Received: 11 May 2023 | Revised: 15 February 2024 | Accepted: 23 February 2024

DOI: 10.1111/mec.17317

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Yosemite toad (*Anaxyrus canorus*) transcriptome reveals interplay between speciation genes and adaptive introgression

Paul A. Maier[1,2] | Amy G. Vandergast[3] | Andrew J. Bohonak[1]

[1]Department of Biology, San Diego State University, San Diego, California, USA

[2]Family TreeDNA, Gene by Gene, Houston, Texas, USA

[3]Western Ecological Research Center, San Diego Field Station, U.S. Geological Survey, San Diego, California, USA

Correspondence
Paul A. Maier, Family TreeDNA, Gene by Gene, 1445 N Loop W, Suite 820, Houston, TX 77008, USA.
Email: paulm@genebygene.com

Funding information
Achievement Rewards for College Scientists Foundation;
Harold & June Memorial Scholarship;
Jordan D. Covin Scholarship

Handling Editor: Sean D Schoville

Abstract

Genomes are heterogeneous during the early stages of speciation, with small 'islands' of DNA appearing to reflect strong adaptive differences, surrounded by vast seas of relative homogeneity. As species diverge, secondary contact zones between them can act as an interface and selectively filter through advantageous alleles of hybrid origin. Such introgression is another important adaptive process, one that allows beneficial mosaics of recombinant DNA ('rivers') to flow from one species into another. Although genomic islands of divergence appear to be associated with reproductive isolation, and genomic rivers form by adaptive introgression, it is unknown whether islands and rivers tend to be the same or different loci. We examined three replicate secondary contact zones for the Yosemite toad (*Anaxyrus canorus*) using two genomic data sets and a morphometric data set to answer the questions: (1) How predictably different are islands and rivers, both in terms of genomic location and gene function? (2) Are the adaptive genetic trait loci underlying tadpole growth and development reliably islands, rivers or neither? We found that island and river loci have significant overlap within a contact zone, suggesting that some loci are first islands, and later are predictably converted into rivers. However, gene ontology enrichment analysis showed strong overlap in gene function unique to all island loci, suggesting predictability in overall gene pathways for islands. Genome-wide association study outliers for tadpole development included LPIN3, a lipid metabolism gene potentially involved in climate change adaptation, that is island-like for all three contact zones, but also appears to be introgressing (as a river) across one zone. Taken together, our results suggest that adaptive divergence and introgression may be more complementary forces than currently appreciated.

KEYWORDS
adaptive introgression, admixture, genomic islands of divergence, hybridization, reproductive isolation, transcriptome

1 | INTRODUCTION

Darwin viewed speciation as a process involving differential adaptation to contrasting environments, by either natural or sexual selection (Darwin, 1859, 1871). However, the neo-Darwinian synthesis placed extreme emphasis on the totality of reproductive isolation between populations in allopatry, regardless of the role selection played (Dobzhansky, 1937; Mayr, 1942). One solution to these conflicting views came via the 'genic species' concept, which imagines incipient species as semipermeable barriers to gene flow, with some adaptively divergent genes playing larger roles in reproductive isolation (Wu, 2001; Wu & Ting, 2004). Under this view, blocks of coadapted gene complexes become larger as advantageous mutations are clustered together by hitchhiking selection, and these genomic 'islands' of divergence are closely involved in the formation of new species (Nosil et al., 2009; Nosil & Feder, 2012; Rundle & Nosil, 2005). This more nuanced perspective acknowledges the ongoing hybridization observed at multiple levels of organization (i.e., from lineages to distant species), and shifts the focus away from entire genomes towards genetic loci as the units of speciation (Abbott et al., 2013). Since the paradigm of genic speciation, numerous studies have reported genomic islands of highly differentiated loci surrounded by relatively homogenous genetic backgrounds (e.g., Ellegren et al., 2012; Malinsky et al., 2015; Turner et al., 2005).

If genes (and not populations of organisms) are the units of speciation, a question naturally arises about the numerous genes that continue to recombine in secondary contact zones: What is the consequence of this recombination on the fate of species? One increasingly common pattern found in young species is that hybrid zones generate recombinant diversity, which increases hybrid fitness (Arnold et al., 2012; Arnold & Hodges, 1995; Arnold & Martin, 2010). Although such heterosis wanes in subsequent hybrid generations, what remains are the extreme hybrid phenotypes that result from transgressive segregation (Rieseberg et al., 1999). Such extreme traits likely result from recombination between species when multiple complementary genes underlie those traits. Normally, a trait may be generated by multiple alleles that oppose each other, but recombination can shuffle those alleles into an order that amplifies or diminishes the trait value beyond what either species had previously displayed. Transgressive segregation is most common when the two species are more inbred, or more genetically divergent (Rieseberg et al., 1999). If it is adaptive, a transgressive hybrid trait can then flow freely into one or the other species by so-called adaptive introgression, to an extent that is limited by the balance between selection and migration (Arnold & Martin, 2009; Rieseberg, 2011; reviewed in Hedrick, 2013).

In contrast to genomic islands of divergence, these adaptive recombinant alleles can be thought of as genomic 'rivers', because they are new adaptations that flow directionally from hybrid zones into one species. Generally, genomic rivers will flow into whichever genomic and/or environmental background where it confers the highest fitness. Hybrid zones are thus excellent places to study the outcome of ongoing speciation: They are barriers to island loci with species-specific adaptations that may lower hybrid fitness, but they filter through river loci according to levels of recombinant adaptation (Barton & Bengtsson, 1986; Barton & Hewitt, 1985; Hewitt, 1988; Martinsen et al., 2001). Many examples of adaptive introgression are known in plants (Castric et al., 2008; Whitney et al., 2006), and animals (Fraïsse et al., 2014; Norris et al., 2015; Song et al., 2011), including in humans (Hawks & Cochran, 2006; Racimo et al., 2015). Although studies of genomic islands and rivers are widespread, it remains unclear whether islands must be different loci than rivers.

In the most extreme scenario (referred to here as 'S1'), islands cannot become rivers, because islands are composed of 'speciation genes' with species-specific adaptations (Figure 1). Hence, recombining them would lead to negative epistasis (Coyne, 1992; Orr et al., 2004). The idea is that certain loci, whether related to reproduction or not, are predisposed to accumulating Bateson–Dobzhansky–Muller (BDM) incompatibilities (Cutter, 2012; Orr & Turelli, 2001) or chromosomal rearrangements (Kirkpatrick & Barton, 2006) that predictably lower hybrid fitness. If divergence islands are predisposed to contain specific genes or mutations, then the same island loci should predictably be found across species divergences. This view of speciation genes has been popular until recently (Nosil & Schluter, 2011). For example, mutations and chromosomal inversions near loci controlling skeletal armour in three-spine sticklebacks have repeatedly evolved between pairs of freshwater and marine populations (Barrett et al., 2008; Jones et al., 2012). 'Magic traits' that are both the subject of divergent selection and cause of non-random mating may add to this predictability (Servedio et al., 2011). For example, the two stickleback phenotypes are associated with body size, which directly leads to assortative mating (McKinnon et al., 2004; Snyder & Dingle, 1989). Under this view, island loci should repeatedly become island loci across various contact zones, because they are genomic regions biased towards predictable reproductive incompatibilities. River loci, under this view, should form a distinct class of genes that is non-overlapping with islands.

There are two primary alternatives to Scenario S1. A second scenario (S2) is that islands reflect adaptations which are phenotypically predictable, but genomically unpredictable, and hence they are not biased towards reproductive incompatibilities. Such flexibility may allow islands to later transform into rivers if local adaptations are recombined by hybridization into novel and even more beneficial adaptations. This scenario is most likely if adaptive phenotypes are polygenic. For example, different genes that each play roles in regulating plumage pigmentation have independently differentiated at multiple crow contact zones (Poelstra et al., 2014; Vijay et al., 2016). Each molecular pathway may lead to the same adaptation, but perhaps be stochastically influenced by genetic draft, background selection, or drift (Cruickshank & Hahn, 2014; Noor & Bennett, 2009; Wolf & Ellegren, 2017). In this scenario, recombining a divergence island is unlikely to substantially lower a hybrid's fitness, because each island has a small effect on fitness. However, recombination is more likely to promote extreme hybrid traits for these island loci because they are found at extreme allele frequencies (Rieseberg et al., 1999). Therefore, adaptive introgression may co-opt some islands into

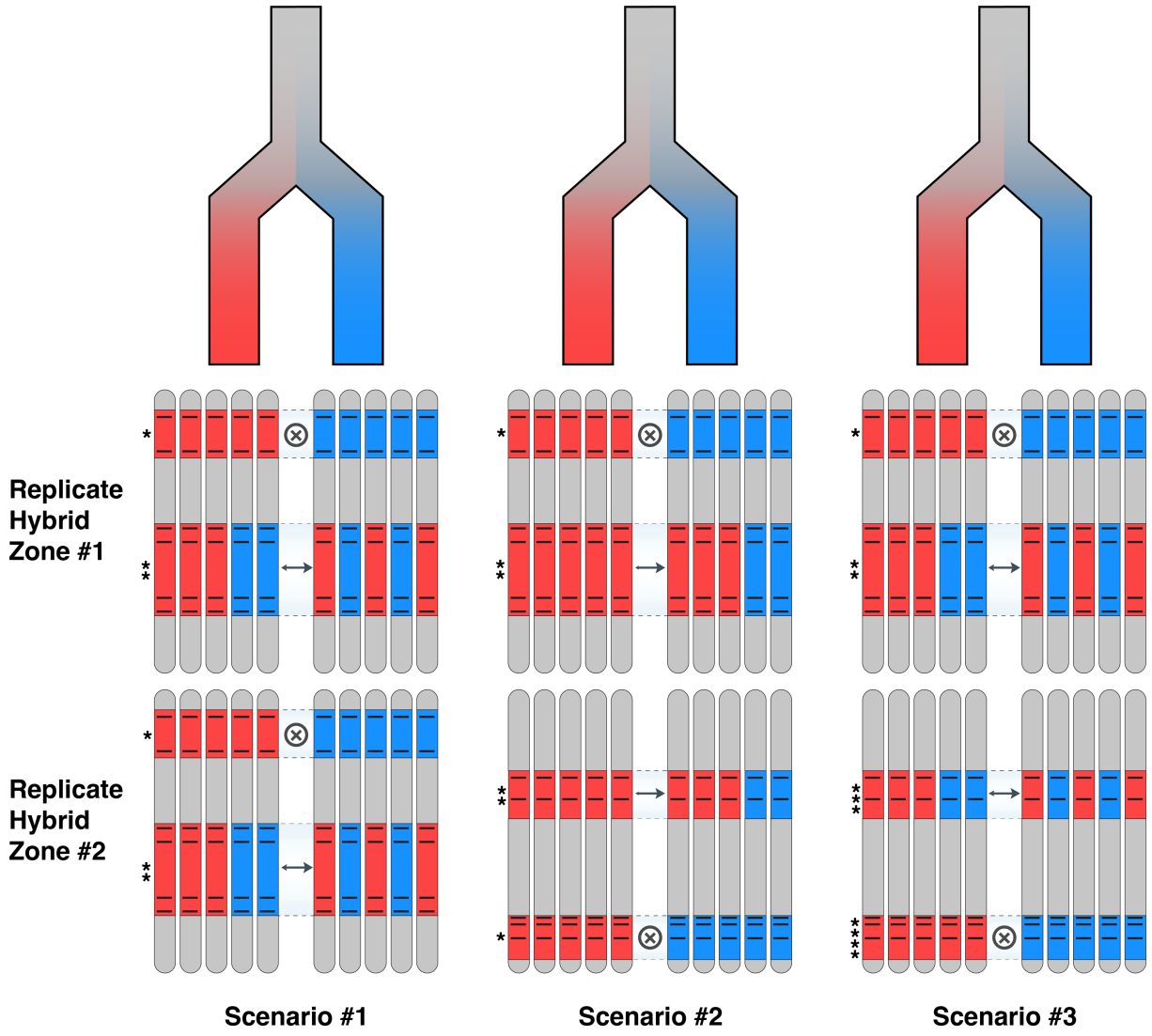

FIGURE 1 Model scenarios. Three scenarios describe alternative processes of island and river formation, as described in the text. In each scenario, two hypothetical hybrid zones are shown, at secondary contact after first diverging into red and blue lineages. Five chromosomes represent individuals in each lineage along a spatial cline. Genomic location is denoted vertically. Red and blue genomic locations are differentiated loci inherited from red/blue lineages, whereas grey locations are undifferentiated. Dashes represent SNPs. Loci with zero gene flow (islands) are denoted by ⊗. Loci experiencing directional introgression (rivers) are denoted by arrows. In Scenario #2, rivers are also formerly islands. Asterisks denote genes with similar functions, but potentially different identity and/or location.

rivers, if the recombinant fitness of these rivers surpasses the fitness of islands they are replacing (Jiggins et al., 2008; Martinsen et al., 2001; Rieseberg et al., 1999). Thus, weakly selected island loci may act as fuel for adaptive introgression, which may then erode and homogenize those islands (Bay & Ruegg, 2017; Clarkson et al., 2014). For example, an insecticide-resistance mutation that was originally a divergence island later introgressed between *Anopheles* mosquito species, with no perceptible effect on reproductive isolation or hybrid fitness (Clarkson et al., 2014). The expectation under this scenario would be that islands show genomic inconsistency yet functional consistency across contact zones, while some loci are both islands and rivers in each contact zone.

In a third possible scenario (S3), neither islands nor rivers predictably emerge in the same genetic loci, because adaptation is not the dominant force affecting divergence or introgression. Every locus could reflect unique histories of interacting evolutionary forces such as mutation, drift, recombination and selection, the so-called '$n = 1$ constraint' (Beaumont & Balding, 2004; Buerkle et al., 2011). This situation could also arise if adaptation—both divergent and introgressive adaptation—is primarily through small-effect polygenes with rampant epistasis (Rockman, 2012). This scenario would predict loci to sort into islands and rivers with no discernible pattern. Of course, S1–S3 are not mutually exclusive or exhaustive of all possible scenarios, but they are meant to explain the three most salient patterns found in nature.

We used a series of recently described contact zones between four lineages of the Yosemite toad (*Anaxyrus canorus*; Maier et al., 2019) to address which of these scenarios best describes

incipient speciation (Figure 2). In this system, three replicate zones of secondary contact exist for lineages that diverged 214–732 kya. Previous work on this system also found that inter-lineage admixture was extensive enough to promote lineage 'fusion' for two of these contact zones, raising the possibility that adaptive introgression is ongoing. Using evolutionary replicates can be a powerful way to address alternative hypotheses, particularly when the focal question involves how predictable a process is in nature (Hoekstra, 2006; Hohenlohe et al., 2010; Losos, 1992). We tested these hypotheses by first identifying markers as putative divergence islands, then identifying putative genomic rivers, and finally assessing whether there is any predictability to genic

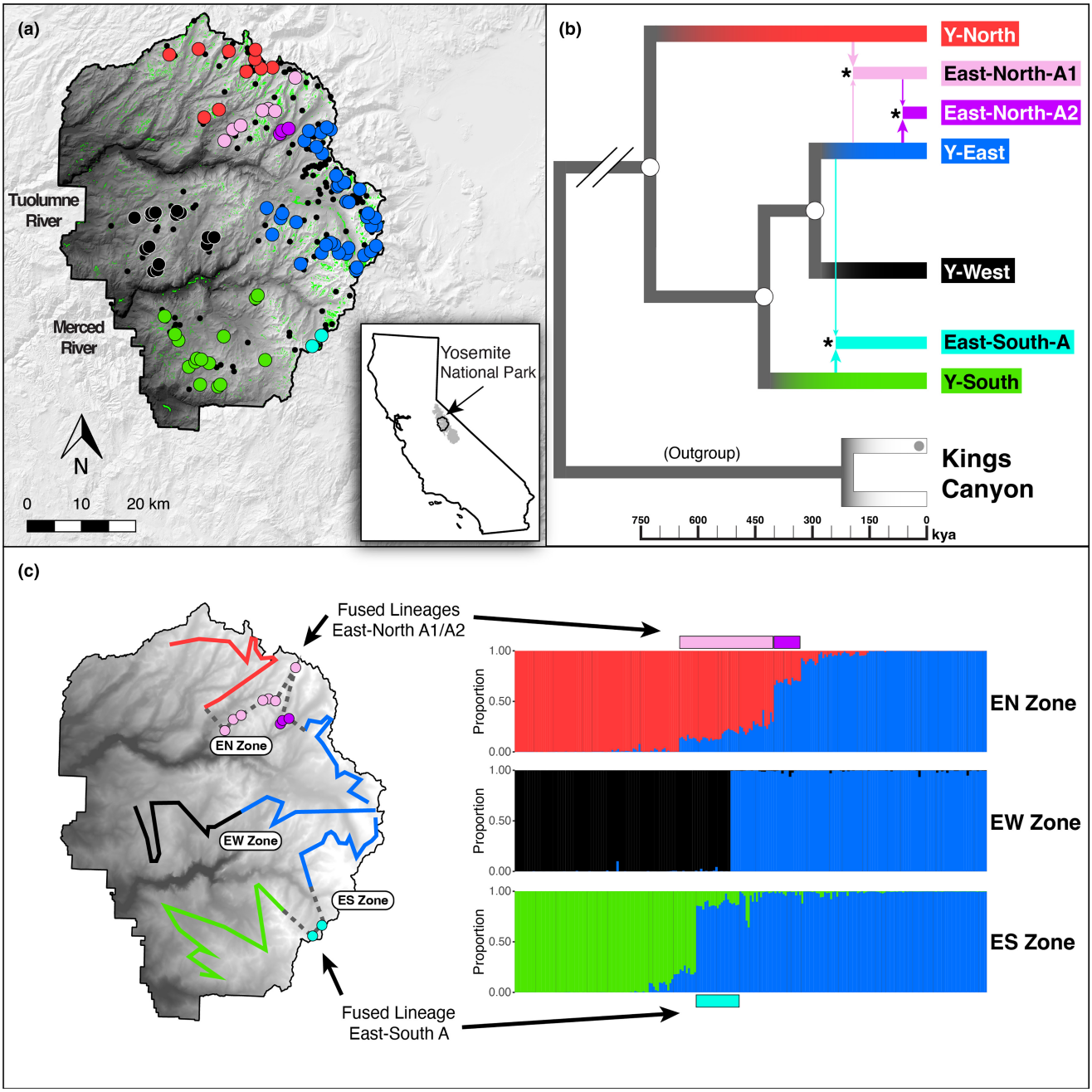

FIGURE 2 Study area and contact zones. (a) Primary study area in Yosemite National Park (YOSE), CA includes approximately 33% of sites known to be occupied by Yosemite toads. Bottom right inset shows the range of Yosemite toads in grey, and the boundaries of YOSE in black. Small green polygons are all meadows within the park (Keeler-Wolfe et al., 2012). Solid black circles indicate all known Yosemite toad meadows identified between 1915 and the present. Large circles indicate the meadows sampled and sequenced in the present study ($n = 90$). Colours correspond to phylogenetic lineages shown in panel (b). Random jitter is added to protect the locations of this threatened species. (b) Previously identified ancestral lineages and their estimated divergence dates (Maier et al., 2019), including four 'pure' lineages, and three 'fused' or 'admixed' lineages (asterisks). (c) Three contact zones used in the study: East-North (EN), East–West (EW), and East-South (ES), with two-way ancestry denoted by STRUCTURE barplots.

patterns across three replicate contact zones. Since genetic predictability can manifest in DNA loci, or in broader protein function (Reid et al., 2000; Stern & Orgogozo, 2008, 2009), we also assembled and annotated a de novo larval transcriptome to address whether genetic patterns are mirrored by gene functional patterns.

Additionally, we asked whether genome-wide patterns of island and river predictability applied to genetic trait loci of adaptive importance for Yosemite toad tadpoles. Specifically, we asked: Are loci underlying growth and development reliably islands, rivers or neither? The Yosemite toad is ostensibly under greatest future threat from a warming climate, because its habit of laying eggs in shallow ephemeral ponds predisposes the species to mass mortality when ponds desiccate (U.S. Fish & Wildlife Service, 2014). These toads appear to have lineage-specific genomic variation adapted to varying levels of winter snowpack and summer rainfall (Maier et al., 2023). Tadpole growth and development are generally seen as a life history trade-off, where faster development comes at the expense of smaller size at metamorphosis (for a comparative analysis, see Richter-Boix et al., 2011). Specialization on shorter hydroperiod ponds generally has a heritable basis, whether adaptations are fixed or plastic (Brady & Griffiths, 2000; Leips et al., 2000; Lind & Johansson, 2007; Morey & Reznick, 2004; Richter-Boix et al., 2006, 2011). Thus, it is reasonable to expect adaptive genetic differences between Yosemite toad lineages experiencing different levels of desiccation.

We tested Scenarios S1–S3 using a previously described double-digest Restriction Site-Associated DNA Sequencing (ddRADseq) data set, a newly developed reference transcriptome, and a tadpole morphometric data set collected from Yosemite National Park. Our analytical workflow used several divergence estimators to ensure that putative islands were not artifacts of any particular method (Cruickshank & Hahn, 2014), and used Bayesian genomic cline analysis (Gompert & Buerkle, 2011) to identify likely rivers. Understanding how adaptation and speciation proceed for this federally threatened species is essential, because long-term population declines and susceptibility to climate change make its future uncertain (Brown et al., 2015; Maier et al., 2022b, 2023; U.S. Fish & Wildlife Service, 2014). Our results will also bear significantly on the nature of speciation generally, by directly testing alternative scenarios of genic speciation in a comparative framework.

2 | MATERIALS AND METHODS

2.1 | Sample selection, molecular methods, ddRAD sequencing and bioinformatics

We used the previously described ddRADseq haplotype data set (Maier et al., 2019) for all analyses of genetic differentiation and introgression. Details about the library preparation, sequencing and bioinformatic parameters used to identify variable loci are described therein. Briefly, a total of 535 individual Yosemite toad (*Anaxyrus canorus*) tadpoles were sampled from 90 meadows across Yosemite National Park (YOSE) to maximize representation across all known breeding locations (Figure 2). Meadows are chosen as units because population boundaries are highly correlated with meadow boundaries (Maier et al., 2022a). Libraries were prepared using a ddRADseq protocol (Peterson et al., 2012; Protocol S1), then sequenced using 2 × 100 bp sequencing on seven lanes of an Illumina HiSeq 2500 (Illumina Inc., San Diego, California, USA). The data set was compiled and analysed with STACKS version 1.19 (Catchen et al., 2011, 2013) and assembled de novo. Wherever quality and coverage thresholds permitted, both reads were concatenated into a ~200 bp sequence, otherwise the first read (~100 bp) was used. Instead of calling one SNP per locus, we used the sequence as the allele for that locus. This gave the data set an enhanced amount of intra-locus information, and potentially made it more sensitive to intra-locus recombination that can occur within admixture zones. Additionally, we applied the following quality filters: a minimum 10× depth of coverage per locus, a minor allele frequency (MAF) of 0.05, heterozygosity less than 0.5 and a missing data frequency of 0.25. The data set contains 3261 polymorphic loci with a mean of 2.29 SNPs/locus and 2.78 haplotypes/locus. Although tadpoles were sampled from different pools wherever possible, COLONY version 2.0.6.4 (Jones & Wang, 2010) was used to remove 173 siblings from the data set.

2.2 | RNA sequencing, transcriptome assembly and annotation

A full de novo transcriptome was constructed to match RAD loci to genes and gene functional information (gene ontologies) where available. For full details of methods, see the Supporting Information S1. Briefly, three tadpoles were collected from throughout YOSE, and libraries were prepared and sequenced using 2 × 100 bp reads on an Illumina HiSeq 2500. A complete transcriptome was assembled using standard methods in Trinity r2014-02-04 (Grabherr et al., 2011; Haas et al., 2013). We annotated the transcriptome using the Trinotate suite included in Trinity, using standard methods (https://trinotate.github.io/). Biallelic SNPs were called across the transcriptome using the GATK version 4.0.1.2 best practices for RNAseq (McKenna et al., 2010). SNPs were then annotated with synonymous/non-synonymous effects and predicted protein changes in two steps: (1) Likely ORFs ≥100 amino acids were reconstructed using TRANSDECODER version 5.0.2 (Haas et al., 2013); (2) SNP effects were annotated using ENSEMBL VEP version 92.1 (McLaren et al., 2016). We identified ddRADseq markers that fall within coding genes by performing a blast-n nucleotide search on full RAD sequences, with the transcriptome as the database. To minimize the possibility of matching paralogs, we only allowed ungapped alignments with no more than three mismatches beyond the known RAD SNPs, with an e-value cut-off of 1×10^{-6}.

2.3 | Defining lineages and admixture zones

Four 'pure' lineages (Y-North, Y-East, Y-South, Y-West) exist in YOSE, and due to the topographic barriers of Merced and Tuolumne Rivers, there are three contact zones (Figure 2; Maier et al., 2019). Those include East-North (hereafter 'EN'), East-West ('EW') and East-South ('ES'). Admixed or 'fused' lineages are found in the EN and ES contact zones. Genomic cline methods can identify introgressed loci based on their departure from genome-wide patterns of admixture, but they require a definition of pure (i.e., parental) allele frequencies. To screen individuals as putatively admixed, we used a four-step process: (1) We ran STRUCTURE version 2.3.4 (Pritchard et al., 2000) five times for 5.0×10^5 steps and 1.0×10^5 burn-in with $K=2$ ancestral populations to find individuals with dual ancestry; (2) we used NEWHYBRIDS version 1.1 (Anderson & Thompson, 2002) to screen for signal of recent admixture; (3) we used the HIEST package (Fitzpatrick, 2012) in R (R Core Team, 2023) to screen for more advanced (>2 generations) admixture; and (4) we performed a principal components analysis (Jombart, 2008) on genotypes to assess whether putatively admixed individuals identified in Steps 1–3 clustered separately and intermediate from pure individuals. Any meadows lacking individuals with evidence of admixture were thereafter considered 'pure'. Sample sizes (individuals; meadows) for contact zones were as follows: EN ($n=184$; $n=39$), ES ($n=162$; $n=42$) and EW ($n=132$; $n=33$). Sample sizes of pure meadows were as follows: EN ($n_1=30$, $n_2=49$; $n_1=6$, $n_2=10$), ES ($n_1=27$, $n_2=61$; $n_1=10$, $n_2=18$) and EW ($n_1=33$, $n_2=41$; $n_1=10$, $n_2=10$).

2.4 | Finding genomic islands of divergence

We defined genomic islands of divergence as loci with unusually high divergence among lineages. Three separate metrics were used to avoid bias that may be associated with any single approach (see Supporting Information S1). First, we performed a hierarchical molecular analysis of variance (AMOVA; Excoffier et al., 1992) on each marker at the level of lineages, while accounting for variance at the level of meadows and individuals. We used the AMOVA function in the R package PEGAS (Paradis, 2010), and extracted Φ_{ST} representing between-lineage variance using the GETPHI function. Second, we calculated D_{XY} (Nei, 1987), an absolute measure of sequence divergence that quantifies the average number of nucleotide differences among lineages. D_{XY} only reflects differences that accumulated since the lineages split (as well as ancestral polymorphism), and thus is unaffected by levels of intra-lineage diversity. We calculated D_{XY} using custom R scripts. Φ_{ST} and D_{XY} provide independent measures of divergence but are correlated with each other. Third, we calculated Slatkin and Maddison's 's' statistic (Slatkin & Maddison, 1989), the number of parsimony steps required for each gene tree to be congruent with the 'true' lineage tree (Figure 2) using DENDROPY version 4.4.0 (Sukumaran & Holder, 2010). This cladistic measure is sensitive to lowered gene flow or increased selection, which should both cause gene trees to be more congruent with the lineage tree. Input gene trees were generated using the BIONJ neighbour joining algorithm in the APE package (Gascuel, 1997; Paradis et al., 2004), with the Kimura 2-parameter model of evolution (Kimura, 1980). For all three metrics, markers were considered statistical outliers if they: (a) exceeded the 95th quantile of marker-wide values, and (b) rejected the null hypothesis of no differentiation, based on 2.5×10^4 permutations at the Bonferroni-corrected $\alpha=.05$ level (adjusting for the number of loci examined in each contact zone). We refer to markers that passed two out of three tests as outliers, and those passing all three tests as stringent outliers (both hereafter 'divergence islands'). Computations were performed using custom R scripts.

2.5 | Finding genomic rivers using genomic cline analysis

We defined genomic rivers as loci with aberrant genomic clines suggestive of adaptive introgression. We estimated locus-specific genomic clines for each contact zone using BGC version 1.03 (Gompert & Buerkle, 2011, 2012). A genomic cline is a function of how, for each locus, the proportion of alleles shifts along an average genomic hybridization gradient (between pure lineages 'A' and 'B'). Locus-specific patterns of ancestry (ϕ_{ih}) can differ from the genome-wide average (h, i.e., hybrid index), if a locus introgresses directionally into one lineage more or less than average. Two parameters are used to describe this 'excess ancestry': cline center (α_i) and cline steepness (β_i):

$$\phi_{ih} = h + 2\left(h + h^2\right)\left(\alpha_i + \beta_i(2h - 1)\right)$$

The cline center parameter α_i indicates increases (positive values) or decreases (negative values) in the 'extent' of ancestry from one specific lineage, for one marker relative to genome-wide expectations. The cline steepness parameter β_i indicates increases (positive values) or decreases (negative values) in the 'rate' of introgression from one specific lineage, representing levels of pairwise linkage disequilibrium for that marker compared with all others. Simulations have shown that selection against hybrids that leads to reproductive isolation, such as underdominance or pairwise epistasis, can impact either parameter (α_i or β_i) (Gompert et al., 2012; Gompert & Buerkle, 2011, 2012). However, adaptive introgression that favours homozygous genotypes by directional selection should only impact α_i (Gompert & Buerkle, 2011). Therefore, we chose to interpret extreme α_i values as evidence for adaptive introgression.

The BGC method uses a Bayesian framework to estimate the probability of an individual with hybrid index h inheriting a gene copy at locus i from the Y-East lineage (defined here as ϕ_{ih}). The Y-East lineage is chosen for convenience, since all contact zones contain it; the probability for a contrasting lineage is defined by $1-\phi_{ih}$. We ran models 3× each for 5.0×10^5 steps and 1.0×10^5 burn-in to check for convergence, sampling every 20th step with default settings, except that MCMC tuning parameters were increased 2× to increase mixing of the chain. If a marker had 95% equal-tailed posterior probability

credible intervals of α_i that did not include the genome-wide average value (zero), it was designated a 'genomic river'.

2.6 | Estimating genome-wide and outlier-specific migration

Given the recent discussion over whether divergence islands truly represent regions that are barriers to migration, we estimated the rates and symmetries of migration. Once island and river loci were identified, we used MIGRATE-N version 3.6.11 (Beerli & Felsenstein, 2001) to compare rates and direction of gene flow for each marker class (all markers, islands, rivers with positive *a*, rivers with negative *a*), and for each contact zone. To reduce the number of parameters, each contact zone was classified into three populations: P1, P2, and admixed, and modelled under a stepping-stone scenario (migration of P1 and P2 through the admixed population). Each model was run 10× each for 2×10^6 steps with 2×10^4 burn-in, using four MCMC chains with static heating.

2.7 | Comparative analyses

We tested the null hypothesis of no association for genetic loci among the different categories of outlier locus (six total—islands vs. rivers, across three contact zones) in several ways. First, we used pairwise Fisher's exact tests to test for locus overlap between each pair of groups, given the number of loci examined and a *p*-value correction for multiple tests. We implemented this using the GENEOVERLAP package in R (https://github.com/shenlab-sinai).

Second, we tabulated gene ontology terms for each outlier locus where that information was available, and performed a Cochran–Mantel–Haenszel (CMH) chi-squared test for count data (Agresti, 2002) to determine whether gene functional categories are conditionally independent in each stratum. Gene ontologies (GO) categorize gene products based on biological processes, cellular components and molecular functions. We used Level 2 of the GO term hierarchy, which gives highly specific categories. First, we tested for independence of GO terms by islands and rivers while accounting for contact zone, and then we tested for independence of GO terms by contact zone, while accounting for marker type. To understand which specific GO terms are sorting by contact zone and marker type, we also performed a hierarchical clustering analysis on the level of GO enrichment or depletion, using the HCLUST package in R. Enrichment was calculated using two-tailed Fisher's exact tests on counts of each GO term compared with expected values by group. We multiplied the sign of the odds ratio by the $-\log_{10}(p-\text{value})$ to get level of enrichment (positive) or depletion (negative; Janoušek et al., 2015; Sánchez et al., 2007). We removed GO terms that were only enriched by a single group (singletons, e.g., only EN Islands) before clustering them.

Finally, we performed another pair of CMH tests (as described above) to test whether severity of SNP effects—for example, missense mutations—differs by either contact zone or marker type. Given the low proportion of RAD markers that successfully matched to gene annotations, we slightly lowered the threshold for defining islands and rivers in tests that relied upon gene or SNP annotations. To accomplish this, we used one-tailed and two-tailed 95% confidence intervals for islands and rivers, respectively.

2.8 | Power analyses

Although we identified island and river loci using 95% quantiles of genome-wide values, this threshold could in principle be increased to more confidently isolate adaptive loci from neutrally evolving ones. However, our comparative analyses should be robust to false-positive results of adaptation (yet vulnerable to false negatives) because patterns of neutral locus overlap across replicate contact zones should be insignificant or random. Nevertheless, we performed power analyses to evaluate whether 95% quantiles are a reasonable threshold for our data set, or whether more stringent cut-offs have sufficient power to test our hypotheses.

For both Fisher's exact tests and CMH tests, we simulated 1000 data sets using loci obtained from quantile thresholds ranging from 0.90 to 0.99, in 0.01 increments. Given that the 95% quantile data set produced significant results (see Section 3), we calculated power $(1-\beta)$ as the proportion of simulated data sets that correctly reject the null hypothesis at the corrected α level of 0.05. For insignificant results, we also estimated the factor increase in genetic markers necessary to reject the null.

2.9 | Loci related to tadpole growth and development: islands or rivers?

We opportunistically recorded the total length and larval stage—field categories of Gosner (1960) stages—for 1725 sampled tadpoles from 97 meadows throughout YOSE during Years 2012–2013. We used these data to conduct a genome-wide association study (GWAS), to identify SNPs likely involved in larval growth and development. We first modelled overall growth-by-development using polynomial regression, where growth was an *n*th degree polynomial of developmental stage. We tested between one and six polynomial terms, and compared model fit using Bayesian Information Criterion (BIC) with the COMPARELM function in the RCOMPANION package, as well as with an ANOVA. Tadpoles were only measured once, so we treated deviation from the overall growth-developmental curve as evidence for a tadpole's phenotypic deviation; residuals for 535 genotyped individuals were modelled against 1302 SNP loci, and a fixed effects model using population structure as a covariate was fit to each SNP with the LMEM.GWASER package (Quero et al., 2018). Population structure was determined using principal component analysis; all significant eigenvectors at the 0.05 significance level were included as covariates to reduce false discovery. Candidate loci were designated

with the SNP effect p-value cut-off of $-\log_{10}\left(5\times10^{-6}\right)$, which we chose to be more conservative than the typical Bonferroni threshold of $-\log_{10}\left(3.8\times10^{-5}\right)$, that is, 0.05/1302. Spatial genetic patterns for candidate loci were visualized using the ADEGENET package (Jombart, 2008).

3 | RESULTS

3.1 | Transcriptome assembly quality

The transcriptome assembly contained 285,751 components (putative 'genes'), with 50% of assembled nucleotides residing in transcripts 1571 bp or longer (defined as the N50). Out of 1485 unique RAD sequences that passed quality thresholds in the three contact zones, 653 successfully matched transcripts. For full details of transcriptome assembly quality, see the Supporting Information S1.

3.2 | Extent of admixture zones

STRUCTURE and NEWHYBRIDS analyses delineated the extent for each of the three secondary contact zones, and along with HIEST/PCA analyses also revealed differences in overall admixture levels (Figures S1–S3). The EN contact zone—also the oldest contact zone—was the most geographically expansive of all three contact zones and showed an even distribution of hybrid index (h) values (Figure S1). Although PCA showed hybrid genotypes intermediate between pure genotypes in all contact zones, EN hybrids were particularly intermixed with pure genotypes. HIEST and PCA showed a mix of putatively recent and advanced hybrids clustering together, supporting an overall picture of recent and widespread admixture in the EN contact zone.

Admixture was not uniform across contact zones. More individuals ($n=105$) and meadows ($n=23$) were defined as admixed for the EN contact zone than for the ES ($n=74$ individuals, $n=14$ meadows) and EW ($n=58$ individuals, $n=13$ meadows) contact zones. These two contact zones showed the opposite pattern, namely narrower zones, with more distinct clustering of hybrid genotypes (Figures S2 and S3). One possible reason for this discrepancy is the difference in contact zone type: These two younger contact zones are low–high-elevation contrasts, whereas the EN contact zone is a high–high contrast (see Maier et al., 2019). However, with a paucity of contact zones to study, it cannot be known for certain whether climatic divergence is the primary force in shaping the dynamics of the EN, ES and EW contact zones.

3.3 | Genomic islands: divergence outliers

We initially removed loci from each contact zone that were invariant in the pure meadows, fell below a MAF of 0.05, or had zone-specific missing data levels >0.5. After applying these thresholds, we retained 1126 EN loci, 1138 ES loci and 925 EW loci (the total number of unique loci across all zones was 1485). The same locus sets were used in all subsequent analyses. Divergence islands (loci passing two of three tests) were attained in numbers roughly proportional to these overall locus counts: EN ($n=37$), ES ($n=43$) and EW ($n=15$). For loci passing all three tests (stringent outliers), these counts were: EN ($n=14$), ES ($n=15$) and EW ($n=1$). Values of hierarchical Φ_{ST} and D_{XY} were positively correlated, especially closer to the upper tails of the distributions (Figure 3). The cladistic measure of divergence 's' had a more nuanced relationship with hierarchical Φ_{ST} and D_{XY} (Figures S4–S6). However, neighbour joining trees based on concatenated divergence islands were strikingly effective at generating the 'correct' lineage phylogeny; in comparison, an equivalent number of loci chosen randomly from the interquartile range of D_{XY} values did not recover the correct topology (Figure 3).

3.4 | Genomic rivers: introgression outliers

Log likelihood values and parameters converged across replicate BGC runs and replicate α and β estimates were highly correlated (ρ of .76–.93), so all chains were combined into a single estimate (Figures S7 and S8). Estimates of α showed an abundance of outliers in the EN contact zone ($n=193$; Figure 4), supporting the pattern of higher admixture found earlier for that contact zone (Figure S1). The other two contact zones had comparatively few α outliers ($n=47$ for ES; $n=15$ for EW). The majority of these α outliers were positive, meaning clines are shifted toward Y-East, and more introgression from Y-East compared with the reverse direction. Overall, the range of α values estimated was larger and more even-tailed for EN (min = −0.62, max = 0.71) compared with the ES (min = −0.37, max = 0.65) and EW (min = −0.38, max = 0.54) ranges.

The parameter β ranged from −0.5 to 0.52, but the 95% CI overlapped zero for all loci. As noted earlier, β is more associated with forms of selection such as underdominance or pairwise epistasis that may confer reproductive isolation, and hence, we solely used α to identify introgression outliers. However, it is interesting to note that α and β had a strong positive correlation in the two younger contact zones (ρ of 0.25–0.30), and a small but significantly negative correlation for the EN contact zone (ρ of −0.11), indicating the possibility that divergence and introgression have a different dynamic in that contact zone (Figures S9–S11). D_{XY} and β were all positively correlated, as expected (ρ of 0.15–0.26). The parameter estimates of α and β are likely accurate because they were repeatable across replicate BGC runs, and because all genomic cline estimates were based on sufficient coverage of hybrid index values (Figure 4).

3.5 | Direction of migration for islands versus rivers

MIGRATE-N results generally supported the overall expectations (Figure 5). Divergence islands as a group were migrational sinks

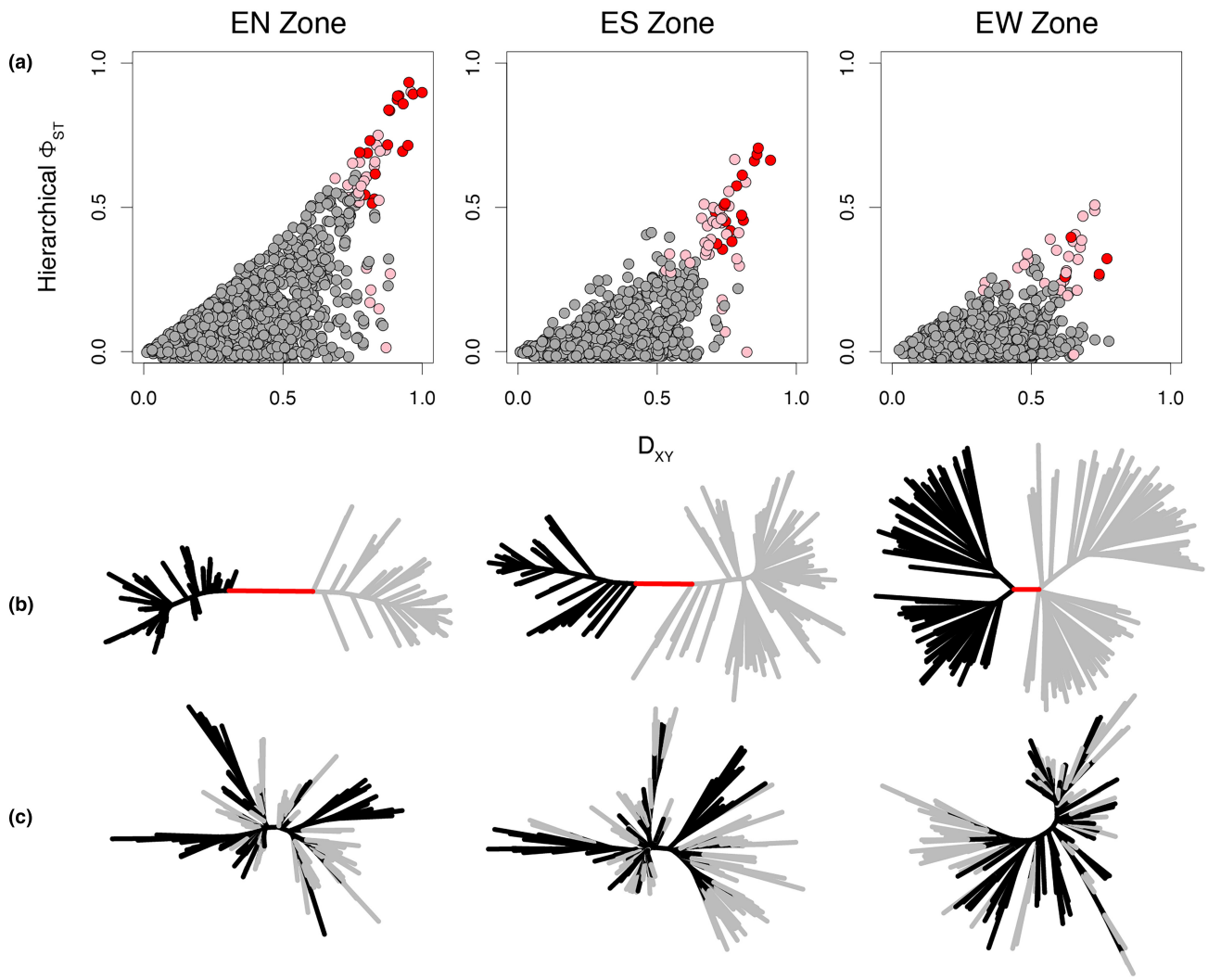

FIGURE 3 Genomic islands of divergence. Loci with extreme differentiation between pure individuals, at three contact zones. (a) Scatterplots of hierarchical Φ_{ST} by D_{XY}, showing loci in the 95th quantile that are significantly different than zero; red = outlier loci for Φ_{ST}, D_{XY}, and Maddison and Slatkin's *'s'*, pink = outliers in two out of three tests. (b, c) Neighbour joining trees from RAD sequences at the level of haplotypes (i.e., 2× number of individuals). Loci used are either islands only (b) or an equivalent number of loci chosen randomly from the interquartile range of all D_{XY} values (c), for comparison. Black and grey branches represent ancestry from the pure lineages in that contact zone. The red branch in (b) highlights the fact that island loci produce reciprocal monophyly of lineages.

(net gene flow into hybrid zones), as would be expected if those markers are somewhat deleterious for hybrids possessing them. In contrast, genomic rivers (divided into negative and positive α values) had asymmetric migration rates into one lineage. For example, positive α rivers (which contain excess Y-East lineage ancestry) showed highest migration between the Y-East lineage and the admixture zone, suggesting introgression from the Y-East lineage. In the negative rivers, where introgression is expected from the other (non-Y-East) lineage, patterns were slightly more nuanced. One river had the same pattern (i.e., EN had elevated Y-North-admixed migration), and the other two had net migration towards Y-East. Collectively, rivers in each admixture zone appeared to act as migration sources, with net migration radiating away from the admixture zone.

3.6 | Results of comparative analyses

Within each contact zone, we found significant overlap between the identity of loci for islands and rivers (Figure 6). Pairwise Fisher's exact tests found odds ratios of 6.88 for the EN contact zone ($p = 7.4 \times 10^{-7}$), 20.46 for the ES contact zone ($p = 7.9 \times 10^{-11}$), and 17.10 for the EW contact zone ($p = 4.6 \times 10^{-2}$), after correcting p-values with the Benjamin–Hochberg method. No such pattern of significant locus overlap was found for islands as a group, or rivers as a group. No other group combinations had significant overlap, given their respective locus counts, and no locus was unique to islands or rivers only. This suggests that within a particular contact zone, highly introgressed loci may derive from highly diverged loci.

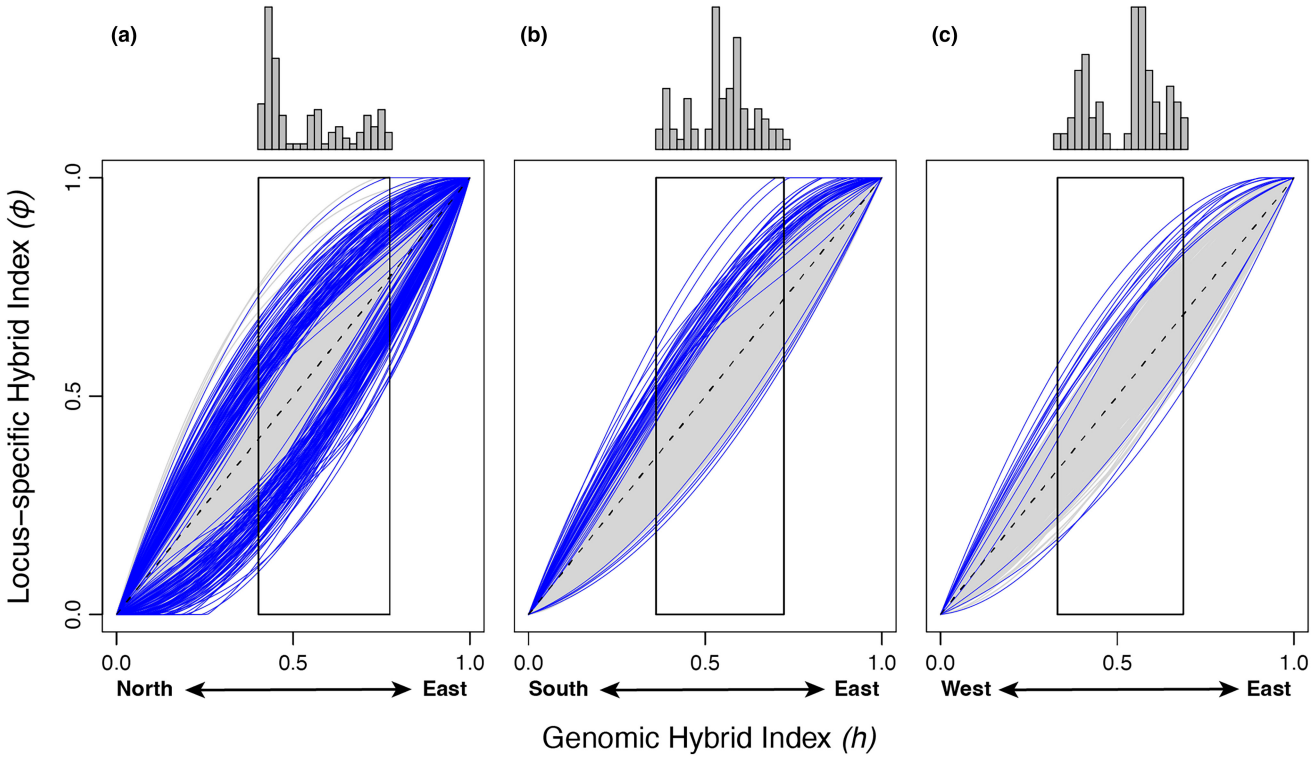

FIGURE 4 Genomic rivers of introgression. Bayesian genomic clines for each contact zone: (a) East-North (EN), (b) East-South (ES), and (c) East-West (EW). All loci passing quality and frequency thresholds are shown, with genomic rivers highlighted in blue. Rivers are loci for which 95% credible intervals of its α estimate exclude zero. Hence, rivers can be extreme positive (above stippled line) or negative (below stippled line) values, for loci containing excess Y-East (positive) or Y-North/Y-South/Y-West (negative) ancestry. A histogram of observed genomic hybrid indices (h) is displayed above each plot, with a box highlighting the region of the plot for which h was observed.

In contrast, when considering only gene functional categories, we found that islands (across contact zones) form a functional group that is significantly different than rivers. The CMH chi-squared test for independence of marker class (i.e., islands vs. rivers) across GO categories was significant ($M^2 = 360.31$, df = 260, $p = 3.71 \times 10^{-5}$), when accounting for differences across contact zones. We used a Woolf test to check that all odds ratios were homogenous across strata, to make sure the CMH test was appropriate; we found no three-way association of odds ratios across strata ($\chi^2 = .007$, df = 2, $p = .9965$), indicating the CMH assumption was not violated. However, post hoc tests showed that no island-river pair was significant within each contact zone (EN: $p = .93$, ES: $p = 1.0$, EW: .06). This is likely due to low counts in pairwise comparisons (mean count: 0.56 per cell). The opposite pattern was not found; namely, a CMH test for independence of contact zones while accounting for marker class was insignificant ($M^2 = 389.93$, df = 520, $p = 1.0$).

This pattern of functional difference between islands and rivers was strongly supported by hierarchical clustering of GO terms (Figure 7). Most GO terms that were enriched in more than one group were enriched for all three sets of island loci. Although many GO terms were enriched for all islands and depleted in rivers, low sample sizes prevented the odds ratios for these depleted river terms from being significantly different than zero. There was less evidence of functional (GO term) overlap between islands and rivers in each contact zone; however, some GO terms did conform to this pattern (Figure 7). In total, 61 specific GO terms differentiated islands from rivers, whereas 36 specific GO terms were unique to islands and rivers of a specific contact zone. Of this latter category, 10, 23 and 3 terms were unique to the EN, ES and EW contact zones, respectively (Table S1). These patterns were based entirely on highly specific (Level 2) GO term categories. Sorted into broader GO categories ('biological process', 'cellular component', 'molecular function'), the functional profiles of 'island-only', 'EN-only', 'ES-only', and 'EW-only' were distinct (CMH test; $M^2 = 25.063$, df = 6, $p = .0003$; Table 1). However, when these latter three groups were lumped into a 'zone-only' grouping, there was no significant difference in functional profile between 'island-only' and 'zone-only' (CMH test; $M^2 = 4.069$, df = 2, $p = .1308$). Overall, this suggests that island loci are functionally the most distinct grouping, yet each contact zone has island/river loci with unique functional patterns.

We also tested whether 'type' of SNP (i.e., the effect of that SNP on gene function) is characteristic of marker class, contact zone, or neither (Figure 8). We used the SNP effect categories defined by VEP: low (no protein change), moderate (e.g., missense), high (e.g., nonsense) and modifier (non-coding, e.g., 3′ UTR variant). Using a CMH chi-squared test for independence, we found a significant difference in counts of SNP type by marker class ($M^2 = 10.363$, df = 3, $p = .0157$), and by contact zone ($M^2 = 19.702$,

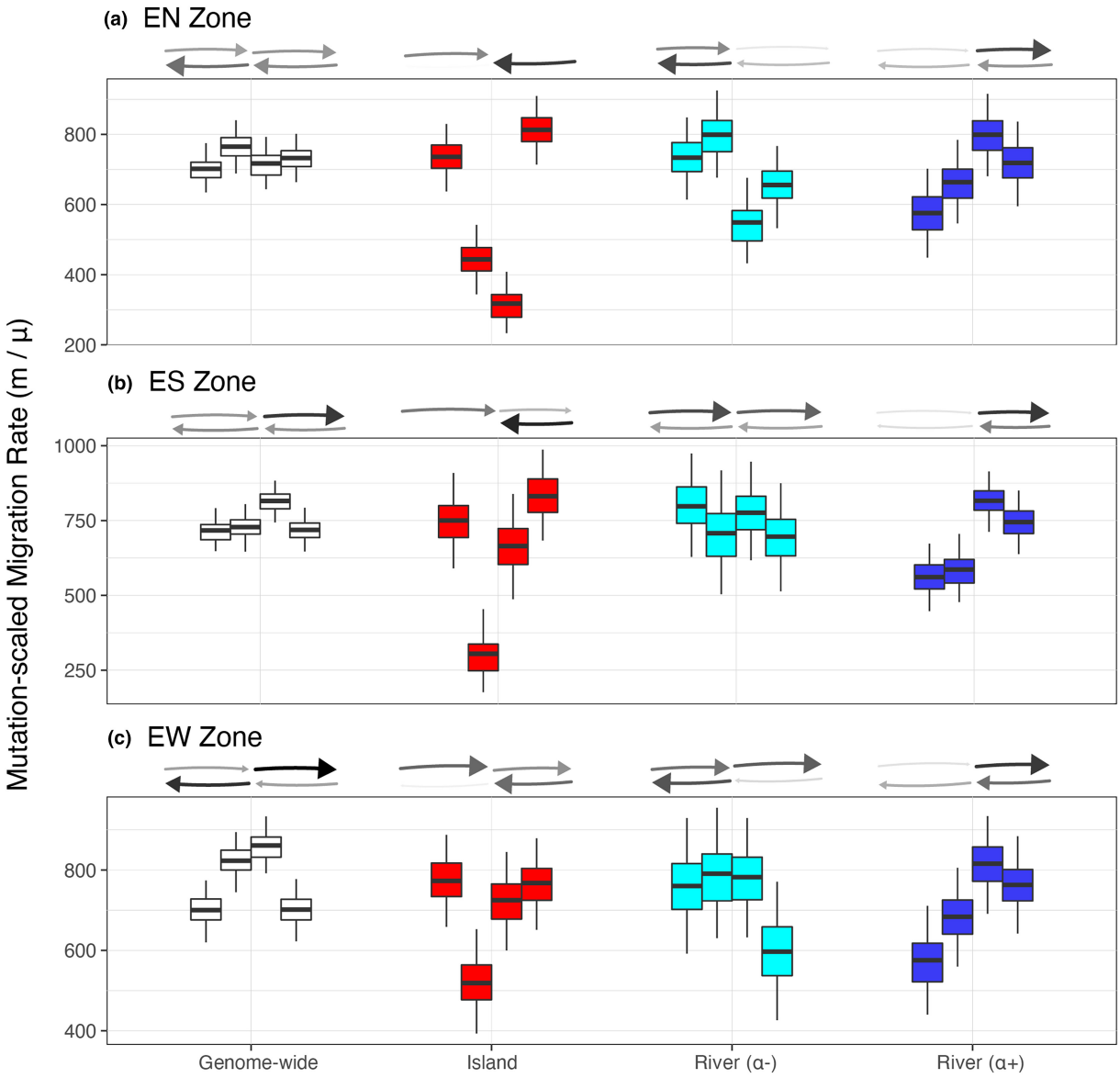

FIGURE 5 Direction of migration for islands versus rivers. Results of MIGRATE-N analyses for each contact zone and marker class separately. Relative to genome-wide patterns, island loci tend to act as migrational sinks, and river loci tend to show directional introgression from one lineage or the other. From left to right, each group of four boxplots summarizes the following migration rates: O → Ad, Ad → O, Ad → E, E → Ad, where E = Y-East, Ad = Admixed, and O = Other (Y-North/Y-South/Y-West, depending on contact zone). Each mutation-scaled migration rate (m / μ) is summarized by median and 95% credible intervals. Arrows indicate strength of migration from O (left) to E (right). Shading of arrows is proportional to corresponding migration rates and scaled by minimum and maximum rates. ′Genome-wide′ rates are based on 100 randomly sampled loci. Rivers are separated into two classes: $\alpha-$ and $\alpha+$, based on the direction of introgression (Figure 4).

df = 6, $p = .0031$). Woolf tests showed that odds ratios were homogenous across strata for both marker class ($\chi^2 = .243$, df = 2, $p = .8857$) and contact zone ($\chi^2 = .024$, df = 1, $p = .8763$), in line with CMH assumptions. To pinpoint which SNP types were most important in differentiating the three contact zones, we pooled the data across both marker types and performed a standard Pearson's chi-squared test. The residuals from the chi-squared test show that a different SNP type predominates in each contact zone: moderate (EN), non-coding (ES) and low (EW). To pinpoint which SNP types were most important in differentiating the two marker types, we pooled the data across all three contact zones and performed a standard Pearson's chi-squared test. The residuals from the chi-squared test show that differences between islands and rivers were driven almost entirely by non-coding (possibly regulatory) and low-effect SNPs.

3.7 | Power analyses

Simulations of power found that our significant results (found using 95% quantile cutoffs) were reproducible for quantile values up to 96% (CMH test) or 97% (Fisher's test) (Figures S12 and

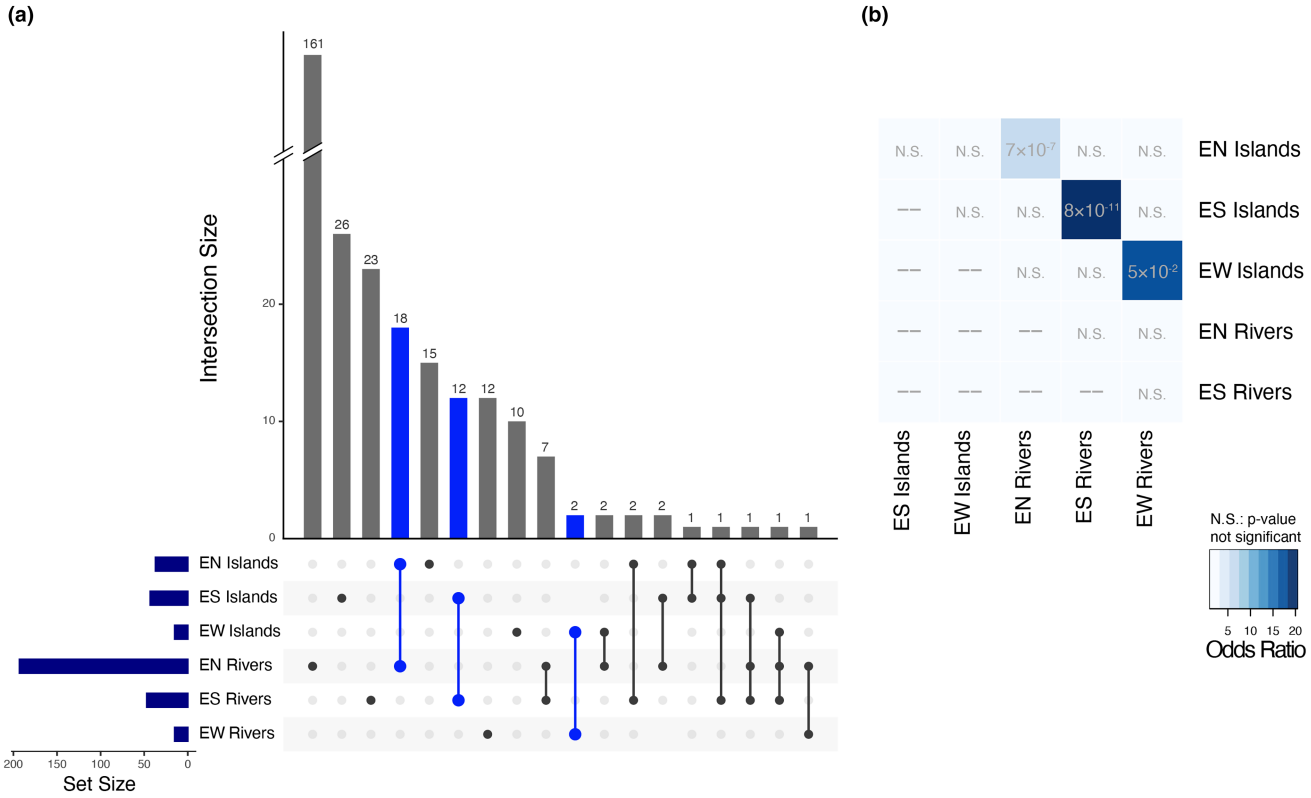

FIGURE 6 Outlier loci overlap primarily by contact zone. Islands and rivers share a significant number of loci within the same contact zone only. (a) Upset plot showing the intersection of six categories of loci (islands vs. rivers, across three contact zones). For each intersection shown in a column, the number of shared loci (intersection size) is shown, based on the total number of loci in each category (set size, shown by histogram on rows). Intersections that are specific to a contact zone (e.g., EN Islands + EN Rivers) are highlighted in blue; no intersections involving all islands or all rivers are observed. (b) Results of pairwise Fisher's exact tests of the null hypothesis that categories do not significantly overlap. Odds ratios are represented by the blue heatmap, and significant *p*-values are shown on their respective box.

S13). Power exceeded 0.95 for these values and then dropped to nearly zero at the 99% level. The only exception was a lower level of maximum power (0.6) for Fisher's tests in the EW zone, which had fewer overlapping markers. To correctly reject the null hypothesis using a stringent 99% quantile cut-off, we estimated the need for at least 1.5–2.75× additional markers. Data set size is limiting partly due to small effect sizes (proportion of overlapping markers) which we estimated to be: 0.09 (EN), 0.19 (ES) and 0.07 (EW).

3.8 | Relation of islands and rivers to tadpole phenotype

The best model of tadpole length by stage was a second-order polynomial curve, as determined by lowest BIC score ($R^2_{adj}=0.481$, $F(2,1702)=791.2$, $p<.001$; Table S2; Figure 9a). Both stage ($p<.001$) and stage^2 ($p<.001$) were significant variables in the model, and a Q–Q plot showed the residuals to be normally distributed. The GWAS between SNP markers and growth-development residuals found 10 RAD loci out of 1302 with $-\log_{10}(p)$ greater than the candidate threshold of $-\log_{10}\left(5\times10^{-6}\right)$ (Figure 9b). A spatial PCA based upon these 10 loci found 91.6% of the spatial genetic variance in the first (65.1%) and second (26.5%) spatial principal components (sPCs). The first sPC is divided sharply at the EN contact zone, while the second sPC shows a pattern of Y-East ancestry that has been spread to several meadows in other lineages, particularly the Y-West but also the Y-South lineage (Figure 9c,d).

None of the 10 loci influencing tadpole phenotype were strictly islands or rivers based on our originally defined thresholds. However, we ranked them by quantiles of D_{XY} and α to assess whether they had potential to follow one pattern or the other (Table 2). One locus had been removed from island/river analysis due to missing data in key meadows. Seven of the other nine were above the 90th quantile for islands, above the 90th quantile for rivers, or below the 10th quantile for rivers in at least one contact zone. Interestingly, one locus (C2952) was nearly an island for all three contact zones, and a river for the EW contact zone (Table 2). The pattern of introgression from Y-East to Y-West that we found for the 10 loci (Figure 9d) fits with the extreme negative α value for C2952 in the EW contact zone. C2952 was one of only two loci that successfully received gene annotation information and was identified as the phosphatidate phosphatase (LPIN3) gene. This gene produces an enzyme that plays a crucial role in multiple lipid metabolic pathways, by catalyzing (regulating) the conversion of triacylglycerols to phospholipids, and vice versa.

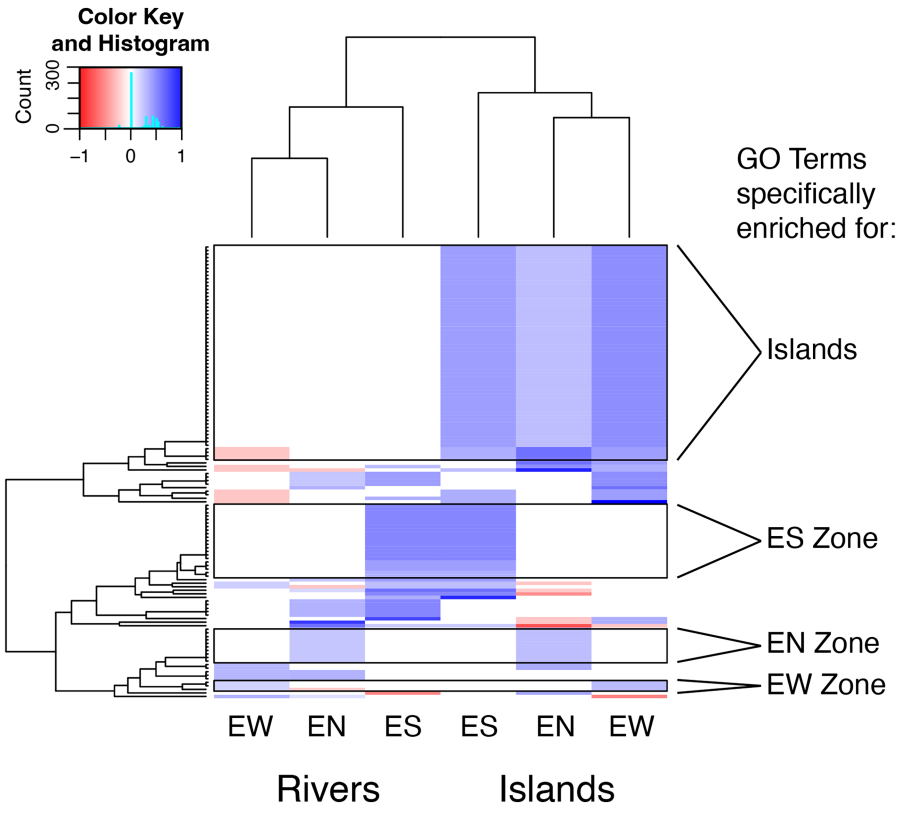

FIGURE 7 Gene ontologies cluster by marker type and contact zone. Most shared gene ontologies (GO) are unique to island loci. Hierarchical clustering analysis based on log-transformed *p*-values of GO term enrichment or depletion in each category of loci. Values representing depleted GO terms are given a (−) sign to polarize enriched/depleted values. Increasing shades of blue show enriched GO terms and increasing shades of red show depleted GO terms. Marker categories distinctly cluster into islands and rivers (top), whereas GO terms form two primary clusters that include (1) the majority of GO terms specifically enriched for islands, and (2) the majority of GO terms specifically enriched for each of the three contact zones (e.g., EN Islands + EN Rivers only). These GO terms are highlighted with four boxes.

4 | DISCUSSION

We found Scenario S2 best explains the pattern of islands/rivers significantly overlapping within each contact zone, without any significant overlap of either islands or rivers collectively (Figure 6). However, at the level of gene ontologies, islands cluster together overwhelmingly (Figure 7), suggesting some predictability to functional gene networks involved in divergent adaptation. Altogether, this means adaptation within lineages and adaptation between lineages are compatible processes, and likely synergistic in some cases. This is best exemplified by our GWAS results, showing metabolic adaptation for tadpoles appearing to involve a gene that is largely fixed between lineages, and yet introgressing from Y-East into Y-West (Figure 9). At the same time, some islands probably bestow hybrids with incompatibilities, given that islands are migrational sinks overall, within admixture zones (Figure 5).

4.1 | Compatibility of adaptive introgression with the speciation process

Divergence islands in Yosemite toads may underly climatic adaptations: previous work suggested that their lineages adapted to different refugial climates during the Pleistocene (Maier et al., 2019). Linked selection of locally adaptive climate mutations can support coadapted gene complexes (Feder, Egan, & Nosil, 2012; Feder, Gejji, et al., 2012; Nosil et al., 2009; Nosil & Feder, 2012; Via, 2012), which may erode if recombination overcomes selection pressure (Barton & Bengtsson, 1986; Samuk et al., 2017; Yeaman, 2013). Our results show that significantly more river loci are co-opted from divergence islands than expected by chance alone (Figure 6), and this pattern is consistently found in all three contact zones. There are several reasons that recombinant adaptation may be fueled by divergent adaptation. Transgressive segregation, or phenotypes more extreme in hybrids than either parent, is most likely for additive alleles that are extremely divergent (Rieseberg et al., 1999, 2003). Extreme recombinant variation can be repeatedly and selectively filtered through the F1 generation, expediting selection at loci that were already of adaptive importance (Hedrick, 2013). This process could be particularly potent for the Yosemite toad, because its extremely small population sizes (mean $N_e \approx 30$; Maier et al., 2019) cause natural selection to operate very inefficiently on novel mutations or standing variation.

As genome-wide rates of divergence and introgression are increasingly measured and compared, other researchers are finding positive relationships between locus-specific levels of divergence

TABLE 1 Summary of gene ontology (GO) terms.

Broader (level 1[a] and 2) GO Category[b]		*Specific* GO Terms[b] *only* enriched in:			
		All islands	EN Zone[c]	ES Zone[c]	EW Zone[c]
BP	Biological regulation	4		1	
BP	Cellular component organization or biogenesis	11	2		
BP	Cellular process	24	3	5	1
BP	Developmental process	17		2	
BP	Growth	2			
BP	Immune system process	1			
BP	Localization		1	2	
BP	Metabolic process	9		2	1
BP	Multi-organism process	1			
BP	Multicellular organismal process	15		1	
BP	Negative regulation of biological process	6		2	
BP	Positive regulation of biological process	4			
BP	Regulation of biological process	11		2	1
BP	reproduction			1	
BP	Reproductive process			1	
BP	Response to stimulus	1		3	
BP	Signalling	1			
BP	Single-organism process	25	3	7	
CC	Cell part	8	3	2	
CC	Extracellular region		1		
CC	Extracellular region part	1	1		
CC	Macromolecular complex	8			
CC	Membrane part				1
CC	Membrane-enclosed lumen	3			
CC	Organelle	8	4	1	
CC	Organelle part	7	3		
MF	Binding	8	1	6	1
MF	Catalytic activity	3		2	
MF	Transcription factor activity, protein binding	3			

Note: GO terms that are unique to a category of locus (e.g., all islands only) are listed. Specific terms are not shown, but rather summarized at the level of broader GO term for brevity. For a complete listing of specific GO terms, see Table S1.

[a]Lev. 1 categories: BP = biological process, CC = cellular component, MF = molecular function.

[b]Specific and broad GO terms do not form a simple hierarchy; a specific GO term (e.g., aortic smooth muscle cell differentiation) may fall into multiple categories at Level 2 (i.e., cellular process, developmental process, and multicellular organismal process, and single-organism process).

[c]GO terms enriched in both islands/rivers of that zone, but no other zones, e.g., EN islands + EN rivers only.

and introgression for many hybrid zones: for example, in butterflies (Gompert et al., 2012), house mice (Janoušek et al., 2015), budworms (Blackburn et al., 2017), thrushes (Bay & Ruegg, 2017), and rattlesnakes (Schield et al., 2017). Analyses of whole-genome data sets have found specific adaptively divergent trait loci that subsequently introgressed across species boundaries. For example, *Heliconius* butterflies among different hybrid zones are highly divergent at large effect loci that control mimicry coloration, and

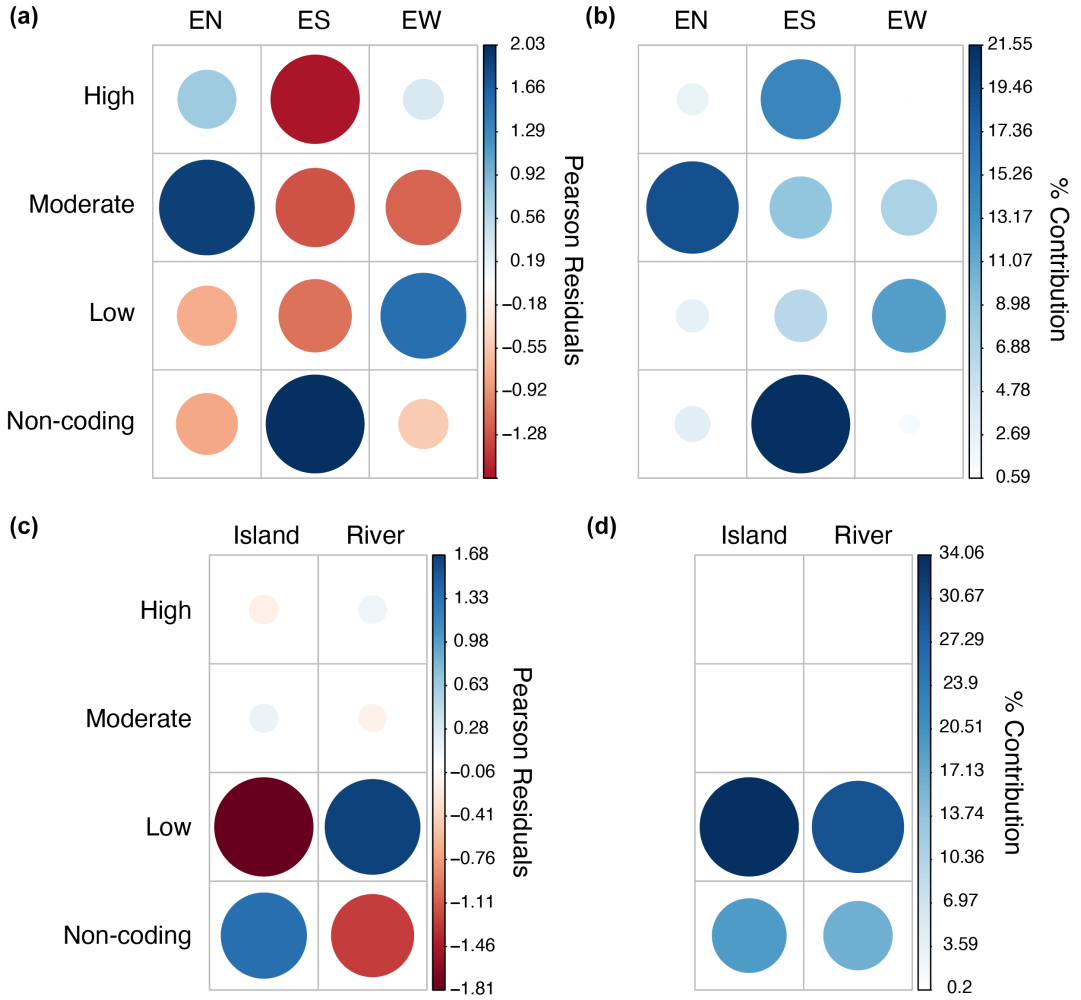

FIGURE 8 SNP effect categories differ by marker type and contact zone. Balloon plots summarize the results of a Cochran–Mantel–Haenszel (CMH) chi-squared test, and a subsequent standard chi-squared test, on the independence of SNP effect category across contact zones (a, b) and marker classes (c, d). Panels (a) and (c) show Pearson residuals of the standard chi-squared test, with increasingly positive values (larger, more blue) suggesting a positive association, and increasingly negative values (larger, more red) suggesting a negative association. Panels (b) and (d) show contribution (%) of each cell to the overall χ^2 score.

yet some of these loci have introgressed adaptively between species pairs (Baxter et al., 2008; Joron et al., 2006; Martin et al., 2012; Nadeau et al., 2014; Pardo-diaz et al., 2012; Reed et al., 2011). A similar pattern has been found for flower coloration genes in monkeyflowers (Stankowski & Streisfeld, 2015), insecticide resistance in mosquitos (Clarkson et al., 2014), and beak morphology in Darwin's finches (Grant, 2015; Grant & Grant, 2014). Clearly, a consilience of examples in nature supports the idea of genomic rivers co-opting existing divergence islands.

Some aspects of theory may predict the opposite pattern, however. There are constraints on the genomic composition of hybrids, and thus on divergence islands amenable to adaptive introgression. For example, it is well known introgression is limited on sex chromosomes, particularly in species with extreme heterogamety (Coyne & Orr, 1989, 1997; Presgraves et al., 2003; Swanson et al., 2001; Turelli & Hoffmann, 1995). This is presumably because BDM incompatibilities accumulate on X or W chromosomes, where disruption of coadapted genes is severely limited by negative epistasis (Presgraves, 2008; Turelli & Orr, 2000). Moreover, when hybrid genomes initially form, portions of the genome are constrained to come from one specific parent (Runemark et al., 2018). However these patterns may only reflect constraints on the most reproductively isolated portions of the genome, and not the vast majority of divergence islands (e.g., Taylor et al., 2014). Ultimately, it may be that only strongly selected islands with epistatic effects on reproduction are refractory to adaptive introgression (Abbott et al., 2013).

It is difficult to estimate what proportion of islands in the Yosemite toad genome may confer reproductive isolation; the ddRADseq data set in this study only represents a small proportion of sites with limited gene annotation available. However, two of our results may shed some light on the answer: (1) collectively, islands in every contact zone are migrational sinks (Figure 5), suggesting some islands may contain speciation genes; and (2) there is a paucity of fully fixed loci, so introgression may overcome reproductive barriers throughout most of the genome. It is also important to note that

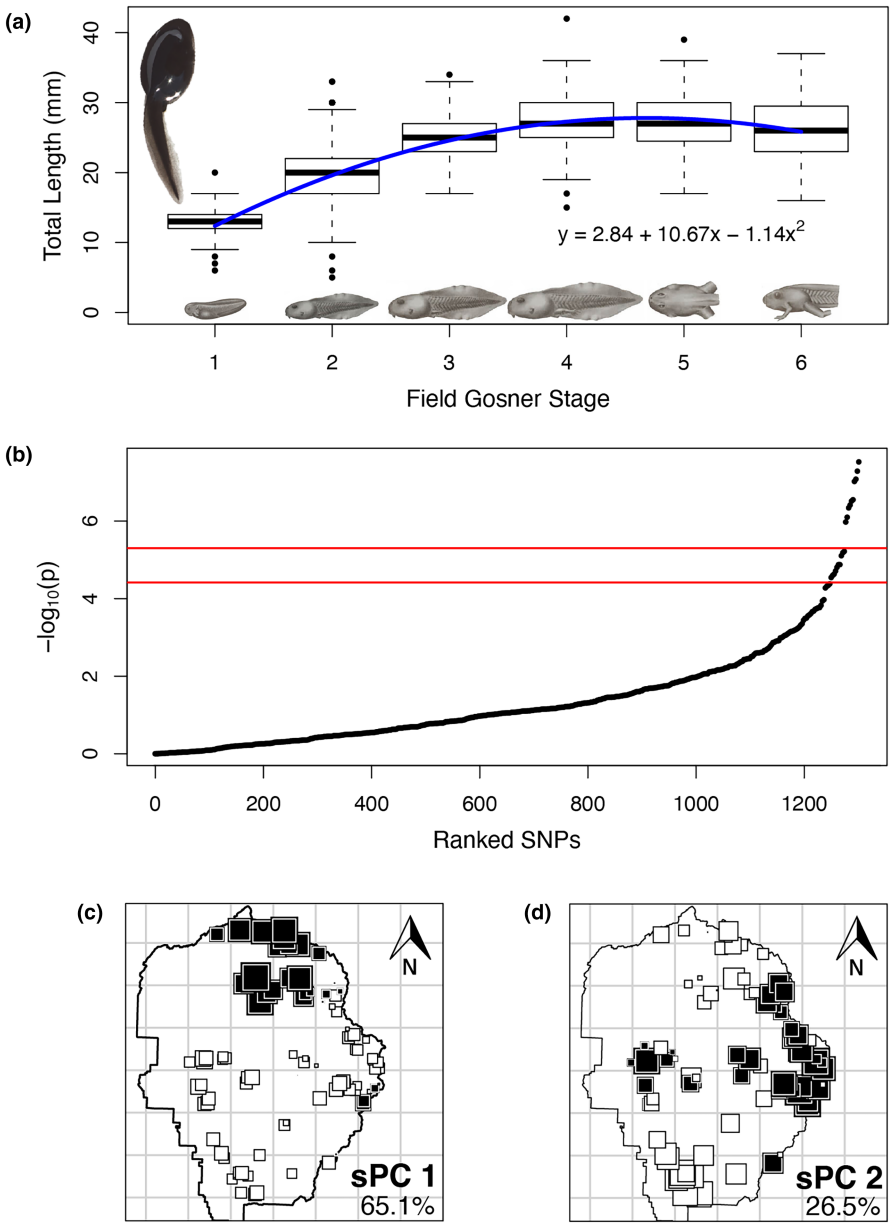

FIGURE 9 Loci associated with tadpole growth and development. Loci underlying tadpole growth and development show spatial patterns that are both lineage-specific (island-like) and lineage-introgressive (river-like). (a) Best polynomial regression model of total tadpole length based on Gosner developmental stage (summarized field version). Boxplots show mean and interquartile range with 95% confidence intervals as whiskers. (b) Results of a genome-wide association study using SNPs as predictors, and residuals from the phenotypic model as the response variable. The two red lines represent 'Bonferroni' and 'candidate' thresholds. (c, d) Spatial genetic patterns for the ten candidate loci, as calculated by a spatial PCA (sPCA), are shown for the first two sPCs. Percent of total variance is shown in the bottom right for each sPC. Squares throughout Yosemite National Park represent sPC scores: more positive (larger/black), more negative (larger/white), or closer to zero (smaller).

'adaptive' introgression is predicated on the assumption that most α outliers are not affected by drift, which is supported by simulations (Gompert & Buerkle, 2011, 2012). Although it remains possible some rivers are flowing stochastically, this does not explain island/river overlap consistently found across all contact zones.

4.2 | Predictability of adaptation

Intriguingly, we found that island loci are united by many common gene ontologies, despite being mostly distinct loci (Figure 7). If there is stochasticity at the level of loci, but predictability at the level of gene networks, this favours the view that islands represent parallel divergent adaptation (Stern, 2013; Stern & Orgogozo, 2008, 2009). Other studies have found similar results. For example, house mouse divergence outliers also cluster by gene ontology (Janoušek et al., 2015), and crow divergence genes affecting plumage coloration show 'parallelism by pathway', not identical speciation genes (Vijay et al., 2016). These patterns would imply that although distinct loci are involved in genetic divergence, they might underlie related traits with polygenic inheritance. This raises the question: how predictable is adaptation in speciation, and generally?

One survey of the published literature estimates the mean probability for a gene to be reused in parallel or convergent adaptation is 0.32–0.55, and declines with age of divergence (Conte et al., 2012). However, this high estimate may partly be biased from genes of large effect preferentially studied, when polygenic traits of smaller effect may dominate most of adaptation (Berg & Coop, 2014; Rockman, 2012; Yeaman, 2015). Even if this estimate were off by an order of magnitude, our results do not match any such pattern of parallel island evolution, at the level of individual RAD loci (Figure 6). It may be that our short-read loci obscure

TABLE 2 Genetic markers underlying tadpole growth and development.

RAD locus	p[a]	Island (D_{XY})			River (α)			Transcript[b]	Mutation[c]	Gene[d]	Protein domains[e]	GO terms[f]
		EN	ES	EW	EN	ES	EW					
C2952	3×10^{-8}	***0.93***	***0.94***	***0.98***	0.28	0.66	***0.02***	Yes	Missense (K/I)	Phosphatidate phosphatase (LPIN3)	• Lipin_N (conserved N-terminus) • LNS2 (plasmid maintenance, respiration)	• Fatty acid metabolism • Phosphatidate phosphatase • Glycerophospholipid biosynthesis • Phosphatidylcholine biosynthesis • Phosphatidylethanolamine biosynthesis • Phospholipid metabolism • Small molecule metabolism • Endoplasmic reticulum membrane
C18725	5×10^{-8}	0.48	0.17	–	0.13	***0.09***	–	Yes	–	–	–	–
C22735	8×10^{-3}	***0.97***	0.28	0.17	0.54	***0.01***	***0.92***	No	–	–	–	–
C9564	1×10^{-7}	–	–	–	–	–	–	Yes	–	–	–	–
C13533	3×10^{-7}	–	0.38	0.11	–	0.62	***0.02***	No	–	–	–	–
I34049	3×10^{-7}	***0.92***	***0.9***	0.77	0.25	0.14	0.47	No	–	–	–	–
I49070	4×10^{-7}	–	***0.95***	0.65	–	0.8	0.78	No	–	–	–	–
I41608	5×10^{-7}	0.81	0.41	0.62	0.83	0.82	0.27	Yes	–	–	–	–
C402	8×10^{-7}	***0.95***	0.89	0.89	0.84	0.62	0.41	Yes	Missense (T/A)	Gastrula zinc finger protein (ZG57)	Zinc finger, C2H2 type	• DNA-binding transcription activator • Regulation of transcription • RNA polymerase II DNA binding
C14559	1×10^{-6}	0.53	–	0.73	0.26	–	0.38	No	–	–	–	–

Note: Profile of the ten markers that were identified by the GWAS. For each RAD locus, quantiles of D_{XY} and α are provided in each contact zone, to illustrate whether that marker is highly fixed (island-like) and/or highly introgressed (river-like). Underlined and italicized numbers highlight quantiles in the 10% tails of the distribution.

[a]p-Value for the SNP coefficient in the GWAS model.

[b]Whether the RAD sequence matched to a transcript in the transcriptome (annotation in subsequent columns as available).

[c]Type of mutation, with alternate amino acid states.

[d]From SwissProt database.

[e]From Pfam database.

[f]Specific gene ontology (GO) terms describing gene function and localization.

parallelism at the level of genes, but this is unlikely given that loci are enzymatically cut from disparate locations in the genome (Davey et al., 2011; Peterson et al., 2012). Another possibility is that young lineages are more prone to parallelism by pathway than by gene. Rewiring of protein interaction networks is known to be an important driver of phenotypic change, by shifting the importance of individual genes without disrupting pleiotropic interactions (Kim et al., 2012; Stern, 2013). This pattern was found in a series of diverging stick insect populations: only a proportion of divergent SNPs were shared among all population pairs and reflected host-plant ecological divergence, whereas the remainder were idiosyncratic SNPs, but shared gene ontologies (Soria-Carrasco et al., 2012). As mentioned above, given this element of stochasticity to island formation, there may be less bias toward developing strong reproductive isolation, thus increasing chances of adaptive introgression at a locus.

Thus, overall predictability in phenotypic evolution might not have a 1:1 relationship with genomic changes, which can manifest as simple mutations, protein interaction shifts (Kim et al., 2012), expression differences (Brawand et al., 2011) or even changes in developmental systems (Verster et al., 2014). This could explain why predictability for single-gene adaptation in bacteria (Gaut, 2015; Roy, 2009; Weinreich et al., 2006) does not scale up to predictability at the level of speciation genes (Cutter, 2015; Gaut, 2015; Mandeville et al., 2015). Our results are consistent with a growing number of examples that predictability in evolution can manifest at various scales in the genes-to-phenotype pathway, but with stochasticity at other scales. We found predictability for gene ontologies of divergent adaptation, some predictability for a process of adaptive introgression co-opting island loci, and stochasticity at the level of individual RAD loci and genic SNP effects (Figures 6–8). One criticism of divergence island studies generally is that islands may simply be regions of low recombination that are unrelated to levels of adaptation and introgression (Cruickshank & Hahn, 2014). However, using semi-relative (Φ_{ST}) as well as absolute (D_{XY}, Maddison and Slatkin's '*s*') measures of divergence is the prescribed method of minimizing that possibility (Cruickshank & Hahn, 2014; Noor & Bennett, 2009; Wolf & Ellegren, 2017). These measures as well as the use of replicated contact zones to test for adaptation gives increased confidence of the results attained.

4.3 | Consequences for Yosemite toad tadpole growth and development

Apparent introgression of the highly differentiated LPIN3 candidate gene (Figure 9) is very interesting, because it may indicate the importance of adaptive introgression for the evolution of desiccation resistance. We add the caveat that our short-read RAD sequences offer limited genic and genomic context, making annotation imperfect and adaptive interpretation speculative. The LPIN3 gene plays a role in regulating lipid metabolism, and so one possibility is it regulates the tradeoff between growth and development in response to a drying environment. Yosemite toad tadpoles in the Y-East lineage have been shown to possess faster development times than any other lineage (P. Maier, unpublished data), and so this trait may be spreading adaptively into the lower elevation Y-West. Lower elevation toads face more intense selective death due to climate change, where other climate-related loci such as MAP3K5 show patterns of adaptation (Maier et al., 2023).

The patterns found for RAD locus C2952—the locus that matched to LPIN3—should be followed up with full gene sequencing to confirm the pattern found. Previous studies have found that expression levels of thyroid receptor genes (TRα and TRβ), as well as thyroid hormone (TH), and corticosterone (CORT) explain the faster development and metamorphosis of certain spadefoot species (Gomez-Mestre et al., 2013; Hollar et al., 2011; Kulkarni & Buchholz, 2012). LPIN3's putative role in regulating Yosemite toad tadpole growth versus development during desiccation is consistent with data showing that faster-developing desert spadefoots reduce their developmental plasticity, and dramatically reduce storage of fat bodies (Kulkarni et al., 2011). Another possibility is that for fast-developing Yosemite toad tadpoles, LPIN3 variants keep growth rate at uniformly low levels. Fast-developing species are known to sacrifice variable growth rates in order to accelerate development rate, at the expense of metamorphosing much smaller (Richter-Boix et al., 2011). Regardless of the exact metabolic role LPIN3 plays, it is part of a larger pattern of likely adaptive introgression for the Yosemite toad.

5 | CONCLUSIONS

Adaptive divergence and introgression may be more complementary forces than currently appreciated. Clearly, they are separate forces that can act antagonistically, but our work supports the view that introgression between incipient species may build off the accumulated differences in a constructive way. We showed evidence that admixture in secondary contact zones is a probable source for new adaptive variation in the Yosemite toad. More specifically, we found evidence that one of the Yosemite toad's lineage-specific adaptations—a metabolic gene underlying tadpole growth and development—may be spreading into lower-elevation regions with stronger climate change pressure. The Yosemite toad is under great future threat from a warming climate, because it is a species that specializes on breeding in shallow, ephemeral snowmelt ponds; shifting snowmelt phenology and more unpredictable precipitation are expected to directly impact its breeding success (Brown et al., 2015; U.S. Fish & Wildlife Service, 2014). It has already been suggested that conservation managers should utilize hybrids—specifically advanced hybrids with well-understood recombinant adaptations to climate—as a resource for conservation genetics (Hamilton & Miller, 2016). The series of contact zones in Yosemite may be a perfect opportunity to use this natural process for conservation of the species, given additional experimental research into the fitness of advanced hybrids in nature.

AUTHOR CONTRIBUTIONS

All authors contributed to spatial study design. PAM and AGV designed the transcriptome sequencing. PAM designed the analytical approaches, secured funding, collected tissue samples, performed the laboratory work and conducted the bioinformatics and analyses, with guidance from AJB and AGV. PAM wrote the manuscript, with input from AJB and AGV.

ACKNOWLEDGEMENTS

This research was supported the Harold & June Memorial, Jordan D. Covin, and ARCS scholarships awarded to PAM, and the USGS Ecosystems Mission Area. The UC Merced Sierra Nevada Institute provided housing and accommodations for fieldwork. All animal handling was performed in accordance with SDSU animal care and use protocol #13-03-001B. All data and sample collection were only completed after obtaining and adhering to the regulations of NPS research permits. The UC Riverside High Performance Computer Cluster was used for bioinformatic analyses. We thank Dean H. Leavitt for comments that improved this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

Raw RNAseq data used for transcriptome assembly are deposited at NCBI GenBank SRA under BioProject PRJNA574353 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA574353). Raw double-digest RADseq data from a previous study are available at NCBI GenBank SRA under BioProject PRJNA558546 (https://www.ncbi.nlm.nih.gov/sra/PRJNA558546). R, Python, and shell scripts for data conversion, tests, and models are available on Dryad (https://doi.org/10.5061/dryad.pg4f4qrvj). Information about all samples used in this study, in addition to supplemental tables, figures, and methods, are available in the Supporting Information S1.

ORCID

Paul A. Maier https://orcid.org/0000-0003-0851-8827
Amy G. Vandergast https://orcid.org/0000-0002-7835-6571
Andrew J. Bohonak https://orcid.org/0000-0002-5429-2331

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., Eroukhmanoff, 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

Agresti, A. (2002). An introduction to categorical data analysis. In *Wiley Series in Probability and Statistics* (2nd ed.). John Wiley & Sons, Inc.

Anderson, E. C., & Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, *160*(3), 1217–1229. https://doi.org/10.1093/genetics/160.3.1217

Arnold, M. L., Ballerini, E. S., & Brothers, A. N. (2012). Hybrid fitness, adaptation and evolutionary diversification: Lessons learned from Louisiana irises. *Heredity*, *108*(3), 159–166. https://doi.org/10.1038/hdy.2011.65

Arnold, M. L., & Hodges, S. A. (1995). Are natural hybrids fit or unfit relative to their parents? *Trends in Ecology & Evolution*, *10*(2), 67–71. https://doi.org/10.1016/S0169-5347(00)88979-X

Arnold, M. L., & Martin, N. H. (2009). Adaptation by introgression. *Journal of Biology*, *8*(9), 9–11. https://doi.org/10.1186/jbiol176

Arnold, M. L., & Martin, N. H. (2010). Hybrid fitness across time and habitats. *Trends in Ecology & Evolution*, *25*(9), 530–536. https://doi.org/10.1016/j.tree.2010.06.005

Barrett, R. D., Rogers, S. M., & Schluter, D. (2008). Natural selection on a major armor gene in threespine stickleback. *Science*, *322*(5899), 2006–2008. https://doi.org/10.1126/science.1159978

Barton, N., & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, *57*(3), 357–376. https://doi.org/10.1038/hdy.1986.135

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

Baxter, S. W., Papa, R., Chamberlain, N., Humphray, S. J., Joron, M., Morrison, C., ffrench-Constant, R. H., McMillan, W. O., & Jiggins, C. D. (2008). Convergent evolution in the genetic basis of Müllerian mimicry in *Heliconius* butterflies. *Genetics*, *180*(3), 1567–1577. https://doi.org/10.1534/genetics.107.082982

Bay, R. A., & Ruegg, K. (2017). Genomic islands of divergence or opportunities for introgression? *Proceedings of the Royal Society of Biology*, *284*(9), 2229–2243. https://doi.org/10.1111/j.1558-5646.2007.00179.x

Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, *13*(4), 969–980. https://doi.org/10.1111/j.1365-294X.2004.02125.x

Beerli, P., & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences*, *98*(8), 4563–4568. https://doi.org/10.1073/pnas.081068098

Berg, J. J., & Coop, G. (2014). A population genetic signal of polygenic adaptation. *PLoS Genetics*, *10*(8), e1004412. https://doi.org/10.1371/journal.pgen.1004412

Blackburn, G. S., Brunet, B. M. T., Muirhead, K., Cusson, M., Béliveau, C., Levesque, R. C., Lumley, L. M., & Sperling, F. A. H. (2017). Distinct sources of gene flow produce contrasting population genetic dynamics at different range boundaries of a *Choristoneura* budworm. *Molecular Ecology*, *26*(23), 6666–6684. https://doi.org/10.1111/mec.14386

Brady, L. D., & Griffiths, R. A. (2000). Developmental responses to pond desiccation in tadpoles of the British anuran amphibians (Bufo bufo, B. Calamita and Rana temporaria). *Journal of Zoology*, *252*(2000), 61–69. https://doi.org/10.1111/j.1469-7998.2000.tb00820.x

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., Albert, F. W., Zeller, U., Khaitovich, P., Grützner, F., Bergmann, S., Nielsen, R., Pääbo, S., & Kaessmann, H. (2011). The evolution of gene expression levels in mammalian organs. *Nature*, *478*(7369), 343–348. https://doi.org/10.1038/nature10532

Brown, C., Hayes, M. P., Green, G. A., Macfarlane, D. C., & Lind, A. J. (2015). Yosemite toad conservation assessment.

Buerkle, C. A., Gompert, Z., & Parchman, T. L. (2011). The n = 1 constraint in population genomics. *Molecular Ecology*, *20*(8), 1575–1581. https://doi.org/10.1111/j.1365-294X.2011.05046.x

Castric, V., Bechsgaard, J., Schierup, M. H., & Vekemans, X. (2008). Repeated adaptive introgression at a gene under multiallelic

balancing selection. *PLoS Genetics*, *4*(8), e1000168. https://doi.org/10.1371/journal.pgen.1000168

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

Catchen, J. M., 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

Clarkson, C. S., Weetman, D., Essandoh, J., Yawson, A. E., Maslen, G., Manske, M., Field, S. G., Webster, M., Antão, T., MacInnis, B., Kwiatkowski, D., & Donnelly, M. J. (2014). Adaptive introgression between *Anopheles* sibling species eliminates a major genomic Island but not reproductive isolation. *Nature Communications*, *5*(1), 4248. https://doi.org/10.1038/ncomms5248

Conte, G. L., Arnegard, M. E., Peichel, C. L., & Schluter, D. (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 5039–5047. https://doi.org/10.1098/rspb.2012.2146

Coyne, J. A. (1992). Genetics and speciation. *Nature*, *355*(6360), 511–515. https://doi.org/10.1038/355511a0

Coyne, J. A., & Orr, A. (1997). Patterns of speciation in *drosophila* revisited. *Evolution*, *51*(1), 295–303. https://doi.org/10.2307/2410984

Coyne, J. A., & Orr, H. A. (1989). Patterns of speciation in *drosophila*. *Evolution*, *43*(2), 362–381. https://doi.org/10.1111/j.1558-5646.1989.tb04233.x

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–3157. https://doi.org/10.1111/mec.12796

Cutter, A. D. (2012). The polymorphic prelude to Bateson-Dobzhansky-muller incompatibilities. *Trends in Ecology & Evolution*, *27*(4), 209–218. https://doi.org/10.1016/j.tree.2011.11.004

Cutter, A. D. (2015). Repeatability, ephemerality and inconvenient truths in the speciation process. *Molecular Ecology*, *24*(8), 1643–1644. https://doi.org/10.1111/mec.13163

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of Favoured races in the struggle for life* (1st ed.). John Murray.

Darwin, C. (1871). *The descent of man, and selection in relation to sex* (1st ed.). John Murray.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499–510. https://doi.org/10.1038/nrg3012

Dobzhansky, T. (1937). *Genetics and the origin of species* (3rd ed.). Columbia University Press.

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., Künstner, A., Mäkinen, H., Nadachowska-Brzyska, K., Qvarnström, A., Uebbing, S., & Wolf, J. B. W. (2012). The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, *491*(7426), 756–760. https://doi.org/10.1038/nature11584

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479–491. https://doi.org/10.1007/s00424-009-0730-7

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Feder, J. L., Gejji, R., Yeaman, S., & Nosil, P. (2012). Establishment of new mutations under divergence and genome hitchhiking. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*(1587), 461–474. https://doi.org/10.1098/rstb.2011.0256

Fitzpatrick, B. M. (2012). Estimating ancestry and heterozygosity of hybrids using molecular markers. *BMC Evolutionary Biology*, *12*(1), 131. https://doi.org/10.1186/1471-2148-12-131

Fraïsse, C., Roux, C., Welch, J. J., & Bierne, N. (2014). Gene-flow in a mosaic hybrid zone: Is local introgression adaptive? *Genetics*, *197*(3), 939–951. https://doi.org/10.1534/genetics.114.161380

Gascuel, O. (1997). BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution*, *14*(7), 685–695. https://doi.org/10.1093/oxfordjournals.molbev.a025808

Gaut, B. S. (2015). Evolution is an experiment: Assessing parallelism in crop domestication and experimental evolution. *Molecular Biology and Evolution*, *32*(7), 1661–1671. https://doi.org/10.1093/molbev/msv105

Gomez-Mestre, I., Kulkarni, S., & Buchholz, D. R. (2013). Mechanisms and consequences of developmental acceleration in tadpoles responding to pond drying. *PLoS One*, *8*(12), 1–12. https://doi.org/10.1371/journal.pone.0084266

Gompert, Z., & Buerkle, C. A. (2011). Bayesian estimation of genomic clines. *Molecular Ecology*, *20*, 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., Nice, C. C., Fordyce, J. A., Forister, M. L., & Buerkle, C. A. (2012). Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution*, *66*(7), 2167–2181. https://doi.org/10.1111/j.1558-5646.2012.01587.x

Gosner, K. (1960). A simplified table for staging anuran embryos and larvae with notes on identification. *Herpetologica*, *16*(3), 183–190.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Trinity: Reconstructing a full-length transcriptome without a genome from RNA-seq data. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883.Trinity

Grant, B. R. (2015). Introgressive hybridization and natural selection in Darwin's finches. *Biological Journal of the Linnean Society*, *117*, 812–822. https://doi.org/10.1111/bij.12702

Grant, P. R., & Grant, B. R. (2014). Synergism of natural selection and introgression in the origin of a new species. *American Naturalist*, *183*(5), 671–681. https://doi.org/10.1086/675496

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. P., V, Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hamilton, J. A., & Miller, J. M. (2016). Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, *30*(1), 33–41. https://doi.org/10.1111/cobi.12574

Hawks, J., & Cochran, G. (2006). Dynamics of adaptive introgression from archaic to modern humans. *PaleoAnthropology*, *2006*, 101–115.

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*, 4606–4618. https://doi.org/10.1111/mec.12415

Hewitt, G. M. (1988). Hybrid zones–natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, *3*(7), 158–167. https://doi.org/10.1016/0169-5347(88)90033-X

Hoekstra, H. E. (2006). Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity*, *97*(3), 222–234. https://doi.org/10.1038/sj.hdy.6800861

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in

threespine stickleback using sequenced RAD tags. *PLoS Genetics*, *6*(2), e1000862. https://doi.org/10.1371/journal.pgen.1000862

Hollar, A. R., Choi, J., Grimm, A. T., & Buchholz, D. R. (2011). Higher thyroid hormone receptor expression correlates with short larval periods in spadefoot toads and increases metamorphic rate. *General and Comparative Endocrinology*, *173*(1), 190–198. https://doi.org/10.1016/j.ygcen.2011.05.013

Janoušek, V., Munclinger, P., Wang, L., Teeter, K. C., & Tucker, P. K. (2015). Functional organization of the genome may shape the species boundary in the house mouse. *Molecular Biology and Evolution*, *32*(5), 1208–1220. https://doi.org/10.1093/molbev/msv011

Jiggins, C. D., Salazar, C., Linares, M., & Mavarez, J. (2008). Hybrid trait speciation and Heliconius butterflies. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *363*(1506), 3047–3054. https://doi.org/10.1098/rstb.2008.0065

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

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), 55–61. https://doi.org/10.1038/nature10944

Jones, O. R., & Wang, J. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, *10*(3), 551–555. https://doi.org/10.1111/j.1755-0998.2009.02787.x

Joron, M., Papa, R., Beltrán, M., Chamberlain, N., Mavárez, J., Baxter, S., Abanto, M., Bermingham, E., Humphray, S. J., Rogers, J., Beasley, H., Barlow, K., ffrench-Constant, R. H., Mallet, J., McMillan, W. O., & Jiggins, C. D. (2006). A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. *PLoS Biology*, *4*(10), 1831–1840. https://doi.org/10.1371/journal.pbio.0040303

Keeler-Wolf, T., Reyes, E. T., Menke, J. M., Johnson, D. N., & Karavidas, D. L. (2012). *Yosemite National Park vegetation classification and mapping project report*. National Park Service report.

Kim, J., Kim, I., Han, S. K., Bowie, J. U., & Kim, S. (2012). Network rewiring is an important mechanism of gene essentiality change. *Scientific Reports*, *2*, 1–7. https://doi.org/10.1038/srep00900

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*(2), 111–120. https://doi.org/10.1007/BF01731581

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, *173*(1), 419–434. https://doi.org/10.1534/genetics.105.047985

Kulkarni, S. S., Gomez-Mestre, I., Moskalik, C. L., Storz, B. L., & Buchholz, D. R. (2011). Evolutionary reduction of developmental plasticity in desert spadefoot toads. *Journal of Evolutionary Biology*, *24*(11), 2445–2455. https://doi.org/10.1111/j.1420-9101.2011.02370.x

Kulkarni, S. S., & Buchholz, D. R. (2012). Beyond synergy: Corticosterone and thyroid hormone have numerous interaction effects on gene regulation in *Xenopus tropicalis* tadpoles. *Endocrinology*, *153*(11), 5309–5324. https://doi.org/10.1210/en.2012-1432

Leips, J., Mcmanus, M. G., & Travis, J. (2000). Response of treefrog larvae to drying ponds: Comparing temporary and permanent pond breeders. *Ecology*, *81*(11), 2997–3008.

Lind, M. I., & Johansson, F. (2007). The degree of adaptive phenotypic plasticity is correlated with the spatial environmental heterogeneity experienced by Island populations of *Rana temporaria*. *Journal of Evolutionary Biology*, *20*(4), 1288–1297. https://doi.org/10.1111/j.1420-9101.2007.01353.x

Losos, J. B. (1992). The evolution of convergent structure in Caribbean Anolis communities. *Systematic Biology*, *41*(4), 403–420. https://doi.org/10.1093/sysbio/41.4.403

Maier, P. A., Vandergast, A. G., & Bohonak, A. J. (2023). Using landscape genomics to delineate future adaptive potential for climate change in the Yosemite toad (*Anaxyrus canorus*). *Evolutionary Applications*, *16*(1), 74–97. https://doi.org/10.1111/eva.13511

Maier, P. A., Vandergast, A. G., Ostoja, S. M., Aguilar, A., & Bohonak, A. J. (2019). Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*). *Evolution*, *73*(12), 2476–2496. https://doi.org/10.1111/evo.13868

Maier, P. A., Vandergast, A. G., Ostoja, S. M., Aguilar, A., & Bohonak, A. J. (2022a). Gene pool boundaries for the Yosemite toad (*Anaxyrus canorus*) reveal asymmetrical migration within meadow neighborhoods. *Frontiers in Conservation Science*, *3*, 1–14. https://doi.org/10.3389/fcosc.2022.851676

Maier, P. A., Vandergast, A. G., Ostoja, S. M., Aguilar, A., & Bohonak, A. J. (2022b). Landscape genetics of a sub-alpine toad: Climate change predicted to induce upward range shifts via asymmetrical migration corridors. *Heredity*, *129*(5), 257–272. https://doi.org/10.1038/s41437-022-00561-x

Malinsky, M., Challis, R. J., Tyers, A. M., Schiffels, S., Terai, Y., Ngatunga, B. P., Miska, E. A., Durbin, R., Genner, M. J., & Turner, G. F. (2015). Genomic islands of speciation separate cichlid ecomorphs in an east African crater lake. *Science*, *350*(6267), 1493–1498. https://doi.org/10.1126/science.aac9927

Mandeville, E. G., Parchman, T. L., Mcdonald, D. B., & Alex, C. (2015). Highly variable reproductive isolation among pairs of *Catostomus* species. *Molecular Ecology*, *24*(8), 1856–1872. https://doi.org/10.5061/dryad.r31sg

Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., Halder, G., Jiggins, C. D., Kronforst, M. R., Long, A. D., McMillan, W. O., & Reed, R. D. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences*, *109*(31), 12632–12637. https://doi.org/10.1073/pnas.1204800109

Martinsen, G. D., Whitham, T. G., Turek, R. J., & Keim, P. (2001). Hybrid populations selectively filter gene introgression between species. *Evolution*, *55*(7), 1325–1335. https://doi.org/10.1111/j.0014-3820.2001.tb00655.x

Mayr, E. (1942). *Systematics and the origin of species: From the viewpoint of a zoologist*. Columbia University Press.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

McKinnon, J. S., Mori, S., Blackman, B. K., David, L., Kingsley, D. M., Jamieson, L., Chou, J., & Schluter, D. (2004). Evidence for ecology's role in speciation. *Nature*, *429*, 294–298. https://doi.org/10.1038/nature02550.1

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl variant effect predictor. *Genome Biology*, *17*(1), 1–14. https://doi.org/10.1186/s13059-016-0974-4

Morey, S. R., & Reznick, D. N. (2004). The relationship between habitat permanence and larval development in California spadefoot toads: Field and laboratory comparisons of developmental plasticity. *Oikos*, *104*(1), 172–190. https://doi.org/10.1111/j.0030-1299.2004.12623.x

Nadeau, N. J., Ruiz, M., Salazar, P., Counterman, B., Medina, J. A., Ortiz-Zuazaga, H., Morrison, A., McMillan, W. O., Jiggins, C. D., & Papa, R. (2014). Population genomics of parallel hybrid zones in the mimetic butterflies, *H. melpomene* and *H. erato*. *Genome Research*, *24*, 1316–1333. https://doi.org/10.1101/gr.169292.113

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press.

Noor, M. A. F., & Bennett, S. M. (2009). Islands of speciation or mirages in the desert examining the role of restricted recombination in

maintaining species. *Heredity*, *103*(6), 439–444. https://doi.org/10.1038/hdy.2009.151

Norris, L. C., Main, B. J., Lee, Y., Collier, T. C., Fofana, A., Cornel, A. J., & Lanzaro, G. C. (2015). Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proceedings of the National Academy of Sciences*, *112*(3), 815–820. https://doi.org/10.1073/pnas.1418892112

Nosil, P., & Feder, J. L. (2012). Genomic divergence during speciation: Causes and consequences. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*, 332–342. https://doi.org/10.1098/rstb.2011.0263

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, *18*(3), 375–402. https://doi.org/10.1111/j.1365-294X.2008.03946.x

Nosil, P., & Schluter, D. (2011). The genes underlying the process of speciation. *Trends in Ecology & Evolution*, *26*(4), 160–167. https://doi.org/10.1016/j.tree.2011.01.001

Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: Accumulating Dobzhansky-muller incompatibilities. *Evolution*, *55*(6), 1085–1094. https://doi.org/10.1111/j.0014-3820.2001.tb00628.x

Orr, H. A., Masly, J. P., & Presgraves, D. C. (2004). Speciation genes. *Current Opinion in Genetics and Development*, *14*(6), 675–679. https://doi.org/10.1016/j.gde.2004.08.009

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412

Paradis, E. (2010). Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Pardo-diaz, 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), 1–13. https://doi.org/10.1371/journal.pgen.1002752

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Müller, I., Baglione, V., Unneberg, P., Wikelski, M., Grabherr, M. G., & Wolf, J. B. W. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, *344*(6190), 1410–1414. https://doi.org/10.1126/science.1253226

Presgraves, D. C. (2008). Sex chromosomes and speciation in *drosophila*. *Trends in Genetics*, *24*(7), 336–343. https://doi.org/10.1016/j.tig.2008.04.007

Presgraves, D. C., Balagopalan, L., Abmayr, S. M., & Orr, H. A. (2003). Adaptive evolution drives divergence of a hybrid inviability gene between two species of *drosophila*. *Nature*, *423*(6941), 715–719. https://doi.org/10.1038/nature01679

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Quero, G., Gutiérrez, L., Monteverde, E., Blanco, P., Pérez de Vida, F., Rosas, J., Fernández, S., Garaycochea, S., McCouch, S., Berberian, N., Simondi, S., & Bonnecarrère, V. (2018). Genome-wide association study using historical breeding populations discovers genomic regions involved in high-quality rice. *The Plant Genome*, *11*(3), 1–12. https://doi.org/10.3835/plantgenome2017.08.0076

R Core Team. (2023). *R: a language and environment for statistical computing*. https://www.r-project.org

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

Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., Jiggins, C. D., Chamberlain, N. L., Kronforst, M. R., Chen, R., Halder, G., Nijhout, H. F., & Mcmillan, W. O. (2011). Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*, *333*, 1137–1141. https://doi.org/10.1126/science.1208227

Reid, S. D., Herbelin, C. J., Bumbaugh, A. C., Selander, R. K., & Whittam, T. S. (2000). Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature*, *406*(6791), 64–67. https://doi.org/10.1038/35017546

Richter-Boix, A., Llorente, G. A., & Montori, A. (2006). A comparative analysis of the adaptive developmental plasticity hypothesis in six Mediterranean anuran species along a pond permanency gradient. *Evolutionary Ecology Research*, *8*(6), 1139–1154.

Richter-Boix, A., Tejedo, M., & Rezende, E. L. (2011). Evolution and plasticity of anuran larval development in response to desiccation: A comparative analysis. *Ecology and Evolution*, *1*(1), 15–25. https://doi.org/10.1002/ece3.2

Rieseberg, L. (2011). Adaptive introgression: The seeds of resistance. *Current Biology*, *21*(15), R581–R583. https://doi.org/10.1016/j.cub.2011.06.038

Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity*, *83*(4), 363–372. https://doi.org/10.1038/sj.hdy.6886170

Rieseberg, L. H., Widmer, A., Arntz, A. M., Burke, J. M., Carr, D. E., Abbott, R. J., & Meagher, T. R. (2003). The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *358*(1434), 1141–1147. https://doi.org/10.1098/rstb.2003.1283

Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution*, *66*(1), 1–17. https://doi.org/10.1111/j.1558-5646.2011.01486.x

Roy, S. W. (2009). Probing evolutionary repeatability: Neutral and double changes and the predictability of evolutionary adaptation. *PLoS One*, *4*(2), e4500. https://doi.org/10.1371/journal.pone.0004500

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Runemark, A., Trier, C. N., Eroukhmanoff, F., Hermansen, J. S., Matschiner, M., Ravinet, M., Elgvin, T. O., & Sætre, G. P. (2018). Variation and constraints in hybrid genome formation. *Nature Ecology & Evolution*, *2*(3), 549–556. https://doi.org/10.1038/s41559-017-0437-7

Samuk, K., Owens, G. L., Delmore, K. E., Miller, S. E., Rennison, D. J., & Schluter, D. (2017). Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Molecular Ecology*, *26*(17), 4378–4390. https://doi.org/10.1111/mec.14226

Sánchez, A., Salicrú, M., & Ocaña, J. (2007). Statistical methods for the analysis of high-throughput data based on functional profiles derived from the gene ontology. *Journal of Statistical Planning and Inference*, *137*(12), 3975–3989. https://doi.org/10.1016/j.jspi.2007.04.015

Schield, D. R., Adams, R. H., Card, D. C., Perry, B. W., Pasquesi, G. M., Jezkova, T., Portik, D. M., Andrew, A. L., Spencer, C. L., Sanchez, E. E., Fujita, M. K., Mackessy, S. P., & Castoe, T. A. (2017). Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. *Ecology and Evolution*, *7*(11), 3951–3966. https://doi.org/10.1002/ece3.2996

Servedio, M. R., Doorn, G. S., Kopp, M., Frame, A. M., & Nosil, P. (2011). Magic traits in speciation: "Magic" but not rare? *Trends in Ecology & Evolution*, *26*(8), 389–397. https://doi.org/10.1016/j.tree.2011.04.005

Slatkin, M., & Maddison, W. P. (1989). A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, *123*(3), 603–613. https://doi.org/10.1093/genetics/123.3.603

Snyder, R. J., & Dingle, H. (1989). Adaptive, genetically based differences in life history between estuary and freshwater threespine

sticklebacks (*Gasterosteus aculeatus* L.). *Canadian Journal of Zoology*, *67*(10), 2448–2454. https://doi.org/10.1139/z89-345

Song, Y., Endepols, S., Klemann, N., Richter, D., Matuschka, F. R., Shih, C. H., Nachman, M. W., & Kohn, M. H. (2011). Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Current Biology*, *21*(15), 1296–1301. https://doi.org/10.1016/j.cub.2011.06.043

Soria-Carrasco, V., Gompert, Z., Comeault, A. A., Farkas, T. E., Parchman, T. L., Johnston, J. S., Buerkle, C. A., Feder, J. L., Bast, J., Schwander, T., Egan, S. P., Crespi, B. J., & Nosil, P. (2012). Stick insect genomes reveal natural selection's role in parallel speciation. *Science*, *344*, 738–742. https://doi.org/10.5061/dryad.9g182

Stankowski, S., & Streisfeld, M. A. (2015). Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 1–9. https://doi.org/10.1098/rspb.2015.1666

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, *14*(11), 751–764. https://doi.org/10.1038/nrg3483

Stern, D. L., & Orgogozo, V. (2008). The loci of evolution: How predictable is genetic evolution? *Evolution*, *62*(9), 2155–2177. https://doi.org/10.1111/j.1558-5646.2008.00450.x

Stern, D. L., & Orgogozo, V. (2009). Is genetic evolution predictable? *Science*, *323*(5915), 746–751. https://doi.org/10.1126/science.1158997

Sukumaran, J., & Holder, M. T. (2010). DendroPy: A python library for phylogenetic computing. *Bioinformatics*, *26*(12), 1569–1571. https://doi.org/10.1093/bioinformatics/btq228

Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., Wolfner, M. F., & Aquadro, C. F. (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in drosophila. *Proceedings of the National Academy of Sciences*, *98*(13), 7375–7379. https://doi.org/10.1073/pnas.131568198

Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone. *Evolution*, *68*(11), 3066–3081. https://doi.org/10.1111/evo.12510

Turelli, M., & Hoffmann, A. A. (1995). Cytoplasmic incompatibility in *Drosophila simulans*: Dynamics and parameter estimates from natural populations. *Genetics*, *140*, 1319–1338. https://doi.org/10.1093/genetics/140.4.1319

Turelli, M., & Orr, H. A. (2000). Dominance, epistasis and the genetics of postzygotic isolation. *Genetics*, *154*(4), 1663–1679. https://doi.org/10.2307/2410324

Turner, T. L., Hahn, M. W., & Nuzhdin, S. V. (2005). Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, *3*(9), 1572–1578. https://doi.org/10.1371/journal.pbio.0030285

U.S. Fish & Wildlife Service. (2014). Endangered and threatened wildlife and plants; endangered status for the Sierra Nevada yellow-legged frog and the northern distinct population segment of the mountain yellow-legged frog, and threatened status for the Yosemite toad: Final rule. *Federal Register*, *79*, 1–56.

Verster, A. J., Ramani, A. K., McKay, S. J., & Fraser, A. G. (2014). Comparative RNAi screens in *C. elegans* and *C. briggsae* reveal the impact of developmental system drift on gene function. *PLoS Genetics*, *10*(2), e1004077. https://doi.org/10.1371/journal.pgen.1004077

Via, S. (2012). Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*(1587), 451–460. https://doi.org/10.1098/rstb.2011.0260

Vijay, N., Bossu, C. M., Poelstra, J. W., Weissensteiner, M. H., Suh, A., Kryukov, A. P., & Wolf, J. B. W. (2016). Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Communications*, *7*, 1–10. https://doi.org/10.1038/ncomms13195

Weinreich, D. M., Delaney, N. F., DePristo, M. A., & Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, *312*(5770), 111–114. https://doi.org/10.1126/science.1123539

Whitney, K. D., Randell, R. A., & Rieseberg, L. H. (2006). Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *The American Naturalist*, *167*(6), 794–807. https://doi.org/10.1086/504606

Wolf, J. B. W., & Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*, *18*(2), 87–100. https://doi.org/10.1038/nrg.2016.133

Wu, C. I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*(6), 851–865. https://doi.org/10.1046/j.1420-9101.2001.00335.x

Wu, C. I., & Ting, C. T. (2004). Genes and speciation. *Nature Reviews Genetics*, *5*(2), 114–122. https://doi.org/10.1038/nrg1269

Yeaman, S. (2013). Hybridization and the porous genome: Patterns of isolation and introgression in manakins. *Molecular Ecology*, *22*(12), 3195–3197. https://doi.org/10.1111/mec.12314

Yeaman, S. (2015). Local adaptation by alleles of small effect. *The American Naturalist*, *186*(S1), S74–S89. https://doi.org/10.1086/682405

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Maier, P. A., Vandergast, A. G., & Bohonak, A. J. (2024). Yosemite toad (*Anaxyrus canorus*) transcriptome reveals interplay between speciation genes and adaptive introgression. *Molecular Ecology*, *00*, e17317. https://doi.org/10.1111/mec.17317