Title: Genomic insights into variation in thermotolerance between hybridizing swordtail fishes

Authors: Cheyenne Payne[1,2], Stephen Bovio[2,3], Daniel Powell[1,2], Theresa Gunn[1,2], Shreya Banerjee[1,2], Victoria Grant[1,2], Gil Rosenthal[2,3], Molly Schumer[1,2,4]

Affiliations:

[1]Department of Biology, Stanford University, Stanford, CA, USA
[2]Centro de Investigaciones Científicas de las Huastecas "Aguazarca", A.C., Calnali, Hidalgo, México
[3]Department of Biology, Texas A&M University
[4]Hanna H. Gray Fellow, Howard Hughes Medical Institute, Stanford, CA, USA

Correspondence: cypayne@stanford.edu and schumer@stanford.edu

Abstract

Understanding how organisms adapt to changing environments is a core focus of research in evolutionary biology. One common mechanism is adaptive introgression, which has received increasing attention as a potential route to rapid adaptation in populations struggling in the face of ecological change, particularly global climate change. However, hybridization can also result in deleterious genetic interactions that may limit the benefits of adaptive introgression. Here, we used a combination of genome-wide quantitative trait locus mapping and differential gene expression analyses between the swordtail fish species *Xiphophorus malinche* and *X. birchmanni* to study the consequences of hybridization on thermotolerance. While these two species are adapted to different thermal environments, we document a complicated architecture of thermotolerance in hybrids. We identify a region of the genome that contributes to reduced thermotolerance in individuals heterozygous for *X. malinche* and *X. birchmanni* ancestry, as well as widespread misexpression in hybrids of genes that respond to thermal stress in the parental species, particularly in the circadian clock pathway. We also show that a previously mapped hybrid incompatibility between *X. malinche* and *X. birchmanni* contributes to reduced thermotolerance in hybrids. Together, our results highlight the challenges of understanding the impact of hybridization on complex ecological traits and its potential impact on adaptive introgression.

Keywords: thermotolerance, hybridization, swordtail fishes, misexpression, molecular ecology

Introduction

Hybridization, or interbreeding between species, is much more common than previously thought and can have diverse genetic and evolutionary consequences [1]. For example, a large body of work has shown that hybridization can facilitate the movement of adaptive alleles between species, promoting ecological adaptation to novel or changing environments [2–16]. In hybridizing species, gene flow may serve as a mechanism of rapid adaptation [7,8,16,17], since adaptive introgression can occur on a shorter timescale than that required for new adaptive mutations to arise within a species [18,19].

While hybridization has played a role in adaptation on evolutionary timescales [19–21], hybridization is thought to occur more frequently under environmental disturbance and stress [12,13,22–24]. As environmental conditions shift due to climate change, understanding the genetic mechanisms that can facilitate rapid adaptation will be critical in predicting whether vulnerable populations will adapt or collapse [18,25]. This in turn requires characterizing the genetic architecture of ecologically relevant traits that distinguish hybridizing species [26–29].

Although there are many examples of adaptive introgression between species [7–9,18], deleterious effects of hybridization are also well-documented and have been studied for decades [30,31]. Hybridization frequently uncovers negative interactions between mutations that have arisen independently in the genomes of the two parental species. These interactions can result in reduced hybrid viability or fertility [32–34], and their costs may outweigh the potential benefits of hybridization as a source of adaptive alleles [1]. Such interactions are commonly known as Bateson-Dobzhansky-Muller incompatibilities (BDMIs; [35]). While BDMIs were originally envisioned to result from incompatible interactions between proteins encoded by two or more genes, recent work has highlighted the diversity of mechanisms through which BDMIs may arise [36–42]. Recently attention has been paid to regulatory BDMIs, which arise from the coevolution of *cis* and *trans* regulatory elements within species that become mismatched and cause misexpression of target genes in hybrids [1,36,43,44] (here, misexpression is defined as expression of genes in hybrids that is much higher or lower than that observed in either parent species).

BDMIs can impact a range of traits, including those relevant for an individual's survival in their environment [45,46]. In fact, both theory and empirical results suggest that BDMIs may frequently arise from divergent adaptation to the environment [47–49]. In addition to the expectation that hybrids may have reduced ecological fitness due to phenotypic intermediacy or dominance of particular parental traits [45,46,50–54], BDMIs can arise at loci underlying ecological traits. Despite their predicted importance, few ecological BDMIs have been identified to date [46] (see [55] and [46] for examples from *Arabidopsis* and sticklebacks), making it difficult to study the tradeoffs between adaptive introgression of ecological traits and selection on ecological BDMIs in hybrids.

An ecological trait that can be used to address this gap and that is of particular interest for predicting how populations may adapt to global climate change is thermal tolerance [56–58]. Though thermal tolerance can be defined in many ways, as global temperatures warm, a relevant thermal tolerance trait is an organism's upper thermal limit (hereafter referred to as

"thermotolerance") [59,60]. Little work to date has explored whether loci that control variation in thermotolerance introgress between hybridizing animal populations or how effective such introgression is as a mechanism of thermal adaptation (though suggestive results have been reported in some species; box turtles [61]; wasps [12]; copepods [62,63]). The deficit in empirical work in this area is likely due to the paired difficulties of mapping the genetic basis of complex traits like thermotolerance and studying them in natural hybrid populations.

Here, we take advantage of a system where natural hybridization is ongoing between species that vary in thermotolerance. Two sister species of swordtail fishes, *Xiphophorus birchmanni* and *X. malinche*, are endemic to rivers in eastern México [64], and their distributions are determined in part by their thermal habitats [65]. *X. malinche* lives in cooler (7-25°C) streams at high elevations, while the more heat tolerant *X. birchmanni* lives downstream in the warmer lowlands (15-35°C; [64]). These species are sympatric in regions where their temperature ranges overlap. Recently, pollution has interfered with species-specific olfactory communication, causing breakdown of mating barriers [66,67]. As a result, natural hybrid zones have formed, with clinal ancestry patterns that mirror thermal gradients, where there is low *X. birchmanni* ancestry and low thermotolerance in the highlands and high *X. birchmanni* ancestry and high thermotolerance in the lowlands [64,65]. Though phenotypic plasticity contributes to this differential tolerance (as shown in [65]), we show here that variation in innate thermotolerance between species is in part genetic. Leveraging this finding, we combine thermotolerance assays with classic quantitative trait locus (QTL) mapping, gene expression analysis, and analysis of ancestry in natural hybrid populations to explore the evolution of thermotolerance in this system. Unexpectedly, we find that individuals that are heterozygous for ancestry in one genomic region have reduced critical thermal maxima, and F_1 hybrids have widespread misexpression of core regulatory genes of the circadian clock, which appear to be associated with proper thermal regulation. Additionally, we uncover a relationship between reduced thermotolerance and a previously mapped hybrid incompatibility.

Methods

Comparison of CT_{max} between X. malinche, X. birchmanni, F_1s, and F_2s and measurement of CT_{max} for QTL mapping

One ecologically-relevant measure of upper thermotolerance in ectotherms is the critical thermal maximum, or CT_{max} [59]. Specifically, the CT_{max} of a fish is the highest temperature it can withstand before it experiences loss of equilibrium and is unable to maintain its balance [59,60]. We tested CT_{max} for *Xiphophorus malinche*, *X. birchmanni,* and F_1 and F_2 hybrids between the two species reared in a common garden environment.

We simultaneously reared *X. malinche* fry born to wild-caught mothers from the Chicayotla locality on the Río Xontla (1003 meters elevation; 20°55'27.24"N 98°34'34.50W), *X. birchmanni* fry from wild-caught mothers from the Coacuilco locality on the Río Coacuilco (320 meters elevation; 21°5'50.85 N, 98°35'19.46 W), and F_1 and F_2 fry generated from these parent populations (Fig. 1A). Specifically, F_1 fry were generated by crossing *X. malinche* (Chicayotla) females to *X. birchmanni* (Coacuilco) males, and F_2 fry were generated by intercrossing previously produced F_1s (Fig. 1B). We note that due to the crossing design, all artificial hybrids in this study harbor *X. malinche* mitochondria; crosses in the reverse direction are largely unsuccessful. All fish were crossed and raised in 2,000 L semi-natural mesocosms at the CICHAZ field station in Calnali, Hidalgo, México. Individuals for all four groups were born between 16 May and 24 May 2016, at which time offspring from each group were randomly assigned to one of three replicate 2,000 L semi-natural mesocosms for a total of 12 tanks (three per class, n = 34 per tank).

To measure variation in CT_{max} between *X. malinche, X. birchmanni*, F_1s, and F_2s, CT_{max} trials were performed in February 2018 using methods similar to Culumber et al [65]. Trials followed procedures approved in Texas A&M IACUC protocol #117419. Briefly, the test fish (eight per trial, mix of males and females from one group), an air bubbler, a standard glass thermometer, and a HOBO temperature logger (Onset) were placed in an enamel container holding 4 L of water at ambient temperature (16.1 ± 0.2°C). The enamel container was nested in a larger container of water which was suspended above a hot plate. Water was heated at a standardized ramp-up rate of 0.3 °C/min until the fish lost equilibrium (following Becker & Genoway [60]). The time and temperature of initial loss of equilibrium (i.e. the first time balance is lost) for each fish was recorded, and the fish was immediately placed in an ambient temperature recovery tank. Because the data departed from the assumption of normality, we used a Mann-Whitney Wilcoxon test to evaluate the effect of genotype on CT_{max} (Table S1).

We repeated these procedures for a larger mapping population of 152 *X. malinche*-*X. birchmanni* artificial hybrids. Due to the difficulty of raising sufficient numbers of individuals in common garden conditions, our mapping population included individuals ranging from F_2-F_4 generations, initially generated from F_1 intercrosses. For each individual, we collected a fin clip from each fish at the end of the CT_{max} trial to perform QTL mapping. In addition to CT_{max} time and temperature, we recorded metadata for each fish to account for potential covariates in mapping. We found that one of the strongest covariates with CT_{max} was rearing tank number (which also corresponded to trial number). Therefore, we combined these covariates into a single variable that we refer to as tank throughout the manuscript (listed as ‘site.tank’ in data files).

DNA extraction and library preparation

Fin clips were added to 96 well plates and DNA was extracted using the Agencourt DNAdvance bead-based kit. The protocol followed that specified by the manufacturer except that we used half-reactions. We quantified extracted DNA using a TECAN microplate reader. After diluting DNA to 2.5 ng/ul, we prepared tagmentation-based libraries for low-coverage whole genome sequencing. DNA was enzymatically sheared using the Illumina Tagment DNA TDE1 Enzyme and Buffer Kits by incubating DNA, buffer and enzyme at 55°C for 5 minutes. Fragmented DNA was amplified in a dual-indexed PCR reaction for 12 cycles and PCR-products were pooled and bead purified with 18% SPRI magnetic beads. Purified libraries were quantified using a Qubit fluorometer and library size distribution was evaluated using an Agilent 4200 Tapestation.

Artificial hybrid QTL mapping sample sequencing and genotyping

Low-coverage whole genome sequence data was collected from these libraries on an Illumina 4000 machine using 150 bp paired-end reads (~0.1-0.3X per basepair coverage). Using the program *ancestryinfer* [68], reads were mapped to both the *X. birchmanni* and *X. malinche* genomes with BWA-MEM [69], and those that showed evidence of mapping bias or that did not map uniquely were discarded. Reads matching each parental allele at ancestry-informative sites were counted from samtools mpileup files [70], and informative sites were thinned to one per read to minimize errors due to mismapping. This data was input into AncestryHMM [71], a hidden Markov model (HMM) based local ancestry inference program that relies on read counts at ancestry informative sites and transition probabilities to infer posterior probabilities for ancestry states (in our case: homozygous *X. birchmanni*, heterozygous, or homozygous *X. malinche*). Past work has shown that this low-coverage whole genome sequencing approach has excellent accuracy for early generation *X. malinche x X. birchmanni* hybrids [68,72]. This analysis yielded posterior probabilities for each ancestry state at ~700,000 ancestry-informative sites across the genome (approximately one per kb).

Because it was convenient for downstream analyses, we converted posterior probabilities at each ancestry informative site to hard genotype calls. For each sample, markers with greater than 0.9 posterior probability for any ancestry state were assigned to that state; markers with less than 0.9 posterior probability for any ancestry state were converted to NAs. Homozygous *X. birchmanni*, heterozygous, and homozygous *X. malinche* ancestry calls were assigned genotypes of 0, 1, and 2 respectively.

CT_{max} QTL mapping analysis

To identify regions of the genome that are associated with variation in thermotolerance, we used a QTL mapping approach. We performed QTL mapping with R/qtl [73] to scan for an association between genotypes at ancestry-informative markers across the genome and the CT_{max} phenotype. For computational efficiency, markers were thinned to retain at most one marker per 20 kb. Ancestry linkage disequilibrium decays over several megabases in early generation hybrids [74]; thus, we do not expect to sacrifice any power to map QTL by performing this thinning. The thinning step resulted in 30,244 ancestry informative markers retained throughout the genome, or ~45 per Mb.

Data were converted to the R/qtl input format using custom scripts (https://github.com/Schumerlab/thermotolerance). Input files for analysis with R/qtl included

CT_{max}, covariate data (e.g. tank), and genotype data for all 152 individuals. Markers with fewer than 80% of individuals genotyped (i.e. less than 120 out of 152) were filtered. Several individuals had high levels of missing data (>25% of markers with an NA ancestry state) and these individuals were removed. After filtering, 144 individuals and 29,652 markers were retained, with ~95% of individuals genotyped at any given marker. Next markers were evaluated for segregation distortion at a Bonferroni corrected p-value < 0.05 using R/qtl's internal commands, and 610 markers on chromosome 13 were dropped, resulting in 29,042 markers for the QTL scan. Recombination frequency and genotype probabilities were calculated using the `est.rf` and `calc.genoprob` functions, respectively.

To select an appropriate model for mapping in R/qtl, we used the R step function to calculate AIC for models incorporating a suite of possible covariates, including the tank variable (tank), hybrid index (the proportion of the genome derived from the *X. malinche* parent), genome-wide ancestry heterozygosity, and sex (e.g. CT_{max} ~ hybrid_index + heterozygosity + tank + sex). We retained all tank variables with a significant effect on CT_{max} (17 total) and used a method called 'one-hot encoding' to recode the tank variable so that the tank variable would be treated categorically by R/qtl; other covariates were not retained in the step function analysis. Even though hybrid index was not retained, we included it in our final mapping model since past work has suggested that failing to include ancestry as a covariate can result in artifacts in QTL analysis [72].

A genome-wide scan with a single-QTL model was performed with the `scanone` function, using the Haley-Knott regression method [75] and the tank and hybrid index covariates, as described above. The 5% and 10% false discovery rate thresholds were estimated with 1,000 permutations (LOD thresholds of 4.72 and 4.33 respectively), where CT_{max} phenotypes were shuffled onto genotypes and a QTL scan conducted 1,000 times to create a null distribution of associations expected by chance. To search for potential interacting QTL, we performed a second scan using the same method, but added genotypes at the chromosome 22 QTL peak as an interaction term in the model (significant thresholds at the 5% and 10% FDR level for the interaction analysis were 9.63 and 8.96, respectively).

Estimating the effect size of detected QTL

We identified one QTL on chromosome 22 and one putative interacting QTL on chromosome 15 that were significantly associated with variation in CT_{max} after controlling for other covariates at the 10% false discovery threshold (see Results). We used two methods to obtain estimates for the effect sizes of these QTL (i.e. the percentage of the variation in CT_{max} explained by each QTL and their interaction). First, we used the drop-one-term analysis from fitting a multiple QTL model with the R/qtl function `fitqtl`, to estimate the effect sizes of the chromosome 22 and 15 QTL on CT_{max}, as well as to estimate the effect size of their interaction. Because effect size estimates are often inflated in QTL studies with low statistical power [76], we also performed simulations to explore the range of possible effect sizes consistent with our empirical results for the main effect QTL on chromosome 22. Methods and results for those simulations are reported in Supporting Information 1.

Multiple tissue thermal stress RNAseq experiment, library preparation, and sequencing

To compare expression of genes across the genome that respond to high temperature in the two parental species and their hybrids, we used an RNAseq-based experimental approach. *X. birchmanni* and *X. malinche* individuals born to wild mothers (collected at the Coacuilco and Chicayotla populations respectively [64]) were raised at 22.5°C (14h light:10h dark cycle). A separate group of *X. malinche* females were crossed with *X. birchmanni* males to generate F_1 hybrids. All individuals were raised in the same lab environment to adulthood before experiments began. Though we cannot discount the potential impact of maternal effects on expression response, all mothers were reared under the same environmental conditions.

For thermal stress experiments, three male individuals from each group were kept at a control temperature of 22.5°C for the duration of the experiment, and three male individuals from each group were subjected to a thermal stress treatment. Males in the treatment trials experienced a temperature ramp-up of 0.3°C/min from 22.5°C to 33.5°C (~30 min duration). Control and treatment trials were run simultaneously between 11:00 AM and 1:00 PM. An air bubbler was used to maintain dissolved oxygen saturation in tank water for the duration of the experiment. Fish from both control and treatment tanks were anaesthetized with Tricaine mesylate diluted in tank water immediately after treatment tanks reached 33.5°C, and brain and liver tissues were dissected and placed in RNAlater. These samples were stored at 4°C for 24 hours and subsequently at -20°C. mRNA was extracted for a total of 36 brain and liver samples with a Qiagen RNeasy MiniPrep Kit. One *X. birchmanni* brain from the 22.5°C treatment and one *X. malinche* brain from the 33.5°C treatment yielded insufficient mRNA for RNAseq library preparation; therefore, these samples were not sequenced. RNAseq libraries were prepared using a KAPA mRNA HyperPrep Kit, pooled, and sequenced on three Illumina HiSeq4000 lanes. To control for batch effects, extraction, library prep, and sequencing batches were designed to include one individual from each biological group. We sequenced three biological replicates per experimental group and collected >30 million 150 bp paired-end reads per sample (Table S2).

Differential gene expression analysis

Genes that are differentially expressed in response to thermal treatment, especially those that respond differently in *X. birchmanni* and *X. malinche,* are candidate genes that may contribute to variation in thermotolerance between species. For differential gene expression comparisons, we aligned RNAseq reads to reference transcriptomes inferred from high-quality *X. malinche* and *X. birchmanni* genome assemblies. For GO and KEGG enrichment analyses, we aligned reads to developed "pseudoreference" transcriptomes for these two species (from Schumer et al [77]) that were based on the genome assembly of the southern platyfish *X. maculatus* [78]. We used these references because *X. maculatus* is widely used as a model in melanoma research, and as a result has a well-annotated genome [78,79] with GO and KEGG pathways associated with each Ensembl gene ID. To reduce mapping bias in differential expression analysis we used a version of these references with within-species polymorphisms masked [77].

Before aligning reads, the program cutadapt and the FastQC wrapper tool Trim Galore! were used to trim Illumina adapter sequences and low-quality bases (Phred score < 30) from reads. All trimmed reads are available under NCBI BioProject PRJNA746324. One F_1 liver sample from the 22.5°C ambient temperature treatment group was removed from downstream

analyses due to unusually low read count (<1500 reads). Reads were then pseudoaligned to the *X. birchmanni* reference transcriptome with *kallisto* [80] and raw transcript counts were imported into the R package DESeq2 [81] for differential gene expression analysis. Genes with zero counts for all samples, extreme outliers (using a Cook's distance cutoff of 0.99), or low mean normalized counts (i.e. genes with counts below an optimized threshold through an internal filtering step in DESeq2) were removed from analysis. This resulted in an analysis of 24,174 genes for both the brain and liver datasets.

To analyze differential expression of these genes, we used a design formula that included sequencing batch, genotype (*X. birchmanni*, *X. malinche*, or F_1), and temperature treatment. Using DESeq2, we normalized gene counts by library size, estimated within-experimental group dispersion, fit a negative binomial generalized linear model, and tested significance with a Wald test. Following these steps, shrunken log-fold changes were calculated using an adaptive shrinkage estimator with a fitted mixture of normal distributions as a prior, derived from the 'ashr' package [82]. Genes were considered to be significantly differentially expressed between groups and treatments at an FDR-adjusted p-value < 0.1. To check for potential bias as a function of the reference sequence used in the pseudoalignment step, we repeated the above steps using the *X. malinche* reference transcriptome. Reassuringly, qualitatively similar results were obtained from this analysis (Supporting Information 2, Fig. S2).

Co-expression network analysis with WGCNA

To identify clusters of interacting genes that respond to temperature treatment, we used the R package WGCNA to evaluate patterns of co-expression in the RNAseq data [83]. Weighted co-expression network analysis clusters genes with highly correlated expression patterns across samples into groups called modules. The expression patterns of modules are summarized by their 'module eigengenes,' defined as PC1 of the expression profiles of genes in the module, which can then be used to test for correlations between gene modules and traits or treatments of interest. This unsupervised approach is particularly powerful for identifying biological pathways whose expression strongly correlates with a specific treatment. In this case, we were most interested in modules that correlated with temperature treatment and with genotype.

WGCNA analysis was performed separately for sets of samples of each tissue type, using raw gene counts obtained from pseudoalignment to the *X. birchmanni* pseudoreference transcriptome (as described in the *Differential gene expression analysis* methods). We used the DESeq2 `varianceStabilizingTransformation` function to normalize raw gene counts by library size and size factors (the median ratio of the geometric mean of a gene over all samples) so that samples had comparable variances. Genes were filtered as described in the previous section, and additionally all genes with zero counts for one or more samples were dropped (out of 19,176 genes, 262 genes from the brain and 905 from the liver were dropped in this step).

As recommended by the WGCNA documentation, we selected a soft-thresholding power to transform the network into a more scale-free topology, which has been shown to better approximate biologically-relevant gene networks [84]. This step is intended to minimize the effect of noise in subsequent clustering steps and avoid using arbitrary thresholds for cluster construction. For each tissue dataset, the soft-thresholding power parameter was chosen by calculating the scale-free topology fit index for a range of powers (1 to 30) and selecting the

asymptote of soft-thresholding power for downstream analysis (here, using the recommended thresholds for scale-free topology fitting index $R^2 > 0.8$ and mean connectivity < 100). This resulted in the selection of soft-thresholding powers of 7 for the brain and 12 for the liver tissue datasets. Using these soft-thresholding values, we constructed single-block unsigned networks using WGCNA's `blockwiseModules` function (see Appendix 1). We used a minimum module size of 20 and an unsigned topological overlap matrix to create a network that clusters genes by strength of co-expression, regardless of whether the correlation in expression is positive or negative.

After co-expression modules were identified using this approach, we sought to determine whether variation in any of these modules correlated with variation in traits of interest. As such, we looked for correlations between the module eigengene and genotype (*X. malinche, X. birchmanni, X. malinche-X. birchmanni* F_1), temperature treatment (22.5°C and 33.5°C), or both. Correlations between traits and modules were calculated using the WGCNA `corPvalueStudent` function, and modules that correlated with genotype, temperature treatment, or both at Student asymptotic p-value < 0.05 were selected for further analysis (31 out of 54 for brain, 16 out of 50 for liver).

GO and KEGG pathway enrichment of differentially expressed genes between temperature treatments

To explore which biological pathways are most affected by temperature treatment, we performed KEGG pathway and Gene Ontology (GO) term enrichment analysis of differentially expressed genes identified using results from the analyses of the brain and liver RNAseq data described above. We asked if there were enriched KEGG pathways in the set of genes that were significantly differentially expressed between *X. birchmanni*, *X. malinche*, and F_1s at each temperature treatment (FDR adjusted p-value < 0.1). To do so, Gene IDs were mapped to Entrez IDs using the *X. maculatus* Ensembl database (version 99), and enriched KEGG pathways for each dataset were generated with the `kegg.gsets` function from the R package GAGE [85].

For GO enrichment, the R packages biomaRt [86] and GOstats [87] were used to extract *X. maculatus* Ensembl IDs and generate a GO gene universe of all genes analyzed with DESeq2, as described above (19,143 genes for brain, 19,176 for liver). We used a hypergeometric test implemented in the R package GSEABase [88] to identify overrepresented GO terms in the set of significantly differentially expressed genes between temperature treatments. We also performed GO analysis of genes in gene modules that correlated with temperature treatment in WGCNA analyses (12 modules for brain, 2 for liver). For both sets of GO analyses, we focused on significantly enriched categories (hypergeometric test p-value < 0.05) where greater than one gene was observed in our focal dataset.

Ancestry of QTL and circadian clock genes in natural populations

We analyzed data from naturally occurring *X. malinche-X. birchmanni* hybrid populations to evaluate evidence for shifts in ancestry at genes under the chromosome 22 QTL, and clock genes that show misexpression in hybrids (see Results). These data from natural hybrid populations have been published in previous studies [72,77], so we only describe it briefly here. We analyzed data collected from the Tlatemaco (n=85) and Acuapa (n=97) hybrid populations in 2017 and 2018, respectively [72,77] (Fig. 1A). We previously collected ~1X

whole genome sequence data from individual hybrids collected from these populations and followed the local ancestry inference approach described above except that we used population-specific priors for admixture proportion and time since initial admixture. This approach resulted in estimates of the posterior probability for ancestry state at ~613-629 million ancestry informative sites throughout the genome in the two populations. Using ancestry informative sites that fell within annotated coding regions, we generated estimates of the average ancestry per gene in both natural hybrid populations. We compared ancestry at genes of interest to the genomic background of each population.

Correlation of caudal spot phenotype and CT_{max} in natural hybrids

Hybrids between *X. malinche* and *X. birchmanni* harbor a number of extreme traits not seen in either of the parental species. One of these is a hybrid incompatibility that causes melanoma in some individuals, originating from a melanocyte spotting pattern on the caudal fin [89]. To evaluate any relationship between this spotted caudal phenotype and CT_{max}, we measured CT_{max} using the methods described above (see *Comparison of CT_{max} between X. malinche, X. birchmanni, F_1s, and F_2s and measurement of CT_{max} for QTL mapping*) for 123 lab-raised natural hybrids born from wild-caught mothers from the Chahuaco Falls population. These 123 natural hybrids were reared to adulthood under common conditions in the lab. Individuals were classified as having one of the following caudal spot phenotypes: no spot, normal spotted caudal, expanded spotted caudal, and 3D melanoma. Past histological work has indicated that individuals with the expanded spotted caudal phenotype have early-stage melanoma [89]. Individuals were assigned a 3D melanoma phenotype if they had melanoma that had completely overtaken the caudal fin (i.e. completely melanized and/or degrading) and/or that was growing laterally off the fin. We used a linear model to test for a correlation between CT_{max} and caudal spot phenotype.

Results

Evidence for a genetic basis for variation in thermotolerance

Given that *X. malinche* and *X. birchmanni* live in different thermal environments (Fig. 1C), we predicted that these species may have adapted to their respective thermal ranges. To determine whether there was a genetic basis for variation in CT_{max} between *X. malinche* and *X. birchmanni*, we reared *X. malinche, X. birchmanni*, F_1, and F_2 hybrid fish in a common garden, and measured their CT_{max} (see Methods). We found that *X. birchmanni* have significantly higher CT_{max} than *X. malinche* (p-value $< 10^{-6}$; see Table S1), and that F_1 and F_2 hybrids exhibit intermediate CT_{max} (Fig. 2A). Though we know that CT_{max} is partially environmentally mediated in this system [65], this result shows that variation in CT_{max} between these species is also partly attributable to genetic differences.

QTL mapping of loci involved in thermotolerance

Given these results, we proceeded to perform QTL mapping to evaluate associations between CT_{max} and ancestry in *X. malinche-X. birchmanni* artificial hybrids raised under common conditions (see Methods). We detected a single QTL associated with CT_{max} at a 10% false discovery rate threshold (Fig. 2B). The 1.5 LOD interval of this QTL spans ~2.5 Mbs and contains 45 genes.

Surprisingly, further examination revealed that the QTL we detected was not associated with species-level differences in CT_{max}. Instead, heterozygous ancestry in this region was associated with an average reduction in CT_{max} of 0.3°C (Fig. 2C). We estimate this QTL to have a moderate effect on the overall variation in CT_{max} in artificial hybrids, explaining approximately 6.9% of the variation (see Methods, Supporting Information 1, and Fig. S1 for simulations evaluating effect size inflation).

The relationship between genotype and phenotype observed at the chromosome 22 QTL is consistent with underdominance. Individuals with either homozygous genotype exhibit approximately the same average CT_{max} whereas individuals heterozygous for ancestry have a significantly reduced CT_{max} on average (Table S5; Fig. 2C). Though distinguishing whether this signal is caused by true underdominance or pseudo-underdominance (i.e. two or more linked loci with dominance and opposing effects in homozygotes; Fig. S3) is not feasible with our data, we discuss this possibility in more detail in Supporting Information 4.

Exploring candidates in the QTL region

There are several possible explanations for the observed signal of underdominance at the chromosome 22 CT_{max} QTL. Because chromosomal inversions are a common genetic cause of underdominance [90–93], we confirmed that there are no inversions under this QTL by aligning *X. malinche* and *X. birchmanni* PacBio assemblies (Fig. S4; Supporting Information 4).

Next, we investigated genes that fell within the QTL region. The 1.5-LOD interval associated with the QTL spans from ~8.6 Mb to 10.1 Mb, overlapping with 45 genes. We investigated functional annotations and patterns of expression of genes in this region, as well as amino acid differences between species (see Supporting Information 3). Because thermotolerance is a complex trait that is impacted by many biological pathways, narrowing

down causal loci under the QTL based on their annotations is not straightforward. We highlight a handful of candidate genes in Table S3 but focus primarily on candidates with notable expression patterns in this section.

Given the CT_{max} phenotypes observed in individuals heterozygous for ancestry at this QTL, we were particularly interested in comparing gene expression patterns in F_1 hybrids to the parental species. Because heterozygotes have reduced CT_{max} on average, we might expect genes controlled by a causal locus in *cis* to be misexpressed in F_1 hybrids (which are heterozygous for ancestry across the whole genome). We evaluated expression patterns in the brain and liver and identified genes that were misexpressed in F_1 hybrids compared to *X. birchmanni* and *X. malinche*, defined here as significantly higher or lower than either of the typical parental expression ranges, in either temperature condition. Of the genes in the QTL interval that were significantly differentially expressed between *X. malinche* and *X. birchmanni* for at least one of the temperature treatments (17 in the brain, 7 in the liver), most mirrored expression levels of one of the parental species or had intermediate expression in F_1s (Table S3). However, four genes in this interval (*p4ha1*, *ndnf*, *tnfaip3,* and *infgr1l*) were misexpressed in F_1s in at least one tissue and temperature condition (Fig. 2D, S8). We also identified genes that responded differently to thermal stress in F_1s compared to parentals (see brain and liver expression results in Tables S6 and S7, respectively). For example, the zinc-finger protein *zfp62* exhibited an exaggerated response to high temperature compared to parental expression responses and the spliceosome subunit *sf3b5* was significantly downregulated at high temperatures in F_1s whereas parental expression remained constant across temperatures (Fig. S8).

Detection of a possible interacting QTL on chromosome 15

Another potential cause of a signal of underdominance at the chromosome 22 QTL could be interactions with other regions of the genome. In particular, in the literature on the evolution of gene regulation, a breakdown in interactions between paired *cis*- and *trans*-acting regulatory elements can explain aberrant expression patterns in F_1 hybrids [36,43,94]. Thus, we performed a second QTL scan, including genotypes at the chromosome 22 QTL peak as an interaction term. Based on this analysis, we recovered a second QTL at the permuted 10% false discovery threshold, spanning ~2.1 Mb on chromosome 15 (Fig. 2E).

While we are cautious of overinterpreting this result given low power in our study, we discuss it briefly here. Intriguingly, we find that artificial hybrids heterozygous at *both* the chromosome 22 and 15 QTL have reduced CT_{max} on average (-0.4°C), but that hybrids heterozygous at the chromosome 22 QTL and homozygous *X. birchmanni* at the chromosome 15 QTL have elevated CT_{max} on average (+0.5°C over hybrids homozygous for the *X. malinche* allele; Table S5; Fig. 2F, S5). We estimate that the combined additive and interaction effects of the chromosome 22 and 15 QTL could explain ~14.8% of the total variation in CT_{max} in the artificial hybrids, although this number is likely an overestimate of their true effect size (see [76] and Supporting Information 1).

We explored evidence for known genetic interactions between genes under the chromosome 22 and chromosome 15 QTLs. The zinc finger protein gene *zbtb18* and the adjacent serine/threonine-protein kinase *akt3*, and the heterogenous nuclear RNA binding protein *hnrnpu*, which fall directly under the chromosome 22 and 15 peaks respectively, are known to interact during neurodevelopment [95]. We discuss what is known about their interactions and other

interacting genes in these regions in more detail in Supporting Information 5 and summarize all genes under the chromosome 15 QTL in Table S4.

Gene expression profiles differ between species and thermal treatment

To broadly survey changes in gene expression between parental species and their F_1 hybrids in response to high temperature, we generated RNAseq data for brain and liver tissue from fish exposed to ambient and high temperature conditions (see Methods). We sampled the brain and liver to survey two tissues that play different roles in organismal homeostasis – energy consumption and detection of and response to environmental changes by the brain, and energy metabolism by the liver.

In addition to using these data to evaluate expression patterns of genes under the QTL region, we analyzed it in a genome-wide context. As expected, we found broad differences in expression between tissues and species (Fig. S6), as well as strong responses to high temperature (Fig. 3A, S7). The vast majority of the variation in expression in our dataset is explained by tissue type (83.5%; Fig. S6); therefore, the tissue datasets were analyzed separately. Of the remaining variation, genotype explained ~23% of the variation in expression between samples and temperature treatment explained ~11% (Fig. 3A). A large number of genes were significantly differentially expressed between *X. birchmanni* and *X. malinche* (FDR adjusted p-value < 0.1) across temperatures in both tissues (brain: 3,357 and 3,121 genes; liver: 2,318 and 1,508 genes at 22.5°C and 33.5°C, respectively). Interestingly, while the number of genes for which expression changed in response to high temperature in *X. birchmanni* and *X. malinche* was similar (brain: 882 and 979 genes; liver: 113 and 38 genes, respectively), ~2.5-3.5x more genes were responsive to high temperature in F_1 brain (2,318) and liver (408) tissues compared to those of parentals. We found that ~3% and ~0.5% of all genes across the genome were misexpressed under at least one of the two tested thermal contexts in F_1 hybrid brains and livers, respectively. Moreover, ~9% and ~3% of the genes that responded to temperature in *X. birchmanni* and/or *X. malinche* were misexpressed under at least one of the thermal contexts in F_1 brains and livers, respectively. Overall, we found that more genes were misexpressed (low or high) in F_1 hybrids at 22.5°C than at 33.5°C in the brain (521 versus 187, respectively), whereas the liver exhibited the opposite pattern (10 versus 96, respectively); however, we are cautious of potential technical factors that could influence these patterns, such as differences in variance between groups at high temperature. Only a handful of genes were misexpressed under both thermal contexts. See Tables S6-7 for complete results for each group, and Table S8 for a summary of F_1 expression patterns.

We report functional categories with enriched expression responses (p-value < 0.05) to thermal stress, including general thermal stimulus and immune response pathways, in Table S9. In the brain, 84 GO terms are enriched in response to temperature in both *X. malinche* and *X. birchmanni*. Notably, response to temperature stimulus and circadian rhythm categories were commonly enriched for both species. Additionally, 77 terms are enriched only in *X. birchmanni*, and 70 terms are enriched just for *X. malinche*. Among those terms that were enriched only in *X. birchmanni* under thermal stress were autophagy and disassembly of mitochondria, negative regulation of biosynthesis and gene expression, and endogenous stimulus response pathways.

KEGG analysis recovered only one biological pathway commonly enriched in the set of genes that were significantly differentially expressed between temperature treatments in *X. malinche* and *X. birchmanni* brains: protein processing in the endoplasmic reticulum (xma04141;

FDR adjusted p-value < 0.1; Table S10). This result may be attributable to the fact that the endoplasmic reticulum plays a key role in the unfolded protein response, which is activated by thermal stress and is key for maintaining homeostasis during stress [96]. Intriguingly, one transcriptional activator of the unfolded protein response, *xbp1*, is significantly upregulated in both *X. malinche* and *X. birchmanni*, but not in F_1 hybrid brains (Fig. S8).

Strikingly, 20 KEGG pathways were significantly enriched in F_1 brains in response to high temperature. Among these enriched pathways were protein processing in the endoplasmic reticulum and signaling pathways (see Table S10 for full list) that induce the transcriptional regulator of the innate immune response, *nfkb1* [97–100]. Interestingly, one of the potential sets of interactors under the chromosome 22 and chromosome 15 QTL are inhibitors of *nfkb1* expression (see Supporting Information 5 for more information). No significantly enriched KEGG pathways were recovered from the set of genes differentially expressed in response to temperature in the liver across groups.

Co-expression network analysis reveals misexpression in F_1 clock genes

We used the co-expression network analysis software WGCNA [83] to identify clusters of co-expressed genes in our RNAseq datasets (Fig. S10, S11). We performed this analysis separately for the two tissue types. In total, 54 and 50 gene co-expression modules were recovered from the brain and liver RNAseq data, 12 and 2 of which were significantly correlated with temperature treatment respectively (p-value < 0.05; Table S11; Fig. 3B, S9). Additionally, four of the 12 brain temperature-correlated modules were significantly correlated with at least one genotype (see Supporting Information 8).

Notably, one temperature-correlated module was shared between tissue types, suggesting that it may represent a cluster of genes globally involved in the thermal stress response. This module is enriched in genes involved in the circadian rhythm and circadian regulation. This finding is notable since circadian clock pathways are impacted by temperature and play a role in thermoregulation and thermal stress response across taxonomic groups [101–104].

Strikingly, several of the circadian clock genes in this shared temperature-associated module are misexpressed in F_1 hybrids, particularly in data collected from the brain in the high temperature treatment (Supporting Information 6-7). The number of misexpressed genes in this module greatly exceeds the number expected by chance (based on permutations, Table S12, Fig. 3D, Supporting Information 7). This suggests that genes in these circadian clock pathways may be commonly misregulated under thermal stress in *X. malinche*-*X. birchmanni* hybrids. Specifically, we find that most of the clock genes in this module are strongly up- or down-regulated in *X. malinche* and *X. birchmanni* brains and livers in response to high temperature. In contrast, at ambient temperature, F_1 clock gene expression tends to be similar to parental expression, but at high temperatures these genes are misexpressed in F_1 brains compared to parental brains (Table S13; Fig. 3C-D). These results hint at a failure to regulate expression of these genes in hybrids. Specifically, much of the misexpression observed in these genes is attributable to the fact that while their expression in parental brains is strongly responsive to the thermal treatment, F_1 expression does not change substantially between temperature treatments. Additionally, some of these genes, such as the transcription factors *dbpb* and *bhlhe41* shown in Fig. 3C and *nr1d2a* and *cipcb* shown in Fig. S8, show patterns of F_1 misregulation under both thermal contexts. We discuss these patterns in more detail in Supporting Information 9.

Ancestry patterns in natural hybrid populations at regions implicated in thermotolerance

Hybrid populations between *X. birchmanni* and *X. malinche* occur across a range of elevations in different river systems [64] and experience different average temperatures [65]. To determine whether there is evidence of selection against a particular ancestry state in natural hybrid populations at the chromosome 22 and 15 CT_{max} QTL and at the clock genes discussed above, we focused on two hybrid populations that occur at elevations closer to those typical of *X. birchmanni* populations (Fig. 1A) and thus experience higher temperatures on average. These populations are the Acuapa and Tlatemaco populations (elevations of 476 and 480 meters, respectively). Notably, while individuals from the Acuapa population derive the majority of their genomes from *X. birchmanni* (~75%; [72]), the parental species with higher thermotolerance, individuals from the Tlatemaco population derive the majority of their genomes from *X. malinche* (~72%; [67]). Thus, regions that have unusually high *X. birchmanni* ancestry in both populations compared with the genome-wide background and that overlap with mapping or expression results may be of particular interest as candidates for loci underlying variation in thermotolerance phenotypes.

Focusing first on the QTL regions, we found that four genes under the chromosome 22 QTL (*akt3, sdccag8,* and *olig3*) and a handful of genes under the chromosome 15 QTL (including *nrxn3a* and one *nrxn3b* isoform) have higher than average *X. birchmanni* ancestry in both low-elevation hybrid populations (Fig. 4A). This shared high *X. birchmanni* ancestry in both populations deviates significantly from expectation (based on permutations, Table S14).

We next evaluated ancestry in both hybrid populations among genes in the circadian clock gene expression module. Notably, two clock genes in this module that are misexpressed in F_1 hybrids, *nr1d2b* and *arntl1a*, have unusually high *X. birchmanni* ancestry in both populations (>89% in both, permuted p-value<0.01; Fig. 4A, Table S14). Interestingly, *nr1d2b* directly represses *arntl1a* expression [105]. Clock genes with strong skews in ancestry in both natural hybrid populations may be adaptive in lower elevation habitats, as this level of ancestry sharing across the two populations is unexpected by chance (see Supporting Information 10). Together, these analyses highlight regions that may be under selection due to their impacts on thermotolerance in natural hybrid populations.

Other phenotypes associated with thermotolerance in hybrids

Given the overall pattern of reduced thermotolerance associated with heterozygous ancestry at the chromosome 22 QTL and aberrant expression of many thermally responsive genes in F_1 hybrids, we wanted to further investigate other possible phenotypic drivers of reduced thermotolerance in hybrids. One trait that is present in hybrids but not in parental individuals of either species is a hybrid incompatibility involving a pigmentation phenotype called the "spotted caudal". While the spotted caudal is a benign melanocyte pigmentation pattern in *X. birchmanni*, it can transform into a malignant melanoma in hybrids (Fig. 4B) with certain genotype combinations ([89]; those with *X. birchmanni* ancestry at the *xmrk* gene and *X. malinche* ancestry at *cd97*).

We found that the spotted caudal phenotype was significantly correlated with CT_{max} in lab-reared offspring from wild mothers collected from a natural *X. malinche-X. birchmanni* hybrid population from Chahuaco Falls (Fig. 1A). In particular, hybrid individuals with an

expanded spot typical of early melanoma as well as hybrids with a more advanced 3D melanoma phenotype had significantly reduced CT_{max} compared to those with a benign spot or no spot (Fig. 4B). This poor performance in hybrids with incompatible genotype combinations highlights one potential mechanism through which underdominance in traits such as thermotolerance could occur. We discuss the implications of this result in more detail in Supporting Information 11.

Discussion

How adaptive traits arise at the genetic level has been a classic question in evolutionary biology for decades. Here, we used a QTL mapping approach to identify loci contributing to variation in thermotolerance in hybrids between the northern swordtail species *X. malinche* and *X. birchmanni*. Mapping CT_{max} QTL in an artificial hybrid population revealed one underdominant QTL spanning ~1.5 Mb on chromosome 22 and a putative interacting QTL on chromosome 15. This finding, along with our gene expression results, points to a breakdown in the response to thermal stress in hybrids, with important implications for understanding the genetic architecture and evolution of ecologically relevant traits in general.

Though more commonly reported in plants ([90,106–108]; but see [109]), underdominant QTL provide insight into genotypes that may be disadvantageous in hybrids. For example, mapping pollen fertility in *Mimulus* has identified hybrid sterility loci in heterozygotes caused by structural rearrangements [90] and mapping in tomatoes has revealed a reduction in fruit size in heterozygotes [108]. Unlike these QTL, which generally appear to have a simple genetic architecture, we find that the QTL on chromosome 22 explains a modest proportion of the total variation in this trait in *X. malinche*-*X. birchmanni* artificial hybrids. This both highlights the complex nature of this trait, and explains why, despite an average signal of reduced CT_{max} in individuals heterozygous at the chromosome 22 QTL (Fig. 2C), most F_1 and F_2 hybrids have a CT_{max} that is intermediate to the parental ranges (Fig. 2A; Supporting Information 4).

What mechanisms drive reduced thermotolerance of heterozygous individuals at the chromosome 22 QTL? One clue comes from gene expression results from *X. malinche, X. birchmanni*, and F_1 hybrids. We see widespread misexpression in F_1 hybrids (approximately 9% and 3% of temperature responsive genes in parental brain and liver, respectively), where heterozygous individuals show expression patterns far outside the range of either parental species, including at genes under the chromosome 22 QTL (Fig. 2D). These aberrant expression patterns likely result from disruption of gene expression networks in hybrids at the molecular level [43], and could lead to phenotypic effects such as the reduced CT_{max} we observe at the chromosome 22 QTL. While well-documented in literature on the evolution of gene regulation [36,94,110–112], these types of misexpression dynamics have only recently been appreciated in the speciation genetics community as a source of hybrid incompatibilities between species [94].

One particularly intriguing example of gene expression misregulation in F_1 hybrids occurs in circadian clock pathways. Overall, we find strong correlations between co-expression patterns of clock genes and temperature treatment in our RNAseq datasets. This finding is consistent with decades of data showing that expression levels of core clock genes are regulated in response to temperature across taxa (for example in plants: [104,113]; flies: [114]; fish: [115–119]; mammals: [120]). This regulatory response is important for maintaining homeostasis and timing

of the biological clock regardless of temperature-induced shifts in basic processes like enzymatic activity [121]. While we observe a strong circadian clock regulatory response to temperature treatment in both *X. malinche* and *X. birchmanni*, we find that an unexpectedly large number of circadian clock genes are misexpressed in F_1 hybrids (permuted p-value$<10^{-6}$), particularly after exposure to high temperature (Fig. 3C-D; Supporting Information 7). The response observed in parent species suggests that proper regulation of these genes is important in thermal stress response in *Xiphophorus*, and enriched misexpression in F_1 hybrids points to a potential breakdown of basic regulatory processes in hybrids. Moreover, multiple pairs of genes that fall under the chromosome 22 and 15 QTLs are known to interact with clock genes. For example, several loci under the QTL regions (*akt3, zbtb18, nrxn3b, tnfaip3*, and *nfkbia*) are co-expressed or interact with the regulatory clock gene *bhlhe40* [122–126]. Future work should address the functional basis of this misregulation as well as whether hybrids exhibit difficulty maintaining homeostasis compared to the two parental species, particularly at a range of rearing temperatures.

Consistent with a role in fitness in natural populations, we see evidence of selection on ancestry at a handful of temperature-associated clock genes. Natural hybrids from the Acuapa and Tlatemaco populations derive the majority of their genomes from *X. birchmanni* and *X. malinche*, respectively, but both reside at *X. birchmanni* typical elevations. Specifically, clock genes *nr1d2b* and *arntl1a* (Fig. 4A) are unusually skewed towards *X. birchmanni* ancestry in both populations. This could indicate an ecological advantage of the *X. birchmanni* alleles at these genes (or selection to resolve misexpression).

Given evidence for poorer performance and widespread misexpression in some hybrid individuals in response to thermal stress, we were curious about the ways that known hybrid incompatibilities interact with the thermal environment. Previous work has shown that in *X. malinche-X. birchmanni* hybrids, the combination of *X. malinche* ancestry at the gene *cd97* and *X. birchmanni* ancestry at the gene *xmrk* results in the formation of a malignant melanoma. This incompatibility appears to reduce fitness in the wild based on population resampling results, but the mechanism is unclear, as individuals can survive for more than 2 years in the lab even with severe melanoma [89]. We found that both 3D melanoma and less severe melanoma are significantly correlated with reduced CT_{max} in *X. malinche-X. birchmanni* hybrids. This hints at a potential ecological fitness consequence for individuals with the melanoma incompatibility and exploring whether this relationship is causal is an exciting future direction (we discuss this result more thoroughly in Supporting Information 11).

We set out to use QTL mapping and differential gene expression analysis to identify the genetic basis of differences in thermotolerance between *X. malinche* and *X. birchmanni,* so that we could identify regions of the genome that may undergo adaptive introgression in response to changing thermal environments. However, our mapping and RNAseq results instead uncovered signals of hybrid breakdown and potential BDMIs. Our results highlight a more general problem with QTL mapping of species-level differences; in some cases, breakdown in the biological processes and traits of interest in hybrids will obscure the differences between the parental species that researchers seek to map. On the other hand, our results provide indirect clues into the expected outcomes for our original questions. Hybrids between *X. malinche* and *X. birchmanni* experience widespread misregulation of genes that respond to thermal treatments in the parental species, and

737 some individuals that harbor heterozygous ancestry at the chromosome 22 QTL or a common
738 hybrid incompatibility between species exhibit markedly reduced thermotolerance. These results
739 suggest that adaptive introgression of as of yet unidentified *X. birchmanni* thermotolerance
740 alleles may not be sufficient to offset the costs of hybridization, and therefore may not lead to
741 higher thermotolerance in *X. malinche* populations. We also note that although we focus on
742 CT_{max} in the present study, *X. malinche* is found in environments with lower temperatures than
743 those experienced by any other *Xiphophorus* species. Studying the genetic architecture of
744 tolerance of cool temperatures in *X. malinche* may provide insight into the pressures driving
745 regulatory divergence between species and misexpression in hybrids.
746
747 Together, this work highlights the potential for ecological incompatibilities to play a role in
748 selection on *X. malinche*-*X. birchmanni* hybrids [46]. Nearly a decade of work has uncovered
749 evidence for genetic incompatibilities between these two species, but most cases that have been
750 evaluated in detail have focused on intrinsic hybrid incompatibilities [89,127]. Our results
751 underscore how shifts in global climate may impact a suite of biological processes and
752 exacerbate or uncover ecological incompatibilities in hybrids. Such potential consequences may
753 limit the success of genetic rescue as an effective strategy for population conservation.
754

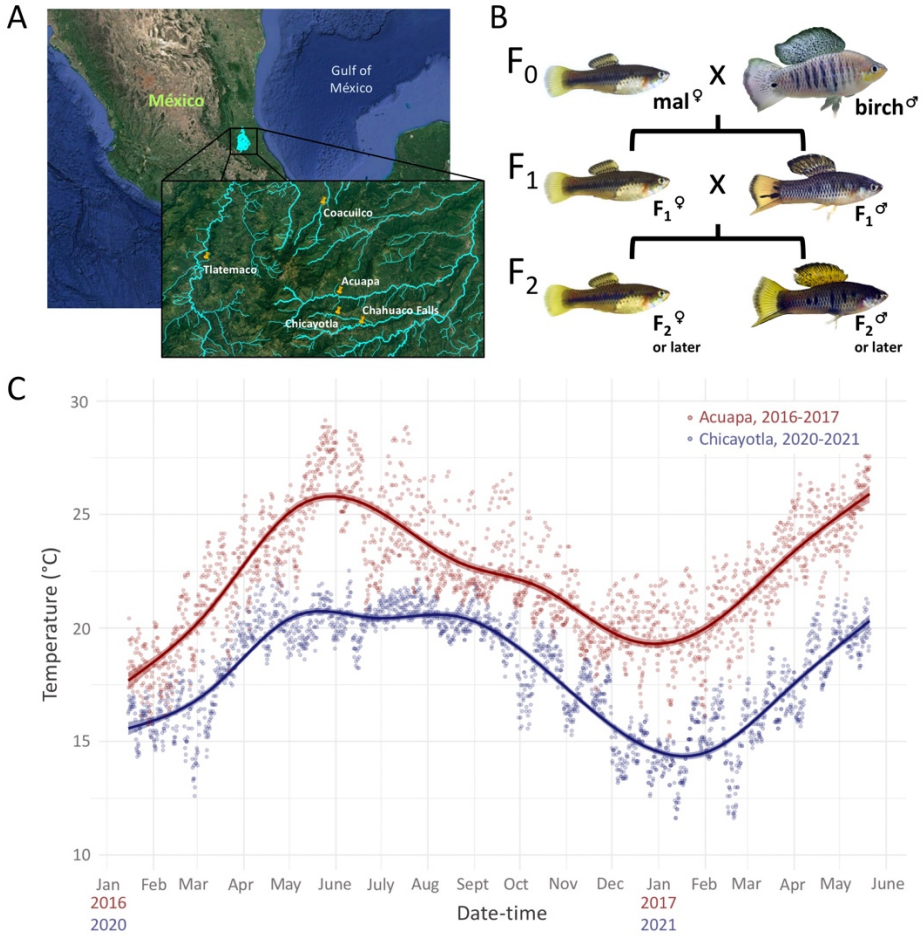


Figure 1.

A. Map adapted from Google Earth showing the five natural populations from which fish were collected for data used in this study. Pure *X. malinche* mothers and *X. birchmanni* fathers used in crosses and for RNAseq experiments were originally collected from the Chicayotla and Coacuilco populations, respectively. Natural hybrids were collected from Chahuaco Falls to evaluate links between hybrid melanoma and CT_{max}, and natural hybrids for analysis of population-level ancestry were collected from the low elevation Acuapa and Tlatemaco hybrid populations.

B. The cross design used to generate individuals for both the mapping and RNAseq datasets. Wild *X. malinche* mothers from Chicayotla and *X. birchmanni* fathers from Coacuilco were crossed create an F_1 population. A subset of F_1s were crossed to generate an artificial hybrid mapping population that was raised in common garden conditions. Other F_1 individuals were raised in the lab and used for the RNAseq thermal stress experiment. Abbreviations: mal – *X. malinche*, birch – *X. birchmanni*.

C. Temperature data collected by HOBO loggers deployed at Acuapa from 2016-2017 and Chicayotla from 2020-2021. Acuapa is a hybrid population that is found at a similar elevation to pure *X. birchmanni* sites (~400 meters versus ~250-300 meters; [64]), and Chicayotla is a site where pure *X. malinche* individuals are found (~1000 meters). Data points were collected four times per day by the loggers. Points and trend lines are shown in red for Acuapa and blue for Chicayotla.

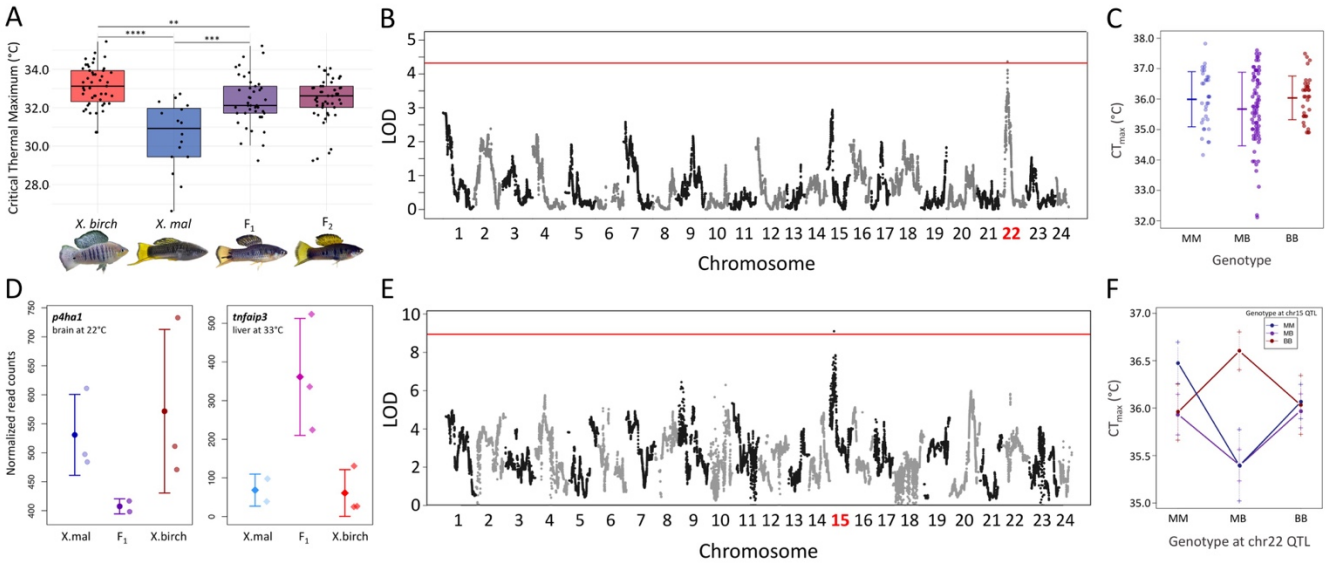


Figure 2.

A. Results of CT_{max} trials on parental and hybrid individuals raised under common garden conditions indicate that variation in thermal tolerance between *X. birchmanni* and *X. malinche* is controlled in part by genetic factors. *X. birchmanni* has a significantly higher CT_{max} than *X. malinche*, and F_1s and F_2s on average have an intermediate CT_{max}. See Table S1 for p-values for statistical comparisons between groups using Mann-Whitney Wilcoxon test.

B. QTL mapping reveals one region on chromosome 22 associated with CT_{max}. The QTL is significant at a 10% false discovery rate threshold, determined by permutations (red line).

C. Artificial hybrids individuals with a heterozygous genotype at the peak associated marker on chromosome 22 have a 0.3°C reduction in CT_{max} on average compared to hybrid individuals homozygous for *X. malinche* or *X. birchmanni* ancestry, which have comparable CT_{max} on average. Bars and whiskers show the CT_{max} means for each genotype and 1 standard deviation. Points represent the CT_{max} of individual hybrids.

D. Of the 45 genes under the CT_{max} QTL on chromosome 22, several show misexpression in F_1s in at least one tissue or thermal context. The two examples shown here are *p4ha1*, which has significantly reduced expression in F_1 brains at ambient temperature, and *tnfaip3*, which has significantly higher expression in F_1 liver tissue under thermal stress (both at FDR adjusted p-value < 0.1). In these expression plots, mean normalized counts at 22.5°C are represented by a circle in a darker color and mean normalized counts at 33.5°C are represented by a diamond in a brighter color. Error bars show one standard deviation of expression.

E. A second QTL scan, adding genotype at the chromosome 22 QTL as an interaction term, uncovered a putative interacting QTL on chromosome 15. This QTL is significant at a 10% false discovery rate threshold, determined by permutations (red line).

F. Interaction plot of the peak associated marker of the chromosome 22 QTL (on the x-axis) and the peak associated marker of the chromosome 15 QTL (in the legend). This analysis shows that a combination of a heterozygous or homozygous *X. malinche* ancestry at the chromosome 15 QTL and a heterozygous genotype at the chromosome 22 QTL is associated with reduced CT_{max}. By contrast combination of homozygous *X. birchmanni* ancestry and heterozygous genotype at the chromosome 15 and 22 QTLs, respectively, is

808 associated with a modest increase in CT_{max} (see Table S5 for adjusted p-values). Bars and
809 whiskers show the mean and 1 standard error.
810

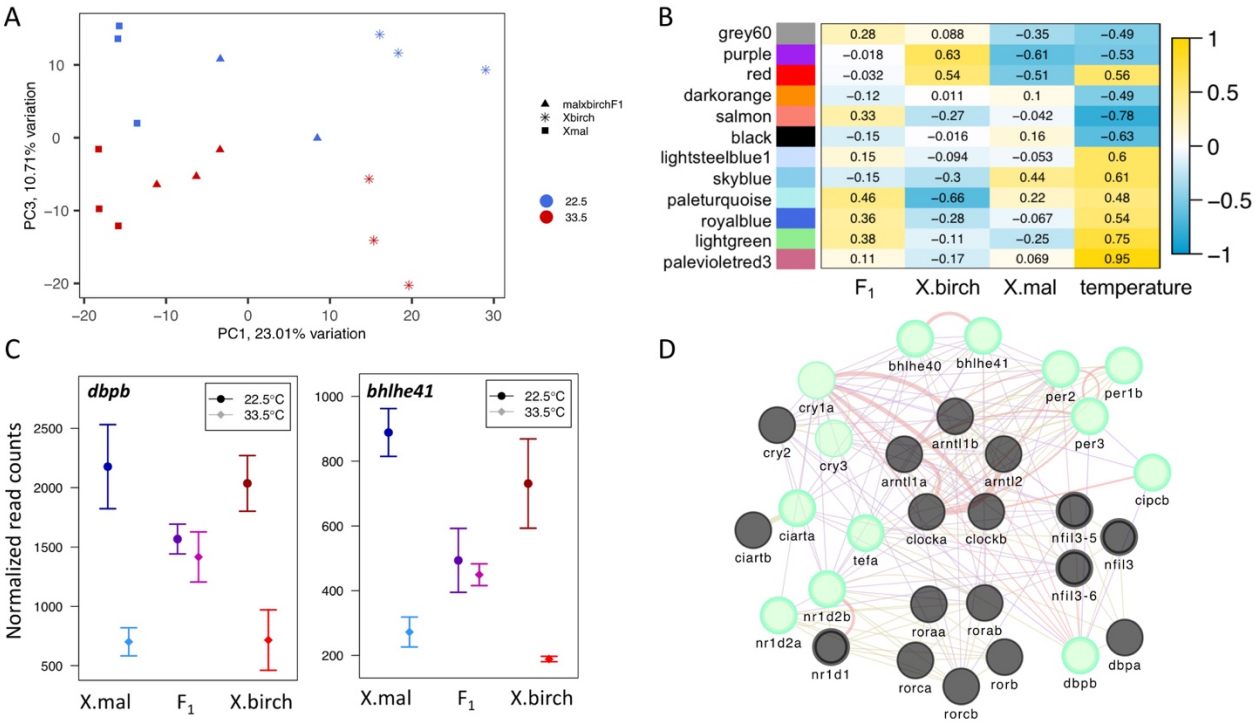


Figure 3.

A. PCA plot of normalized gene count data in the brain for all 17 individuals for which RNAseq data was collected. Individuals clearly separate by genotype and temperature treatment along PC1 and PC3 respectively. Genotype explained 23.01% of the variation in overall expression and temperature treatment explained 10.71%. PC2, which is not shown here, explained 19.22% of the variation in expression and was most strongly correlated with sequencing batch.

B. Weighted gene co-expression analysis uncovered 12 temperature-associated modules in the brain (shown here) and 2 in the liver (Fig. S9). Traits are listed on the x-axis, and color blocks and labels on the y-axis represent the WGCNA module. Pearson's correlation coefficients are listed for each module and trait, with box color corresponding to the strength of the correlation (yellow spectrum for a positive trait-module correlation, blue spectrum for a negative trait-module correlation).

C. Several clock genes that were identified in the circadian clock gene-enriched module, including *dbpb* and *bhlhe41*, are misexpressed under both ambient and high temperature conditions in F_1 brains. Interestingly, the mechanism of misexpression may be due to a failure of F_1 hybrids to respond to temperature change. *X. birchmanni* and *X. malinche* strongly downregulate both genes in response to high temperature, while F_1s do not.

D. The gene network for core circadian clock genes in the *Xiphophorus* genome, predicted by GeneMania [128] and visualized with Cytoscape [129]. The structure of the network is colored based on the nature of evidence of each interaction, including direct interactions between genes (red), co-expression (purple), and shared domains (yellow). Genes that are misexpressed in F_1 brains at high temperature in our study are highlighted in bright green, and genes that appeared in the circadian clock gene expression module identified by WGCNA are shown with a bold outline.

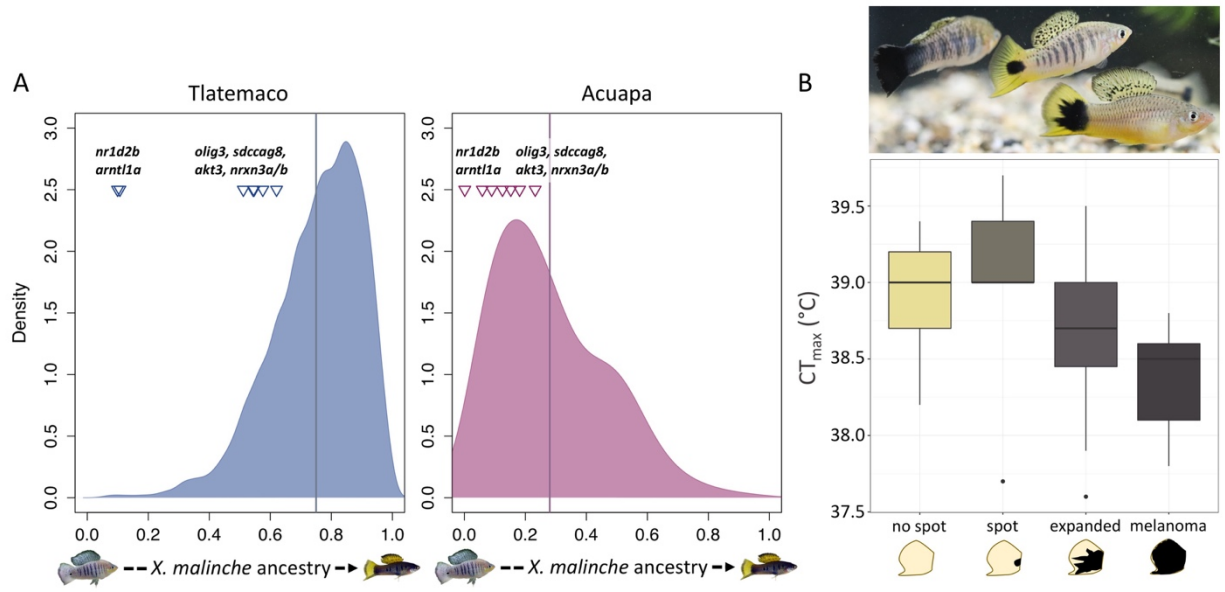


Figure 4.

A. Ancestry at regions implicated in thermal stress response compared to the genome-wide ancestry distributions in two natural hybrid populations that occur at low elevations. Individuals from the Tlatemaco population derive on average ~75% of their genome from the *X. malinche* parent species and individuals from the Acuapa population derive on average ~28% of their genome from the *X. malinche* parent (genome-wide means represented by solid lines). Conversely, a handful of genes under the chromosome 22 (*olig3, sdccga8, akt3*) and 15 (*nrxn3a, nrxn3b*) QTL and two clock genes (*nr1d2b, arntl1a*) have unusually high *X. birchmanni* ancestry in both populations, raising the possibility that there may be positive selection for *X. birchmanni* ancestry at these genes in low elevation populations (see Table S14 for p-values from permutations).

B. The top image shows three Chahuaco Falls hybrids, from left to right, with 3D melanoma, normal spotted caudal, and expanded spot phenotypes. Boxplots show CT_{max} of lab-reared Chahuaco Falls hybrids, split by spotted caudal phenotype. Lab-reared individuals with expanded spot and 3D melanoma phenotypes have significantly lower CT_{max} compared to individuals with no spot or a normal spotting pattern.

Acknowledgements

We thank Lisa Couper, Pablo Delclos, Justin Havird, Erik Iverson, Katya Mack, and members of the Schumer and Rosenthal labs for helpful feedback and discussion. We are grateful to the Mexican federal government for permission to collect samples. We thank Stanford University and the Stanford Research Computing Center for providing computational support for this project. This work was supported by NIGMS Center of the National Institutes of Health under award number T32GM007276, a Human Frontiers in Science Program grant to MS (RGY0081), a SICB GIAR to CP, and an NSF GRFP to VG.

869 **References**

870 1. Moran BM, Payne C, Langdon Q, Powell DL, Brandvain Y, Schumer M. 2021 The genomic
871 consequences of hybridization. *eLife* **10**, e69016. (doi:10.7554/eLife.69016)

872 2. Anderson E, Stebbins GL. 1954 Hybridization as an Evolutionary Stimulus. *Evolution* **8**, 378–
873 388. (doi:10.2307/2405784)

874 3. Stebbins GL. 1959 The synthetic approach to problems of organic evolution. *Cold Spring*
875 *Harb. Symp. Quant. Biol.* **24**, 305–311. (doi:10.1101/sqb.1959.024.01.028)

876 4. Lewontin RC, Birch LC. 1966 Hybridization as a Source of Variation for Adaptation to New
877 Environments. *Evolution* **20**, 315–336. (doi:10.1111/j.1558-5646.1966.tb03369.x)

878 5. Arnold, Michael. 1997 *Natural Hybridization and Evolution*. Oxford, New York: Oxford
879 University Press.

880 6. Barton NH. 2001 The role of hybridization in evolution. *Mol. Ecol.* **10**, 551–568.
881 (doi:10.1046/j.1365-294x.2001.01216.x)

882 7. Hedrick PW. 2013 Adaptive introgression in animals: examples and comparison to new
883 mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* **22**, 4606–4618.
884 (doi:10.1111/mec.12415)

885 8. Stelkens RB, Brockhurst MA, Hurst GDD, Greig D. 2014 Hybridization facilitates
886 evolutionary rescue. *Evol. Appl.* **7**, 1209–1217. (doi:10.1111/eva.12214)

887 9. Suarez-Gonzalez A, Lexer C, Cronk QCB. 2018 Adaptive introgression: a plant perspective.
888 *Biol. Lett.* **14**, 20170688. (doi:10.1098/rsbl.2017.0688)

889 10. Mitchell N, Owens GL, Hovick SM, Rieseberg LH, Whitney KD. 2019 Hybridization
890 speeds adaptive evolution in an eight-year field experiment. *Sci. Rep.* **9**, 1–12.
891 (doi:10.1038/s41598-019-43119-4)

892 11. Rieseberg LH *et al.* 2003 Major ecological transitions in wild sunflowers facilitated by
893 hybridization. *Science* **301**, 1211–1216. (doi:10.1126/science.1086949)

894 12. Krehenwinkel H, Tautz D. 2013 Northern range expansion of European populations of
895 the wasp spider Argiope bruennichi is associated with global warming–correlated genetic
896 admixture and population-specific temperature adaptations. *Mol. Ecol.* **22**, 2232–2248.
897 (doi:https://doi.org/10.1111/mec.12223)

898 13. Song Y, Endepols S, Klemann N, Richter D, Matuschka F-R, Shih C-H, Nachman MW,
899 Kohn MH. 2011 Adaptive Introgression of Anticoagulant Rodent Poison Resistance by
900 Hybridization between Old World Mice. *Curr. Biol.* **21**, 1296–1301.
901 (doi:10.1016/j.cub.2011.06.043)

902 14. Dasmahapatra KK *et al.* 2012 Butterfly genome reveals promiscuous exchange of
903 mimicry adaptations among species. *Nature* **487**, 94–98. (doi:10.1038/nature11041)

904 15. Besansky NJ, Krzywinski J, Lehmann T, Simard F, Kern M, Mukabayire O, Fontenille
905 D, Touré Y, Sagnon N. 2003 Semipermeable species boundaries between Anopheles gambiae
906 and Anopheles arabiensis: Evidence from multilocus DNA sequence variation. *Proc. Natl.*
907 *Acad. Sci.* **100**, 10818–10823. (doi:10.1073/pnas.1434337100)

908 16. Grant PR, Grant BR. 2010 Conspecific versus heterospecific gene exchange between
909 populations of Darwin's finches. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 1065–1076.
910 (doi:10.1098/rstb.2009.0283)

911 17. Baskett M, Gomulkiewicz R. 2011 Introgressive hybridization as a mechanism for
912 species rescue. *Theor. Ecol.* **4**, 223–239. (doi:10.1007/s12080-011-0118-0)

913 18. Oziolor EM, Reid NM, Yair S, Lee KM, VerPloeg SG, Bruns PC, Shaw JR, Whitehead
914 A, Matson CW. 2019 Adaptive introgression enables evolutionary rescue from extreme
915 environmental pollution. *Science* **364**, 455–457. (doi:10.1126/science.aav4155)

916 19. Jones MR *et al.* 2018 Adaptive introgression underlies polymorphic seasonal camouflage
917 in snowshoe hares. *Science* **360**, 1355–1358. (doi:10.1126/science.aar5273)

918 20. Zhang X, Witt KE, Bañuelos MM, Ko A, Yuan K, Xu S, Nielsen R, Huerta-Sanchez E.
919 2021 The history and evolution of the Denisovan-EPAS1 haplotype in Tibetans. *Proc. Natl.*
920 *Acad. Sci.* **118**. (doi:10.1073/pnas.2020803118)

921 21. Racimo F, Sankararaman S, Nielsen R, Huerta-Sánchez E. 2015 Evidence for archaic
922 adaptive introgression in humans. *Nat. Rev. Genet.* **16**, 359–371. (doi:10.1038/nrg3936)

923 22. Garroway CJ, Bowman J, Cascaden TJ, Holloway GL, Mahan CG, Malcolm JR, Steele
924 MA, Turner G, Wilson PJ. 2010 Climate change induced hybridization in flying squirrels.
925 *Glob. Change Biol.* **16**, 113–121. (doi:10.1111/j.1365-2486.2009.01948.x)

926 23. Wellenreuther M, Tynkkynen K, Svensson EI. 2010 Simulating Range Expansion: Male
927 Species Recognition and Loss of Premating Isolation in Damselflies. *Evolution* **64**, 242–252.
928 (doi:10.1111/j.1558-5646.2009.00815.x)

929 24. Muhlfeld CC, Kovach RP, Jones LA, Al-Chokhachy R, Boyer MC, Leary RF, Lowe WH,
930 Luikart G, Allendorf FW. 2014 Invasive hybridization in a threatened species is accelerated
931 by climate change. *Nat. Clim. Change* **4**, 620–624. (doi:10.1038/nclimate2252)

932 25. Chunco AJ. 2014 Hybridization in a warmer world. *Ecol. Evol.* **4**, 2019–2031.
933 (doi:10.1002/ece3.1052)

934 26. Bay R, Rose N, Logan C, Palumbi S. 2017 Genomic models predict successful coral
935 adaptation if future ocean warming rates are reduced. *Sci. Adv.* **3**, e1701413.
936 (doi:10.1126/sciadv.1701413)

27. Schiffers K *et al.* 2014 Landscape structure and genetic architecture jointly impact rates of niche evolution. *Ecography* **37**, 1218–1229. (doi:10.1111/ecog.00768)

28. Thomas L, Rose NH, Bay RA, López EH, Morikawa MK, Ruiz-Jones L, Palumbi SR. 2018 Mechanisms of Thermal Tolerance in Reef-Building Corals across a Fine-Grained Environmental Mosaic: Lessons from Ofu, American Samoa. *Front. Mar. Sci.* **4**. (doi:10.3389/fmars.2017.00434)

29. Kramer K, van der Werf B, Schelhaas M-J. 2015 Bring in the genes: genetic-ecophysiological modeling of the adaptive response of trees to environmental change. With application to the annual cycle. *Front. Plant Sci.* **0**. (doi:10.3389/fpls.2014.00742)

30. Maheshwari S, Barbash DA. 2011 The Genetics of Hybrid Incompatibilities. *Annu. Rev. Genet.* **45**, 331–355. (doi:10.1146/annurev-genet-110410-132514)

31. Arnold ML, Martin NH. 2010 Hybrid fitness across time and habitats. *Trends Ecol. Evol.* **25**, 530–536. (doi:10.1016/j.tree.2010.06.005)

32. Bateson W. 2009 Heredity and Variation in Modern Lights. *Hered. Var. Mod. Lights* (doi:10.1017/CBO9780511693953.006)

33. Dobzhansky T. 1937 *Genetics and the Origin of Species*.

34. Muller HJ. 1942 Isolating mechanisms, evolution, and temperature. *Biol. Symp.* **6**, 71–125.

35. Orr HA. 1996 Dobzhansky, Bateson, and the Genetics of Speciation. *Genetics* **144**, 1331–1335.

36. Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL. 2005 Compensatory cis-trans Evolution and the Dysregulation of Gene Expression in Interspecific Hybrids of Drosophila. *Genetics* **171**, 1813–1822. (doi:10.1534/genetics.105.047449)

37. Meiklejohn CD, Coolon JD, Hartl DL, Wittkopp PJ. 2014 The roles of cis- and trans-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression. *Genome Res.* **24**, 84–95. (doi:10.1101/gr.156414.113)

38. Meiklejohn CD *et al.* 2018 Gene flow mediates the role of sex chromosome meiotic drive during complex speciation. *eLife* **7**, e35468. (doi:10.7554/eLife.35468)

39. Werren JH. 2011 Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci.* **108**, 10863–10870. (doi:10.1073/pnas.1102343108)

40. Presgraves DC. 2010 The molecular evolutionary basis of species formation. *Nat. Rev. Genet.* **11**, 175–180. (doi:10.1038/nrg2718)

969 41. Frank SA. 1991 Divergence of Meiotic Drive-Suppression Systems as an Explanation for
970 Sex-Biased Hybrid Sterility and Inviability. *Evolution* **45**, 262–267. (doi:10.1111/j.1558-
971 5646.1991.tb04401.x)

972 42. Hurst LD, Pomiankowski A. 1991 Causes of sex ratio bias may account for unisexual
973 sterility in hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**,
974 841–858.

975 43. Landry CR, Hartl DL, Ranz JM. 2007 Genome clashes in hybrids: insights from gene
976 expression. *Heredity* **99**, 483–493. (doi:10.1038/sj.hdy.6801045)

977 44. Mack KL, Campbell P, Nachman MW. 2016 Gene regulation and speciation in house
978 mice. *Genome Res.* **26**, 451–461. (doi:10.1101/gr.195743.115)

979 45. Arnegard ME *et al.* 2014 Genetics of ecological divergence during speciation. *Nature*
980 **511**, 307–311. (doi:10.1038/nature13301)

981 46. Thompson KA *et al.* 2021 Genetic evidence for environment-dependent hybrid
982 incompatibilities in threespine stickleback. *bioRxiv* , 2021.06.24.449805.
983 (doi:10.1101/2021.06.24.449805)

984 47. Schluter D. 2001 Ecology and the origin of species. *Trends Ecol. Evol.* **16**, 372–380.
985 (doi:10.1016/S0169-5347(01)02198-X)

986 48. Dettman JR, Sirjusingh C, Kohn LM, Anderson JB. 2007 Incipient speciation by
987 divergent adaptation and antagonistic epistasis in yeast. *Nature* **447**, 585–588.
988 (doi:10.1038/nature05856)

989 49. Dettman JR, Anderson JB, Kohn LM. 2008 Divergent adaptation promotes reproductive
990 isolation among experimental populations of the filamentous fungus Neurospora. *BMC Evol.*
991 *Biol.* **8**, 35. (doi:10.1186/1471-2148-8-35)

992 50. Thompson KA, Urquhart-Cronish M, Whitney KD, Rieseberg LH, Schluter D. 2021
993 Patterns, Predictors, and Consequences of Dominance in Hybrids. *Am. Nat.* **197**, E72–E88.
994 (doi:10.1086/712603)

995 51. Hatfield T, Schluter D. 1999 Ecological Speciation in Sticklebacks: Environment-
996 Dependent Hybrid Fitness. *Evolution* **53**, 866–873. (doi:10.1111/j.1558-5646.1999.tb05380.x)

997 52. Cooper BS, Sedghifar A, Nash WT, Comeault AA, Matute DR. 2018 A Maladaptive
998 Combination of Traits Contributes to the Maintenance of a Drosophila Hybrid Zone. *Curr.*
999 *Biol.* **28**, 2940-2947.e6. (doi:10.1016/j.cub.2018.07.005)

1000 53. Delmore KE, Irwin DE. 2014 Hybrid songbirds employ intermediate routes in a
1001 migratory divide. *Ecol. Lett.* **17**, 1211–1218. (doi:10.1111/ele.12326)

1002 54. Nosil P. 2007 Divergent host plant adaptation and reproductive isolation between
1003 ecotypes of Timema cristinae walking sticks. *Am. Nat.* **169**, 151–162. (doi:10.1086/510634)

1004 55. Bjerkan KN *et al.* 2020 Genetic variation and temperature affects hybrid barriers during
1005 interspecific hybridization. *Plant J.* **101**, 122–140. (doi:10.1111/tpj.14523)

1006 56. Ofori BY, Stow AJ, Baumgartner JB, Beaumont LJ. 2017 Influence of adaptive capacity
1007 on the outcome of climate change vulnerability assessment. *Sci. Rep.* **7**, 12979.
1008 (doi:10.1038/s41598-017-13245-y)

1009 57. Foo SA, Byrne M. 2016 Chapter Two - Acclimatization and Adaptive Capacity of
1010 Marine Species in a Changing Ocean. In *Advances in Marine Biology* (ed BE Curry), pp. 69–
1011 116. Academic Press. (doi:10.1016/bs.amb.2016.06.001)

1012 58. Calosi P, Bilton DT, Spicer JI. 2008 Thermal tolerance, acclimatory capacity and
1013 vulnerability to global climate change. *Biol. Lett.* **4**, 99–102. (doi:10.1098/rsbl.2007.0408)

1014 59. Cowles RB (Raymond B, Bogert CM (Charles M. 1944 A preliminary study of the
1015 thermal requirements of desert reptiles. Bulletin of the AMNH ; v. 83, article 5. *Reptile*
1016 *thermal requirements*

1017 60. Becker CD, Genoway RG. In press. Evaluation of the critical thermal maximum for
1018 determining thermal tolerance of freshwater fish. , 12.

1019 61. Martin BT, Douglas MR, Chafin TK, Placyk JS, Birkhead RD, Phillips CA, Douglas ME.
1020 2020 Contrasting signatures of introgression in North American box turtle (Terrapene spp.)
1021 contact zones. *Mol. Ecol.* **29**, 4186–4202. (doi:10.1111/mec.15622)

1022 62. Griffiths JS, Kawji Y, Kelly MW. 2020 An Experimental Test of Adaptive Introgression
1023 in Locally Adapted Populations of Splash Pool Copepods. *Mol. Biol. Evol.*
1024 (doi:10.1093/molbev/msaa289)

1025 63. Pereira RJ, Barreto FS, Burton RS. 2014 Ecological Novelty by Hybridization:
1026 Experimental Evidence for Increased Thermal Tolerance by Transgressive Segregation in
1027 Tigriopus Californicus. *Evolution* **68**, 204–215. (doi:https://doi.org/10.1111/evo.12254)

1028 64. Culumber ZW, Fisher HS, Tobler M, Mateos M, Barber PH, Sorenson MD, Rosenthal
1029 GG. 2011 Replicated hybrid zones of Xiphophorus swordtails along an elevational gradient.
1030 *Mol. Ecol.* **20**, 342–356. (doi:10.1111/j.1365-294X.2010.04949.x)

1031 65. Culumber ZW, Shepard DB, Coleman SW, Rosenthal GG, Tobler M. 2012 Physiological
1032 adaptation along environmental gradients and replicated hybrid zone structure in swordtails
1033 (Teleostei: Xiphophorus). *J. Evol. Biol.* **25**, 1800–1814. (doi:10.1111/j.1420-
1034 9101.2012.02562.x)

1035 66. Fisher HS, Wong BBM, Rosenthal GG. 2006 Alteration of the chemical environment
1036 disrupts communication in a freshwater fish. *Proc. R. Soc. B Biol. Sci.* **273**, 1187–1193.
1037 (doi:10.1098/rspb.2005.3406)

1038 67. Schumer M, Powell DL, Delclós PJ, Squire M, Cui R, Andolfatto P, Rosenthal GG. 2017
1039 Assortative mating and persistent reproductive isolation in hybrids. *Proc. Natl. Acad. Sci.* **114**,
1040 10936–10941. (doi:10.1073/pnas.1711238114)

1041 68. Schumer M, Powell DL, Corbett-Detig R. 2020 Versatile simulations of admixture and
1042 accurate local ancestry inference with mixnmatch and ancestryinfer. *Mol. Ecol. Resour.* **20**,
1043 1141–1151. (doi:10.1111/1755-0998.13175)

1044 69. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows–Wheeler
1045 transform. *Bioinformatics* **25**, 1754–1760. (doi:10.1093/bioinformatics/btp324)

1046 70. Li H. 2011 A statistical framework for SNP calling, mutation discovery, association
1047 mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*
1048 **27**, 2987–2993. (doi:10.1093/bioinformatics/btr509)

1049 71. Corbett-Detig R, Nielsen R. 2017 A Hidden Markov Model Approach for Simultaneously
1050 Estimating Local Ancestry and Admixture Time Using Next Generation Sequence Data in
1051 Samples of Arbitrary Ploidy. *PLOS Genet.* **13**, e1006529. (doi:10.1371/journal.pgen.1006529)

1052 72. Powell DL, Payne C, Banerjee SM, Keegan M, Bashkirova E, Cui R, Andolfatto P,
1053 Rosenthal GG, Schumer M. 2021 The Genetic Architecture of Variation in the Sexually
1054 Selected Sword Ornament and Its Evolution in Hybrid Populations. *Curr. Biol.* **31**, 923-
1055 935.e11. (doi:10.1016/j.cub.2020.12.049)

1056 73. Broman KW, Wu H, Sen S, Churchill GA. 2003 R/qtl: QTL mapping in experimental
1057 crosses. *Bioinforma. Oxf. Engl.* **19**, 889–890. (doi:10.1093/bioinformatics/btg112)

1058 74. Schumer M, Cui R, Powell DL, Dresner R, Rosenthal GG, Andolfatto P. 2014 High-
1059 resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species.
1060 *eLife* **3**, e02535. (doi:10.7554/eLife.02535)

1061 75. Haley CS, Knott SA. 1992 A simple regression method for mapping quantitative trait loci
1062 in line crosses using flanking markers. *Heredity* **69**, 315–324. (doi:10.1038/hdy.1992.131)

1063 76. Xu S. 2003 Theoretical basis of the Beavis effect. *Genetics* **165**, 2259–2268.

1064 77. Schumer M *et al.* 2018 Natural selection interacts with recombination to shape the
1065 evolution of hybrid genomes. *Science* **360**, 656–660. (doi:10.1126/science.aar3684)

1066 78. Schartl M *et al.* 2013 The genome of the platyfish, Xiphophorus maculatus , provides
1067 insights into evolutionary adaptation and several complex traits. *Nat. Genet.* **45**, 567–572.
1068 (doi:10.1038/ng.2604)

1069 79. Wittbrodt J, Adam D, Malitschek B, Mäueler W, Raulf F, Telling A, Robertson SM,
1070 Schartl M. 1989 Novel putative receptor tyrosine kinase encoded by the melanoma-inducing
1071 Tu locus in Xiphophorus. *Nature* **341**, 415–421. (doi:10.1038/341415a0)

1072 80. Bray NL, Pimentel H, Melsted P, Pachter L. 2016 Near-optimal probabilistic RNA-seq
1073 quantification. *Nat. Biotechnol.* **34**, 525–527. (doi:10.1038/nbt.3519)

1074 81. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion
1075 for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. (doi:10.1186/s13059-014-0550-8)

1076 82. Stephens M. 2017 False discovery rates: a new deal. *Biostatistics* **18**, 275–294.
1077 (doi:10.1093/biostatistics/kxw041)

1078 83. Langfelder P, Horvath S. 2008 WGCNA: an R package for weighted correlation network
1079 analysis. *BMC Bioinformatics* **9**, 559. (doi:10.1186/1471-2105-9-559)

1080 84. Zhang B, Horvath S. 2005 A general framework for weighted gene co-expression
1081 network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, Article17. (doi:10.2202/1544-6115.1128)

1082 85. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. 2009 GAGE: generally
1083 applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161.
1084 (doi:10.1186/1471-2105-10-161)

1085 86. Durinck S, Spellman PT, Birney E, Huber W. 2009 Mapping identifiers for the
1086 integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**,
1087 1184–1191. (doi:10.1038/nprot.2009.97)

1088 87. Falcon S, Gentleman R. 2007 Using GOstats to test gene lists for GO term association.
1089 *Bioinforma. Oxf. Engl.* **23**, 257–258. (doi:10.1093/bioinformatics/btl567)

1090 88. Morgan, Martin, Falcon S, Gentleman R. In press. GSEABase: Gene set enrichment data
1091 structures and methods version 1.52.1 from Bioconductor. See https://rdrr.io/bioc/GSEABase/
1092 (accessed on 5 August 2021).

1093 89. Powell DL *et al.* 2020 Natural hybridization reveals incompatible alleles that cause
1094 melanoma in swordtail fish. *Science* **368**, 731–736. (doi:10.1126/science.aba5216)

1095 90. Stathos A, Fishman L. 2014 Chromosomal rearrangements directly cause underdominant
1096 F1 pollen sterility in Mimulus lewisii–Mimulus cardinalis hybrids. *Evolution* **68**, 3109–3119.
1097 (doi:10.1111/evo.12503)

1098 91. White MJD. 1973 *Animal cytology and evolution*. Cambridge [England: University Press.

1099 92. Kirkpatrick M, Barton N. 2006 Chromosome Inversions, Local Adaptation and
1100 Speciation. *Genetics* **173**, 419–434. (doi:10.1534/genetics.105.047985)

1101 93. Wellenreuther M, Bernatchez L. 2018 Eco-Evolutionary Genomics of Chromosomal
1102 Inversions. *Trends Ecol. Evol.* **33**, 427–440. (doi:10.1016/j.tree.2018.04.002)

1103 94. Mack KL, Nachman MW. 2017 Gene Regulation and Speciation. *Trends Genet. TIG* **33**,
1104 68–80. (doi:10.1016/j.tig.2016.11.003)

95. Depienne C *et al.* 2017 Genetic and phenotypic dissection of 1q43q44 microdeletion syndrome and neurodevelopmental phenotypes associated with mutations in ZBTB18 and HNRNPU. *Hum. Genet.* **136**, 463–479. (doi:10.1007/s00439-017-1772-0)

96. Walter P, Ron D. 2011 The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science* **334**, 1081–1086. (doi:10.1126/science.1209038)

97. Oeckinghaus A, Ghosh S. 2009 The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb. Perspect. Biol.* **1**, a000034. (doi:10.1101/cshperspect.a000034)

98. Zhang H-X *et al.* 2013 Rig-I regulates NF-κB activity through binding to Nf-κb1 3'-UTR mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6459–6464. (doi:10.1073/pnas.1304432110)

99. Kawai T, Akira S. 2007 Signaling to NF-κB by Toll-like receptors. *Trends Mol. Med.* **13**, 460–469. (doi:10.1016/j.molmed.2007.09.002)

100. Saxena M, Yeretssian G. 2014 NOD-Like Receptors: Master Regulators of Inflammation and Cancer. *Front. Immunol.* **0**. (doi:10.3389/fimmu.2014.00327)

101. Glaser FT, Stanewsky R. 2005 Temperature Synchronization of the Drosophila Circadian Clock. *Curr. Biol.* **15**, 1352–1363. (doi:10.1016/j.cub.2005.06.056)

102. Nagashima K, Matsue K, Konishi M, Iidaka C, Miyazaki K, Ishida N, Kanosue K. 2005 The involvement of Cry1 and Cry2 genes in the regulation of the circadian body temperature rhythm in mice. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **288**, R329–R335. (doi:10.1152/ajpregu.00395.2004)

103. Low KH, Lim C, Ko HW, Edery I. 2008 Natural variation in the splice site strength of a clock gene and species-specific thermal adaptation. *Neuron* **60**, 1054–1067. (doi:10.1016/j.neuron.2008.10.048)

104. Blair EJ, Bonnot T, Hummel M, Hay E, Marzolino JM, Quijada IA, Nagel DH. 2019 Contribution of time of day and the circadian clock to the heat stress responsive transcriptome in Arabidopsis. *Sci. Rep.* **9**, 4814. (doi:10.1038/s41598-019-41234-w)

105. Ikeda R *et al.* 2019 REV-ERBα and REV-ERBβ function as key factors regulating Mammalian Circadian Output. *Sci. Rep.* **9**, 10171. (doi:10.1038/s41598-019-46656-0)

106. Gustafsson ALS, Skrede I, Rowe HC, Gussarova G, Borgen L, Rieseberg LH, Brochmann C, Parisod C. 2014 Genetics of Cryptic Speciation within an Arctic Mustard, Draba nivalis. *PLOS ONE* **9**, e93834. (doi:10.1371/journal.pone.0093834)

107. García-Lara S, Khairallah MM, Vargas M, Bergvinson DJ. 2009 Mapping of QTL Associated with Maize Weevil Resistance in Tropical Maize. *Crop Sci.* **49**, 139–149. (doi:10.2135/cropsci2007.06.0326)

108. Brekke TD, Stroud JA, Shaw DS, Crawford S, Steele KA. 2019 QTL mapping in salad tomatoes. *Euphytica* **215**, 115. (doi:10.1007/s10681-019-2440-3)

1140 109. Kenney-Hunt JP *et al.* 2008 Pleiotropic Patterns of Quantitative Trait Loci for 70 Murine
1141 Skeletal Traits. *Genetics* **178**, 2275–2288. (doi:10.1534/genetics.107.084434)

1142 110. Turner LM, White MA, Tautz D, Payseur BA. 2014 Genomic Networks of Hybrid
1143 Sterility. *PLOS Genet.* **10**, e1004162. (doi:10.1371/journal.pgen.1004162)

1144 111. Go AC, Civetta A. 2020 Hybrid Incompatibilities and Transgressive Gene Expression
1145 Between Two Closely Related Subspecies of Drosophila. *Front. Genet.* **11**, 1544.
1146 (doi:10.3389/fgene.2020.599292)

1147 112. Metzger BPH, Wittkopp PJ, Coolon JosephD. 2017 Evolutionary Dynamics of
1148 Regulatory Changes Underlying Gene Expression Divergence among Saccharomyces Species.
1149 *Genome Biol. Evol.* **9**, 843–854. (doi:10.1093/gbe/evx035)

1150 113. Wang Y *et al.* 2020 Light- and temperature-entrainable circadian clock in soybean
1151 development. *Plant Cell Environ.* **43**, 637–648. (doi:10.1111/pce.13678)

1152 114. Kaneko H, Head LM, Ling J, Tang X, Liu Y, Hardin PE, Emery P, Hamada FN. 2012
1153 Circadian Rhythm of Temperature Preference and Its Neural Control in Drosophila. *Curr.*
1154 *Biol.* **22**, 1851–1857. (doi:10.1016/j.cub.2012.08.006)

1155 115. Lahiri K, Vallone D, Gondi SB, Santoriello C, Dickmeis T, Foulkes NS. 2005
1156 Temperature Regulates Transcription in the Zebrafish Circadian Clock. *PLOS Biol.* **3**, e351.
1157 (doi:10.1371/journal.pbio.0030351)

1158 116. Lahiri K, Froehlich N, Heyd A, Foulkes NS, Vallone D. 2014 Developmental Stage-
1159 Specific Regulation of the Circadian Clock by Temperature in Zebrafish. *BioMed Res. Int.*
1160 **2014**, 930308. (doi:10.1155/2014/930308)

1161 117. Long Y, Li L, Li Q, He X, Cui Z. 2012 Transcriptomic Characterization of Temperature
1162 Stress Responses in Larval Zebrafish. *PLoS ONE* **7**, e37209.
1163 (doi:10.1371/journal.pone.0037209)

1164 118. Jesus TF, Grosso AR, Almeida-Val VMF, Coelho MM. 2016 Transcriptome profiling of
1165 two Iberian freshwater fish exposed to thermal stress. *J. Therm. Biol.* **55**, 54–61.
1166 (doi:10.1016/j.jtherbio.2015.11.009)

1167 119. Chappuis S, Ripperger JA, Schnell A, Rando G, Jud C, Wahli W, Albrecht U. 2013 Role
1168 of the circadian clock gene Per2 in adaptation to cold temperature. *Mol. Metab.* **2**, 184–193.
1169 (doi:10.1016/j.molmet.2013.05.002)

1170 120. Okamoto-Mizuno K, Mizuno K. 2012 Effects of thermal environment on sleep and
1171 circadian rhythm. *J. Physiol. Anthropol.* **31**, 14. (doi:10.1186/1880-6805-31-14)

1172 121. Sweeney BM, Hastings JW. 1960 Effects of temperature upon diurnal rhythms. *Cold*
1173 *Spring Harb. Symp. Quant. Biol.* **25**, 87–104. (doi:10.1101/sqb.1960.025.01.009)

1174 122. Lin A, Wang RT, Ahn S, Park CC, Smith DJ. 2010 A genome-wide map of human
1175 genetic interactions inferred from radiation hybrid genotypes. *Genome Res.* **20**, 1122–1132.
1176 (doi:10.1101/gr.104216.109)

1177 123. Wang IX, Ramrattan G, Cheung VG. 2015 Genetic variation in insulin-induced kinase
1178 signaling. *Mol. Syst. Biol.* **11**, 820. (doi:10.15252/msb.20156250)

1179 124. Alizadeh AA *et al.* 2000 Distinct types of diffuse large B-cell lymphoma identified by
1180 gene expression profiling. *Nature* **403**, 503–511. (doi:10.1038/35000501)

1181 125. Rosenwald A *et al.* 2001 Relation of gene expression phenotype to immunoglobulin
1182 mutation genotype in B cell chronic lymphocytic leukemia. *J. Exp. Med.* **194**, 1639–1647.
1183 (doi:10.1084/jem.194.11.1639)

1184 126. Innocenti F *et al.* 2011 Identification, replication, and functional fine-mapping of
1185 expression quantitative trait loci in primary human liver tissue. *PLoS Genet.* **7**, e1002078.
1186 (doi:10.1371/journal.pgen.1002078)

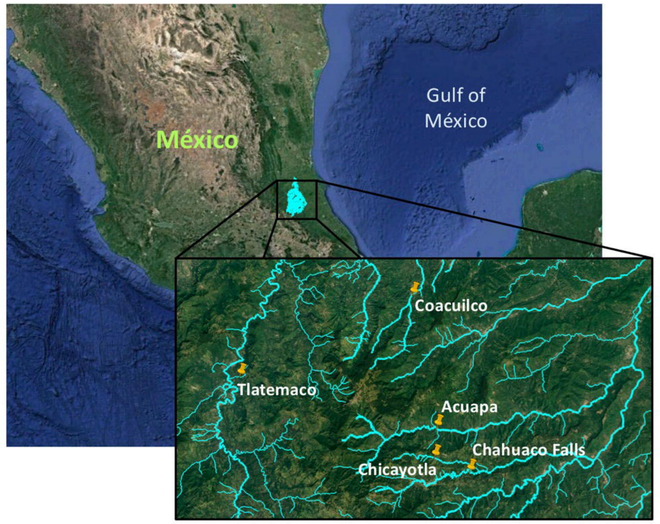
1187 127. Moran BM *et al.* 2021 A Lethal Genetic Incompatibility between Naturally Hybridizing
1188 Species in Mitochondrial Complex I. *bioRxiv* , 2021.07.13.452279.
1189 (doi:10.1101/2021.07.13.452279)

1190 128. Warde-Farley D *et al.* 2010 The GeneMANIA prediction server: biological network
1191 integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38**, W214–
1192 W220. (doi:10.1093/nar/gkq537)

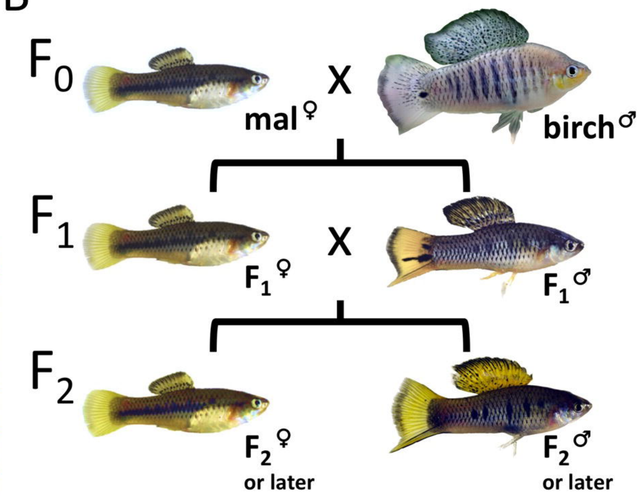
1193 129. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski
1194 B, Ideker T. 2003 Cytoscape: A Software Environment for Integrated Models of Biomolecular
1195 Interaction Networks. *Genome Res.* **13**, 2498–2504. (doi:10.1101/gr.1239303)

1196

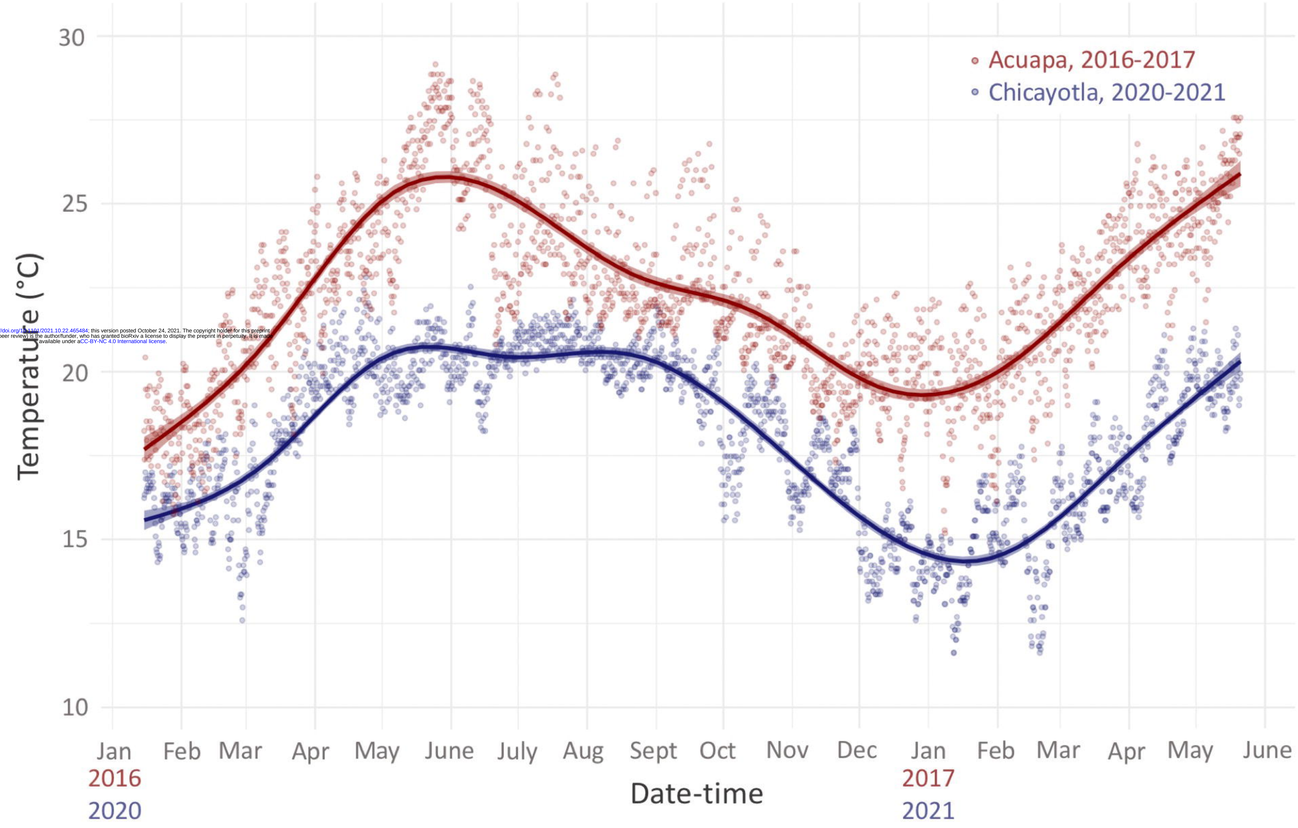
A

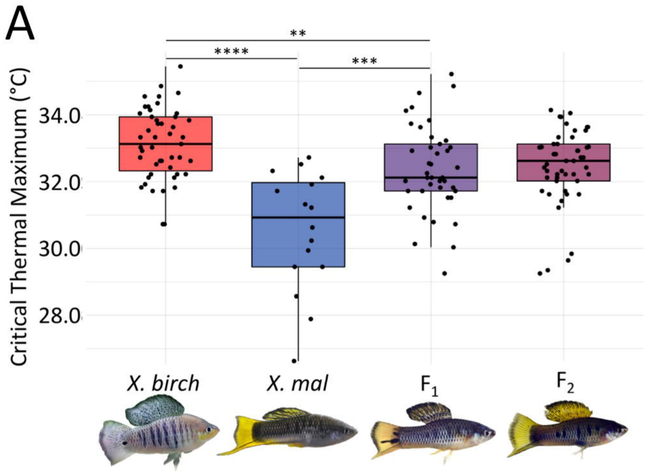


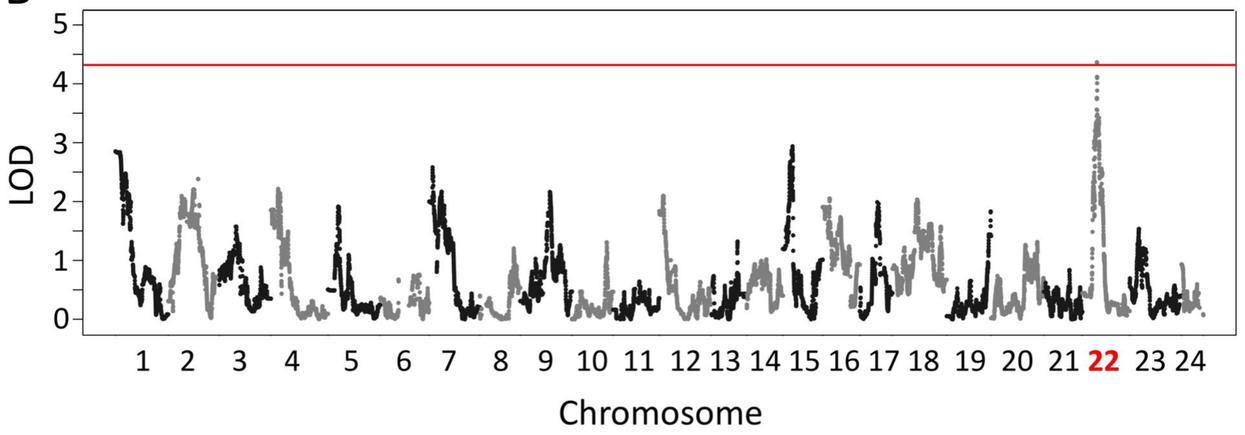
B

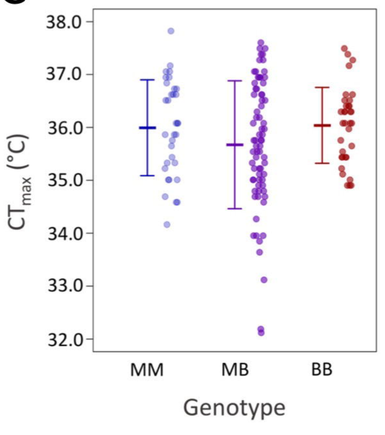


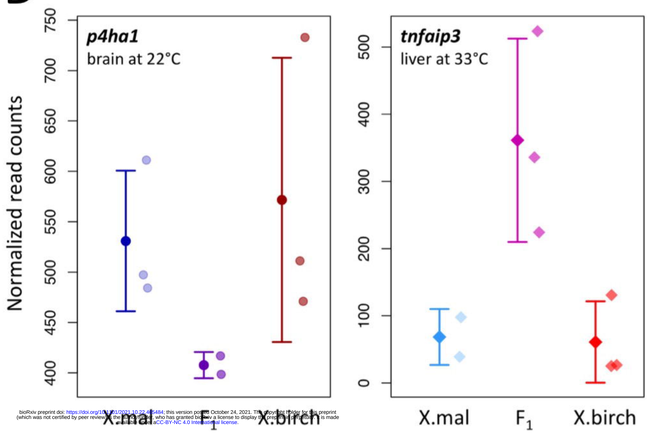
C

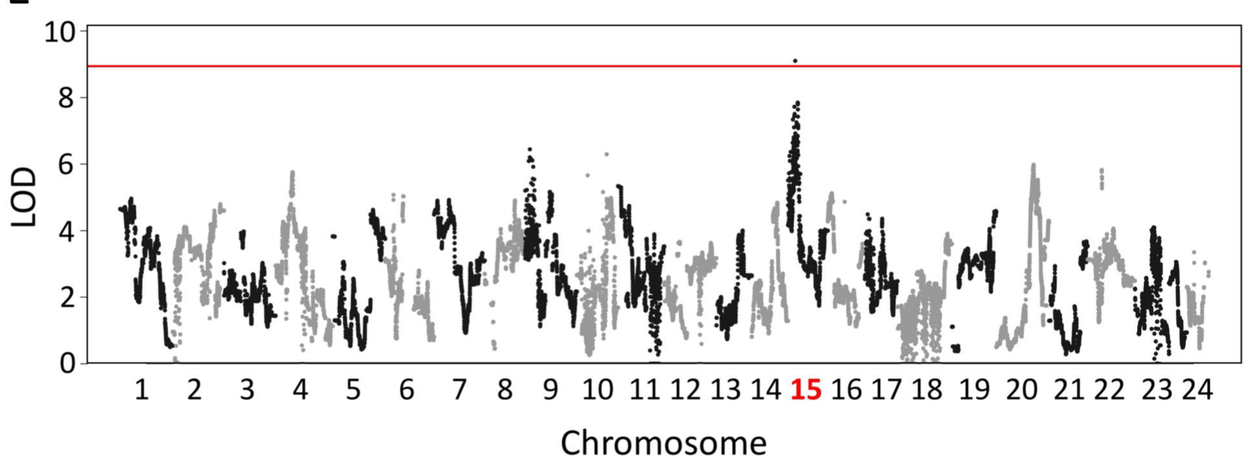


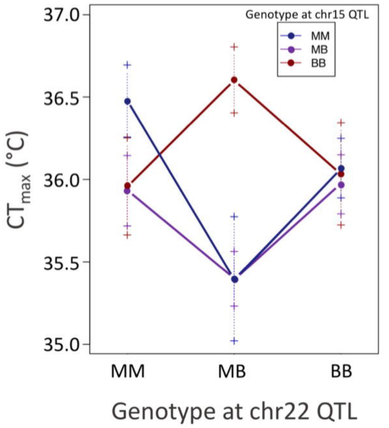


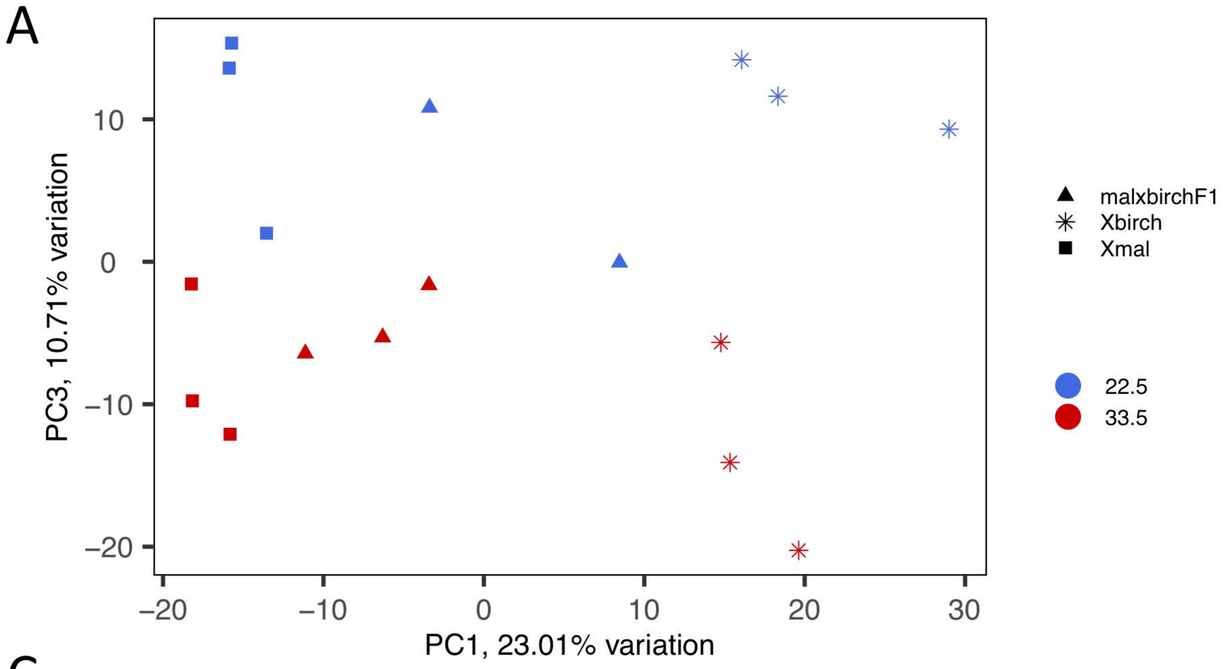


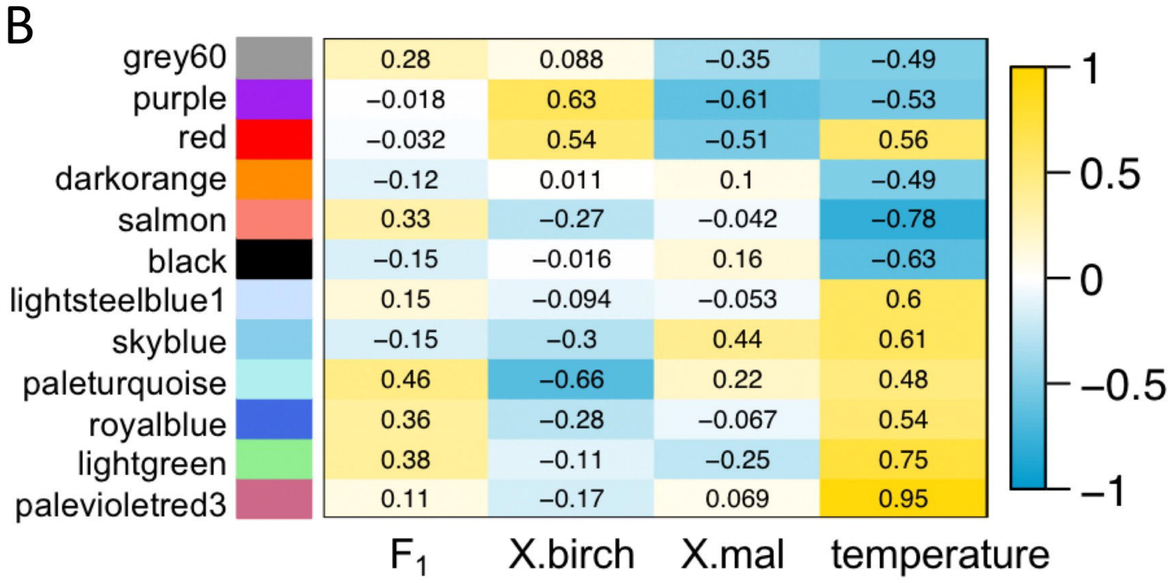


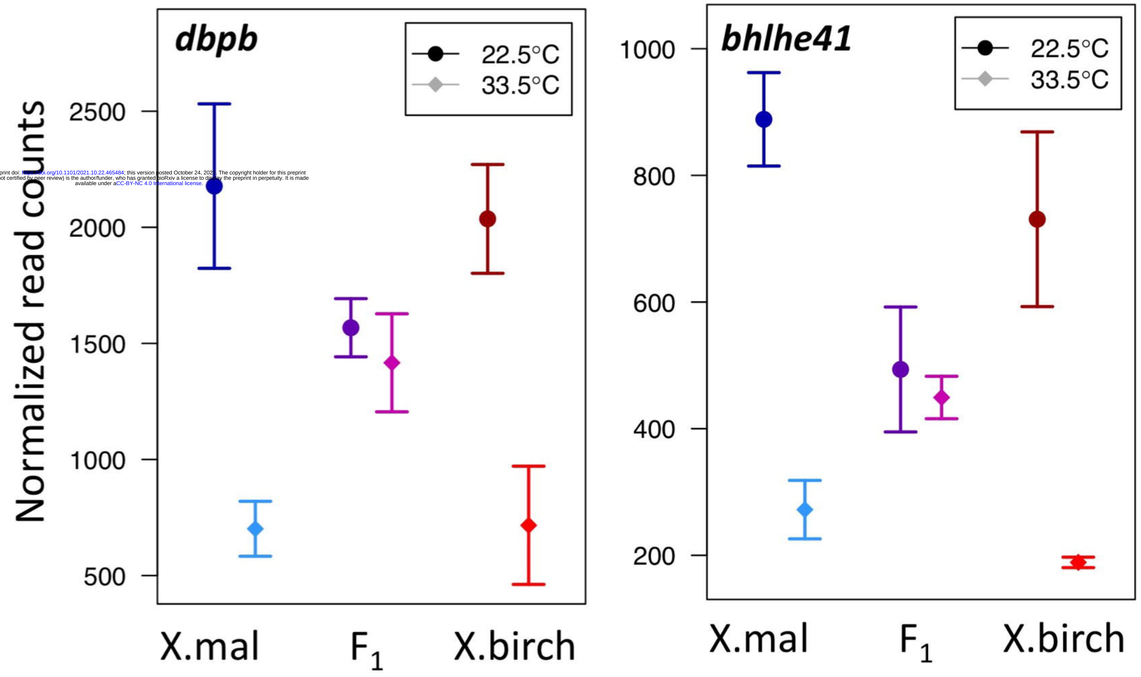


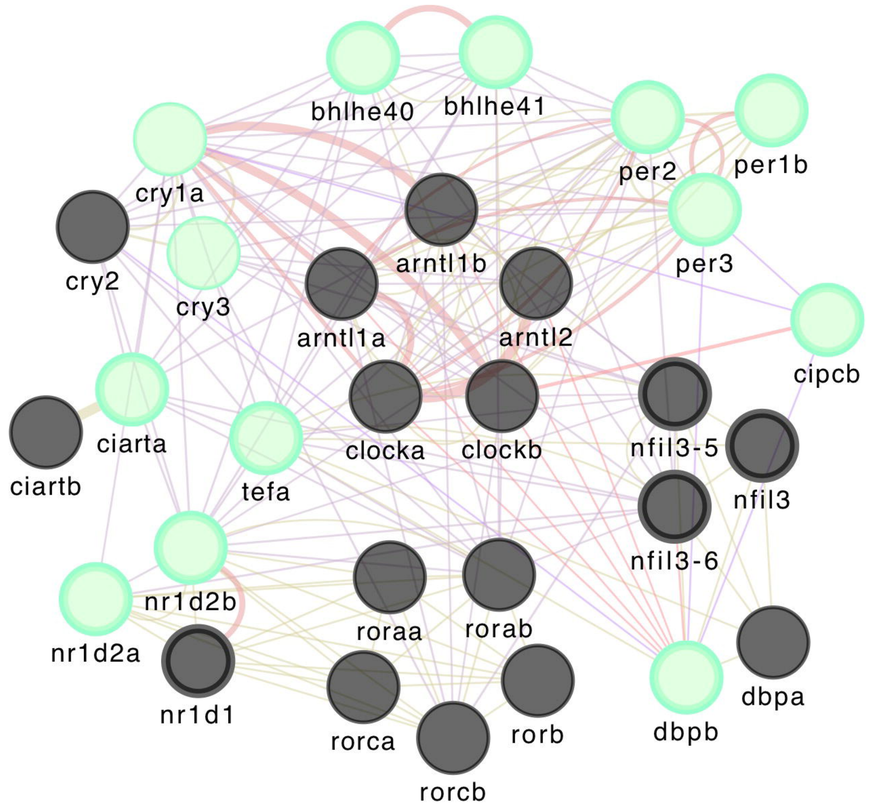





A
PC3, 10.71% variation
PC1, 23.01% variation
10
0
−10
−20
−20
−10
0
10
20
30
malxbirchF1
Xbirch
Xmal
22.5
33.5


B
grey60 0.28 0.088 −0.35 −0.49
purple −0.018 0.63 −0.61 −0.53
red −0.032 0.54 −0.51 0.56
darkorange −0.12 0.011 0.1 −0.49
salmon 0.33 −0.27 −0.042 −0.78
black −0.15 −0.016 0.16 −0.63
lightsteelblue1 0.15 −0.094 −0.053 0.6
skyblue −0.15 −0.3 0.44 0.61
paleturquoise 0.46 −0.66 0.22 0.48
royalblue 0.36 −0.28 −0.067 0.54
lightgreen 0.38 −0.11 −0.25 0.75
palevioletred3 0.11 −0.17 0.069 0.95
F1
X.birch
X.mal
temperature
1
0.5
0
−0.5
−1


C
dbpb
22.5°C
33.5°C
Normalized read counts
2500
2000
1500
1000
500
X.mal
F1
X.birch
bhlhe41
22.5°C
33.5°C
1000
800
600
400
200
X.mal
F1
X.birch


D
bhlhe40
bhlhe41
per2
per1b
cry1a
per3
cry2
arntl1b
cry3
arntl1a
arntl2
cipcb
ciarta
clocka
clockb
nfil3-5
ciartb
tefa
nfil3
nfil3-6
nr1d2b
roraa
rorab
nr1d2a
dbpa
nr1d1
rorca
rorb
dbpb
rorcb