

ORIGINAL RESEARCH

Phenotypic and genetic diversity in aposematic Malagasy poison frogs (genus *Mantella*)

Karina Klonoski^{1,2}  | Ke Bi^{2,3} | Erica Bree Rosenblum^{1,2}¹Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, California²Museum of Vertebrate Zoology, University of California, Berkeley, Berkeley, California³Computational Genomics Resource Laboratory (CGRL), California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, Berkeley, California**Correspondence**Karina Klonoski, Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA.
Email: kklonosk@berkeley.edu**Funding information**

National Geographic Society, Grant/Award Number: 9469-14; Museum of Vertebrate Zoology at UC Berkeley, Grant/Award Number: No number provided; Mohamed bin Zayed Species Conservation Fund, Grant/Award Number: 150511309

Abstract

Intraspecific color variation has long fascinated evolutionary biologists. In species with bright warning coloration, phenotypic diversity is particularly compelling because many factors, including natural and sexual selection, contribute to intraspecific variation. To better understand the causes of dramatic phenotypic variation in Malagasy poison frogs, we quantified genetic structure and color and pattern variation across three closely related species, *Mantella aurantiaca*, *Mantella crocea*, and *Mantella milotympanum*. Although our restriction site-associated DNA (RAD) sequencing approach identified clear genetic clusters, they do not align with current species designations, which has important conservation implications for these imperiled frogs. Moreover, our results suggest that levels of intraspecific color variation within this group have been overestimated, while species diversity has been underestimated. Within major genetic clusters, we observed distinct patterns of variation including: populations that are phenotypically similar yet genetically distinct, populations where phenotypic and genetic breaks coincide, and populations that are genetically similar but have high levels of within-population phenotypic variation. We also detected admixture between two of the major genetic clusters. Our study suggests that several mechanisms—including hybridization, selection, and drift—are contributing to phenotypic diversity. Ultimately, our work underscores the need for a reevaluation of how polymorphic and polytypic populations and species are classified, especially in aposematic organisms.

KEYWORDScolor evolution, hybridization, *Mantella*, poison frogs, polymorphism, population genomics

1 | INTRODUCTION

The origin and maintenance of intraspecific color variation and its consequences for speciation have long captivated evolutionary biologists (e.g., Huxley, 1955; Endler, 1980; Gray & McKinnon, 2007;

Corl, Davis, Kuchta, & Sinervo, 2010; Hugall & Stuart-Fox, 2012). Examples of color variation in nature are widespread and have been documented across many diverse taxonomic groups, including reptiles, fish, mammals, birds, and invertebrates (e.g., Sandoval, 1994; Sinervo, Bleay, & Adamopoulou, 2001; Galeotti, Rubolini, Dunn, &

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

Fasola, 2003; Hoekstra, Drumm, & Nachman, 2004; Maan et al., 2008). Previous work has indicated that color variation observed within or between populations may be mediated by natural selection (e.g., apostatic selection, divergent selection in different substrate or light environments as in Sandoval, 1994; Hoekstra et al., 2004), sexual selection (e.g., mate choice, variation in male mating strategies as in Sinervo & Lively, 1996; Kingston, Rosenthal, & Ryan, 2003), genetic drift (Hoffman, Schueler, Jones, & Blouin, 2006), or some combination of these factors (e.g., Endler, 1991; Oxford, 2005; Reynolds & Fitzpatrick, 2007). Intraspecific color variation can also potentially give rise to new species, especially when natural or sexual selection reduces gene flow between alternative morphs (e.g., Rosenblum & Harmon, 2011).

Examples of color variation in aposematic organisms—where conspicuous warning signals advertise toxicity or unpalatability to predators (Poulton, 1890)—are particularly compelling. Aposematic colors are often highly contrasting, variable, and potentially exhibit trade-offs between natural and sexual selection (Crothers & Cummings, 2013; Estrada & Jiggins, 2008; Jiggins, Naisbit, Coe, & Mallet, 2001; Nokelainen, Hegna, Reudler, Lindstedt, & Mappes, 2012; Reynolds & Fitzpatrick, 2007; Stevens & Ruxton, 2012; Summers, Symula, Clough, & Cronin, 1999). Historically, variation in aposematic signals has been considered perplexing from a theoretical perspective because phenotypic diversity is expected to be highly constrained in such systems due to positive frequency-dependent selection via predation (Briolat et al., 2018; Endler & Greenwood, 1988; Mallet & Joron, 1999). Once a predator has learned to associate toxicity with a particular phenotype, protection should be conferred to those organisms displaying a similar phenotype, encouraging uniformity in warning coloration. According to theoretical predictions of predator avoidance learning, novel phenotypes should be unrecognizable to predators as toxic and thus quickly removed from populations (Guilford & Dawkins, 1993; Mallet & Barton, 1989; Müller, 1879). In recent years, however, studies have identified many biotic and abiotic factors—including intraspecific communication, parasite load, temperature and variability in predator learning and sensory abilities—that contribute to variation in aposematic signals (reviewed in Briolat et al., 2018). Growing awareness of the variety of selective factors influencing aposematic coloration has led scientists to encourage a more holistic approach when investigating diversity in warning coloration and to consider the range of factors that may be at play in maintaining phenotypic variation (Briolat et al., 2018).

Studies of color variation within aposematic species have traditionally focused on systems demonstrating either multiple color morphs within a population (i.e., polymorphism) or geographic color variation among populations (i.e., polytypism) (reviewed in Briolat et al., 2018). In systems with high color variability, however, determining whether color variants represent different species, different populations, or different morphs within populations is difficult particularly when genetic structure is not well resolved. Distinguishing between species, populations, and morphs can be especially challenging in phenotypically diverse poison frog groups, where high rates of phenotypic variation can confound our understanding of

species delimitations (Posso-Terranova & Andrés, 2018; Roland et al., 2017; Tarvin, Powell, Santos, Ron, & Cannatella, 2017).

In many instances, the inability to distinguish whether color variation occurs within populations, between populations, or between species is further compounded by nonexistent or limited genetic datasets, which lack the resolution needed to clarify relationships among populations. Tarvin et al. (2017) recently demonstrated that the level of interspecific mitochondrial divergence among four distinct poison frog species was comparable to the divergence levels observed between populations considered to be a single polymorphic species (Hauswaldt, Ludwig, Vences, & Pröhl, 2011). Noting the limitations of mitochondrial data in resolving species boundaries, the authors explicitly called for genome-level studies, in combination with information on phenotypic diversity and natural history, to understand relationships in these complicated systems (Tarvin et al., 2017). The power of more comprehensive genetic datasets to resolve genetic structure in poison frogs has been demonstrated in recent studies where Neotropical poison frogs considered to be a single species in fact contained multiple genetic lineages that potentially represent new species (Posso-Terranova & Andrés, 2018; Roland et al., 2017).

While the relationship between phenotypic and genetic diversity has been extensively studied in the Neotropical poison frogs (e.g., Wang & Summers, 2010; Twomey et al., 2013; Roland et al., 2017; Tarvin et al., 2017), there is an entirely separate radiation of poison frogs in which color diversity has never been examined. Endemic to Madagascar, the *Mantella* genus describes sixteen species of toxic, diurnal frogs exhibiting variable coloration and pattern both within and among species (Glaw & Vences, 2007). Commonly called Malagasy poison frogs, the bright coloration displayed by many species within this group is presumed to be aposematic. The toxic skin alkaloids found in *Mantella* species are believed to be derived from arthropod prey (Daly, Andriamaharavo, Andriantsiferana, & Myers, 1996; Daly, Garraffo, Hall, & Cover, 1997), similar to the Neotropical poison frogs, and variation in alkaloid composition has been observed among species, populations, and habitats (Andriamaharavo et al., 2015; Daly et al., 2008). The mechanism of chemical defense is hypothesized to have evolved convergently in Neotropical and Malagasy poison frogs (Clark, Raxworthy, Rakotomalala, Sierwald, & Fisher, 2005). Despite their high degree of phenotypic variation and apparent similarity to the Neotropical poison frogs (Heying, 2001b; Rojas, 2016), little is known about the natural history, ecology, and genetic background of Malagasy poison frogs.

Within the *Mantella* genus, one complex of three closely related species, *Mantella aurantiaca*, *Mantella crocea*, and *Mantella milotympanum*, demonstrates a particularly high degree of variability in conspicuous coloration and pattern. Found in the rainforests of central-eastern Madagascar, the geographic range of all three species is highly restricted and patchy in its distribution (Bora et al., 2008). Among and within populations, there is exceptional phenotypic variation both in dorsal coloration, which ranges from red to green at the extremes, and in patterning elements, which are variable in the degree of ventral spotting and black banding present on the side. Yet,

any attempt to understand the phenotypic diversity in this group is hindered by the unresolved taxonomy of these species. Previous genetic work is limited to a handful of mitochondrial DNA and allozyme studies, which have yielded somewhat confusing results (Chiari et al., 2004; Schaefer, Vences, & Veith, 2002; Vences, Chiari, Raharivoloniaina, & Meyer, 2004; Vences, Hille, & Glaw, 1998). *Mantella aurantiaca*, *M. milotympanum*, and *M. crocea* are thought to fall within the *Mantella madagascariensis* group, one of five clades within the *Mantella* genus, though their position in this group is controversial (Schaefer et al., 2002; Vences et al., 2004). Population genetic studies have detected high degrees of haplotype sharing between *M. milotympanum* and *M. crocea*, resulting in the hypothesis that these two species are conspecific (Chiari et al., 2004; Vences et al., 2004). Additionally, frog populations displaying patterning that is intermediate between that of *M. crocea* and *M. milotympanum* exist in the wild and are referred to as *M. cf. milotympanum* in the literature (Chiari et al., 2004). Evidence of haplotype sharing between *M. aurantiaca* and *M. crocea* (Chiari et al., 2004; Vences et al., 2004) has also prevented taxonomic resolution within this group, and species designations remain controversial.

Given the lack of resolution in previous molecular studies, it is apparent that a high-resolution genetic dataset is needed to both clarify relationships within this group and to determine whether observed color variants represent distinct species or morphs. In this

study, we used a restriction site-associated (RAD) sequencing approach, in combination with multiple matrix regression analysis, to compare variation in dorsal coloration and side and ventral patterning with genetic and geographic distance across the entire known range of three species of Malagasy poison frog. Specifically, our objectives were to (a) clarify genetic structure among populations of *M. crocea*, *M. milotympanum*, and *M. aurantiaca*, (b) quantify variation in dorsal coloration and side and ventral patterning both among and within populations, and (c) describe the relationship between genetic diversity, phenotypic diversity, and geographic distance for all major genetic clusters within this three-species complex. This study—the first quantitative and objective exploration of color diversity in the *Mantella* genus—not only provides a foundation for future studies of color evolution in Malagasy poison frogs but also identifies several critical issues that should be more thoroughly considered in any investigation of aposematic organisms presumed to be polymorphic or polytypic.

2 | MATERIALS AND METHODS

2.1 | Field sampling

We sampled three closely related species, currently named *M. aurantiaca*, *M. crocea*, and *M. milotympanum*, throughout their entire known

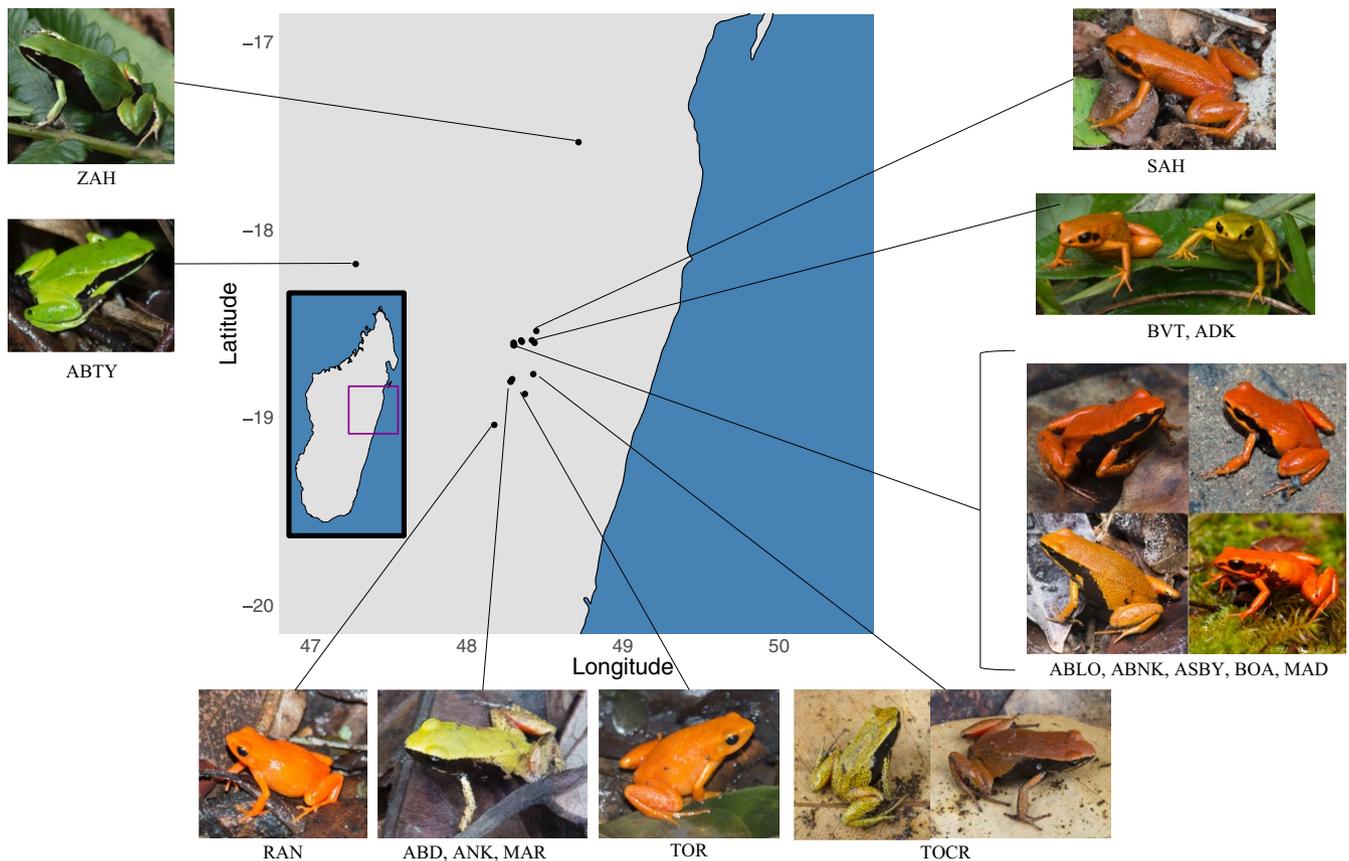


FIGURE 1 Sampling localities for 16 populations of *Mantella crocea*, *Mantella aurantiaca*, and *Mantella milotympanum* across central-eastern Madagascar. Representative individuals from a population or group of populations are pictured next to population labels. Images shown here do not represent the entire range of observed phenotypic diversity

range in central-eastern Madagascar (Figure 1). We also sampled individuals containing an intermediate phenotype between *M. crocea* and *M. milotympanum*, hereafter referred to as *M. cf. milotympanum*, following previous nomenclature in the literature. Overall, we sampled 88 frogs from 16 populations. Fieldwork was conducted during the rainy breeding season over 3 years: January–February 2014, January–February 2015, and November 2015–January 2016. We captured the frogs by hand and transported them back to a field laboratory where all data collection occurred. We collected digital photographs and toe clips from 11 *M. aurantiaca* individuals from two populations, 34 *M. crocea* individuals from six populations, 19 *M. milotympanum* individuals from three populations, and 24 *M. cf. milotympanum* individuals from five populations. Several studies have demonstrated that toe-clipping has no significant impact on frog survival, body condition, or growth (reviewed in Perry, Wallace, Perry, Curzer, & Muhlberger, 2011). Toe clips were preserved in salt-saturated DMSO and stored at room temperature. All data were collected on the same day that frogs were captured. Frogs were held overnight and released to their site of capture the following morning. All animal handling procedures were approved by the Animal Care and Use Committee at the University of California at Berkeley (AUP-2015-01-7083). Collection and exportation of samples were performed under permits issued by the Direction Générale des Forêts, Direction de la Conservation de la Biodiversité et du Système des Aires Protégées, and Ministère de l'Environnement, de l'Ecologie et des Forêts in Madagascar (collection permits: 315/13/MEF/SG/DGF/DCB.SAP/SCB, 335/14/MEF/SG/DGF/DCB.SAP/SCB, and 336/14/MEF/SG/DGF/DCB.SAP/SCB; export permits: 051C-EA02/MG14, 048C-EA02/MG15, and 002C-EA01/MG16).

2.2 | Quantification of phenotypic variation

To characterize frog coloration and pattern, we photographed the dorsal, ventral, and side surfaces of all frogs in a standardized manner following a protocol modified from Stevens, Párraga, Cuthill, Partridge, and Troscianko (2007). Photographs of all frogs were taken after transportation to the field laboratory, but before any other handling occurred. Frogs were always photographed in natural light during the hours of 1:00–5:00 p.m. using a Pentax K-30 digital single-lens reflex camera fitted with a Pentax 18–135 mm lens. All frogs were photographed on a white paper background with a scale bar and a white–gray–black standard present (QPCard 101; gray standard reflectance value of 18%). Two individuals (ABNK09 and RAN11) were excluded from our phenotypic analysis due to unsuitable photographs.

To quantify dorsal coloration, we first used the Image Calibration and Analysis Toolbox (Troscianko & Stevens, 2015), utilized within ImageJ v1.51 (Schneider, Rasband, & Eliceiri, 2012), to generate aligned and normalized images from our RAW photographs, which enabled us to objectively measure color and pattern. After standardization, we selected two regions of interest on the frog's dorsal surface in which to quantify color: one 3 × 3 mm square toward the back of the frog dorsum, and one 3 × 1 mm rectangle behind the

frog's right arm. Regions of interest on the frog's dorsum were manually selected in order to avoid any glare present in photographs as well as any injuries on the frog's dorsal surface that resulted in discoloration. After selecting regions of interest, we used the Batch Image Analysis function of the toolbox to extract the red color values, green color values, and blue color values (hereafter referred to as RGB values) for each frog. R values, G values, and B values were averaged for all pixels within the regions of interest. Color values were averaged separately for each color channel (R, G, and B).

To transform RGB values into measurements relevant to a vertebrate visual system, we followed the protocol of Krohn and Rosenblum (2016), modified from Endler (2012) and McKay (2013). Briefly, we calculated three axes from our RGB values corresponding to a red-green channel, $(R - G)/(R + G)$, a blue-green channel, $(G - B)/(G + B)$, and a luminance channel, which we defined as untransformed R values. Because luminance is processed separately in vertebrates (Endler, 2012; Endler & Mielke, 2005), we used our other two axes, $(R - G)/(R + G)$ and $(G - B)/(G + B)$, to plot dorsal coloration as a point in a two-dimensional color space based on a red-green channel and a blue-green channel. In our two-dimensional color space, $(R - G)/(R + G)$ represented the x-axis and $(G - B)/(G + B)$ represented the y-axis. From this color space, we calculated chroma and hue values following Krohn and Rosenblum (2016). Chroma was calculated as the hypotenuse of the triangle formed by the x and y axes, and hue was calculated as the angle between the hypotenuse and the x-axis. We used these values of chroma, hue, and luminance to characterize dorsal coloration.

To quantify side and ventral pattern, we again used the Image Calibration and Analysis Toolbox (Troscianko & Stevens, 2015) to generate aligned and normalized images from RAW photographs. Next, we manually outlined the ventral and side surfaces of frogs on the standardized images using the polygon and brush selection tools. We selected the entire ventral and side surfaces to obtain a comprehensive measure of overall pattern on each surface. After manually selecting the regions of interest, we used the scale bars present in each photograph to scale all images to the same number of pixels per millimeter (side surfaces = 19 px/mm; ventral surfaces = 18.6 px/mm), which is necessary for pattern analysis. We performed a granularity analysis, which is based on Fast Fourier bandpass filtering, using the Image Calibration and Analysis Toolbox implemented in ImageJ. For our side pattern analysis, we specified Fast Fourier Transform Bandpass filters at 16 levels, starting at two pixels and increasing by a multiple of the square root of two until 430 pixels. For our ventral pattern analysis, we specified Fast Fourier Transform Bandpass filters at 14 levels, starting at two pixels and increasing by a multiple of the square root of two until 193 pixels. From our granularity analysis, we generated descriptive summary statistics to estimate pattern contrast, pattern diversity, and luminance contrast for both side and ventral pattern. Granularity analysis has been increasingly used to measure pattern in a variety of organisms (e.g., Barbosa et al., 2008; Stoddard & Stevens, 2010) and draws on basic characteristics of early-stage visual processing present across diverse animals (Campbell & Rodson, 1968; Godfrey, Lythgoe, &

Rumball, 1987; Pérez-Rodríguez, Jovani, & Stevens, 2017; Stoddard & Stevens, 2010). Conversely, color adjacency analysis, a method which has previously been used to quantify pattern in poison frogs, does not represent visual processing of pattern (Pérez-Rodríguez et al., 2017).

Our dataset included both males and females (58 males, 26 females, two juveniles). Preliminary analyses did not reveal an effect of sex on frog coloration or pattern, so sexes were lumped for the analyses presented here. However, our study design did not explicitly aim to quantify sexual dimorphism, and future studies can investigate this question with targeted sampling.

2.3 | Quantification of genetic variation

We extracted DNA from toe clips using Qiagen DNeasy extraction kits (Qiagen, Valencia, CA, USA) generally following the manufacturer's protocol with two modifications: 4 μ l of 1 mg/ml RNase A was added to each sample after lysis, and DNA was eluted in 45 μ l of 1 \times LTE buffer to maximize concentration. Prior to library preparation, we checked the quality of extracted DNA using agarose gels and quantified DNA using a Qubit 2.0 Fluorometer (ThermoFisher, Waltham, MA, USA).

We constructed a restriction site-associated DNA (RAD) library following the protocol of Ali et al. (2016), without the targeted bait capture step, otherwise referred to as the "bestRAD" protocol. During preparation of our bestRAD library, we digested 50 ng of DNA from each individual with SbfI-HF (New England Biolabs, Ipswich, MA, USA) restriction enzyme, performed size selection with magnetic beads, and amplified our library using 12 cycles of polymerase chain reaction (PCR). We sequenced our library on two lanes of the Illumina HiSeq 4000 at the U.C. Davis Genome Center with 150 bp paired-end reads.

2.4 | RADseq data processing

To process RADseq data, we used pipelines implemented in a customized PERL workflow that also utilized various external programs (pipelines can be accessed through <https://github.com/CGRL-QB3-UCBerkeley/RAD>). We first demultiplexed raw fastq reads using internal barcode sequences and allowing for one mismatch in barcode sequence. We removed demultiplexed reads that did not include the expected restriction enzyme cut site at the beginning of the read, again allowing for one mismatch in cut site sequence, and also removed exact duplicates using Super Deduper (<https://github.com/dstreeett/Super-Deduper>). To filter reads, we used Cutadapt (Martin, 2011) and Trimmomatic (Bolger, Lohse, & Usadel, 2014) to trim adapter contamination and low quality reads. We removed filtered reads that were shorter than 50 bp. After cleaning and filtering reads, we used cd-hit (Fu, Niu, Zhu, Wu, & Li, 2012; Li & Godzik, 2006) to cluster forward reads of each individual at 95% similarity, retaining only those clusters with at least two supported reads. For each cluster, we used the sequence identified as representative by cd-hit as our marker. Retained markers were next masked

for repetitive elements, low complexities, and short repeats with Ns using RepeatMasker (Smit, Hubley, & Green, 2014) with "frog" as a database. Post-masking, we removed markers where more than 30% of nucleotides were Ns. To remove potential paralogs present within each individual, we used Blastn (Altschul, Gish, Miller, Myers, & Lipman, 1990) to compare all clustered loci against themselves, and subsequently eliminated any locus that matched a locus other than itself. Next, remaining RAD markers from each individual were combined and clustered for all individuals using cd-hit. We used a similarity threshold of 90% to select for markers containing at least 50 nucleotides and shared by at least 60% of all individuals. This served as our reference genome for all samples, and we subsequently aligned each individual's cleaned sequence reads to this reference using Novoalign (<http://www.novocraft.com>). We only retained those reads that mapped uniquely to the reference. Using Picard (<http://www.picard.sourceforge.net>) and GATK (McKenna et al., 2010), we added read groups and performed realignment around indels. To generate quality control information in VCF format, we used SAMtools/BCftools (Li et al., 2009), after which data were further filtered using a custom method, SNPcleaner (<https://github.com/tplinderoth/ngsQC/tree/master/snpCleaner>; Bi et al., 2013), that was slightly modified to remove sites around indels and implemented in our pipelines. Additionally, we filtered out markers where more than two alleles were called at any site, and also masked sites that were within 10 bp of any indel. Only individual sites falling within the first to ninety-ninth percentile of overall coverage among all samples were retained. We also removed SNPs failing to pass a one-tailed HWE exact test ($1e-4$) or showing strong base quality bias ($1e-100$). To avoid bias resulting from excessive missing data, we only used sites present in at least 60% of individuals with at least 3 \times coverage in our downstream analyses.

2.5 | Population genomic analyses

We used genotype likelihoods instead of genotype calls whenever possible to account for uncertainty in our data. Because our data showed low to medium coverage (1.8–13.9 \times , with an average of 5.4 \times), SNP and genotype calls based on allele counts could potentially cause bias or introduce noise in downstream analyses (Johnson & Slatkin, 2008; Lynch, 2008). Genotype likelihoods were calculated in an empirical Bayesian framework using ANGSD (<http://www.popgen.dk/angsd/index.php/ANGSD>; Korneliussen, Albrechtsen, & Nielsen, 2014), a software that is specialized for analyzing low to medium coverage next-generation sequencing data. The majority of downstream analyses conducted in ANGSD are performed based on likelihood of site allele frequencies, genotype likelihoods, or genotype posterior probabilities. For analyses that required SNP or genotype calling, we only included high-confidence variants (identified using a likelihood ratio test to determine variable sites with p -values $<1e-6$ and genotype posterior probabilities >0.95) where at least 80% of individuals showed at least 3 \times coverage.

To characterize population structure for all samples, we first performed a Principal Components Analysis (PCA) of the covariance

matrix of posterior genotype probabilities implemented in ngsTools (<http://github.com/mfumagalli/ngsTools>; Fumagalli, Vieira, Linderoth, & Nielsen, 2014). Next, we used a neighbor-joining network (NeighborNet) analysis based on uncorrelated p -distances in Splitstree (Huson, 1998; Huson & Bryant, 2006) to visualize population structure. We adhered to the stringent thresholds mentioned above to call a set of high-quality variants, which were used to compute a genetic distance matrix for Splitstree using the Adegenet package in R (Jombart, 2008). We also quantified the population structure of all samples using NGSadmix (Skotte, Korneliusen, & Albrechtsen, 2013), which relies on genotype likelihoods. Because there were sixteen populations present in our study, we estimated individual admixture proportions with the number of clusters ranging from one to seventeen ($K = 1-17$), with ten replicates per K value. We then used the Evanno method (Evanno, Regnaut, & Goudet, 2005) to identify the most likely K value.

To characterize fine-scale population structure, we performed an NGSadmix analysis within each major genetic cluster. We again estimated individual admixture proportions with the number of clusters ranging from one to one more than the total number of populations ($K = 1-3$ for Cluster A; $K = 1-7$ for Cluster B; $K = 1-9$ for Cluster C; $K = 1-15$ for candidate hybrid populations). We ran NGSadmix with ten replicates for each K value and used the Evanno method to determine the most likely K value. We assigned admixed populations to the NGSadmix group from which more than 50% of its admixture was drawn. Finally, we calculated F_{ST} values for all possible pairwise population comparisons using the realSFS function of ANGSD.

2.6 | Integration of genomic and phenotypic datasets

We used Mantel and partial Mantel tests to investigate the relationship between genetic, geographic, and phenotypic distance for each major cluster in our study. Some concerns have been raised over the use of Mantel tests in population genetics, especially in regards to inflated type I error rates for partial Mantel tests when spatial autocorrelation is present (Diniz-Filho et al., 2013; Guillot & Rousset, 2013; Legendre & Fortin, 2010). Therefore, we rely most heavily on pairwise comparisons, and we are cautious in our interpretation of partial Mantel results. In addition, we focus on comparisons across major genetic groupings, where any potential biases should be comparable. Although we are conservative throughout about linking pattern to process, our results highlight complexes with interesting spatial patterns of genetic and phenotypic variation.

Our regression analysis was performed on 86 individuals and excluded the two samples indicated above. To generate our geographic distance matrix, we used the Geographic Distance Matrix Generator (Ersts, 2018). To generate our genetic distance matrix, we used ngsDist (Vieira, Lassalle, Korneliusen, & Fumagalli, 2016) to estimate pairwise genetic distances using genotype posterior probabilities. To quantify phenotypic distances among individuals, we generated three separate distance matrices: a dorsal coloration distance matrix, a side pattern distance matrix, and a ventral pattern distance

matrix. To generate our dorsal coloration distance matrix, we first standardized digital photographs, extracted and averaged RGB values from regions of interest, and transformed RGB values to a two-dimensional color space as described above. We then calculated the Euclidean distance between points in this conceptual color space to generate measures of pairwise distance in dorsal coloration. To generate our side and ventral pattern distance matrices, we used differences in luminance (with R as the luminance channel) to characterize the pairwise distances in pattern among individuals. We designated R as the luminance channel because many vertebrates are believed to use long-wavelength sensitive (LWS) cones to detect achromatic signals (Endler & Mielke, 2005; Osorio & Vorobyev, 2005). After standardizing digital photographs, selecting regions of interest, and scaling all pictures as described above, we calculated the number of pixels that fell into each of 95 luminance bins ranging from 0% to 100% reflectance for each frog's surfaces separately (ventral and side). Next, we used the toolbox's Luminance Distribution Difference Calculator to compare the luminance distribution histograms among each pair of frogs and to generate pairwise measures of difference in luminance distribution, which we used as our measure of variation in ventral and side pattern. This methodology follows the recommendation of the toolbox's user guide because pattern differences in this study system are characterized by discrete patches of high and low luminance values. All Mantel and partial Mantel tests were performed in R version 3.4.3 using the vegan package (R Core Team, 2017), and each test used Pearson's method of correlation and performed 999 permutations.

3 | RESULTS

Our sequencing efforts yielded a high-quality dataset. After filtering reads for intact barcode sequences and restriction enzyme sites, the number of reads per sample ranged from 0.86 to 13.8 million, with an average of 4.48 million reads per sample. After self-blasting, we obtained an average of 43,559 RAD loci per individual, with the number of loci ranging from 15,491 to 59,488 per sample. The final pseudo-reference genome, which was generated by clustering loci across all samples and included those loci shared by at least 60% of the samples, contained 35,113 loci. The average coverage per individual was $5.4\times$ and ranged from $1.8\times$ to $13.9\times$. After raw variant filtering, we retained 2,284,942 sites derived from 21,733 loci for ANGSD analyses. Above 65% of the samples in our dataset have at least $3\times$ coverage at these loci.

We found that *M. aurantiaca*, *M. crocea*, and *M. milotympanum* populations were highly structured with three distinct groups emerging. Our PCA, based on 2,284,942 sites, revealed three main clusters: Cluster A contained the two most geographically remote *M. crocea* populations, Cluster B contained the remaining *M. crocea* populations and the two *M. aurantiaca* populations, and Cluster C contained all *M. milotympanum* and all *M. cf. milotympanum* populations (Figure 2a). These three clusters were supported by our NGSadmix analysis, which also identified three

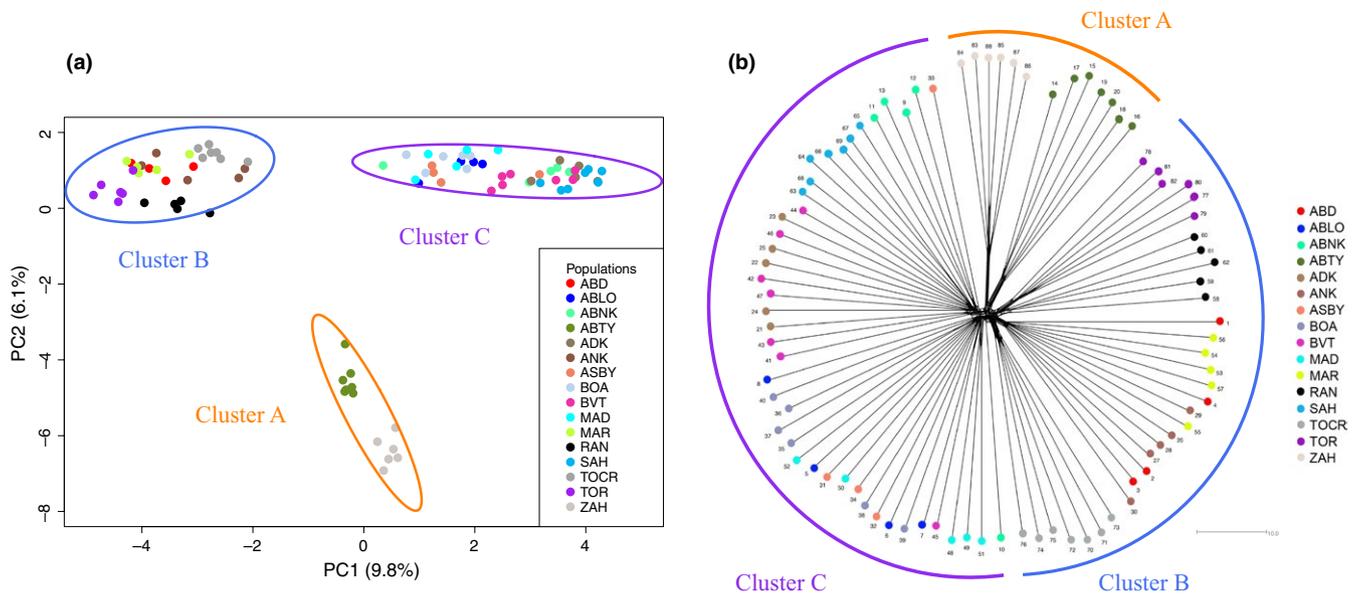


FIGURE 2 Patterns of genetic variation among individuals from all sampling localities. Plots show PCA (panel a) and Splitstree diagram (panel b) for all individuals, with colors denoting sampling localities. Ellipses on the PCA plot represent the 95% confidence intervals around the mean of each group identified in our NGSadmixture analysis (Clusters A, B, and C). Semi-circles on Splitstree diagram denote groups identified in NGSadmixture analysis and again correspond to Clusters A, B, and C. NeighborNet analysis in panel (b) was based on *p*-distances in Splitstree

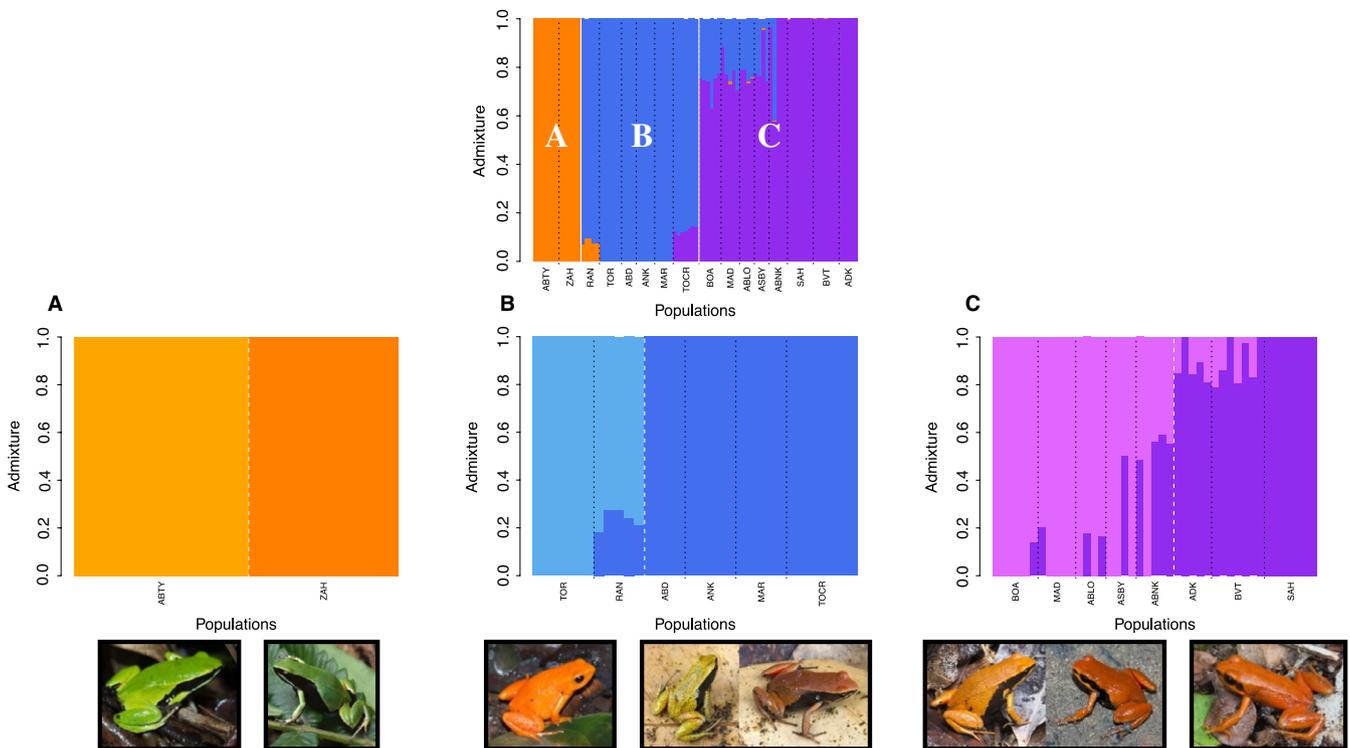


FIGURE 3 Population structure determined by NGSadmixture analysis. The top plot shows NGSadmixture clustering for $K = 3$ (the most likely K) when all samples are included. Dashed black lines separate geographic sampling locations, while solid white lines separate groups identified in NGSadmixture analysis (designated here as Clusters A, B, and C). NGSadmixture sub-plots show fine-scale population structure within each major genetic cluster (A, B, and C) for $K = 2$ (determined to be the most likely K for each cluster). Within each sub-plot, black dashed lines separate geographic sampling locations, while dashed white lines separate groups identified in NGSadmixture analysis. Below each sub-plot are pictures of representative individuals from the population or populations contained within each NGSadmixture group

groups ($K = 3$ based on Evanno et al., 2005 method) corresponding to the same population groupings (Figure 3). Our splitsTree analysis, based on 14,367 high-quality SNPs, further supported this general pattern of genetic variation partitioning (Figure 2b), though there were notable differences in the degree of genetic admixture present within clusters. The admixture detected in our NGSadmix and splitsTree analysis was concordant with our pairwise population F_{ST} value comparisons, which were lower among admixed populations (Figure 4). Still, F_{ST} values were relatively high for all pairwise population comparisons, including those that were admixed. It is important to note that geographic and genetic distance were not equivalent across major genetic clusters. While populations within some clusters were both geographically and genetically disparate (as in Cluster A), other populations were geographically more proximate with varying degrees of genetic distance between them (as in Clusters B and C).

Within Cluster A, both our PCA and NGSadmix analysis indicated that each population was a genetically distinct entity. Our NGSadmix analysis identified two groups ($K = 2$ based on Evanno et al., 2005 method) corresponding to each population (Figure 3). These populations were also clearly differentiated in our splitsTree analysis,

with no admixture occurring between them (Figure 2b). In fact, the pairwise F_{ST} value between these two populations (0.43) was among the highest in our dataset (Figure 4). Within this cluster, Mantel and partial Mantel tests confirmed that genetic distance was most highly correlated with geographic distance ($r = 0.9444$; p -value = 0.001; Table 1). Although dorsal coloration and side pattern were also correlated with genetic variation, neither phenotypic trait remained significantly associated with genetic distance after accounting for the effect of geographic distance in partial Mantel tests (side pattern: $r = -0.1835$; p -value = 0.994; dorsal coloration: $r = 0.1033$, p -value = 0.129). Additionally, the degree of phenotypic variability was much lower in this cluster than in the others, as side pattern was the only phenotypic trait that was significantly correlated with geographic distance when accounting for genetic distance ($r = 0.3006$; p -value = 0.008). Phenotypically, populations within this cluster were relatively uniform, and there was little variation in either dorsal coloration or in side and ventral pattern (Figure 5).

Within Cluster B, we also found evidence for two distinct groups based on our PCA and NGSadmix analysis. Our NGSadmix analysis identified two groups ($K = 2$ based on Evanno et al., 2005 method), one containing the two *M. aurantiaca* populations and one containing

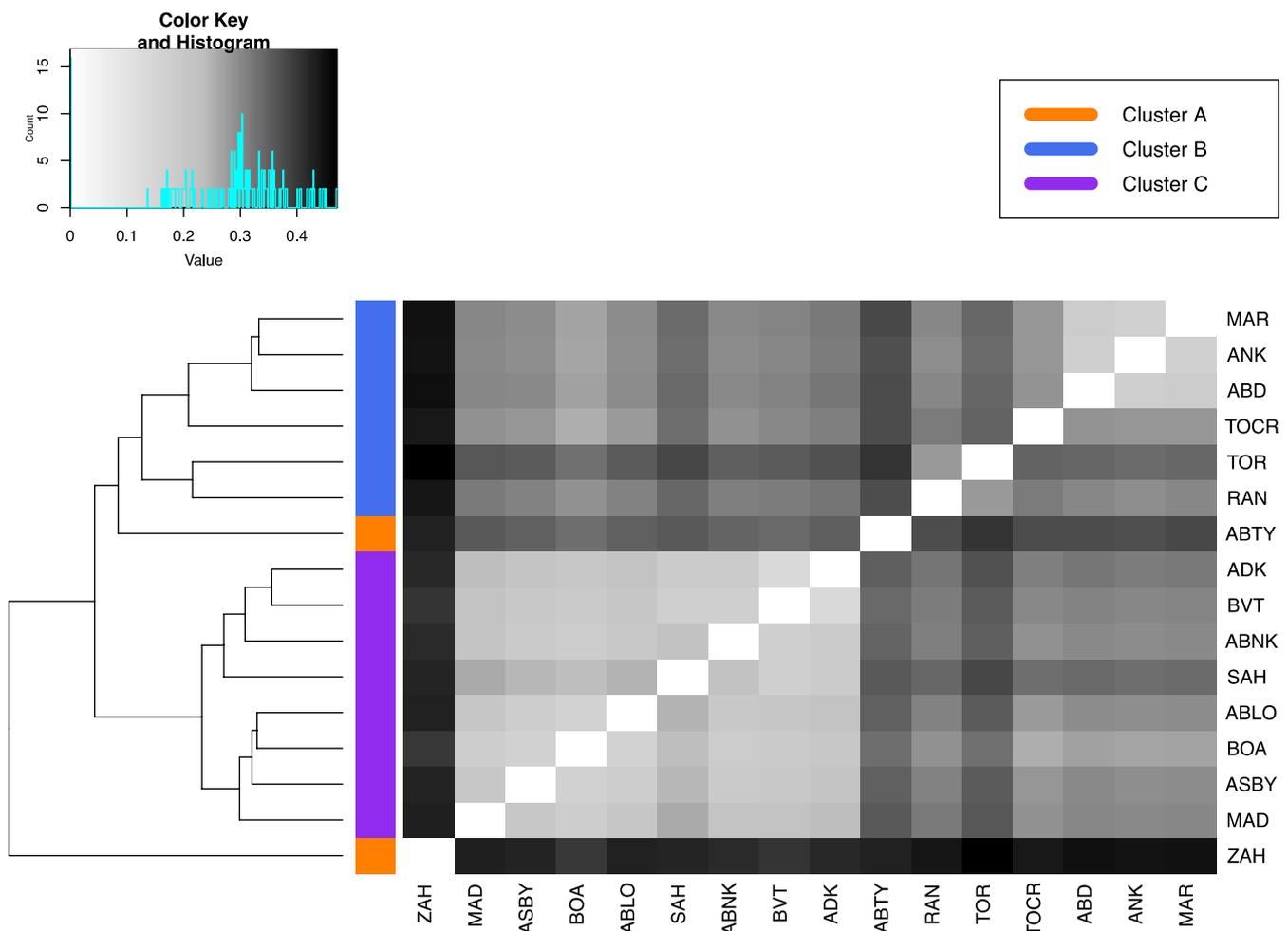


FIGURE 4 Dendrogram and heat map of pairwise F_{ST} values for all sampling localities. Lighter cells denote lower pairwise F_{ST} values (less genetic differentiation), while darker cells denote higher pairwise F_{ST} values (more genetic differentiation)

TABLE 1 Matrix regression results of variables that correlate with genetic distance within major genetic clusters

Group	X matrix	Mantel r	p
Cluster A	Dorsal coloration	0.3089	0.019*
	Geographic dist.	0.9444	0.001* [†]
	Side pattern	0.3066	0.006*
	Ventral pattern	0.0529	0.258
Cluster B	Dorsal coloration	0.3715	0.001* [†]
	Geographic dist.	0.3414	0.011*
	Side pattern	0.2880	0.002*
	Ventral pattern	0.3467	0.001*
Cluster C	Dorsal coloration	-0.0202	0.537
	Geographic dist.	0.1038	0.018* [†]
	Side pattern	-0.1280	0.897
	Ventral pattern	-0.0956	0.876
Candidate hybrids	Dorsal coloration	-0.0012	0.472
	Geographic dist.	0.1391	0.050* [†]
	Side pattern	-0.0386	0.740
	Ventral pattern	0.0471	0.153

Notes. "Candidate hybrids" refer to the subset of samples that includes admixed individuals and the clusters from which this admixture was drawn (Clusters B and C).

*Significant p -values. [†]The variable with the highest Mantel r statistic.

the four remaining *M. crocea* populations (Figure 3), though there was some admixture between one *M. crocea* and one *M. aurantiaca* population. Our splitsTree analysis supported this general pattern and also showed admixture among the three geographically closest *M. crocea* populations (Figure 2b). Pairwise F_{ST} values among these three admixed *M. crocea* populations ranged from 0.17 to 0.18, which is low in comparison to the range of values in our dataset (Figure 4). The F_{ST} values between any of these three admixed *M. crocea* populations and the fourth remaining *M. crocea* population ranged from 0.28 to 0.29 (Figure 4). Comparatively, values between either *M. aurantiaca* population or *M. crocea* populations ranged from 0.30 to 0.35 (Figure 4). Within this group, Mantel and partial Mantel tests confirmed that genetic distance was most highly correlated with variation in dorsal coloration ($r = 0.3715$; p -value = 0.001; Table 1). Even after accounting for the effects of geographic distance in a partial Mantel test, dorsal coloration remained significantly associated with genetic variation ($r = 0.2398$; p -value = 0.002). Variation in ventral patterning also remained significantly correlated with genetic distance after controlling for geographic distance ($r = 0.2152$; p -value = 0.011). All three quantified phenotypic traits were significantly correlated with geography after controlling for genetic distance (dorsal coloration: $r = 0.4589$; p -value = 0.001; side pattern: $r = 0.5174$; p -value = 0.001; ventral pattern: $r = 0.437$; p -value = 0.001). Phenotypic discrepancies between the *M. crocea* and *M. aurantiaca* groups were most evident in dorsal coloration, particularly in dorsal chroma and hue (Figure 5). Although genetic distance was correlated with variation in side pattern, ventral pattern, and

geographic distance, none of these correlations remained significant when accounting for the effects of dorsal coloration, the variable with the highest r value in Mantel tests.

Within Cluster C, both our PCA analysis and our NGSadmixture analysis indicated two distinct groups. Our NGSadmixture analysis revealed two groups ($K = 2$ based on Evanno et al., 2005 method), one containing all *M. milotympanum* populations and one containing all *M. cf. milotympanum* populations (Figure 3). While our splitsTree analysis supported these two main groupings, there was a high degree of genetic admixture among populations in this cluster (Figure 2b) and populations were less well-defined. The F_{ST} values between populations in this cluster were the lowest out of all pairwise population comparisons in this dataset, ranging from 0.14 to 0.25 (Figure 4). Despite the high degree of phenotypic variability in populations of this cluster, genetic variation was not correlated with any of the phenotypic traits that we quantified (Table 1). Genetic distance was only significantly correlated with geographic distance ($r = 0.1038$; p -value = 0.018; Table 1). All three phenotypic traits, however, remained significantly correlated with geographic distance when accounting for genetic distance (dorsal coloration: $r = 0.1353$; p -value = 0.009; side pattern: $r = 0.2067$; p -value = 0.002; ventral pattern: $r = 0.387$; p -value = 0.001). Despite this spatial segregation of phenotypes, we observed an exceptionally high degree of variation within populations in this group (Figure 5). For side and ventral pattern, in particular, within-population variation was highest and discrete groupings were not apparent (Figure 5). Although there was less within-population variation in dorsal chroma and hue, there was still not an obvious phenotypic split between *M. milotympanum* and *M. cf. milotympanum* populations (Figure 5). Overall, phenotypes in this group demonstrated much higher degrees of intrapopulation variation with less discrete groups emerging.

One of the most interesting findings was evidence of possible introgression between *M. crocea/aurantiaca* and *M. milotympanum* populations. To further investigate this phenomenon, we repeated our NGSadmixture and Mantel analysis on the subset of samples that included the admixed populations and the two clusters from which this admixture was drawn (Cluster B and Cluster C). Our NGSadmixture results yielded a consistent signature of admixture in all *M. cf. milotympanum* populations that we sampled (Figure 6). In the wild, *M. cf. milotympanum* individuals demonstrate a side and ventral pattern that is intermediate between that observed in *M. crocea* and *M. milotympanum* (Figure 6). Despite the existence of these phenotypically and genotypically intermediate populations, genetic distance was not correlated with any of the phenotypic traits that we quantified, but was significantly correlated with geographic distance ($r = 0.1391$; p -value = 0.050; Table 1). However, variation in dorsal coloration, side pattern, and ventral pattern was all significantly correlated with geographic distance even when controlling for the effects of genetic distance with partial Mantel tests (dorsal coloration: $r = 0.2585$; p -value = 0.001; side pattern: $r = 0.3873$; p -value = 0.001; ventral pattern: $r = 0.3233$; p -value = 0.001). Notably, candidate hybrids, or *M. cf. milotympanum* individuals, demonstrated

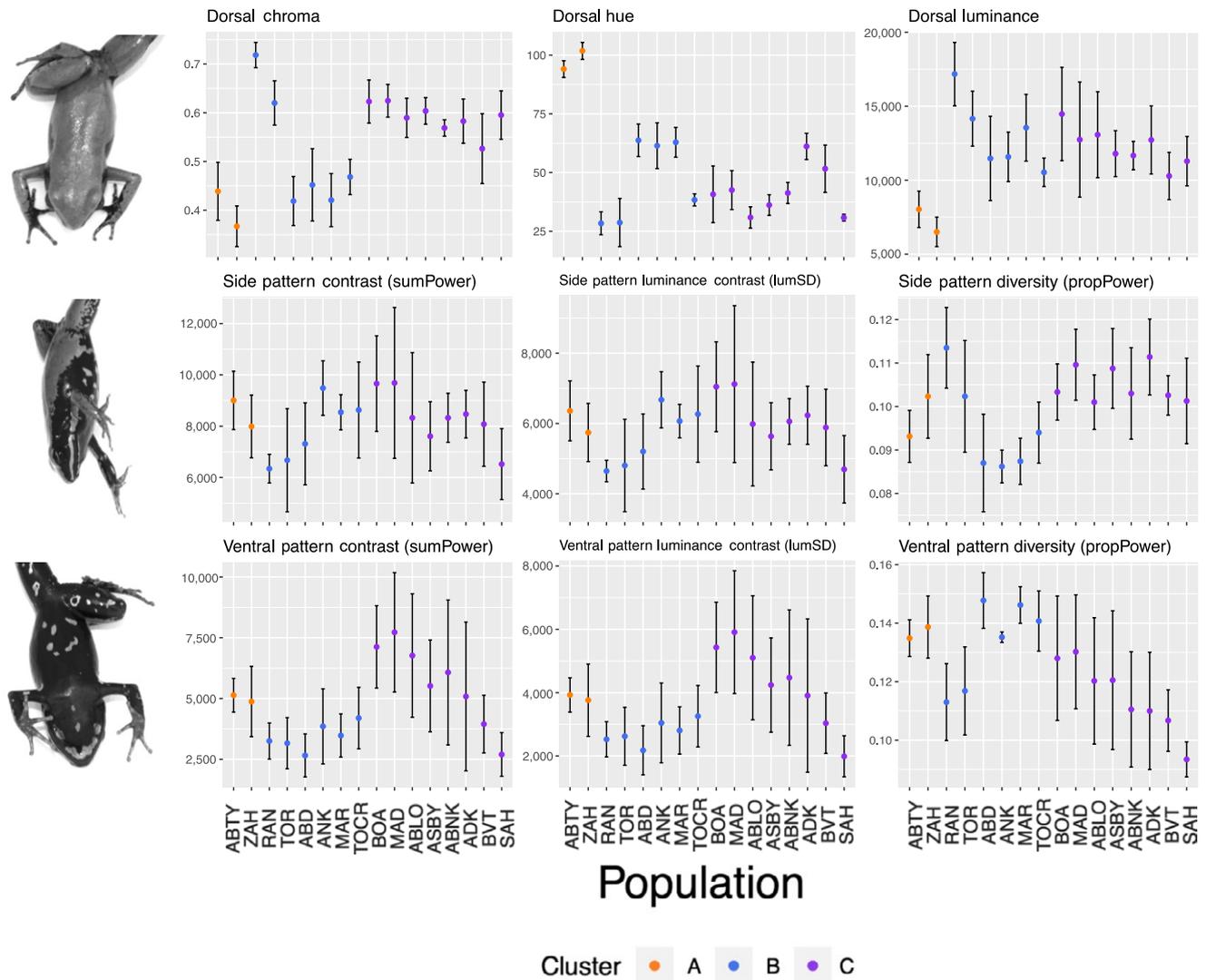


FIGURE 5 Variation in dorsal coloration (chroma, hue, and luminance) and side and ventral pattern (pattern contrast, luminance contrast, and pattern diversity) across all sampling localities. Mean values and standard deviation for each population are depicted in the graphs

higher within-population variation in patterning characteristics, and also displayed novel pattern phenotypes not present in either “parental” species (Figures 5 and 6). Within each candidate hybrid population, we observed a spectrum of pattern variants rather than discrete morphs.

4 | DISCUSSION

Our investigation of color diversity and genetic structure in a complex of three closely related species of Malagasy poison frog revealed that the level of intraspecific color variation in this species complex has likely been overestimated, while the occurrence of distinct species has been underestimated. Although not in alignment with current species designations, we found evidence for several clear genetic clusters, each demonstrating a distinctive pattern of genetic and phenotypic variation, suggesting that a number of mechanisms are contributing to color evolution in this complex. Below,

we discuss the relevance of our findings for taxonomy, conservation, and the evolutionary processes contributing to phenotypic variation in aposematic organisms. Additionally, we consider the implications of our findings for characterizing morphs in aposematic systems.

4.1 | Taxonomic resolution and implications for conservation

Populations within this complex were highly structured at relatively small spatial scales, with distinct genetic groups emerging. Unexpectedly, genetic clusters identified in this study do not correspond to current species designations. Specifically, our results indicated that frog populations that have thus far been considered to be a green color variant of the *M. crocea* species form their own genetic cluster (Cluster A) and are in fact distinct from the other *M. crocea* populations included in this study (Figures 2 and 3). On the basis of this work, we suggest that what has previously been considered to be a green morph of *M. crocea* likely constitutes a new species, and

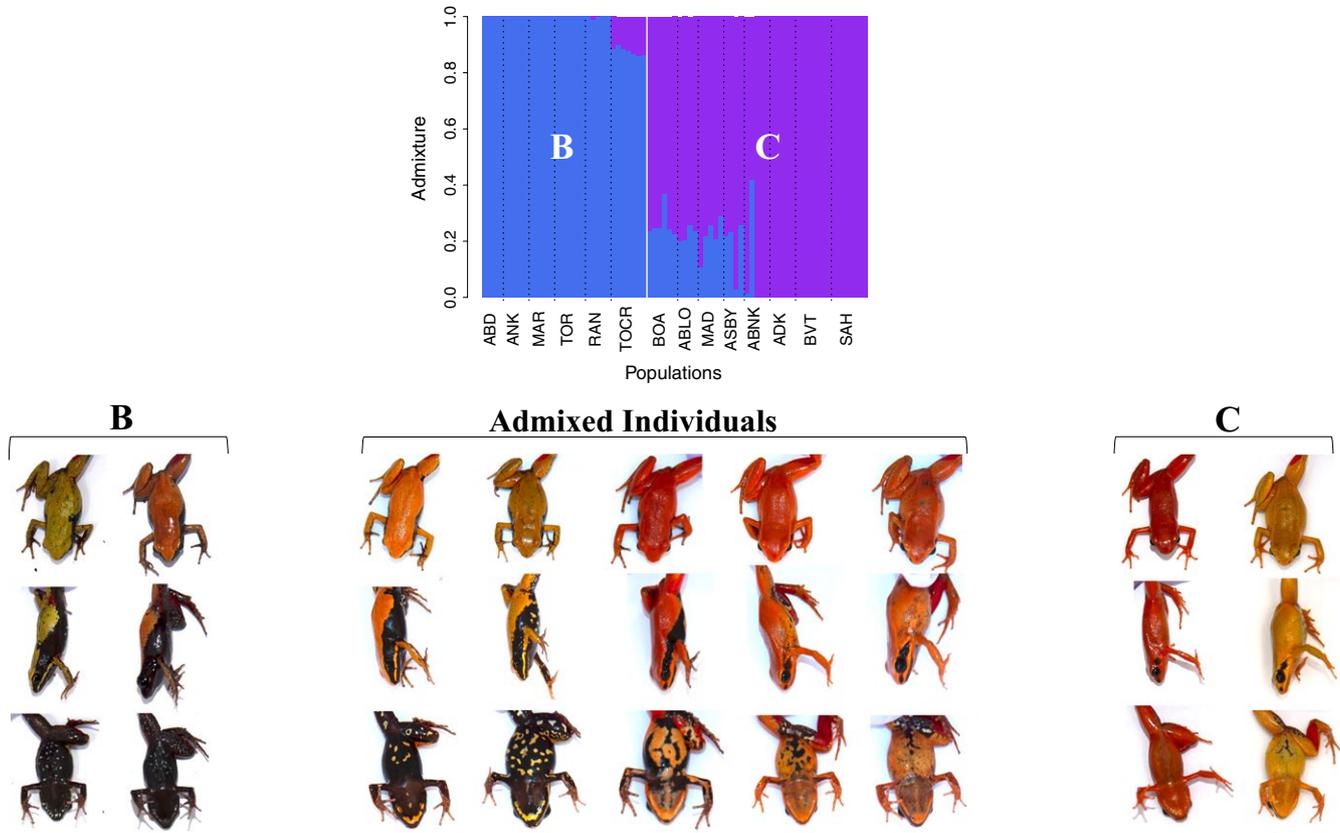


FIGURE 6 Genetic and phenotypic admixture among candidate hybrid populations. Top plot displays NGSadmix results for the subset of samples from Clusters B and C only. Dashed black lines separate geographic sampling locations, while the solid white line separates groups identified in NGSadmix analysis. Pictures under “B” and “C” headings show representative specimens from populations within each cluster (B or C), while pictures under “Admixed Individuals” show representative variation observed in individuals admixed between Clusters B and C. Individuals that are classified as *Mantella aurantiaca* are not pictured here under Cluster B. All variants displayed under “Admixed Individuals” come from a single population and demonstrate the spectrum of variation that occurs in candidate hybrid populations. Candidate hybrid populations displayed intermediate phenotypes primarily in terms of side and ventral patterning elements

we recommend further investigation of morphology, acoustics and behavior to delineate species boundaries with more certainty. The remaining (non-green) *M. crocea* populations form a distinct cluster with the two *M. aurantiaca* populations (Cluster B; Figures 2 and 3). Although previous work has found evidence of haplotype sharing between *M. crocea* and *M. aurantiaca* (Chiari et al., 2004; Vences et al., 2004), it is surprising that populations of these two species are identified as a cohesive cluster in our analysis, as they differ greatly in color, pattern and body shape. Our work also demonstrates the validity of the *M. milotympanum* species (Cluster C; Figures 2 and 3). Previous studies have hypothesized *M. milotympanum* to be a color variant of *M. crocea* (Chiari et al., 2004; Vences et al., 2004), but our results demonstrate that *M. milotympanum* is genetically distinct from the clusters that contain what is currently called *M. crocea* (Clusters A and B).

Finally, we identified candidate hybrid populations (genetically admixed between Clusters B and C) that displayed intermediate genotypes and pattern phenotypes (Figures 3 and 6). Because earlier studies have hypothesized that *M. milotympanum* and *M. crocea* are conspecific, these intermediates were referred to as

M. cf. milotympanum in the literature and assumed to be another phenotypic variant (Chiari et al., 2004; Vences et al., 2004). However, we demonstrate that *M. milotympanum* is its own distinct genetic entity, and likely experienced admixture with *M. crocea* at some point in the past. Although *M. milotympanum* and *M. crocea* do not currently live in sympatry to our knowledge, it is certainly feasible that these species either lived sympatrically in the past or have experienced secondary contact with each other, given the strikingly high degree of deforestation in Madagascar and the highly fragmented nature of remaining *Mantella* populations.

The taxonomic resolution provided by our study has significant consequences for conservation efforts, given that *M. aurantiaca* and *M. milotympanum* are currently classified as critically endangered, while *M. crocea* is classified as vulnerable (Andreone et al., 2005). Overall, populations were highly genetically structured with substantial phenotypic variation, indicating that there may be a number of distinct units to consider in management efforts. More importantly, genetic differentiation was not consistently correlated with phenotypic variation, emphasizing the importance of integrating both phenotypic and genetic information in prioritizing units for

conservation. Based on our findings, we offer several recommendations for future management efforts. Because *M. milotympanum* represents a distinct species and is not a color morph of *M. crocea*, its status as critically endangered warrants high conservation priority. Future conservation endeavors will need to address the existence of populations admixed between *M. crocea* and *M. milotympanum*, and to determine where these populations fit in a conservation plan. The conservation status of green *M. crocea* populations should be reassessed, as these populations likely constitute a new species rather than a color morph of an already described species. Finally, within each major genetic cluster, our results suggest that there may be two distinct species or subspecies (Figure 3). We recommend that these subclusters be considered as distinct units for Cluster A (the green *M. crocea* cluster, where genetic differentiation is exceptionally high) and Cluster B (the *M. crocea*-*M. aurantiaca* cluster, where phenotypic differentiation is exceptionally high). Due to the highly isolated and fragmented nature of *Mantella* populations, further intensive survey efforts will be important to identify any additional extant populations.

Overall, while it is premature to delineate new species on the basis of our analyses—especially considering the significant conservation implications in this system—our study confirms that there are at least three genetically distinct groups that do not correspond to current species descriptions. At the least, it seems clear that Cluster A (composed of green “*M. crocea*”) should be considered a distinct entity and prioritized given its extremely limited distribution (i.e., known from only two isolated locations). Due to the likely hybridization that we detected between Clusters B and C, and the fine-scale population structure that we observed within each cluster, the status of Clusters B and C is less clear. Rather than revising taxonomy prematurely, we recommend additional studies on gene flow and migration, characteristics of frog calls, and mating behavior before species boundaries can be clarified with any certainty.

4.2 | Divergent patterns of genetic and phenotypic diversity

Our findings suggest that a variety of processes at different spatial and genetic scales are likely contributing to differentiation in this Malagasy poison frog complex. Within major genetic clusters, we found evidence of highly regionalized patterns of phenotypic and genetic diversity (Figure 3). In our regression analysis, while genetic distance was most highly correlated with geographic distance for Clusters A and C, dorsal coloration was most highly correlated with genetic distance within Cluster B (Table 1). We also identified likely *M. crocea*-*M. milotympanum* hybrid populations (populations genetically admixed between Clusters B and C) that displayed intermediate genotypes and novel pattern phenotypes (with especially high within-population pattern variation; Figures 3, 5 and 6). Hybridization has been hypothesized to be an important mechanism of generating novel phenotypes in other poison frog systems (Medina, Wang, Salazar, & Amézquita, 2013) and may play a similar role here. Based on our findings, patterns of color evolution are likely

influenced by variable patterns of drift, selection, and hybridization across major genetic groups in this Malagasy poison frog complex. Regional diversity in patterns of genetic and phenotypic variation has been found in other frog systems, including the red-eyed treefrog *Agalychnis callidryas*, and lends support to the hypothesis that modes of diversification can vary substantially at relatively small spatial scales (Robertson & Zamudio, 2009).

While studies of phenotypic variation in Malagasy poison frogs are extremely limited, research on aposematic signal evolution in Neotropical poison frogs suggests that both natural and sexual selection likely contribute to phenotypic diversity (e.g., Reynolds & Fitzpatrick, 2007; Noonan & Comeault, 2009; Cummings & Crothers, 2013; Yang, Richards-Zawacki, Devar, & Dugas, 2016). Previous work in *Oophaga pumilio* has demonstrated that male dorsal coloration is an important component of female mating preferences (Maan & Cummings, 2008, 2009) and that dorsal brightness is an important signal in male-male competition (Crothers, Gering, & Cummings, 2011), while also confirming that conspicuous coloration is an honest signal to predators (Maan & Cummings, 2011). In fact, studies within this system have even suggested that natural and sexual selection may operate on different aspects of frog coloration, as variation in male brightness, an important component of assortative mating and male-male interactions, is not visible to avian predators (Crothers & Cummings, 2013). Additionally, there is growing evidence that pattern traits in poison frogs may also be under selection (Barnett, Michalis, Scott-Samuel, & Cuthill, 2018; Rojas & Endler, 2013; Wollenberg, Lötters, Mora-Ferrer, & Veith, 2008). Studies of intrapopulation pattern variation in the poison frog *Dendrobates tinctorius* have demonstrated that certain pattern traits are correlated with movement behavior, suggesting that patterning elements may also be important in determining how aposematic organisms are perceived by predators (Rojas, Devillechabrolle, & Endler, 2014). In fact, recent work in *D. tinctorius* indicates that pattern and color function as an aposematic signal to predators at close range, but are perceived as cryptic when viewed from longer distances (Barnett et al., 2018). Future work on phenotypic diversity in Malagasy poison frogs should draw on this body of literature and consider the relative roles of natural and sexual selection in shaping phenotypes, as well as the relative contributions of color and pattern to predator avoidance.

While Malagasy and Neotropical poison frogs demonstrate interesting parallels with each other, unique characteristics of the *Mantella* system render it particularly interesting from an evolutionary perspective. For example, while studies indicate that birds are particularly important predators of Neotropical poison frogs (Maan & Cummings, 2011; Saporito, Zuercher, Roberts, Gerow, & Donnelly, 2007), the only published instances of predation in Malagasy poison frogs were by snakes (*Thamnosophis* sp. in *M. aurantiaca*, and *Acrantophis madagascariensis* and *Compsophis laphystius* in *Mantella laevigata*) and lizards (*Zonosaurus* sp. in *M. aurantiaca* and *Zonosaurus madagascariensis* in *M. laevigata*) (Heying, 2001a; Hutter, Andriampenanomana, Razafindraibe, Rakotoarison, & Scherz, 2018; Jovanovic, Vences, Safarek, Rabemananjara, & Dolch, 2009). Although the extent to which snakes and lizards utilize

visual signals—especially in relation to olfactory cues—in predation is not fully understood (Maan & Cummings, 2011; Willink, García-Rodríguez, Bolaños, & Pröhl, 2014), lizards have been found to select for lower conspicuousness in the poison frog *Oophaga granulifera* (Willink et al., 2014). Consequently, there is reason to believe that selective pressures resulting from predation—presumably a major driving force in shaping phenotypic variation in aposematic organisms—may be substantially different for Malagasy poison frogs. Thus, the Malagasy poison frogs represent an important unique and comparative system for developing generality about the selective factors influencing color evolution. To understand the processes generating the interesting and variable patterns of phenotypic diversity, there is a pronounced need for research on the ecology and life history of these frogs. Fundamental research on *Mantella* predation, migration, mating behavior, diet, toxicity, and habitat quantification will be essential in formulating explicit hypotheses regarding color evolution.

4.3 | Broader implications for defining species and morphs in phenotypically diverse aposematic organisms

Our results demonstrate that characterization of morphs based solely on observed phenotypic variation, especially when genetic structure is unresolved, can lead to an overestimation of the degree of polymorphism and/or polytypism that occurs in aposematic systems. Our findings are consistent with other studies in Neotropical poison frogs where sophisticated genetic datasets, utilized either in isolation or paired with morphological and ecological data, have revealed that a single species likely includes several distinct lineages (Posso-Terranova & Andrés, 2018; Roland et al., 2017). Thus, overestimation of intraspecific variation and underestimation of species diversity in phenotypically diverse lineages may be more widespread than previously appreciated. In addition to potentially leading to erroneous evolutionary inferences, underestimating species diversity can also have important conservation implications if newly identified lineages are highly restricted in their distribution and/or exist outside of protected areas (Posso-Terranova & Andrés, 2018), as is the case for this complex of Malagasy poison frogs.

In this study, although our high-resolution genomic dataset clarified species boundaries at the highest level, the way in which fine-scale population structure is interpreted has significant implications for how hypotheses of color evolution are framed in this system. Within Cluster B, for example, if *M. aurantiaca* and *M. crocea* are considered to be color morphs of the same species, then one interpretation of our results is that differences in dorsal coloration are driving reproductive isolation and may be a potential mechanism for incipient speciation in this group, as hypothesized in another polymorphic poison frog complex (Wang & Summers, 2010). If considered to be separate species, however, then the dramatic phenotypic differences that coincide with genetic variation may serve to reinforce species recognition and to prevent hybridization. Our study demonstrates that even when genetic structure is well understood, distinguishing

between species, subspecies and morphs is not always a straightforward process and may require additional information on behavior, acoustics and mating preferences, among other traits.

For aposematic organisms presumed to be polymorphic or polytypic, our results emphasize that more thorough deliberation is necessary not only in the delineation of species but also in the characterization of morphs. Refining our understanding of what constitutes a morph, and when species or populations are polymorphic and/or polytypic, may be a necessary first step in this regard. Below, we highlight three conceptual areas identified on the basis of our findings that merit more careful consideration in the identification and description of phenotypic morphs.

Although phenotypes are multifaceted, morphs are often defined on the basis of one charismatic trait. Consequently, designating variants of one phenotypic axis as morphs is premature and fails to account for variation across multiple traits. In our study, patterns of variation in coloration and pattern traits were not always concordant; interpopulation differences in dorsal coloration were much more discrete than differences in pattern, although both varied among populations (Figure 5). Within most candidate hybrid populations, dorsal coloration was relatively uniform, but patterning elements varied almost continuously along a spectrum (Figures 5 and 6). These findings highlight the difficulties of characterizing distinct morphs when there is variation along multiple, potentially correlated phenotypic axes. Similar findings have been reported in other poison frog species, where continuous pattern variation was observed within populations (Rojas & Endler, 2013). In such instances, explicitly describing and quantifying variation—whether it is continuous or discrete—in multiple phenotypic traits is essential. We recommend caution in designating discrete morphs when there may be continuous phenotypic variation, especially along multiple phenotypic axes.

In aposematic organisms, designation of morphs is often based on human-observed phenotypic variation. This raises the question: is there really that much diversity in aposematic signals, when considered from the relevant predator and conspecific visual perspectives? In our study, although populations varied in terms of dorsal chroma and hue, dorsal luminance was largely similar across most populations (Figure 5). Interpreted in light of evidence suggesting that high luminance contrast can serve as an effective warning signal to predators (Prudic, Skemp, & Papaj, 2007), and that dorsal brightness is an important component of conspecific signaling in poison frogs (Crothers et al., 2011; Maan & Cummings, 2009), biologically meaningful color diversity in this system may be much less than expected. Although we observed high degrees of pattern variation on the side and ventral surfaces of frogs in our study (Figures 5 and 6), the role of this variation is unknown. Recent work in poison frogs linking patterning elements to movement behavior and detectability by predators (Barnett et al., 2018; Rojas, Devillechabrolle et al., 2014) indicates that pattern variation may be equally, or even more, biologically relevant than color variation. Further, evidence that the detectability of different color pattern variants is influenced by the existing light environment (Rojas, Rautiala, & Mappes, 2014) underscores the importance of incorporating information on both predator

sensory abilities and ambient lighting conditions into the characterization of phenotypic variation. Moving forward, it is important to determine which aspects of coloration and/or pattern are perceived by predators and conspecifics. Then, morphs should be defined on the basis of these biologically relevant visual perspectives to accurately capture the degree of diversity in aposematic warning signals.

Finally, our study underscores the need for a reassessment of how morphs are characterized in evolutionary biology. Determining whether color variants represent different species or morphs requires, at the least, sophisticated genomics datasets and an explicit integration of multidimensional phenotypic data. At a broad conceptual scale, our results clearly indicate that levels of intraspecific color variation in this complex of Malagasy poison frogs have been overestimated. At the same time, our results underscore the difficulty in delineating species and morphs at a fine scale in highly complex systems, particularly when genetic and phenotypic breaks are not predictably correlated (Figure 3). Even with robust genomic data, defining morphs can still be challenging, especially when species boundaries are not straightforward. Polymorphism and polytypism cannot be studied if it is not clear whether discrete morphs exist and whether observed variation is within populations, among populations, or between species.

So, what is a morph? Evolutionary biologists often use the term “morph” without a formal conceptual or operational definition. The research community has thought carefully over decades about how to define and delineate species (e.g., Mayr, 1942; Wiley, 1978; Mallet, 1995; Sites & Marshall, 2004; de Queiroz, 2007). The same attention has not been paid to the morph concept (but see Teasdale, Stevens, & Stuart-Fox, 2013). Not only do we need a more standardized definition of morph but also operational criteria for delineating morphs when there is phenotypic variation along multiple axes and/or when species boundaries are unclear or dynamic through time.

Our purpose here is not to provide a revised definition of “morph” but rather to highlight the importance of contextualizing phenotypic variation appropriately. When a novel phenotypic variant is discovered, rather than prematurely being described as a new species or morph, it should serve as a launching point for comprehensive studies that integrate phenotypic information with genomic datasets (and, ideally, information on acoustics and behavior). Describing novel phenotypic variants within a species should require—at the least—a high-resolution genetic dataset to confirm intraspecific relationships. In addition, rather than characterizing human-observed variants as morphs, we recommend that researchers specify the facet of phenotype measured, the sensory perspective used to quantify phenotype, and the level of biological organization where variation is observed. The way we define and delineate morphs is relevant in many systems, but carries particular significance in aposematic organisms, where novel color and pattern variants are regularly found. If we are overestimating intraspecific color variation and underestimating species diversity—as was found in this study—the classification of populations and species as polymorphic or polytypic may need to be reevaluated. Ultimately, it is time to give as much consideration to the conceptual and operational morph delineation as has been given to species.

ACKNOWLEDGMENTS

We thank S. Ndriantsoa for his invaluable assistance in the field. We also thank M. Vences and D. Edmonds for sharing their bountiful expertise on Malagasy poison frogs. We are grateful to numerous local guides and to R. Klonoski for their assistance and guidance in the field. For logistical support in Madagascar, we thank the Université d'Antananarivo and ICTE/MICET. This work was carried out under a collaboration agreement between MICET, the Département de Biologie Animale of the Université d'Antananarivo, and the Ministère de l'Environnement et des Forêts, Direction du Système des Aires Protégées. We also thank the Direction Générale des Forêts, Ministère de l'Environnement et des Forêts, and Madagascar National Parks for granting us research and export permits. We thank the Rosenblum lab for helpful feedback on the manuscript, and J. Troscianko and A. Krohn for helpful conversations on quantifying color and pattern. Funding was provided by The Mohamed bin Zayed Species Conservation Fund (Project #150511309), the National Geographic Society (Grant Project #9469-14) and the Museum of Vertebrate Zoology at the University of California at Berkeley. All animal work was approved by the Animal Care and Use Committee at the University of California at Berkeley (AUP-2015-01-7083). Publication made possible in part by support from the Berkeley Research Impact Initiative (BRII) sponsored by the UC Berkeley Library.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

KK designed the research with input from EBR. KK performed the field work and lab work. KK and KB analyzed the data. KK and EBR wrote the paper with contributions from KB.

DATA ACCESSIBILITY

RADseq data will be deposited in the NCBI Sequence Read Archive (SRA).

ORCID

Karina Klonoski  <https://orcid.org/0000-0001-7128-7078>

REFERENCES

- Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (rapture): Flexible and efficient sequence-based genotyping. *Genetics*, *202*, 389–400. <https://doi.org/10.1534/genetics.115.183665>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andreone, F., Cadle, J. E., Cox, N., Glaw, F., Nussbaum, R. A., Raxworthy, C. J., ... Vences, M. (2005). Species review of

- amphibian extinction risks in Madagascar: Conclusions from the Global Amphibian Assessment. *Conservation Biology*, 19(6), 1790–1802. <https://doi.org/10.1111/j.1523-1739.2005.00249.x>
- Andriamaharavo, N. R., Garraffo, H. M., Spande, T. F., Giddings, L. A., Vieites, D. R., Vences, M., & Saporito, R. A. (2015). Individual and geographic variation of skin alkaloids in three swamp-forest species of Madagascar poison frogs (*Mantella*). *Journal of Chemical Ecology*, 41(9), 837–847. <https://doi.org/10.1007/s10886-015-0616-4>
- Barbosa, A., Mäthger, L. M., Buresch, K. C., Kelly, J., Chubb, C., Chiao, C., & Hanlon, R. T. (2008). Cuttlefish camouflage: The effects of substrate contrast and size in evoking uniform, mottle or disruptive body patterns. *Vision Research*, 48(10), 1242–1253. <https://doi.org/10.1016/j.visres.2008.02.011>
- Barnett, J. B., Michalis, C., Scott-Samuel, N. E., & Cuthill, I. C. (2018). Distance-dependent defensive coloration in the poison frog *Dendrobates tinctorius*, Dendrobatidae. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25), 6416–6421.
- Bi, K., Linderoth, T., Vanderpool, D., Good, J. M., Nielsen, R., & Moritz, C. (2013). Unlocking the vault: Next generation museum population genomics. *Molecular Ecology*, 22(24), 6018–6032. <https://doi.org/10.1111/mec.12516>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bora, P., Dolch, R., Jenkins, R., Jovanovic, O., Rabemananjara, F. C., Randrianirina, J., ... Randrianelona, R. (2008). Geographical distribution of three species of Malagasy poison frogs of high conservation priority: *Mantella aurantiaca*, *M. crocea* and *M. milotympanum*. *Herpetology Notes*, 1, 39–48.
- Briolat, E. S., Burdfield-Steel, E. R., Paul, S. C., Rönkä, K. H., Seymoure, B. M., Stankowich, T., & Stuckert, A. M. M. (2018). Diversity in warning coloration: Selective paradox or the norm? *Biological Reviews*, 1–27.
- Campbell, F. W., & Robson, J. G. (1968). Application of fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197(3), 551–566. <https://doi.org/10.1113/jphysiol.1968.sp008574>
- Chiari, Y., Vences, M., Vieites, D. R., Rabemananjara, F., Bora, P., Ramilijaona Ravoahangimalala, O., & Meyer, A. (2004). New evidence for parallel evolution of colour patterns in Malagasy poison frogs (*Mantella*). *Molecular Ecology*, 13(12), 3763–3774. <https://doi.org/10.1111/j.1365-294X.2004.02367.x>
- Clark, V. C., Raxworthy, C. J., Rakotomalala, V., Sierwald, P., & Fisher, B. L. (2005). Convergent evolution of chemical defense in poison frogs and arthropod prey between Madagascar and the Neotropics. *Proceedings of the National Academy of Sciences of the United States of America*, 102(33), 11617–11622. <https://doi.org/10.1073/pnas.0503502102>
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Corl, A., Davis, A. R., Kuchta, S. R., & Sinervo, B. (2010). Selective loss of polymorphic mating types is associated with rapid phenotypic evolution during morphic speciation. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 4254–4259. <https://doi.org/10.1073/pnas.0909480107>
- Crothers, L. R., & Cummings, M. E. (2013). Warning signal brightness variation: Sexual selection may work under the radar of natural selection in populations of a polytypic poison frog. *The American Naturalist*, 181(5), E116–E124.
- Crothers, L., Gering, E., & Cummings, M. (2011). Aposematic signal variation predicts male-male interactions in a polymorphic poison frog. *Evolution*, 65(2), 599–605. <https://doi.org/10.1111/j.1558-5646.2010.01154.x>
- Cummings, M. E., & Crothers, L. R. (2013). Interacting selection diversifies warning signals in a polytypic frog: An examination with the strawberry poison frog. *Evolutionary Ecology*, 27(4), 693–710. <https://doi.org/10.1007/s10682-013-9648-9>
- Daly, J. W., Andriamaharavo, N. R., Andriantsiferana, M., & Myers, C. W. (1996). Madagascar poison frogs (*Mantella*) and their skin alkaloids. *American Museum Novitates*, 3177, 1–34.
- Daly, J. W., Garraffo, H. M., Hall, G. S., & Cover Jr, J. F. (1997). Absence of skin alkaloids in captive-raised Madagascar mantelline frogs (*Mantella*) and sequestration of dietary alkaloids. *Toxicon*, 35(7), 1131–1135. [https://doi.org/10.1016/S0041-0101\(96\)00212-7](https://doi.org/10.1016/S0041-0101(96)00212-7)
- Daly, J. W., Garraffo, H. M., Spande, T. F., Giddings, L. A., Saporito, R. A., Vieites, D. R., & Vences, M. (2008). Individual and geographic variation of skin alkaloids in three species of Madagascar poison frogs (*Mantella*). *Journal of Chemical Ecology*, 34(2), 252–279. <https://doi.org/10.1007/s10886-007-9396-9>
- de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, 56(6), 879–886. <https://doi.org/10.1080/10635150701701083>
- Diniz-Filho, J. A. F., Soares, T. N., Lima, J. S., Dobrovolski, R., Landeiro, V. L., de Campos Telles, M. P., ... Bini, L. M. (2013). Mantel test in population genetics. *Genetics and Molecular Biology*, 36(4), 475–485. <https://doi.org/10.1590/S1415-47572013000400002>
- Endler, J. A. (1980). Natural selection on color patterns in *Poecilia reticulata*. *Evolution*, 34(1), 76–91.
- Endler, J. A. (1991). Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions. *Vision Research*, 31(3), 587–608. [https://doi.org/10.1016/0042-6989\(91\)90109-1](https://doi.org/10.1016/0042-6989(91)90109-1)
- Endler, J. A. (2012). A framework for analysing colour pattern geometry: Adjacent colours. *Biological Journal of the Linnean Society*, 107(2), 233–253. <https://doi.org/10.1111/j.1095-8312.2012.01937.x>
- Endler, J. A., & Greenwood, J. J. D. (1988). Frequency-dependent predation, crypsis and aposematic coloration. *Philosophical Transactions of the Royal Society B*, 319(1196), 505–523.
- Endler, J. A., & Mielke Jr, P. W. (2005). Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, 86(4), 405–431. <https://doi.org/10.1111/j.1095-8312.2005.00540.x>
- Ersts, P. J. (2018). *Geographic distance matrix generator (version 1.2.3)*. American Museum of Natural History, Center for Biodiversity and Conservation. Retrieved from http://biodiversityinformatics.amnh.org/open_source/gdmg
- Estrada, C., & Jiggins, C. D. (2008). Interspecific sexual attraction because of convergence in warning coloration: Is there a conflict between natural and sexual selection in mimetic species? *Journal of Evolutionary Biology*, 21(3), 749–760.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Fumagalli, M., Vieira, F. G., Linderoth, T., & Nielsen, R. (2014). ngsTools: Methods for population genetics analyses from next-generation sequencing data. *Bioinformatics*, 30(10), 1486–1487. <https://doi.org/10.1093/bioinformatics/btu041>
- Galeotti, P., Rubolini, D., Dunn, P. O., & Fasola, M. (2003). Colour polymorphism in birds: Causes and functions. *Journal of Evolutionary Biology*, 16(4), 635–646. <https://doi.org/10.1046/j.1420-9101.2003.00569.x>
- Glaw, F., & Vences, M. (2007). *A field guide to the amphibians and reptiles of Madagascar* (3rd ed., p. 496). Cologne, Germany: Vences and Glaw.
- Godfrey, D., Lythgoe, J. N., & Rumball, D. A. (1987). Zebra stripes and tiger stripes: The spatial frequency distribution of the pattern compared to that of the background is significant in display and crypsis. *Biological Journal of the Linnean Society*, 32(4), 427–433. <https://doi.org/10.1111/j.1095-8312.1987.tb00442.x>

- Gray, S. M., & McKinnon, J. S. (2007). Linking color polymorphism maintenance and speciation. *Trends in Ecology and Evolution*, 22(2), 71–79. <https://doi.org/10.1016/j.tree.2006.10.005>
- Guilford, T., & Dawkins, M. S. (1993). Receiver psychology and the design of animal signals. *Trends in Neurosciences*, 16(11), 430–436. [https://doi.org/10.1016/0166-2236\(93\)90068-W](https://doi.org/10.1016/0166-2236(93)90068-W)
- Guillot, G., & Rousset, F. (2013). Dismantling the mantel tests. *Methods in Ecology and Evolution*, 4(4), 336–344. <https://doi.org/10.1111/2041-210x.12018>
- Hauswaldt, J. S., Ludewig, A. K., Vences, M., & Pröhl, H. (2011). Widespread co-occurrence of divergent mitochondrial haplotype lineages in a Central American species of poison frog (*Oophaga pumilio*). *Journal of Biogeography*, 38(4), 711–726. <https://doi.org/10.1111/j.1365-2699.2010.02438.x>
- Heying, H. E. (2001a). *Mantella laevigata* (climbing Mantella). Aborted Predation. *Herpetological Review*, 32, 34.
- Heying, H. E. (2001b). Social and reproductive behavior in the Madagascan poison frog, *Mantella laevigata*, with comparisons to the dendrobatids. *Animal Behaviour*, 61(3), 567–577.
- Hoekstra, H. E., Drumm, K. E., & Nachman, M. W. (2004). Ecological genetics of adaptive color polymorphism in pocket mice: Geographic variation in selected and neutral genes. *Evolution*, 58(6), 1329–1341. <https://doi.org/10.1111/j.0014-3820.2004.tb01711.x>
- Hoffman, E. A., Schueler, F. W., Jones, A. G., & Blouin, M. S. (2006). An analysis of selection on a colour polymorphism in the northern leopard frog. *Molecular Ecology*, 15(9), 2627–2641. <https://doi.org/10.1111/j.1365-294X.2006.02934.x>
- Hugall, A. F., & Stuart-Fox, D. (2012). Accelerated speciation in colour-polymorphic birds. *Nature*, 485(7400), 631–634. <https://doi.org/10.1038/nature11050>
- Huson, D. H. (1998). SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics*, 14(1), 68–73. <https://doi.org/10.1093/bioinformatics/14.1.68>
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–267. <https://doi.org/10.1093/molbev/msj030>
- Hutter, C. R., Andriampemanana, Z. F., Razafindraibe, J. H., Rakotoarison, A., & Scherz, M. D. (2018). New dietary data from *Compsophis* and *Alluaudina* species (Squamata: Lamprophiidae: Pseudoxyrhophiinae), and implications for their dietary complexity and evolution. *Journal of Natural History*, 52(39–40), 2497–2510.
- Huxley, J. (1955). Morphism in birds. In A. Portmann, & E. Sutter (Eds.), *Acta XI Congressus Internationalis Ornithologicae* (pp. 309–328). Basel, Switzerland: Birkhäuser Verlag.
- Jiggins, C. D., Naisbit, R. E., Coe, R. L., & Mallet, J. (2001). Reproductive isolation caused by colour pattern mimicry. *Nature*, 411(6835), 302–305. <https://doi.org/10.1038/35077075>
- Johnson, P. L., & Slatkin, M. (2008). Accounting for bias from sequencing error in population genetic estimates. *Molecular Biology and Evolution*, 25(1), 199–206. <https://doi.org/10.1093/molbev/msm239>
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jovanovic, O., Vences, M., Safarek, G., Rabemananjara, F. C. E., & Dolch, R. (2009). Predation upon *Mantella aurantiaca* in the Torotorofotsy wetlands, central-eastern Madagascar. *Herpetology Notes*, 2, 95–97.
- Kingston, J. J., Rosenthal, G. G., & Ryan, M. J. (2003). The role of sexual selection in maintaining a colour polymorphism in the pygmy sword-tail, *Xiphophorus pygmaeus*. *Animal Behaviour*, 65(4), 735–743. <https://doi.org/10.1006/anbe.2003.2110>
- Korneliusson, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, 15(1), 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Krohn, A. R., & Rosenblum, E. B. (2016). Geographic color variation and physiological color change in eastern collared lizards (*Crotaphytus collaris*) from Southern New Mexico, USA. *Herpetologica*, 72(4), 318–323.
- Legendre, P., & Fortin, M. J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, 10(5), 831–844. <https://doi.org/10.1111/j.1755-0998.2010.02866.x>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, and 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>
- Lynch, M. (2008). Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. *Molecular Biology and Evolution*, 25(11), 2409–2419. <https://doi.org/10.1093/molbev/msn185>
- Maan, M. E., & Cummings, M. E. (2008). Female preferences for aposematic signal components in a polymorphic poison frog. *Evolution*, 62(9), 2334–2345. <https://doi.org/10.1111/j.1558-5646.2008.00454.x>
- Maan, M. E., & Cummings, M. E. (2009). Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45), 19072–19077. <https://doi.org/10.1073/pnas.0903327106>
- Maan, M. E., & Cummings, M. E. (2011). Poison frog colors are honest signals of toxicity, particularly for bird predators. *The American Naturalist*, 179(1), E1–E14.
- Maan, M. E., Eshuis, B., Haesler, M. P., Schneider, M. V., van Alphen, J. J. M., & Seehausen, O. (2008). Color polymorphism and predation in a Lake Victoria cichlid fish. *Copeia*, 2008(3), 621–629. <https://doi.org/10.1643/CE-07-114>
- Mallet, J. (1995). A species definition for the modern synthesis. *Trends in Ecology and Evolution*, 10(7), 294–299. [https://doi.org/10.1016/0169-5347\(95\)90031-4](https://doi.org/10.1016/0169-5347(95)90031-4)
- Mallet, J., & Barton, N. H. (1989). Strong natural selection in a warning-color hybrid zone. *Evolution*, 43(2), 421–431. <https://doi.org/10.1111/j.1558-5646.1989.tb04237.x>
- Mallet, J., & Joron, M. (1999). Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and Systematics*, 30(1), 201–233. <https://doi.org/10.1146/annurev.ecolsys.30.1.201>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet. Journal*, 17(1), 10–12.
- Mayr, E. (1942). *Systematics and the origin of species* (p. 334). New York, NY: Columbia University Press.
- McKay, B. D. (2013). The use of digital photography in systematics. *Biological Journal of the Linnean Society*, 110(1), 1–13. <https://doi.org/10.1111/bij.12086>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., ... DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20, 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Medina, I., Wang, I. J., Salazar, C., & Amézquita, A. (2013). Hybridization promotes color polymorphism in the aposematic harlequin poison frog. *Oophaga histrionica*. *Ecology and Evolution*, 3(13), 4388–4400.
- Müller, F. (1879). *Ituna* and *Thyridia*: A remarkable case of mimicry in butterflies. *Transactions of the Entomological Society of London*, 1879, xx–xxix.
- Nokelainen, O., Hegna, R. H., Reudler, J. H., Lindstedt, C., & Mappes, J. (2012). Trade-off between warning signal efficacy and mating success in the wood tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), 257–265. <https://doi.org/10.1098/rspb.2011.0880>

- Noonan, B. P., & Comeault, A. A. (2009). The role of predator selection on polymorphic aposematic poison frogs. *Biology Letters*, 5(1), 51–54. <https://doi.org/10.1098/rsbl.2008.0586>
- Osorio, D., & Vorobyev, M. (2005). Photoreceptor spectral sensitivities in terrestrial animals: Adaptations for luminance and colour vision. *Proceedings of the Royal Society B: Biological Sciences*, 272(1574), 1745–1752.
- Oxford, G. S. (2005). Genetic drift within a protected polymorphism: Enigmatic variation in color-morph frequencies in the candy-stripe sider. *Enoplognatha ovata*. *Evolution*, 59(10), 2170–2184.
- Pérez-Rodríguez, L., Jovani, R., & Stevens, M. (2017). Shape matters: Animal colour patterns as signals of individual quality. *Proceedings of the Royal Society B: Biological Sciences*, 284(1849), 20162446. <https://doi.org/10.1098/rspb.2016.2446>
- Perry, G., Wallace, M. C., Perry, D., Curzer, H., & Muhlberger, P. (2011). Toe clipping of amphibians and reptiles: Science, ethics, and the law. *Journal of Herpetology*, 45(4), 547–555.
- Posso-Terranova, A., & Andrés, J. A. (2018). Multivariate species boundaries and conservation of harlequin poison frogs. *Molecular Ecology*, 27(17), 3432–3451. <https://doi.org/10.1111/mec.14803>
- Poulton, E. B. (1890). *The colours of animals: Their meaning and use, especially considered in the case of insects*. London, UK: Paul, Trench, Trübner & Co. Ltd.
- Prudic, K. L., Skemp, A. K., & Papaj, D. R. (2007). Aposematic coloration, luminance contrast, and benefits of conspicuousness. *Behavioral Ecology*, 18(1), 41–46.
- Reynolds, R. G., & Fitzpatrick, B. M. (2007). Assortative mating in poison-dart frogs based on an ecologically important trait. *Evolution*, 61(9), 2253–2259. <https://doi.org/10.1111/j.1558-5646.2007.00174.x>
- Robertson, J. M., & Zamudio, K. R. (2009). Genetic diversification, vicariance, and selection in a polytypic frog. *Journal of Heredity*, 100(6), 715–731. <https://doi.org/10.1093/jhered/esp041>
- Rojas, B. (2016). Behavioural, ecological, and evolutionary aspects of diversity in frog colour patterns. *Biological Reviews*, 92(2), 1059–1080. <https://doi.org/10.1111/brv.12269>
- Rojas, B., Devillechabrolle, J., & Endler, J. A. (2014). Paradox lost: Variable colour-pattern geometry is associated with differences in movement in aposematic frogs. *Biology Letters*, 10(6), 20140193.
- Rojas, B., & Endler, J. A. (2013). Sexual dimorphism and intra-population colour pattern variation in the aposematic frog *Dendrobates tinctorius*. *Evolutionary Ecology*, 27(4), 739–753. <https://doi.org/10.1007/s10682-013-9640-4>
- Rojas, B., Rautiala, P., & Mappes, J. (2014). Differential detectability of polymorphic warning signals under varying light environments. *Behavioural Processes*, 109, 164–172.
- Roland, A. B., Santos, J. C., Carriker, B. C., Caty, S. N., Tapia, E. E., Coloma, L. A., & O'Connell, L. A. (2017). Radiation of the polymorphic Little Devil poison frog (*Oophaga sylvatica*) in Ecuador. *Ecology and Evolution*, 7(22), 9750–9762.
- Rosenblum, E. B., & Harmon, L. J. (2011). "Same same but different": Replicated ecological speciation at White Sands. *Evolution*, 65(4), 946–960. <https://doi.org/10.1111/j.1558-5646.2010.01190.x>
- Sandoval, C. P. (1994). Differential visual predation on morphs of *Timema cristinae* (Phasmatodeae:Timemidae) and its consequences for host range. *Biological Journal of the Linnean Society*, 52(4), 341–356. <https://doi.org/10.1111/j.1095-8312.1994.tb00996.x>
- Saporito, R. A., Zuercher, R., Roberts, M., Gerow, K. G., & Donnelly, M. A. (2007). Experimental evidence for aposematism in the Dendrobatid poison frog *Oophaga pumilio*. *Copeia*, 2007(4), 1006–1011. [https://doi.org/10.1643/0045-8511\(2007\)7\[1006:EEFAIT\]2.0.CO;2](https://doi.org/10.1643/0045-8511(2007)7[1006:EEFAIT]2.0.CO;2)
- Schaefer, H. C., Vences, M., & Veith, M. (2002). Molecular phylogeny of Malagasy poison frogs, genus *Mantella* (Anura: Mantellidae): Homoplastic evolution of colour pattern in aposematic amphibians. *Organisms Diversity & Evolution*, 2(2), 97–105.
- 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. <https://doi.org/10.1038/nmeth.2089>
- Sinervo, B., Bleay, C., & Adamopoulou, C. (2001). Social causes of correlational selection and the resolution of a heritable throat color polymorphism in a lizard. *Evolution*, 55(10), 2040–2052. <https://doi.org/10.1111/j.0014-3820.2001.tb01320.x>
- Sinervo, B., & Lively, C. M. (1996). The rock-paper-scissors game and the evolution of alternative male strategies. *Nature*, 380(6571), 240–243. <https://doi.org/10.1038/380240a0>
- Sites Jr, J. W., & Marshall, J. C. (2004). Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics*, 35, 199–227. <https://doi.org/10.1146/annurev.ecolsys.35.112202.130128>
- Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, 195, 693–702. <https://doi.org/10.1534/genetics.113.154138>
- Smit, A. F. A., Hubley, R., & Green, P. (2014). *RepeatMasker Open-4.0*. Retrieved from <http://www.repeatmasker.org>
- Stevens, M., Párraga, C. A., Cuthill, I. C., Partridge, J. C., & Troscianko, T. S. (2007). Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, 90(2), 211–237. <https://doi.org/10.1111/j.1095-8312.2007.00725.x>
- Stevens, M., & Ruxton, G. D. (2012). Linking the evolution and form of warning coloration in nature. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), 417–426. <https://doi.org/10.1098/rspb.2011.1932>
- Stoddard, M. C., & Stevens, M. (2010). Pattern mimicry of host eggs by the common cuckoo, as seen through a bird's eye. *Proceedings of the Royal Society B: Biological Sciences*, 277, 1387–1393. <https://doi.org/10.1098/rspb.2009.2018>
- Summers, K., Symula, R., Clough, M., & Cronin, T. (1999). Visual mate choice in poison frogs. *Proceedings of the Royal Society B: Biological Sciences*, 266(1434), 2141–2145. <https://doi.org/10.1098/rspb.1999.0900>
- Tarvin, R. D., Powell, E. A., Santos, J. C., Ron, S. R., & Cannatella, D. C. (2017). The birth of aposematism: High phenotypic divergence and low genetic diversity in a young clade of poison frogs. *Molecular Phylogenetics and Evolution*, 109, 283–295. <https://doi.org/10.1016/j.ympev.2016.12.035>
- Teasdale, L. C., Stevens, M., & Stuart-Fox, D. (2013). Discrete colour polymorphism in the tawny dragon lizard (*Ctenophorus decresii*) and differences in signal conspicuousness among morphs. *Journal of Evolutionary Biology*, 26(5), 1035–1046.
- Troscianko, J., & Stevens, M. (2015). Image calibration and analysis toolbox – a free software suite for objectively measuring reflectance, colour and pattern. *Methods in Ecology and Evolution*, 6(11), 1320–1331. <https://doi.org/10.1111/2041-210X.12439>
- Twomey, E., Yeager, J., Brown, J. L., Morales, V., Cummings, M., & Summers, K. (2013). Phenotypic and genetic divergence among poison frog populations in a mimetic radiation. *PLoS ONE*, 8(2), e55443.
- Vences, M., Chiari, Y., Raharivoloniaina, L., & Meyer, A. (2004). High mitochondrial diversity within and among populations of Malagasy poison frogs. *Molecular Phylogenetics and Evolution*, 30(2), 295–307. [https://doi.org/10.1016/S1055-7903\(03\)00217-3](https://doi.org/10.1016/S1055-7903(03)00217-3)
- Vences, M., Hille, A., & Glaw, F. (1998). Allozyme differentiation in the genus *Mantella* (Amphibia: Anura: Mantellinae). *Folia Zoologica*, 47(4), 261–274.
- Vieira, F. G., Lassalle, F., Korneliussen, T. S., & Fumagalli, M. (2016). Improving the estimation of genetic distances from next-generation sequencing data. *Biological Journal of the Linnean Society*, 117(1), 139–149. <https://doi.org/10.1111/bij.12511>
- Wang, I. J., & Summers, K. (2010). Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology*, 19(3), 447–458. <https://doi.org/10.1111/j.1365-294X.2009.04465.x>

- Wiley, E. O. (1978). The evolutionary species concept reconsidered. *Systematic Biology*, 27(1), 17–26.
- Willink, B., García-Rodríguez, A., Bolaños, F., & Pröhl, H. (2014). The interplay between multiple predators and prey colour divergence. *Biological Journal of the Linnean Society*, 113(2), 580–589. <https://doi.org/10.1111/bij.12355>
- Wollenberg, K. C., Lötters, S., Mora-Ferrer, C., & Veith, M. (2008). Disentangling composite colour patterns in a poison frog species. *Biological Journal of the Linnean Society*, 93(3), 433–444. <https://doi.org/10.1111/j.1095-8312.2007.00906.x>
- Yang, Y., Richards-Zawacki, C. L., Devar, A., & Dugas, M. B. (2016). Poison frog color morphs express assortative mate preferences in

allopatry but not sympatry. *Evolution*, 70(12), 2778–2788. <https://doi.org/10.1111/evo.13079>

How to cite this article: Klonoski K, Bi K, Rosenblum EB. Phenotypic and genetic diversity in aposematic Malagasy poison frogs (genus *Mantella*). *Ecol Evol*. 2019;9:2725–2742. <https://doi.org/10.1002/ece3.4943>