Early life-stage thermal resilience is determined by climate-linked regulatory variation

Joaquin C. B. Nunez*, Sumaetee Tangwancharoen[♮], Kylie M. Finnegan[Φ], Luke A. Proud, and Brent L. Lockwood*

Dept. of Biology, University of Vermont, Burlington, Vermont, USA

*Corresponding authors: Brent.Lockwood@uvm.edu, Joaquin.Nunez@uvm.edu

[♮]Current address: Chulalongkorn University, Bangkok, Thailand
Φ Current address: Dept. of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

Abstract:

Despite decades of research in environmental change, we know relatively little about the genetics of environmentally influenced traits across the life cycle of species with complex life histories. Previously, we reported that natural variation in heat tolerance is life-stage specific in *Drosophila melanogaster*, suggesting that thermal selection predominantly targets the early embryonic life stage. Here, we used advanced introgression and pooled whole-genome resequencing to map the genomic basis of enhanced embryonic heat tolerance in a neotropical line of *D. melanogaster*. We identified two loci on chromosomes 2R and X that were consistently targeted by 16 generations of thermal selection across six replicate introgressions. We compared alleles in these regions to published datasets of natural variation from North America and Europe using the DEST dataset. This analysis revealed that two SNPs associated with embryonic heat tolerance exhibited both clinal and seasonal patterns, with the seasonal variation significantly correlated with environmental variability in average precipitation and temperature variance across space and time. The two SNPs lie in the putative regulatory regions of the genes *SP70* and *sog*, and individuals with different alleles at these loci exhibited disparate gene expression responses to heat stress. Overall, our results suggest that loci that influence embryonic heat tolerance are under selection in nature. Our study extends previous work in developmental genetics of *Drosophila* by characterizing the genomics of an ecologically relevant developmental trait in natural populations.

Significance Statement:

By comparing populations from distinct environments, we can determine the factors important for responses to environmental change. We used laboratory-based selection and whole-genome sequencing to uncover the genomic basis of embryonic heat tolerance in tropical and temperate fruit flies. We compared our results to published genomes of wild-collected flies from multiple continents and seasons. Strikingly, variants of two genes that influence embryonic heat tolerance are correlated with precipitation and temperature in natural populations. Further, different genetic variants produce distinct patterns of gene expression in response to heat stress. This is strong evidence that embryonic heat tolerance is under selection in nature and should be considered when forming predictions about responses to environmental change in species with complex life cycles.

Introduction

For nearly a century, evolutionary biologists have worked to describe how populations become differentiated when powerful forces, like purifying selection and migration, work to homogenize them (1–5). This fundamental question has led to a substantial body of work, including investigations of local adaptation and spatially varying selection, which constitute some of the strongest evidence that selection structures natural populations (6–9). Species that span environmental gradients have been powerful model systems for the study of evolutionary and ecological genetics (10–12), in part because abiotic factors like temperature exert strong selective pressures that can be quantified (13–15) and, thus, the causative selective environments can be characterized.

The vinegar fruit fly, *Drosophila melanogaster*, is an excellent model for studying adaptation across ecological gradients. In addition to its well-established genomic tools and rich experimental literature, its natural populations possess characteristics that make them ideal for investigating the genetic basis of adaptation (16). For instance, *D. melanogaster* adults can overwinter in temperate regions, allowing for the establishment of resident populations that undergo local adaptation to their specific environments (17–22). This process contributes to the formation of adaptive clines along latitudinal and longitudinal gradients, many of which have been extensively studied (10, 23, 24). Among these, temperature clines are particularly well-documented (25), with notable examples along the east coasts of North America (26) and Australia (27).

Despite extensive research on the topic, uncovering the genetic basis of adaptation along temperature gradients remains challenging for several reasons. First, selection operates across both spatial and temporal scales. For instance, numerous studies have demonstrated *D. melanogaster*'s ability to rapidly adapt to seasonal changes through bursts of directional selection (i.e., a process known as adaptive tracking; (19, 21, 28–34). These adaptive dynamics are shaped by multiple, co-varying selection pressures across space and time, complicating the identification and experimental tracking of these changes (22, 35, 36). Second, thermal adaptation studies in *Drosophilids* show that genomic ancestry, rather than a shared selection history, is the primary determinant of adaptive potential (37). As a result, even when thermal tolerance genes are identified at one location, they may not play a role at other sites. Lastly, selection acts throughout the species' life history (38), influencing various stages of development and survival. Indeed, our previous work has shown that thermal adaptation, particularly heat tolerance, is life-stage dependent and primarily occurs during the early embryonic stage, with little correlation between embryonic and adult heat tolerance (39). This highlights the importance of assessing selection at the appropriate life stage to accurately measure its impact.

Given the complexity of this biology, identifying both the causative agents of selection and the genes that respond to them requires a comprehensive integration of ecological and evolutionary insights with rigorous experimental validation. This approach is crucial for addressing several key questions: What are the primary agents of selection? At which stages of life history does selection act? And, most importantly, what are the causal loci responsible for thermal adaptation? In this paper, we tackle these questions by combining classical quantitative genetics methods with genomics, transcriptomics, and simulations to uncover the genetic basis of embryonic thermal tolerance in flies. Specifically, we utilized an advanced introgression-backcrossing design (40–42) to map the genomic basis of embryonic heat tolerance in a neotropical line of *D. melanogaster* by comparing lines derived from Vermont, USA (**VT**; temperate habitat) and from Saint Kitts, Caribbean (**SK**; tropical habitat). Then, to identify potential candidate loci, we conducted population genetic simulations and integrated our

findings with previously published genomic datasets, including the *Drosophila* Evolution over Space and Time (**DEST**) dataset, which encompasses samples from multiple temperate populations worldwide (20, 22), as well as with published findings from the *Drosophila* Genomic Reference Panel (**DGRP**, (43)). To further explore putative functional patterns at these loci, we also quantified gene expression.

Our results identified regions on chromosomes 2R and X as the genetic basis for embryonic heat tolerance in *D. melanogaster*. Additionally, our analysis using the DEST dataset uncovered putative regulatory loci in the *SP70* gene on chromosome 2R, which encodes a serine-type endopeptidase, and the *sog* gene on chromosome X, which encodes a morphogen involved in dorsal-ventral patterning and gastrulation during embryogenesis. Transcriptomic data revealed that alleles of these loci were associated with differential expression to heat stress and that expression responses of these two genes to heat stress were anticorrelated. Notably, both mutations exhibited strong latitudinal clines in North America and showed significant footprints of seasonal selection, with consistent genetic patterns across habitats in both Europe and the Americas. Phenotypic analysis of DGRP lines also shows that these two loci influence heat tolerance, further supporting our findings. Based on these results, early embryonic development may underpin thermal adaptation in the wild. Overall, our study enhances our understanding of how local adaptation to thermal stress occurs and underscores the importance of considering spatial, temporal, and life-stage-specific factors in uncovering the genetic basis of adaptation.

Results

The environmental context of tropical and temperate flies

This study investigates the genetic architecture underlying embryonic heat tolerance in *Drosophila*, building on previously reported phenotypic differences in embryonic survival between tropical and temperate populations (39) (see **Fig. 1A-inset**). To address this, we conducted an adaptive introgression experiment involving reciprocal crosses between flies from Vermont (VT; a temperate habitat with low embryonic heat tolerance) and Saint Kitts (SK; a tropical habitat with high embryonic heat tolerance; **Fig. 1A**; see metadata in **Table S1**) over 16 generations (F16). Every other generation, ~1-h-old embryos were exposed to a heat shock treatment for 1 hour, causing approximately 80% mortality and selecting for heat-tolerant alleles. In the alternative generation, offspring were backcrossed to the VT background after each generation of experimental evolution (i.e., 9 generations of selection and 7 generations of VT backcrossing). In principle, this approach should produce a genome largely representative of the VT parental population, except for genomic regions contributing to embryonic heat shock tolerance.

We selected VT and SK as parental populations for our study for two primary reasons: first, they were the same populations analyzed by (39), facilitating direct comparison with prior findings; and second, because they originate from ecologically distinct environments along the east coast of the Americas, characterized by fundamentally different environmental pressures. Indeed, long-term climate data from the past 25 years highlights Vermont's colder, more humid environment compared to the warmer, relatively drier conditions of Saint Kitts. However, the major distinction between these locations lies in their ecological variability. Vermont experiences pronounced seasonal fluctuations, with high variance in temperature, humidity, and precipitation, whereas Saint Kitts maintains stable humidity and temperature levels throughout the year (**Figs. 1B-C**). Notably, both habitats show opposite patterns of precipitation across years (**Fig. 1D**). These patterns are particularly pronounced when we restrict the weather variation to just the period from May to November, the period when temperate flies are active. (i.e., for the

remainder of the year, temperate populations enter a state of metabolic quiescence to survive the colder months; see blue bands in **Figs. 1B-C**).

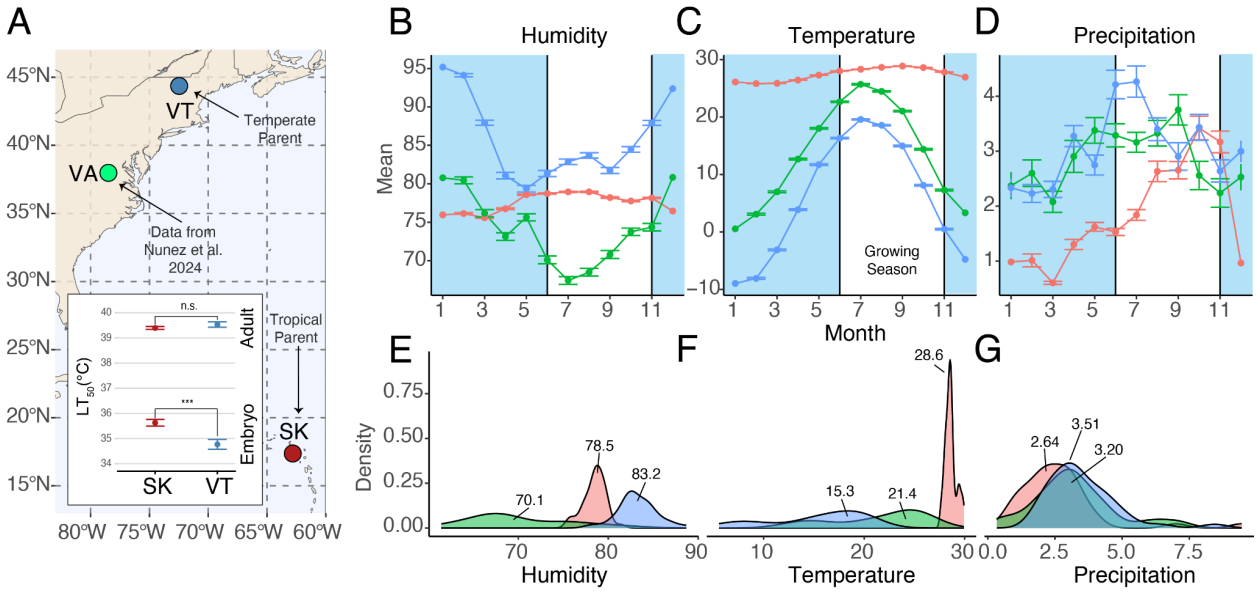

Figure 1. The ecological context of temperate and tropical *Drosophila*. (A) Map of the eastern seaboard of North America indicating the location of the parental tropical (SK) and temperate (VT) lines. The location of samples from Virginia (VA) from Nunez et al. 2024 is also shown. **(A-inset)** LT_{50} of adults and embryos estimated from survival after acute heat stress, as reported in Lockwood et al. 2018. ***Wald test, P = 0.0003. **(B)** Average monthly humidity (%) over 25 years of data (2015-2025). The blue bands indicate the overwintering period of temperate *Drosophila*. **(C)** Same as B, but plot shows mean air temperature (°C) measured at 2 m. **(D)** Same as B, but plot shows mean precipitation (mm/day; bias-corrected including snow). **(E)** Distribution of mean humidity values (2015-2025) within the period of May-November (i.e., growing season in temperate habitats) for the three localities. **(F)** Same as E, but for temperature. **(G)** Same as E, but for precipitation. Mean values are shown in E-F.

Genome-wide patterns of genetic variation across parental and introgressed lines

After 16 generations of experimental evolution, the introgressed populations exhibited elevated embryonic heat tolerance (**Fig. S1**; Welch's t-test, P = 0.013). To investigate the genomic basis of this response, we performed pooled whole-genome sequencing (Pool-Seq; (44)) on six replicate F16 introgressed populations: in three, VT was the maternal parent, and in the other three, SK was the maternal parent. Sequencing was also performed on the original VT and SK parental strains. Sequencing results of both the parental pools and the six F16 introgression pools have deep levels of coverage (mean = 87.6X, sd = 14.1X) and high mapping quality scores (mean = 56.6, sd = 0.249). Considered individually (see **Fig. S2A**), the parental lines and the replicate pools 1 and 2 for F16 crosses have similar coverages (78X-82X). Pools from the third F16 replicate have slightly lower, yet deep, levels of coverage (66X-69X). Given pools were sequenced on individual flow cells, these coverage differences suggest a mild technical effect for the third replicate of our crosses. Yet, this is not a concern for downstream analyses since our analytical tools account for coverage differences by design (see materials and methods). The estimated levels of PCR and optical duplicates filtered during mapping was ~0.25% (sd = 0.03%). Combined, these results indicate that our sequencing efforts yielded high-quality data.

After applying a 5% minimum threshold for missing data, we discovered 865,020 SNPs segregating among all pools. Of these, 159,641 are found in 2L, 205,515 in 2R, 187,745 in 3L, 203,632 in 3R, and 108,859 in X. Analyses of nucleotide diversity (π) showed lower π in our parental lines (**Fig. S1B**; ANOVA; $F_{2,1085}$ = 63.12, P = 2.0×10^{-16}) relative to our introgressed lines. For example, parental inbred lines in VT and SK had π levels ranging from 0.10-0.11. In contrast, F16 lines had doubled the level of π (i.e., 0.239 and 0.186; for the F16 initiated with SK-mothers and VT-mothers, respectively). These data revealed two salient results: first, the SK

and VT parental lines harbor residual heterozygosity. Second, 7 generations of introgression did not return π to the original levels of the parental VT lines, a likely byproduct of the segregating residual heterozygosity from the initial SK x VT cross.

Patterns of adaptive introgression of the F16 lines

We conducted principal component analysis on all lines (F16 introgression and parents). We conducted this analysis across the genome and individual chromosomes. Our results reveal a general pattern in PC space, with PC 1 across all analyses differentiating the SK line from all other lines (VT and F16s). These findings are consistent with our backcrossing design, whereby the F16 offspring have primarily a VT genomic background, regardless of the maternal line used in the first cross (red and green points in **Fig. 2A**). PC 2, on the other hand, showcases differences among lines relative to the VT parental background. We observe that, for most chromosomes, and genome-wide, the F16 pools and the VT parental pool show high levels of genetic similarity as they cluster near each other in PC 2. Notably, this pattern is not observed in chromosome 2R, where both SK/VT parents and the F16 pools cluster separately in PC space (*see* panel for 2R **Fig. 2A**).

Next, to investigate whether particular regions of the F16 genomes show differential patterns of introgression, relative to the parental lines, we estimated local PCAs using a sliding window approach (window size = 0.1 Mb; step = 50 Kb). We then estimated the mean euclidean distance (i.e., d), using PCs 1 and 2, across F16 pools relative to each parent (i.e., d_{SK} and d_{VT}). Accordingly, the ratio of d_{SK} and d_{VT} is a proxy for genetic distance, whereby $\text{Log}_2(d_{SK}/d_{VT}) > 0$ indicates higher similarity to the VT parent and $\text{Log}_2(d_{SK}/d_{VT}) < 0$ indicates higher similarity to the SK parent (**Fig. 2B**). As expected, most of the genome shows high similarity to the VT parent. Notably, we observed 29 windows showing signatures of SK introgression (6 in 2L, 10 in 2R, 7 in 3R, and 6 in X; *see* red arrows in **Fig. 2B**).

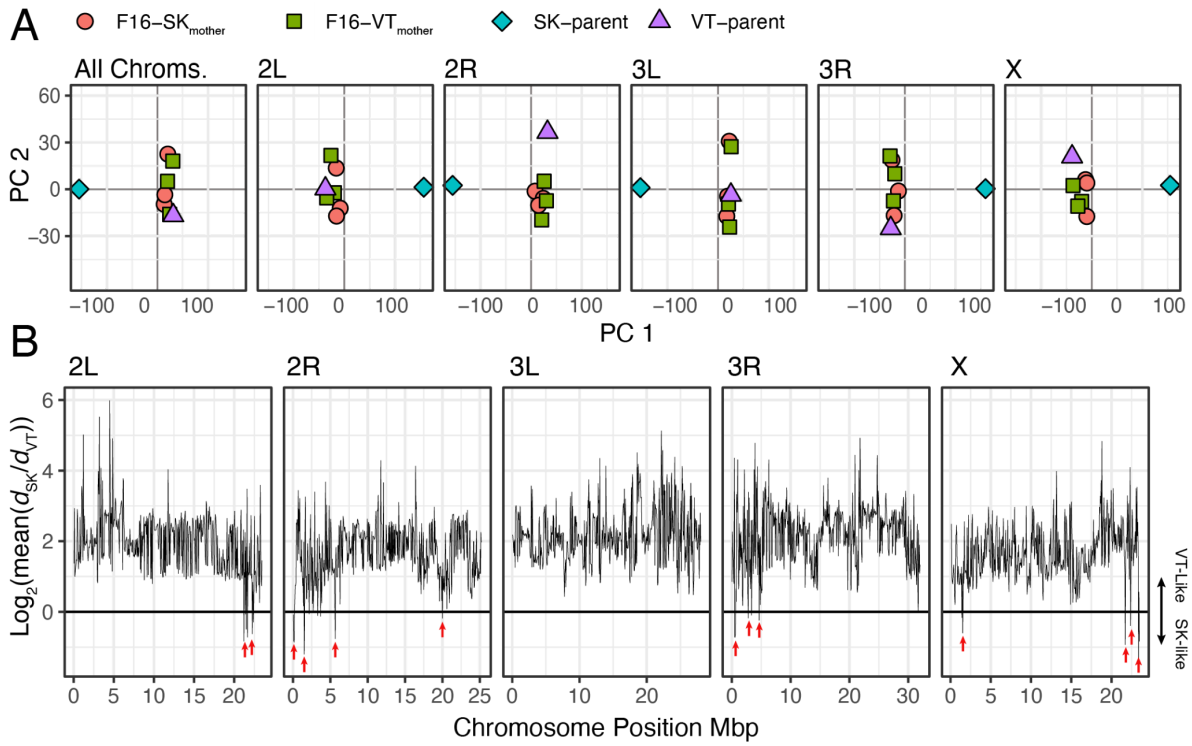

Figure 2. Genomic patterns of adaptive introgression. (A) Principal component analysis using parental and F16 lines. PCAs were estimated using all chromosomes, as well as for each chromosome separately. The percentage of variance explained (PVE) by PC 1, across analyses are 84.3% (All), 90.8% (2L), 85.9% (2R), 76.9% (3L), 85.4% (3R), 89.7% (X). The PVE by PC 2, across

analyses are 4.22 (All), 3.33 (2L), 6.28 (2R), 8.18 (3L), 5.65 (3R), 3.17 (X). **(B)** Estimates of genetic distance of F16 pools relative to the parental genomes of the VT and SK parent $Log_2(d_{SK}/d_{VT})$. The red arrows indicate regions of the genome with signatures of SK introgression.

The genomic basis of embryonic heat tolerance

We used a two-pronged approach to uncover the genetic basis of embryonic heat tolerance. First, we characterized levels of genetic differentiation ($\boldsymbol{F_{ST}}$) across pools compared to permutations. Second, we used a series of Fisher's Exact tests (**FET**) across our replicates compared to the pools with the parental background. We compared the results of these analyses to the output of population genetic simulations (using SLiM; (45)) parameterized based on real features from the genome and the recombination map of *D. melanogaster* as well as the allele frequencies of the parental lines (see *Materials and Methods*). We explicitly simulated the expected levels of genetic variation of our crossing introgression scheme under a neutral (i.e., drift only) scenario.

Combined, these approaches helped us narrow down regions of interest for embryonic heat tolerance. For the first analysis, we characterized levels of F_{ST} across windows with similar levels of recombination in all F16 replicates. We estimated the pairwise F_{ST} between each F16 pool (that started with VT or SK mothers) relative to the parental VT line. We compared the observed F_{ST} values relative to 95% of all simulations (**Fig. 3A; Fig. S3A**). Overall, we observe 23 windows across the genome where all F16 pools outperform 95% of simulations (Black rectangles in **Fig. S3A**). For example, we identified 2 windows in 2L, 8 in 2R, 2 in 3L, 1 in 3R, and 10 in X.

Next, we assessed whether any particular window across the genome was enriched for F_{ST} values in the top 1% of the empirical distribution. To do this, we ranked normalized all SNP-wise F_{ST} estimates across the genome and we calculated, using a binomial test (see *Materials and Methods*), whether any window harbors more SNPs with F_{ST} values in the top 1% of the distribution than expected by chance (i.e., under a null expectation of 1%). We also compared the output of this analysis to our neutral introgression simulations (**Fig. S3B**). Considering the output of these two analyses (F_{ST} vs simulations and Top 1% F_{ST} enrichment) reveals two regions, one in 2R and the other in X, that are consistently enriched in both tests across all F16 replicates. Accordingly, a 1.5 Mb window in 2R (19,115,753-20,615,753; **Fig. 3A** left), near the right breakpoint of the cosmopolitan inversion In(2R)NS, and a 1 Mb region in X (15,123,410-16,123,410; **Fig. 3A** right).

Next, we conducted an FET analysis on a "synthetic pool" created by aggregating all F16 samples, maximizing statistical power across replicates. This analysis identified 391 SNPs across the genome with significant FET *P*-values (0.01; Bonferroni corrected). However, these SNPs were not evenly distributed; most were concentrated on chromosome 2R (293 SNPs) and the X chromosome (67 SNPs), while no significant hits were found on 2L, and only 22 and 9 SNPs were detected on 3R and 3L, respectively.

Notably, the top FET hits aligned with the genomic windows identified in our F_{ST} analysis (purple bars in **Fig. 3B**). These regions also exhibited a high density of SNPs that were nominally significant (*P*-value < 1%) in individual FET analyses at the pool level (**Fig. 3C**). The windows of interest showed a pattern of high concordance for the FET analyses at the level of individual SNP level, with 5.6% on 2R and 10.5% on X, compared to just 0.76% across the rest of the genome. Even within their respective chromosomes, these windows displayed substantially higher concordance (expected for 2R = 0.88%, X = 0.65%). Furthermore, the window on 2R at 19.5Mb was also identified as a key region in the introgression local PCA analysis (see **Fig. 2B**). The strong agreement across our PCA-, F_{ST}-, and FET-based approaches further supports the regions in 2R and X as candidate loci for embryonic heat tolerance.

The window of interest in 2R, at 19.5Mb, contains 157 genes and harbors 12,949 SNPs out of which 197 were associated with heat tolerance. Of these, 20 are missense mutations, 55

are synonymous variants, and 21 occur in UTRs. On the other hand, the window in X at 15.5Mb, contains 60 genes and harbors 2,673 SNPs and 58 were associated with heat tolerance. Of these, 6 are missense mutations, 29 are synonymous variants, and 8 occur in UTRs (**Fig. 3D; Tables S2**). We performed GO enrichment analysis using outlier SNPs within key genomic windows on 2R and the X chromosome. In 2R, one significant term related to odorant binding (Q = 0.081) was driven by outliers in *Obp56/57* paralogs. Additional terms related to heat shock protein binding showed low P-values but did not meet the Q-value threshold. On the X chromosome, 10 enriched GO terms (Q < 0.1) were identified, including those associated with euchromatin regulation, tRNA ligase activity, and centriole assembly, largely driven by genes *Myb*, *AlkB*, *CG8097*, *ArgRS*, and *CCT6* (see **Text S1, Tables S3-S5**).

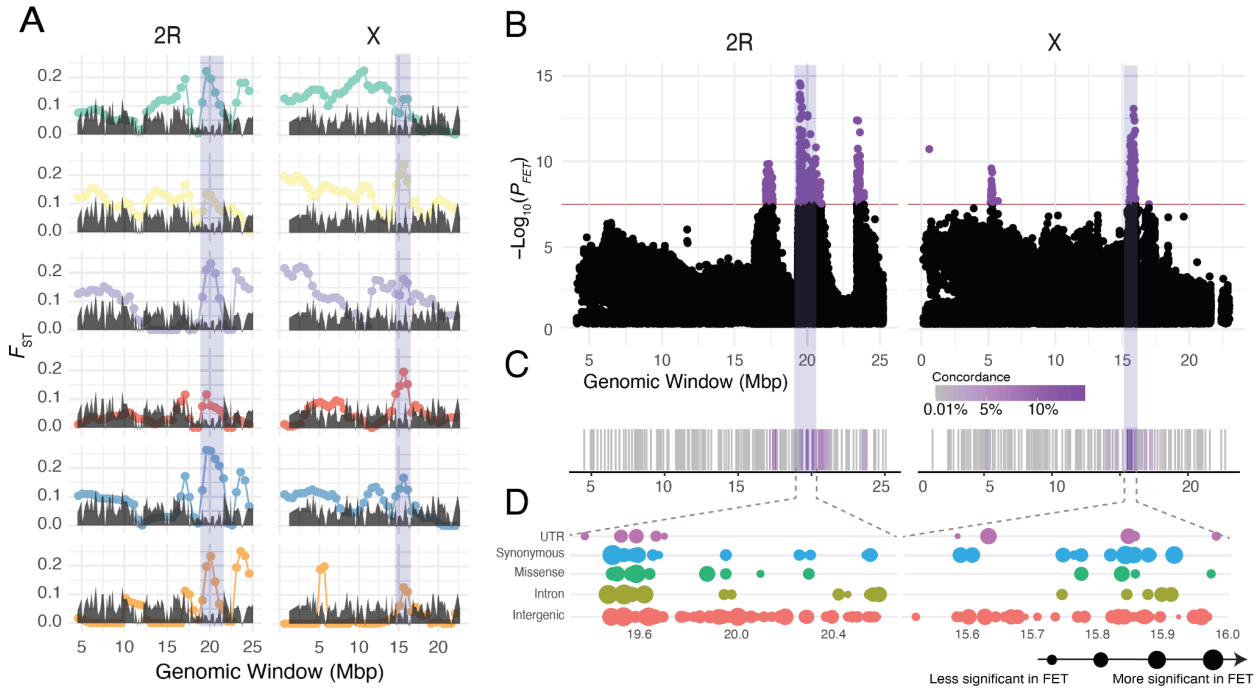

Figure 3. Genome scan for targets of embryonic thermal adaptation. (A) Sliding window analysis for F_{ST} outliers across our introgression crosses relative to neutral introgression simulations. **(B)** Fisher's exact test (FET) P-value (Bonferroni corrected) for the synthetic pool composed of all F16 replicates. Significant regions are shown in purple. The red line represents a threshold of ~10^{-8} (equivalent to a Bonferroni cutoff for 0.01). **(C)** The plot showcases the degree of per-SNP concordance among individual FET tests (for each F16 pool) across the genome. Regions where significant SNP co-occur with windows of interest identified in tests shown in panels A and B are shown with darker shades of purple. The highlighted vertical band between panels C and D show windows with SNPs showing 100% concordance across F16 pools. **(D)** Functional annotations for the SNPs within windows of high concordance.

Outlier co-localization with seasonal and latitudinal adaptive scans in autosomes

We compared our findings with previously published ecological genomic datasets (see *Materials and Methods*) to assess whether any SNPs within our candidate windows for embryonic heat tolerance had been reported as clinal (Q-value < 0.05; (30)) or as seasonal hits (P-value < 0.004; (19)). Because most ecological genetics studies primarily focus on autosomal chromosome variation, we first restricted our co-localization analysis to SNPs occurring in 2R. In total, we identified 11 SNPs associated with both clinality and embryonic thermal tolerance (**Fig. 4A**) and two SNPs linked to seasonality (**Fig. 4B**). Strikingly, one SNP, 2R: 20,551,633 (C→A on the forward strand; G→T on the reverse strand), is significant in both clinal and seasonal studies (**Fig. 4C**) as well as in our embryonic study. To underscore the rarity of this finding, we calculated the degree of overlap among all genome-wide outlier loci across three datasets.

Remarkably, 2R: 20,551,633 is the only mutation across the entire genome that is shared among all three datasets, highlighting the potential biological significance of this overlap.

In our experimental lines, we observed that the SNP is fixed for the "C" allele in the SK parent and that it is polymorphic in the VT parent, yet A is the major allele ($AF_{C\text{-allele}}$ = 0.598). As for the F16s, the introgressed pools show a mean allele frequency of the A allele of 0.18. This observed frequency is significantly lower than the expected allele frequency ($p_{expected}$ = 0.59; Δp = 0.41) based on the predicted dynamics of back-crossing introgression lines into the VT genetic background. In the absence of selection, the expected level of introgression after seven back-crosses is minimal, such that F16 pools would be expected to exhibit an allele frequency shift of approximately 0.021% (see p. 406 of (41)). Thus, the observed allele frequency shift of 41.7% is strong evidence that this locus on chromosome 2R played a major role in embryonic heat tolerance during adaptive introgression.

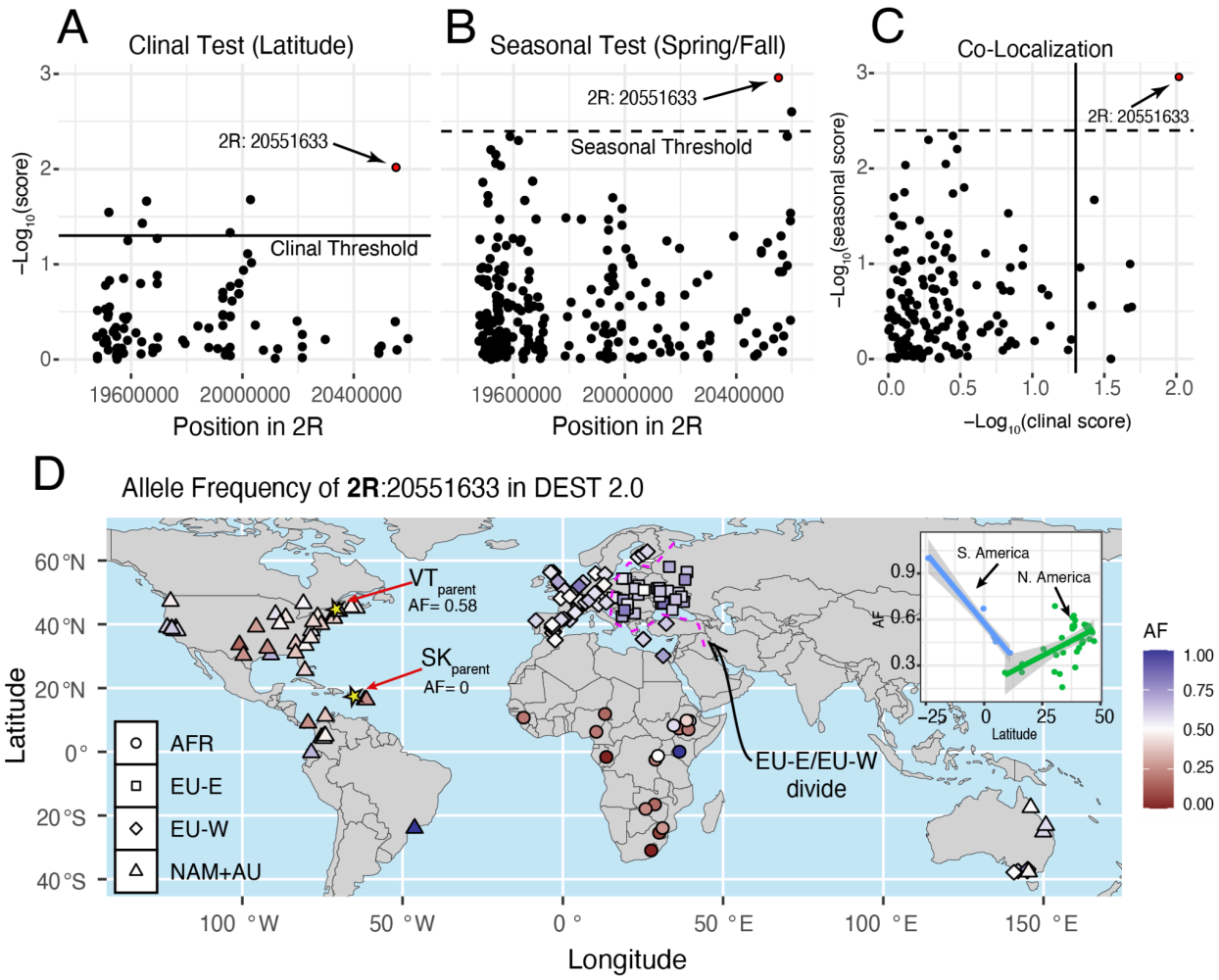

Figure 4. Footprints of seasonality and clinality at the window of interest in 2R. (A) Colocalization analysis of outlier SNPs in the FET analysis with clinal SNPs from (30), the horizontal line is the significance threshold indicated by the authors. **(B)** Colocalization analysis of outlier SNPs in the FET analysis with seasonal SNPs from (19), the horizontal line is the significance threshold indicated by the authors. **(C)** SNPs that are jointly FET outliers, seasonal outliers, and clinal outliers. **(D)** Global allele frequencies at the top SNP in 2R (position 20,551,633; frequency of A shown) from the DEST 2.0 dataset**.** Shapes indicate known demographic clusters in *D. melanogaster,* including the EU-E EU-W divide reported in (20)**. (D-inset)** Patterns of allele clinality in North (green) and South (blue) America.

Seasonal and clinal patterns for the 2R chromosome

We used allele frequency data from the DEST 2.0 dataset (22) to explore the patterns of latitudinal variation across North America and Europe (**Fig. 4D**). Our results show significant patterns of clinality in the Americas. For example, in North America, the allele has a positive correlation with latitude (r = 0.540, P-value = 0.0012; **Fig 4D-inset**, see green line). In contrast, the allele is negatively correlated with latitude in South America (r = -0.9816, P-value = 8.70×10^{-5}; **Fig 4D-inset**, see blue line). These observations indicate that, in the Americas, the A allele is more common in temperate habitats, a result that is consistent with the mutation frequencies in our experimental lines. We also investigated the spatial patterns of variation of our top SNP across non-American habitats. At the continental level, there is a marked difference between Europe and Africa. In Europe, the A allele is predominantly the major allele across the continent (mean AF = 0.601), whereas in Sub-Saharan Africa, the ancestral habitat of the species, the A allele is the minor allele across the continent (AF = 0.358). Within Europe, we detected no correlation with latitude (r = -0.0146, P = 0.878) but a significant correlation with longitude (r = 0.226, P = 0.035).

To further examine seasonal selection patterns at our top SNP in 2R, we performed a locus-specific test for seasonality (see *Materials and Methods*). Using data from a densely sampled population from Virginia (USA; **Fig 1A,** green dot**;** (21)), we applied a generalized linear model to assess the changes in allele frequency in relation to three environmental variables (i.e., temperature, humidity, precipitation). In this analysis, the environmental variables were summarized across time windows ranging from 7 days to 90 days (3 months) prior to sample collection. We compared the results of these models to 100 random permutations. Our results indicate that the variance of temperature, 0-45 days prior to collection ($T_{SD:0\text{-}45}$) is the best predictor for our mutation of interest ($\beta_{seasonal}$ = -0.089, $P_{seasonal}$ = 0.0053). Furthermore, this model beats 96% of permutations (**Fig. 5A**).

The timeframe of this analysis (~45 days) suggests that populations where the A allele is at low frequency are likely descendants of flies that experienced high temperature variation at least three generations prior (i.e., experience selection favoring the C allele). To visualize these patterns, we plotted the relationship between allele frequency and $T_{SD:0\text{-}45}$ for the Virginia population, another temperate habitat similar to Vermont (**Figs. 1B-G**). Additionally, we examined two other populations from DEST with high-resolution seasonal sampling: Odessa, Ukraine (2013-2021) and Yesiloz, Turkey (2015-2021). Our results show that allele frequencies vary in relationship to $T_{SD:0\text{-}45}$ in Charlottesville (r = -0.2985, P = 0.072; **Fig 5B**), and Odesa (r = -0.4627, P = 0.026; **Fig 5C**), but not in Yesiloz (r = -0.038, P = 0.82; **Fig 5D**).

We visualized the seasonal dynamics of our top SNP on 2R by plotting weather data (**Fig. 5E**) relative to the estimates of allele frequencies (**Fig. 5H**) from Virginia. These plots reveal two notable behaviours in the DEST data. First, it showcases changes in allele frequency following the onset of the late-fall cold-snaps (i.e., multiple days with T < 5°C; **Fig. 5G**), yet the direction of these changes is not consistent across years. Second, the SNP shows a consistent pattern of summer selection whereby the A allele decreases in frequency (**Fig. 5H**; also see **Fig. S4**) following exposure to heat stress (i.e., multiple days with T > 32°C, as per the thermal limit model (19); see **Fig. 5F**). Note that this is the same direction of allele frequency change that we observed in our adaptive introgression, as described above. Overall, these data suggests that 2R:20,551,633 (or a closely linked variant) is likely an ecologically important mutation.

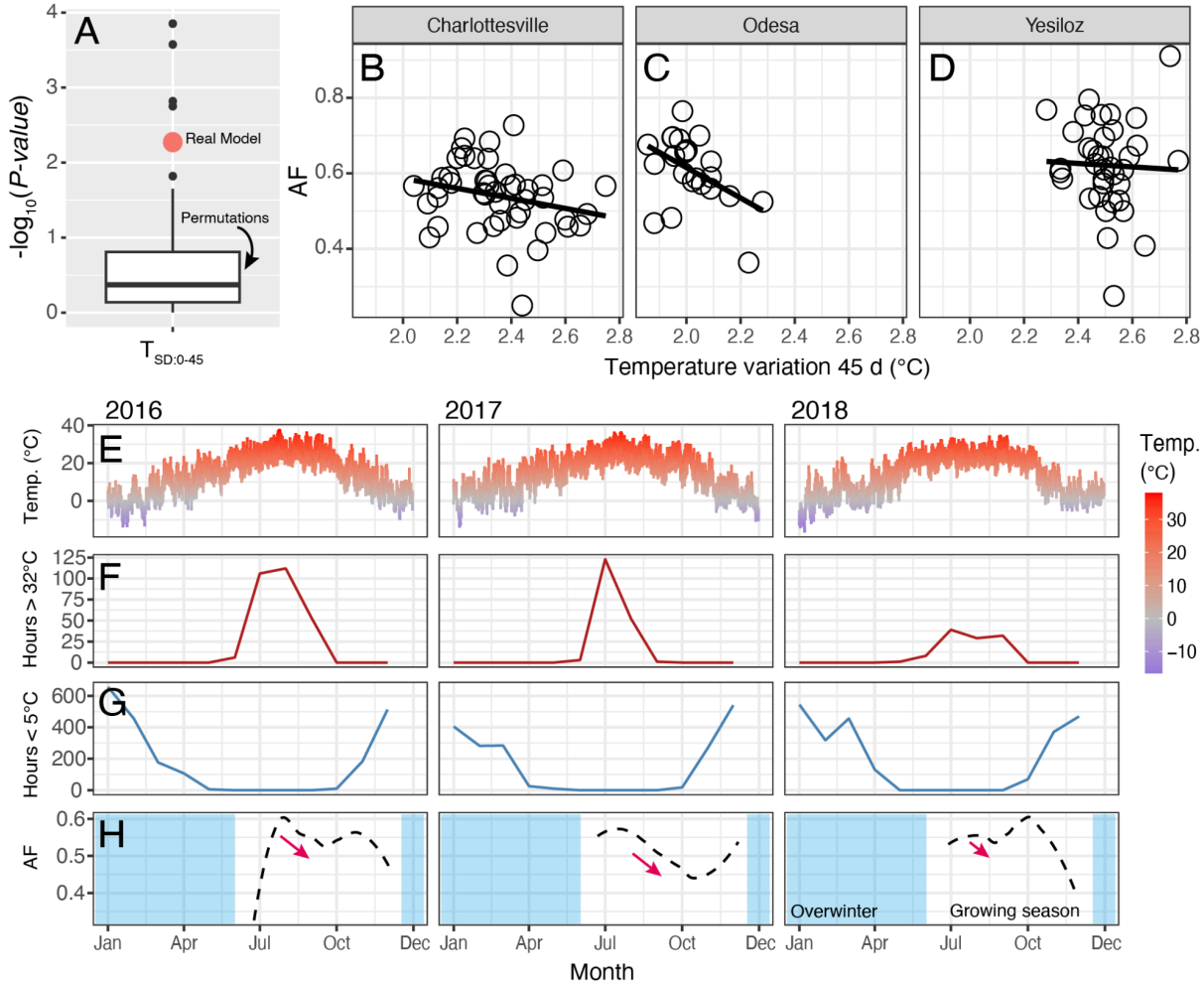

Figure 5. Dynamics of seasonal selection at the top SNP in 2R. (A) Result of the generalized linear model at the top SNP in 2R using genomic data from Virginia derived from Nunez *et al.* 2024. The real estimate for the locus is shown in red, and the outcome of 100 permutations is captured by the boxplot. The data shown represents the model summarizing temperature variance 0-45 days prior to collection. **(B)** Allele frequency relative to temperature variance 0-45d in Charlottesville, Virginia. **(C)** Same as B, but in Odesa, Ukraine. **(D)** Same as B, but in Yesiloz, Turkey. **(E)** Daily mean temperature in Charlottesville, Virginia from 2016-2018. This weather data is matched to the genomic data from Nunez *et al.* 2024. **(F)** Number of hours per month above 32°C. This is a proxy for summer-like conditions. **(G)** Number of hours per month below 5°C. This is a proxy for winter-like conditions. **(H)** Allele frequency trajectories of the A allele across the growing seasons 2016-2018. The overwintering period is shown in blue.

Seasonal and clinal patterns for the X chromosome

Given that ecological scans often overlook variation in non-autosomal regions, we next used data from Charlottesville, VA, to perform ecological association analyses on the region of interest on the X chromosome, with the goal of identifying outlier loci associated with both clinal and seasonal variation in North America (see *Materials and Methods*). Our analyses identified six loci exhibiting significant signals in both the seasonal and clinal test (**Fig. S5A**). Evaluating the pattern of introgression in our F16 pools revealed that three of these six loci showed clear patterns of adaptive introgression (positions: 15,602,941; 15,607,604; and 15,847,814; **Fig. S6**). Notably, these outliers occur in a region of reduced VT ancestry. While the outlier region on the X chromosome did not cross zero, it mirrored the directional trend toward SK ancestry observed on 2R, suggesting a shared underlying pattern of introgression (see **Fig. S5B**). Of these three loci, mutation X:15,602,941 emerges as a candidate locus as it shows both strong signals of seasonal evolution ($\beta_{seasonal}$ = 0.0846, $P_{seasonal}$ = 0.0153 [beats 97% of permutations]; best model: precipitation 0-90 days prior to collection), and significant correlations with latitude in both North

America (r = -0.4716, P = 0.0048) and Europe (r = -0.2156, P = 0.0502). This locus also shows a longitudinal pattern in Europe (r = -0.1852, P = 0.0936; **Fig. S5D**). The other two loci do not show clinal patterns in Europe.

The candidate mutation in X showed two alleles (A/G), with "G" fixed in SK but only present at low frequencies (~0.05) in VT. In the introgression experiment, the locus increased to an average of 0.44 (shift due to selection: Δp = 0.39). To put these patterns in an ecological context, we explored the allele frequency dynamics in the VA seasonal data relative to the environmental variable ($Precip_{max.}$0-90d). Our analysis revealed a positive association between the frequency of the G allele and maximum precipitation in Virginia (r = 0.430, P = 0.0035; **Fig. S5B;** A similar trend was observed in Odesa (**Fig. S5C)**, though lacking statistical significance). Furthermore, we observed that for our NASA weather measurements (2015-2025), in both VT and SK, precipitation and temperature show positive, albeit weak, correlations (SK, r = 0.087, $P < 2.2\text{x}10^{-16}$; VT, r = 0.101, $P < 2.2\text{x}10^{-16}$), suggesting that, at the collection sites of our parental lines, increases in precipitation are often accompanied by increases in temperature. In summary, while the association with thermal selection is less evident for X:15,602,941, the observed clinal and seasonal variation at this locus, paralleling that of 2R:20,551,633, suggests it has ecological relevance.

The evolutionary context of our top SNPs in 2R and X

To frame the findings from our adaptive introgression experiment, we integrated findings from ecological genomics analyses using the DEST dataset. Through this approach, we identified two high-confidence SNPs within our focal regions (2R:20,551,633 and X:15,602,941) that display both seasonal and clinal variation, making them compelling targets of selection for embryonic thermal tolerance in the wild. In the case of 2R, the tropical allele ("C") is the ancestral allele relative to four outgroup species (*D. simulans, D. yakuba, D. sechellia, and D. erecta*) and it is the major allele in Africa (**Fig 4D**). For the case of X, the temperate allele ("A") is the ancestral allele relative to two outgroup species (*D. simulans,* and *D. sechellia*) and is the major allele in Africa (**Fig. S5D**), though, the tropical (neotropical; "G") state is reported for *D. yakuba*, and *D. erecta*. This is a notable finding because, on 2R, the temperate allele predominates in Europe and the tropical allele is common in Africa, creating a pattern where signals of selection are confounded by demographic history. In contrast, the tropical (neotropical) allele on the X chromosome reaches high frequency exclusively in the Americas, suggesting that selection in North America has reshaped its distribution, overriding the expected demographic signature.

While both mutations showed a response to selection in our experiment, we did not detect any meaningful linkage disequilibrium (LD) between them in VA ($r^2_{\text{X-2R SNPs}}$ = 0.0002) or in the DGRP panel ($r^2_{\text{X-2R SNPs}}$ = 0.010). For the particular case of our 2R candidate, given the close physical proximity of our top SNP to the cosmopolitan inversion *In(2R)NS*, we investigated the extent of potential linkage between these loci. To this end, we calculated LD within a genomic window of approximately 10^7 bp in the DGRP panel, including a set of established inversion markers for *In(2R)NS* as described by (46). The results showed that, despite its physical proximity to the inversion, the top SNP in 2R segregates independently of *In(2R)NS* (**Fig. S7**).

We also inferred the relative ages of the mutation and the inversion using a phased dataset from wild Virginia populations (21). Our results suggest that both the 2R and X mutations are old ($TMRCA_{\text{2R SNP}}$= 214,384.8 ya; $TMRCA_{\text{X SNP}}$= 200,395.1 ya), predating the most recent out-of-Africa migration by at least 200,000 years. In the particular case of 2R, the top SNP emerged roughly three times earlier than the *In(2R)NS* inversion (Mean $TMRCA_{\text{inv. markers}}$ = 62,303.9 ya).

Phenotypic associations for our top SNP within the DGRP panel

To explore the potential broader phenotypic significance of our top SNPs, we analyzed GWAS data compiled by (21), which includes over 200 phenotypes from the DGRP. We identified phenotypes significantly associated with the top introgression SNPs in 2R and X, using an exploratory threshold (P < 0.1). For 2R, the data exploration yielded 10 associated traits, including larval anoxia tolerance, adult CT_{max} (in males), larval survival under chlorantraniliprole exposure, adult female survival after DDT exposure (1, 2, and 6 hours), thorax size (females), eye area (females), fecundity under low yeast conditions (females), and adult attraction to citral (males and females), helional (females), hexanal (females), and L-carvone (male). For X, it yielded 13 phenotypes, including: starvation resistance (males), acetophenone attraction (males and females), activity levels (females), larval anoxia tolerance, response to bacterial infections (female), ethyl-Butyrate attraction (males), flight prowess (females), adult CT_{max} (male and female), larval chlorantraniliprole resistance, L-carvone attraction (female), methyl salicylate attraction (males and females), paraquat resistance (male and female), and pupal case size.

We further explored the association between our two top SNPs and adult CT_{max} (data from (47)) by calculating genotype-specific differences in trait values using *Wolbachia*-free DGRP data. Wolbachia-positive lines were excluded, as this symbiont is known to significantly alter thermal phenotypes in *Drosophila*, potentially confounding the detection of genetic effects (48, 49). Indeed, the phenotype data from the DGRP revealed a significant genetic effect on adult male CT_{max} for alternative homozygotes at our top SNP on 2R (ANOVA; $F_{1,51}$ = 4.89, P = 0.0314). A similar effect was observed for female CT_{max} at our top SNP on the X chromosome (ANOVA; $F_{1,51}$ = 3.745, P = 0.0582), a weaker effect was also seen for male data for the same SNP (ANOVA; $F_{1,51}$ = 2.866, P = 0.0967). In line with the lack of linkage between these two loci, we did not detect any significant interaction effects (ANOVA [2R-by-X interaction], P_{males} = 0.708, $P_{females}$ = 0.953). Notably, when visualizing the patterns of variation in adult CT_{max}, we observe that the concordance of thermal tolerance phenotypes between embryos and adults for tropical (2R: C/C, X: A/A) and temperate (2R: A/A, X: G/G) alleles is reversed between the two markers (**Fig. 6A**). For instance, at the 2R locus, the tropical allele confers higher thermal tolerance in both embryos and adults (e.g., in males, mean $CT_{max(Tropical)}$ = 39.87°C, mean $CT_{max(Temperate)}$ = 39.63°C [ΔCT_{max} = 0.24°C]). In contrast, at the X locus, the tropical allele is associated with higher thermal tolerance in embryos, but lower thermal tolerance in adults (e.g., in males, mean $CT_{max(Tropical)}$ = 39.65°C, mean $CT_{max(Temperate)}$ = 39.84°C [ΔCT_{max} = -0.19°C]). Together, these findings indicate that while adult CT_{max} and embryonic thermal tolerance are potentially correlated, the direction of the genetic correlation may differ among loci, potentially leading to instances of locus-specific synergistic or antagonistic pleiotropy (50, 51).

Embryonic expression profiles reveal functional genes on 2R and X

Thus far, we have characterized the patterns of genetic variation surrounding our focal SNPs on chromosomes 2R and X. However, the functional genes linked to these top SNPs remain unknown. To address this, we employed a two-pronged approach to identify candidate genes associated with these loci. First, as these mutations were originally identified through an embryonic heat shock experiment, we utilized developmental transcriptomic data from the *modENCODE* project (52) to prioritize candidate genes expressed during the relevant developmental window. To accomplish this, we extracted gene expression levels at 0–2 hours post-egg laying for all genes annotated to each SNP using SnpEff. According to the annotation file for *D. melanogaster,* the top SNP in 2R is putatively associated with 6 genes including a 5' UTR in *SP70* (*CG13430*), a synonymous variant in *BORCS7* (*CG18065*; Leu76Leu), and an upstream variant relative to genes *CG30148*, *Obp57e*, *Obp57d*, *Cpr57A*, and *CG30148* (**Fig. 6B**). The *modENCODE* data (52) allow us to rule out *Obp57d, Obp57e,* and CG30148 genes since they show no signs of embryonic gene expression at 0-2 hours (Average expression = 0 Reads Per Kilobase per Million mapped reads; **RPKM; Table S6**). *Cpr57A* shows moderate expression (~3.65 RPKM). Notably, *SP70 and BORCS7* show high expression during early

embryonic development (6.26 and 7.25 RPKM respectively). Notably, our top hit in chromosome X has a single annotation: an intergenic SNP in the immediate vicinity of the *sog* (short gastrulation) gene. The *modENCODE* data show low levels of expression of *sog* at the 0–2 development window (2.0 RPKM).

Second, we generated an independent expression dataset by performing RNA-seq on mRNA extracted from ~1-hour-old embryos derived from the same populations used in the adaptive introgression experiment (VT, SK, and the F16 offspring). Embryos were exposed to either a 45-minute control treatment at 25°C or a heat shock at 34°C. For the 2R locus, our results indicate that *SP70* was highly expressed and displayed contrasting heat shock responses in the VT and SK parental lines, with the expression of F16 offspring showing the same norm of reaction as the SK parent (**Fig. 6C**). This pattern is consistent with selection for the tropical pattern of expression of *SP70* in the F16 embryos. In contrast, while *Cpr57A* also exhibited high expression with opposing responses between the parental lines, the F16 offspring showed an expression pattern more similar to VT, indicating that the expression of *Cpr57A* was not selected upon during adaptive introgression (**Fig. 6D**). Additionally, despite *BORCS7* (CG18065) showing high expression in the *modENCODE* dataset, it exhibited no detectable expression in our RNA-seq experiment. At the X-linked locus, *sog*, the expression pattern mirrored that of SP70: the F16 offspring displayed expression levels similar to the SK parent, consistent with selection for gene expression patterns that matched the more heat-tolerant SK embryos (**Fig. 6E**).

Top SNPs in SP70 and sog likely regulate gene expression in early-stage embryos

Based on the collective evidence presented throughout this study, we hypothesize that *SP70* (2R), which is predicted to encode a serine-type endopeptidase, and *sog* (X), which encodes a secreted developmental morphogen, house the key players of the genetic basis for embryonic heat tolerance in the wild. For the particular case of 2R, the data suggest that our top SNP (located in the 5' UTR of the gene) plays a direct regulatory role on *SP70* (**Fig. 6A**). This hypothesis is further supported by the fact that this 5' UTR SNP is located within the binding motif of the transcription factor Medea, and Medea binding to this motif has been verified by ChiP-seq (Flybase; (53)). Medea is a transcription factor involved in dorso-ventral patterning in the early stage embryo (54). Notably, aside from a single intergenic SNP, this 5' UTR variant is the only functional SNP associated with SP70 (**Fig. 6F; Table S7**).

In the case of X, the relationship between patterns of gene expression (i.e., **Fig. 6E**) and the top SNP is complex. The top seasonal and clinal SNP (X:15,602,941), per se, is an intergenic SNP near the 3' end of *sog* (**Fig. 6G**). The only annotation for this SNP is that of an origin point for the DNA replication program of *D. melnogaster* (**Table S7**). Yet, unlike SP70, *sog* contains several SNPs that are significant in the introgression FET analysis, of which one SNP (X:15,648,874) is in an enhancer known to regulate *sog* expression (*sog_Odz03*; (55, 56)). Although X:15,602,941 and X:15,648,874, are not linked in either the DGRP (r^2 = 0.013) or in Virginia (r^2 = 0.009), both loci are fixed in the SK parent, present at low frequency in the VT parent, and underwent the same pattern of strong adaptive introgression during the experiment (ΔAF = 0.35). The *sog_Odz03* mutation is not clinal in North America (P = 0.99), though it shows evidence of latitudinal and longitudinal clinality in Europe (r_{Lat} = -0.26, P = 0.018; r_{Long} = -0.46, P = 1.46x10^{-5}), as well as a marginal signature of seasonality in Virginia ($\beta_{seasonal}$ = -0.254, $P_{seasonal}$ = 0.12 [beats 85% of permutations]; best model: Humidity variation 0-15 days prior to collection).

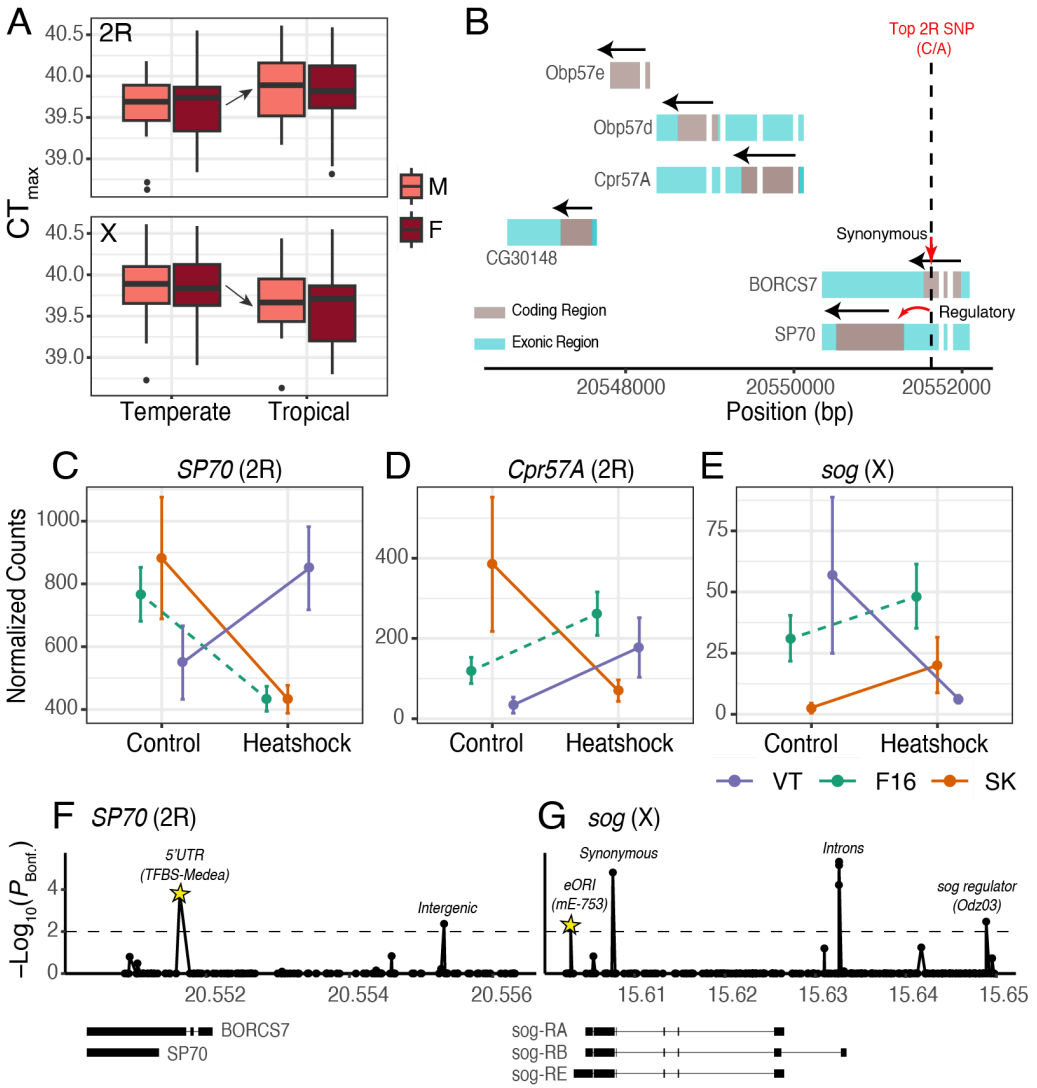

Figure 6. Genes and phenotypes associated with the top SNPs in 2R and X. (A) Phenotypic patterns of adult CT_{max} associated with the top SNPs in chromosomes 2R (top) and X (bottom) in the DGRP. The x-axis shows lines homozygous for either the temperate or tropical alleles. Data from (47). **(B)** Gene models for all coding loci associated with our top SNP in 2R. **(C-E)** Levels of transcript expression (normalized read counts) for SK, VT, and F16 embryos, at 25°C (Control) or after a heat shock of 34°C. Genes associated with our top SNPs in 2R (*SP70* and *Cpr57A*), and in X (*sog*) are shown. Note that RNAs of *Obp57e*, *Obp57d*, *CG30148*, and *BORCS7* were not expressed in our dataset. **(F)** SNPs in *SP70* that are significant in the introgression FET analysis. The top seasonal and clinal SNP is indicated with a star. Feature annotations are also shown. **(G)** Same as F, but for *sog*. See Table S7 for feature annotations of regions in (F) and (G).

Discussion

Here, we used a combinatorial multi-omics approach to identify loci underlying life-stage-specific adaptation in *D. melanogaster*. Neotropical and temperate North American flies have distinct embryonic heat tolerance, and with lab-based selection, introgression, and pooled whole-genome sequencing, we mapped the genomic basis of embryonic heat tolerance to two regions of major effect on chromosomes 2R and X. We used colocalization analyses with the vast DEST population genomics dataset to further characterize these genomic regions of interest, and we identified 2 SNPs that are both seasonal and clinal across populations sampled from multiple continents. These SNPs lie in the putative regulatory regions of the genes *SP70*

and *sog*. With RNA sequencing, we discovered that these two genes exhibited gene expression responses to heat stress that were correlated to variation in embryonic heat tolerance, suggesting that selection acts on gene expression to orchestrate adaptive responses to environmental variability in fly embryos. Altogether, we present strong evidence that (i) *SP70* and *sog* are important genes for climate adaptation and (ii) that embryonic heat tolerance is an ecologically relevant trait. These are novel findings that open the opportunity to investigate previously uncharacterized mechanisms of adaptation and emphasize the importance of ontogenetic investigations in the field of ecological and evolutionary genomics and physiology.

Mutations that are both seasonal and clinal

The *SP70* and *sog* genes are notable for several reasons. Beyond the strong signal detected in our mapping experiment (**Fig. 3**), colocalization analysis incorporating data from the DEST dataset indicates that mutations in these two genes exhibit genomic footprints of both clinal (**Fig. 4 and S5D;** (10)), and seasonal adaptation (**Figs. 5 and S5A**; (36)). These findings offer three insights into the genetic basis of adaptation in nature.

First, our data revealed significant allele clinality across multiple spatial scales. At the continental level, mutations in both *SP70* and *sog* exhibit parallel clinal patterns in North America and Europe. However, these patterns must be interpreted within the context of *Drosophila*'s complex demographic history, where North American populations resulted from a secondary encounter between African and European fly populations (22, 57). This history may explain the observed allele frequency clines, which could reflect historical mixing rather than direct selection. Yet, our genomic mapping showed that the mutations improve embryo survival following heat shock. This suggests that the clines may be adaptive, potentially shaped by environmental factors rather than purely demographic processes. These findings resonate with previous studies on populations from the east coasts of North America and Australia, where temperature gradients have been linked to clinal adaptation across hundreds of loci (58, 59).

Second, in addition to the strong signals of latitudinal clinality, our genomic mapping results provide experimental validation for two loci undergoing seasonal adaptive tracking (28). This is a noteworthy finding given that decades of ecological genomics research have leveraged *Drosophila* life history to explore the targets and dynamics of adaptive tracking in the wild (19, 21, 30, 33, 34, 60–66). Yet, while this body of work has shown that adaptive tracking influences thousands of *Drosophila* genes, few candidate loci have been experimentally validated or directly linked to ecologically important phenotypes (but see (35, 67–70)). Our findings address this gap by providing empirical evidence for specific adaptive loci closely tied to embryonic survival. For example, in our embryonic introgression experiment, we observed a pronounced allelic shift in *SP70* from the 'temperate' allele (A) towards the 'tropical' allele (C) after 16 generations of intermittent embryonic heat selection. Notably, this pattern was also evident in our follow up analyses using wild fly population data from DEST. Indeed, in our DEST analyses (e.g., from Virginia, Ukraine, and Turkey), the A allele becomes increasingly prevalent over the summer (**Fig. 5H**) tracking changes in the variance of temperature 1.5 months prior to collection (**Fig. 5A**). However, this pattern becomes erratic toward the end of the growing season, likely due to increased drift triggered by the onset of winter and overwintering bottlenecks (21).

Third, these results suggest that the alleles of *SP70* and *sog* exhibit consistent responses to environmental variables across both spatial and temporal scales, in line with the expectations of the seasonal trade-off model (29, 71). This physiological model, centered on temperature-driven trade-offs, proposes that cold conditions favor traits related to somatic maintenance, while warmer conditions select for traits that enhance fecundity (31, 32, 72–74). Consequently, loci associated with these adaptive shifts are expected to show context-dependent responses to temperature variation. Surprisingly, despite these predictions, empirical studies have revealed that only a small fraction of genomic loci (~3.7%) exhibit parallel seasonal and clinal changes (75). As a result, identifying loci with consistent ecological

responses has remained a challenge, with only a few notable exceptions (76, 77). In this context, our discovery of *SP70* and *sog* highlights the power of experiments that integrate classical tools in quantitative genetics with large genomic datasets to identify and validate ecologically important loci. In a landscape where finding loci with consistent ecological responses is akin to finding a needle in a haystack, the ability to pinpoint such targets represents a significant step forward in understanding the genetic basis of adaptation.

Selection acts on an early life stage

In addition to shedding light on spatial and temporal selection patterns, our work provides key insights into the timing of selection across the life history. Specifically, we propose that embryonic selection may be a crucial stage for adaptation. Unlike mobile adults, embryos are spatially constrained and unable to escape environmental stressors, making them particularly vulnerable during this critical developmental stage (78, 79). This stark contrast in mobility between early and late life stages shapes selective pressures throughout the fly life cycle, potentially leading to scenarios where the same environmental stressor drives distinct selection dynamics at different life stages. This phenomenon can give rise to complex ontogenetic histories of selection, such as ontogenetic decoupling (80) or antagonistic pleiotropy (50, 51), both processes that have been documented across a wide range of taxa, including insects (38, 81–83), amphibians (84, 85), birds and mammals (86, 87), and marine invertebrates (79, 88). Indeed, the data we present herein indicate that genetic correlations across life stages are locus dependent—e.g., the genetic association of *SP70* alleles to heat tolerance is concordant across life stages, yet *sog* alleles exhibit anti-correlated associations with heat tolerance in embryos vs. adults. This may help explain why we previously observed an overall lack of genetic correlation between embryonic and adult heat tolerance (39): on average, the allelic effects of different genes could offset one another. Although the evolutionary genetics of thermal adaptation is inherently complex, by examining thermal responses across multiple life stages, we uncover genotype-to-phenotype relationships that persist beyond laboratory conditions and into natural environments. In doing so, we gain deeper insight into the thermal selective landscape and its potential role in shaping the evolution of thermal tolerance traits.

Overall, the data presented herein supports our previous work (39) and suggest that early life stages play a key role in the ecological physiology of insects. The observation that embryonic heat tolerance genes are correlated with environmental data is strong evidence that embryonic heat tolerance is of ecological importance. Further, if thermal traits are genetically and/or physiologically decoupled across life stages, then the most thermally sensitive life stages could be the target of thermal selection that ultimately establish the upper thermal limits of a species more broadly (89). This has important implications for how species respond to environmental change, particularly when embryos are the most thermally sensitive life stage and incapable of behavioral thermoregulation (90).

Alternatively, natural selection could act on another trait or life stage entirely, and our current characterization of embryonic heat tolerance, although a direct response to our lab-based selection, may represent a correlated response to selection in nature. The vast majority of previous work on heat tolerance in *Drosophila* has focused on the adult life stage, with many studies that describe the genomic basis of adult thermal tolerance (47, 91, 92) and others that demonstrate clinal variation in adult heat tolerance (25, 73, 93). But despite this large body of work, few studies have connected genotype and phenotype of adult heat tolerance in an adaptive context (but see (94)), as we have done here for embryonic heat tolerance. Other thermal traits, such as thermal sensitivity of male fertility (95), correlate more strongly with environmental temperature than adult heat tolerance. In addition, thermal selection likely acts on adult behavioral thermoregulation in *D. subobscura (96)*. There is even evidence that selection acts on the circadian rhythm of egg laying in *D. melanogaster*, such that populations that experience extreme daytime heat only lay eggs at night (97). To our knowledge, there is no

evidence of selection on female choice of oviposition site in egg laying behavior (98). Further, the thermal properties of necrotic fruit (99) and *in situ* temperature data (13) suggest that the thermal microenvironment of the *Drosophila* embryo is highly variable and likely to exceed the temperatures that cause thermal stress on a regular basis (39, 100, 101). Thus, because embryos are immobile, they are likely to be at the mercy of the microenvironment in which they are laid. Accordingly, we suggest that embryonic thermal traits are a heretofore underappreciated aspect of thermal adaptation. In fact, it is plausible, in the context of genetic correlations, that adult thermal traits track environmental change as correlated responses to selection on embryonic traits, and this may be the reason for the comparatively weak correlation of adult heat tolerance and environmental temperature (39, 102).

The physiological target of selection

Our data, when viewed in the physiological context of *Drosophila* embryogenesis, may elucidate the molecular targets of thermal selection. We characterized heat tolerance in embryos that were less than 2 hours old, and at this stage of development, the early embryo is setting up its body axis and beginning to specify primordial cell types in preparation for tissue and organ formation, such as gastrulation. Gastrulation is a complex process regulated by a myriad of interacting genetic factors (103, 104) to orchestrate changes in cell shape and cell migration and ultimately form the gut of the organism (105, 106). Disruptions to gastrulation lead to developmental arrest (104). Importantly, gastrulation takes place within an hour of the developmental window in which we characterized embryonic heat tolerance (107). Based on this developmental context, the functions of our two major effect loci (*SP70* and *sog*), and how these genes were expressed in heat-tolerant vs. heat-sensitive embryos, we hypothesize that regulation of gastrulation is the physiological target of thermal selection.

Previous work suggests that *SP70* and *sog* interact indirectly to regulate gastrulation (108–110), with *SP70* promoting gastrulation and *sog* inhibiting it. Although the function of the SP70 protein has not been characterized in *D. melanogaster*, the human ortholog, TMPRSS4, induces the epithelial-to-mesenchymal transition (EMT) (111), which is a pro-gastrulation process that facilitates mesoderm formation via cell specification and migration (110). The interaction between *SP70* and *sog* is evidenced by the fact that *sog* directly inhibits the secreted factor *Dpp*, and *Dpp* activates the pMad/Medea transcription factor complex (54, 109). Intriguingly, our top SNP in *SP70* lies in the binding motif of pMad/Medea (53, 112). These findings, along with our RNA expression data, suggest that *sog* inhibits the expression of *SP70*, a pro-gastrulation signal.

Recall that *sog* and *SP70* showed anti-correlated expression patterns, with heat-tolerant embryos increasing the abundance of *sog* and decreasing the abundance of *SP70* RNAs after heat stress. These expression patterns suggest that heat-tolerant embryos slow or inhibit gastrulation under conditions of acute heat stress, perhaps to delay gastrulation until after recovery from heat stress. Whereas, heat-sensitive embryos do the opposite; by decreasing the abundance of *sog* and increasing the abundance of *SP70*, they likely promote gastrulation upon heat shock. Promoting gastrulation during heat stress may be disastrous due to (i) the potential for thermally induced disruptions to macromolecular structures in cell membranes (113) and the proteins (114) that facilitate gastrulation and (ii) misexpression of other genes involved in the spatial control of gastrulation that could lead to gut malformation (115, 116).

Why, then, do the alleles that confer the pro-gastrulation response under heat shock conditions segregate in nature? We hypothesize that more benign increases in temperature, if they don't lead to thermal denaturation of lipid membranes and proteins, could create ideal conditions for gastrulation to occur and thereby speed up development time (117). Shorter development time may be advantageous because it limits the amount of time that flies exist in a vulnerable life stage (78), as well as shortening the generation time. While it is beyond the scope of the present study to test our molecular physiological model of embryonic heat

tolerance mediated by gastrulation signals, future work will capitalize on this unique opportunity to uncover the mechanism of adaptation to environmental variability in embryos.

Conclusion

We utilized complementary approaches in quantitative genetics and ecological genomics to connect phenotype to genotype and to further connect genetics to patterns in nature. To our knowledge, this is the first genomic mapping of an ecologically relevant trait in an early life stage of *Drosophila*. It is rare to find genetic loci that exhibit both spatial and temporal patterns of adaptation in nature (75). Yet, here we report on two such genes (i.e., *SP70* and *sog*), which reside on separate chromosomes, providing strong support for our conjecture that embryonic heat tolerance is an ecologically relevant trait. Further, our gene expression data, along with previously published work (54), indicate that these two genes functionally interact to coordinate early developmental events (e.g., gastrulation) in the face of environmental variability. Altogether, this study presents a novel perspective on the multi-faceted ecological and evolutionary physiology of the complex life cycle of *D. melanogaster*.

Materials and Methods

Fly lines: We procured two parental lines with compatible inversions, one tropical line from the island of Saint Kitts (SK; Stock: 14021-0231.34) in the Caribbean and one temperate line from Vermont (VT), USA (Stock name: VTECK8). Lines were maintained under common-garden conditions fed with cornmeal-yeast-molasses medium at 25°C and 12L:12D for several generations prior to introgression mapping. ***Environmental data:*** We obtained hourly estimates for three environmental parameters, temperature, humidity, and precipitation, from the NASA POWER dataset (118). The data covered the period from January 1, 2015, to January 1, 2025, and were downloaded for three locations: sites corresponding to our parental lines (VT [latitude: 44.36272, longitude: -72.47167] and SK [latitude: 17.37319, longitude: -62.80985]) and a site in Virginia (VA) from (21)latitude: 37.97900, longitude: -78.48970(21). We queried 3 variables: relative humidity (RH2M), temperature at 2 meters (T2M), and mean precipitation (PRECTOTCORR). ***Introgression mapping*****:** To map the genomic basis of variation in embryonic heat tolerance, we conducted an advanced introgression with selection and backcrossing design (40–42) with 6 replicates per cross. We did reciprocal crosses, such that three replicate crosses used SK founder females and the other three replicate crosses used VT founder females. We followed the protocol of (39) except that embryos were heat-shocked for 1 hour instead of 45 min. The heat shock temperature was 35.2$\pm$0.05°C (mean$\pm$s.d.), which induces 80% mortality (20$\pm$10% survival). At generation 16, we again heat-shocked embryos and collected 30 adult females from the survivors for pooled whole-genome sequencing. ***DNA sequencing, read mapping and SNP calling:*** DNA was extracted with a phenol-chloroform extraction protocol (119). Each introgression replicate was sequenced on the Illumina NovaSeq 6000 using a configuration of 2x150 (insert size 350). DNA reads were mapped to a hologenome for *D. melanogaster* using the self-contained mapping pipeline developed for the DEST dataset (20). ***Estimation of genetic diversity, PCA, and introgression analyses:*** Analyses of nucleotide diversity (π) were done using npstat v1.0 (120). To assess introgression at a global level, we used FactoMineR (121) to calculate PCA on the genomic data including the two parental pools and all six of the F16 pools using a sliding window scheme (window size = 0.1 Mb; step = 50 Kb). We quantified the genomic background of each F16 line as either more “VT-like” or “SK-like” by calculating the ratio between the mean euclidean distance (i.e., *d*), in PCs 1 and 2, of each F16 pool relative to the SK or VT parent. ***Simulating neutral introgression:*** We used computer simulations to assess whether the observed patterns of

genetic variation in our F16 crosses differed from simple expectations of genetic drift. This model assumed that no introgression took place and thus that all patterns of genetic variation resulted from genetic drift. To this end, we used the forward-in-time simulator SLiM [v4.2.2] (45), and a non-Wright-Fisher framework (122), to simulate 16 generations of drift with rounds of back-crossing into a parental line (similar to the real experimental design). We simulated 100 replicates for each recombination window across the genome for a total of 60,800 "neutral-introgression" simulations. We also simulated a final step where genotypes are measured by applying pool-seq-like noise to the simulation's outcome. ***Identifying regions enriched for large F_{ST} outliers:*** We assessed signals of genetic differentiation using the fixation index (F_{ST}) statistic, as defined for pool-seq in Hubert et al. (123), implemented in the *poolfstat* [v3.0] package in R (124). Analysis tools in *poolfstat* are optimized for pool-seq data and explicitly account for coverage differences among pools through the "effective coverage" of the pool (n_e) statistic as described by (22, 58, 125). We evaluated the signal of the F_{ST} statistic at the level of genomic windows based on recombination rates. We report the enrichment analysis of F_{ST} values in the top 1% across the genome. ***Characterizing footprints of adaptive introgressions using the Fisher's Exact Test:*** We identified regions of interest resulting from our adaptive introgression experiment by calculating a Fisher's Exact Test between the parental and the F16 introgressed lines (both with VT and SK mothers) and the parental Vermont line. We ran this analysis for all alleles in the genome (mAF > 0.05) considering both the allele count and overall coverage to construct the contingency table. In the context of this test, we defined a locus of interest as any mutation showing a significant result in a Fisher's Exact Test within a region identified by our simulations as deviating from the expected pattern of genetic drift. ***Colocalization with other datasets:*** We assessed whether any of our putative targets of adaptive introgression have been reported in previous studies of ecological adaptation. We did this by comparing the FET *Q*-values reported in our study for the windows of interest with the *P*-values from clinal (30) and seasonal (19) datasets. ***Tests for clinality and seasonality:*** We used the DEST 2.0 (20) dataset in order to study the patterns of genetic variation from the top SNPs emanating from the adaptive introgression and colocalization study. Accordingly, we fit two types of analyses: clinal and seasonal tests. To test for clinality, we calculated the correlation between latitude and the allele frequency at each locus of interest. On the other hand, the seasonal models were performed using a generalized linear model framework using a binomial error structure. As for the dependent variable, we used the same summary statistics as in the clinal models summaries, yet explicitly summarized across 15, 30, 45, 60, 75, and 90 days prior to the collection date of each sample. We conducted the seasonal test using a framework similar to that of Nunez et al. 2024 (21). ***Gene expression:*** We conducted RNAseq to assess the extent to which candidate genes exhibited differential expression based on genetic background and exposure to heat stress. Following the introgression mapping, we subjected each replicate introgression to another round of heat selection (80% mortality) and isolated three mated females per replicate cross, which were used to establish three isofemale lines. To assess the gene expression response of embryos to heat stress, and following the protocol of (100), we subjected 0-1 h old embryos to an acute temperature treatment of 45 minutes at 25°C or 34°C. A total of 100 embryos per condition (i.e., VT parent, SK parent, and F16) were pooled for subsequent RNA extraction and sequencing. Samples were sequenced in a 2x150 configuration using Novaseq 6000 (S4 Flowcell, Illumina). We used R (v. 4.2.3) and the package "DESeq2" (v. 1.38.3) to normalize read counts and quantify gene expression (126). A complete and detailed description of the methods in this paper can be found in the ***Extended Materials and Methods*** supplemental section.

Acknowledgements

We thank Melissa Pespeni, Reid Brennan, Sara Helms Cahan, Kristi Montooth, Colin Meiklejohn, Thomas O'Leary, and John Pool for helpful discussions. JCBN acknowledges the Henderson-Harris fellowship program at the University of Vermont for institutional support, and the Vermont Advanced Computing Center (VACC; https://www.uvm.edu/vacc) for providing computational resources that contributed to this publication.

Funding

This work was supported by NSF grant IOS-1750322 to BLL, and by Start-up funds from the University of Vermont to JCBN.

Data availability

Code for all analyses presented in this paper is available in Github at https://github.com/bllockwo/embryonic_introgression_mapping_VTtropical_2020. The raw DNA reads for the parental and introgression lines are available in the National Center for Biotechnology Information (NCBI), short read archive (SRA), SRR ids: SRR34395643-SRR34395650 (in PRJNA1285949). The raw RNA reads can be found as SRR ids: SRR34364632-SRR34364655 (in PRJNA1285997). Information regarding the DEST data set can be found at (https://dest.bio/), and the pipeline is available in Github at https://github.com/DEST-bio/DESTv2.

References

1. T. Dobzhansky, Speciation as a Stage in Evolutionary Divergence. *Am. Nat.* **74**, 312–321 (1940).

2. E. Mayr, Speciation and selection. *Proc. Am. Philos. Soc.* **93**, 514–519 (1949).

3. L. H. Rieseberg, S. A. Church, C. L. Morjan, Integration of populations and differentiation of species: Research review. *New Phytol.* **161**, 59–69 (2004).

4. S. Yeaman, M. C. Whitlock, The genetic architecture of adaptation under migration-selection balance: The genetic architecture of local adaptation. *Evolution* **65**, 1897–1911 (2011).

5. T. Maruki, Z. Ye, M. Lynch, Evolutionary genomics of a subdivided species. *Mol. Biol. Evol.* **39**, msac152 (2022).

6. H. Levene, Genetic equilibrium when more than one ecological niche is available. *Am. Nat.* **87**, 331–333 (1953).

7. P. Hedrick, M. Ginevan, E. Ewing, Genetic polymorphism in heterogeneous environments.

Annual Review of Ecology, Evolution, and Systematics **7**, 1–32 (1976).

8. P. W. Hedrick, Genetic polymorphism in heterogeneous environments: The age of genomics. *Annu. Rev. Ecol. Evol. Syst.* **37**, 67–93 (2006).

9. P. W. Hedrick, Genetic variation in a heterogeneous environment. V. Spatial heterogeneity in finite populations. *Genetics* **89**, 389–401 (1978).

10. J. R. Adrion, M. W. Hahn, B. S. Cooper, Revisiting classic clines in Drosophila melanogaster in the age of genomics. *Trends Genet.* **31**, 434–444 (2015).

11. Y. F. Li, J. C. Costello, A. K. Holloway, M. W. Hahn, "Reverse ecology" and the power of population genomics. *Evolution* **62**, 2984–2994 (2008).

12. P. S. Schmidt, *et al.*, Ecological genetics in the North Atlantic: environmental gradients and adaptation at specific loci. *Ecology* **89**, S91–107 (2008).

13. M. E. Feder, N. Blair, H. Figueras, Natural thermal stress and heat shock protein expression in Drosophila larvae and pupae. *Funct. Ecol.* **11**, 90–100 (1997).

14. M. W. Denny, W. W. Dowd, L. Bilir, K. J. Mach, Spreading the risk: Small-scale body temperature variation among intertidal organisms and its implications for species persistence. *J. Exp. Mar. Bio. Ecol.* **400**, 175–190 (2011).

15. L. P. Miller, W. W. Dowd, Multimodal *in situ* datalogging quantifies inter-individual variation in thermal experience and persistent origin effects on gaping behavior among intertidal mussels (*Mytilus californianus*). *J. Exp. Biol.* jeb.164020 (2017). https://doi.org/10.1242/jeb.164020.

16. T. Flatt, Life-History Evolution and the Genetics of Fitness Components in Drosophila melanogaster. *Genetics* **214**, 3–48 (2020).

17. P. T. Ives, The genetic structure of American populations of Drosophila melanogaster. *Genetics* **30**, 167–196 (1945).

18. P. T. Ives, Further Genetic Studies of the South Amherst Population of Drosophila melanogaster. *Evolution* **24**, 507 (1970).

19. H. E. Machado, *et al.*, Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in Drosophila. *Elife* **10** (2021).

20. M. Kapun, *et al.*, Drosophila Evolution over Space and Time (DEST) - A New Population Genomics Resource. *Mol. Biol. Evol.* (2021). https://doi.org/10.1093/molbev/msab259.

21. J. C. B. Nunez, *et al.*, A cosmopolitan inversion facilitates seasonal adaptation in overwintering Drosophila. *Genetics* **226** (2024).

22. J. C. B. Nunez, *et al.*, Footprints of worldwide adaptation in structured populations of*D. melanogaster*through the expanded DEST 2.0 genomic resource. *Molecular Biology and Evolution* 1–28 (2025). https://doi.org/10.1093/molbev/msaf132.

23. L. Mateo, G. E. Rech, J. González, Genome-wide patterns of local adaptation in Western European Drosophila melanogaster natural populations. *Sci. Rep.* **8**, 16143 (2018).

24. G. De Jong, Z. Bochdanovits, Latitudinal clines in Drosophila melanogaster: body size, allozyme frequencies, inversion frequencies, and the insulin-signalling pathway. *J. Genet.* **82**, 207–223 (2003).

25. A. A. Hoffmann, A. Anderson, R. Hallas, Opposing clines for high and low temperature resistance in Drosophila melanogaster. *Ecol. Lett.* **5**, 614–618 (2002).

26. D. K. Fabian, *et al.*, Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. *Mol. Ecol.* **21**, 4748–4769 (2012).

27. A. A. Hoffmann, A. R. Weeks, Climatic selection on genes and traits after a 100 year-old invasion: A critical look at the temperate-tropical clines in Drosophila melanogaster from eastern Australia. *Genetica* **129**, 133–147 (2007).

28. C. A. Botero, F. J. Weissing, J. Wright, D. R. Rubenstein, Evolutionary tipping points in the capacity to adapt to environmental change. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 184–189 (2015).

29. P. S. Schmidt, D. R. Conde, Environmental heterogeneity and the maintenance of genetic variation for reproductive diapause in Drosophila melanogaster. *Evolution* **60**, 1602–1611 (2006).

30. A. O. Bergland, E. L. Behrman, K. R. O'Brien, P. S. Schmidt, D. a. Petrov, Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in *Drosophila*. *PLoS Genet.* **10**, e1004775 (2014).

31. E. L. Behrman, S. S. Watson, K. R. O'Brien, M. S. Heschel, P. S. Schmidt, Seasonal variation in life history traits in two Drosophila species. *J. Evol. Biol.* **28**, 1691–1704 (2015).

32. S. Rajpurohit, *et al.*, Spatiotemporal dynamics and genome-wide association genome-wide association analysis of desiccation tolerance in Drosophila melanogaster. *Mol. Ecol.* **27**, 3525–3540 (2018).

33. S. M. Rudman, *et al.*, Direct observation of adaptive tracking on ecological time scales in Drosophila. *Science* **375**, eabj7484 (2022).

34. M. C. Bitter, *et al.*, Continuously fluctuating selection reveals fine granularity of adaptation. *Nature* 1–8 (2024). https://doi.org/10.1038/s41586-024-07834-x.

35. E. L. Behrman, *et al.*, Rapid seasonal evolution in innate immunity of wild Drosophila melanogaster. *Proc. Biol. Sci.* **285** (2018).

36. O. L. Johnson, R. Tobler, J. M. Schmidt, C. D. Huber, Fluctuating selection and the determinants of genetic variation. *Trends Genet.* **39**, 491–504 (2023).

37. K. A. Otte, V. Nolte, F. Mallard, C. Schlötterer, The genetic architecture of temperature adaptation is shaped by population ancestry and not by selection regime. *Genome Biol.* **22**, 211 (2021).

38. J. a. Coyne, J. Bundgaard, T. Prout, Geographic Variation of Tolerance to Environmental Stress in Drosophila pseudoobscura. *Am. Nat.* **122**, 474–488 (1983).

39. B. L. Lockwood, T. Gupta, R. Scavotto, Disparate patterns of thermal adaptation between

life stages in temperate vs. tropical Drosophila melanogaster. *J. Evol. Biol.* **31**, 323–331 (2018).

40. J. E. Pool, Genetic mapping by bulk segregant analysis in Drosophila: Experimental design and simulation-based inference. *Genetics* **204**, 1295–1306 (2016).

41. M. Lynch, B. Walsh, *Genetics and analysis of quantitative traits* (Sinauer Sunderland, MA, 1998).

42. E. J. Earley, C. D. Jones, Next-generation mapping of complex traits with phenotype-based selection and introgression. *Genetics* **189**, 1203–1209 (2011).

43. T. F. C. Mackay, *et al.*, The Drosophila melanogaster genetic reference panel. *Nature* **482**, 173–178 (2012).

44. C. Schlötterer, R. Tobler, R. Kofler, V. Nolte, Sequencing pools of individuals - mining genome-wide polymorphism data without big funding. *Nat. Rev. Genet.* **15**, 749–763 (2014).

45. B. C. Haller, P. W. Messer, SLiM 4: Multispecies Eco-evolutionary modeling. *Am. Nat.* **201**, E127–E139 (2023).

46. M. Kapun, H. van Schalkwyk, B. McAllister, T. Flatt, C. Schlötterer, Inference of chromosomal inversion dynamics from Pool-Seq data in natural and laboratory populations of Drosophila melanogaster. *Mol. Ecol.* **23**, 1813–1827 (2014).

47. M. C. Lecheta, *et al.*, Integrating GWAS and Transcriptomics to Identify the Molecular Underpinnings of Thermal Stress Responses in Drosophila melanogaster. *Front. Genet.* **11**, 658 (2020).

48. S. V. Serga, O. M. Maistrenko, N. P. Matiytsiv, A. M. Vaiserman, I. A. Kozeretska, Effects of Wolbachia infection on fitness-related traits in Drosophila melanogaster. *Symbiosis* **83**, 163–172 (2021).

49. M. T. J. Hague, C. N. Caldwell, B. S. Cooper, Pervasive effects of Wolbachia on host temperature preference. *MBio* **11** (2020).

50. G. C. Williams, Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11**, 398–411 (1957).

51. D. A. Roff, Contributions of genomics to life-history theory. *Nat. Rev. Genet.* **8**, 116–125 (2007).

52. M. Muers, Functional genomics: the modENCODE guide to the genome: Functional genomics. *Nat. Rev. Genet.* **12**, 80 (2011).

53. A. Öztürk-Çolak, *et al.*, FlyBase: updates to the Drosophila genes and genomes database. *Genetics* **227**, iyad211 (2024).

54. M. B. O'Connor, D. Umulis, H. G. Othmer, S. S. Blair, Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. *Development* **133**, 183–193 (2006).

55. A. Ozdemir, *et al.*, High resolution mapping of Twist to DNA in Drosophila embryos: Efficient functional analysis and evolutionary conservation. *Genome Res.* **21**, 566–577 (2011).

56. S. M. Gallo, *et al.*, REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Res.* **39**, D118–23 (2011).

57. A. O. Bergland, R. Tobler, J. González, P. Schmidt, D. Petrov, Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in Drosophila melanogaster. *Mol. Ecol.* **25**, 1157–1174 (2016).

58. B. Kolaczkowski, A. D. Kern, A. K. Holloway, D. J. Begun, Genomic differentiation between temperate and tropical Australian populations of Drosophila melanogaster. *Genetics* **187**, 245–260 (2011).

59. J. A. Reinhardt, B. Kolaczkowski, C. D. Jones, D. J. Begun, A. D. Kern, Parallel Geographic Variation in Drosophila melanogaster. *Genetics* **197**, 361–373 (2014).

60. T. Dobzhansky, Genetics of natural populations IX. Temporal changes in the composition of populations of Drosophila pseudoobscura. *Genetics* **28**, 162–186 (1943).

61. T. Dobzhansky, GENETICS OF NATURAL POPULATIONS. XIV. A RESPONSE OF CERTAIN GENE ARRANGEMENTS IN THE THIRD CHROMOSOME OF *DROSOPHILA PSEUDOOBSCURA* TO NATURAL SELECTION. *Genetics* **32**, 142–160 (1947).

62. T. Dobzhansky, F. J. Ayala, Temporal frequency changes of enzyme and chromosomal polymorphisms in natural populations of *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 680–683 (1973).

63. J. Boulétreau-Merle, P. Fouillet, O. Terrier, Seasonal variations and balanced polymorphisms in the reproductive potential of temperate D. Melanogaster populations. *Entomol. Exp. Appl.* **43**, 39–48 (1987).

64. F. Rodríguez-Trelles, G. Alvarez, C. Zapata, Time-series analysis of seasonal changes of the *O* inversion polymorphism of *Drosophila subobscura*. *Genetics* **142**, 179–187 (1996).

65. E. L. Rezende, *et al.*, Climate change and chromosomal inversions in Drosophila subobscura. *Clim. Res.* **43**, 103–114 (2010).

66. L. Olazcuaga, *et al.*, Rapid and transient evolution of local adaptation to seasonal host fruits in an invasive pest fly. *Evol. Lett.* **6**, 490–505 (2022).

67. A. Glaser-Schmitt, M. J. Wittmann, T. J. S. Ramnarine, J. Parsch, Sexual antagonism, temporally fluctuating selection, and variable dominance affect a regulatory polymorphism in *Drosophila melanogaster*. *Mol. Biol. Evol.* **38**, 4891–4907 (2021).

68. R. Cogni, *et al.*, Variation in*Drosophila melanogaster*central metabolic genes appears driven by natural selection both within and between populations. *Proc. Biol. Sci.* **282**, 20142688 (2015).

69. A. B. Paaby, A. O. Bergland, E. L. Behrman, P. S. Schmidt, A highly pleiotropic amino acid polymorphism in the*Drosophila*insulin receptor contributes to life-history adaptation. *Evolution* **68**, 3395–3409 (2014).

70. P. S. Schmidt, *et al.*, An amino acid polymorphism in the *couch potato* gene forms the basis for climatic adaptation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16207–16211 (2008).

71. T. Flatt, G. V. Amdam, T. B. L. Kirkwood, S. W. Omholt, Life-history evolution and the polyphenic regulation of somatic maintenance and survival. *Q. Rev. Biol.* **88**, 185–218 (2013).

72. P. S. Schmidt, L. Matzkin, M. Ippolito, W. F. Eanes, Geographic variation in diapause incidence, life-history traits, and climatic adaptation in Drosophila melanogaster. *Evolution* **59**, 1721–1732 (2005).

73. P. S. Schmidt, A. B. Paaby, Reproductive Diapause and Life-History Clines in North American Populations of Drosophila Melanogaster. *Evolution* **62**, 1204–1215 (2008).

74. P. A. Erickson, *et al.*, Unique genetic signatures of local adaptation over space and time for diapause, an ecologically relevant complex trait, in Drosophila melanogaster. *PLoS Genet.* **16**, e1009110 (2020).

75. Y. Yu, A. O. Bergland, Distinct signals of clinal and seasonal allele frequency change at eQTLs in *Drosophila melanogaster*. *Evolution* **76**, 2758–2768 (2022).

76. M. Kapun, D. K. Fabian, J. Goudet, T. Flatt, Genomic Evidence for Adaptive Inversion Clines in Drosophila melanogaster. *Mol. Biol. Evol.* **33**, 1317–1336 (2016).

77. R. Cogni, *et al.*, The intensity of selection acting on the couch potato gene--spatial-temporal variation in a diapause cline: Spatial-temporal variation in diapause Cline. *Evolution* **68**, 538–548 (2014).

78. J. G. Kingsolver, L. B. Buckley, Ontogenetic variation in thermal sensitivity shapes insect ecological responses to climate change. *Curr Opin Insect Sci* **41**, 17–24 (2020).

79. L. L. M. Pandori, C. J. B. Sorte, The weakest link: sensitivity to climate extremes across life stages of marine invertebrates. *Oikos* **128**, 621–629 (2019).

80. J. Collet, S. Fellous, Do traits separated by metamorphosis evolve independently? Concepts and methods. *Proc. Biol. Sci.* **286**, 20190445 (2019).

81. C. J. Austin, A. J. Moehring, Local thermal adaptation detected during multiple life stages across populations of Drosophila melanogaster. *J. Evol. Biol.* (2019). https://doi.org/10.1111/jeb.13530.

82. J. G. Kingsolver, *et al.*, Complex life cycles and the responses of insects to climate change. *Integr. Comp. Biol.* **51** (2011).

83. P. J. Freda, *et al.*, Transcriptomic and functional genetic evidence for distinct ecophysiological responses across complex life cycle stages. *J. Exp. Biol.* **225** (2022).

84. K. Ruthsatz, K. H. Dausmann, M. A. Peck, J. Glos, Thermal tolerance and acclimation capacity in the European common frog (Rana temporaria) change throughout ontogeny. *J. Exp. Zool. A Ecol. Integr. Physiol.* **337**, 477–490 (2022).

85. R. Dupré, J. Petranka, Ontogeny of temperature selection in larval amphibians. *Copeia* **1985**, 462–467 (1985).

86. A. Nord, S. Giroud, Lifelong effects of thermal challenges during development in birds and mammals. *Front. Physiol.* **11**, 419 (2020).

87. D. Webb, Thermal tolerance of avian embryos: A review. *The Condor* **89**, 874–898 (1987).

88. A. P. Rebolledo, C. M. Sgrò, K. Monro, Thermal performance curves reveal shifts in optima, limits and breadth in early life. *J. Exp. Biol.* **223** (2020).

89. M. J. Angilletta, M. H. Zelic, G. J. Adrian, a. M. Hurliman, C. D. Smith, Heat tolerance during embryonic development has not diverged among populations of a widespread species (Sceloporus undulatus). *Conservation Physiology* **1**, cot018 (2013).

90. L. B. Buckley, J. C. Ehrenberger, M. J. Angilletta, Thermoregulatory behaviour limits local adaptation of thermal niches and confers sensitivity to climate change. *Funct. Ecol.* **29**, 1038–1047 (2015).

91. T. J. Morgan, T. F. C. Mackay, Quantitative trait loci for thermotolerance phenotypes in Drosophila melanogaster. *Heredity* **96**, 232–242 (2006).

92. A. R. Anderson, J. E. Collinge, A. A. Hoffmann, M. Kellett, S. W. McKechnie, Thermal tolerance trade-offs associated with the right arm of chromosome 3 and marked by the hsr-omega gene in Drosophila melanogaster. *Heredity (Edinb.)* **90**, 195–202 (2003).

93. C. M. Sgro, *et al.*, A comprehensive assessment of geographic variation in heat tolerance and hardening capacity in populations of *Drosophila melanogaster* from Eastern Australia. *J. Evol. Biol.* **23**, 2484–2493 (2010).

94. B. R. Bettencourt, I. Kim, A. A. Hoffmann, M. E. Feder, Response to Natural and Laboratory Selection at the Drosophila HSP70 Genes. *Evolution* **56**, 1796–1801 (2002).

95. S. R. Parratt, *et al.*, Temperatures that sterilize males better match global species distributions than lethal temperatures. *Nat. Clim. Chang.* **11**, 481–484 (2021).

96. R. B. Huey, M. Pascual, Partial thermoregulatory compensation by a rapidly evolving invasive species along a latitudinal cline. *Ecology* **90**, 1715–1720 (2009).

97. R. Allemand, J. R. David, The circadian rhythm of oviposition in *Drosophila melanogaster*: a genetic latitudinal cline in wild populations. *Experientia* **32**, 1403–1404 (1976).

98. M. E. Feder, N. Blair, H. Figueras, Oviposition site selection : unresponsiveness of Drosophila to cues of potential thermal stress. *Anim. Behav.* **53**, 585–588 (1997).

99. M. W. Denny, Air and water: the biology and physics of life's media. (1993).

100. E. E. Mikucki, T. S. O'Leary, B. L. Lockwood, Heat tolerance, oxidative stress response tuning and robust gene activation in early-stage Drosophila melanogaster embryos. *Proc. Biol. Sci.* **291**, 20240973 (2024).

101. R. Leemans, *et al.*, Quantitative transcript imaging in normal and heat-shocked Drosophila embryos by using high-density oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12138–12143 (2000).

102. L. B. Buckley, R. B. Huey, Integrative and Comparative Biology How Extreme Temperatures Impact Organisms and the Evolution of their Thermal Tolerance. *Integr. Comp. Biol.* 1–12 (2016). https://doi.org/10.1093/icb/icw004.

103. E. Wieschaus, Positional information and cell fate determination in the early Drosophila embryo. *Curr. Top. Dev. Biol.* (2016).

104. A. C. Martin, The Physical Mechanisms of Drosophila Gastrulation: Mesoderm and Endoderm Invagination. *Genetics* **214**, 543–560 (2020).

105. M. Leptin, B. Grunewald, Cell shape changes during gastrulation in Drosophila. *Development* **110**, 73–84 (1990).

106. A. McMahon, W. Supatto, S. E. Fraser, A. Stathopoulos, Dynamic analyses of Drosophila gastrulation provide insights into collective cell migration. *Science* **322**, 1546–1550 (2008).

107. V. E. Foe, B. M. Alberts, Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. *J. Cell Sci.* **61**, 31–70 (1983).

108. S. B. Zusman, D. Sweeton, E. F. Wieschaus, short gastrulation, a mutation causing delays in stage-specific cell shape changes during gastrulation in Drosophila melanogaster. *Dev. Biol.* **129**, 417–427 (1988).

109. G. Pyrowolakis, B. Hartmann, B. Müller, K. Basler, M. Affolter, A simple molecular complex mediates widespread BMP-induced repression during Drosophila development. *Dev. Cell* **7**, 229–240 (2004).

110. E. Gheisari, M. Aakhte, H.-A. J. Müller, Gastrulation in Drosophila melanogaster: Genetic control, cellular basis and biomechanics. *Mech. Dev.* **163**, 103629 (2020).

111. A. L. de Aberasturi, A. Calvo, TMPRSS4: an emerging potential therapeutic target in cancer. *Br. J. Cancer* **112**, 4–8 (2015).

112. V. K. Jenkins, A. Larkin, J. Thurmond, “Using FlyBase: A Database of Drosophila Genes and Genetics” in *Drosophila: Methods and Protocols*, C. Dahmann, Ed. (Springer US, 2022), pp. 1–34.

113. J. R. Hazel, Thermal Adaptation in Biological-Membranes - Is Homeoviscous Adaptation the Explanation. *Annu. Rev. Physiol.* **57**, 19–42 (1995).

114. G. N. Somero, B. L. Lockwood, L. Tomanek, *Biochemical adaptation: response to environmental challenges, from life’s origins to the Anthropocene* (Sinauer Associates, Incorporated …, 2017).

115. W. Driever, V. Siegel, C. Nüsslein-Volhard, Autonomous determination of anterior structures in the early Drosophila embryo by the bicoid morphogen. *Development* **109**, 811–820 (1990).

116. D. J. Sutherland, M. Li, X.-Q. Liu, R. Stefancsik, L. A. Raftery, Stepwise formation of a SMAD activity gradient during dorsal-ventral patterning of the Drosophila embryo. *Development* **130**, 5705–5716 (2003).

117. M. Ashburner, K. G. Golic, R. S. Hawley, *Drosophila: A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, 1989).

118. A. Sparks, Nasapower: A NASA POWER global meteorology, surface solar energy and climatology data client for R. *J. Open Source Softw.* **3**, 1035 (2018).

119. H. Bastide, J. D. Lange, J. B. Lack, A. Yassin, J. E. Pool, A variable genetic architecture of melanic evolution in Drosophila melanogaster. *Genetics* **204**, 1307–1319 (2016).

120. L. Ferretti, S. E. Ramos-Onsins, M. Pérez-Enciso, Population genomics from pool sequencing. *Mol. Ecol.* **22**, 5561–5576 (2013).

121. S. Lê, J. Josse, F. Husson, FactoMineR: AnRPackage for Multivariate Analysis. *J. Stat. Softw.* **25**, 1–18 (2008).

122. B. C. Haller, P. W. Messer, SLiM 3: Forward genetic simulations beyond the Wright-Fisher model. *Mol. Biol. Evol.* **36**, 632–637 (2019).

123. V. Hivert, R. Leblois, E. J. Petit, M. Gautier, R. Vitalis, Measuring genetic differentiation from pool-seq data. *Genetics* **210**, 315–330 (2018).

124. M. Gautier, R. Vitalis, L. Flori, A. Estoup, f-statistics estimation and admixture graph construction with Pool-Seq or allele count data using the R package poolfstat. *Mol. Ecol. Resour.* **22**, 1394–1416 (2022).

125. A. F. Feder, D. A. Petrov, A. O. Bergland, LDx: estimation of linkage disequilibrium from high-throughput pooled resequencing data. *PLoS One* **7**, e48588 (2012).

126. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).