

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

2

3 **Authors:** Cheyenne Payne^{1,2}, Stephen Bovio^{2,3}, Daniel Powell^{1,2}, Theresa Gunn^{1,2}, Shreya
4 Banerjee^{1,2}, Victoria Grant^{1,2}, Gil Rosenthal^{2,3}, Molly Schumer^{1,2,4}

5

6 **Affiliations:**

7 ¹Department of Biology, Stanford University, Stanford, CA, USA

8 ²Centro de Investigaciones Científicas de las Huastecas “Aguazarca”, A.C., Calnali, Hidalgo,
9 México

10 ³Department of Biology, Texas A&M University

11 ⁴Hanna H. Gray Fellow, Howard Hughes Medical Institute, Stanford, CA, USA

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

13

14

15 **Abstract**

16

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

34

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

36 Introduction

37

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

45

46 While hybridization has played a role in adaptation on evolutionary timescales [19–21],
47 hybridization is thought to occur more frequently under environmental disturbance and stress
48 [12,13,22–24]. As environmental conditions shift due to climate change, understanding the
49 genetic mechanisms that can facilitate rapid adaptation will be critical in predicting whether
50 vulnerable populations will adapt or collapse [18,25]. This in turn requires characterizing the
51 genetic architecture of ecologically relevant traits that distinguish hybridizing species [26–29].

52

53 Although there are many examples of adaptive introgression between species [7–9,18],
54 deleterious effects of hybridization are also well-documented and have been studied for decades
55 [30,31]. Hybridization frequently uncovers negative interactions between mutations that have
56 arisen independently in the genomes of the two parental species. These interactions can result in
57 reduced hybrid viability or fertility [32–34], and their costs may outweigh the potential benefits
58 of hybridization as a source of adaptive alleles [1]. Such interactions are commonly known as
59 Bateson-Dobzhansky-Muller incompatibilities (BDMIs; [35]). While BDMIs were originally
60 envisioned to result from incompatible interactions between proteins encoded by two or more
61 genes, recent work has highlighted the diversity of mechanisms through which BDMIs may arise
62 [36–42]. Recently attention has been paid to regulatory BDMIs, which arise from the
63 coevolution of *cis* and *trans* regulatory elements within species that become mismatched and
64 cause misexpression of target genes in hybrids [1,36,43,44] (here, misexpression is defined as
65 expression of genes in hybrids that is much higher or lower than that observed in either parent
66 species).

67

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

77

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

82 “thermotolerance”) [59,60]. Little work to date has explored whether loci that control variation
83 in thermotolerance introgress between hybridizing animal populations or how effective such
84 introgression is as a mechanism of thermal adaptation (though suggestive results have been
85 reported in some species; box turtles [61]; wasps [12]; copepods [62,63]). The deficit in
86 empirical work in this area is likely due to the paired difficulties of mapping the genetic basis of
87 complex traits like thermotolerance and studying them in natural hybrid populations.
88

89 Here, we take advantage of a system where natural hybridization is ongoing between species that
90 vary in thermotolerance. Two sister species of swordtail fishes, *Xiphophorus birchmanni* and *X.*
91 *malinche*, are endemic to rivers in eastern México [64], and their distributions are determined in
92 part by their thermal habitats [65]. *X. malinche* lives in cooler (7-25°C) streams at high
93 elevations, while the more heat tolerant *X. birchmanni* lives downstream in the warmer lowlands
94 (15-35°C; [64]). These species are sympatric in regions where their temperature ranges overlap.
95 Recently, pollution has interfered with species-specific olfactory communication, causing
96 breakdown of mating barriers [66,67]. As a result, natural hybrid zones have formed, with clinal
97 ancestry patterns that mirror thermal gradients, where there is low *X. birchmanni* ancestry and
98 low thermotolerance in the highlands and high *X. birchmanni* ancestry and high thermotolerance
99 in the lowlands [64,65]. Though phenotypic plasticity contributes to this differential tolerance (as
100 shown in [65]), we show here that variation in innate thermotolerance between species is in part
101 genetic. Leveraging this finding, we combine thermotolerance assays with classic quantitative
102 trait locus (QTL) mapping, gene expression analysis, and analysis of ancestry in natural hybrid
103 populations to explore the evolution of thermotolerance in this system. Unexpectedly, we find
104 that individuals that are heterozygous for ancestry in one genomic region have reduced critical
105 thermal maxima, and F₁ hybrids have widespread misexpression of core regulatory genes of the
106 circadian clock, which appear to be associated with proper thermal regulation. Additionally, we
107 uncover a relationship between reduced thermotolerance and a previously mapped hybrid
108 incompatibility.
109
110

111 **Methods**

112

113 *Comparison of CT_{max} between *X. malinche*, *X. birchmanni*, F_{1S} , and F_{2S} and measurement of*
114 *CT_{max} for QTL mapping*

115

116 One ecologically-relevant measure of upper thermotolerance in ectotherms is the critical
117 thermal maximum, or CT_{max} [59]. Specifically, the CT_{max} of a fish is the highest temperature it
118 can withstand before it experiences loss of equilibrium and is unable to maintain its balance
119 [59,60]. We tested CT_{max} for *Xiphophorus malinche*, *X. birchmanni*, and F_1 and F_2 hybrids
120 between the two species reared in a common garden environment.

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

134 To measure variation in CT_{max} between *X. malinche*, *X. birchmanni*, F_{1S} , and F_{2S} , CT_{max}
135 trials were performed in February 2018 using methods similar to Culumber et al [65]. Trials
136 followed procedures approved in Texas A&M IACUC protocol #117419. Briefly, the test fish
137 (eight per trial, mix of males and females from one group), an air bubbler, a standard glass
138 thermometer, and a HOBO temperature logger (Onset) were placed in an enamel container
139 holding 4 L of water at ambient temperature ($16.1 \pm 0.2^\circ\text{C}$). The enamel container was nested in
140 a larger container of water which was suspended above a hot plate. Water was heated at a
141 standardized ramp-up rate of $0.3^\circ\text{C}/\text{min}$ until the fish lost equilibrium (following Becker &
142 Genoway [60]). The time and temperature of initial loss of equilibrium (i.e. the first time balance
143 is lost) for each fish was recorded, and the fish was immediately placed in an ambient
144 temperature recovery tank. Because the data departed from the assumption of normality, we used
145 a Mann-Whitney Wilcoxon test to evaluate the effect of genotype on CT_{max} (Table S1).

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

155

156

157 *DNA extraction and library preparation*

158

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

168

169 *Artificial hybrid QTL mapping sample sequencing and genotyping*

170

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

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

191

192 *CT_{max} QTL mapping analysis*

193

194 To identify regions of the genome that are associated with variation in thermotolerance,
195 we used a QTL mapping approach. We performed QTL mapping with R/qlt [73] to scan for an
196 association between genotypes at ancestry-informative markers across the genome and the CT_{max}
197 phenotype. For computational efficiency, markers were thinned to retain at most one marker per
198 20 kb. Ancestry linkage disequilibrium decays over several megabases in early generation
199 hybrids [74]; thus, we do not expect to sacrifice any power to map QTL by performing this
200 thinning. The thinning step resulted in 30,244 ancestry informative markers retained throughout
201 the genome, or ~45 per Mb.

202

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

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

213 To select an appropriate model for mapping in R/qtl, we used the R step function to
214 calculate AIC for models incorporating a suite of possible covariates, including the tank variable
215 (tank), hybrid index (the proportion of the genome derived from the *X. malinche* parent),
216 genome-wide ancestry heterozygosity, and sex (e.g. $CT_{\max} \sim \text{hybrid_index} + \text{heterozygosity} +$
217 $\text{tank} + \text{sex}$). We retained all tank variables with a significant effect on CT_{\max} (17 total) and used a
218 method called 'one-hot encoding' to recode the tank variable so that the tank variable would be
219 treated categorically by R/qtl; other covariates were not retained in the step function analysis.
220 Even though hybrid index was not retained, we included it in our final mapping model since past
221 work has suggested that failing to include ancestry as a covariate can result in artifacts in QTL
222 analysis [72].

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

232 233 *Estimating the effect size of detected QTL*

234
235 We identified one QTL on chromosome 22 and one putative interacting QTL on
236 chromosome 15 that were significantly associated with variation in CT_{\max} after controlling for
237 other covariates at the 10% false discovery threshold (see Results). We used two methods to
238 obtain estimates for the effect sizes of these QTL (i.e. the percentage of the variation in CT_{\max}
239 explained by each QTL and their interaction). First, we used the drop-one-term analysis from
240 fitting a multiple QTL model with the R/qtl function `fitqtl`, to estimate the effect sizes of the
241 chromosome 22 and 15 QTL on CT_{\max} , as well as to estimate the effect size of their interaction.
242 Because effect size estimates are often inflated in QTL studies with low statistical power [76],
243 we also performed simulations to explore the range of possible effect sizes consistent with our
244 empirical results for the main effect QTL on chromosome 22. Methods and results for those
245 simulations are reported in Supporting Information 1.

246
247
248
249

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

251

252 To compare expression of genes across the genome that respond to high temperature in
253 the two parental species and their hybrids, we used an RNAseq-based experimental approach. *X.*
254 *birchmanni* and *X. malinche* individuals born to wild mothers (collected at the Coacuilco and
255 Chicayotla populations respectively [64]) were raised at 22.5°C (14h light:10h dark cycle). A
256 separate group of *X. malinche* females were crossed with *X. birchmanni* males to generate F₁
257 hybrids. All individuals were raised in the same lab environment to adulthood before
258 experiments began. Though we cannot discount the potential impact of maternal effects on
259 expression response, all mothers were reared under the same environmental conditions.

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

277

278 *Differential gene expression analysis*

279

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

292 Before aligning reads, the program cutadapt and the FastQC wrapper tool Trim Galore!
293 were used to trim Illumina adapter sequences and low-quality bases (Phred score < 30) from
294 reads. All trimmed reads are available under NCBI BioProject PRJNA746324. One F₁ liver
295 sample from the 22.5°C ambient temperature treatment group was removed from downstream

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

303 To analyze differential expression of these genes, we used a design formula that included
304 sequencing batch, genotype (*X. birchmanni*, *X. malinche*, or F₁), and temperature treatment.
305 Using DESeq2, we normalized gene counts by library size, estimated within-experimental group
306 dispersion, fit a negative binomial generalized linear model, and tested significance with a Wald
307 test. Following these steps, shrunken log-fold changes were calculated using an adaptive
308 shrinkage estimator with a fitted mixture of normal distributions as a prior, derived from the
309 'ashr' package [82]. Genes were considered to be significantly differentially expressed between
310 groups and treatments at an FDR-adjusted p-value < 0.1. To check for potential bias as a function
311 of the reference sequence used in the pseudoalignment step, we repeated the above steps using
312 the *X. malinche* reference transcriptome. Reassuringly, qualitatively similar results were obtained
313 from this analysis (Supporting Information 2, Fig. S2).

314

315 *Co-expression network analysis with WGCNA*

316

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

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

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

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

350 After co-expression modules were identified using this approach, we sought to determine
351 whether variation in any of these modules correlated with variation in traits of interest. As such,
352 we looked for correlations between the module eigengene and genotype (*X. malinche*, *X.*
353 *birchmanni*, *X. malinche*-*X. birchmanni* F₁), temperature treatment (22.5°C and 33.5°C), or both.
354 Correlations between traits and modules were calculated using the WGCNA
355 `corPvalueStudent` function, and modules that correlated with genotype, temperature
356 treatment, or both at Student asymptotic p-value < 0.05 were selected for further analysis (31 out
357 of 54 for brain, 16 out of 50 for liver).

358
359 *GO and KEGG pathway enrichment of differentially expressed genes between temperature*
360 *treatments*

361
362 To explore which biological pathways are most affected by temperature treatment, we
363 performed KEGG pathway and Gene Ontology (GO) term enrichment analysis of differentially
364 expressed genes identified using results from the analyses of the brain and liver RNAseq data
365 described above. We asked if there were enriched KEGG pathways in the set of genes that were
366 significantly differentially expressed between *X. birchmanni*, *X. malinche*, and F₁s at each
367 temperature treatment (FDR adjusted p-value < 0.1). To do so, Gene IDs were mapped to Entrez
368 IDs using the *X. maculatus* Ensembl database (version 99), and enriched KEGG pathways for
369 each dataset were generated with the `kegg.gsets` function from the R package GAGE [85].

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

379
380 *Ancestry of QTL and circadian clock genes in natural populations*

381
382 We analyzed data from naturally occurring *X. malinche*-*X. birchmanni* hybrid
383 populations to evaluate evidence for shifts in ancestry at genes under the chromosome 22 QTL,
384 and clock genes that show misexpression in hybrids (see Results). These data from natural
385 hybrid populations have been published in previous studies [72,77], so we only describe it briefly
386 here. We analyzed data collected from the Tlatemaco (n=85) and Acuapa (n=97) hybrid
387 populations in 2017 and 2018, respectively [72,77] (Fig. 1A). We previously collected ~1X

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

396

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

398

399 Hybrids between *X. malinche* and *X. birchmanni* harbor a number of extreme traits not
400 seen in either of the parental species. One of these is a hybrid incompatibility that causes
401 melanoma in some individuals, originating from a melanocyte spotting pattern on the caudal fin
402 [89]. To evaluate any relationship between this spotted caudal phenotype and CT_{max} , we
403 measured CT_{max} using the methods described above (see *Comparison of CT_{max} between *X.**
404 *malinche, *X. birchmanni*, F_{1s} , and F_{2s} and measurement of CT_{max} for QTL mapping*) for 123 lab-
405 raised natural hybrids born from wild-caught mothers from the Chahuaco Falls population. These
406 123 natural hybrids were reared to adulthood under common conditions in the lab. Individuals
407 were classified as having one of the following caudal spot phenotypes: no spot, normal spotted
408 caudal, expanded spotted caudal, and 3D melanoma. Past histological work has indicated that
409 individuals with the expanded spotted caudal phenotype have early-stage melanoma [89].
410 Individuals were assigned a 3D melanoma phenotype if they had melanoma that had completely
411 overtaken the caudal fin (i.e. completely melanized and/or degrading) and/or that was growing
412 laterally off the fin. We used a linear model to test for a correlation between CT_{max} and caudal
413 spot phenotype.

414

415

416 **Results**

417

418 *Evidence for a genetic basis for variation in thermotolerance*

419

420 Given that *X. malinche* and *X. birchmanni* live in different thermal environments (Fig.
421 1C), we predicted that these species may have adapted to their respective thermal ranges. To
422 determine whether there was a genetic basis for variation in CT_{max} between *X. malinche* and *X.*
423 *birchmanni*, we reared *X. malinche*, *X. birchmanni*, F_1 , and F_2 hybrid fish in a common garden,
424 and measured their CT_{max} (see Methods). We found that *X. birchmanni* have significantly higher
425 CT_{max} than *X. malinche* (p-value < 10^{-6} ; see Table S1), and that F_1 and F_2 hybrids exhibit
426 intermediate CT_{max} (Fig. 2A). Though we know that CT_{max} is partially environmentally mediated
427 in this system [65], this result shows that variation in CT_{max} between these species is also partly
428 attributable to genetic differences.

429

430 *QTL mapping of loci involved in thermotolerance*

431

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

437 Surprisingly, further examination revealed that the QTL we detected was not associated
438 with species-level differences in CT_{max} . Instead, heterozygous ancestry in this region was
439 associated with an average reduction in CT_{max} of 0.3°C (Fig. 2C). We estimate this QTL to have
440 a moderate effect on the overall variation in CT_{max} in artificial hybrids, explaining approximately
441 6.9% of the variation (see Methods, Supporting Information 1, and Fig. S1 for simulations
442 evaluating effect size inflation).

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

450

451 *Exploring candidates in the QTL region*

452

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

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

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

465 Given the CT_{max} phenotypes observed in individuals heterozygous for ancestry at this
466 QTL, we were particularly interested in comparing gene expression patterns in F_1 hybrids to the
467 parental species. Because heterozygotes have reduced CT_{max} on average, we might expect genes
468 controlled by a causal locus in *cis* to be misexpressed in F_1 hybrids (which are heterozygous for
469 ancestry across the whole genome). We evaluated expression patterns in the brain and liver and
470 identified genes that were misexpressed in F_1 hybrids compared to *X. birchmanni* and *X.*
471 *malinche*, defined here as significantly higher or lower than either of the typical parental
472 expression ranges, in either temperature condition. Of the genes in the QTL interval that were
473 significantly differentially expressed between *X. malinche* and *X. birchmanni* for at least one of
474 the temperature treatments (17 in the brain, 7 in the liver), most mirrored expression levels of
475 one of the parental species or had intermediate expression in F_1 s (Table S3). However, four
476 genes in this interval (*p4ha1*, *ndnf*, *tnfaip3*, and *infgr11*) were misexpressed in F_1 s in at least one
477 tissue and temperature condition (Fig. 2D, S8). We also identified genes that responded
478 differently to thermal stress in F_1 s compared to parentals (see brain and liver expression results
479 in Tables S6 and S7, respectively). For example, the zinc-finger protein *zfp62* exhibited an
480 exaggerated response to high temperature compared to parental expression responses and the
481 spliceosome subunit *sf3b5* was significantly downregulated at high temperatures in F_1 s whereas
482 parental expression remained constant across temperatures (Fig. S8).

483

484 *Detection of a possible interacting QTL on chromosome 15*

485

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

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

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

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

509

510 *Gene expression profiles differ between species and thermal treatment*

511

512 To broadly survey changes in gene expression between parental species and their F₁
513 hybrids in response to high temperature, we generated RNAseq data for brain and liver tissue
514 from fish exposed to ambient and high temperature conditions (see Methods). We sampled the
515 brain and liver to survey two tissues that play different roles in organismal homeostasis – energy
516 consumption and detection of and response to environmental changes by the brain, and energy
517 metabolism by the liver.

518 In addition to using these data to evaluate expression patterns of genes under the QTL
519 region, we analyzed it in a genome-wide context. As expected, we found broad differences in
520 expression between tissues and species (Fig. S6), as well as strong responses to high temperature
521 (Fig. 3A, S7). The vast majority of the variation in expression in our dataset is explained by
522 tissue type (83.5%; Fig. S6); therefore, the tissue datasets were analyzed separately. Of the
523 remaining variation, genotype explained ~23% of the variation in expression between samples
524 and temperature treatment explained ~11% (Fig. 3A). A large number of genes were
525 significantly differentially expressed between *X. birchmanni* and *X. malinche* (FDR adjusted p-
526 value < 0.1) across temperatures in both tissues (brain: 3,357 and 3,121 genes; liver: 2,318 and
527 1,508 genes at 22.5°C and 33.5°C, respectively). Interestingly, while the number of genes for
528 which expression changed in response to high temperature in *X. birchmanni* and *X. malinche* was
529 similar (brain: 882 and 979 genes; liver: 113 and 38 genes, respectively), ~2.5-3.5x more genes
530 were responsive to high temperature in F₁ brain (2,318) and liver (408) tissues compared to those
531 of parentals. We found that ~3% and ~0.5% of all genes across the genome were misexpressed
532 under at least one of the two tested thermal contexts in F₁ hybrid brains and livers, respectively.
533 Moreover, ~9% and ~3% of the genes that responded to temperature in *X. birchmanni* and/or *X.*
534 *malinche* were misexpressed under at least one of the thermal contexts in F₁ brains and livers,
535 respectively. Overall, we found that more genes were misexpressed (low or high) in F₁ hybrids
536 at 22.5°C than at 33.5°C in the brain (521 versus 187, respectively), whereas the liver exhibited
537 the opposite pattern (10 versus 96, respectively); however, we are cautious of potential technical
538 factors that could influence these patterns, such as differences in variance between groups at high
539 temperature. Only a handful of genes were misexpressed under both thermal contexts. See
540 Tables S6-7 for complete results for each group, and Table S8 for a summary of F₁ expression
541 patterns.

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

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

553 FDR adjusted p-value < 0.1; Table S10). This result may be attributable to the fact that the
554 endoplasmic reticulum plays a key role in the unfolded protein response, which is activated by
555 thermal stress and is key for maintaining homeostasis during stress [96]. Intriguingly, one
556 transcriptional activator of the unfolded protein response, *xbp1*, is significantly upregulated in
557 both *X. malinche* and *X. birchmanni*, but not in F₁ hybrid brains (Fig. S8).

558 Strikingly, 20 KEGG pathways were significantly enriched in F₁ brains in response to
559 high temperature. Among these enriched pathways were protein processing in the endoplasmic
560 reticulum and signaling pathways (see Table S10 for full list) that induce the transcriptional
561 regulator of the innate immune response, *nfkb1* [97–100]. Interestingly, one of the potential sets
562 of interactors under the chromosome 22 and chromosome 15 QTL are inhibitors of *nfkb1*
563 expression (see Supporting Information 5 for more information). No significantly enriched
564 KEGG pathways were recovered from the set of genes differentially expressed in response to
565 temperature in the liver across groups.

566

567 *Co-expression network analysis reveals misexpression in F₁ clock genes*

568

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

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

581 Strikingly, several of the circadian clock genes in this shared temperature-associated
582 module are misexpressed in F₁ hybrids, particularly in data collected from the brain in the high
583 temperature treatment (Supporting Information 6-7). The number of misexpressed genes in this
584 module greatly exceeds the number expected by chance (based on permutations, Table S12, Fig.
585 3D, Supporting Information 7). This suggests that genes in these circadian clock pathways may
586 be commonly misregulated under thermal stress in *X. malinche*-*X. birchmanni* hybrids.

587 Specifically, we find that most of the clock genes in this module are strongly up- or down-
588 regulated in *X. malinche* and *X. birchmanni* brains and livers in response to high temperature. In
589 contrast, at ambient temperature, F₁ clock gene expression tends to be similar to parental
590 expression, but at high temperatures these genes are misexpressed in F₁ brains compared to
591 parental brains (Table S13; Fig. 3C-D). These results hint at a failure to regulate expression of
592 these genes in hybrids. Specifically, much of the misexpression observed in these genes is
593 attributable to the fact that while their expression in parental brains is strongly responsive to the
594 thermal treatment, F₁ expression does not change substantially between temperature treatments.
595 Additionally, some of these genes, such as the transcription factors *dbpb* and *bhlhe41* shown in
596 Fig. 3C and *nr1d2a* and *cipcb* shown in Fig. S8, show patterns of F₁ misregulation under both
597 thermal contexts. We discuss these patterns in more detail in Supporting Information 9.

598

599

600 *Ancestry patterns in natural hybrid populations at regions implicated in thermotolerance*

601

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

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

621 We next evaluated ancestry in both hybrid populations among genes in the circadian
622 clock gene expression module. Notably, two clock genes in this module that are misexpressed in
623 F₁ hybrids, *nr1d2b* and *arntl1a*, have unusually high *X. birchmanni* ancestry in both populations
624 (>89% in both, permuted p-value<0.01; Fig. 4A, Table S14). Interestingly, *nr1d2b* directly
625 represses *arntl1a* expression [105]. Clock genes with strong skews in ancestry in both natural
626 hybrid populations may be adaptive in lower elevation habitats, as this level of ancestry sharing
627 across the two populations is unexpected by chance (see Supporting Information 10). Together,
628 these analyses highlight regions that may be under selection due to their impacts on
629 thermotolerance in natural hybrid populations.

630

631 *Other phenotypes associated with thermotolerance in hybrids*

632

633 Given the overall pattern of reduced thermotolerance associated with heterozygous
634 ancestry at the chromosome 22 QTL and aberrant expression of many thermally responsive
635 genes in F₁ hybrids, we wanted to further investigate other possible phenotypic drivers of
636 reduced thermotolerance in hybrids. One trait that is present in hybrids but not in parental
637 individuals of either species is a hybrid incompatibility involving a pigmentation phenotype
638 called the “spotted caudal”. While the spotted caudal is a benign melanocyte pigmentation
639 pattern in *X. birchmanni*, it can transform into a malignant melanoma in hybrids (Fig. 4B) with
640 certain genotype combinations ([89]; those with *X. birchmanni* ancestry at the *xmrk* gene and *X.*
641 *malinche* ancestry at *cd97*).

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

645 expanded spot typical of early melanoma as well as hybrids with a more advanced 3D melanoma
646 phenotype had significantly reduced CT_{max} compared to those with a benign spot or no spot (Fig.
647 4B). This poor performance in hybrids with incompatible genotype combinations highlights one
648 potential mechanism through which underdominance in traits such as thermotolerance could
649 occur. We discuss the implications of this result in more detail in Supporting Information 11.

650

651 **Discussion**

652

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

661

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

672

673 What mechanisms drive reduced thermotolerance of heterozygous individuals at the
674 chromosome 22 QTL? One clue comes from gene expression results from *X. malinche*, *X.*
675 *birchmanni*, and F_1 hybrids. We see widespread misexpression in F_1 hybrids (approximately 9%
676 and 3% of temperature responsive genes in parental brain and liver, respectively), where
677 heterozygous individuals show expression patterns far outside the range of either parental
678 species, including at genes under the chromosome 22 QTL (Fig. 2D). These aberrant expression
679 patterns likely result from disruption of gene expression networks in hybrids at the molecular
680 level [43], and could lead to phenotypic effects such as the reduced CT_{max} we observe at the
681 chromosome 22 QTL. While well-documented in literature on the evolution of gene regulation
682 [36,94,110–112], these types of misexpression dynamics have only recently been appreciated in
683 the speciation genetics community as a source of hybrid incompatibilities between species [94].

684

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

691 of the biological clock regardless of temperature-induced shifts in basic processes like enzymatic
692 activity [121]. While we observe a strong circadian clock regulatory response to temperature
693 treatment in both *X. malinche* and *X. birchmanni*, we find that an unexpectedly large number of
694 circadian clock genes are misexpressed in F₁ hybrids (permuted p-value < 10⁻⁶), particularly after
695 exposure to high temperature (Fig. 3C-D; Supporting Information 7). The response observed in
696 parent species suggests that proper regulation of these genes is important in thermal stress
697 response in *Xiphophorus*, and enriched misexpression in F₁ hybrids points to a potential
698 breakdown of basic regulatory processes in hybrids. Moreover, multiple pairs of genes that fall
699 under the chromosome 22 and 15 QTLs are known to interact with clock genes. For example,
700 several loci under the QTL regions (*akt3*, *zbtb18*, *nrxn3b*, *tnfaip3*, and *nfkbia*) are co-expressed
701 or interact with the regulatory clock gene *bhlhe40* [122–126]. Future work should address the
702 functional basis of this misregulation as well as whether hybrids exhibit difficulty maintaining
703 homeostasis compared to the two parental species, particularly at a range of rearing
704 temperatures.

705
706 Consistent with a role in fitness in natural populations, we see evidence of selection on ancestry
707 at a handful of temperature-associated clock genes. Natural hybrids from the Acuapa and
708 Tlatemaco populations derive the majority of their genomes from *X. birchmanni* and *X.*
709 *malinche*, respectively, but both reside at *X. birchmanni* typical elevations. Specifically, clock
710 genes *nr1d2b* and *arntl1a* (Fig. 4A) are unusually skewed towards *X. birchmanni* ancestry in
711 both populations. This could indicate an ecological advantage of the *X. birchmanni* alleles at
712 these genes (or selection to resolve misexpression).

713
714 Given evidence for poorer performance and widespread misexpression in some hybrid
715 individuals in response to thermal stress, we were curious about the ways that known hybrid
716 incompatibilities interact with the thermal environment. Previous work has shown that in *X.*
717 *malinche*-*X. birchmanni* hybrids, the combination of *X. malinche* ancestry at the gene *cd97* and
718 *X. birchmanni* ancestry at the gene *xmrk* results in the formation of a malignant melanoma. This
719 incompatibility appears to reduce fitness in the wild based on population resampling results, but
720 the mechanism is unclear, as individuals can survive for more than 2 years in the lab even with
721 severe melanoma [89]. We found that both 3D melanoma and less severe melanoma are
722 significantly correlated with reduced CT_{max} in *X. malinche*-*X. birchmanni* hybrids. This hints at a
723 potential ecological fitness consequence for individuals with the melanoma incompatibility and
724 exploring whether this relationship is causal is an exciting future direction (we discuss this result
725 more thoroughly in Supporting Information 11).

726
727 We set out to use QTL mapping and differential gene expression analysis to identify the genetic
728 basis of differences in thermotolerance between *X. malinche* and *X. birchmanni*, so that we could
729 identify regions of the genome that may undergo adaptive introgression in response to changing
730 thermal environments. However, our mapping and RNAseq results instead uncovered signals of
731 hybrid breakdown and potential BDMs. Our results highlight a more general problem with QTL
732 mapping of species-level differences; in some cases, breakdown in the biological processes and
733 traits of interest in hybrids will obscure the differences between the parental species that
734 researchers seek to map. On the other hand, our results provide indirect clues into the expected
735 outcomes for our original questions. Hybrids between *X. malinche* and *X. birchmanni* experience
736 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

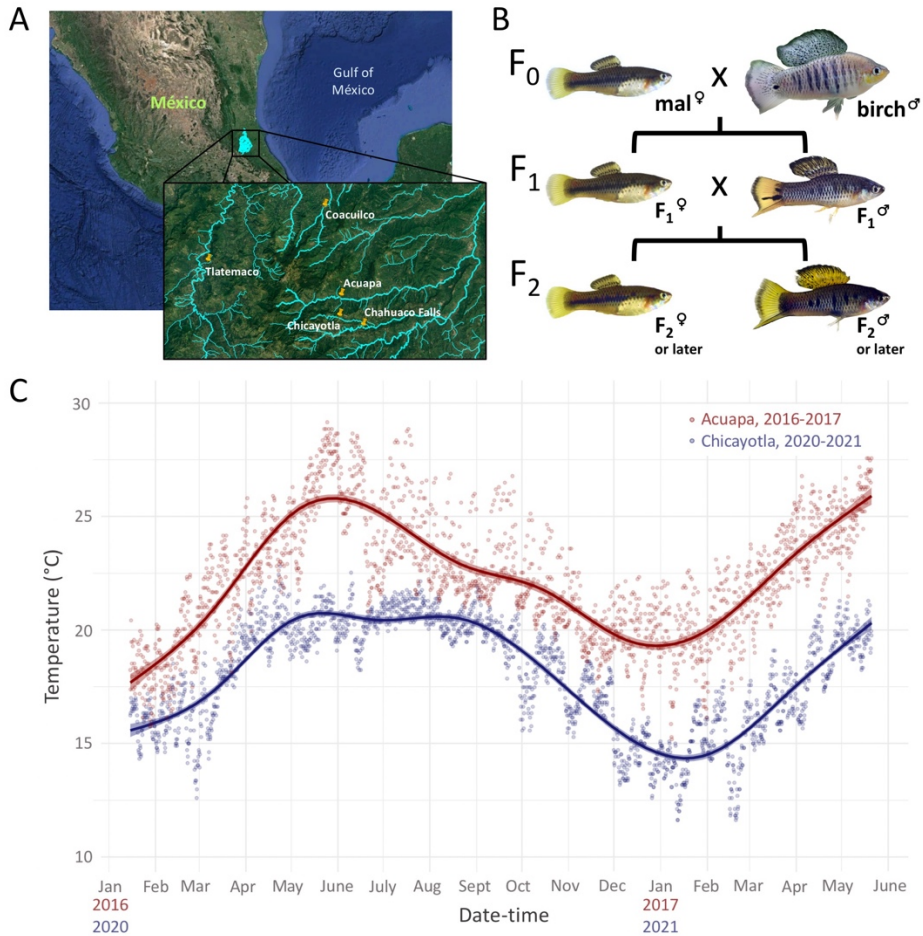


Figure 1.

- 755
756
757 **A.** Map adapted from Google Earth showing the five natural populations from which fish
758 were collected for data used in this study. Pure *X. malinche* mothers and *X. birchmanni*
759 fathers used in crosses and for RNAseq experiments were originally collected from the
760 Chicayotla and Coacuilco populations, respectively. Natural hybrids were collected from
761 Chahuaco Falls to evaluate links between hybrid melanoma and CT_{max}, and natural
762 hybrids for analysis of population-level ancestry were collected from the low elevation
763 Acuapa and Tlatemaco hybrid populations.
- 764 **B.** The cross design used to generate individuals for both the mapping and RNAseq datasets.
765 Wild *X. malinche* mothers from Chicayotla and *X. birchmanni* fathers from Coacuilco
766 were crossed create an F₁ population. A subset of F₁s were crossed to generate an
767 artificial hybrid mapping population that was raised in common garden conditions. Other
768 F₁ individuals were raised in the lab and used for the RNAseq thermal stress experiment.
769 Abbreviations: mal – *X. malinche*, birch – *X. birchmanni*.
- 770 **C.** Temperature data collected by HOBO loggers deployed at Acuapa from 2016-2017 and
771 Chicayotla from 2020-2021. Acuapa is a hybrid population that is found at a similar
772 elevation to pure *X. birchmanni* sites (~400 meters versus ~250-300 meters; [64]), and
773 Chicayotla is a site where pure *X. malinche* individuals are found (~1000 meters). Data
774 points were collected four times per day by the loggers. Points and trend lines are shown
775 in red for Acuapa and blue for Chicayotla.

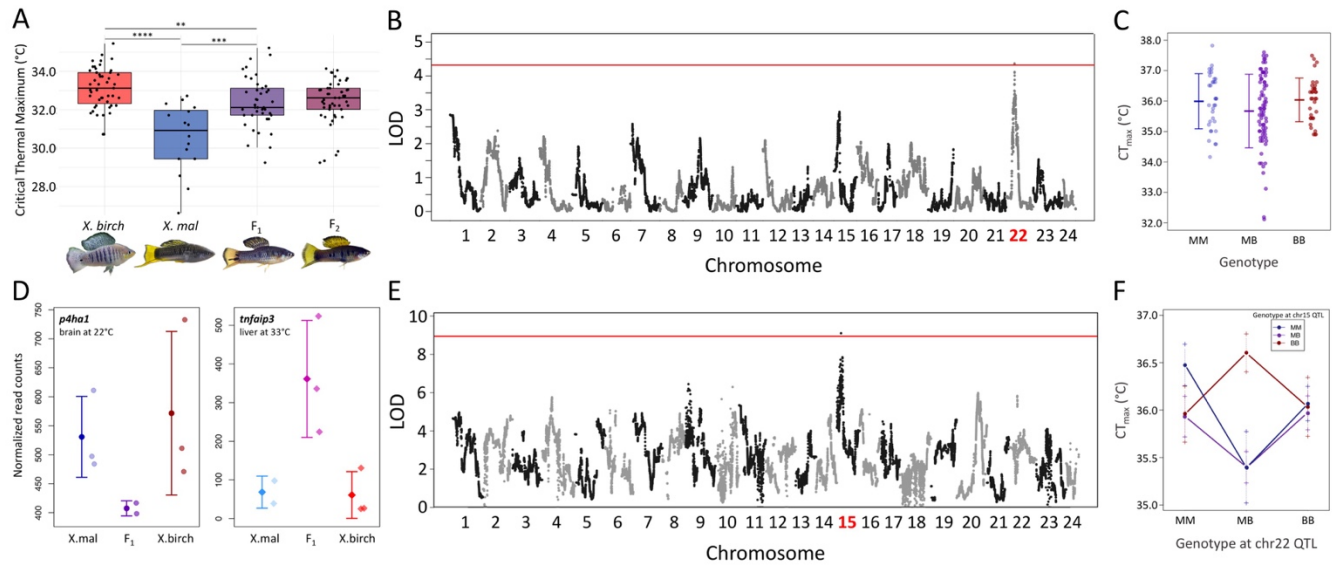


Figure 2.

- 776
777
778 **A.** Results of CT_{max} trials on parental and hybrid individuals raised under common garden
779 conditions indicate that variation in thermal tolerance between *X. birchmanni* and *X.*
780 *malinche* is controlled in part by genetic factors. *X. birchmanni* has a significantly higher
781 CT_{max} than *X. malinche*, and F₁s and F₂s on average have an intermediate CT_{max}. See
782 Table S1 for p-values for statistical comparisons between groups using Mann-Whitney
783 Wilcoxon test.
- 784 **B.** QTL mapping reveals one region on chromosome 22 associated with CT_{max}. The QTL is
785 significant at a 10% false discovery rate threshold, determined by permutations (red line).
786 **C.** Artificial hybrids individuals with a heterozygous genotype at the peak associated marker
787 on chromosome 22 have a 0.3°C reduction in CT_{max} on average compared to hybrid
788 individuals homozygous for *X. malinche* or *X. birchmanni* ancestry, which have
789 comparable CT_{max} on average. Bars and whiskers show the CT_{max} means for each
790 genotype and 1 standard deviation. Points represent the CT_{max} of individual hybrids.
- 791 **D.** Of the 45 genes under the CT_{max} QTL on chromosome 22, several show misexpression in
792 F₁s in at least one tissue or thermal context. The two examples shown here are *p4hal*,
793 which has significantly reduced expression in F₁ brains at ambient temperature, and
794 *tnfaip3*, which has significantly higher expression in F₁ liver tissue under thermal stress
795 (both at FDR adjusted p-value < 0.1). In these expression plots, mean normalized counts
796 at 22.5°C are represented by a circle in a darker color and mean normalized counts at
797 33.5°C are represented by a diamond in a brighter color. Error bars show one standard
798 deviation of expression.
- 799 **E.** A second QTL scan, adding genotype at the chromosome 22 QTL as an interaction term,
800 uncovered a putative interacting QTL on chromosome 15. This QTL is significant at a
801 10% false discovery rate threshold, determined by permutations (red line).
802 **F.** Interaction plot of the peak associated marker of the chromosome 22 QTL (on the x-axis)
803 and the peak associated marker of the chromosome 15 QTL (in the legend). This analysis
804 shows that a combination of a heterozygous or homozygous *X. malinche* ancestry at the
805 chromosome 15 QTL and a heterozygous genotype at the chromosome 22 QTL is
806 associated with reduced CT_{max}. By contrast combination of homozygous *X. birchmanni*
807 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

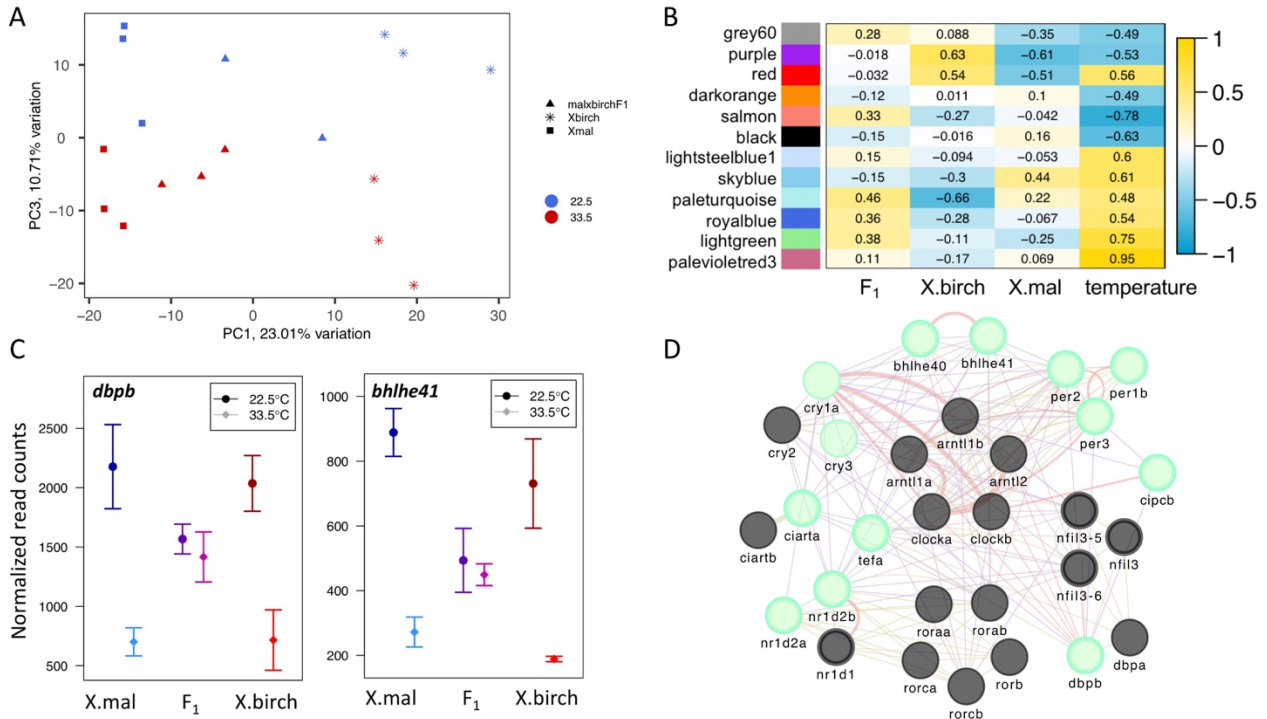
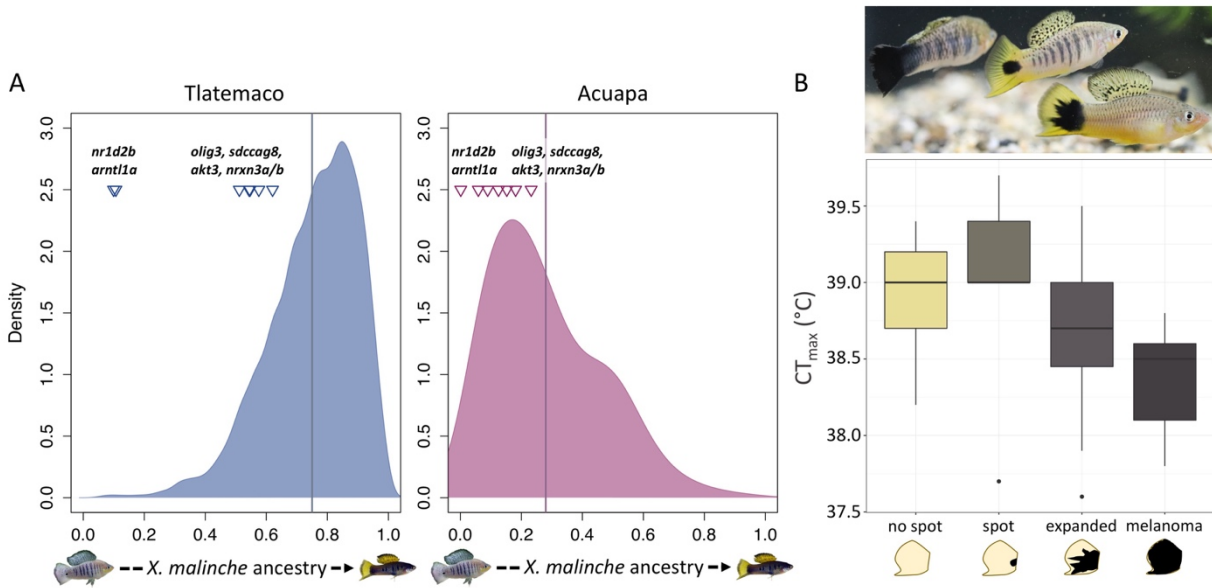


Figure 3.

- 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 F₁ brains. Interestingly, the mechanism of misexpression may be due to a failure of F₁ hybrids to respond to temperature change. *X. birchmanni* and *X. malinche* strongly downregulate both genes in response to high temperature, while F₁s 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 F₁ 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.

811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837



838

839

Figure 4.

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

- 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*, *sdccag8*, *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.

860 **Acknowledgements**

861

862 We thank Lisa Couper, Pablo Delclos, Justin Havird, Erik Iverson, Katya Mack, and members of
863 the Schumer and Rosenthal labs for helpful feedback and discussion. We are grateful to the
864 Mexican federal government for permission to collect samples. We thank Stanford University
865 and the Stanford Research Computing Center for providing computational support for this
866 project. This work was supported by NIGMS Center of the National Institutes of Health under
867 award number T32GM007276, a Human Frontiers in Science Program grant to MS (RGY0081),
868 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)

- 937 27. Schiffrers K *et al.* 2014 Landscape structure and genetic architecture jointly impact rates
938 of niche evolution. *Ecography* **37**, 1218–1229. (doi:10.1111/ecog.00768)
- 939 28. Thomas L, Rose NH, Bay RA, López EH, Morikawa MK, Ruiz-Jones L, Palumbi SR.
940 2018 Mechanisms of Thermal Tolerance in Reef-Building Corals across a Fine-Grained
941 Environmental Mosaic: Lessons from Ofu, American Samoa. *Front. Mar. Sci.* **4**.
942 (doi:10.3389/fmars.2017.00434)
- 943 29. Kramer K, van der Werf B, Schelhaas M-J. 2015 Bring in the genes: genetic-
944 ecophysiological modeling of the adaptive response of trees to environmental change. With
945 application to the annual cycle. *Front. Plant Sci.* **0**. (doi:10.3389/fpls.2014.00742)
- 946 30. Maheshwari S, Barbash DA. 2011 The Genetics of Hybrid Incompatibilities. *Annu. Rev.*
947 *Genet.* **45**, 331–355. (doi:10.1146/annurev-genet-110410-132514)
- 948 31. Arnold ML, Martin NH. 2010 Hybrid fitness across time and habitats. *Trends Ecol. Evol.*
949 **25**, 530–536. (doi:10.1016/j.tree.2010.06.005)
- 950 32. Bateson W. 2009 Heredity and Variation in Modern Lights. *Hered. Var. Mod. Lights*
951 (doi:10.1017/CBO9780511693953.006)
- 952 33. Dobzhansky T. 1937 *Genetics and the Origin of Species*.
- 953 34. Muller HJ. 1942 Isolating mechanisms, evolution, and temperature. *Biol. Symp.* **6**, 71–
954 125.
- 955 35. Orr HA. 1996 Dobzhansky, Bateson, and the Genetics of Speciation. *Genetics* **144**,
956 1331–1335.
- 957 36. Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL. 2005
958 Compensatory cis-trans Evolution and the Dysregulation of Gene Expression in Interspecific
959 Hybrids of *Drosophila*. *Genetics* **171**, 1813–1822. (doi:10.1534/genetics.105.047449)
- 960 37. Meiklejohn CD, Coolon JD, Hartl DL, Wittkopp PJ. 2014 The roles of cis- and trans-
961 regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene
962 expression. *Genome Res.* **24**, 84–95. (doi:10.1101/gr.156414.113)
- 963 38. Meiklejohn CD *et al.* 2018 Gene flow mediates the role of sex chromosome meiotic drive
964 during complex speciation. *eLife* **7**, e35468. (doi:10.7554/eLife.35468)
- 965 39. Werren JH. 2011 Selfish genetic elements, genetic conflict, and evolutionary innovation.
966 *Proc. Natl. Acad. Sci.* **108**, 10863–10870. (doi:10.1073/pnas.1102343108)
- 967 40. Presgraves DC. 2010 The molecular evolutionary basis of species formation. *Nat. Rev.*
968 *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 JJ. 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)

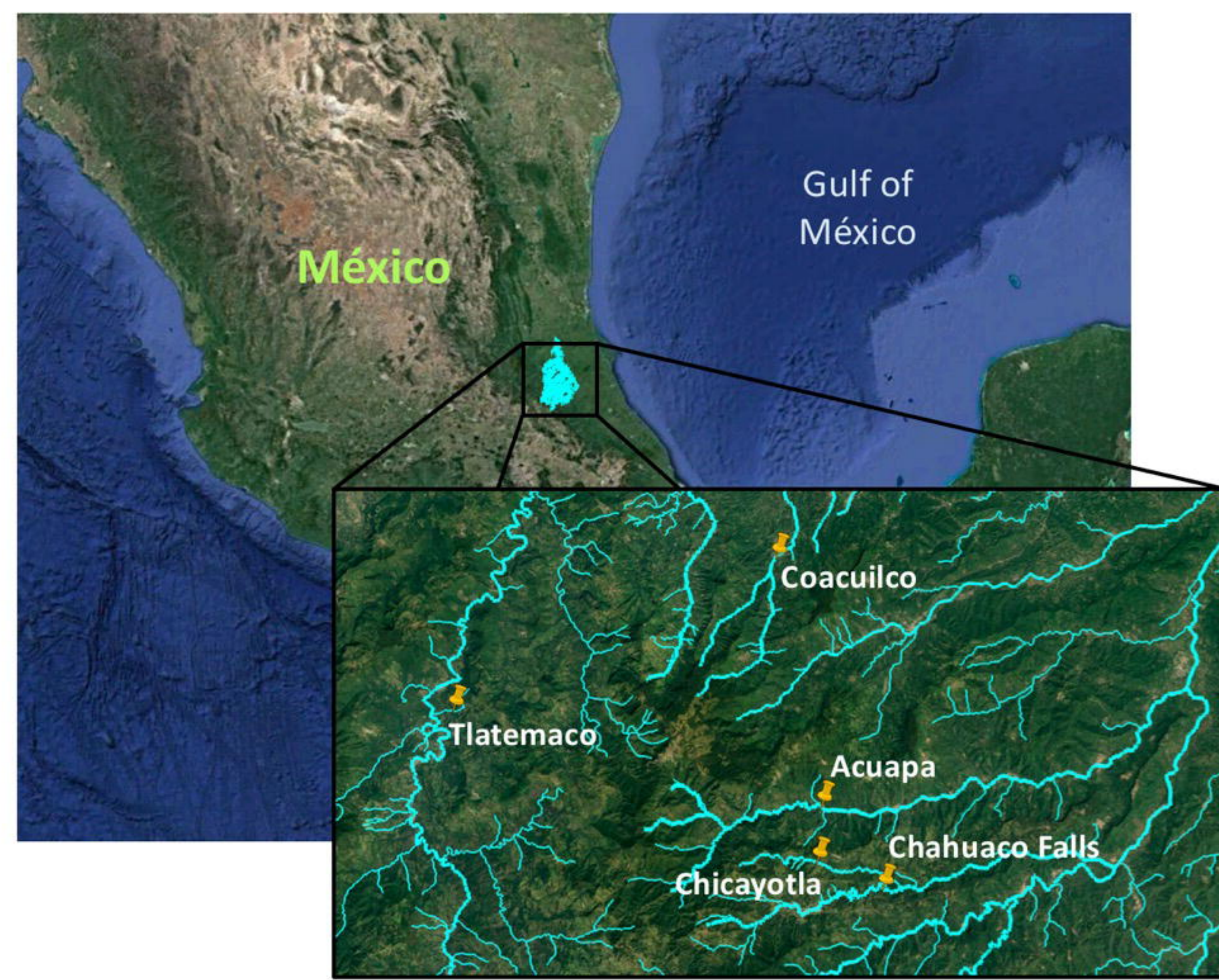
- 1072 80. Bray NL, Pimentel H, Melsted P, Pachter L. 2016 Near-optimal probabilistic RNA-seq
1073 quantification. *Nat. Biotechnol.* **34**, 525–527. (doi:10.1038/nbt.3519)
- 1074 81. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion
1075 for RNA-seq data with DESeq2. *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)

- 1105 95. Depienne C *et al.* 2017 Genetic and phenotypic dissection of 1q43q44 microdeletion
1106 syndrome and neurodevelopmental phenotypes associated with mutations in ZBTB18 and
1107 HNRNPU. *Hum. Genet.* **136**, 463–479. (doi:10.1007/s00439-017-1772-0)
- 1108 96. Walter P, Ron D. 2011 The Unfolded Protein Response: From Stress Pathway to
1109 Homeostatic Regulation. *Science* **334**, 1081–1086. (doi:10.1126/science.1209038)
- 1110 97. Oeckinghaus A, Ghosh S. 2009 The NF-kappaB family of transcription factors and its
1111 regulation. *Cold Spring Harb. Perspect. Biol.* **1**, a000034. (doi:10.1101/cshperspect.a000034)
- 1112 98. Zhang H-X *et al.* 2013 Rig-I regulates NF- κ B activity through binding to Nf- κ B1 3'-UTR
1113 mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6459–6464. (doi:10.1073/pnas.1304432110)
- 1114 99. Kawai T, Akira S. 2007 Signaling to NF- κ B by Toll-like receptors. *Trends Mol. Med.* **13**,
1115 460–469. (doi:10.1016/j.molmed.2007.09.002)
- 1116 100. Saxena M, Yeretssian G. 2014 NOD-Like Receptors: Master Regulators of Inflammation
1117 and Cancer. *Front. Immunol.* **0**. (doi:10.3389/fimmu.2014.00327)
- 1118 101. Glaser FT, Stanewsky R. 2005 Temperature Synchronization of the Drosophila Circadian
1119 Clock. *Curr. Biol.* **15**, 1352–1363. (doi:10.1016/j.cub.2005.06.056)
- 1120 102. Nagashima K, Matsue K, Konishi M, Iidaka C, Miyazaki K, Ishida N, Kanosue K. 2005
1121 The involvement of Cry1 and Cry2 genes in the regulation of the circadian body temperature
1122 rhythm in mice. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **288**, R329–R335.
1123 (doi:10.1152/ajpregu.00395.2004)
- 1124 103. Low KH, Lim C, Ko HW, Edery I. 2008 Natural variation in the splice site strength of a
1125 clock gene and species-specific thermal adaptation. *Neuron* **60**, 1054–1067.
1126 (doi:10.1016/j.neuron.2008.10.048)
- 1127 104. Blair EJ, Bonnot T, Hummel M, Hay E, Marzolino JM, Quijada IA, Nagel DH. 2019
1128 Contribution of time of day and the circadian clock to the heat stress responsive transcriptome
1129 in Arabidopsis. *Sci. Rep.* **9**, 4814. (doi:10.1038/s41598-019-41234-w)
- 1130 105. Ikeda R *et al.* 2019 REV-ERB α and REV-ERB β function as key factors regulating
1131 Mammalian Circadian Output. *Sci. Rep.* **9**, 10171. (doi:10.1038/s41598-019-46656-0)
- 1132 106. Gustafsson ALS, Skrede I, Rowe HC, Gussarova G, Borgen L, Rieseberg LH,
1133 Brochmann C, Parisod C. 2014 Genetics of Cryptic Speciation within an Arctic Mustard,
1134 *Draba nivalis*. *PLOS ONE* **9**, e93834. (doi:10.1371/journal.pone.0093834)
- 1135 107. García-Lara S, Khairallah MM, Vargas M, Bergvinson DJ. 2009 Mapping of QTL
1136 Associated with Maize Weevil Resistance in Tropical Maize. *Crop Sci.* **49**, 139–149.
1137 (doi:10.2135/cropsci2007.06.0326)
- 1138 108. Brekke TD, Stroud JA, Shaw DS, Crawford S, Steele KA. 2019 QTL mapping in salad
1139 tomatoes. *Euphytica* **215**, 115. (doi:10.1007/s10681-019-2440-3)

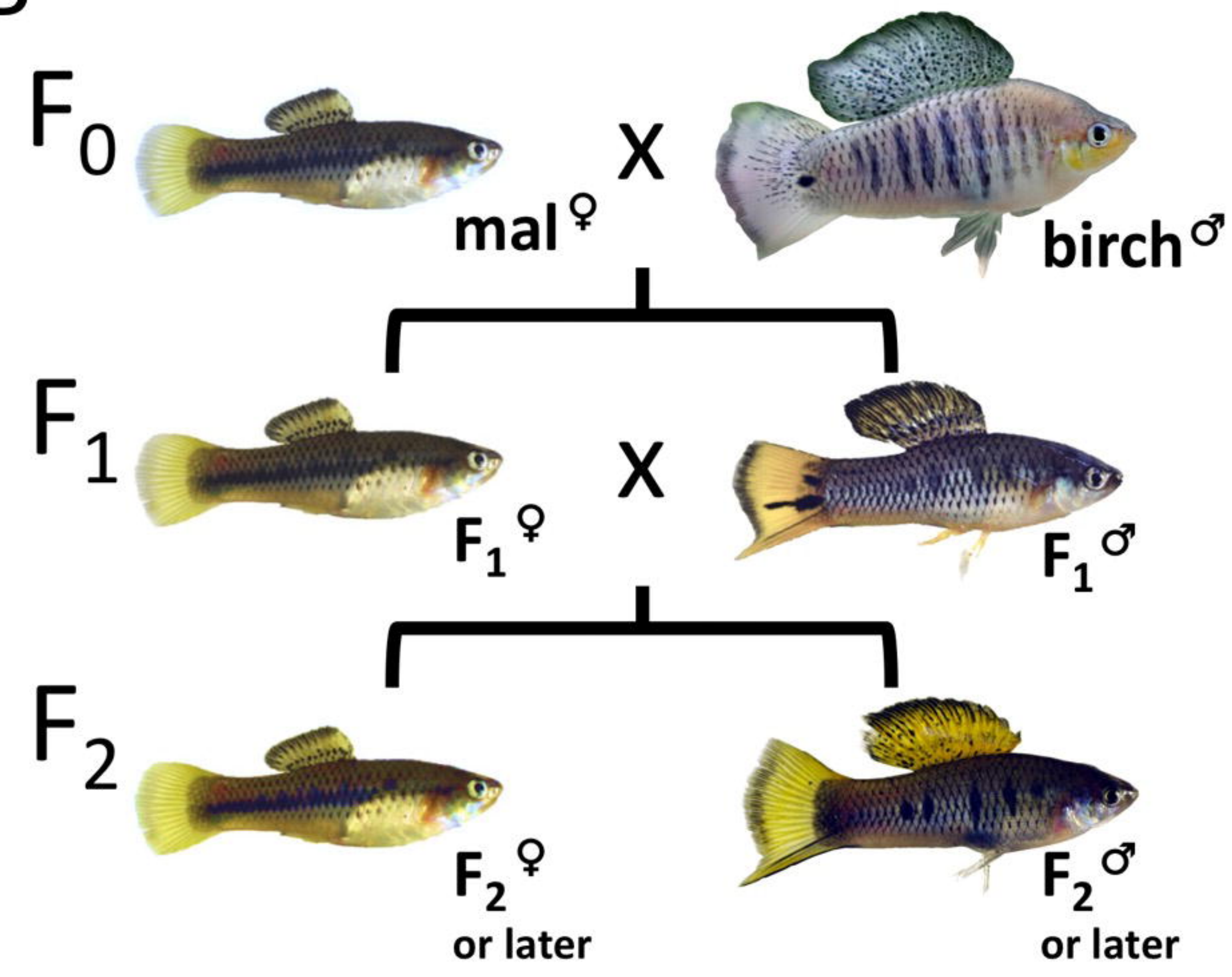
- 1140 109. Kenney-Hunt JP *et al.* 2008 Pleiotropic Patterns of Quantitative Trait Loci for 70 Murine
1141 Skeletal Traits. *Genetics* **178**, 2275–2288. (doi:10.1534/genetics.107.084434)
- 1142 110. Turner LM, White MA, Tautz D, Payseur BA. 2014 Genomic Networks of Hybrid
1143 Sterility. *PLOS Genet.* **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

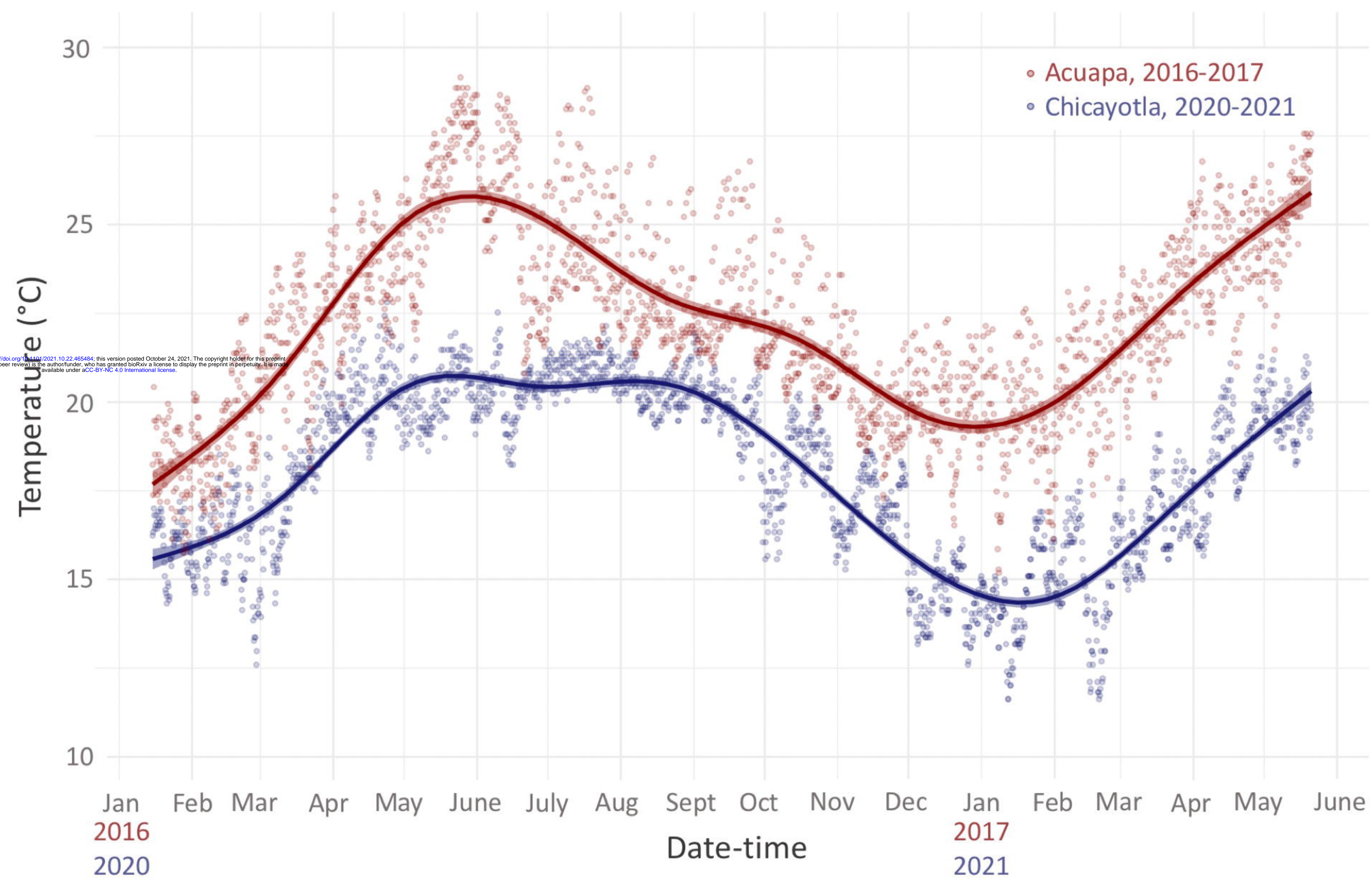
A

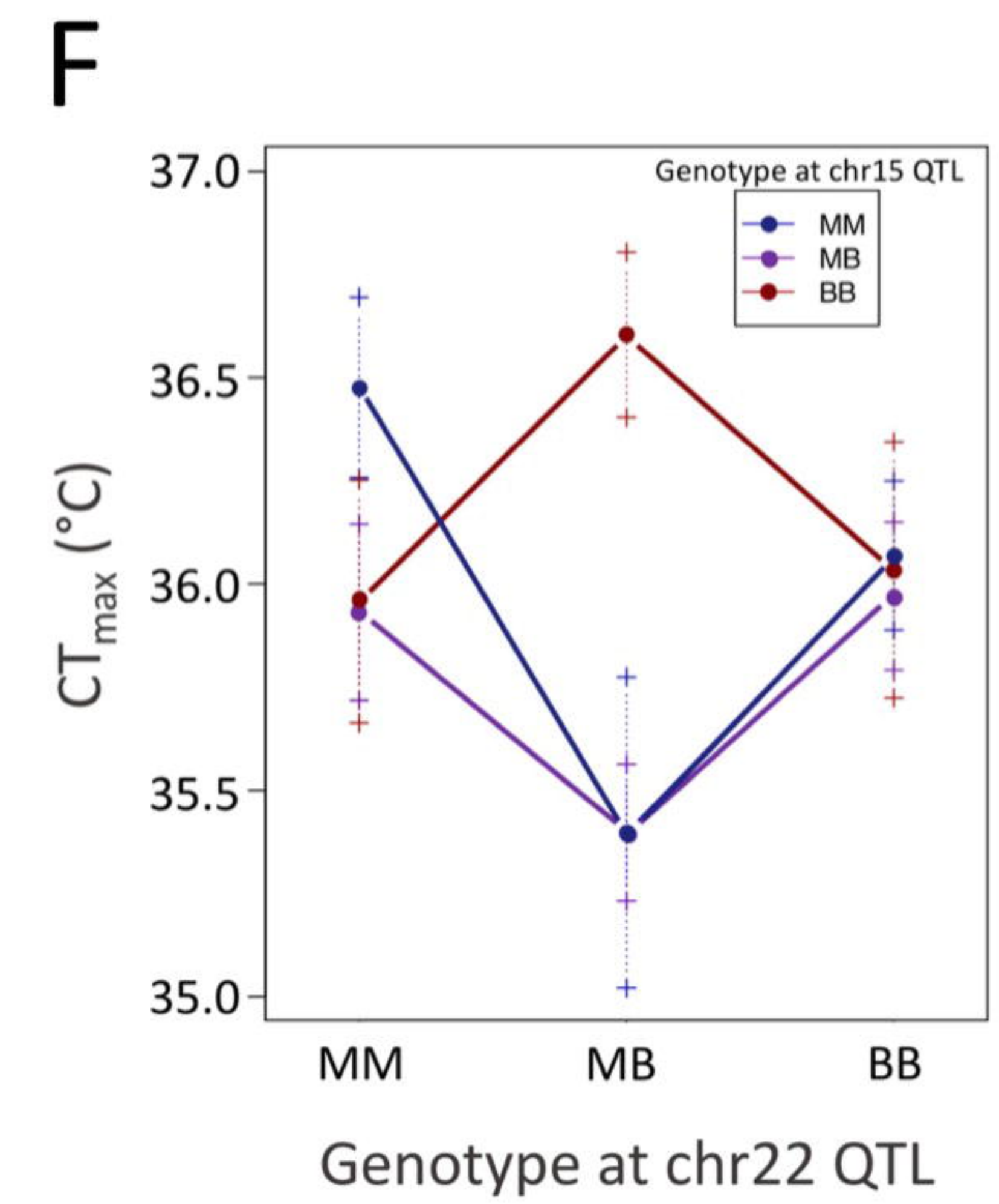
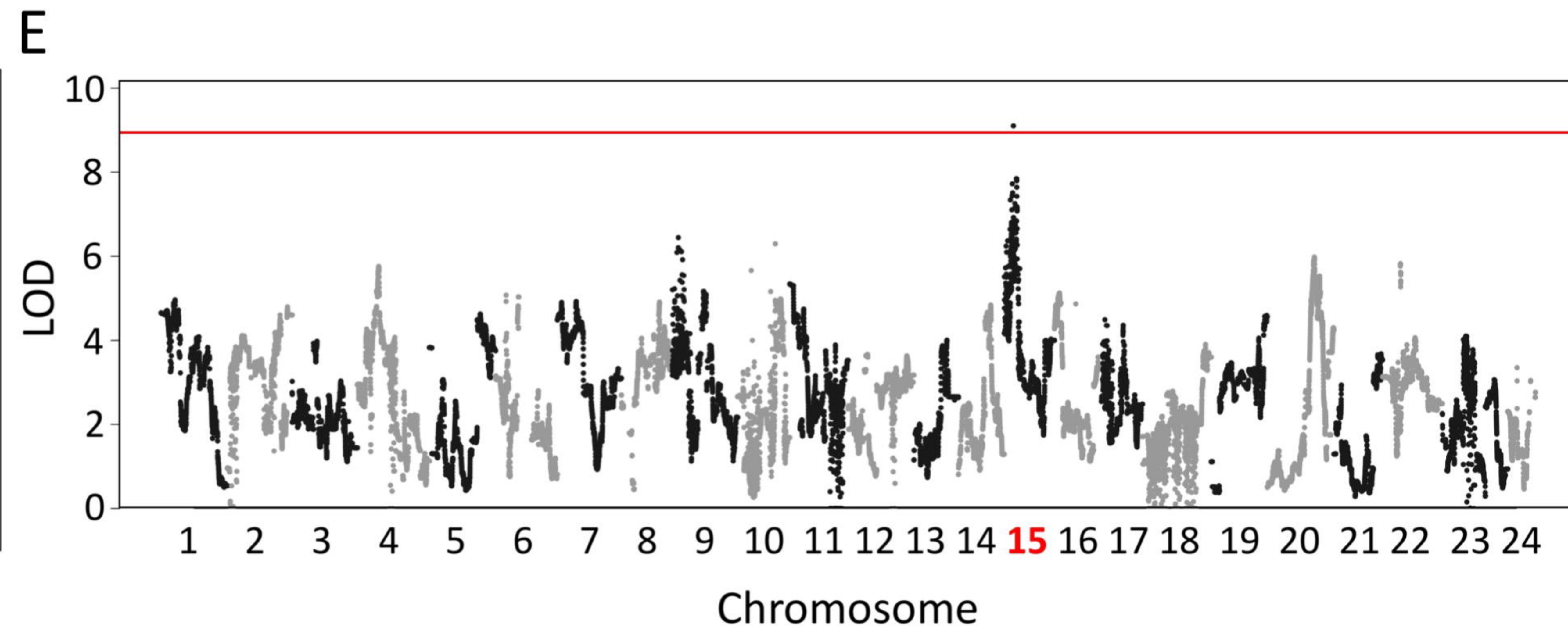
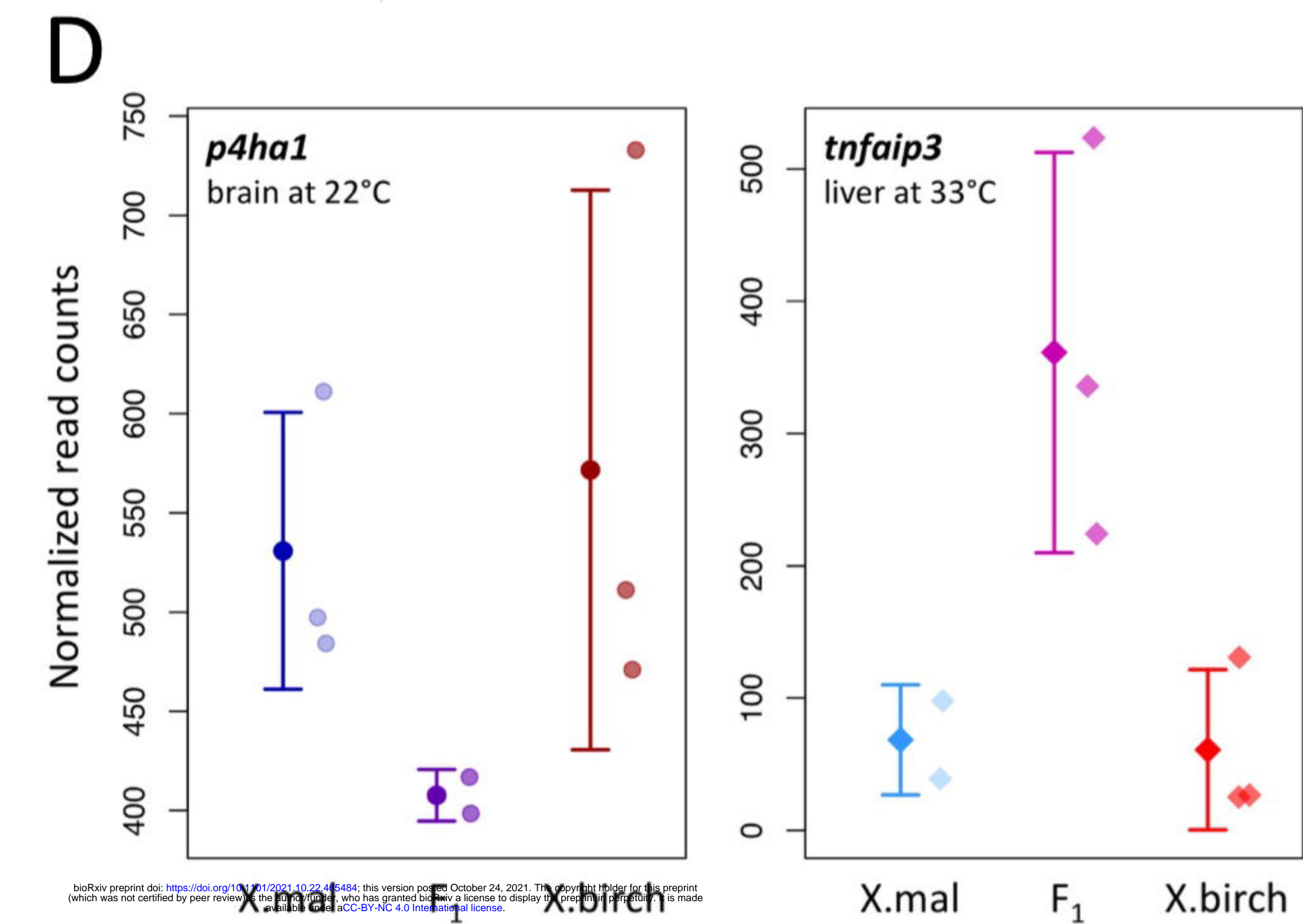
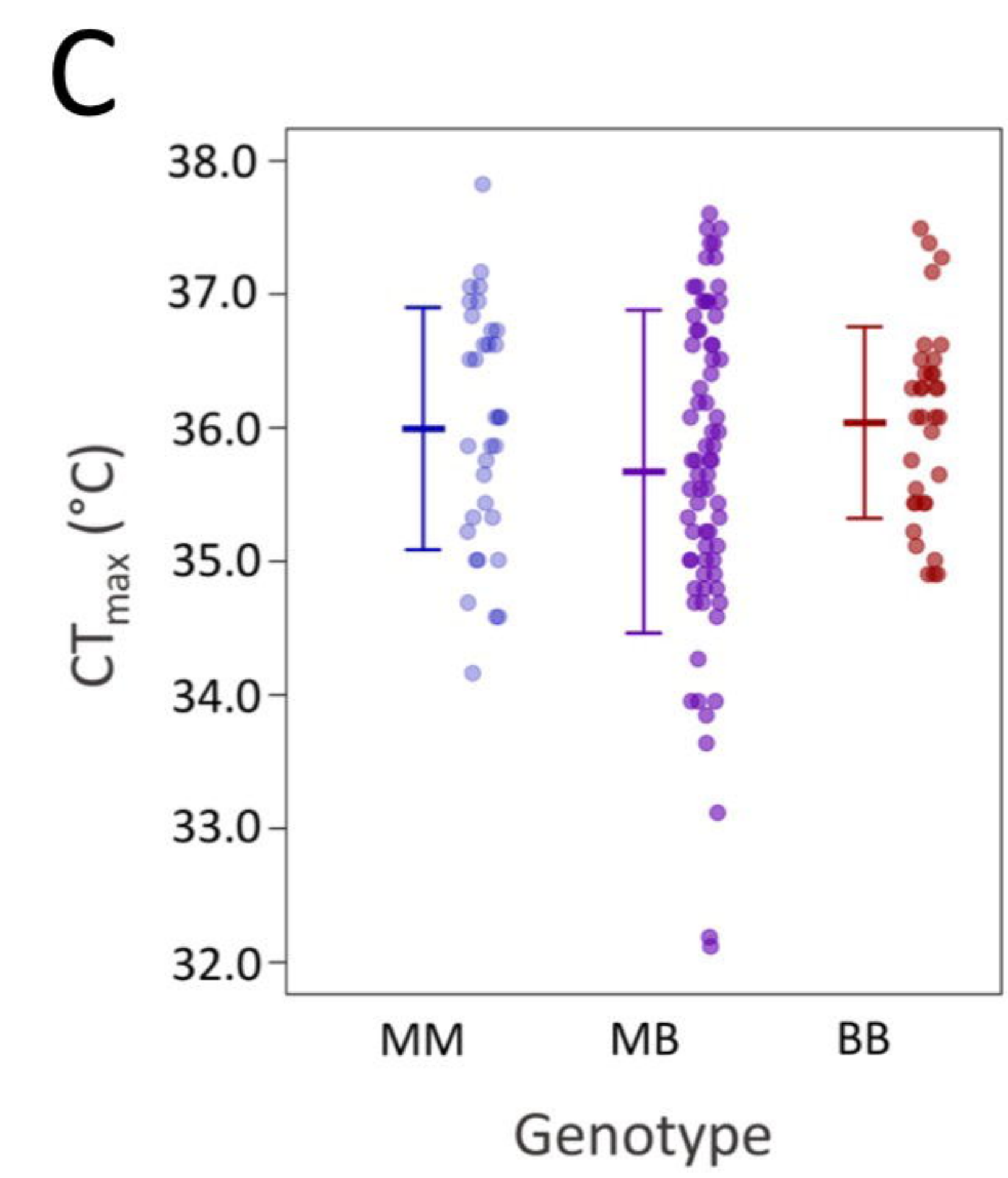
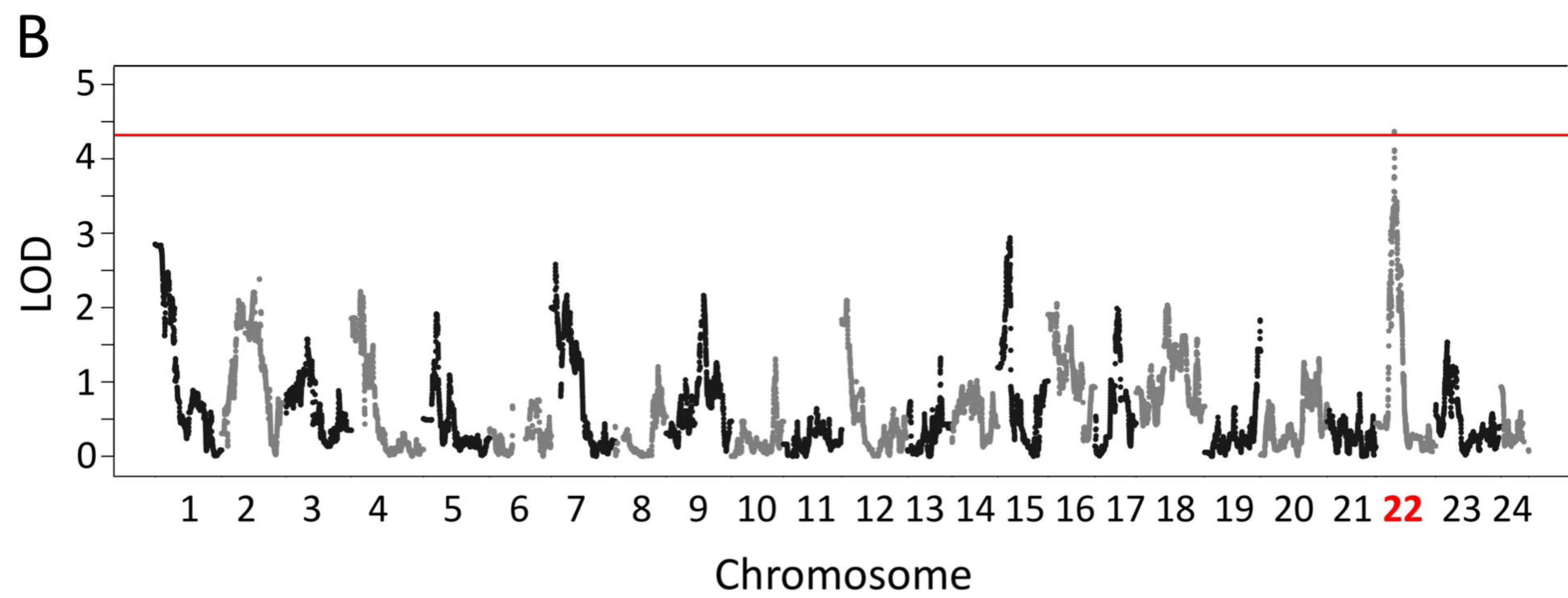
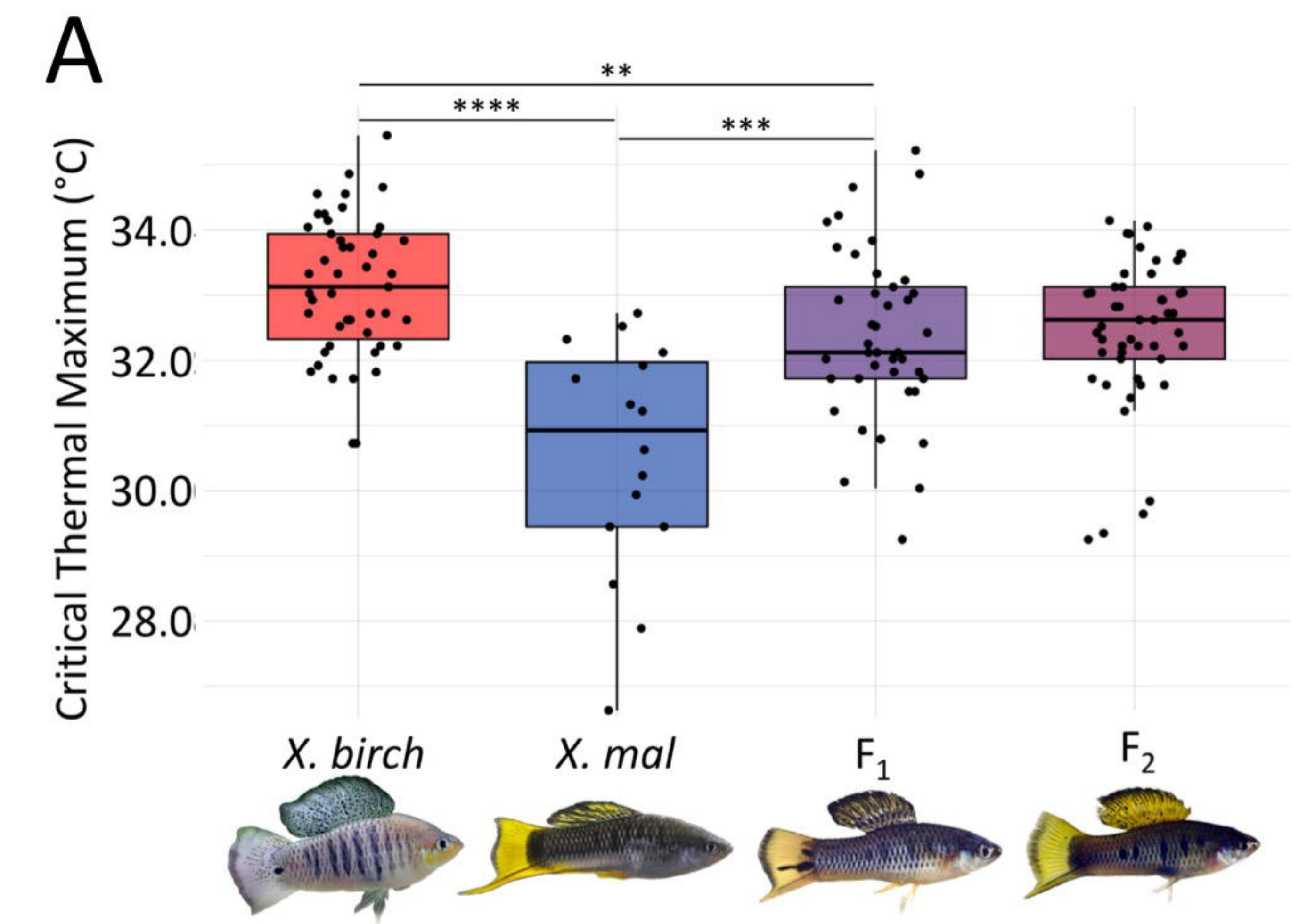


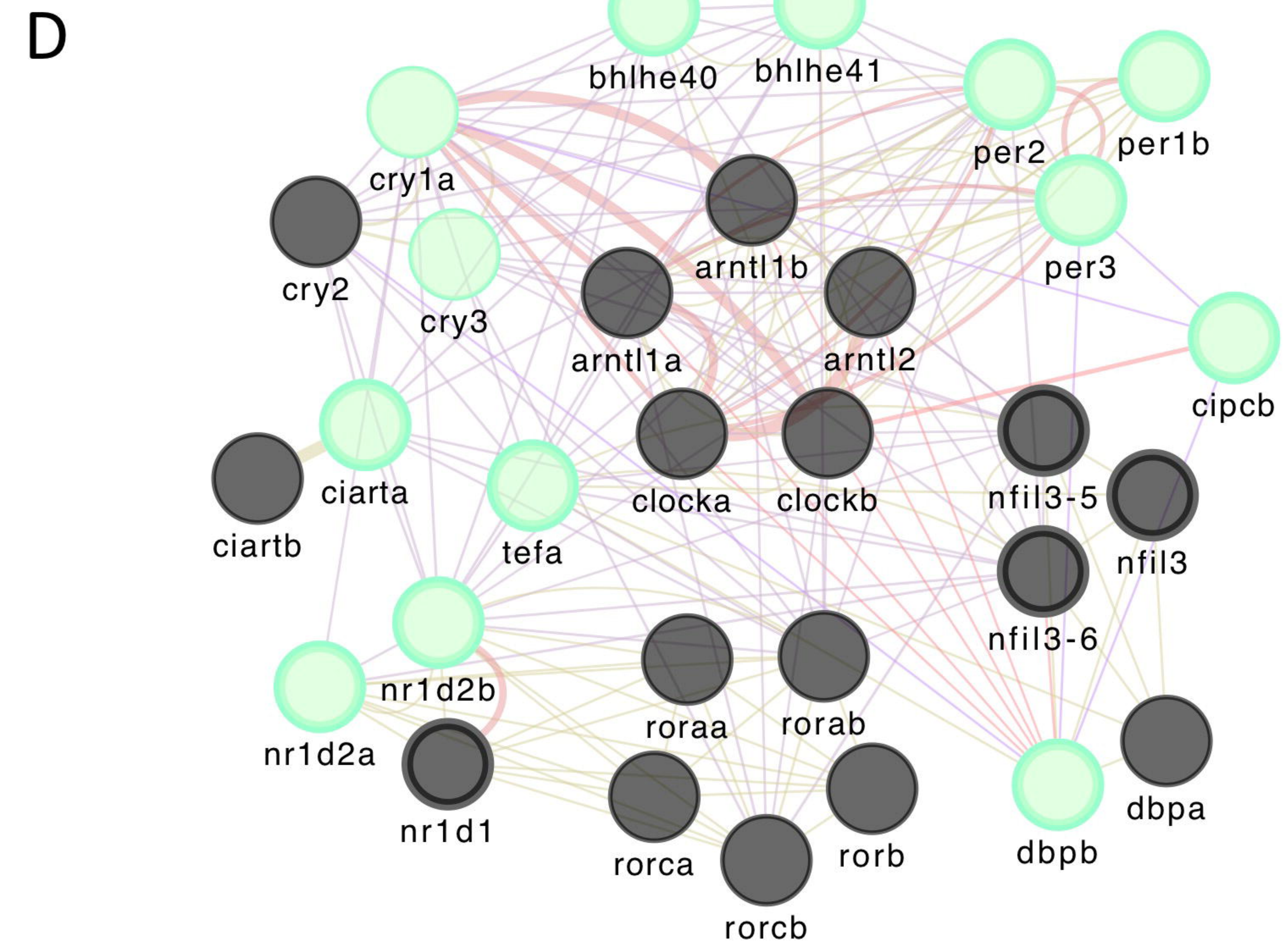
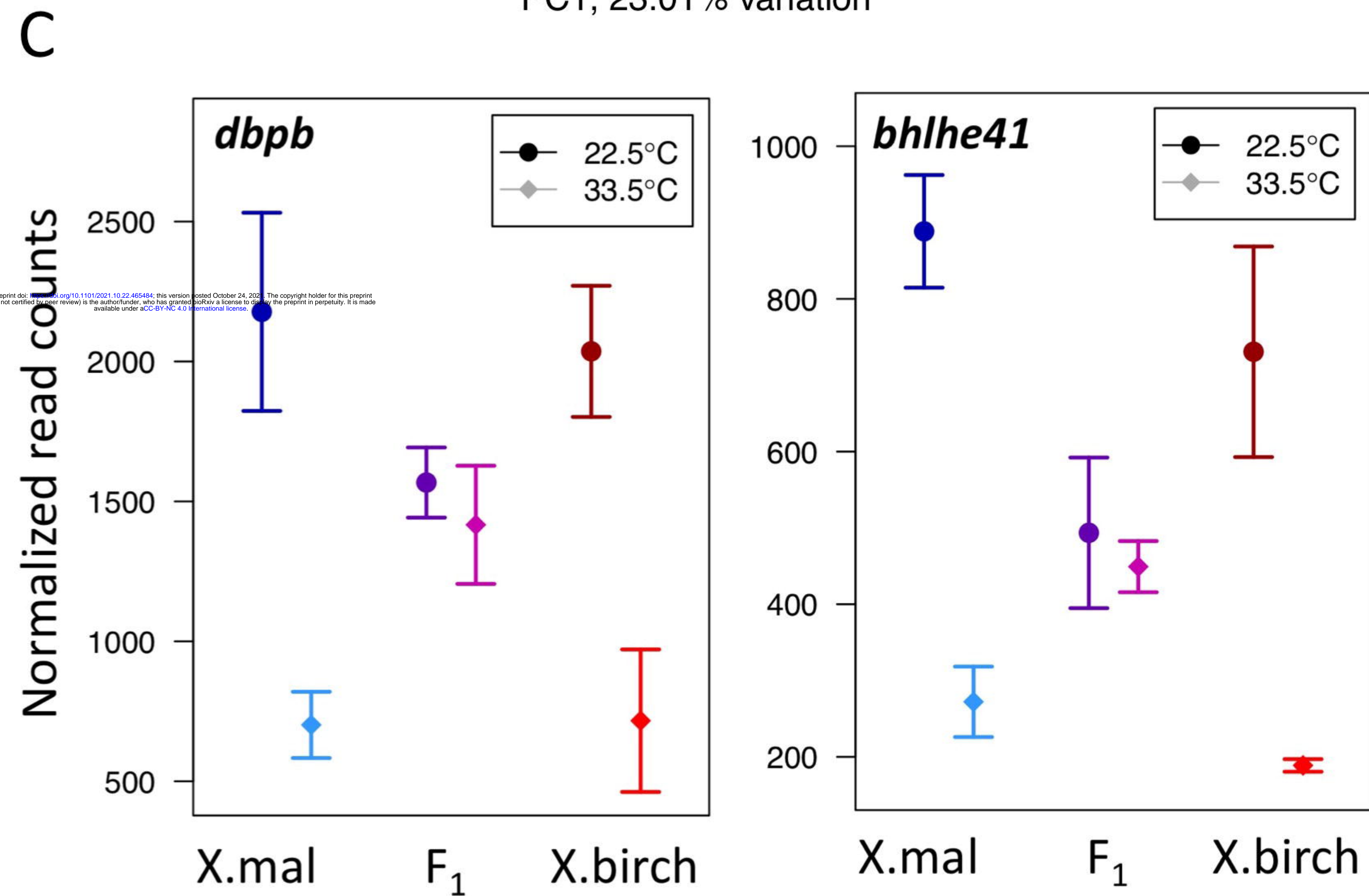
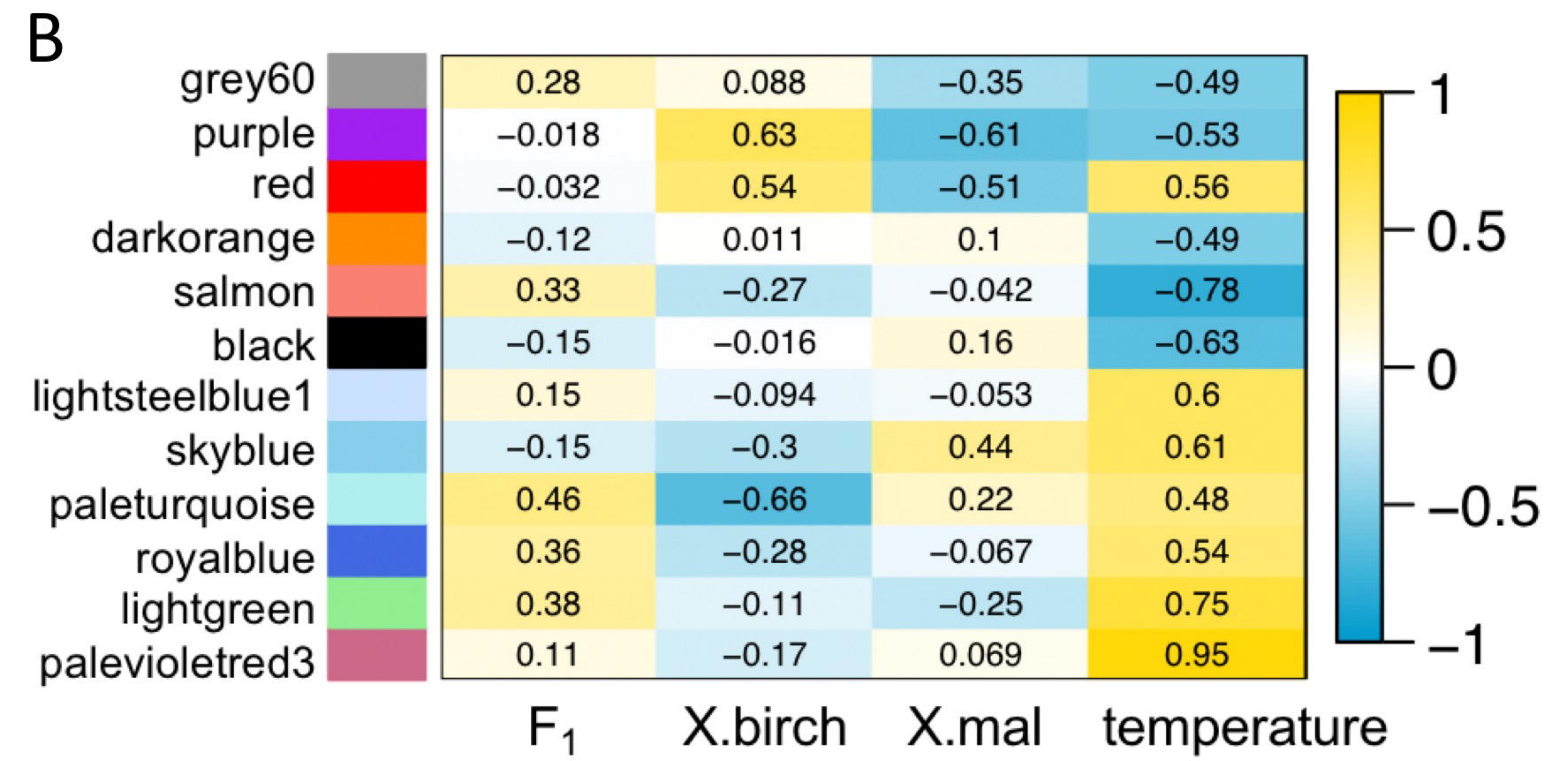
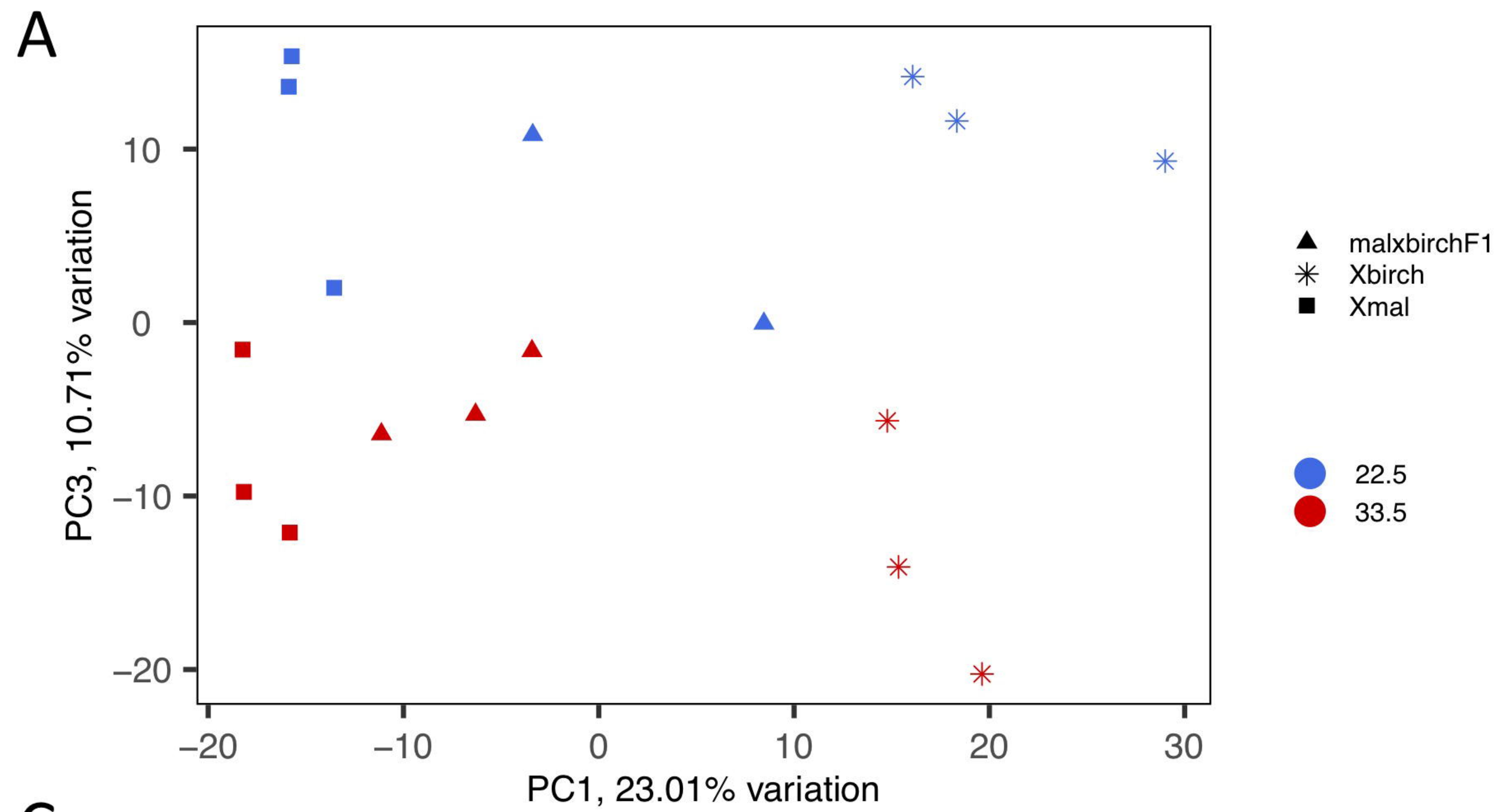
B



C

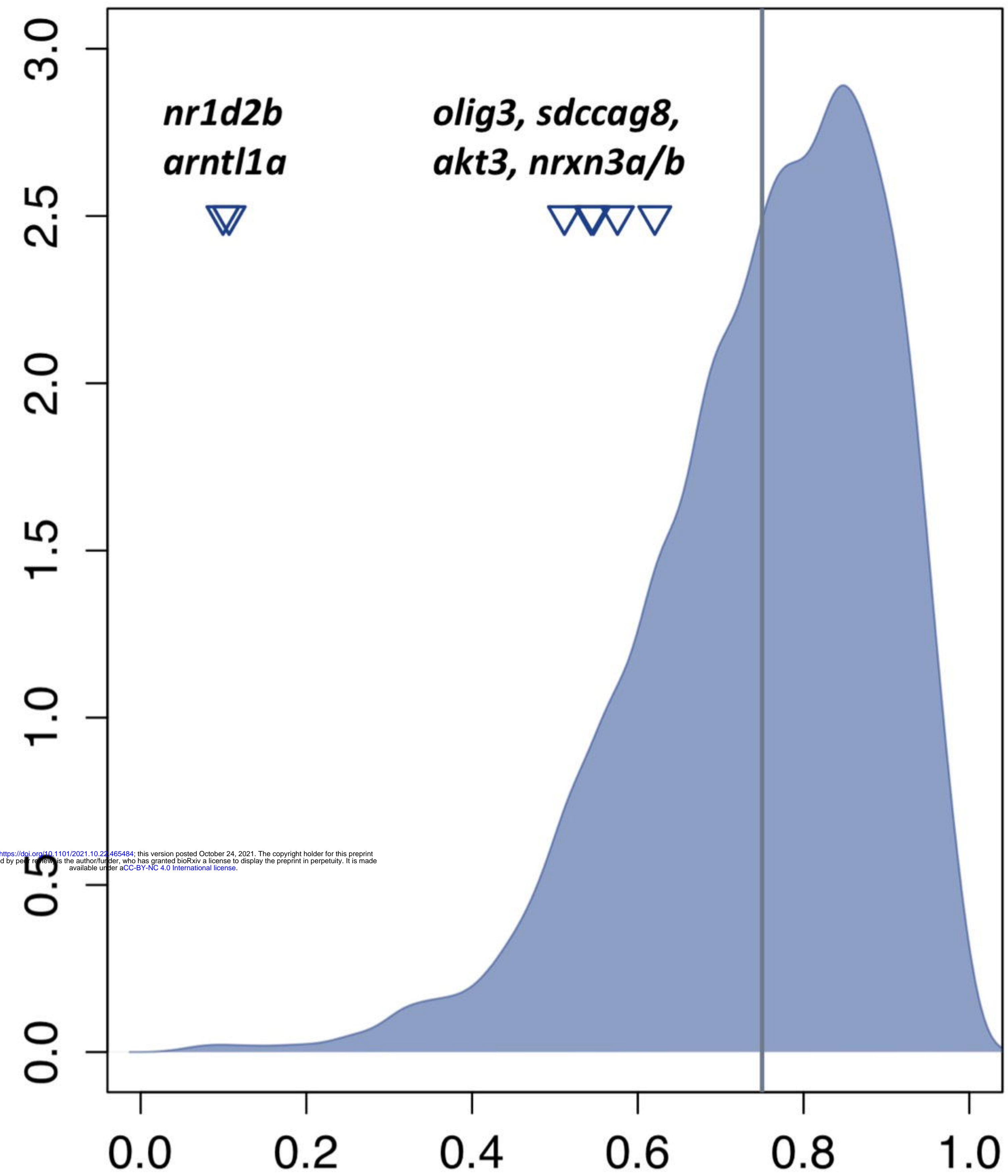




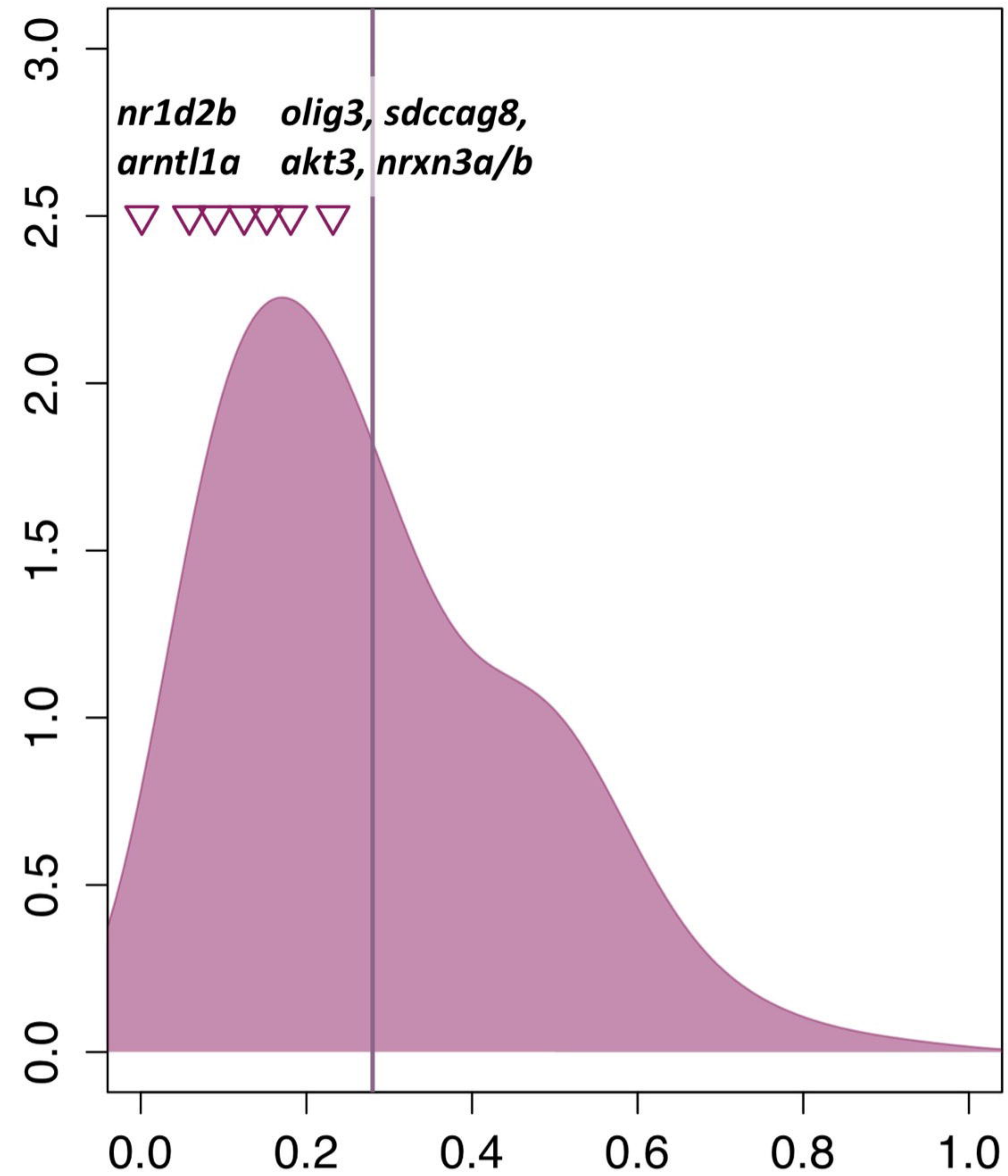


A

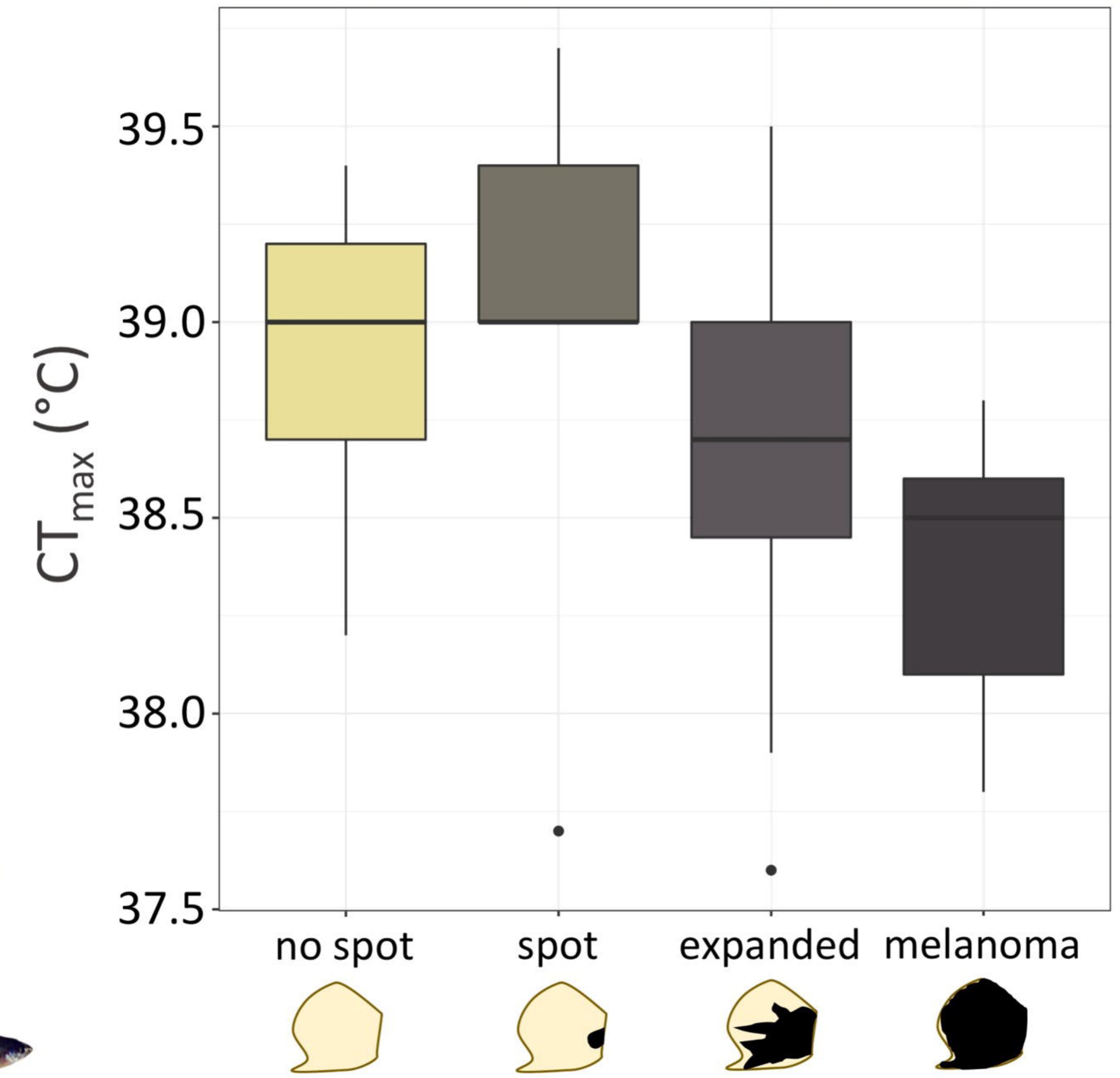
Tlatemaco



Acuapa



B



bioRxiv preprint doi: <https://doi.org/10.1101/2021.10.22.465484>; this version posted October 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

