

Genetic and phenotypic evidence of a contact zone between divergent colour morphs of the iconic red-eyed treefrog

Maria Akopyan^{1,2}  | Zachariah Gompert³  | Karina Klonoski⁴ | Andres Vega⁵ |
 Kristine Kaiser² | Rachel Mackelprang² | Erica Bree Rosenblum⁴  |
 Jeanne M. Robertson² 

¹Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA

²Department of Biology, California State University, Northridge, CA, USA

³Department of Biology, Utah State University, UT, USA

⁴Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA

⁵AMBICOR, Tibás, Costa Rica

Correspondence

Maria Akopyan, Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA.

Email: ma2256@cornell.edu

Funding information

California State University, Northridge

Abstract

Hybrid zones act as natural laboratories where divergent genomes interact, providing powerful systems for examining the evolutionary processes underlying biological diversity. In this study, we characterized patterns of genomic and phenotypic variation resulting from hybridization between divergent intraspecific lineages of the Neotropical red-eyed treefrog (*Agalychnis callidryas*). We found genetic evidence of a newly discovered contact zone and phenotypic novelty in leg colour—a trait suspected to play a role in mediating assortative mating in this species. Analysis of hybrid ancestry revealed an abundance of later-generation F_n individuals, suggesting persistence of hybrids in the contact zone. Hybrids are predominantly of southern ancestry but are phenotypically more similar to northern populations. Genome-wide association mapping revealed QTL with measurable effects on leg-colour variation, but further work is required to dissect the architecture of this trait and establish causal links. Further, genomic cline analyses indicated substantial variation in patterns of introgression across the genome. Directional introgression of loci associated with different aspects of leg colour are inherited from each parental lineage, creating a distinct hybrid colour pattern. We show that hybridization can generate new phenotypes, revealing the evolutionary processes that potentially underlie patterns of phenotypic diversity in this iconic polytypic frog. Our study is consistent with a role of hybridization and sexual selection in lineage diversification, evolutionary processes that have been implicated in accelerating divergence in the most phenotypically diverse species.

KEYWORDS

coloration, gene flow, genomics, hybridization, introgression, treefrogs

1 | INTRODUCTION

Genetic exchange between differentiated lineages (i.e., hybridization) is an important evolutionary process that can promote or constrain adaptation and speciation (Abbott et al., 2016; Anderson & Stebbins, 1954; Arnold, 1997; Gompert et al., 2012;

Grant, 1981; Harrison, 1990; Mallet, 2005; Nolte et al., 2009). For decades, studies on hybrid zones have provided insights into the basis and process of speciation (Barton & Hewitt, 1985; Harrison & Larson, 2016; Hewitt, 2001). Hybrid zones are also relevant for understanding adaptation because gene flow between divergent lineages can facilitate sharing of adaptive alleles (Arnold, 2006;

Fitzpatrick et al., 2009) or provide novel gene combinations that promote adaptation by generating new phenotypes for selection to act upon (Ebersbach et al., 2020; Mallet, 2007; Mavárez & Linares, 2008; Seehausen & Schluter, 2004). Increasing accessibility of population genomic data provides an exciting opportunity to examine the genomic patterns underlying hybridization and its role in these fundamental evolutionary processes (Gompert et al., 2017).

Studies of hybrid zones in the genomic era typically focus on identifying heterogeneous patterns of introgression across the genome (Harrison & Larson, 2016; Ravinet et al., 2017). Because hybrid zones constitute a semipermeable barrier to gene flow, they provide a unique opportunity to test combinations of genotypes from divergent lineages (Harrison & Larson, 2014). The distribution of genotypic variation in hybrid zones is directly influenced by natural selection and can therefore be used to identify genomic regions associated with adaptation or reproductive isolation (Baack & Rieseberg, 2007; Hedrick, 2013). Patterns of introgression are also shaped by recombination, which determines the extent to which selection affects neutral loci linked to causal variants (Martin et al., 2019). Linkage among loci permits detecting patterns of introgression by examining loci linked to causal variants, without knowledge of true causal variants. Genomic regions with restricted levels of introgression compared to the genome-wide average may be linked to loci associated with reproductive isolation, whereas regions with enhanced introgression might be associated with universally adaptive variants (Harrison & Larson, 2016; Payseur, 2010). However, differential introgression among loci need not be a result of selection but can also result from stochastic processes (Fitzpatrick et al., 2009; Gompert et al., 2012). A more powerful method for distinguishing the signature of selection is the combined analysis of the genetic and phenotypic consequences of hybridization, with a focus on divergent phenotypes that are putatively involved in adaptation or reproductive isolation.

The introgression of alleles across a hybrid zone can be (and often is) asymmetric, whereby traits and their associated genes transition more extensively from one parental population into another than in the opposite direction (Harrison & Larson, 2016). Many examples of asymmetric introgression come from hybrid-zone studies examining traits involved in local adaptation (Bay & Ruegg, 2017; Lamichhaney et al., 2015; Pardo-Díaz et al., 2012; Racimo et al., 2015; Walsh et al., 2018). Traits evolving via sexual selection, on the other hand, are typically thought to promote reproductive isolation and are therefore expected to restrict introgression (Bridle & Butlin, 2002). However, adaptive introgression of sexual traits across species boundaries has been observed, for instance, due to mate preference or competition (Lipshutz et al., 2019; Stein & Uy, 2006; While et al., 2015). A consequence of such introgression is the generation of hybrid phenotypes not present in either parental population, which can provide novel adaptive potential for hybrid genotypes. Ecological adaptation of novel hybrid phenotypes is well demonstrated (Lexer et al., 2003; Pereira et al., 2014; Stelkens & Seehausen, 2009), but

examples of novel hybrid phenotypes evolving via sexual selection are limited (but see Jiggins et al., 2001; Medina et al., 2013; Melo et al., 2009). This bias may be due to a historical focus on the consequences of hybridization between divergent species, whereas sexual selection is predicted to be most prevalent when taxa are at an intermediate stage of divergence (Arnold et al., 1996). Thus, examining hybridizing intraspecific lineages can improve our understanding of the role of sexual selection in facilitating introgression and the evolution of phenotypic diversity.

Red-eyed treefrog populations exhibit substantial phenotypic variation across a broad, primarily central American distribution (Duellman, 2001), as well as microgeographic variation among populations in Costa Rica and Panama. Many of the traits known to differ among red-eyed treefrog populations are social signals that probably evolved via sexual selection, including body size and colour pattern of flanks and legs (Robertson & Greene, 2017; Robertson & Robertson, 2008), male advertisement call and female courtship behavior (Akopyan et al., 2018), and skin peptides (Davis et al., 2016). Along the Caribbean coast of Costa Rica and Panama, red-eyed treefrog populations exhibit high levels of phenotypic differentiation despite low genetic divergence (Robertson et al., 2009; Robertson & Zamudio, 2009). All red-eyed treefrogs found in northeastern Costa Rica have blue legs (northern phenotype; Figure 1; sites 1–4), while individuals at the border of eastern Costa Rica and Panama have multicolour legs (blue and red; southern phenotype; Figure 1; sites 9–11). Previous work based on microsatellite analysis showed that the two colour morphs also represent two genetic demes, with evidence of admixture between them, particularly at sampling boundaries (Figure 1 site 4 vs. 9; Robertson et al., 2009). These patterns suggest a role for both gene flow and selection in mediating phenotypic diversification among red-eyed treefrog populations.

In June 2015, we discovered and sampled a centrally located region in Costa Rica between the previously described northern and southern sites. This putative contact zone includes four sampling localities (Figure 1; sites 5–8), each containing multiple phenotypes: blue legs (as in the north), red/blue legs (as in the south), and a predominant intermediate purple-leg morph. The discovery of this polymorphic region containing novel phenotypes presents the opportunity to examine the evolutionary processes that facilitate phenotypic and genetic divergence. To explore the role of hybridization in lineage diversification in red-eyed treefrogs, we characterize patterns of phenotypic and genomic variation across this newly discovered contact zone. We use genome-wide sampling of single nucleotide polymorphisms (SNPs) identified within restriction site-associated DNA (RAD) tags to estimate population genetic structure and hybrid ancestry. We also quantify and compare leg colour across the contact zone, perform a genome-wide association study to identify genetic variants statistically associated with leg colour variation, and determine whether loci associated with colour pattern exhibit exceptional patterns of introgression. This work is an important step for the interpretation of the historical processes that may have led to these patterns of genetic and phenotypic diversity as well as their evolutionary consequences.

FIGURE 1 Population sampling of *Agalychnis callidryas* in Costa Rica and Panama. Pins indicate population localities. Digital drawings of the posterior aspect of frog legs show a simplified version of the geographic variation and distribution in colour pattern. Red-eyed treefrogs found in the north have blue legs (sites 1–4), while individuals from the south have multicolour legs (blue/red; sites 9–11). The contact zone between the north and the south includes four sampling localities (sites 5–8). These populations each contain multiple phenotypes: blue legs (as in the north), red/blue legs (as in the south), and an intermediate purple leg morph. Actual photographs of frog legs shown in Figure 2

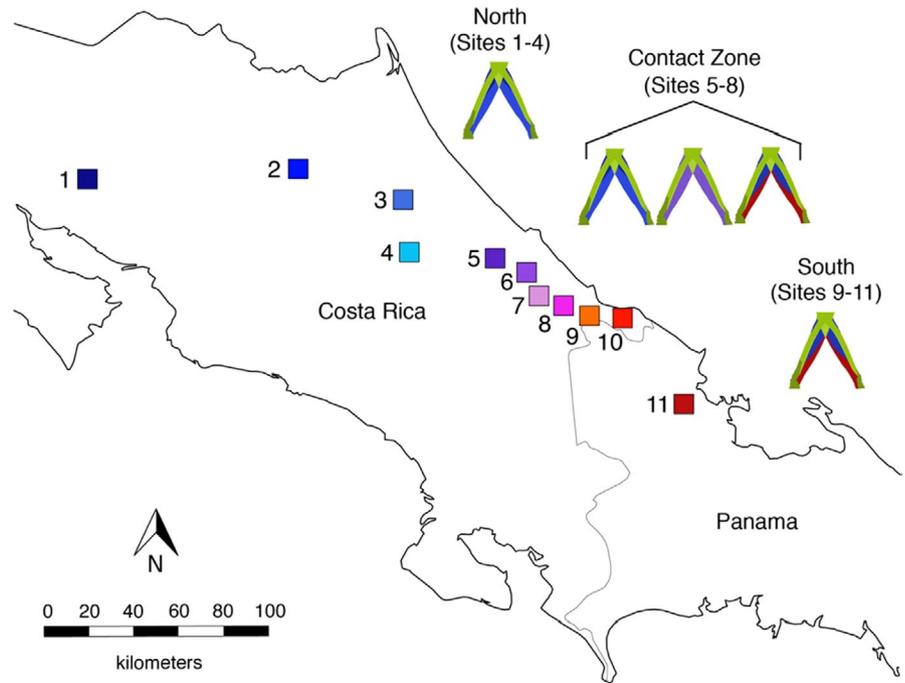


TABLE 1 Sample size and location for each sampling site

Region	Site	Site name	Province, country	<i>n</i> (colour)	<i>n</i> (SNP)	<i>n</i> (GWAS)	Latitude	Longitude
North	1	Tilarán	Guanacaste, CR	16	5	0	10.51622	-84.96017
	2	La Selva	Heredia, CR	26	10	15 ^a	10.43281	-84.00807
	3	EARTH University	Limón, CR	0	5	0	10.23685	-83.56718
	4	Guayacan	Limón, CR	0	10	0	10.01343	-83.56513
Contact Zone	5	Liverpool	Limón, CR	20	20	20	9.92806	-83.17503
	6	Bananito	Limón, CR	40	30	30	9.82970	-83.05202
	7	Armenia	Limón, CR	20	20	20	9.74051	-83.00606
	8	Dandiri	Limón, CR	20	20	20	9.71153	-82.85757
South	9	Cahuita	Limón, CR	0	9	0	9.71898	-82.81483
	10	Manzanillo	Limón, CR	19	10	10	9.63322	-82.65567
	11	Almirante	Bocas del Toro, PA	12	9	8	9.19802	-82.34455
				<i>n</i> = 173	<i>n</i> = 148	<i>n</i> = 123		

Note: See Figure 1 for site location on map.

^aIncludes *n* = 10 individuals assigned a population average from nongenotyped individuals.

2 | MATERIALS AND METHODS

2.1 | Sampling

We genotyped 148 individuals from 11 sites and measured colour pattern in 173 individuals from eight sites throughout Costa Rica and Panama. A total of 113 individuals from seven sites had both genotype and phenotype data. Tissue samples and photographs from northern and southern populations were previously sampled by Robertson et al. (2009). Sampling sites included in this study can be seen in Figure 1, and site information including sampling sizes are detailed in Table 1. In 2015, we haphazardly sampled an additional 100 individuals from four contact-zone sites (*n* = 20–30 per site)

during the breeding season (May–August). We captured and took digital photographs of adult frogs, which were later scored for colour pattern (described below). Toe clips were also collected and stored in 100% ethanol for use in genetic analyses. All frogs were released at their capture site. Animal handling procedures were approved by the Animal Care and Use Committee at California State University, Northridge (IACUC 1415-007a).

2.2 | Quantifying phenotypic variation

We measured phenotypic variation in red-eyed treefrog leg colour using digital photographs of 173 individuals, including 73

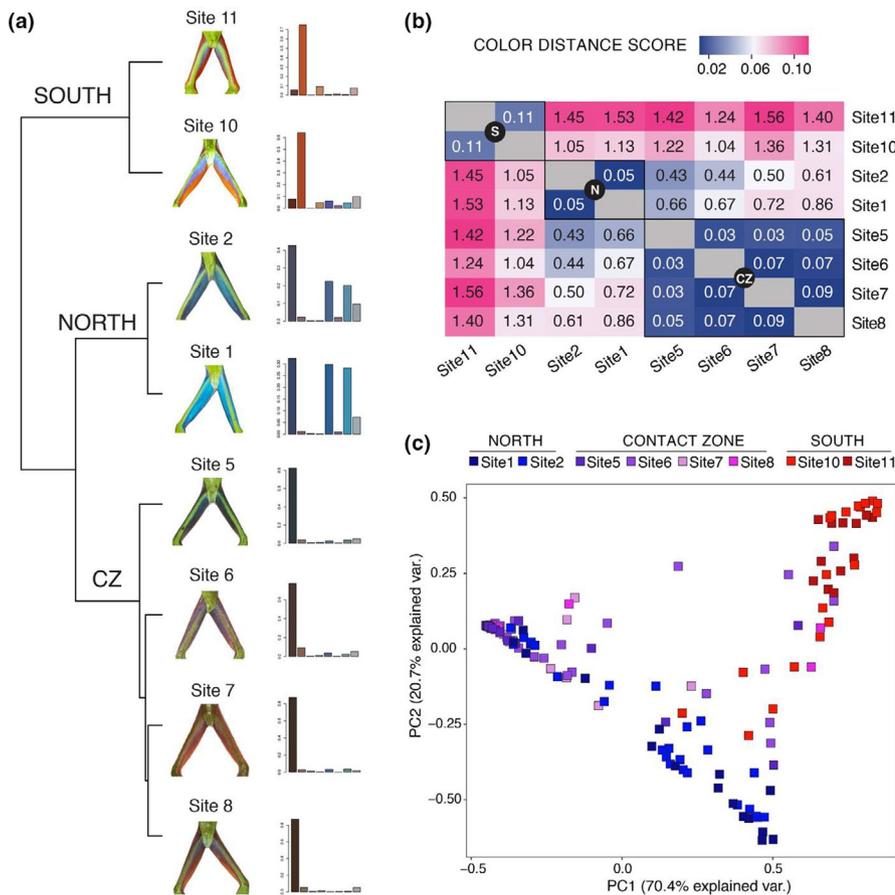


FIGURE 2 Colour similarity analysis of leg colour. (a) Dendrogram based on cluster analysis of the colour distance matrix clustered by similarity using χ^2 distance in colordistance. Branch lengths are proportional to the χ^2 distances between sites. Images of frog legs depict a representative individual from each sampling site, and colour histograms show the proportion of pixels in each of eight colour bins averaged for each site. Histogram bins are coloured by the average colour of the pixels in each bin. (b) Heatmaps of the colour distance matrix where dark blue is more similar, and magenta is more dissimilar. (c) Scatterplot summarizing patterns of leg-colour variation based on a principal component analysis of the percent of pixels in colour bins. Points denote individuals and are coloured based on site locality

individuals from four previously studied populations and 100 individuals from the four contact-zone sites. All frogs were photographed indoors on a grey background using a Canon PowerShot SX160 digital camera fit with a 16x optical zoom lens (28–35 mm). We first imported photographs of each frog into Adobe Photoshop CC v. 18.0.0 (San Jose, CA) to correct for ambient light by reference to a black-white-grey standard placed in each photo (QPCard 101) and cropped images to only include the posterior surface of the thigh where colour pattern varies (Robertson & Robertson, 2008). Images were then loaded into R (v. 3.5.2) using package colordistance (Weller & Westneat, 2019), which treats images as a set of three-dimensional points, where each pixel is a coordinate in RGB (red, green, and blue) colour space. Pixel coordinates were binned, and the resulting data included an RGB triplet and percent of pixels for each of eight bins for each individual. To compare leg colour among our eight sampled sites, we calculated an average distribution for each site and generated a colour distance matrix using the χ^2 method of comparing histogram distributions in colordistance. We conducted a principal component analysis to visualize and infer patterns of leg-colour variation among regions based on the percent values for the eight bins, and conducted a discriminant function analysis to test whether region of origin is predicted for each individual based on the percentage of pixels in each of the colour bins in R (v. 3.5.2).

2.3 | DNA extraction and library preparation

DNA was extracted from toe clips using a DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) generally following the manufacturer's instructions, but modified as described below. We added 4 μ l of 1 mg/ml RNase A to each sample after lysis to remove RNA contamination. DNA was eluted in 60 μ l of AE buffer to maximize concentration. DNA was quantified using a Qubit 2.0 Fluorometer (ThermoFisher, Waltham, MA), and the quality was assessed by visualization on an agarose gel. Two RAD libraries were prepared following the protocol outlined in Ali et al. (2016) using 50 ng of input DNA from each individual. Each library was sequenced on one lane of an Illumina HiSeq 4000 (2 \times 150 bp paired end).

2.4 | Quality filtering and SNP calling

Raw reads were flipped to salvage the reads that were read in the wrong direction (i.e., reads that started with the enzyme cut site on the reverse read instead of the forward read) using a custom perl script. PCR clones were identified and removed by comparing paired-end reads with the clone_filter program from the Stacks (v. 1.34) pipeline (Catchen et al., 2013). Next, we used process_radtags in Stacks to discard low-quality reads (phred quality score

<10; 0.7% of reads removed) and reads with ambiguous barcodes or RAD cut sites (0.8% of reads removed). The remaining 98.5% of reads were demultiplexed to produce fastq files for each individual. All reads from the 148 individuals were pooled and used for a de novo assembly in *ustacks*. The minimum depth of coverage required to create a stack (i.e., an identical set of sequences) was set to three (*-m* option). The maximum number of pairwise differences allowed between any two stacks within a locus (i.e., contig) was set to four (*-M* option).

Because we lacked a reference genome, we assembled a partial de novo pseudo-reference genome using the catalogue of all loci across all populations generated by *Stacks* with a custom perl script (as in Gompert et al., 2014), requiring a minimum of $N = 74$ depth (i.e., half of the individuals in our data set must have that stack). We then aligned all sequenced reads to this pseudo-reference genome using the *aln* and *samse* algorithms in *BWA* v. 0.7.15-r1142 (Li & Durbin, 2009). We allowed for a maximum edit distance of five, with a read trimming parameter of 10, the seed set to 20, and a maximum edit distance in the seed of two. We used *SAMTOOLS* v. 1.3.1 (Li et al., 2009) to compress, sort and index the individual alignments. Next, we identified SNPs from the sequence data by calculating the Bayesian posterior probability that each nucleotide was variable with *SAMTOOLS* and *BCFTOOLS* v. 1.3.1. When identifying variants, we used the full prior with θ set to 0.001 and designated genetic variants at nucleotide sites where we had sequence data for at least 80% of the sampled individuals and where the posterior probability of the sequence data under a null model that the nucleotide was invariant was <0.01. We further filtered our data to remove sequence reads with low depth of coverage (<2x per individual) or a minor allele frequency less than 0.5%. Finally, we used the resulting VCF file to estimate posterior genotype probabilities from genotype likelihoods and an allele frequency prior assuming Hardy-Weinberg equilibrium, then called genotypes at variants with data from at least five individuals using a custom perl script. We used the posterior mean genotype as a point estimate for genotypes based on the posterior genotype probabilities obtained from a population allele frequency model. Input files and scripts are available at <https://doi.org/10.6084/m9.figshare.11923017.v2>.

2.5 | Population structure, hybrid ancestry, and gene flow

We calculated genetic diversity as the observed frequency of heterozygotes (H_o) within populations, then calculated pairwise genetic distance (F_{ST}) between populations in R (v. 3.5.2) using package *ade4* (Jombart, 2008). To visualize and infer patterns of population structure, we conducted a PCA on the centered, unscaled matrix of genotype point estimates. We ran *STRUCTURE* v. 2.3.4 (Pritchard et al., 2000) 10 times from $K = 1$ to $K = 11$ with the burnin length of 100,000 generations followed by 200,000 Monte Carlo Markov chain (MCMC) iterations, using the admixture model with correlated allele frequencies. The optimal K value was estimated by

the maximum value of ΔK following the Evanno method (Evanno et al., 2005), implemented in *STRUCTURE HARVESTER* (Earl, 2012). Alignment of cluster assignments across replicate analyses was then conducted in *CLUMPP* v. 1.1.2 (Jakobsson & Rosenberg, 2007).

We further quantified population structure and examined genetic admixture using *ENTROPY* (v. 1.2) (Gompert et al., 2014), a Bayesian method similar to the admixture model implemented in *STRUCTURE* that deals with uncertainty in genotypes and includes a more complex model for estimating the proportion of ancestry of a set of individuals from contributing populations. To determine the ancestry of the putative red-eyed treefrog hybrids in the contact zone, we estimated the fraction of loci that combine ancestry from these two parental populations in *Entropy* (interpopulation ancestry, denoted Q_{12}). We fit the model with $k = 2$ source populations, as we were interested in two divergent lineages and their hybrids, and ran MCMC chain with 10,000 steps, a burnin of 2,000 iterations, and a thinning interval of four.

To determine whether the system represents primary (no history of geographic isolation) or secondary (contact after isolation) divergence with gene flow, we quantified patterns of gene flow and isolation-by-distance among lineages using two complementary approaches. First, we fit a Bayesian linear mixed model to determine whether genetic differences between populations were best explained by geographic distances, boundaries between nominal lineages (north and south), or both (as in Gompert et al., 2014). We ran a Bayesian regression analysis based on the mixed model framework to account for the correlated error structure inherent in pairwise observations such as genetic distances. We fit a model for logit-transformed F_{ST} as a function of log geographic distance (great circle distance) and taxon distance (defined as 0 for populations from the same nominal lineage and one for populations from different nominal lineages) using MCMC via the *rjags* (v. 4.8) interface with *JAGS* (v. 4.3). Full models were compared to models with only geographic or taxon distance using deviance information criterion (DIC). We ran three chains for each model, each with 10,000 iterations, a 2,000 iteration burnin, and a thinning interval of five.

We also estimated relative effective migration rates among the populations based on a population allele frequency covariance matrix, which we calculated from all 58,697 SNPs using the program *eems* v. 0.0.0.9 (Petkova et al., 2016). This method identifies regions in space with low or high gene flow relative to a simple two-dimensional stepping-stone isolation-by-distance model. We assumed 500 demes on a triangular grid, and fit the model using an MCMC chain with 3 million sampling iterations, 1 million burnin iterations, and a thinning interval of 10,000. If the study system can be described by a simple isolation-by-distance model, this would suggest primary divergence with limited dispersal within a single taxon, whereas evidence of restricted gene flow between these two entities would suggest secondary contact and admixture between the north and south.

Next, we asked whether linkage disequilibrium (LD) was elevated in the contact zone, as predicted by theory (Barton, 1982). We calculated the Pearson correlation coefficient between

genotypes (as a metric of LD) using a set of ancestry informative SNPs, that is, SNPs with an allele frequency difference of at least 0.3 between the north and south reference populations. We polarized genotypes such that positive Pearson correlations (i.e., positive LD) coincides with an association between alleles more common in the north and south reference populations (i.e., coupling alleles), and negative LD coincides with positive associations between the north and south alleles (i.e., repulsion alleles). Ongoing hybridization should result in higher positive estimates of LD for both linked and unlinked markers due to admixture LD. Correlations were calculated separately for populations with a sample size of at least 10 individuals, and again after combining into the three regions (north, contact zone, south) in R (v. 3.5.2).

2.6 | Genomic cline analysis

We fit a Bayesian genomic cline model to quantify genome-wide variation in introgression across the contact zone. This method estimates clines in ancestry for individual genetic loci (e.g., SNPs) along a genome-average admixture gradient (Gompert & Buerkle, 2011; Gompert et al., 2017). As such, it can be applied in cases where hybrid zones are confined to a single geographic locality, or do not have a well defined geographic admixture gradient. Deviations between genome-average introgression and introgression for each locus are measured with two cline parameters, α and β . Cline parameter α denotes an increase (for positive values of α) or decrease (for negative values of α) in the probability of ancestry from reference (parental) lineage 1 relative to null expectations from an individual's hybrid index. Genomic cline parameter β describes an increase (positive values) or decrease (negative values) in the rate of transition from parental lineage 0 ancestry to parental lineage 1 ancestry along the genome-wide admixture gradient. When placed in a geographic context, α is equivalent to twice the shift in cline center, and β measures the decrease (or increase when $\beta < 0$) in cline width relative to the average (Gompert et al., 2017; Szymura & Barton, 1986). Cline parameters can be affected by genetic drift and selection in hybrids (Gompert et al., 2012, 2017).

Genomic clines were estimated for the contact-zone populations (sites 5–8, $N = 90$) with sites 1–4 (north) and 9–11 (south) used as reference parental populations 1 and 2. Clines were fit using BGC (version 1.04b) (Gompert & Buerkle, 2012). We fit the model for a set of ancestry-informative SNPs, specifically, SNPs with an allele frequency difference of at least 0.2 between the north and south reference populations. We identified 8002 SNPs that met this criterion. The model accounts for uncertainty in genotypes, and to take advantage of this, we used the genotype likelihoods output from BCFTOOLS. We ran five MCMC chains to estimate the cline parameters. Hybrid indexes were estimated during a prerun stage that fixed cline parameters at 0. This allows for direct (via the Gibbs algorithm) samplings of hybrid indexes from the conditional posteriors. This was done with a 1,000 step

burnin followed by 3,000 sampling steps. Following this, the main chains were run. We ran a 5,000-step burnin followed by 25,000 sampling steps, which were thinned by only retaining every fifth sample. Posterior estimates were obtained based on the combined output from the five MCMC chains.

2.7 | Genome-wide association mapping of colour variation

We conducted genome-wide association mapping to identify associations between genetic variants (SNPs) and leg colour (percent of pixels in each of eight RGB bins). This top-down approach of identifying candidate loci is a first step to establishing a phenotype-genotype-selection link by describing the genetic architecture of a trait of interest. Due to low sample sizes, we analysed all individuals for which we had genetic and phenotypic data, from both parental and contact-zone populations. Of the 148 individuals genotyped for this study, we had photographs of 113 individuals for phenotypic analysis (Table 1). To deal with small sample sizes of phenotypic data from the blue-legged northern individuals, we measured leg colour in 10 additional nongenotyped individuals (Table 1). We then averaged the percent of pixels in each of the eight bins across these 10 individuals and assigned the mean values to 10 genotyped individuals from the same site, resulting in 123 phenotyped individuals for association mapping. A major challenge of analysing distantly related populations is the risk of spurious genotype-phenotype associations due to population structure (Freedman et al., 2004; Lander & Schork, 1994). To reduce the risk of false positives, we conducted association mapping with raw phenotype values, and with residual values after regressing phenotype on the first principal component of the genotype matrix to account for the confounding effects of population structure (Price et al., 2010). Here, PC1 corresponds with the genome-average admixture gradient. We implemented Bayesian sparse linear mixed models (BSLMMs) in GEMMA (Zhou & Stephens, 2012), which combines linear-mixed models with sparse regression models allowing for various genetic architectures (e.g., many small effect genes versus few large effect genes). This model also controls for more subtle population structure by including a relatedness matrix as a covariate (Zhou & Stephens, 2012). We performed a Markov-chain Monte Carlo (MCMC) chain consisting of five million steps, with the first million discarded as burnin. Parameter values were recorded every 40 steps. We ran nine independent MCMC chains for each colour bin. The output includes estimates aspects of the genetic architecture for each (colour) trait, specifically of the proportion of phenotypic variation that can be explained by the combined effects of SNPs with near infinitesimal and measurable (i.e., detectable) phenotypic effects (proportion of total variance explained; PVE), the proportion of the explained phenotypic variation resulting from measurable-effect SNPs alone (PGE), and the number of genetic variants with measurable effects on the phenotype of interest (n -gamma). Additionally, estimates of individual SNP effects are obtained, along with the posterior inclusion probability for each SNP,

that is the posterior probability that the SNP is associated with the trait of interest.

2.8 | Clines for colour-associated SNPs

We next asked whether the number of colour-associated SNPs differed between the sets of loci with excess northern ($\alpha > 0$) versus excess southern ($\alpha < 0$) ancestry in the contact zone. We would predict such a pattern if selection on colour was (at least in part) responsible for patterns of introgression in the contact zone. We focused on the set of SNPs with patterns of introgression that deviated from the genome-wide average based on genomic cline parameter α (that is, where the 95% equal-tail probability intervals for α excluded 0) as there was no compelling evidence of nonzero values for cline parameter β . We summarized this signal of association by summing posterior inclusion probabilities across SNPs with $\alpha > 0$ or $\alpha < 0$; this represents an estimate of the number of colour QTL included in the set of exceptional introgression SNPs (Gompert et al., 2019; Guan & Stephens, 2011; Lucas et al., 2018). We then calculated the difference in the number of QTL for $\alpha > 0$ minus $\alpha < 0$ SNPs. This observed value is the nominal excess in the number of QTL for northern alleles (or negative for southern alleles). Null expectations were obtained by randomly and repeatedly (1,000 times) sampling an equal number of SNPs and from those calculating the summed posterior inclusion probability for northern versus southern introgression loci. We only considered SNPs with an allele frequency difference of at least 0.2 between the north and south reference populations (i.e., ancestry informative SNPs). We ran these randomization tests in R (v. 3.4.2).

3 | RESULTS

3.1 | Phenotypic variation

Phenotypic analyses revealed differentiation in leg colour among geographic regions (Figure 2). PCA of leg colour based on the percent of pixels in eight colour bins showed that the first two principal components accounted for 91.1% of the total variance (Figure 2c). Principal component (PC) 1 has a high negative loading for percent of pixels in the first bin (-0.8547), while PC2 has a high positive loading for percent of pixels in the second bin (0.7174). Contact zone populations were separated from southern populations along PC1, and separation between northern and southern populations is seen along PC2. Analyses of colour similarity revealed that sites from the south group together but have the highest distance score of any between-site comparison within a region and are the most different compared to the other two regions (Figure 2). The two northern sites with blue legs also group together and are closer in colour to the contact-zone sites, which also cluster together but with varying levels of colour similarity between sites. The discriminant function analysis of leg colour

TABLE 2 Classification matrix of individuals based on colour data

Percent in colour bins	Actual			n
	North (%)	Contact zone (%)	South (%)	
Predicted North	71.4	3.0	0.0	42
Predicted Contact Zone	28.6	91.0	12.9	100
Predicted South	0.0	6.0	87.1	31

Notes: Diagonal values (in bold) indicate percentage correctly classified. Assignment based on colour data was inferred from linear discriminant analysis of percent pixels in each of eight colour bins.

based on percent of pixels in each of the bins correctly assigned the majority of individuals to natal regions (Table 2). When the algorithm misclassified individuals from the northern region, it assigned them to the contact zone, while misclassified individuals from the contact zone were assigned to both the north and south. All individuals from the south were accurately assigned.

3.2 | Genetic variation

The number of raw reads per individual are detailed in Table S1. Our final genetic data set included 58,697 SNPs with an average of 15.4x depth of coverage per individual per SNP (± 17.3 SD). Genome-wide measures of genetic diversity for all nucleotides indicate lower levels of observed heterozygosity in the north and south (averaged across populations: H_O (north) = 0.11, H_O (south) = 0.123, $SE = 0.001$), compared to elevated levels of observed heterozygosity in contact-zone populations ($H_O = 0.135$, $SE < 0.001$). Measures of pairwise genetic distance (F_{ST}) between the 11 sites ranged from 0.0–0.237 (Table 3). F_{ST} estimates are greatest between sites at opposite ends of the cline and lowest between contact-zone sites (F_{ST} not significantly greater than zero between all pairwise comparisons of contact-zone sites). The PCA (Figure 3a) shows genetic clustering that mirrors the geographic distribution of populations (Figure 1). Specifically, individuals from the north cluster along PC1, and separation between the south and contact zone occurs along PC2.

Bayesian structure analysis identified two genetic clusters according to ΔK ($K = 2$; Figure S1), supporting the presence of two main lineages. Red-eyed treefrogs from the contact zone have unique genotypes that include admixture from the north and south lineages, but with a greater contribution from the south (Figure 3b). Populations in the north and south are assigned to a single genetic deme each, with admixture from the contact zone. Populations in the north and south show a clinal change in deme membership. Because we used the full SNP data set to estimate admixture and ancestry, patterns of population structure within parental lineages are reflected in our results (Figure 3). We found evidence of restricted gene flow between the north and south and show that barriers to gene flow between these two entities differ from a simple

TABLE 3 Pairwise F_{ST} estimates (below diagonal) and pairwise geographic distance (km) (above diagonal) for 11 sampling sites

		North				Contact zone				South		
	Site	1	2	3	4	5	6	7	8	9	10	11
North	1		105	156	163	206	223	231	247	251	271.0	322
	2	0.067		53.0	67.3	107	124	134	150	153.0	173	229
	3	0.094	0.018		24.9	55.0	72.4	82.7	97.3	101	120	177
	4	0.118	0.045	0.013		43.8	59.9	68.4	84.6	88.6	108	162
Contact zone	5	0.169	0.101	0.064	0.026		17.4	27.9	42.4	45.9	65.8	122
	6	0.144	0.080	0.044	0.015	0.001		11.1	25.1	28.8	48.7	105
	7	0.166	0.098	0.061	0.025	0.000	0.000		16.6	21.1	40.3	94.5
	8	0.145	0.079	0.043	0.015	0.001	0.000	0.002		4.8	23.8	80.3
South	9	0.194	0.125	0.086	0.042	0.005	0.007	0.004	0.008		19.9	77.7
	10	0.214	0.140	0.098	0.053	0.010	0.015	0.012	0.015	0.010		59.3
	11	0.237	0.166	0.128	0.079	0.030	0.035	0.032	0.036	0.029	0.024	

Note: Bolded F_{ST} estimates are not significantly different from zero.

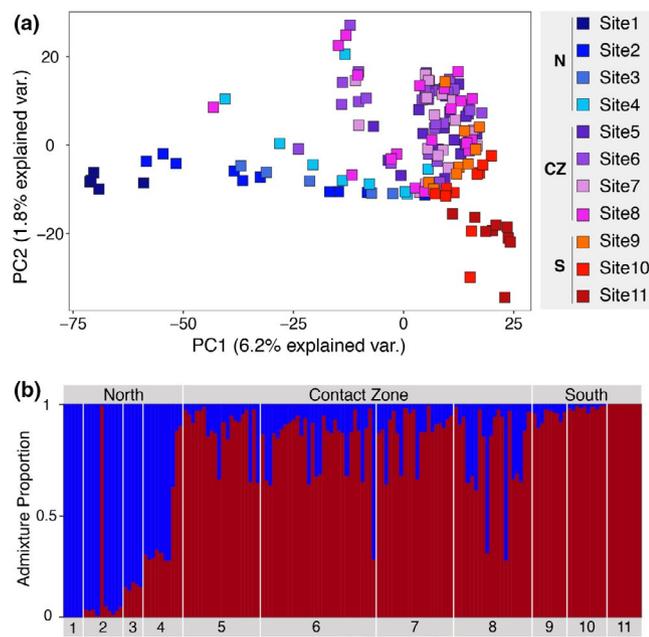


FIGURE 3 (a) Summary of population genetic structure based on principal component axis one (PC1) and axis two (PC2). These axes explain 6.2% (PC1) and 1.8% (PC2) of the variation in genotype estimates for 58,697 variable nucleotides. Points denote individuals and are coloured based on site locality. Individuals from the north load on PC1, and separation of the south and contact zone occurs along PC2. (b) Barplot of admixture proportions for $K = 2$. Each bar corresponds to an individual, and each coloured segment depicts the proportion of an individual's genome inherited from one of the two inferred source populations. White vertical lines separate sampling sites

isolation-by-distance pattern (Figures S2 and S3). Genetic distances between the two lineages are higher than expected from geographic distances alone, and reflect a pattern of isolation-by-distance within each lineage (Figure S2). Similarly, estimates of effective migration rates show that gene flow from the north and south into the contact

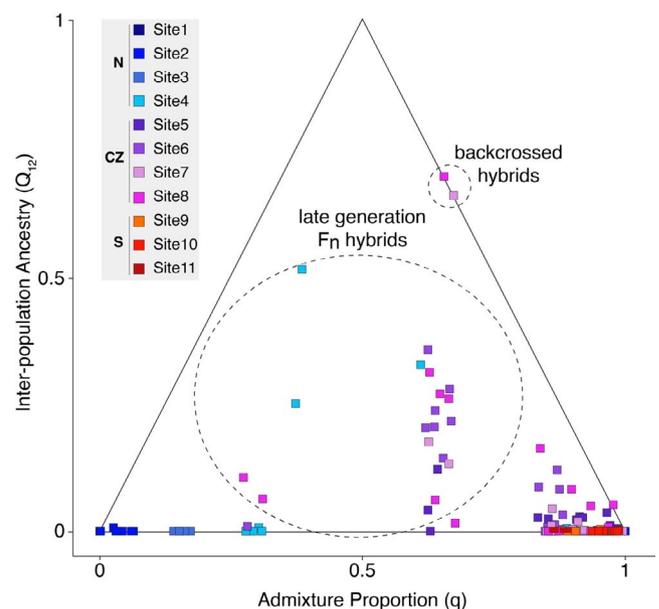


FIGURE 4 Scatterplot of admixture proportion (q) and the fraction of loci at which individuals have ancestry from both parental regions (Q_{12}). Points denote individuals and are coloured based on site locality. Individuals from the north are assigned to ancestor 0 and the south is assigned to ancestor 1; individuals that fall along the x-axis (intermediate admixture proportions, with no interpopulation ancestry) are interpreted as geographic isolation by distance. The majority of individuals in the contact zone are F_n hybrids, with only two backcrossed hybrids

zone is reduced relative to geography (Figure S3). This provides support for secondary contact and admixture between the north and south, rather than primary divergence with limited dispersal within a single taxon.

Furthermore, we found evidence of elevated linkage disequilibrium in two of the four contact-zone populations compared to parental populations (Figure S4). For all populations we analysed (sample

size of at least 10), positive genotypic correlations indicate a positive association between pairs of alleles that were more common in each of the parental populations (i.e., coupling LD). With the exception of site 8 ($r = 0.05$, $SE = 0.05$), contact-zone populations exhibit more positive correlations, with the highest level of LD found in site 7 ($r = 0.41$, $SE = 0.07$). Genetic correlations measured after grouping all populations show a similar distribution of LD (Figure S4), with slightly elevated estimates as expected when combining populations.

3.3 | Hybrid ancestry

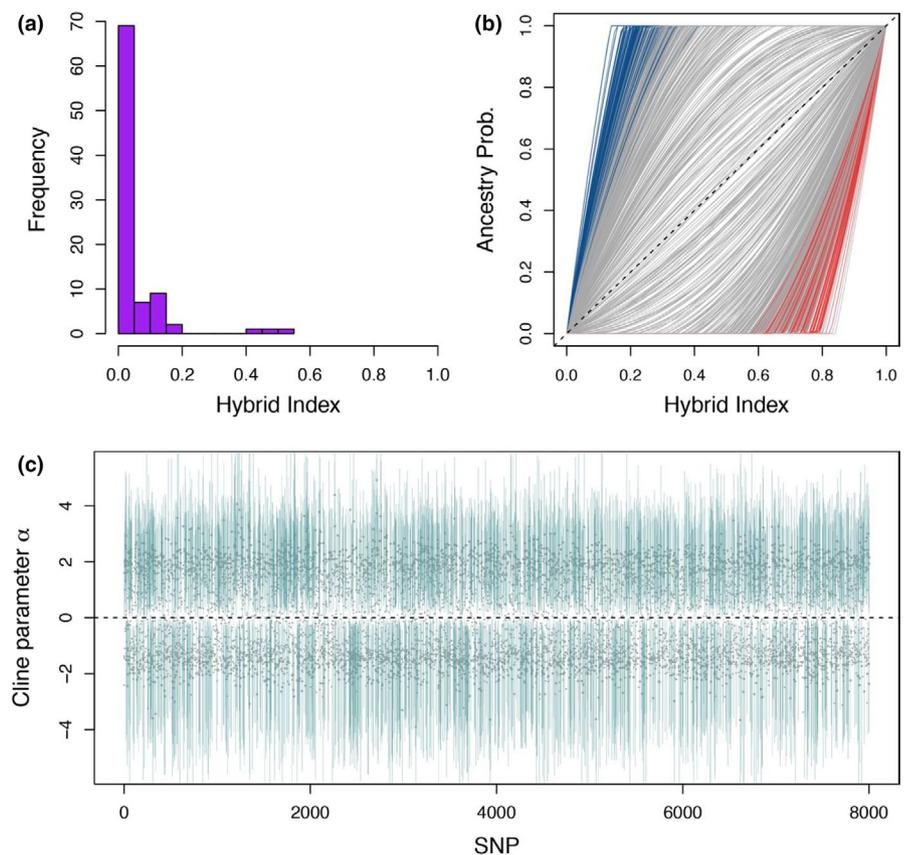
The combination of admixture proportion (q) and interpopulation ancestry (Q_{12}) (Figure 4) summarizes the genomic composition of hybrids at the contact zone. We found two hybrid frogs backcrossed with parentals in the south, which were identified as individuals with maximal Q_{12} for a given q (Gompert et al., 2014). The remaining individuals from the contact zone represent later-generation F_n hybrids (i.e., hybrids that are not offspring of a backcross with parental populations). These hybrids span a large range of admixture proportions, more similar to that of southern populations, probably resulting from historical backcrossing with parentals from the south. Some of these individuals have very low interpopulation ancestry despite their intermediate admixture proportions, which suggests they lack recent nonadmixed ancestors (interspecific) ancestry decays towards zero when hybrids mate with hybrids over multiple generations; Gompert, 2016). Three individuals from the north also appear

to be F_n hybrids. These individuals were from site 4 (Figure 1), which is adjacent to the first contact-zone site (site 5). The remaining individuals from the north and south regions span a range of admixture proportions with minimal interpopulation ancestry. The absence of interpopulation ancestry in these populations suggests that their non-zero admixture proportions reflect population structure within these lineages (i.e., a pattern of isolation by distance within taxa).

3.4 | Differential introgression

Patterns of introgression varied across the genome (Figure 5). Of the 8,002 ancestry-informative SNPs, we had credible evidence of excess northern ancestry at 1518 SNPs (i.e., lower bound of the 95% CI for $\alpha > 0$). In other words, for 1,518 SNPs we detected increased introgression of northern ancestry (alleles) into the contact zone relative to genome-average admixture. Similarly, we found credible evidence of excess southern ancestry at 978 SNPs (i.e., increased introgression of southern alleles into the contact zone; upper bound of the 95% CI for $\alpha < 0$) (Figure 5c). Moreover, many estimates of α were large (mean $|\alpha| = 1.42$), suggesting meaningful differences in introgression among loci. For example, based on the point estimates of α , for an individual with a hybrid index of 0.5, the probability of northern ancestry is ≈ 1 at 3,038 SNPs and ≈ 0 at 2,972 SNPs. Perhaps more relevant, for an individual with a hybrid index of 0.04 (the mean in the hybrid zone, which corresponds with mostly southern ancestry with 4% northern introgression), one

FIGURE 5 (a) Histogram showing distribution of hybrid indices in the contact zone. (b) Plot of estimated genomic clines for 8,002 ancestry-informative SNPs. Solid lines denote estimated probability of northern ancestry for a SNP. Blue lines denote cases of excess northern ancestry (95% CIs for $\alpha > 0$) and red lines denote excess southern ancestry (95% CIs for $\alpha < 0$). The dashed line gives the null expectation based on genome-wide admixture, and grey lines denote clines not credibly different from the genome-average. (c) Plot showing distribution of estimates for cline parameter α across ancestry-informative SNPs. Points denote posterior medians, and vertical lines indicate 95% credible intervals. A dashed horizontal line shows $\alpha = 0$



SNP has a 50% chance of northern ancestry (>10 times the genome-average expectation), 130 SNPs have at least a 25% chance of northern ancestry (>5 times the genome-average expectation), and 3,650 SNPs are completely devoid of northern ancestry (i.e., fixed for southern alleles). In contrast to our results based on cline parameter α , we failed to detect credible evidence of restricted introgression at individual SNPs relative to genome-wide average admixture (95% CIs for β spanned 0 for all SNPs). These genomic analyses were generally consistent with geographic clines in allele frequencies (Figure S5).

3.5 | Genetics of colour and introgression of colour loci

The SNP genetic data explained a substantial percentage of the variation in the eight colour bin traits (Figure S6). For example, point estimates of PVE (posterior median) on the raw colour bin scores ranged from 68% to 99%, and in most cases the credible intervals were narrow (lower bounds of the 95% CI > 75% for all but colour bin 1). Somewhat less variation was explained for the residual colour bin scores (point estimates ranged from 61% to 98% and lower bounds of the 95% CIs ranged from 30% to 95%). Much of the variation explained by genetics was attributable to loci with measurable effects (PGE was large, Figure S6). Moreover, estimates of the number of QTL (n -gamma) were relatively precise with point estimates suggesting an oligogenic architecture of most colour-bin traits, except bin 1 (e.g., 5–13 QTL for the residual scores except for bin 1, with a point estimate of 49 QTL; Figure 6). Furthermore, while some pairs of traits appear to be correlated, we found that mostly different SNPs

are associated with the different colour-bin traits (Figures S7 and S8), suggesting independent effects of traits.

We detected an excess of colour QTL among northern or southern excess directional introgression SNPs (i.e., those with 95% credible intervals for cline parameter α that excluded 0) for several colour traits (Figure 7). Similar results were obtained for raw and residual colour bin scores, so we present the numbers for the raw scores below. In particular, colour bin 1 QTL were found more among northern introgression SNPs than southern introgression SNPs (obs. difference = 0.94, percentile of null = 99.6%). In contrast, colour bin 5, 6 and 7 QTL were found more among southern introgression SNPs than northern introgression SNPs (bin 5, obs. difference = -0.46, percentile of null < 0.1%; bin 6, obs. difference = -0.06, percentile of null = 0.5%; bin 7, obs. difference = -1.02, percentile of null < 0.1%). QTL for colour bins 2, 3, 4 and 8 were more evenly distributed. The directions of these effects are consistent with the overall colour distribution in the contact zone relative to the northern and southern populations (compare Figure 7 and Figure 2a).

4 | DISCUSSION

We provide evidence of hybridization between intraspecific lineages of the polytypic red-eyed treefrog. In our examination of genomic and phenotypic variation across a newly discovered contact zone, we found that hybrids exhibit unique phenotypes and genotypes that arose through admixture of parental populations from the north and south. We detected elevated linkage disequilibrium in the contact zone, and evidence of directional genomic introgression from both parental populations, consistent with ongoing hybridization

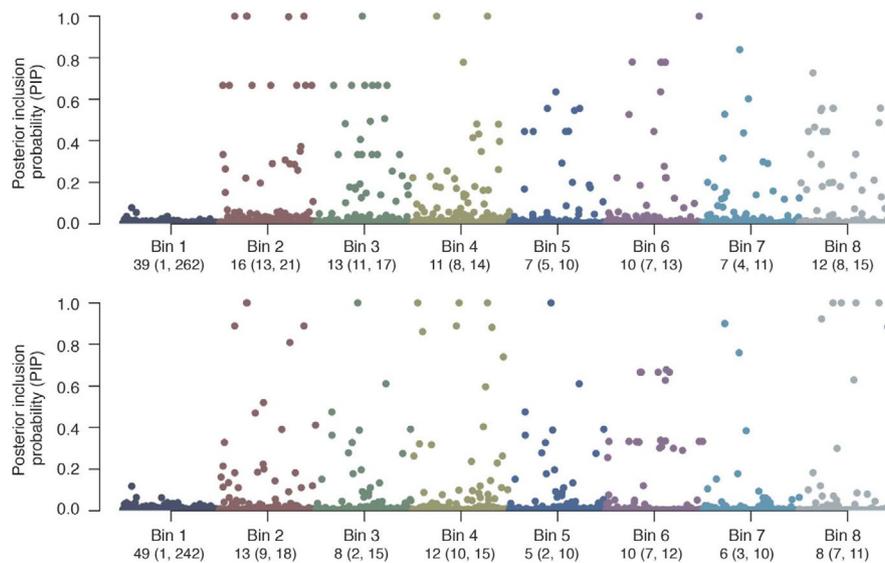


FIGURE 6 Genome-wide association mapping of leg colour. Posterior inclusion probabilities (PIPs) shown for all SNPs for each of eight colour bins from mapping raw phenotype values (a), and residual values (b) after regressing phenotype on the first principal component of the genotype matrix to account for the confounding effects of population structure. The number of QTL for colour bin traits is denoted below the x-axis. Medians and 95% credible intervals (in parentheses) of the posterior distribution for the number of QTL (n -gamma) are given for each colour bin. Points are coloured by the average colour of the pixels in each bin

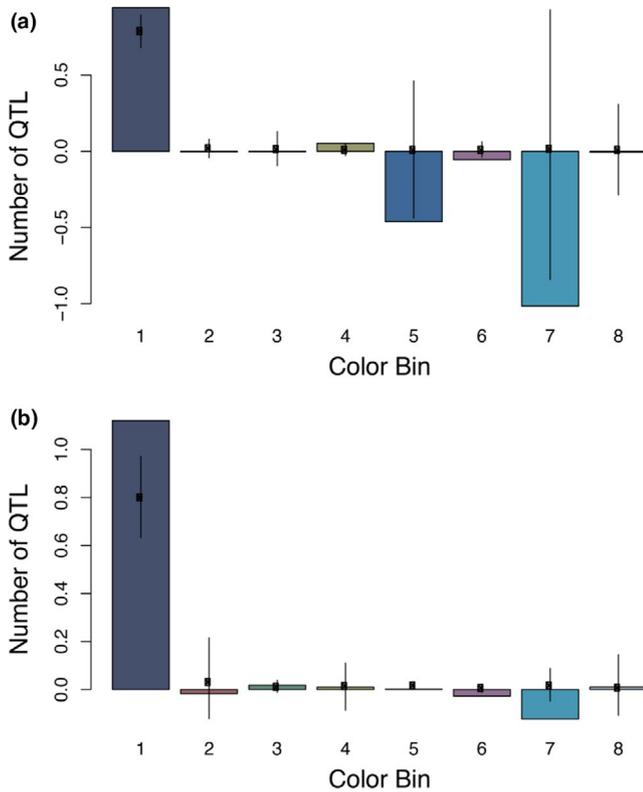


FIGURE 7 Difference in the number of QTL between SNPs with excess ancestry based on the analysis of the raw phenotype values (a) and of the residuals controlling for genetic PC1 (b). Positive values indicate an excess in the number of QTL for northern alleles, and negative values indicate an excess in the number of QTL for southern alleles. QTL associated with colour bin 1 were found more among northern introgression SNPs than southern introgression SNPs, whereas colour bin 5, 6 and 7 QTL were found more among southern introgression SNPs than northern introgression SNPs. QTL for colour bins 2, 3, 4 and 8 were more evenly distributed

between parental forms. Directional introgression of loci associated with different aspects of leg colour are inherited from each parental lineage, creating a distinct hybrid colour pattern. Our findings are compatible with a hybrid origin for the new morph and emphasize an important role for hybridization in generating the substantial phenotypic variation observed in red-eyed treefrogs.

We found that leg colour was oligogenic, with much of the variation in colour bin traits explained by loci with measurable effects. Our genetic data explained a substantial percentage of variation in all but one of colour bin traits analysed. This finding is consistent with other studies that have found colour pattern differences produced by few genes with little environmental effect (Kapan et al., 2006; Nijhout, 2003; Steiner et al., 2007). Nonetheless, we cannot rule out some confounding effect of population structure on our results, despite our efforts to control for this.

Moreover, the excess introgression for loci associated with variation in leg colour suggests selection, at least in part, mediates patterns of introgression in the contact zone. Because demographic processes are captured by genome-average introgression (hybrid

index), exceptional, directional introgression for colour loci is unlikely to be explained by drift or demography and is more likely attributable to selection. Interestingly, QTL associated with colour bin 1 (dark blue) were found more among northern introgression SNPs than southern introgression SNPs, while QTL associated with colour bins 5–7 (light blue and purple) showed the opposite pattern. This suggests that introgression of loci associated with different aspects of leg colour is creating a novel phenotype combination from parental forms, which is also reflected in our colour similarity analysis and matches our visual observations.

Colour patterns, which vary among populations, are demonstrated to serve as a social signal for red-eyed treefrogs (Kaiser et al., 2018; Robertson & Greene, 2017). Indeed, divergent populations of the red-eyed treefrog mate assortatively, with female preference for the local phenotype (Akopyan et al., 2018; Jacobs et al., 2016). The populations examined in this study contain multiple phenotypes, allowing us to examine the consequences of contact between these divergent forms. We find that hybrids in the contact zone are mostly restricted to abundant later-generation F_n individuals, with the majority exhibiting leg colours distinct from either parental form. Future studies should investigate whether hybrids preferentially mate with hybrid forms. If so, they could persist as a lineage and may eventually constitute a novel hybrid species via pre-mating isolation.

The consequences of hybridization on acoustic communication in red-eyed treefrogs remain to be tested. Previous studies show population variation in male advertisement calls (Akopyan et al., 2018), and that female preference for local mates is enhanced when visual and acoustic signaling are combined (Kaiser et al., 2018). Unlike colour pattern, however, acoustic communication is thought to play a limited role in generating phenotypic diversity in frogs (Chen & Wiens, 2020; Vargas-Salinas & Amézquita, 2013). Instead, acoustic communication can maintain distinct lineages by reinforcing pre-mating isolation when selection against hybrids is strong (Hoskin et al., 2005). Given the abundance of late-generation hybrids and our failure to detect restricted introgression of loci across the contact zone relative to genome-wide average admixture, we suspect that selection against hybrids plays a limited role in reproductively isolating these populations.

Sexual and natural selection can interact to accelerate divergence of hybrid forms. The synergistic interplay between natural and sexual selection in hybrid zones has driven rapid speciation of some of the most phenotypically diverse taxa (e.g., heliconian butterflies and dendrobatid frogs; reviewed in Rojas et al., 2018). In these examples, coloration serves a dual purpose in predator deterrence and mate attraction. While the role of aposematic coloration in nocturnal species is not well understood, some evidence supports the possibility of aposematism as a defence mechanism in red-eyed treefrogs (Sazima, 1974). Red-eyed treefrogs secrete host-defense polypeptides on their skin, and skin-peptide profiles vary among populations, in some cases covarying with leg-colour patterns (Davis et al., 2016). The production or sequestering of toxic skin secretions is often associated with bright warning coloration in diurnal frogs

(Daly et al., 1987). If leg colour, in addition to acting as a social signal, also serves as an antipredator signal, then patterns of colour variation are probably shaped by both natural and sexual selection. Additional study is required to test whether bright coloration serves as an antipredator signal to better understand the selective pressures shaping variation in red-eyed treefrog leg colour.

Environmental selection can also play a role in red-eyed treefrog diversification. Hybrids may exhibit increased fitness at the edges of the contact zone due to ecological opportunity created by available habitat that cannot be occupied by either of the parental populations (Devis et al., 1997; Martin et al., 2006; Svensson et al., 2006). Indeed, a biogeographic barrier between northern and contact-zone populations has been proposed for red-eyed treefrogs (Robertson & Vega, 2011), with the Reventazón and Pacuare Rivers in Limón Province potentially isolating sites 3 and 4. In a previous study, the magnitude of phenotypic divergence in coloration across this putative barrier exceeded the extent of genetic differentiation, indicating the possibility of strong localized selection (Robertson et al., 2009; Robertson & Vega, 2011). This break occurs along the Caribbean lowland forest that transitions into a series of floodplain valleys and coincides with the geographic range limits of some anuran and beetle taxa (Kohlmann et al., 2002; Robertson & Vega, 2011). An examination of the biotic and abiotic factors at this microgeographic scale may reveal the subtle and important differences that could act as selective agents (Brito et al., 2010; Milne et al., 2003; Picker, 1985).

Hybridization is recognized as one of the fundamental evolutionary processes fueling adaptive radiations and, in some cases, may contribute to the creation of biological diversity (Abbott et al., 2013; Meier et al., 2017; Seehausen, 2004). Here, we report evidence of hybridization resulting in the evolution of a unique hybrid form. Future work quantifying fitness and reproductive isolation of hybrid forms relative to parentals is an important next step for examining the potential evolutionary outcomes of this hybridization. Consistent with a growing body of literature on the role of hybridization in evolution, our findings challenge the traditional view of this process as disruptive or an evolutionary dead end (Darwin, 1859; Mayr, 1942), and instead highlight its role in facilitating the evolution of diversity.

ACKNOWLEDGEMENTS

Thank you to all of the backers of our Experiment.com crowdfunding campaign (doi.org/10.18258/4416) for their generosity and interest in this project. We also thank the Ministerio de Ambiente y Energía (MINAE-CONAGEBio) of Costa Rica for research and export permits (R-053-2015 and 2015-CR1678). This research was funded by California State University, Northridge (CSUN), and graduate research grants awarded to Maria Akopyan, including the Leslie and Terry Cutler Scholarship Endowment, Association of Retired Faculty Memorial Award, Dr Bob Luszczak, DDS Graduate Scholarship in Biology, and CSUN Thesis Support Grant. Finally, we thank and acknowledge the reviewers and editorial team for their very thoughtful feedback, which greatly improved the manuscript.

AUTHOR CONTRIBUTIONS

M.A., A.V., K.Kaiser., E.B.R., and J.M.R. conceived and designed the study. M.A., and A.V. conducted field sampling in 2015. M.A., and K.Klonoski. performed the molecular laboratory work. R.M., and E.B.R. provided laboratory and computational resources for generating genetic data. M.A., and Z.G. conducted data analysis. M.A. drafted the manuscript and all authors contributed to final refinements of the manuscript.

DATA AVAILABILITY STATEMENT

Raw data, input files, and scripts are available at <https://doi.org/10.6084/m9.figshare.11923017.v2>.

ORCID

Maria Akopyan  <https://orcid.org/0000-0001-7956-3196>

Zachariah Gompert  <https://orcid.org/0000-0003-2248-2488>

Erica Bree Rosenblum  <https://orcid.org/0000-0002-7384-8032>

Jeanne M. Robertson  <https://orcid.org/0000-0003-2642-6280>

REFERENCES

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Abbott, R. J., Barton, N. H., & Good, J. M. (2016). Genomics of hybridization and its evolutionary consequences. *Molecular Ecology*, 25(11), 2325–2332. <https://doi.org/10.1111/mec.13685>
- Akopyan, M., Kaiser, K., Vega, A., Savant, N. G., Owen, C. Y., Dudgeon, S. R., & Robertson, J. M. (2018). Melodic males and flashy females: Geographic variation in male and female reproductive behavior in red-eyed treefrogs (*Agalychnis callidryas*). *Ethology*, 124(1), 54–64. <https://doi.org/10.1111/eth.12705>
- Anderson, E., & Stebbins, G. L. Jr (1954). Hybridization as an evolutionary stimulus. *Evolution*, 8(4), 378–388. <https://doi.org/10.1111/j.1558-5646.1954.tb01504.x>
- Arnold, M. L. (1997). *Natural hybridization and evolution*. Oxford University Press.
- Arnold, M. L. (2006). *Evolution through genetic exchange*. Oxford University Press.
- Arnold, S. J., Verrell, P. A., & Tilley, S. G. (1996). The evolution of asymmetry in sexual isolation: A model and a test case. *Evolution*, 50(3), 1024–1033. <https://doi.org/10.1111/j.1558-5646.1996.tb02343.x>
- Baack, E. J., & Rieseberg, L. H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, 17(6), 513–518. <https://doi.org/10.1016/j.gde.2007.09.001>
- Barton, N. H. (1982). The structure of the hybrid zone in *Uroderma bilobatum* (Chiroptera: Phyllostomatidae). *Evolution*, 36(4), 863–866.
- Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, 16(1), 113–148. <https://doi.org/10.1146/annurev.es.16.110185.000553>
- Bay, R. A., & Ruegg, K. (2017). Genomic islands of divergence or opportunities for introgression? *Proceedings of the Royal Society B: Biological Sciences*, 284(1850), 20162414. <https://doi.org/10.1098/rspb.2016.2414>
- Bridle, J. R., & Butlin, R. K. (2002). Mating signal variation and bimodality in a mosaic hybrid zone between *Chorthippus* grasshopper species. *Evolution*, 56(6), 1184–1198. <https://doi.org/10.1111/j.0014-3820.2002.tb01431.x>

- Brito, J., Lizana, M., Martínez-Freiría, F., & do Amaral, J. P. (2010). Spatial and temporal segregation allows coexistence in a hybrid zone among two Mediterranean vipers (*Vipera aspis* and *V. latastei*). *Amphibia-Reptilia*, 31(2), 195–212. <https://doi.org/10.1163/156853810791069001>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Chen, Z., & Wiens, J. J. (2020). The origins of acoustic communication in vertebrates. *Nature Communications*, 11(1), 1–8. <https://doi.org/10.1038/s41467-020-14356-3>
- Daly, J. W., Myers, C. W., & Whittaker, N. (1987). Further classification of skin alkaloids from neotropical poison frogs (Dendrobatidae), with a general survey of toxic/noxious substances in the amphibia. *Toxicon*, 25(10), 1023–1095. [https://doi.org/10.1016/0041-0101\(87\)90265-0](https://doi.org/10.1016/0041-0101(87)90265-0)
- Darwin, C. (1859). *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. Oxford University Press.
- Davis, L. R., Klonoski, K., Rutschow, H. L., Van Wijk, K. J., Sun, Q. I., Haribal, M. M., Saporito, R. A., Vega, A., Rosenblum, E. B., Zamudio, K. R., & Robertson, J. M. (2016). Host defense skin peptides vary with color pattern in the highly polymorphic red-eyed treefrog. *Frontiers in Ecology and Evolution*, 4, 97. <https://doi.org/10.3389/fevo.2016.00097>
- Devis, N., Aiello, A., Mallet, J., Pomiankowski, A., & Silberglied, R. E. (1997). Speciation in two neotropical butterflies: Extending Haldane's rule. *Proceedings of the Royal Society B: Biological Sciences*, 264(1383), 845–851. <https://doi.org/10.1098/rspb.1997.0118>
- Duellman, W. E. (2001). *The Hylid Frogs of Middle America*, 2, (pp. 1–1180). Ithaca, NY: Society for the Study of Amphibians and Reptiles.
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ebersbach, J., Posso-Terranova, A., Bogdanowicz, S., Gómez-Díaz, M., García-González, M. X., Bolívar-García, W., & Andrés, J. (2020). Complex patterns of differentiation and geneflow underly the divergence of aposematic phenotypes in *Oophaga* poison frogs. *Molecular Ecology*, <https://doi.org/10.1111/mec.15360>
- 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>
- Fitzpatrick, B. M., Johnson, J. R., Kump, D. K., Shaffer, H. B., Smith, J. J., & Voss, S. R. (2009). Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. *BMC Evolutionary Biology*, 9(1), 176. <https://doi.org/10.1186/1471-2148-9-176>
- Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N., Gabriel, S. B., Topol, E. J., Smoller, J. W., Pato, C. N., Pato, M. T., Petryshen, T. L., Kolonel, L. N., Lander, E. S., Sklar, P., Henderson, B., Hirschhorn, J. N., & Altshuler, D. (2004). Assessing the impact of population stratification on genetic association studies. *Nature Genetics*, 36(4), 388–393. <https://doi.org/10.1038/ng1333>
- Gompert, Z. (2016). A continuous correlated beta process model for genetic ancestry in admixed populations. *PLoS One*, 11(3), e0151047. <https://doi.org/10.1371/journal.pone.0151047>
- Gompert, Z., Brady, M., Chalyavi, F., Saley, T. C., Philbin, C. S., Tucker, M. J., Forister, M. L., & Lucas, L. K. (2019). Genomic evidence of genetic variation with pleiotropic effects on caterpillar fitness and plant traits in a model legume. *Molecular Ecology*, 28(12), 2967–2985. <https://doi.org/10.1111/mec.15113>
- Gompert, Z., & Buerkle, C. A. (2011). Bayesian estimation of genomic clines. *Molecular Ecology*, 20(10), 2111–2127. <https://doi.org/10.1111/j.1365-294X.2011.05074.x>
- Gompert, Z., & Buerkle, C. A. (2012). bgc: Software for Bayesian estimation of genomic clines. *Molecular Ecology Resources*, 12(6), 1168–1176. <https://doi.org/10.1111/1755-0998.12009.x>
- Gompert, Z., Lucas, L. K., Buerkle, C. A., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2014). Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology*, 23(18), 4555–4573. <https://doi.org/10.1111/mec.12811>
- Gompert, Z., Mandeville, E. G., & Buerkle, C. A. (2017). Analysis of population genomic data from hybrid zones. *Annual Review of Ecology, Evolution, and Systematics*, 48, 207–229. <https://doi.org/10.1146/annurev-ecolsys-110316-022652>
- Gompert, Z., Parchman, T. L., & Buerkle, C. A. (2012). Genomics of isolation in hybrids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1587), 439–450. <https://doi.org/10.1098/rstb.2011.0196>
- Grant, V. (1981). *Plant speciation* (pp. 1–564). New York, NY: Columbia University Press. <https://doi.org/10.7312/gran92318>
- Guan, Y., & Stephens, M. (2011). Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *The Annals of Applied Statistics*, 5, 1780–1815. <https://doi.org/10.1214/11-AOAS455>
- Harrison, R. G. (1990). Hybrid zones: Windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, 7, 69–128.
- Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, 105(S1), 795–809. <https://doi.org/10.1093/jhered/esu033>
- Harrison, R. G., & Larson, E. L. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Molecular Ecology*, 25(11), 2454–2466. <https://doi.org/10.1111/mec.13582>
- Hedrick, P. W. (2013). Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, 22(18), 4606–4618. <https://doi.org/10.1111/mec.12415>
- Hewitt, G. M. (2001). Speciation, hybrid zones and phylogeography—or seeing genes in space and time. *Molecular Ecology*, 10(3), 537–549. <https://doi.org/10.1046/j.1365-294x.2001.01202.x>
- Hoskin, C. J., Higgie, M., McDonald, K. R., & Moritz, C. (2005). Reinforcement drives rapid allopatric speciation. *Nature*, 437(7063), 1353–1356. <https://doi.org/10.1038/nature04004>
- Jacobs, L. E., Vega, A., Dudgeon, S. R., Kaiser, K., & Robertson, J. M. (2016). Local not vocal: Assortative female choice in divergent populations of red-eyed treefrogs, *Agalychnis callidryas* (Hylidae: Phyllomedusinae). *Biological Journal of the Linnean Society*, 120, 171–178. <https://doi.org/10.1111/bij.12861>
- Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jiggins, C. D., Linares, M., Naisbit, R. E., Salazar, C., Yang, Z. H., & Mallet, J. (2001). Sex-linked hybrid sterility in a butterfly. *Evolution*, 55(8), 1631–1638. <https://doi.org/10.1111/j.0014-3820.2001.tb00682.x>
- 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>
- Kaiser, K., Boehlke, C., Navarro-Pérez, E., Vega, A., Dudgeon, S., & Robertson, J. M. (2018). Local preference encoded by complex signaling: Mechanisms of mate preference in the red-eyed treefrog (*Agalychnis callidryas*). *Behavioral Ecology and Sociobiology*, 72(12), 182. <https://doi.org/10.1007/s00265-018-2597-0>
- Kapan, D. D., Flanagan, N. S., Tobler, A., Papa, R., Reed, R. D., Gonzalez, J. A., & Heckel, D. G. (2006). Localization of Müllerian mimicry genes on a dense linkage map of *Heliconius erato*. *Genetics*, 173(2), 735–757. <https://doi.org/10.1534/genetics.106.057166>

- Kohlmann, B., Wilkinson, J., & Lulla, K. (2002). *Costa Rica from space/Costa Rica desde el espacio* (pp. 1-227). San José, Costa Rica: Neotrópica Foundation Publishers.
- Lamichhane, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martínez-Barrio, A., Promerová, M., Rubín, C.-J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T., & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, *518*(7539), 371–375. <https://doi.org/10.1038/nature14181>
- Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, *265*(5181), 2037–2048. <https://doi.org/10.1126/science.8091226>
- Lexer, C., Welch, M. E., Raymond, O., & Rieseberg, L. H. (2003). The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): Selection on transgressive characters in a novel hybrid habitat. *Evolution*, *57*(9), 1989–2000. <https://doi.org/10.1111/j.0014-3820.2003.tb00379.x>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMTOOLS. *Bioinformatics*, *25*(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Lipshutz, S. E., Meier, J. I., Derryberry, G. E., Miller, M. J., Seehausen, O., & Derryberry, E. P. (2019). Differential introgression of a female competitive trait in a hybrid zone between sex-role reversed species. *Evolution*, *73*(2), 188–201. <https://doi.org/10.1111/evo.13675>
- Lucas, L. K., Nice, C. C., & Gompert, Z. (2018). Genetic constraints on wing pattern variation in Lycaeides butterflies: A case study on mapping complex, multifaceted traits in structured populations. *Molecular Ecology Resources*, *18*(4), 892–907. <https://doi.org/10.1111/1755-0998.12777>
- Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, *20*(5), 229–237. <https://doi.org/10.1016/j.tree.2005.02.010>
- Mallet, J. (2007). Hybrid speciation. *Nature*, *446*(7133), 279–283. <https://doi.org/10.1038/nature05706>
- Martin, N. H., Bouck, A. C., & Arnold, M. L. (2006). Detecting adaptive trait introgression between *Iris fulva* and *I. brevicaulis* in highly selective field conditions. *Genetics*, *172*(4), 2481–2489. <https://doi.org/10.1534/genetics.105.053538>
- Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLOS Biology*, *17*(2), e2006288. <https://doi.org/10.1371/journal.pbio.2006288>
- Mavárez, J., & Linares, M. (2008). Homoploid hybrid speciation in animals. *Molecular Ecology*, *17*(19), 4181–4185. <https://doi.org/10.1111/j.1365-294X.2008.03898.x>
- Mayr, E. (1942). *Systematics and the origin of species*. New York, NY: Columbia University Press.
- 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. <https://doi.org/10.1002/ece3.794>
- Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications*, *8*(1), 1–11. <https://doi.org/10.1038/ncomms14363>
- Melo, M. C., Salazar, C., Jiggins, C. D., & Linares, M. (2009). Assortative mating preferences among hybrids offers a route to hybrid speciation. *Evolution*, *63*(6), 1660–1665. <https://doi.org/10.1111/j.1558-5646.2009.00633.x>
- Milne, R. I., Terzioglu, S., & Abbott, R. J. (2003). A hybrid zone dominated by fertile F1s: Maintenance of species barriers in *Rhododendron*. *Molecular Ecology*, *12*(10), 2719–2729. <https://doi.org/10.1046/j.1365-294X.2003.01942.x>
- Nijhout, H. F. (2003). Polymorphic mimicry in *Papilio dardanus*: Mosaic dominance, big effects, and origins. *Evolution & Development*, *5*(6), 579–592. <https://doi.org/10.1046/j.1525-142X.2003.03063.x>
- Nolte, A. W., Gompert, Z., & Buerkle, C. A. (2009). Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, *18*(12), 2615–2627. <https://doi.org/10.1111/j.1365-294X.2009.04208.x>
- Pardo-Díaz, C., Salazar, C., Baxter, S. W., Merot, C., Figueiredo-Ready, W., Joron, M., McMillan, W. O., & Jiggins, C. D. (2012). Adaptive introgression across species boundaries in *Heliconius* butterflies. *PLOS Genetics*, *8*(6), e1002752. <https://doi.org/10.1371/journal.pgen.1002752>
- Payseur, B. A. (2010). Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Molecular Ecology Resources*, *10*(5), 806–820. <https://doi.org/10.1111/j.1755-0998.2010.02883.x>
- Pereira, R. J., Barreto, F. S., & Burton, R. S. (2014). Ecological novelty by hybridization: Experimental evidence for increased thermal tolerance by transgressive segregation in *Tigriopus californicus*. *Evolution*, *68*(1), 204–215. <https://doi.org/10.1111/evo.12254>
- Picker, M. D. (1985). Hybridization and habitat selection in *Xenopus gilli* and *Xenopus laevis* in the south-western Cape Province. *Copeia*, 574–580. <https://doi.org/10.2307/1444746>
- Price, A. L., Zaitlen, N. A., Reich, D., & Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics*, *11*(7), 459–463. <https://doi.org/10.1038/nrg2813>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.
- Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, *16*(6), 359–371. <https://doi.org/10.1038/nrg3936>
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*(8), 1450–1477. <https://doi.org/10.1111/jeb.13047>
- Robertson, J. M., Duryea, M. C., & Zamudio, K. R. (2009). Discordant patterns of evolutionary differentiation in two Neotropical treefrogs. *Molecular Ecology*, *18*(7), 1375–1395. <https://doi.org/10.1111/j.1365-294X.2009.04126.x>
- Robertson, J. M., & Greene, H. W. (2017). Bright colour patterns as social signals in nocturnal frogs. *Biological Journal of the Linnean Society*, *121*(4), 849–857. <https://doi.org/10.1093/biolinnean/blx021>
- Robertson, J. M., & Robertson, A. D. (2008). Spatial and temporal patterns of phenotypic variation in a Neotropical frog. *Journal of Biogeography*, *35*(5), 830–843. <https://doi.org/10.1111/j.1365-2699.2007.01824.x>
- Robertson, J. M., & Vega, A. (2011). Genetic and phenotypic variation in a colourful treefrog across five geographic barriers. *Journal of Biogeography*, *38*(11), 2122–2135. <https://doi.org/10.1111/j.1365-2699.2011.02548.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., Burdfield-Steel, E., De Pasqual, C., Gordon, S., Hernández, L., Mappes, J., Nokelainen, O., Rönkä, K., & Lindstedt, C. (2018). Multimodal aposematic signals and their emerging role in mate attraction. *Frontiers in Ecology and Evolution*, *6*, 93. <https://doi.org/10.3389/fevo.2018.00093>
- Sazima, I. (1974). Experimental predation on the leaf-frog *Phyllomedusa rohdei* by the water snake *Liophis miliaris*. *Journal of Herpetology*, *8*(4), 376–377. <https://doi.org/10.2307/1562910>

- Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, *19*(4), 198–207. <https://doi.org/10.1016/j.tree.2004.01.003>
- Seehausen, O., & Schluter, D. (2004). Male–male competition and nuptial–colour displacement as a diversifying force in Lake Victoria cichlid fishes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1546), 1345–1353. <https://doi.org/10.1098/rspb.2004.2737>
- Stein, A. C., & Uy, J. A. C. (2006). Unidirectional introgression of a sexually selected trait across an avian hybrid zone: A role for female choice? *Evolution*, *60*(7), 1476–1485. <https://doi.org/10.1111/j.0014-3820.2006.tb01226.x>
- Steiner, C. C., Weber, J. N., & Hoekstra, H. E. (2007). Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLOS Biology*, *5*(9), <https://doi.org/10.1371/journal.pbio.0050219>
- Stelkens, R., & Seehausen, O. (2009). Genetic distance between species predicts novel trait expression in their hybrids. *Evolution*, *63*(4), 884–897. <https://doi.org/10.1111/j.1558-5646.2008.00599.x>
- Svensson, E. I., Eroukhmanoff, F., & Friberg, M. (2006). Effects of natural and sexual selection on adaptive population divergence and premating isolation in a damselfly. *Evolution*, *60*(6), 1242–1253. <https://doi.org/10.1111/j.0014-3820.2006.tb01202.x>
- Szymura, J. M., & Barton, N. H. (1986). Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina orientalis* and *B. orientalis*, near Cracow in southern Poland. *Evolution*, *40*(6), 1141–1159. <https://doi.org/10.1111/j.1558-5646.1986.tb05740.x>
- Vargas-Salinas, F., & Amézquita, A. (2013). Stream noise, hybridization, and uncoupled evolution of call traits in two lineages of poison frogs: *Oophaga histrionica* and *Oophaga lehmanni*. *PLoS One*, *8*(10), e77545. <https://doi.org/10.1371/journal.pone.0077545>
- Walsh, J., Kovach, A. I., Olsen, B. J., Shriver, W. G., & Lovette, I. J. (2018). Bidirectional adaptive introgression between two ecologically divergent sparrow species. *Evolution*, *72*(10), 2076–2089. <https://doi.org/10.1111/evo.13581>
- Weller, H. I., & Westneat, M. W. (2019). Quantitative color profiling of digital images with earth mover's distance using the R package colordistance. *PeerJ*, *7*, e6398. <https://doi.org/10.7717/peerj.6398>
- While, G. M., Michaelides, S., Heathcote, R. J. P., MacGregor, H. E. A., Zajac, N., Beninde, J., Carazo, P., Pérez i de Lanuza, G., Sacchi, R., Zuffi, M. A. L., Horváthová, T., Fresnillo, B., Schulte, U., Veith, M., Hochkirch, A., & Uller, T. (2015). Sexual selection drives asymmetric introgression in wall lizards. *Ecology Letters*, *18*(12), 1366–1375. <https://doi.org/10.1111/ele.12531>
- Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, *44*(7), 821–824. <https://doi.org/10.1038/ng.2310>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Akopyan M, Gompert Z, Klonoski K, et al. Genetic and phenotypic evidence of a contact zone between divergent colour morphs of the iconic red-eyed treefrog. *Mol Ecol* 2020;00:1–15. <https://doi.org/10.1111/mec.15639>