:traitPC1_B.units -0.003471   -0.1155    0.1051
775.0
       traitPC1_B:traitPC1_B.units   0.288536    0.1191    0.5035
753.3

        Location effects: cbind(PC1_A, PC1_P, PC1_B) ~ trait - 1

                  post.mean l-95% CI u-95% CI eff.samp  pMCMC
       traitPC1_A   -2.0913  -2.2609  -1.9099    911.5 <1e-04 ***
       traitPC1_P   -0.7378  -0.9185  -0.5494   1522.3 <1e-04 ***
       traitPC1_B    4.1689   3.7659   4.5462    259.9 <1e-04 ***
       ---
       Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1
```

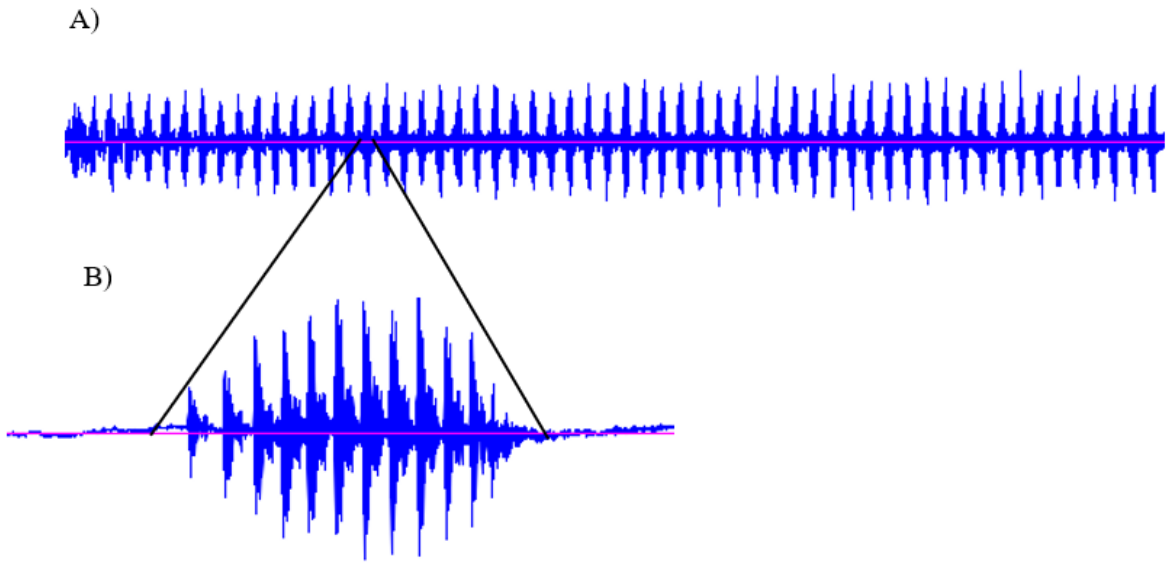
Appendix C Supplementary Material for Chapter 4

Appendix C.1 Supplementary Methods Figures



Appendix Figure 1 Color Sampling Scheme

Placement of the 20 by 20 pixel boxes on the dorsal and ventral sides of the frogs in ImageJ. A) shows an example of a dorsal side with heavy spotting. The boxes were placed to avoid containing as much of the melanistic pattern as possible. B) shows an example of a ventral side with the neck measurement avoiding the male's dark grey throat patch.



Appendix Figure 2 *O. pumilio* Call Diagram

Calls of *O. pumilio* visualized as a waveform in Raven 1.5.0 (Bioacustics Research Program 2014). A) shows a series of notes within a single call bout. B) shows a series of pulses within a single note.

Appendix C.2 Supplementary Results Figures and Tables

Appendix Table 1 Color PCA Results Dorsal

PCA contribution of variables for dorsal side coloration.

Variable	Dim 1	Dim 2
R avg	0.5797	-0.8125
G avg	-0.89101	-0.3555
B avg	-0.9376	-0.1646

Appendix Table 2 Color PCA Results Ventral

PCA contribution of variables for ventral side coloration.

Variable	Dim 1	Dim 2
R avg	0.1908	-0.9816
G avg	-0.9862	-0.1050
B avg	-0.9881	-0.0847

Appendix Table 3 Call Analysis PCA Results

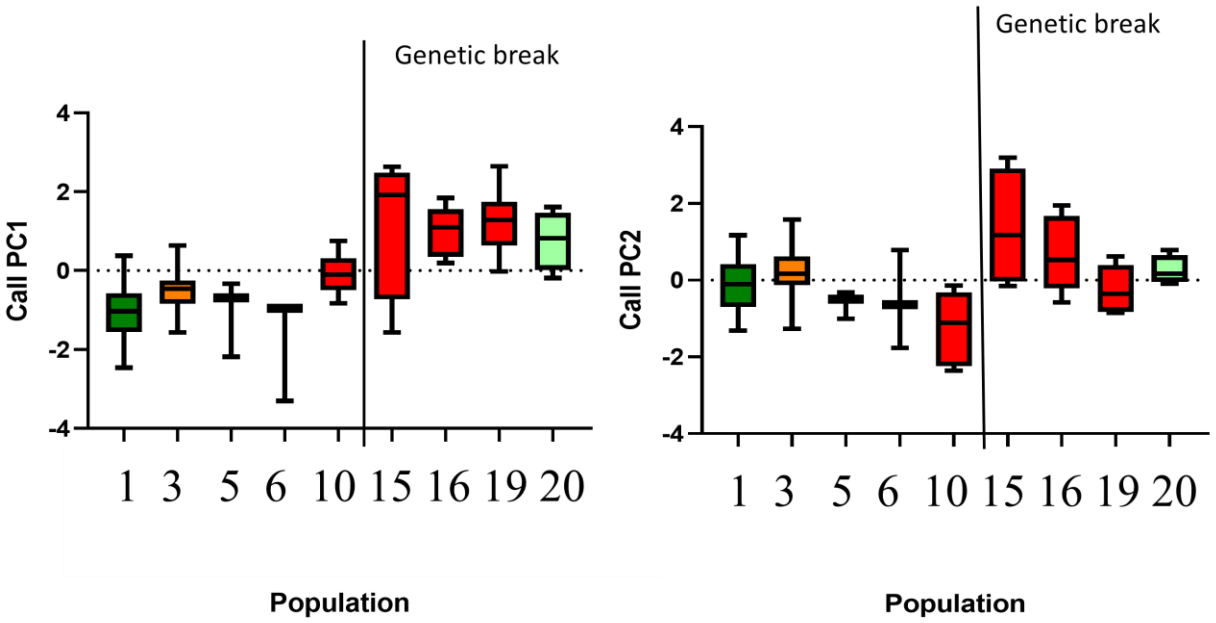
PCA contribution of variables for call analysis

	Dim 1	Dim 2
pulse rate	-0.5449	0.2551
dominant frequency	0.7208	0.2337
call duration	-0.3929	0.7115
note rate	0.6696	0.6696
duty cycle	-0.1069	-0.1069

Appendix Table 4 Microsatellite Locus Statistics

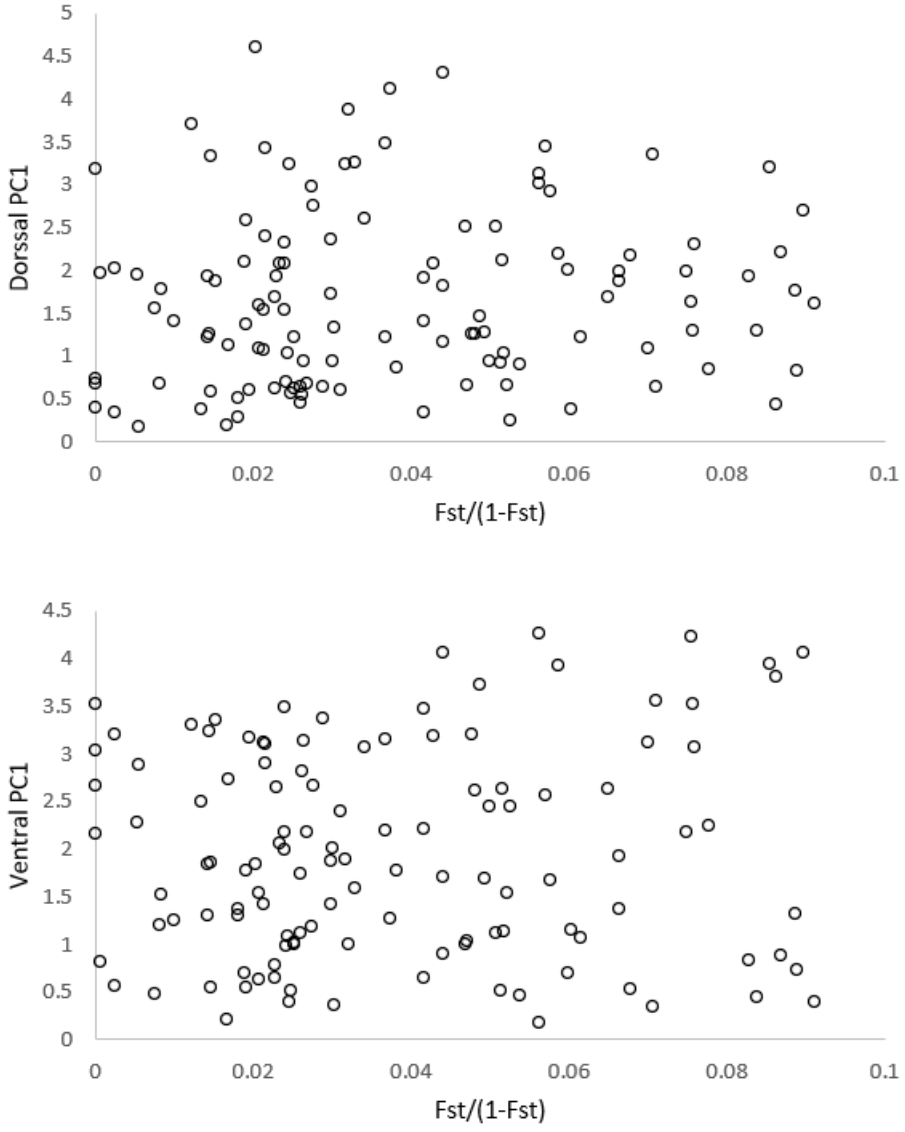
Summary of microsatellite locus statistics. Variables included are, number of alleles (NA), number of individuals (N), observed heterozygosity (Hobs), expected heterozygosity (HExp), polymorphic information content (PIC), significance of Hardy-Weinberg Equilibrium test (HWE), frequency of null alleles (F(Null)), Missing data-number of individuals with missing data for the locus

Locus	N_A	N	HObs	HExp	PIC	HWE	F(Null)	Missing data
C11	26	179	0.665	0.917	0.908	***	0.1592	19
E3	24	118	0.864	0.94	0.933	NS	0.0399	80
F1	26	183	0.825	0.931	0.924	**	0.0589	15
H5	29	161	0.745	0.944	0.938	***	0.1172	37
C3	17	159	0.491	0.898	0.886	***	0.2925	39
D4	19	169	0.58	0.889	0.878	***	0.2121	29
B8	24	167	0.796	0.931	0.924	NS	0.0763	31
B9	17	167	0.479	0.92	0.911	***	0.3136	31



Appendix Figure 3 Call Data by Collection Location

Average values per population for the call analysis for PC1 and PC2. PC1. Dominant frequency and duty cycle were associated with PC1 while call duration and note length were associated with PC2. Population 1 is from Colon, population 3 is from Solarte, populations 5-19 are from Bastimentos, and population 20 is from Popa and this is reflected in the color of the bar.



Appendix Figure 4 Color vs Fst/(1-Fst)

Principal component analysis dimension 1 population differences graphed against corresponding Fst/(1-Fst) values. Top shows dorsal PC1 Mahalanobis distance graphed against

pairwise Fst/(1-Fst) values and bottom shows ventral PC1 Mahalanobis distance graphed against Fst/(1-Fst). PCA for both dorsal and ventral had only one significant eigenvector

Bibliography

Allio, R., S. Donega, N. Galtier, and B. Nabholz. 2017. Large variation in the ratio of mitochondria to nuclear mutation rate across animals: implications for genetic diversity and the use fo mitochondrial DNA as a molecular marker. *Molecular Biology and Evolution* 34 (11): 2762-2772.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410.

Anderson, R. P., and C. O. Handley Jr. 2002. Dwarfism in insular sloths: biogeography, selection, and evolutionary rate. *Evolution* 56(5):1045-1058.

Banci, K. R., A. Eterovic, P. S. Marinho, and O. A. Marques. 2020. Being a bright snake: Testing aposematism and mimicry in a neotropical forest. *Biotropica* 52(6): 1229-1241.

Bates, D., M. Maechler, B. Bolker, S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1-48.

Bennett, D. C., and M. L. Lamoreux. 2003. The color loci of mice-a genetic century. *Pigment Cell Research* 16: 333-344.

Bioacoustics Research Program. 2014. Raven Pro: Interactive Sound Analysis Software (Version 1.5) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology. Available from http://www.birds.cornell.edu/raven.

Blanckenhorn, W. U., C. Mühlhäuser, C. Morf, T. Reusch, and M. Reuter. 2000. Female choice, female reluctance to mate and sexual selection on body size in the dung fly Sepsis cynipsea. *Ethology* 106(7): 577-593.

Bloch, N. I. 2015. Evolution of opsin expression in birds driven by sexual selection and habitat.

Proceedings of the Royal Society B 282: 20142321.

Bloch, N. I., T. D. Price, and B. S. W. Chang. 2015. Evolutionary dynamics of RH2 opsins in birds domonstrate an episode fo accelerated evolution in the New World warblers (*Setophaga*). *Molecular Ecology* 24(10): 2449-2462.

Borg, Å. A., E. Forsgren, and T. Amundsen. 2006. Seasonal change in female choice for male size in the two-spotted goby. *Animal Behaviour* 72(4): 763-771.

Boughman, J. W. 2001. Divergent sexual selection enhances reproductive isolation in sticklebacks. *Letters to Nature* 411:944-948.

Bowmaker, J. K., 2008. Evolution of vertebrate visual pigments. *Vision research* 48(20): 2022-2041.

Brooks, R., and J. A. Endler. 2001. Female guppies agree to differ: phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution* 55(8): 1644-1655.

Brown, J. L., M. E. Maan, M. E. Cummings, and K. Summers. 2010. Evidence for selection on coloration in a Panamanian poison frog: a coalescent-based approach. *Journal of Biogeography* 37: 891-901.

Bybee, S. M., K. K. Johnson, E. J. Gehring, M. F. Whiting, and K. a. Crandall. 2012. All the better to see you with: a review of odonate color vision with transcriptomic insight into the odonate eye. *Organisms, Diversity, and Evolution* 12:241-250.

Coyne, J. A. and H. A. Orr. 2004. *Speciation*. Sinauer Associates Inc., Sunderland, MA.

Cronin, T. W., S. Johnson, N. J. Marshall, and E. J. Warrant. 2014. *Visual Ecology*. Princeton University Press, Princeton, NJ.

Dalziell, A. H., and J. A. Welbergen. 2016. Elaborate mimetic vocal displays by female superb

lyrebirds. *Frontiers in Ecology and Evolution* 4: 34.

de Villemereuil, P. 2012. Estimation of a biological trait heritability using the animal model. How to use the MCMCglmm R package, 1-36.

Dobson, A. E., D. J. Schmidt, and J. M. Hughes. 2019. Heritability of plumage colour morph variation in a wild population of promiscuous, long-lived Australian magpies. *Heredity* 123(3): 349-358.

Dugas, M. B., and C. L. Richards-Zawacki. 2015. A captive breeding experiment reveals no evidence of reproductive isolation among lineages of a polytypic poison frog. *Biological Journal of the Linnean Society* 116(1): 52-62.

Dugas, M. B., and C. L. Richards-Zawacki. 2016. Conspicuous and cryptic morphs of a polytypic poison frog differ in reproductive output because of differences in tadpole performance, not parental effort. *Ethology Ecology & Evolution* 28(4): 441-451.

Dugas, M. B., J. Yeager, and C. L. Richards-Zawacki. 2013. Carotenoid supplementation enhances reproductive success in captive strawberry poison frogs (Oophaga pumilio). *Zoo Biology* 32(6): 655-658.

Endler, J. A. 1992. Signals, signal conditions and the direction of evolution. *The American Naturalist* 139:S125-S153.

Endler, J. A., and A. L. Basolo. 1997. Sensory ecology, receiver biases and sexual selection. *TREE* 13: 415-420.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.

Fain, G.L., H. R. Matthews, M. C. Cornwall, and Y. Koutalos. 2001. Adaptation in vertebrate photoreceptors. *Physiological reviews* 81(1): 117-151.

Finkbeiner, S. D., A. D. Briscoe, and R. D. Reed. 2014. Warning signals are seductive: relative contributions of color and pattern to predator avoidance and mate attraction in *Heliconius* butterflies. *Evolution* 68(12): 3410-3420.

Freeborn, L. 2020. The genetic, cellular, and evolutionary basis of skin coloration in the highly polymorphic poison frog, *Oophaga pumilio.* Unpublished PhD Dissertation. University of Pittsburgh

Gegenfurtner, K. R., and D. C. Kiper. 2003. Color vision. *Annual Review of Neuroscience* 26:181-206.

Google. N.D. [Map of Bocas del Toro, Panama]. Retrieved 26 May, 2016.

Gray, S. M., and J. S. McKinnon. 2007. Linking color polymorphism maintenance and speciation. *Trends in ecology & evolution 22*(2): 71-79.

Hadfield J. D. 2010. "MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package." *Journal of Statistical Software* 33(2): 1–22. https://www.jstatsoft.org/v33/i02/.

Hailman, J. P. 1976. Oildroplets in the eyes of adult anuran amphibians: a comparative survey. *Journal of Morphology*, 148(4): 453-468.

Hansen, T. F. 2006. The evolution of genetic architecture. *Annual Review of Ecology, Evolution, and Systematics* 123-157.

Hargrave, Paul A., J. H. McDowell, D. R. Curtis, J. K. Wang, E. Juszczak, S. Fong, J. K. M. Rao, and P. Argos. 1983. The structure of bovine rhodopsin. *Biophysics of structure and mechanism* 9(4): 235-244.

Hasegawa, A., M. Soma, and T. Hasegawa. 2011. Male traits and female choice in Java sparrows: preference for large body size. Ornithological Science, 10(1), 73-80.

Hartman, S. J., I. Menon, K. Haug-Collet, and N. J. Colley. 2001. Expression of rhodopsin and arrestin during the light-dark cycle in Drosophila. *Molecular Vision* 7:95-100.

Hauswaldt, J. S., A-K. Ludewig, S. Hagemann, H. Pröhl, M. Vences. Ten microsatellite loci fot eh strawberry poison frog (*Oophaga pumilio*). *Conservation Genetics.* DOI 10.1007/s10592-009-9859-3.

Hegna, R. H., R. A. Saporito, and M. A. Donnelly. 2013. Not all colors are equal: predation and color polytypism in the aposematic poison frog *Oophaga pumilio*. *Evolutionary Ecology* 27(5): 831-845.

Hegyi, G., J. Török, and L. Tóth. 2002. Qualitative population divergence in proximate determination of a sexually selected trait in the collared flycatcher. Journal of *Evolutionary Biology* 15(5): 710-719.

Hindar K., and B. Jonsson. 1993. Ecological polymkorphism in Arctic charr. *Biological Journal of the Linnean Societey.* 48:63-74.

Howard, R. D., and A. G. Kluge. 1985. Proximate mechanisms of sexual selection in wood frogs. *Evolution* 39(2): 260-277.

Huang, S-C., T-H. Chiou, J. Marshall, and J. Reinhard. 2014. Spectral sensitivities and color signals in a polymorphic damselfly. *Plos One* 9(1):e87972.

Huang, W., and T. F. Mackay. 2016. The genetic architecture of quantitative traits cannot be inferred from variance component analysis. *PLoS genetics* 12(11): e1006421.

Jefferson, L., and R. Harvey. 2006. Accommodating color blind computer users. *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility* pp 40-47.

Kalinowski, S.T., M. L. Taper, and T. C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099-1106.

Kamphuis, W., C. Cailotto, F. Dijk, A. Bergen, and R. M. Buijs, 2005. Circadian expression of clock genes and clock-controlled genes in the rat retina. *Biochemical and biophysical research communications* 330(1): 18-26.

Kearns, J. A., E. J. Cargill, L. A. Clark, S. I. Candille, T. G. Barryere, M. Olivier, G. Lust, R. J. Todhunter, S. M. Schmutz, K. E. Murphy, and G. S. Barsh. 2007. Linkage and segregation analysis of black and brindle coat color in domestic dogs. *Genetics* 176: 1679-1689.

Kearse, M., R. Moir, A. Wilson., S. Stones-Havas, S. Cheung, M., Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, R. Mentjies, & A. Drummond, 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12): 1647-1649.

Kiriazis, J., and C. N. Slobodchikoff, C. N. 2006. Perceptual specificity in the alarm calls of Gunnison's prairie dogs. *Behavioural Processes* 73(1): 29-35.

Kokko, H., R. Brooks, M. D. Jennions, and J. Morley. 2003. The evolution of mate choice and mating biases. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(1515): 653-664.

Kopelman, N. M; J. Mayzel; M. Jakobsson; N. A. Rosenberg, I. Mayrose. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15(5): 1179-1191,

Koressaar, T., and M. Remm. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23:1289–1291.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolultionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.

Kunte, K. 2009. The diversity and evolution of Batesian mimicry in *Papilio* swallowtail butterflies. *Evolution* 63(10): 2707-2716.

Loyau, A., M. Saint Jalme, C. Cagniant, and G. Sorci. 2005. Multiple sexual advertisements honestly reflect health status in peacocks (*Pavo cristatus*). *Behavioral Ecology and Sociobiology* 58(6): 552-557.

Li, P., S. Temple, Y. Gao, T. J. Haimberger, C. W. Hawryshyn, and L. Li. 2005. Circadian rhythms of behavioral cone sensitivity and long wavelength opsin mRNA expression: a correlation study in zebrafish. *The Journal of Experimental Biology* 208: 497-504.

Liebman, P. A., and G. Entine. 1968. Visual pigments of frog and tadpole (*Rana pipiens*). *Vision Research* 8(7): 761-IN7.

Maan, M. E. and M. E. Cummings. 2008. Female preferences for aposematic signal components in a polymorphic poison frog. *Evolution* 62: 2334-2345.

Maan, M. E. and M. E. Cummings. 2012. Poison frog colors are honest signals of toxicity, particularly for bird predators. *The American Naturalist* 179:E1–E14.

Maan, M. E., K. D. Hofker, J. J. M. van Alphen, and O. Seehausen. 2006. Sensory drive in cichlid speciation. *The American Naturalist* 167: 947-954.

Maan, M. E. and O. Seehausen. 2011. Ecology, sexual selection and speciation. *Ecology Letters* doi: 10.1111/j.1461-0248.2011.01606.x.

Mackay, T. F. 2001. The genetic architecture of quantitative traits. *Annual Review of Genetics* 35: 303-339.

Mappes, J., N. Marples, and J. A. Endler. 2005. The complex business of survival by aposematism.

Trends in Ecology & Evolution 20(11): 598-603.

McGinnis, J. F., J. P. Whelan, and L. A. Donoso. 1992. Transient, cyclic changes in mouse visual cell gene products during the light-dark cycle. *Journal of Neuroscience Research*, 31(3): 584-590.

McLean, C. A. and D. Stuart-Fox. 2014. Geographic variation in animal colour polymorphisms and its role inn speciation. *Biological Reviews* 89: 860-873.

Morales, J., S. Y. Kim, E. Lobato, S. Merino, G. Tomás, J. Martínez-de la Puente, and J. Moreno. 2010. On the heritability of blue-green eggshell coloration. *Journal of Evolutionary Biology* 23(8): 1783-1791.

Myers, C. W., and J. W. Daly. 1976. Preliminary evaluation of skin toxins and vocallizations in taxonomic and evolutionary studies of poison-dart frogs (Dendrobatidae). *Bulletin of the American Museum of Natural History* 157:174-262.

Nali, R. C., K. R. Zamudio, C. F. Haddad, and C. P. Prado. 2014. Size-dependent selective mechanisms on males and females and the evolution of sexual size dimorphism in frogs. *The American Naturalist* 184(6): 727-740.

Ng, J., E. L. Landeed, R. M. Logsdon, and R. E. Glor. 2012. Correlation between *Anolis* Lizard dewlap phenotype and environmental variation indicates adaptive divergence of a signal important to sexual selection and species recognition. *Evolution* 67(2):573-582.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* 18(3): 375-402.

Oliveira, R. F., A. V. Canario, and M. S. Grober. 2001. Male sexual polymorphism, alternative reproductive tactics, and androgens in combtooth blennies (Pisces: Blenniidae). *Hormones and Behavior* 40(2): 266-275.

Panhuis, T. M., R. Butlin, M Zuk, and T. Tregenza. 2001. Sexual selection and speciation *Trends in Ecology and Evolution* 16: 364-371.

Patricelli, G. L., and J. L. Blickley. 2006. Avian communication in urban noise: Causes and consequences of vocal adjustment. *The Auk* 123(3): 639-649.

Pfaffl, M. W. 2001. A new mathematical model for relative quantitation in real-time RT-PCR. Nucleic Acids Research 29(9).

Pfennig, D. W., and S. P. Mullen. 2010. Mimics without models: causes and consequences of allopatry in Batesian mimicry complexes. *Proceedings of the Royal Society B: Biological Sciences* 277(1694): 2577-2585.

Potti, J., and D. Canal. 2011. Heritability and genetic correlation between the sexes in a songbird sexual ornament. *Heredity* 106(6): 945-954.

Pröhl, H. And O. Berke. 2001. Spatial distributions of male and female strawberry poison frogs and their relation to female reproductive resources. *Oecologia* 129:534-542.

Pröhl, H., S. Hagemann, J. Karsch, and G. Höbel. 2007. Geographic variation in male sexual signals in strawberry poison frogs (*Dendrobates pumilio*). *Ethology* 113:825-837.

Price, T. D. 2017. Sensory drive, color, and color vision. *The American Naturalist* 190(2): 157-170.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rasmussen, R. 2001. Quantification on the LightCycler. In *Rapid cycle real-time PCR*. Springer, Berlin, Heidelberg, 2001, pg 21-34.

Raymond M. & F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. *Heredity* 86: 248-249.

Reynolds, R. G. and B. M. Fitzpatrick. 2007. Assortative mating in poison-dart frogs based on an ecologically important trait. *Evolution* 61: 2253-2259.

Richards-Zawacki, C. L. and M. E. Cummings. 2010. Intraspecific reproductive character displacement in a polymorphic poison dart frog, *Dendrobates pumilio*. *Evolution* 65: 259-267.

Richards-Zawacki, C. L., I. J. Wang, and K. Summers. 2012. Mate choice and the genetic basis for colour variation in a polymorphic dart frog: inferences from a wild pedigree *Molecular Ecology* 21: 3879-3892.

Richards-Zawacki, C. L., J. Yeager, and H. P. Bart. 2013. No evidence for differential survival or predation between sympatric color morphs of an aposematic poison frog. *Evolutionary Ecology* 27(4): 783-795.

Rieder, S.. 2009. Molecular tests for coat colours in horses. *Journal of Animal Breeding Genetics* 126: 415-424.

Ritchie, M. G. 2007. Sexual selection and speciation. *Annual Review of Ecology Evolution and Systematics* 38:79-102.

Robinson, P. R., J. Buczyłko, H. Ohguro, and K. Palczewski. 1994. Opsins with mutations at the site of chromophore attachment constitutively activate transducin but are not phosphorylated by rhodopsin kinase. *Proceedings of the National Academy of Sciences* 91(12): 5411-5415.

Rogers, R. L., L. Zhou, C. Chu, R. Marquez, A. Corl, T. Linderoth, L. Freeborn, M. D. MacManes, Z. Xiong, J. Zheng, C. Guo, X. Xun, M. R. Kronforst, K. Summers, Y. Wu, H. Yang, C. L.

Richards-Zawacki, G. Zhang, and R. Nielsen. 2018. Genomic takeover by transposable elements in the strawberry poison frog. *Molecular Biology and Evolution* 35:2913–2927.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145:1219-1228.

Rousset, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resources* 8: 103-106.

Rudh, A., B. Rogell, O. Hastad, and A. Qvarnstrom. 2010. Rapid population divergence linken with co-variation between coloration and sexual display in strawberry poison frogs. *Evolution* 65(5): 1271-1282.

Ryan, M. J., and M. E. Cummings. 2013. Perceptual biases and mate choice. *Annual Review of Ecology, Evolution, and Systematics* 44: 437-459.

Ryan, M. J., C. M Pease, M. R. Morris. 1992. A genetic polymorphism in the swordtail *Xiphophorus nigrensis*: testing the prediction of equal fitness. *The American Naturalist* 139(1): 21-31.

Sandkam, B., C. M. Young, and F. Breden. 2015a. Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). *Molecular Ecology* 24: 596-609.

Sandkam, B. A., C. M. Young, F. M. W. Breden, G. R. Bourne, F. Breden. 2015b. Color vision varies more among populations than among species of live-bearing fish from South America. *BMC Evolutionary Biology* 15:225.

Saporito, R. A., R. Zuercher, M. Roberts, K. G. Gerow, and M. A. Donnelly. 2007. Experimental evidence for aposematism in the Dedrobatid poison frog *Oophaga pumilio*. *Coperia* 2007(4): 1006-1011.

Schartl, M., L. Larue, M.. Goda, M. W. Bosenberg, H. Hashimoto, and R. N. Kelsh. 2015. What is a vertebrate pigment cell?. *Pigement Cel Melanoma Research* 29:8-14.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak and A. Cardona. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods* 9(7): 676–682. doi:10.1038/nmeth.2019.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature methods* 9(7): 671-675.

Servedio, M. R., and J. W. Boughman. 2017. The role of sexual selection in local adaptation and speciation. *Annual Review of Ecology, Evolution, and Systematics* 48:84-109.

Siddiqi, A., T. W. Cronin, E. R. Loew, M. Vorobyev, and K. Summers. 2004. Interspecific and intraspecific views of color signals in the strawberry poison frog *Dendrobates pumilio*. *Journal of Experimental Biology* 207: 2471-2485.

Smith, A. R., L. D'Annunzio, A. E. Smith, A. Sharma, C. M. Hofmann, N. J. Marshall, and K. L Carleton. 2011. Intraspecific cone opsin expression variation in the cichlids of Lake Malawi *Molecular Ecology* 20:299-310.

Stapley, J., C. Wordley, and J. Slate. 2010. No evidence of genetic differtiation bewteen anoles with different dewlap color patterns. *Journal of Heredity* 102:118-124.

Summers. K., T. W. Cronin, and T. Kennedy. 2003. Variation in spectral reflectance among populations of *Dendrobates pumilio*, the strawberry poison frog, in the Bocas del Toro Archipelago, Panama. *Journal of Biogeography* 30:35-53.

Tkalčič, M. and J. F. Tasič. 2003. Colour spaces-perceptual, historical, and application background. IEEE (1): 304-308).

Tsubaki, Y. 2003. The genetic polymorphism linked to mate-securing strategies in the male damselfly *Mnais costalis* Selys (Odonata: Calopterygidae). *Population Ecology* 45(3): 263-266.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Res*. 40:1–12.

Uy, J. A. C., and G. Borgia. 2000. Sexual selection drives rapid divergence in bowerbird display traits. *Evolution 54*(1): 273-278.

Van Oosterhout, C., W. F. Hutchinson, D. P.Wills, & P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources*, *4*(3): 535-538.

Vidya, T. N. C., and R Sukumar. 2005. Social and reproductive behaviour in elephants. *Current Science* 1200-1207.

Visscher, P. M., W. G. Hill, and N. R. Wray. 2008. Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics* 9(4): 255-266.

Wan, Y. C., M. J. Navarrete, L. A. O'Connell, L. H. Uricchio, A. Roland, M. E. Maan, S. R. Ron, M. Betancourth-Cundar, M. R. Pie, K. A. Howell, C. L. Richards-Zawacki, M. E. Cummings, D. C. Cannatella, J. C. Santos, R. D. Tarvin. *In review*. Diurnality shapes the visual opsin genes of colorful Neotropical frogs

Wang, I. J. and H. B. Shaffer. 2008. Rapid color evolution in an aposematic species: a phylogenetic analysis of color variation in the strikingly polymorphic strawberry poison-dart frog. *Evolution* 62: 2742-2759.

Wang, I. J. and K. J. Summers 2010. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* 19: 447-458.

Weaver, R. J., E. S. Santos, A. M. Tucker, A. E. Wilson, and G. E. Hill, G. E. 2018. Carotenoid metabolism strengthens the link between feather coloration and individual quality. *Nature communications* 9(1): 1-9.

Williams, L. J., A. J. King, and C. Mettke-Hofmann. 2012. Colourful characters: head colour reflects personality in a social bird, the Gouldian finch, *Erythrura gouldiae*. *Animal Behavior* 84(1): 159-165.

Yang, Y., C. L. Richards-Zawacki, A. Devar, M. B. Dugas. 2016. Poison frog color morphs express assortative mate preferences in allopatry but not sympatry. *Evolution* 70: 2778-2788.

Yang, Y., M. B. Dugas, H. J. Sudekum, S. N. Murphy, and C. L. Richards-Zawacki. 2018. Male–male aggression is unlikely to stabilize a poison frog polymorphism. *Journal of Evolutionary Biology*, *31*(3), 457-468.

Yang, Y. S. Blomenkamp, M. B. Dugas, C. L. Richards-Zawacki, and H. Pröhl. 2019a. Mate choice versus mate preference: Inferences about color-assortative mating differ between field and lab assays of poison frog behavior. *The American Naturalist* 193(4): 589-607.

Yang, Y., M. R. Servedio, C. L. Richards-Zawacki. 2019b. Imprinting sets the stage for speciation. *Nature* 574 (7776): 99-102.

Yeager, J. 2015. Causes and consequences of warning color variation in a polytypic poison frog. Unpublished PhD dissertation. Tulane University.

Yewers, M. S., C. A. McLean, A. Moussalli, D. Stuart-Fox, T. D. Bennett, and B. Knott. 2015. Spectral sensitivity of cone photoreceptors and opsin expression in two colour-divergent lineages of the lizard *Ctenophorus decresii*. *The Journal of Experimental Biology* 218:1556-1563.

Yokoyama. 2008 Evolution of dim-light and color vison pigments. *Annual Reviews of Genomics and Human Genetics* 9:259-282.

Yokoyama, S, and F. B. Radlwimmer. 1998. The “five sites” rule and the evolution of red and green color vision in mammals. *Molecular Biology and Evolution* 15(5): 560-567.

Yu, C. J., Y. Gao, P. Li, and L. Li. 2007. Synchronizing multiphasic circadian rhythms of rhodopsin promoter expression in rod photoreceptor cells. *Journal of Experimental Biology* 210(4): 676-684.

Zhang, L., X. Ma, J. Jiang, and X. Lu. 2012. Stronger condition dependence in female size explains altitudinal variation in sexual size dimorphism of a Tibetan frog. *Biological Journal of the Linnean Society* 107:558-565.

Zimova, M., K. Hackländer, J. M. Good, J. Melo-Ferreira, P. C. Alves, and L. S. Mills. 2018. Function and underlying mechanisms of seasonal colour moulting in mammals and birds: what keeps them changing in a warming world?. *Biological Reviews* 93(3): 1478-1498.