



Complex evolutionary processes maintain an ancient chromosomal inversion

Patrik Nosil^{a,1}, Victor Soria-Carrasco^b, Romain Villoutreix^b , Marisol De-la-Mora^{a,c} , Clarissa F. de Carvalho^a , Thomas Parchman^d, Jeffrey L. Feder^e , and Zachariah Gompert^{f,1}

Edited by Dolph Schlüter, The University of British Columbia, Vancouver, Canada; received January 12, 2023; accepted May 12, 2023

Genome re-arrangements such as chromosomal inversions are often involved in adaptation. As such, they experience natural selection, which can erode genetic variation. Thus, whether and how inversions can remain polymorphic for extended periods of time remains debated. Here we combine genomics, experiments, and evolutionary modeling to elucidate the processes maintaining an inversion polymorphism associated with the use of a challenging host plant (Redwood trees) in *Timema* stick insects. We show that the inversion is maintained by a combination of processes, finding roles for life-history trade-offs, heterozygote advantage, local adaptation to different hosts, and gene flow. We use models to show how such multi-layered regimes of balancing selection and gene flow provide resilience to help buffer populations against the loss of genetic variation, maintaining the potential for future evolution. We further show that the inversion polymorphism has persisted for millions of years and is not a result of recent introgression. We thus find that rather than being a nuisance, the complex interplay of evolutionary processes provides a mechanism for the long-term maintenance of genetic variation.

balancing selection | ecological genetics | gene flow | insect-plant interactions | local adaptation

Genetic variation is the ultimate fuel for evolution. However, many forms of natural selection (e.g., directional and purifying selection) and random genetic drift are expected to result in the loss of genetic variation, depleting the reservoir of fuel for evolution. Whether and how genetic variation can be maintained over long periods of time thus remains a central question in biology (1–8). We address this question here by studying the maintenance of an ancient chromosomal inversion. Since their discovery by Sturtevant ~100 y ago (9), chromosomal inversions have been central to the development of evolutionary biology. In this regard, inversions served as the first genetic markers and motivated ideas by Dobzhansky et al. concerning co-adapted gene complexes and balancing selection (10, 11). For example, inversions can create “supergenes” by capturing multilocus adaptive allele combinations (12–14). This mechanism underlies several modern theories of adaptation that involve suppressed recombination (15, 16). Inversions also serve as powerful models for studying the maintenance of genetic variation, because their age can be estimated and they are often subject to natural selection (17–19).

Although inversions are now known to vary along environmental clines and to be associated with adaptive traits (10, 11, 18, 20–22), studies that directly estimate selection on inversions are few, some notable exceptions aside (12, 23, 24). Thus, the mode and strength of selection on inversions remains poorly quantified, making it difficult to infer how and why inversion polymorphism is maintained. For example, forms of balancing selection, including heterozygote advantage (i.e., overdominance) and negative frequency-dependent selection, can maintain inversion polymorphisms (11, 25), especially if strong enough to counteract drift (*SI Appendix*, Fig. S1). But this is not true of many other forms of selection. Similarly, the role of gene flow in maintaining polymorphism requires further study (8, 26–28).

Determining the age of an inversion is also important for explaining the maintenance of inversion polymorphisms. For example, one hypothesis is that the inversion is young and still in the process of sweeping to fixation. In other words, it could be that the inversion will not be maintained as polymorphic in the long term. If the inversion polymorphism is found to be ancient such that this “young inversion” hypothesis is refuted, then studies of the processes maintaining variation, particularly natural selection, are required to explain the inversion polymorphism (Fig. 1A). Here we combine field data, genomics, experimental estimates of fitness, and evolutionary modeling to elucidate the processes driving the long-term maintenance of an inversion polymorphism with fitness consequences across populations using different hosts (Fig. 1).

Significance

How genetic variation, the fuel for evolution, is maintained is one of the central questions in biology. This is an especially striking question for chromosomal inversions (a change in the structure of an organism's genome), as inversions are often subject to natural selection, which can erode variation. We studied stick insects that have an inversion that affects growth and survival on two alternative host plants, Redwood trees versus a flowering plant. We found that this inversion has been present for millions of years, and that a suite of factors, including environmental heterogeneity and gene exchange, contribute to the persistence of this polymorphism. Our results show how a complex interplay of evolutionary processes offers resilience against the loss of variation allowing for the potential for future evolution.

Author contributions: P.N., J.L.F., and Z.G. designed research; P.N., R.V., M.D.-I.M., C.F.D.C., and T.P. performed research; V.S.-C. and Z.G. analyzed data; and P.N. and Z.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: patrik.nosil@cefe.cnrs.fr or zach.gompert@usu.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2300673120/DCSupplemental>.

Published June 13, 2023.

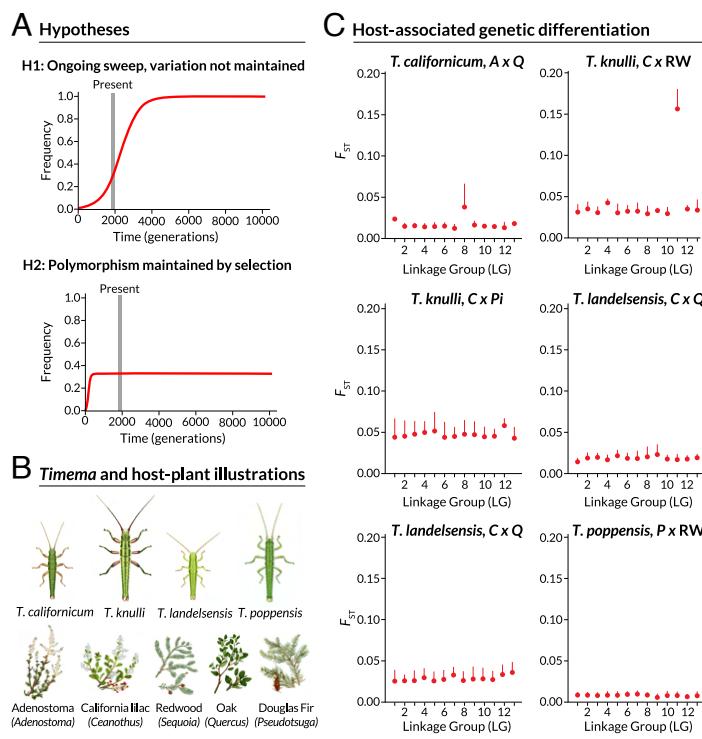


Fig. 1. Conceptual overview and tests for host-associated differentiation. Panel (A) illustrates our two alternative hypotheses of (H1) an ongoing sweep and (H2) polymorphism maintained by selection. Under the first hypothesis the inversion is young and in the process of sweeping; variation will not be maintained. Under the second hypothesis balancing selection promotes the long-term maintenance of inversion polymorphism. (B) Shows illustrations of *Timema* stick insects and their host plants, for the taxa studied here. Panel (C) summarizes genome-wide genetic differentiation for parapatric *Timema* populations on different hosts. Points denote mean F_{ST} for each of 13 *T. cristinae* linkage groups with horizontal lines extending to the 75th percentile of F_{ST} for that linkage group. Host abbreviations are A = *Adenostoma*, C = *Ceanothus*, P = *Pseudotsuga menziesii* (Douglas Fir), Pi = *Pinus* (pine), Q = *Quercus* (oak), and RW = *Sequoia sempervirens* (Redwood). See *SI Appendix*, Fig. S2 for more detailed patterns of genetic differentiation. Illustrations here and in other figures by Rosa Ribas.

Our study system is the genus *Timema*, a group of plant-feeding stick insects distributed throughout southwestern North America (Fig. 1). *Timema* are well studied for their cryptic colors and patterns, which help them avoid predation by visual predators such as birds and lizards (29, 30). These traits are highly heritable and controlled by a modest number (~ 5) of linked loci on linkage group (LG) 8 (LG8 hereafter), which often exhibit strongly reduced recombination due to structural genomic features including chromosomal inversions and deletions (17, 31, 32). *Timema* are also known to use a particularly wide range of host-plant species, including both conifers and flowering plants (i.e., angiosperms) (33). This host-plant use, in the context of local adaptation (i.e., growth and survival on different hosts; “performance” hereafter), is our focus here. Notably, the genetic basis of performance variation in *Timema* was previously unknown, but as we report here also involves a chromosomal inversion (on a different chromosome from color, LG11).

Results

Genome Scans Reveal Exceptional Host-Associated Differentiation on Linkage Group 11. During the 30-million year diversification of the *Timema* genus, host shifts have occurred frequently between plant families (within conifers and within flowering plants), and several times even between these plant divisions (33). Indeed, *Timema* are broadly generalized in diet, often feeding on multiple plant families in nature and surviving in the lab on novel hosts (34). One exception involves the use of Redwood (*Sequoia sempervirens*); very few *Timema* species and populations

are known to use Redwood in nature—only *T. knulli* and *T. poppensis*—and most exhibit poor performance on this host in laboratory experiments (34).

We thus initiated our investigation by quantifying patterns of genetic differentiation for the sexual species of *Timema* that live in the vicinity of Redwood in northern California and that use multiple hosts in nature. Specifically, we study *T. californicum* and *T. landelsensis*, which do not use Redwood, *T. poppensis*, which is specialized on conifers including Redwood in some localities, and *T. knulli*, which uses both Redwood and angiosperms. In this context, *T. knulli* is of particular interest as it is polymorphic in host-plant use, living on Redwood (a conifer) as well as other more commonly-used hosts such as *Ceanothus* (an angiosperm). In contrast, *T. poppensis* uses only conifer hosts. We tested for host-associated genetic differentiation using published genotyping-by-sequencing (GBS) data (35, 36). Our core interest was whether the use of a certain host was associated with genetic differentiation, and if so whether this was genome-wide or restricted to individual chromosomes. Due to the known strong effects of geographic isolation on genetic structure in *Timema* (37), we restricted our survey to the six pairwise comparisons involving nearby populations using different hosts (broadly speaking, “parapatry,” *SI Appendix*, Table S1, Fig. 1, *SI Appendix*, S2). This revealed that genetic differentiation between parapatric, conspecific populations was generally weak. The exception to this trend was LG11 for populations of *T. knulli* using *Ceanothus* versus Redwood: LG11 was strongly differentiated in this comparison. We thus focused our study on *T. knulli*, with particular reference to the use of Redwood.

Redwood *T. knulli* Populations Are Distinguished by a Chromosomal Inversion. The results above were based on mapping GBS reads to the published *T. cristinae* reference genome (35, 36). *Timema knulli* is known from cytological work to have one chromosome pair fewer than *T. cristinae* (38), and we suspected structural variation on LG11 within *T. knulli* based on our initial analyses (*SI Appendix*, Fig. S2B). Thus, to increase the accuracy and precision of the current work and test explicitly for structural variation, we generated a high-quality de novo reference genome assembly for *T. knulli*. We did so using an individual collected from Redwood and a combination of PacBio and Illumina reads with Chicago and Hi-C technology for scaffolding. The *T. knulli* genome comprised 12 large scaffolds corresponding to the 13 known *T. cristinae* chromosomes, but with a fusion between *T. cristinae* chromosomes 1 and 3 (we refer to the fused chromosome as chromosome 1 and retain the *T. cristinae* linkage group numbering for the other chromosomes; total assembly length = 1,322,373,696 base pairs; scaffold N50 = 83,614,905 base pairs) (*SI Appendix*, Table S2 and Fig. S3). We then used this reference genome for further population genetic and trait mapping analyses, with new data collected to allow larger sample sizes for *T. knulli* than what were available from published data.

Using new GBS data from 138 *T. knulli* collected on *Ceanothus* and Redwood (*SI Appendix*, Table S3) we detected a large block of differentiation (e.g., highly accentuated F_{ST}) on chromosome 11, whose boundaries were delimited using a Hidden Markov Model (HMM) approach applied to the results of a principal components analysis (PCA) (Fig. 2). This block spanned genomic positions 13,093,370 to 43,606,674 on chromosome 11 (~30 mega-base pairs, mbps). This 30 mbp genomic region included 876 of the genome-wide total of 36,055 annotated genes, with a striking spike in gene density near the region boundaries (see *SI Appendix*, Fig. S4 and *Datasets S1, S2*, and *S3* for details). We hereafter refer to this region as the “*Perform*” locus, as polymorphism at this regions was associated with performance variation in an experiment reported below. A PCA of SNPs within the *Perform* locus revealed three genetic clusters segregating within populations (Fig. 2). In contrast, PCA of genome-wide genetic variation exhibited structure by geography. This result is consistent with the *Perform* locus being a structural genomic variant that segregates within populations, differs in frequency among populations (as we reported in more detail below, one allele, denoted P_{RW} , is at 84% frequency on Redwood but only at 34% frequency on *Ceanothus* where the alternative allele, denoted

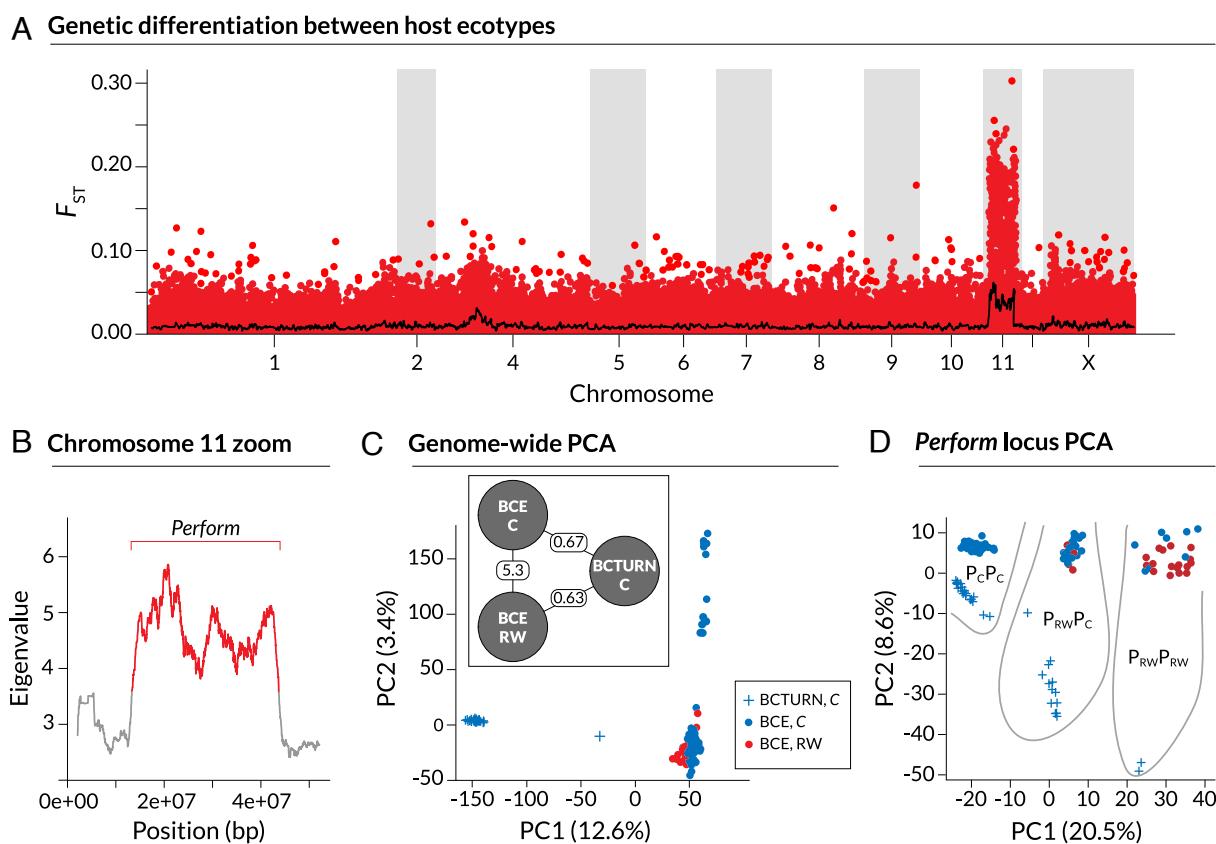


Fig. 2. Genetic differentiation and structure associated with Redwood feeding in *Timema knulli*. These results are based on the new reference genome for *T. knulli*. (A) Manhattan plot of F_{ST} between stick insects collected on *Ceanothus* versus Redwood at the BCE locality. Points denote F_{ST} for individual SNPs; the solid line denotes mean F_{ST} in 100 SNP sliding windows. *Timema knulli* chromosomes are used here (chromosome 3 from *T. cristinae* is fused to chromosome 1; X = the X sex chromosome). Panel (B) shows (square roots of) eigenvalues for the first principal component of genetic variation in *T. knulli* (excluding BCTURN, an allopatric *Ceanothus* population) in 100 SNP overlapping, sliding windows along chromosome 11. Colors denote alternative states as identified by a Hidden Markov model (HMM), with red denoting the elevated eigenvalue state and defining the bounds for the “*Perform*” locus on chromosome 11 (text for details). Panels (C) and (D) show summaries of genetic variation in *T. knulli* based on principal components analysis (PCA) for all SNPs not on chromosome 11 (C) and for the *Perform* locus only (D). Values for the first two principal components are shown with colors and symbols denoting locations and hosts. The inset in (C) is a schematic for the model used to infer neutral rates of gene flow among populations: BCE C (on *Ceanothus*), BCE RW (on Redwood) and BCTURN C (on *Ceanothus*). Point estimates of Nm , that is the number of migrants exchanged per generation, are shown on lines connecting the populations, and are consistent with a pattern of isolation by geographic distance. The ellipses in (D) delimit *Perform* locus genotypes based on PC clusters.

P_C , is more common), and exhibits reduced recombination between the two chromosomal variants.

To more formally test for the existence of a chromosomal inversion on *T. knulli* chromosome 11, we aligned the *T. knulli* genome with published chromosome-level assemblies of *T. cristinae* and *T. chumash* genomes (31, 32). These genome alignments identified an inversion on chromosome 11 in the Redwood *T. knulli* genome relative to both *T. cristinae* and *T. chumash*. Most critically, the breakpoints of this inversion coincided with the identified bounds of the *Perform* locus (Fig. 3). In contrast, this genomic region was co-linear between *T. cristinae* and *T. chumash*. The collective results are most consistent with the *Perform* locus being a polymorphic chromosomal inversion in *T. knulli*. We combined additional population genomic analyses and whole genome sequencing to explicitly test this

hypothesis. Specifically, we identified and analyzed patterns of linkage disequilibrium (LD) and heterozygosity for different *Perform* genotypes corresponding to different clusters in PC space (Fig. 2D). As predicted if *Perform* is a segregating inversion within *T. knulli*, we found elevated LD across *Perform* when considering $P_{RW}P_{RW}$ and $P_C P_C$ homozygotes together (Figs. 3 D–F, SI Appendix, S5), and elevated heterozygosity precisely in the *Perform* locus region in $P_{RW}P_C$ heterozygotes (Fig. 3 G–I). Interestingly, heterozygosity was especially low within *Perform* for $P_C P_C$ homozygotes, consistent with a possible selective sweep on this genomic background. Lastly, we gathered nanopore long-read DNA sequence data from a second *T. knulli* collected on *Ceanothus*. This revealed a large inversion (9,706,606 to 48,357,002 bps on chromosome 11) relative to the Redwood *T. knulli* genome (SI Appendix, Fig. S6). This inversion spanned

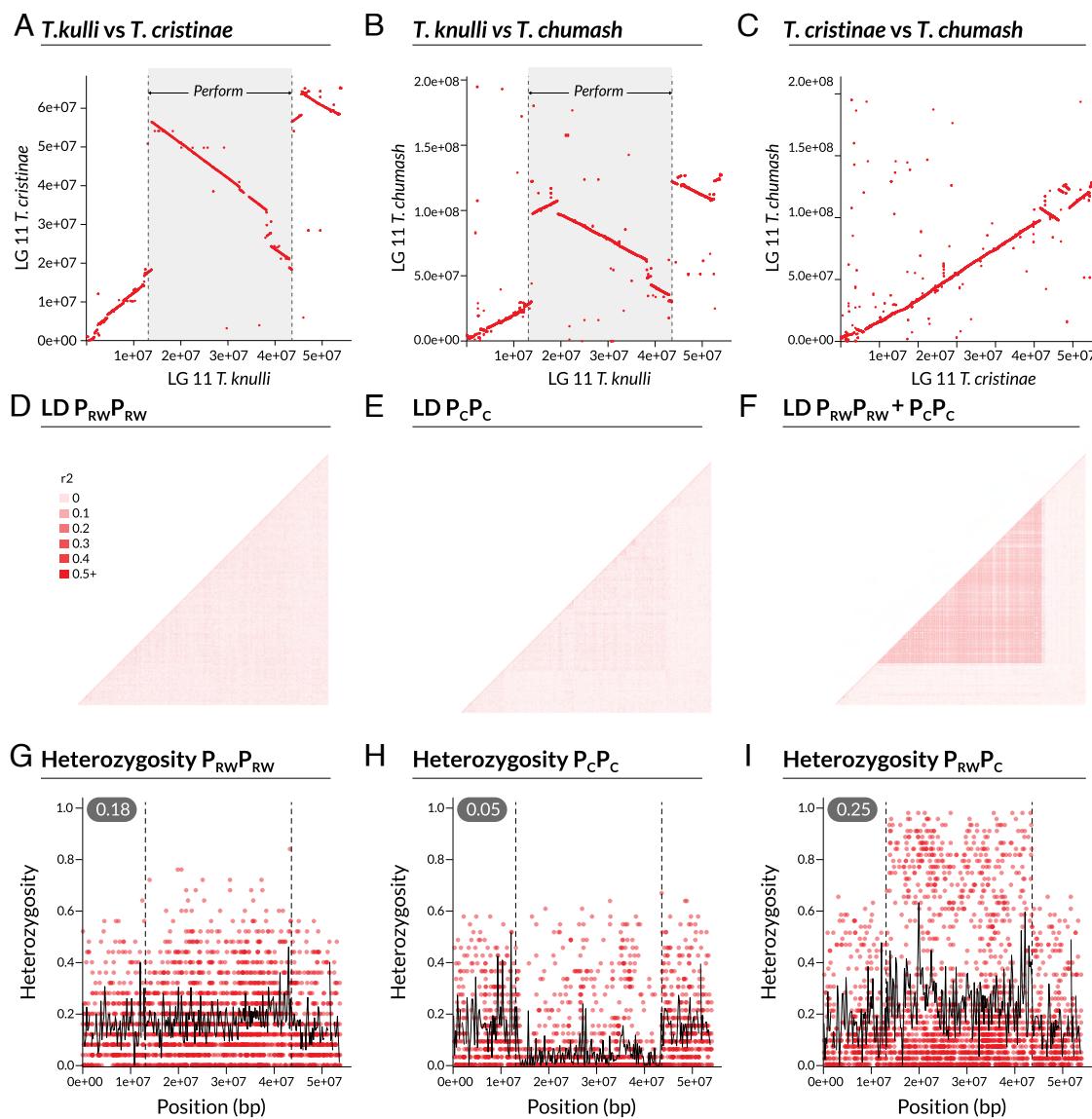


Fig. 3. Genome alignments and population genomic evidence that *Perform* is an inversion. Dot plots show alignments of chromosome 11 for *T. knulli* and *T. cristinae* (A), *T. knulli* and *T. chumash* (B), and *T. cristinae* and *T. chumash* (C). Red line segments denote aligned genome regions with the orientation of the alignment shown by the direction of the lines. The bounds of the *Perform* locus in the *T. knulli* genome are denoted by the gray shaded region. A large inversion coinciding with the *Perform* locus is evident between *T. knulli* and both *T. cristinae* (A) and *T. chumash* (B), but no such inversion is found for *T. cristinae* vs. *T. chumash*. Panels (D–F) show patterns of pairwise linkage disequilibrium (LD) across chromosome 11 in *T. knulli* individuals homozygous for the *Perform* P_{RW} allele (D), homozygous for the P_C allele (E), and the combined sample of individuals homozygous for either allele (F). Panels (G–I) depict mean observed heterozygosity for each *Perform* genotype: $P_{RW}P_{RW}$ (G), $P_C P_C$ (H), and $P_{RW}P_C$ (I). Points denote mean observed heterozygosity for individual SNPs, solid lines denote mean observed heterozygosity for 200 kbp windows, and the vertical dashed lines delimit the *Perform* locus.

the *Perform* locus and the inversion boundaries identified between species (but did not correspond precisely with these boundaries). Collectively, these genome alignments and population genomic results are consistent with expectations for the inversion (or more complex structural variation including an inversion) also segregating within *T. knulli*.

The *Perform* Locus Inversion Affects Performance on Different Hosts. We next considered the evolutionary processes potentially maintaining the inversion polymorphism. Specifically, to connect the inversion polymorphism to fitness, we tested if performance on *Ceanothus* and Redwood is affected by the *Perform* locus. Such an association with fitness would refute strict neutrality with regard to the evolution of inversion frequencies. To do so, we collected *T. knulli* and reared them in the laboratory on either *Ceanothus* or Redwood, measuring growth and survival (notably these are the same individuals analyzed above to delimit the *Perform* locus). We focus our analyses here on specimens from the vicinity of the locality BCE, where *T. knulli* uses both *Ceanothus* and Redwood (we thus exclude population BCTURN, which uses only *Ceanothus*) (*SI Appendix*, Tables S3 and S4). These experiments revealed that the *Perform* locus explains appreciable and significant variation in both growth and survival (see details in the next paragraph), but with a trade-off between these fitness components that suggests the possibility of an overall heterozygote advantage (i.e., overdominance) on *Ceanothus* (Fig. 4, *SI Appendix*, Figs. S1 and S7).

Specifically, our experiments revealed that one allele at the *Perform* locus, P_{RW} , was associated with increased growth on both *Ceanothus* and Redwood (this is the allele at a high-frequency on

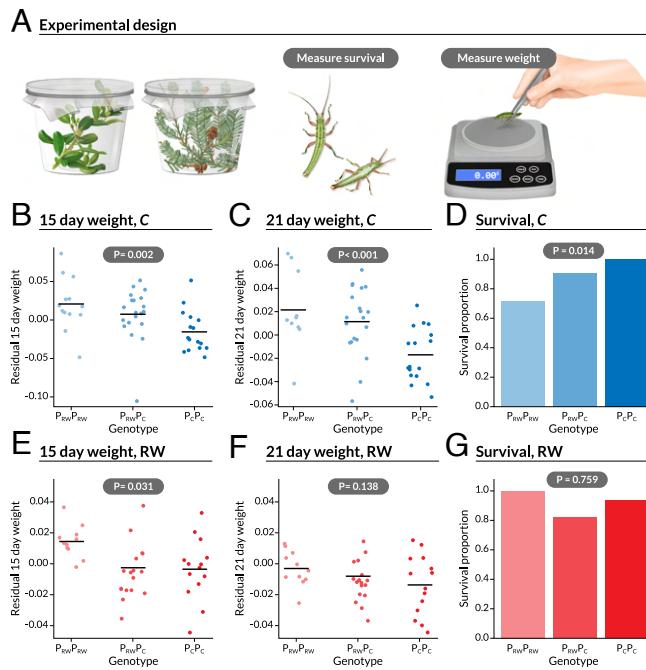


Fig. 4. Summary of the rearing and genetic mapping experiments. Panel (A) illustrates the experimental design. Panels (B) and (C) show 15- and 21-d weight for *T. knulli* reared on *Ceanothus* based on their *Perform* genotype ($P_{RW}P_{RW}$, P_CP_{RW} , and P_CP_C). Points denote individuals (with a small jitter applied to the x-axis), horizontal lines give means for each genotype. The P -value for the null hypothesis of no effect of *Perform* is shown. A barplot (D) shows survival proportions on *Ceanothus* along with the P -value for the null model of no effect of genotype on survival. Analogous results are shown for *T. knulli* on Redwood (RW, *Sequoia*) in (E) (15-d weight), (F) (21-d weight) and (G) (survival).

Redwood; linear regression on residuals after removing effect of sex; *Ceanothus* 15 d weight, $\beta = 0.018$, $r^2 = 0.178$, $P = 0.002$; *Ceanothus* 21 d weight, $\beta = 0.020$, $r^2 = 0.233$, $P < 0.001$; Redwood 15 d weight, $\beta = 0.0086$, $r^2 = 0.109$, $P = 0.031$; Redwood 21 d weight, $\beta = 0.0072$, $r^2 = 0.053$, $P = 0.138$) (see *SI Appendix*, Tables S5 and S6 for model comparison and evaluation). Critically, this same allele negatively affected survival on *Ceanothus* (linear regression, $\beta = -0.14$, $r^2 = 0.115$, $P = 0.014$), representing a host-specific life-history trade-off (see *SI Appendix*, Table S7 for model comparison and evaluation). Notably, there was suggestive evidence that this latter result was sex-dependent, with P_{RW} most markedly decreasing male survival (for individuals homozygous for this allele, 86% of females survived but only 57% of males survived; *SI Appendix*, Table S8, Fig. S7). Comparable results were observed using generalized linear models (GLM) for survival rather than simple linear regression (GLM survival on *Ceanothus*, $\beta = -1.71$, $P = 0.031$), demonstrating that the results are robust to methods of analysis. Thus, there is a fitness trade-off between growth and survival at the *Perform* locus. This trade-off raises the possibility that this locus could exhibit an overall heterozygote advantage across life history stages, at least on *Ceanothus*, even if the heterozygote always has intermediate fitness within life history stages. We evaluate this hypothesis with population genetic models below.

The Inversion Is Maintained by Complex Selection and Gene Flow. The results above suggest that genetic variation at the *Perform* locus could be maintained, in part, due to life-history trade-offs that vary with host and possibly sex resulting in an overall heterozygote advantage (a form of balancing selection, *SI Appendix*, Fig. S1). Moreover, selection appears to be shifted between populations feeding on different hosts. Specifically, the P_{RW} allele that confers higher growth (but reduced survival on *Ceanothus*) is at higher frequency in nature on Redwood (84%) than on *Ceanothus* (34%). Thus, there is a marked (~50%) allele frequency difference between populations on different hosts. We suspect that this reflects the previously documented difficulties *Timema* have using Redwood in laboratory experiments (34); use of Redwood favors P_{RW} (the “growth” allele) to make “a go of it” on this challenging host (at the same time P_{RW} does not appear to compromise survival on Redwood per se). In contrast, growing on *Ceanothus* is easy for *Timema* such that the survival advantage is more important than a growth benefit, leading to a high frequency of the allele associated with increased survival on *Ceanothus* (P_C). Thus, a shift in selection appears to result in divergent allele frequencies (i.e., adaptation) between hosts. Thus, divergent selection combined with gene flow between hosts could also play a role in maintaining variation within populations, especially on Redwood. In principle, divergent selection with gene flow could maintain variation even without heterozygote advantage, that is via a balance between directional selection that acts in divergent directions between hosts and gene flow (*SI Appendix*, Fig. S1). We next used evolutionary modeling to quantify these possibilities and their effects on the maintenance of variation. We did not consider models of underdominance, as underdominance is unlikely to contribute to the maintenance of inversion polymorphism (39).

Specifically, we used approximate Bayesian computation (ABC) to estimate the probability of population genetic models that included genetic drift and gene flow (as inferred from putative neutral loci; see Fig. 2C) and either heterozygote advantage or divergent (between hosts) directional selection on the *Perform* locus (Fig. 5 A and B). We modeled evolution

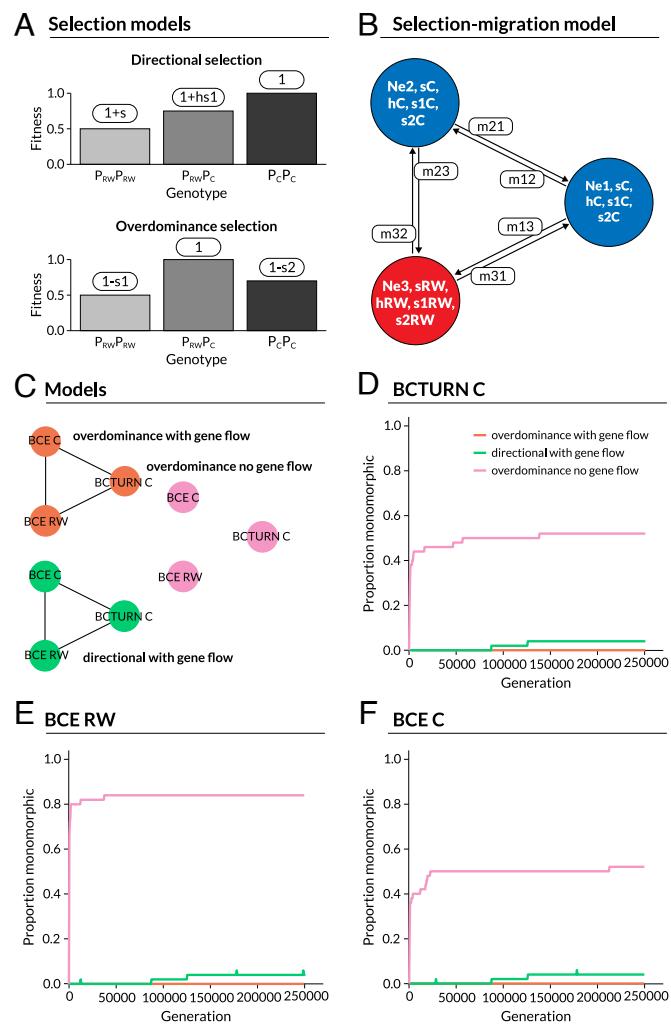


Fig. 5. Summary of the approximate Bayesian computation (ABC) model and inferences and simulations testing the effects of selection and gene flow on the maintenance of polymorphism at the *Perform* locus. Panel (A) illustrates the fitness schemes and definitions of selection coefficients under directional selection versus heterozygote advantage. Specifically, for directional selection s denotes the difference in relative fitness for alternative homozygotes and h gives the heterozygote effect (with $0 < h < 1$), whereas for heterozygote advantage s_1 and s_2 denote the reductions in fitness for homozygotes relative to the heterozygote. Panel (B) summarizes the demographic component of the model. Colored circles correspond with populations with colors denoting host, red = Redwood and blue = *Ceanothus*. Populations have distinct effective population sizes (N_e) and selection coefficients (either s and h or s_1 and s_2) dictated by host (RW = Redwood or C = *Ceanothus*). Asymmetric gene flow is allowed as indicated by the migration edges. Panel (C) illustrates the three models—heterozygote advantage with gene flow (our best model), divergent directional selection with gene flow, and heterozygote advantage without gene flow—that we consider for the three focal populations analyzed with the ABC model, BCE on *Ceanothus* (BCE C), BCE on Redwood (BCE RW) and BCTURN (an allopatric *Ceanothus* population). Panels (D)–(F) show the proportion of replicate simulations in which variation at *Perform* was lost over time, that is the proportion of monomorphic replicates at time, in BCTURN (D), BCE RW (E) and BCE C (F).

of the *Perform* locus inversion alleles, not the DNA sequence variation within this genomic region. We did this because of the evidence for selection on the inversion alleles from our experiment and our interest in the maintenance of this inversion polymorphism rather than on nucleotide variation within the inversion. These models included adjacent (i.e., parapatric) *Ceanothus* (BCE C) and Redwood (BCE RW) populations and an allopatric *Ceanothus* population (BCTURN)

(*SI Appendix*, Table S3), with the latter being important to help parse the roles of heterozygote advantage versus gene flow in maintaining variation. Models with heterozygote advantage were most probable on both *Ceanothus* (posterior probability = 0.897) and Redwood (posterior probability = 0.683), and a model of (divergent) directional selection on both hosts was very unlikely (posterior probability = 0.023) (*SI Appendix*, Fig. S8). Under the most probable model of heterozygote advantage on both hosts (posterior probability = 0.603), relative fitnesses of *Perform* homozygotes (when heterozygote fitness is set to 1.0) were 0.81 for the $P_{RW}P_{RW}$ homozygote (i.e., the homozygote for the allele conferring the growth advantage) and 0.94 for the $P_C P_C$ homozygote on *Ceanothus* versus 0.98 and 0.64 for the $P_{RW}P_{RW}$ and $P_C P_C$ homozygotes on Redwood (*SI Appendix*, Fig. S8). Thus, the population genetic model-fitting analysis strongly supports heterozygote advantage on *Ceanothus*, consistent with the life-history trade-off observed on this host in the rearing experiment. The model-fitting analysis also suggests possible heterozygote advantage on Redwood, though this was not evident from the experiment and the estimated relative fitnesses of the heterozygote (1.0) and the fitter homozygote (0.98 for $P_{RW}P_{RW}$) were very similar.

The Combination of Processes Buffers Populations Against the Loss of Variation. We have shown that gene flow and heterozygote advantage (at least on *Ceanothus*) together can explain the observed polymorphism at the *Perform* locus, but it is unclear whether both processes are necessary for the maintenance of variation in this system. In other words, does this combination of processes maintain variation that would be lost with either process in isolation? To address this question, we simulated evolution under our best model of heterozygote advantage (on both hosts) and gene flow and under two counterfactual models—one with gene flow and divergent directional selection between hosts and one with heterozygote advantage but no gene flow (all broadly considered forms of balancing selection as they maintain variation). Thus, whereas all three models result in a form of balancing selection, the latter two models eliminate heterozygote advantage or gene flow, respectively (i.e., each leaves out one process that can help maintain variation). For all models, we used selection coefficients estimated from the ABC analysis (assuming either heterozygote advantage or divergent directional selection) and gene flow inferred from neutral models based on genome-wide SNP data (except where gene flow was set to 0).

Replicate simulations, each spanning 250,000 generations, showed that the heterozygote advantage with gene flow model routinely maintains variation and predicts the observed data extremely well (Fig. 5, *SI Appendix*, Fig. S9). Divergent selection with gene flow also maintained variation over this time interval in all but a few simulations, but failed to recover the observed *Perform* allele frequencies as well as the heterozygote advantage (overdominance) with gene flow model (*SI Appendix*, Fig. S9). Finally, variation was lost in many of the heterozygote advantage without gene flow simulations. Thus, these simulations suggest that gene flow among populations feeding on different hosts and experiencing different selection pressures is important for the long-term maintenance of variation at *Perform*, and that gene flow combined with heterozygote advantage is particularly effective at generating balancing selection and preventing the loss of polymorphism.

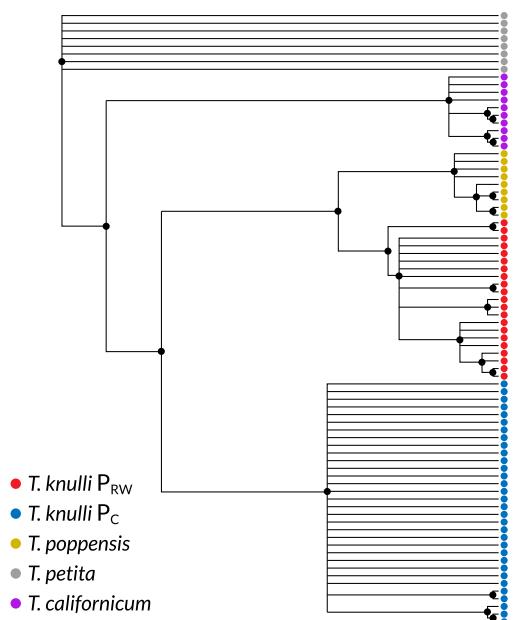
The Chromosomal Inversion Is Ancient. Lastly, we estimated the age of the Redwood inversion to test the hypothesis that

it might be young and in the act of sweeping rather than a polymorphism maintained over the long-term (Fig. 1). To do so, we first used a phylogenetic approach to estimate the divergence time between the *T. knulli* chromosomal variants. Our inferences were based on SNP data within the *Perform* locus for *T. knulli*,

T. poppensis, *T. petita* and *T. californicum* and species divergence time estimates from a published, time-calibrated phylogeny (35). This revealed that the inversion is ancient, inconsistent with the young and sweeping hypothesis. Specifically, the divergence time between Redwood and *Ceanothus* chromosomal variants in *T. knulli* was estimated as 7.5 million years ago, MYA hereafter (90% equal-tail probability intervals [ETPI] = 3.4–13.5 MYA) (Fig. 6). We also generated a complementary estimate of this divergence time using a population genetic approach based on the site-frequency spectrum and allowing for recombination between inversion haplotypes. Our estimate of the divergence time using this approach implemented in $\delta\text{a}\delta\text{i}$ (40) was 5.0 MYA (95% block-jackknife CI lower bound = 1.9 MYA) (see *SI Appendix*, Fig. S10 for model performance and *SI Appendix*, Table S9 and Fig. S11 for model parameter estimates), which is broadly consistent with the phylogenetic estimate above.

Furthermore, our results from the phylogenetic analysis suggest that the deep divergence between Redwood and *Ceanothus* alleles in *T. knulli* is not due to recent introgression from the closely related species *T. poppensis*, which feeds on Redwood and other confiers, or from another unknown or even extinct species. First, in terms of possible introgression from *T. poppensis*, the *T. poppensis* *Perform* DNA sequences were indeed more closely related to the inverted Redwood *T. knulli* alleles (P_{RW}) than the *Ceanothus* *T. knulli* alleles (P_C), but the divergence time between *T. poppensis* and *T. knulli* Redwood alleles was 4.7 MYA (90% ETPI = 2.1 to 9.8 MYA). This corresponds roughly to the previously inferred divergence time between these two species based on genome-wide SNP data (4.1 MYA, 90% ETPI = 2.3 to 6.7 MYA) (35). Thus, while *T. poppensis* appears to share DNA sequence similarity at the *Perform* locus with the *T. knulli* Redwood alleles (P_{RW}) and there is uncertainty regarding whether the origin of the inverted Redwood allele predates the split between *T. knulli* and *T. poppensis* (the posterior probability of this is 0.60), our results suggest that the Redwood allele in *T. knulli* diverged from *T. poppensis* millions of years ago. In terms of introgression from other unknown or extinct species, our phylogenetic analyses yielded estimated times to the common ancestor of P_{RW} and P_C haplotypes within *T. knulli* of 3.4 MYA (90% ETPI = 1.5 to 7.5 MYA) and 2.5 MYA (90% ETPI = 1.2 to 5.4 MYA), respectively. This means that sequence variation within each of these chromosomal variants is millions of years old within *T. knulli*. Thus, even if one of the inversion variants introgressed from another species, our results still suggest that the polymorphism has been maintained within *T. knulli* for a million years or more.

A Tree for *Perform*



B Divergence times

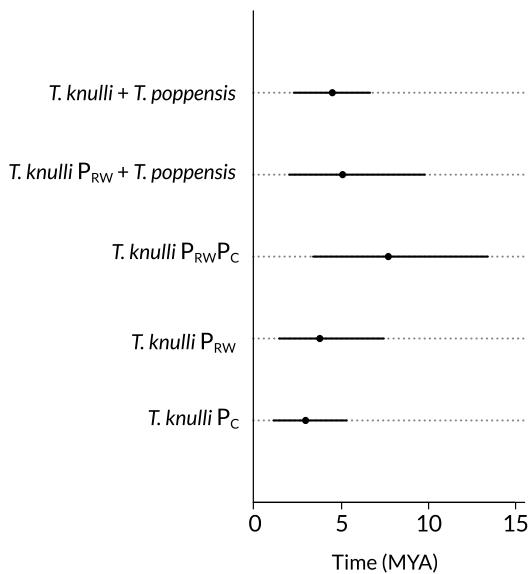


Fig. 6. Evidence that *Perform* is an ancient inversion. Panel (A) shows the phylogeny for the *Perform* locus estimated with BEAST2. Colored points indicate taxa and inversion alleles (for *T. knulli* only) (P_{RW} and P_C). Bifurcations with posterior probability > 0.5 are shown with pie charts colored to denote posterior probabilities. Panel (B) shows the corresponding Bayesian posterior distributions for divergence times for *T. knulli* and *T. poppensis* based on genome-wide SNP data, for the *T. knulli* P_{RW} chromosomal variant and *T. poppensis* based on SNPs within the *Perform* locus, and for SNPs within the *T. knulli* P_{RW} and P_C *Perform* chromosomal variants together and each alone. Points and horizontal lines denote posterior medians and 95% equal-tail probability intervals [ETPIs], respectively.

Discussion. Genetic variation is the ultimate fuel for evolution, but it remains unclear whether and how it can be maintained for extended periods of time (1–3, 5, 7, 8). The maintenance of variation is particularly puzzling given that drift and many forms of natural selection tend to erode variation, depleting evolution's fuel reservoir. Our results have broad implications for understanding the long-term maintenance of genetic variation, and the capacity to adapt to challenging environments. We discovered an ancient chromosomal inversion that likely facilitates the use of a challenging host plant (Redwood) and has been maintained for millions of years in *Timema* stick insects. We combined genomics, experiments, and evolutionary modeling to elucidate the processes maintaining variation, finding a role for life-history trade-offs, heterozygote advantage, local adaptation to different hosts, and gene flow among populations (we ruled out recent introgression from other species). We then used models

to show how such multi-layered regimes of selection and gene flow provide resilience that buffers populations against the loss of genetic variation, maintaining future evolutionary potential.

Genetic and ecological mechanisms for the origin and maintenance of inversion polymorphisms have been identified in a number of species. For example, recent work in *Heliconius numata* butterflies showed that inversions can evolve heterozygote advantage through the accumulation of deleterious mutation load resulting in associative overdominance (19). Other systems provide evidence of inversion polymorphisms maintained by life-history trade-offs, which are sometimes sex specific (24). Yet other systems suggest environmental heterogeneity plays a role in maintaining inversions, especially in terms of inversion clines or inversion polymorphism at the species level (20, 22, 41). Each of these mechanisms acting in isolation can contribute to the maintenance of inversion polymorphism. Our results add to this body of knowledge by explicitly demonstrating how multiple mechanisms operating within and across populations combine to enhance the persistence of inversion polymorphisms. Such a combination of processes, including gene flow, could be especially important for the persistence of inversion polymorphism over very long time periods given the potential for stochastic loss, especially in small populations.

Beyond the maintenance of variation, our results have implications for understanding local adaptation. This process is a hallmark of evolution and is known to be common, but its dynamics remain poorly understood because studying such dynamics often requires genetic analyses of adaptive mutations, whose identification has only recently become more feasible (42–44). In this context, adaptation might involve the fixation of mutations that are beneficial in a new environment. This is a “directional selection” hypothesis, often invoked in classical population genetics thinking and models (45, 46). Alternatively, adaptation may involve shifts in allele-frequencies rather than fixation, representing disruption of a pre-existing evolutionary balance (47–50). Such shifts might stem from standing genetic variation, and could occur via changes in the weight of balancing selection that maintains alternative alleles, slightly towards one allele or the other. This is a “shifting balancing selection” hypothesis, often emphasized in the ecological genetics literature (11). Our results are also broadly consistent with this latter hypothesis. Still, the functional significance of the *Perform* inversion, including the mechanisms responsible for the host-specific effects on performance, remains to be resolved. For example, the inversion could contribute to adaptation by suppressing recombination among linked loci that affect performance thereby creating a supergene, or the breakpoint mutations of the inversion could themselves be responsible for the observed fitness effects of this structural variant (15, 32, 51, 52). Additionally, the low level of heterozygosity for SNPs within the *P_C* allele of *Perform* suggests the possibility of a soft, partial or ongoing selective sweep limited to this genetic background (a recent and complete hard sweep would be inconsistent with the inferred age of this chromosomal variant), but this hypothesis remains to be tested.

Our results are also relevant for understanding the spatial context of evolution, namely the potential for gene flow and recombination between populations. Specifically, at spatial scales allowing gene flow, recombination will occur between populations. This can result in the breakdown of adaptive gene combinations, frustrating the ability of divergent selection to generate multi-locus local adaptation (53). Thus, factors that reduce recombination, such as chromosomal inversions, are predicted to evolve when gene flow occurs (15, 54). Gene

flow is also relevant as it can modulate the degree to which alleles can move around in space and time, as increasingly documented in cases of adaptive introgression (21, 55–57). We here demonstrated a key role for gene flow in the maintenance of genetic variation, but further work is required to test its role in the origin of inversion polymorphism.

In conclusion, although inversion evolution has received much recent attention (18–22, 32, 58), studies that directly elucidate the processes affecting inversions are still few (12, 23, 24). This makes it difficult to connect data and theory, and precludes objective evaluation of ideas that have emerged over the last century concerning the evolutionary dynamics and role of inversions. Studies estimating selection on inversions are needed, and we provided such a study here, thereby elucidating how genetic variation can be maintained for millions of years. We find that rather than being a nuisance, complexity of evolutionary processes can generate resilience that buffers populations against the loss of variation. Further studies of the maintenance of ancient genetic variants, including inversions, are required to solidify the general importance of combinations of multiple evolutionary processes for maintaining genetic variation.

Materials and Methods

Measuring Host-Associated Genetic Differentiation. We used previously published single nucleotide polymorphism (SNP) data, obtained by genotyping-by-sequencing (GBS), to quantify host plant-associated genetic differentiation in four *Timema* species—*T. californicum*, *T. knulli*, *T. landelsensis*, and *T. poppensis*. All four are sexual species from the monophyletic “Northern” *Timema* clade that live in the vicinity of Redwood and use multiple hosts in nature (35). We focused our analyses on six pairwise comparisons of nearby (i.e., parapatric) populations on different hosts (*SI Appendix, Table S1*). Genomic data from these populations were originally described by (35). Here, we used SNPs and associated genotype likelihoods (from vcf files) generated through a more recent re-analysis of these genomic data by ref. 36. See <https://github.com/zgompert/TimemaRW> for code for this and other analyses described below.

We first estimated allele frequencies in each population at each of 1139 to 8548 genome-wide SNPs (*SI Appendix, Table S1*). This was done using the program estpEM (version 0.1) (59) (Dryad, <https://doi.org/10.5061/dryad.nq67q>), which implements the expectation-maximization (EM) algorithm from (60) to estimate allele frequencies while accounting for uncertainty in genotypes as expressed by genotype likelihoods. We used a convergence tolerance of 0.001 and allowed for a maximum of 40 EM iterations. Then, for each pair of populations, we computed $F_{ST} = (H_T - H_S)/H_T$ for each SNP, and then summarized the distribution of F_{ST} for each linkage group (as defined by the *T. cristinae* genome to which these data were aligned) by computing the mean and various percentiles. Here, H_S and H_T denote the expected heterozygosities for the (sub)populations and the total, respectively. These calculations were performed in R (version 4.0.2).

Generating the *T. knulli* Genome and Assigning Chromosome Numbers. We generated a de novo reference genome for *T. knulli* using a combination of PacBio and Illumina reads from Chicago and Hi-C genomic libraries. DNA extraction, library preparation, DNA sequencing, and de novo genome assembly were performed by Dovetail Genomics (now Cantata Bio). Specifically, 105.4 Gbp of PacBio data (~75.2 × coverage) were generated over two SMRT cells and used to build an initial assembly with the Falcon assembler (with default options). This initial assembly was further improved and scaffolded using the HiRise assembler (with default options). To do so, 58.8 Gbp of Chicago DNA sequence data and 72.9 Gbp of Hi-C DNA sequence data were generated on a HiSeqX machine (150 bp paired-end reads). A total of three female stick insects were used for the assembly, and the individuals were chosen based on a preliminary analysis (i.e., PCA on genotypes obtained by GBS) that suggested they were homozygous for the *Perform* *P_{RW}* allele. The final assembly created using Dovetail’s HiRise

Assembly pipeline comprised 1,322,373,696 base pairs (bps) with an N50 of 83,614,905 bps. Based on BUSCO version 4.0.5 with the eukaryota_odb10 database (70 species, 255 BUSCOs), the assembly included 216 complete BUSCOs (212 single copy and four duplicated; 84.7%), 15 fragmented BUSCOs (5.9%) and 24 missing BUSCOs (9.4%). We then used the BRAKER2 pipeline to annotate this genome (61); see "Genome annotation" in *SI Appendix* for additional details.

Much of our recent work in *Timema*, including the analyses of host-associated genetic differentiation described in the previous section, has relied on a *T. cristinae* reference genome and associated linkage map, with each linkage group comprising multiple moderately large scaffolds (version 1.3c2; this genome comes from a melanic stick insect; see refs. 7 and 31). We wished to identify chromosomes (scaffolds) homologous to the *T. cristinae* linkage groups in our *T. knulli* genome for consistency in chromosome (linkage group) names and numbering. To do this, we first compared the *T. cristinae* reference plus linkage map to a more recent yet published *T. cristinae* genome from a green striped stick insect, which was constructed based on proximity ligation of DNA in chromatin and reconstituted chromatin (Hi-C) and comprised 13 large scaffolds, each corresponding to one of the 13 *T. cristinae* chromosomes (32). We then used cactus (version 1.0.0) to align the *T. knulli* genome to the green striped *T. cristinae* genome (62, 63). There was a one-to-one correspondence between *T. cristinae* and *T. knulli* chromosomes with one exception, *T. cristinae* chromosomes 1 and 3 were represented by a single fused chromosome in *T. knulli* (hereafter chromosome 1), consistent with cytological work showing that *T. knulli* has one fewer chromosome than *T. cristinae* (38). Thus, we were able to map our older *T. cristinae* linkage map numbers to the large scaffolds (i.e., chromosomes) in *T. knulli*. See "Assigning chromosome numbers" in *SI Appendix* for additional details.

Timema knulli Sample Collection. In spring 2019 (March 16–18), we collected 138 *T. knulli* for population genomic analyses and for use in a performance rearing experiment (described below). Most stick insects were collected at one of two localities—BCTURN, where *Ceanothus* is the main host and *Timema* are not found on Redwood ($N = 37$), and BCE where both *Ceanothus* and Redwood are hosts ($N = 68$ and $N = 24$, respectively) (*SI Appendix, Table S1*). Ten additional *T. knulli* were collected from three localities near BCE; BCOG ($N = 1$ on *Ceanothus*), BCSH ($N = 1$ on Redwood) and BCXD ($N = 8$ on *Ceanothus*). Stick insects were collected in sweep nets by beating host plants with a stick, as in past work (35, 64). Captured insects were placed in plastic tubes, and kept in a cooler with ice for 1 to 2 d during transplantation to the laboratory for use in the performance experiment, as detailed below.

DNA Extraction, Library Preparation and Sequencing. After the performance experiment (see details below), we isolated DNA from each of 138 *T. knulli*. Frozen legs from each individual were ground into powder form using a Qiagen TissueLyser (Qiagen Inc., Valencia, CA). Genomic DNA was then extracted using Qiagen DNeasy Blood and Tissue kits, using a protocol with slightly altered incubation temperatures and times. We used a reduced-representation technique (i.e., genotyping-by-sequencing or GBS) to construct DNA sequencing libraries following the protocol detailed in ref. 65. Genomic DNA from each individual was digested with two restriction endonucleases, MseI (four base recognition site) and EcoRI (six base recognition site). Illumina adaptors with unique 8 to 10 bp DNA barcodes for each individual were ligated to EcoRI cut sites, and a base Illumina adaptor was ligated to MseI cut sites. Barcoded fragment libraries were then PCR amplified using Illumina primers and a high-fidelity proofreading polymerase (Iproof, BioRad, Hercules, CA). PCR products were pooled into a single library which was then quality screened using an Agilent BioAnalyzer automated electrophoresis device. To reduce the portion of the genome targeted for sequencing, the reduced-representation library was then size-selected for DNA fragments 350 to 450 bp in length using a Pippin Prep quantitative electrophoresis unit (Sage Science, Beverly, MA) at the University of Texas Genome Sequencing and Analysis Facility (UTGSAF). The size-selected library was then sequenced using S2 chemistry and a single lane on an Illumina NovaSeq 4000 at UTGSAF.

DNA Sequence Alignment, Variant Calling, Filtering and Genotype Estimation. We aligned the newly acquired *T. knulli* GBS reads to our new *T. knulli* reference genome. This was done with the aln and samse algorithms

from bwa (version 0.7.17-r1188) (66). For alignment, we set the maximum number of allowed mismatches to 4, allowed only 2 mismatches in the first 20 bp of the alignment, trimmed bases with quality scores <10, and only output alignments for reads with a single, best alignment. We then used samtools (version 1.5) to compress, sort and index the alignments (67). Next, we used samtools (version 1.5) and bcftools (version 1.6) for variant calling (67). Here, we used the consensus caller (-c), applied the recommended mapping quality adjustment for Illumina data (-C 50), and only output SNPs when the probability of all individuals being homozygous for the reference allele conditional on the data was <0.01. We then used a series of Perl scripts to filter the variant set. Specifically, we only retained SNPs that met the following criteria: 2× minimum coverage per individual, a minimum of 10 reads supporting the non-reference allele, Mann-Whitney *P*-values for base quality, mapping quality and read position rank-sum tests >0.005, a minimum ratio of variant confidence to non-reference read depth of 2, a minimum mapping quality of 30, no more than 20% of individuals with missing data, only two alleles observed, and coverage not exceeding 3 SDs of the mean coverage (at the SNP level). We did not thin SNPs based on LD or physical proximity. Filtering left us with 64,650 SNPs for further analysis.

We then used the (ad)mixture model implemented in entropy (version 1.2) to obtain Bayesian estimates of genotypes (68, 69). This model uses a mixture prior on genotypes for each locus and individual based on co-estimated allele frequencies from a series of k hypothetical source populations similar to the admixture model from ref. 70. The model also accounts for uncertainty in genotypes arising from finite sequence coverage and possible sequencing errors as captured by the genotype likelihoods computed with bcftools. We estimated genotypes using Markov chain Monte Carlo (MCMC) and assuming either 2 or 3 source populations (i.e., our estimates integrate over these two possibilities). We ran 10 MCMC chains total (5 each for 2 and 3 source populations), each comprising 8,000 steps, a 5,000 step burnin and a thinning interval of 3. MCMC output was visually inspected to ensure (probable) convergence of the chains to the posterior distribution. Bayesian genotype estimates were then obtained by taking the posterior mean of the number of non-reference alleles (0, 1, or 2) for each locus and individual (these estimates are not constrained to integer values).

Delineating the Perform Locus. We used principal component analysis (PCA) to delineate the region of *T. knulli* chromosome 11 associated with host-plant use (feeding on *Ceanothus* versus Redwood), i.e., the *Perform* locus. First, we conducted separate PCA ordinations of the genetic data (centered but not standardized genotype matrixes) for the 62,093 SNPs not on chromosome 11 and the 2557 SNPs on chromosome 11. Only the PCA of chromosome 11 showed host-associated genetic structure, and thus we then focused on chromosome 11. To localize the portion of chromosome 11 exhibiting this pattern, we performed PCAs for 100-SNP sliding windows along chromosome 11. We summarized each PCA by the eigenvalue associated with the first eigenvector. Larger values coincide with greater genetic structure along this first PCA axis. All PCAs were done with the prcomp function in R (version 4.0.2). Visual inspection of the eigenvalues indicated a broad peak of high eigenvalues (accentuated structure) spanning much of chromosome 11. We fit a Hidden Markov model to the eigenvalues in R with the HiddenMarkov package (version 1.8.13) (71). We allowed for two hidden states, which we initialized with expected values equal to the 25th and 75th percentiles of the empirical (square root) eigenvalue distribution across chromosome 11. We assumed a normal distribution for the observed eigenvalues with SD initialized at half the empirical SD. We then estimated the hidden state means, SD, and transitions between hidden states using the Baum-Welch algorithm (i.e., we set initial values for means and standard deviations but these were then refined with the Baum-Welch algorithm) (72). For this, we allowed a maximum of 500 iterations and set the tolerance to $1e^{-4}$. This procedure identified high (mean square root eigenvalue = 4.6, SD = 0.45) and low (mean square root eigenvalue = 2.8, SD = 0.29) states. We then used the Viterbi algorithm for decoding, that is for inferring the most likely hidden state for each 100 SNP window (73). A single contiguous set of 100 SNP windows was assigned to the high state, which we hereafter refer to as the *Perform* locus. This region (i.e., the *Perform* locus) includes base positions 13,093,370 to 43,606,674 (i.e., ~30 megabases) of *T. knulli* chromosome 11.

Determining *Perform* is a Chromosomal Inversion. We used a series of comparative genome alignments to test the hypothesis that the *Perform* locus is an inversion. Specifically, we performed pairwise whole-genome alignments for our de novo chromosome-level reference genomes for *T. knulli* (described in this paper), *T. cristinae* (the green striped morph) (32), and *T. chumash* (31). Repetitive genomic regions were masked prior to genome alignment using RepeatMasker (version 4.0.7) and a *Timema* repeat library from (32). We ran RepeatMasker using the slow/sensitive search (-s) with the NCBI engine. We then used cactus (version 1.0.0) to align each pair of genomes (62, 63). cactus creates genome alignment graphs, which can represent genome rearrangements and copy number variation. We then used HAL (Hierarchical Alignment) tools (version 2.1) to extract synteny blocks from the genome graphs. This was specifically done with HalSynteny with the default lower bound for synteny blocks of 5,000 bps (74). We then constructed sequence alignment dot plots from the synteny blocks using R (version 4.0.2) to visualize inversions and other structural variation between species. These patterns of structural variation were compared to the bounds of the *Perform* locus, delimited within *T. knulli* as described above using the PCA approach. The comparative alignments described above demonstrated that *Perform* coincides with an inversion in the Redwood *T. knulli* allele relative to *T. cristinae* and *T. chumash*. We used Oxford Nanopore long-read sequencing (75) and additional population genomic analyses of LD and heterozygosity to verify that this genomic region is a segregating inversion within *T. knulli* (as strongly suggested by the PCA results) similar to ref. 21. See “Nanopore sequencing and structural variant calling in *T. knulli*” and “Population genomic analyses of LD and heterozygosity” in *SI Appendix* for details.

Performance Experiment. We conducted a laboratory experiment to test for a potential effect of the *Perform* locus on performance (here growth and survival) in *T. knulli* reared on *Ceanothus* or Redwood. For this experiment, each of the 138 *T. knulli* collected (see “*T. knulli* sample collection” above) were placed individually in 500 mm plastic containers, with air holes for breathing punched into the lid containers using a needle. Each stick insect was then fed fresh plant material from either *Ceanothus* or Redwood every second day (when survival was recorded, see below). Host-plant treatment was determined randomly and was independent of the host from which the stick insect was collected. We then measured weight and survival at 15 and 21 d as metrics of performance, and survival (dead or alive) was monitored every second day for the course of the 21-d experiment.

Testing for Associations Between *Perform* and *T. knulli* Performance. We next tested for an association between *Perform* genotype and weight and survival on *Ceanothus* and Redwood during the performance experiment. We used PCA and k-means clustering to assign *Perform* genotypes (following, e.g., ref. 17). Specifically, we performed a PCA of the SNP genotypes for SNPs within the *Perform* locus; this was done on the centered but not standardized genotype matrix. We then used k-means clustering with three centers to assign each individual to a cluster based on the first PC from the ordination of SNPs in the *Perform* locus. We then fit models for 15 and 21 d weight (linear models) and survival (generalized linear model with binomial response and logit link) on each host plant as a function of *Perform* genotype (i.e., we fit distinct models for each of the two host-plant treatments), source host plant, and the interaction between these variables. Here, genotype corresponds to the assigned cluster number with homozygous clusters coded as 0 and 2 and the heterozygous (intermediate on PC1) cluster coded as 1 (17). We removed the effects of sex and developmental stage (juvenile versus adult) on weight prior to the analyses by using the residuals from linear regression of weight on these factors, but included sex and genotype sex interactions as additional factors in the survival analysis. We dropped *T. knulli* from BCTURN to avoid possible confounding effects of population structure. Models were fit in R with the lm and glm functions (version 4.0.2). Alternative models with subsets of variables were compared using Akaike information criterion (AIC).

Modeling Gene Flow and Selection. We used approximate Bayesian computation (ABC) to fit and compare alternative models for selection with gene flow in the *T. knulli*-*Ceanothus*-Redwood system (76, 77). Details of this analysis are described in the “ABC inference of gene flow and selection” in

SI Appendix. Briefly, our approach involved first estimating (putative neutral) migration rates *Nm* (number of migrants per generation) between our three main populations: BCE C (BCE on *Ceanothus*), BCE RW (BCE on Redwood; parapatric with BCE C) and BCTURN (on *Ceanothus*, allopatric with respect to BCE C and BCE RW). We then fit ABC models for selection on *Perform*, with gene flow based on our estimates of gene flow. The selection models allowed for directional selection or heterozygote advantage (overdominance) on each host.

Additional Simulations Testing if a Combination of Processes Buffers Populations Against the Loss of Genetic Variation. We conducted an additional set of forward-time simulations of evolution to determine whether and to what extent our best fit model (heterozygote advantage with gene flow) maintained variation at the *Perform* locus (i.e., we conducted a predictive check of this model) and how this compared to two counterfactual models—one with directional selection and gene flow and one with heterozygote advantage and no gene flow. We used the same general model described above. For the heterozygote advantage with gene flow simulations, we sampled effective population size and gene flow parameters from the same prior distributions used for the aforementioned ABC analysis and then sampled selection coefficients from the posterior distributions inferred from ABC. For the heterozygote advantage without gene flow simulations, we did the same thing, except we set the migration rates to 0. Lastly, for directional selection with gene flow, we sampled selection coefficients from a posterior inferred from the ABC model when only considering the directional selection model (i.e., forcing directional selection). We ran 50 simulations (50 samples from the prior or posterior distributions depending on the parameters) under each of the three models with each running for 250,000 generations. Initial *Perform* allele frequencies were set to 0.5 for all simulations. These simulations were conducted in R (version 4.1.3). We then compared the outcome of these simulations to the observed variability at the *Perform* locus.

Dating the Chromosomal Inversion. We first used a phylogenetic approach to estimate the divergence time between the *Perform* chromosomal variants, i.e., alleles (as in ref. 17). For this, we used GBS data from 138 *T. knulli* described above, along with 69 newly sequenced *T. petita* (from site 101S, latitude = 35.73°N, longitude = 121.31°W), and 329 *T. poppensis* and 86 *T. californicum* originally described in ref. 32. These data were aligned to the *T. knulli* reference genome using the bwa aln algorithm (version 0.7.17-r1198) and alignments were compressed, sorted and indexed with samtools as described above for the *T. knulli* samples (66, 67). We identified variable nucleotides (SNPs) across this full set of samples but only within the *Perform* locus using samtools (version 1.5) and bcftools (version 1.6). Other than considering only the *Perform* locus, variant calling options and subsequent filtering were as described above for *T. knulli*. We then determined the number of invariant bases of each type (A, C, G, or T) within the *Perform* locus, as this information is part of the phylogenetic model. Specifically, using the samtools depth command, we determined coverage for each individual at each site within *Perform* that was not called as a SNP (even before filtering). We counted the site as invariant if we had data for at least 80% of the individuals with a mean coverage of at least 2 × per individual. This resulted in 789 variable sites (SNPs) and 18,425, 11,610, 12,007, and 18,570 invariant As, Cs, Gs, and Ts, respectively. We then used Perl scripts to convert the variant file to a nexus alignment and to choose a subset of individuals for phylogenetic analysis (the conversion scripts are from ref. 17 and are available from GitHub, <https://github.com/zgompe/TimemaRW>). Specifically, for the outgroup taxa *T. californicum*, *T. poppensis* and *T. petita*, we chose the 8 to 10 (10 for *T. petita* only) individuals with the least missing data for the aligned SNPs, and for *T. knulli* we retained 33 individuals from BCE C (host = *Ceanothus*) and 25 from the parapatric population BCE RW (host = Redwood).

We then used BEAST2 (version 2.6.6) (78) to estimate the divergence times between the *Perform* chromosomal variants in *T. knulli*. We encoded information on the invariant sites using the constantSiteWeights option. We fit the GTR sequence evolution model with rate heterogeneity that approximated a gamma distribution using four rate categories. We assumed a relaxed log-normal clock (79) with a coalescent extended Bayesian skyline tree prior (80). Following (17),

we fit a gamma distribution to the previously inferred divergence time for all four of our taxa—*T. knulli*, *T. petita*, *T. californicum*, and *T. poppensis*—using the `fittdist` function in R. This gives a gamma with $\alpha = 10.8509$ and $\beta = 0.973$, which has a mean of 11.5 million years and SD of 3.4 million years. We used this as the prior on the root divergence time and thus as a calibration point for our key divergence time of interest, that is between the two chromosomal variants in *T. knulli*. Our input XML file (`tknnulli_perform_og.xml`) is available from GitHub (<https://github.com/zgompert/TimemaRW>). We estimated the tree and associated divergence times based on 3 chains each comprising 10 million iterations. Posteriors were summarized in R.

Second, we estimated the divergence time in a population genetic context with the diffusion approximation approach implemented in `δαδι` (40). We specifically followed an approach inspired by ref. 81, which modeled recombination between subgenomes (in polyploids) as being analogous to gene flow between populations. We focused on the BCE population and designated two “populations,” each comprising individuals homozygous for one of the *Perform* inversion alleles. See “Dating the chromosomal inversion with $\delta\alpha\delta i$ ” in *S1 Appendix* for details.

Data, Materials, and Software Availability. DNA sequence data are available from the NCBI SRA ([PRJNA967016](https://doi.org/10.1101/267016)) (82). All other data, including data

from the rearing experiment, are available from Dryad (<https://doi.org/10.5061/dryad.1vhmgqzd>) (83). Scripts and computer code used for core analyses are available from GitHub (<https://github.com/zgompert/TimemaRW>) (84).

ACKNOWLEDGMENTS. This study is part of a project that has received funding from the European Research Council (ERC) to PN, under the European Union’s Horizon 2020 research and innovation program (Grant agreement No. 770826 EE-Dynamics). Z.G. was supported by the US NSF (DEB 1844941). The support and resources from the Center for High-Performance Computing at the University of Utah are gratefully acknowledged. We thank Mark Readdie and the Landels-Hill Big Creek Reserve for sampling access, under permit application 21361. Figure illustrations were provided by Rosa Ribas. This manuscript was improved by comments from Claire Mérot and anonymous reviewers, and by discussion with Tom Reimchen.

Author affiliations: ^aCEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul Valéry Montpellier 3, Montpellier 34090, France; ^bJohn Innes Centre, Norwich, NR4 7UH, UK; ^cEscuela Nacional de Estudios Superiores, Unidad Juriquilla, Universidad Nacional Autónoma de México, Querétaro 76230, México; ^dDepartment of Biology, University of Nevada, Reno, NV 89557; ^eDepartment of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556; and ^fDepartment of Biology, Utah State University, Logan, UT 84322

- R. C. Lewontin, J. L. Hubby, A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* **54**, 595 (1966).
- M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge University Press, 1983).
- J. H. Gillespie, *The Causes of Molecular Evolution* (Oxford University Press on Demand, 1994), vol. 2.
- T. E. Reimchen, Predator-induced cyclical changes in lateral plate frequencies of *Gasterosteus*. *Behaviour* **132**, 1079–1094 (1995).
- M. J. Wittmann, A. O. Bergland, M. W. Feldman, P. S. Schmidt, D. A. Petrov, Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9932–E9941 (2017).
- D. A. Marques, F. C. Jones, F. Di Palma, D. M. Kingsley, T. E. Reimchen, Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nat. Ecol. Evol.* **2**, 1128–1138 (2018).
- P. Nosil *et al.*, Natural selection and the predictability of evolution in Timema stick insects. *Science* **359**, 765–770 (2018).
- Z. Gompert, A. Springer, M. Brady, S. Chaturvedi, L. K. Lucas, Genomic time-series data show that gene flow maintains high genetic diversity despite substantial genetic drift in a butterfly species. *Mol. Ecol.* **30**, 4991–5008 (2021).
- A. H. Sturtevant, A case of rearrangement of genes in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **7**, 235–237 (1921).
- T. Dobzhansky, Adaptive changes induced by natural selection in wild populations of *Drosophila*. *Evolution* **1**, 1–16 (1947).
- E. B. Ford, “Ecological genetics” in *Ecological Genetics* (Springer, 1977), pp. 1–11.
- D. B. Lowry, J. H. Willis, A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8**, e1000500 (2010).
- E. M. Tuttle *et al.*, Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350 (2016).
- E. R. Hager *et al.*, A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science* **377**, 399–405 (2022).
- M. Kirkpatrick, N. Barton, Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434 (2006).
- A. A. Hoffmann, L. H. Rieseberg, Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* **39**, 21 (2008).
- D. Lindtke *et al.*, Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Mol. Ecol.* **26**, 6189–6205 (2017).
- D. Ayala *et al.*, Association mapping desiccation resistance within chromosomal inversions in the African malaria vector *Anopheles gambiae*. *Mol. Ecol.* **28**, 1333–1342 (2019).
- P. Jay *et al.*, Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nat. Genet.* **53**, 288–293 (2021).
- E. Durmaz, C. Benson, M. Kapun, P. Schmidt, T. Flatt, An inversion supergene in *Drosophila* underpins latitudinal clines in survival traits. *J. Evol. Biol.* **31**, 1354–1364 (2018).
- M. Todesco *et al.*, Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature* **584**, 602–607 (2020).
- E. L. Koch *et al.*, Genetic variation for adaptive traits is associated with polymorphic inversions in *Littorina saxatilis*. *Evol. Lett.* **5**, 196–213 (2021).
- D. Ayala, R. F. Guerrero, M. Kirkpatrick, Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. *Evolution* **67**, 946–958 (2013).
- C. Mérot, V. Llaurens, E. Normandeau, L. Bernatchez, M. Wellenreuther, Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.* **11**, 1–11 (2020).
- V. Llaurens, A. Whibley, M. Joron, Genetic architecture and balancing selection: The life and death of differentiated variants. *Mol. Ecol.* **26**, 2430–2448 (2017).
- J. Felsenstein, The theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.* **10**, 253–280 (1976).
- S. Yeaman, A. Jarvis, Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proc. R. Soc. B: Biol. Sci.* **273**, 1587–1593 (2006).
- J. M. Reid *et al.*, Immigration counter-acts local micro-evolution of a major fitness component: Migration-selection balance in free-living song sparrows. *Evol. Lett.* **5**, 48–60 (2021).
- P. Nosil, Reproductive isolation caused by visual predation on migrants between divergent environments. *Proc. R. Soc. London, Ser. B: Biol. Sci.* **271**, 1521–1528 (2004).
- P. Nosil, B. J. Crespi, Experimental evidence that predation promotes divergence in adaptive radiation. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9090–9095 (2006).
- P. Nosil *et al.*, Ecology shapes epistasis in a genotype–phenotype–fitness map for stick insect colour. *Nat. Ecol. Evol.* **4**, 1673–1684 (2020).
- R. Villoutreix *et al.*, Large-scale mutation in the evolution of a gene complex for cryptic coloration. *Science* **369**, 460–466 (2020).
- M. Muschick, V. Soria-Carrasco, J. L. Feder, Z. Gompert, P. Nosil, Adaptive zones shape the magnitude of preexisting reproductive isolation in Timema stick insects. *Philos. Trans. R. Soc. B* **375**, 20190541 (2020).
- C. Larose, S. Rasmann, T. Schwander, Evolutionary dynamics of specialisation in herbivorous stick insects. *Ecol. Lett.* **22**, 354–364 (2019).
- R. Riesch *et al.*, Transitions between phases of genomic differentiation during stick-insect speciation. *Nat. Ecol. Evol.* **1**, 1–13 (2017).
- S. Chaturvedi *et al.*, Climatic similarity and genomic background shape the extent of parallel adaptation in Timema stick insects. *Nat. Ecol. Evol.* **6**, 1952–1964 (2022).
- P. Nosil *et al.*, Genomic consequences of multiple speciation processes in a stick insect. *Proc. R. Soc. B* **279**, 5058–5065 (2012).
- T. Schwander, B. J. Crespi, Multiple direct transitions from sexual reproduction to apomictic parthenogenesis in Timema stick insects. *Evolution* **63**, 84–103 (2009).
- R. Faria, K. Johannesson, R. K. Butlin, A. M. Westram, Evolving inversions. *Trends Ecol. Evol.* **34**, 239–248 (2019).
- R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009).
- 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).
- R. D. Barrett *et al.*, Linking a mutation to survival in wild mice. *Science* **363**, 499–504 (2019).
- S. L. Archambeault, L. R. Bärtschi, A. D. Merminod, C. L. Peichel, Adaptation via pleiotropy and linkage: association mapping reveals a complex genetic architecture within the stickleback *Eda* locus. *Evol. Lett.* **4**, 282–301 (2020).
- R. Villoutreix *et al.*, Testing for fitness epistasis in a transplant experiment identifies a candidate adaptive locus in Timema stick insects. *Philos. Trans. R. Soc. B* **377**, 20200508 (2022).
- R. A. Fisher, *The Genetical Theory of Natural Selection* (Oxford, 1930).
- P. C. Sabeti *et al.*, Positive natural selection in the human lineage. *Science* **312**, 1614–1620 (2006).
- T. Reimchen, P. Nosil, Temporal variation in divergent selection on spine number in threespine stickleback. *Evolution* **56**, 2472–2483 (2002).
- M. K. Burke *et al.*, Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* **467**, 587–590 (2010).
- J. K. Pritchard, A. Di Rienzo, Adaptation—Not by sweeps alone. *Nat. Rev. Genet.* **11**, 665–667 (2010).
- S. M. Rudman *et al.*, Direct observation of adaptive tracking on ecological time scales in *Drosophila*. *Science* **375**, eabj7484 (2022).
- R. Villoutreix *et al.*, Inversion breakpoints and the evolution of supergenes. *Mol. Ecol.* **30**, 2738–2755 (2021).
- L. Zhang, R. Reifová, Z. Halenková, Z. Gompert, How important are structural variants for speciation? *Genes* **12**, 1084 (2021).
- J. Felsenstein, Skepticism towards *Santa rosalia*, or why are there so few kinds of animals? *Evolution* **35**, 124–138 (1981).

54. S. Yeaman, M. C. Whitlock, The genetic architecture of adaptation under migration-selection balance. *Evolution* **65**, 1897–1911 (2011).
55. K. K. Dasmahapatra *et al.*, Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).
56. E. Huerta-Sánchez *et al.*, Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194–197 (2014).
57. I. Giska *et al.*, Introgression drives repeated evolution of winter coat color polymorphism in hares. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24150–24156 (2019).
58. S. M. Schaal, B. C. Haller, K. E. Lotterhos, Inversion invasions: When the genetic basis of local adaptation is concentrated within inversions in the face of gene flow. *Philos. Trans. R. Soc. B* **377**, 20210200 (2022).
59. V. Soria-Carrasco *et al.*, Stick insect genomes reveal natural selection's role in parallel speciation. *Science* **344**, 738–742 (2014).
60. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
61. T. Brua, K. J. Hoff, A. Lomsadze, M. Stanke, M. Borodovsky, BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database *NAR Genomics Bioinf.* **3**, lqaa108 (2021).
62. B. Paten *et al.*, Cactus: Algorithms for genome multiple sequence alignment. *Genome Res.* **21**, 1512–1528 (2011).
63. J. Armstrong *et al.*, Progressive cactus is a multiple-genome aligner for the thousand-genome era. *Nature* **587**, 246–251 (2020).
64. C. P. Sandoval, The effects of the relative geographic scales of gene flow and selection on morph frequencies in the walking-stick *Timema cristinae*. *Evolution* **48**, 1866–1879 (1994).
65. T. L. Parchman *et al.*, Genome-wide association genetics of an adaptive trait in lodgepole pine. *Mol. Ecol.* **21**, 2991–3005 (2012).
66. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
67. H. Li *et al.*, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
68. Z. Gompert *et al.*, Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Mol. Ecol.* **23**, 4555–4573 (2014).
69. V. Shastry *et al.*, Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Mol. Ecol. Res.* **21**, 1434–1451 (2021).
70. J. K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
71. D. Harte, *HiddenMarkov: Hidden Markov Models* (Statistics Research Associates, Wellington, New Zealand, 2021), R package version 1.8-13.
72. L. E. Baum, T. Petrie, G. Soules, N. Weiss, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164–171 (1970).
73. G. D. Forney, The Viterbi algorithm. *Proc. IEEE* **61**, 268–278 (1973).
74. K. Krasheninnikova *et al.*, halSynteny: A fast, easy-to-use conserved synteny block construction method for multiple whole-genome alignments. *GigaScience* **9**, gja047 (2020).
75. H. Lu, F. Giordano, Z. Ning, Oxford Nanopore MinION sequencing and genome assembly. *Genomics, Proteomics Bioinf.* **14**, 265–279 (2016).
76. M. A. Beaumont, W. Zhang, D. J. Balding, Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025–2035 (2002).
77. S. A. Sisson, Y. Fan, M. Beaumont, *Handbook of Approximate Bayesian Computation* (CRC Press, 2018).
78. R. Bouckaert *et al.*, BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10**, e1003537 (2014).
79. A. J. Drummond, S. Y. W. Ho, M. J. Phillips, A. Rambaut, Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).
80. J. Heled, A. J. Drummond, Bayesian inference of population size history from multiple loci. *BMC Evol. Biol.* **8**, 1–15 (2008).
81. P. D. Blischak, M. Sajan, M. S. Barker, R. N. Gutenkunst, Demographic history inference and the polyploid continuum. *bioRxiv [Preprint]* (2022). <https://doi.org/10.1101/2022.09.15.508148> (Accessed 29 September 2022).
82. P. Nosil, Z. Gompert, Timema knulli GBS data. *NCBI SRA*. <https://www.ncbi.nlm.nih.gov/bioproject/967016>. Deposited 3 May 2023.
83. P. Nosil *et al.*, Data from: Complex evolutionary processes maintain an ancient chromosomal inversion. *Dryad, Dataset*. <https://doi.org/10.5061/dryad.1vhmgqzd>. Deposited 1 June 2023.
84. Z. Gompert, TimemaRW. *GitHub*. <https://github.com/zgompert/TimemaRW>. Deposited 1 March 2023.