Climatic similarity and genomic background shape the extent of parallel adaptation in *Timema* stick insects

Evolution can repeat itself, resulting in parallel adaptations in independent lineages occupying similar environments. Moreover, parallel evolution sometimes, but not always, uses the same genes. Two main hypotheses have been put forth to explain the probability and extent of parallel evolution. First, parallel evolution is more likely when shared ecologies result in similar patterns of natural selection in different taxa. Second, parallelism is more likely when genomes are similar because of shared standing variation and similar mutational effects in closely related genomes. Here we combine ecological, genomic, experimental and phenotypic data with Bayesian modelling and randomization tests to quantify the degree of parallelism and its relationship with ecology and genetics. Our results show that the extent to which genomic regions associated with climate are parallel among species of *Timema* stick insects is shaped collectively by shared ecology and genomic background. Specifically, the extent of genomic parallelism decays with divergence in climatic conditions (that is, habitat or ecological similarity) and genomic similarity. Moreover, we find that climate-associated loci are likely subject to selection in a field experiment, overlap with genetic regions associated with cuticular hydrocarbon traits and are not strongly shaped by introgression between species. Our findings shed light on when evolution is most expected to repeat itself.
It is now known that evolution can repeat itself but does not always do so\textsuperscript{9,10}. Parallelism has been documented at the genetic level, with striking cases of parallel evolution involving single genes of major effect both within and among species\textsuperscript{11–13}. For example, the \textit{Ectodysplasin} gene controlling body armour has repeatedly been used by numerous populations of stickleback fish during freshwater adaptation\textsuperscript{14}. Likewise, the \textit{Agouti} and \textit{Mc1R} genes control coloration in diverse organisms\textsuperscript{15–17}. In contrast to these studies of major effect genes, parallelism is less understood when evolution involves many genes of smaller diversification\textsuperscript{4–6}. For example, instances of repeated or parallel evolution in response to similar environmental pressures can provide evidence of evolution by natural selection. In contrast, idiosyncratic outcomes can support a role for chance or contingency in evolution and indicate constraints on the power of selection. The predictability of evolution also has practical implications, for example, for forecasting organismal responses to natural and human-induced environmental change\textsuperscript{7}, the planning of breeding programmes, and the design of medicines and strategies to combat disease spread\textsuperscript{8}.

It is now known that evolution can repeat itself but does not always do so\textsuperscript{9,10}. Parallelism has been documented at the genetic level, with striking cases of parallel evolution involving single genes of major effect both within and among species\textsuperscript{11–13}. For example, the \textit{Ectodysplasin} gene controlling body armour has repeatedly been used by numerous populations of stickleback fish during freshwater adaptation\textsuperscript{14}. Likewise, the \textit{Agouti} and \textit{Mc1R} genes control coloration in diverse organisms\textsuperscript{15–17}. In contrast to these studies of major effect genes, parallelism is less understood when evolution involves many genes of smaller diversification\textsuperscript{4–6}. For example, instances of repeated or parallel evolution in response to similar environmental pressures can provide evidence of evolution by natural selection. In contrast, idiosyncratic outcomes can support a role for chance or contingency in evolution and indicate constraints on the power of selection. The predictability of evolution also has practical implications, for example, for forecasting organismal responses to natural and human-induced environmental change\textsuperscript{7}, the planning of breeding programmes, and the design of medicines and strategies to combat disease spread\textsuperscript{8}.

\section*{Fig. 1 | Schematics to summarize the analyses conducted in this study.}
\textbf{a}, Diagram showing the approach to quantify overlap of top climate-associated SNP windows between a given pair of species (species 1 and species 2). Red dots denote climate-associated SNP windows for each species. We then quantify overlap in these windows between a given set of species that can share \textquoteleft2 or more\textquoteright, \textquoteleft3 or more\textquoteright and \textquoteleft4 or more\textquoteright SNPs (\textit{N}). \textbf{b}, Parallelism: diagram showing the approach to quantify excess overlap of top climate-associated SNP windows for multiple species. \textbf{c}, Experimental comparison: diagram showing two steps to identify excess overlap in climate-associated SNP windows and those that changed in an elevation-dependent manner during an experiment. Here we first identify loci/genomic regions associated with the greatest allele-frequency change in an elevation-dependent manner in an experiment as those that show exception change compared with a null expectation (denoted as \textit{X}, in dashed green line). Second, we compare whether these regions (\textit{X}) show excess overlap with the climate-associated SNP windows (\textit{N}). \textbf{d}, CHC comparison: diagram showing two steps to identify excess overlap in climate-associated SNP windows and genomic regions associated with CHCs. First, we identify loci/genomic regions associated with the greatest effect on CHC traits (denoted as \textit{C}). Second, we compare whether these regions (\textit{C}) show excess overlap with the climate-associated SNP windows (\textit{N}).
effect, although studies of genome-wide variation are beginning to fill this gap\textsuperscript{20,22}. However, evolution is not always parallel. Indeed, the probability and extent of parallelism decline as the time of divergence increases between taxa\textsuperscript{20,24}. Although this decline is well established, its probable causes are potentially complex (that is, time itself is not the causal agent controlling parallelism; rather, factors such as climate and genetics are probably involved, as outlined below and as we test here) and remain poorly resolved, particularly beyond experimental evolution experiments in microbes\textsuperscript{23,24}. Our goal here is to elucidate the factors shaping the extent of parallel evolution in the wild, focusing on quantifying parallelism at the genome-wide level.

In this context, two general hypotheses have been put forth, which are not mutually exclusive. First, parallel evolution is more likely when shared ecologies result in similar patterns of natural selection in different taxa, such as ecotypes or divergent lineages (the ‘shared ecology’ hypothesis)\textsuperscript{27,28}. Shared aspects of environmental variation can decline with time since divergence, as species (or even populations or ecotypes) come to occupy different geographic areas or as local environments change over time, thus reducing parallelism at both phenotypic and genotypic levels\textsuperscript{28,29}. Second, parallelism is expected to be more likely when genomes are similar because pools of standing variation, new mutations that arise and the effects of these mutations will tend to be more similar in closely related genomes (the ‘shared genetics’ hypothesis); we use this term to also encompass the role of gene regulation and development\textsuperscript{30}. Epistatic interactions might be particularly important here because the effects of new mutations are dependent on the mutations that preceded them.

Both ecological (that is, habitat and climatic) and genetic similarity are expected to decline with time and there is support for both hypotheses\textsuperscript{27,30–32}. However, few studies have simultaneously examined ecology and genetics, particularly in wild populations, such that the relative contribution of the two factors remains unclear. Parsing these contributions is important because it is required to test the roles of selection (that is, shared ecology) and constraint (that is, shared genetics) in evolution\textsuperscript{30–32}. Here we combine ecological data, genomic analyses, a field experiment and genetic mapping to ascertain the genetic extent and causes of parallel adaptation to climate, thus testing the shared ecology and genetics hypotheses. Rather than focusing on time per se, we conduct analyses that jointly consider the degree of climatic and genetic divergence between taxa to parse their relative contributions to explaining the degree of parallel evolution observed.

Our study system is wingless, univoltine, herbivorous stick insects in the genus *Timema*, many species of which are endemic to California, USA\textsuperscript{33}. These insects are best-studied for their cryptic colours and colour patterns, which are controlled by the same genetic region (termed *MelStripes*) in all species studied so far\textsuperscript{34–39}. *Timema* colouration thus provides a striking example of highly parallel evolution at the level of a single, largely non-recombining gene region that could be considered akin to a major effect locus. However, adaptation often involves many genes, including those with alleles of minor effect, arrayed throughout the genome\textsuperscript{40,41}, where the probability of parallel genetic evolution is less clear\textsuperscript{42}. In this context, we study a novel ecological dimension in *Timema*, namely climate, motivated by the fact that adaptation to varying climatic (abiotic) conditions of the environment can be polygenic, and the genus *Timema* inhabits variable habitats in California. For example, the occupied habitats of *Timema* range from sea-level to mountainous regions, and from arid semi-deserts near the Mexican border to wet evergreen forests in northern California\textsuperscript{42}. Moreover, there is climatic variation both within and among species, with several species being distributed along elevational gradients (ranging from 10 m to -2,800 m)\textsuperscript{40}. This creates an opportunity to test the role of climatic variables, such as precipitation and temperature (known

---

**Fig. 2** | **Map of species ranges and plots for within-species variation in climate**

The second diagram shows our prediction for the ‘shared genetics’ hypothesis where we expect a decay in parallelism with an increase in genetic distance. We use these two hypotheses to study the decay of parallelism. c–e, Boxplots of PC variation for the first three principal components (PC1, PC2, PC3) for the eight species included in the study (N = 1,420 individuals from 53 localities).
to be important determinants of selection in many organisms (52,53), in driving parallel evolution in *Timema*.

For this study, we test the shared ecology and genetics hypotheses in *Timema* to identify climate-associated gene regions within species that show a range of divergence times of up to tens of millions of years (here, generations). We assess the contribution of shared ecology and genetics to genomic parallelism by comparing the proportions of the genome that exhibit repeated genotype–climate association. We then bolster the evidence that climate-associated gene regions are likely subject to selection by using a field experiment and genetic mapping of cuticular hydrocarbons. Our collective results yield a comprehensive evaluation of genome-wide parallel evolution in the context of an environmental pressure of high current interest (that is, climate) and in a system where comparison can be made to parallelism seen at a single, major locus (that is, *Mel-Stripe*) (Fig. 1).

### Results

#### Climatic variation within and among species

We studied 8 *Timema* species across 53 geographic localities (*N* = 1,420 individuals) (Fig. 2 and Supplementary Table 1). We used 22 bioclimatic variables describing precipitation and temperature variation, which are known drivers of selection in many systems (52) including *Timema* (54). Due to high correlations among the studied climate variables, we performed an ordination using principal component analysis (PCA) of the climate variables for all populations included in the study (see Fig. 2a for species range map). This revealed that most of the variation in climate variables was explained by the first three principal components (PC) (Total = 92.2%, PC1 = 51.7%, PC2 = 24.4% and PC3 = 16.1%), which we hereafter focus on and refer to as PC1, PC2 and PC3 (see Supplementary Table 2 for PC loadings, and Extended Data Fig. 1).

One-way analysis of variance (ANOVA) revealed significant among-species variation for all three PCs (PC1: variance component 12.1%, DF = 7, *F* value = 104.5, *P* value < 0.0001; PC2: variance component 3.2%, DF = 7, *F* value = 6.803, *P* value < 0.001; PC3: variance component 3.1%, DF = 7, *F* value = 28.07, *P* value < 0.0001). We also detected clear within-species variation (range of median PC scores across the eight species included in the study). Results are shown along the 13 linkage groups. The two values in parentheses above each panel are the number of SNP windows in the top 10% quantile (‘windows’), followed by the number of linkage groups with at least 1 SNP window in the top 10% quantile (‘LGs’).
species were −3.0 and 5.8 for PC1, −2.5 to 6.5 for PC2, and −1.6 to 3.5 for PC3; Fig. 2c,d). We next used these three PCs to identify genomic regions associated with climate within species, which is a prerequisite for testing parallelism among species.

**Identifying climate-associated genomic regions**

We first identified the genomic regions most strongly associated with climatic variation within each of the eight species. To do so, we analysed single-nucleotide polymorphisms (SNPs) obtained through previous genotyping-by-sequencing (GBS) of natural populations. Since our data included species that are considerably diverged from each other, the number and fine-scale genomic positions of SNPs called for each species were different. This could be due to different evolutionary histories of the restriction sites targeted for the sequencing and genome-level divergence of species from the genome of T. cristinae. To account for this variation, we focused on 100 kilobase (Kb) SNP windows to allow subsequent comparisons among species (N=9,487 windows in each species, across the eight study species, minimum SNPs per window 1, mean SNPs per window 1.78).

Within each species, we quantified SNP–climate associations for each of the three climate PCs using BayPass (version 1.2). The association of each SNP with population-average PC variables was assessed using Bayes Factors (BF), which for a given SNP compares the marginal likelihoods of models with zero versus non-zero regression coefficients. For each species, we then calculated the median of the marginal likelihoods of models with zero versus non-zero regres-

---

**Parallel evolution of climate-associated genomic regions**

We next quantified the extent to which climate-associated SNP windows were parallel (that is, the same) across the eight species of *Timema* that we studied. Here we are interested in identifying and quantifying genomic parallelism on the basis of the 100 Kb SNP windows spread across the genome (‘genomic parallelism’ hereafter). Critically, we tested whether windows exhibited excess overlap across species relative to overlap expected by chance, that is, whether the same SNP windows show association with climate PCs between 3, 4, 5, 6, 7 or 8 species (Fig. 1b). To do so, we conducted randomization tests to quantify excess overlap of windows relative to expectations for multispecies comparisons (Fig. 1a). As an example, an x-fold enrichment of 2.0 in the genomic parallelism analyses would indicate that overlap of climate-associated SNP windows for a given comparison was two times higher than expected by chance based on the mean of the null. For this, we focused on windows with the greatest (top 10%) association of each SNP with population-average PC variables was assessed using Bayes Factors (BF), which for a given SNP compares the marginal likelihoods of models with zero versus non-zero regres-

---

**Genomic parallelism declines predictably between species**

We next tested the extent to which the shared ecology and shared genetics hypotheses could account for the degree of genomic parallelism observed with climate across *Timema* species (Fig. 2b). Shared ecology would cause a higher degree of parallelism due to similar selective pressures from similar climate conditions experienced by taxa (that is, PCs 1–3) (Fig. 2b, ‘shared ecology hypothesis’). On the other hand, shared genetics would cause a higher degree of parallelism due to a higher extent of gene re-use associated with variation retained from a common ancestor (Fig. 2b, ‘shared genetics hypothesis’). Here we quantified genomic parallelism as the degree of excess overlap of climate-associated SNP windows relative to null expectations for pairwise comparisons. We estimated climatic similarity between pairs of species using climatic data and genetic similarity based on a previously published genome-level phylogeny. We then fit Bayesian linear mixed models to explicitly compare models where the degree of parallelism is determined by climatic similarity, genetic similarity, or both. Notably, this mixed model approach accounts for the non-independence of pairwise distances (see ref. 29 for details). Specifically, for each climatic PC variable, we modelled parallelism as the x-fold enrichment of three or more species was −1.6× more than expected by chance (observed = 43, expected = 26, x-fold enrichment 1.63, P value < 0.01; Fig. 4) and almost 5× more than expected by chance for four or more species, excess overlap was −4× more than expected by chance (observed = 4, expected = 1.03, x-fold enrichment 3.87, P value = 0.02; Extended Data Fig. 4). For PC2, excess overlap of SNP windows with the largest median BF among three or more species was about −1.5× more than expected by chance (observed = 42, expected = 26.41, x-fold enrichment 1.59, P value < 0.01; Extended Data Fig. 5), and for four or more species, excess overlap was −4× more than expected by chance (observed = 5, expected = 1.19, x-fold enrichment 4.17, P value = 0.007; Extended Data Fig. 5). Lastly, for PC3 the excess overlap of climate-associated SNP windows among three or more species was −1.6× more than expected by chance (observed = 43, expected = 26, x-fold enrichment 1.63, P value < 0.01; Fig. 4) and almost 5× more than expected by chance for four or more species (observed = 5, expected = 1.10, x-fold enrichment 4.53, P value = 0.006; Fig. 4). Additional tests for historical and contemporary gene flow revealed that introgression and gene flow were not responsible for this parallelism (see Supplementary methods and results; Fig. S5a and Supplementary Figs. 7–9).
between a given pair of species (hereafter referred to as ecology, indicating ‘climatic divergence’), genetic distance, which was calculated as pairwise phylogenetic distances for a given pair of *Timema* species (hereafter referred to as genes, indicating ‘genome-wide divergence’), or both. The fit of models with or without ecology or genetics was compared using deviance information criterion (DIC) (Fig. 5b, and Extended Data Figs. 6b and 7b), which is a metric of predictive performance.

Our analyses revealed evidence for the effects of both ecology and genes on the extent of genomic parallelism, with details that varied among the climate PCs (Fig. 5c–d, and Supplementary Data Figs. 6a–c and 7a–c for PC1 and PC2, respectively). For PC3, the best fit was obtained for the full model (ecology and genes), with similar, negative effects on parallelism observed for ecology (standardized $\beta = -0.47$, 95% CI $= -0.80$ to $-0.14$) and genes (standardized $\beta = -0.55$, 95% CI $= -0.87$ to $-0.21$) (Fig. 5e and Supplementary Table 3). For PC1, the genes-only model was the best model (standardized $\beta = -0.53$, 95% CI $= -0.82$ to $-0.25$; Extended Data Fig. 6d and Supplementary Table 3). The second-best model was the full model, but this included a positive rather than negative effect of climatic distance on parallelism. Lastly, for PC2 the best model was a null model of no effect of genes or ecology on parallelism (Extended Data Fig. 7d and Supplementary Table 9). The results thus provide variable support for both the shared ecology and shared genetics hypotheses depending on the climate PC, with the association being strongest for PC3.

**Comparison of parallelism results with permuted data sets**

We next conducted permutation analyses that randomized the climatic data before implementing BayPass. We did so to ask whether the patterns of observed genomic parallelism and its decay could have been inflated by unaccounted aspects of the genetic data, such as shared SNP density in specific genomic regions, allele-frequency distributions or linkage disequilibrium, affecting some genomic regions more than others. To generate null expected distributions for climate-associated SNP windows, we therefore initially permuted PC climatic values across populations within species, thereby randomizing the relation between the environmental variables and any potential unaccounted-for feature(s) in gene regions affecting parallelism. We generated 10 such permuted data sets, hereafter referred to as ‘permuted data sets.’ We then repeated the analysis for each of the 10 permuted data sets, for each species separately, exactly as described for the observed data set. First, we reran BayPass using each of the permuted data sets and for each species. Second, we quantified the degree of genomic parallelism by making multispecies comparisons. Third, we conducted our Bayesian linear mixed models to test for the effect of ecology and genetics on the decay of genomic parallelism.

For all 3 PCs, the 10 permuted data sets showed no evidence for the decay in parallelism with increased ecological or genetic distance seen in the actual data set (Supplementary Figs. 1–3). However, the permuted data sets indicated significant x-fold enrichments of multiple species sharing climate-associated SNP windows (Supplementary Figs. 4–6). In certain instances, the parallelism involved 4 or more species, as we found significant x-fold excesses in 3 of the 10 permuted data sets for PC1, 6 of 10 for PC2, and 4 of 10 for PC3 (Supplementary Figs. 4–6). These results suggest that aspects of the genetic data could generate apparent parallelisms of gene regions responding to environmental variables across species. However, for PC3 which displayed the strongest association of climate and genetics with parallelism, the x-fold excesses in the 4 or more species comparisons in the 10 permuted data sets did not approach the level observed in the original data (Supplementary Fig. 6). Most importantly, as noted above, for the 10 permuted data sets, the pattern of excess parallelism was random across species with respect to its relationship with climatic and genome-wide divergence. Our core test of the shared ecology and shared genetic hypotheses thus appears highly robust. Having tested these hypotheses, we next tested for additional evidence, beyond genomic parallelism, that the climate-associated SNP windows have been affected by natural selection.

**Climate-associated regions experience natural selection**

To bolster the evidence that climate-associated SNP windows are enriched for genetic variants experiencing natural selection, we tested whether these windows exhibited exceptional patterns of allele-frequency change in a published transplant-and-sequence field experiment (Fig. 1c).

The transplant experiment used a block design to measure 8 d survival and associated genome-wide allele-frequency change during this period in 500 *T. cristinae* transplanted to 10 experimental bushes comprising two host plants occurring along a gradient of higher elevations than the source population for the experiment (see ref. 61 for further details). Distances between plants within a block ranged from 6 to 10 m and distances between blocks ranged from 12 to 30 m. A previous analysis of this experiment documented evidence of selection associated with elevation, which is relevant as the samples of species analysed for the current study of parallelism were distributed along elevational gradients ranging from 10 m to $-2,800$ m. Here, as a metric of possible elevation (environment)-dependent selection, we calculated the Pearson correlation between transplant elevation and allele-frequency change caused by mortality during the transplant experiment. We found that the 100 Kb windows exhibiting patterns of allele-frequency change most strongly associated with elevation in the transplant experiment coincided modestly but significantly with climate-associated SNP windows. Specifically, when focusing on the windows with the greatest (top 10%) correlation between change and elevation in the experiment and with the greatest (top 10%) climate association in nature, windows associated with all three climate PCs corresponded with those where change was most strongly associated with elevation $-1.2$–$1.3$ times more than expected under the null hypothesis of independence (constrained randomization test controlling for SNP density within windows based on 1,000 randomizations; PC1: observed = 108 shared windows, $P = 0.005$; PC2: observed = 101 shared windows, $P = 0.015$; PC3: observed = 105 shared windows, $P = 0.021$) (Fig. 6). Similar patterns were observed when more extreme top percentiles were considered and when using an unconstrained randomization test (Supplementary Table 4). These patterns are consistent with the hypothesis that multiple genetic variants in these windows are subject to selection in nature.
Additionally, we found that climate-associated SNP windows overlapped more than expected with regions associated with phenotypic variation in genetic mapping analyses of cuticular hydrocarbons (CHCs), specifically pentacosanes in females (Supplementary methods and results; Fig. 1d and Supplementary Tables 5–8), which studies of insects have shown can contribute to climate adaptation \(^{62,63}\). This, combined with the results presented above, suggests a polygenic basis for climatic adaptation in *T. cristinae*, with at least a modest correspondence between our top climate-associated windows and the actual loci involved in climate adaptation.
Discussion and Conclusion

We used GBS data from 1,420 individuals across 8 species combined with data from field transplant and genome-wide association mapping for cuticular hydrocarbons to show that adaptation to climate occurs in parallel across species but as a function of the climatic and genomic divergence between species. Our results inform five fundamental issues in biology, namely: the repeatability of evolution, variation in the degree of parallelism based on the climate variables considered, the effect of ecology and genetics on parallelism, technical aspects pertaining to the study of parallelism and the processes promoting parallelism. We treat these issues in turn below.

First, we show that evolution in response to climate occurs in parallel among eight species and that parallelism probably involves multiple SNPs. These findings fill a gap in our knowledge of parallel evolution because many studies, including past work in *Timema*14,15,47, have mostly focused on parallelism driven by single genes or specific regions of the genome. Our results agree with other cases of parallel or convergent climate adaptation that are also driven by polygenic interactions. Overall, our study demonstrates that repeatability of evolution can be driven by numerous genetic paths, but the magnitude of repeatability can be highly variable, specifically when considering inter-species comparisons.

Second, our results reveal notable variation in the degree of parallelism across the three PCs, which we use as composite climate variables. We attribute the variation in the degree of parallelism to *Timema* species occupying variable environmental niches in their geographic distributions, which can cause environmentally heterogeneous selection. Furthermore, each PC is composed of different climatic variables. Therefore, the level of genomic association and in turn parallelism would vary on the basis of the PC (and climatic variables) being considered. For example, precipitation (which is one of the top loading variables on PC1 and PC2) can affect variability in selection in space52 and has also been shown to drive thermoregulatory evolution in *Timema*.54

Other unaccounted factors can influence response to climate, such as microclimate variation on the spatial scale that *Timema* species occupy, and nonlinear gene–climate associations. All these factors together

---

**Fig. 6** Evidence for excess overlap between 100 Kb windows associated with climate in nature and those that changed in an elevation-dependent manner during an experiment. 

**a**. Scatterplot showing the mean correlation between change and elevation during an experiment versus the median Bayes factor measuring SNP–climate (PC3) association in nature for *T. cristinae* for 100 Kb windows. Points denoting windows in the top 10% for change–elevation correlations are shown in orange, those in the top 10% for SNP–climate associations are shown in blue, and those in the top 10% for both are in purple (other windows are shown in grey points). We are interested in the top right corner of the plot, that is, the purple points denoting that windows were exceptional (top 10%) in the experiment and in nature, and we used a randomization test to ask whether more windows fall in this category than expected by chance. 

**b–c.** Null expectations for the number of windows in the top 10% for the experiment and in nature based on climate PCs 1, 2 and 3, respectively. The null distribution from the constrained randomization test in each case is denoted by the grey density plot, whereas the observed value is shown with a vertical purple line. The $P$ value for the null hypothesis of no association between SNP–climate and change–elevation correlations is reported in each panel.
could contribute to the variable degree of parallelism observed across the three PCs, emphasizing that the genomic basis of adaptation to climate in *Timema* is predictable to some extent, yet complex.

Third, our results reveal that parallelism decays with climatic and genome-wide divergence, suggesting that both shared ecology and shared genetics can affect parallel evolution. Thus, the parallelism we observe in *Timema* can be partly attributed to selection pressures exerted on insects inhabiting similar niches. In addition, genetic similarity increases the chances for shared standing genetic variation in closely related taxa to allow for gene reuse in response to similar environmental pressures. Finally, parallelism was also associated with allele-frequency changes underlie parallelism were also associated with allele-frequency changes (two units). Our study demonstrates that both these aspects can affect parallelism, with a perhaps more consistent effect of genetics due to patterns of ecological variation being more complex among species compared with genetics.

Fourth, our approach involving permuted data sets highlights important issues concerning analytical aspects of parallelism tests. We found no evidence for the observed decay in parallelism with climatic or genome-wide divergence in permuted data sets conducted before or following an analysis with BayPass. Overall, these findings, in combination with the experiment and CHC results, provide support that the documented parallelism in genomic association with climate reflects a contribution from selection. However, we also note that our analyses using permuted data sets generated instances with significant x-fold excesses in the numbers of gene regions displaying parallelism above null expectations. Our findings thus concur with previous studies using simulation-based approaches showing that false positives can be detected due to unaccounted aspects of the genetic data. Therefore, we suggest that these associations should be interpreted with caution, and studies identifying genomic association with climatic variables warrant additional cross-validation of findings, as performed here.

Fifth, our collective results inform how two core evolutionary processes, namely introgression/gene flow and selection, might affect parallelism. We show that parallel evolution and adaptation to climate occurs despite limited or minimal gene flow among *Timema* species. While introgression can facilitate parallel adaptation to similar environmental pressures through the sharing of novel genetic material, a lack of introgression or gene flow drives independent instances of adaptation and the role of selection in driving parallel evolution at the genomic level. Ancestral genetic variation can also underlie parallelism due to similar selection pressures driving phenological similarity not just for newly formed and partially reproductively isolated host races, but also for distantly related sibling species. Additionally, while a study on divergent conifers has indicated that conserved genomic regions can drive convergent adaptation to climate, another study on distinct genetic clusters of *Arabidopsis lyrata* (two lineages) shows that parallelism in genomic association to climate is detectable at the gene but not the SNP level. Both these systems also have minimal gene flow. In comparison, a study on replicate pairs of threespine sticklebacks implies a substantial role for the environment and gene flow in affecting parallelism. In summary, our study shows how local adaptation among species with minimal between-species gene flow can occur and consequently be crucial for predicting evolution in response to rapidly changing environments and climate. Furthermore, our results bolster evidence for selection beyond a correlative genome scan because we found that the genomic regions that underlie parallelism were also associated with allele-frequency changes in a manipulative field experiment and climatically relevant CHCs traits. Thus, together these results suggest that allele reuse through standing genetic variation, new mutations and selection can all be powerful drivers of parallel local adaptation.

Methods

Below we describe details of our methods and analyses, and we provide a graphic summary in Fig. of the main text.

Samples and DNA sequences from natural populations

For this study, we analysed GBS data from 1,420 *Timema* stick insects belonging to 9 species from 33 localities: 6 *T. bartmani* populations (N = 195 individuals), 3 *T. californicum* populations (N = 77 individuals), 12 *T. chumash* populations (N = 358 individuals), 6 *T. cristinae* populations (N = 205 individuals), 5 *T. knulli* populations (N = 89 individuals), 4 *T. landelsensis* populations (N = 125 individuals), 12 *T. podura* populations (N = 255 individuals) and 5 *T. poppensis* populations (N = 116 individuals) (Supplementary Table 1). The GBS data for this study have been previously published in a study of the speciation continuum in *Timema*. DNA sequence data, the reference genome, experimental data and CHC data used in this study are associated with the previously published studies. The associated DNA sequence data have been archived on NCBI's SRA (Accession: PRJNA356405, ID: 356405). The genomic data in the transplant experiment used for genetic mapping of cuticular hydrocarbons are independent from these published data and are described in detail below.

Sequence alignment and variant calling

To incorporate variants typed for individuals of each species, we built a consensus reference sequence for each species similar to ref. 44,47. To do this, we first aligned all reads from all our samples to the *T. cristinae* reference genome (draft version 0.3) using the MEM algorithm of BWA (version 0.7.17-r1188). We ran BWA MEM with a minimum seed length of 15 (k), internal seeds of longer than 20 bp, and only output alignments with a quality score ≥30 (-T). We then used SAMTOOLS (version 1.5) to view, sort and index the alignments. We called variants using SAMTOOLS and BCFTOOLS (version 1.6). For variant calling, we used the mapping quality adjustment of 50 (-C), skipped alignments with mapping quality 0, skipped bases with base quality <13 and ignored insertion-deletion polymorphisms. We then set the prior on SNPs to 0.001 (-P) and called SNPs when the posterior probability that the nucleotide was invariant was <0.01 (-p). We then performed 2 rounds of filtering to retain final sets of SNPs. In the first round, we filtered the initial set of SNPs to retain only those with sequence data for at least 80% of the individuals, a mean sequence depth of 2 per individual, at least 4 reads of the alternative allele, a minimum quality score of 30, a minimum (overall) minor allele frequency of at least 5% and no more than 0.01% of the reads in the reverse orientation. In the second round of filtering, we removed SNPs with excessive coverage (2 standard deviations above the mean) or tight clustering (within 5 base pairs (bp) of each other). This left us with the following number of SNPs for each species: 10,036 SNPs for *T. bartmani*, 14,953 SNPs for *T. californicum*, 20,478 SNPs for *T. chumash*, 3,43,746 SNPs for *T. cristinae*, 25,835 SNPs for *T. knulli*, 21,314 SNPs for *T. landelsensis*, 21,986 SNPs for *T. podura* and 18,237 SNPs for *T. poppensis*.

We used these filtered variants for each species to construct consensus reference sequences for each species using the CONSENSUS algorithm of BCFTOOLS (version 1.6)44. We then used the consensus reference of each species to redo alignments of GBS sequences of individuals for each species separately. Following this, we repeated variant calling and two rounds of variant filtering as described above. This left us with the following number of SNPs for each species: 3,074 SNPs for *T. bartmani*, 7,858 SNPs for *T. californicum*, 4,172 SNPs for *T. chumash*, 196,252 SNPs for *T. cristinae*, 11,139 SNPs for *T. knulli*, 8,548 SNPs for *T. landelsensis*, 6,000 SNPs for *T. podura* and 7,157 SNPs for *T. poppensis*. We used this second set of SNPs for all downstream analyses.

Climate variables and SNP by climate association

We used 22 climate variables associated with our 53 study localities (Supplementary Table 2), these variables having been extracted from
we assessed coarse-grain (100 Kb) genomic parallelism by analysing enrichment of 2.0 would indicate that twice as many climate-associated by chance (x-fold enrichments and its decay could have been inflated (unexpectedly high numbers) due to unaccounted aspects of the genetic data. We tested this by permuting environmental variables (that is, PC scores) before running BayPass rather than just permuting BF across species. Our expectation was that a high number of false positives with the permuted environmental variables would raise a warning against the results obtained from the observed data. We generated and analysed 10 permuted
Climate-associated SNP windows and field-experiment-associated genetic regions

We quantified overlap between climate-associated SNP windows and genetic regions identified in ref. 24. This was done using an expectation-maximization (EM) algorithm as implemented in the programme estpEM (version 0.1), with a tolerance of 0.001 and a maximum of 50 EM iterations. We then used these estimates to compute allele-frequency change between the start and end of the experiment. Then, for each SNP we calculated the Pearson correlation between allele-frequency change and the elevation at each of the 10 transplant sites. Finally, we determined the average correlation between change and elevation for the 100 Kbp windows across the genome. Windows with fewer than 4 SNPs were ignored. These steps were done using R (version 3.4)26.

We then calculated the number of 100 Kbp windows that were among the top 10% for both elevation-dependent change during the experiment (highest average absolute correlation) and for climate association (highest average BF for each climate PC). We used a constrained randomization procedure to generate null expectations for such concordance between change and climate-association windows, using a separate randomization for each PC. Specifically, we randomized mean change metrics across windows, but only among windows with similar SNP densities (10 equally sized bins were used for this purpose). This was done because we observed a positive correlation between SNP density and mean change—elevation correlations per window (Pearson R = 0.069, 95% CI = 0.047–0.091, P < 0.001), and we wanted to control for this. Null distributions and P values were based on 1,000 randomizations and are reported for each climate PC.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The climate data and PCA scores used for analyses, a list of accession numbers of sequences used in this study, genotype likelihood files, variant calling format files and associated scripts, input file and scripts for programmes such as Entropy, BayPass and Treemix are available on the DRYAD repository (https://doi.org/10.5061/dryad.51c59zwbr).

References

1. Gould, S. J. Wonderful Life (Radius, 1990).
2. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. Proc. Natl Acad. Sci. USA 105, 7899–7906 (2008).
3. Meyers, P. J. et al. Can the genomics of ecological speciation be predicted across the divergence continuum from host races to species? A case study in Rhagoletis. Phil. Trans. R. Soc. Lond. B 375, 20190534 (2020).
4. Stern, D. L. & Orgogozo, V. Is genetic evolution predictable? Science 323, 746–751 (2009).
5. Langerhans, R. B. Predicting evolution with generalized models of divergent selection: a case study with poeciliid fish. Integr. Comp. Biol. 50, 1167–1184 (2010).
6. Losos, J. B. Convergence, adaptation, and constraint. Evolution 65, 1827–1840 (2011).
7. Waldvogel, A.-M. et al. Evolutionary genomics can improve prediction of species’ responses to climate change. Evol. Lett. 4, 4–18 (2020).
8. Lieberman, T. D. et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nat. Genet. 43, 1275–1280 (2011).
9. Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30-year study of Darwin’s finches. Science 296, 707–711 (2002).
10. Bolnick, D. I. et al. (Non)parallel evolution. Annu. Rev. Ecol. Evol. Syst. 49, 303–330 (2018).
11. Elmner, K. R. & Meyer, A. Adaptation in the age of ecological genomics: insights from parallelism and convergence. Trends Ecol. Evol. 26, 298–306 (2011).
12. Stern, D. L. The genetic causes of convergent evolution. Nat. Rev. Genet. 14, 751–764 (2013).
13. Greenway, R. et al. Convergent evolution of conserved mitochondrial pathways underlies repeated adaptation to extreme environments. Proc. Natl Acad. Sci. USA 117, 16424–16430 (2020).
14. Colosimo, P. F. et al. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307, 1928–1933 (2005).
15. Kingsley, E. P., Manceau, M., Wiley, C. D. & Hoekstra, H. E. Melanism in Peromyscus is caused by independent mutations in Agouti. PLoS ONE 4, e6435 (2009).
16. Manceau, M. et al. Convergence in pigmentation at multiple levels: mutations, genes and function. Phil. Trans. R. Soc. Lond. B 365, 2439–2450 (2010).
17. Linnen, C. R. et al. Adaptive evolution of multiple traits through multiple mutations at a single gene. Science 339, 1312–1316 (2013).
18. Barrett, R. D. H. & Schluter, D. Adaptation from standing genetic variation. Trends Ecol. Evol. 23, 38–44 (2008).
19. Papadopoulos, A. S. T. et al. Rapid parallel adaptation to anthropogenic heavy metal pollution. Mol. Biol. Evol. 38, 3724–3736 (2021).

Computer code is available at https://github.com/karwaan/Timema_climate_adaptation_genomics. All associated data files and scripts for specific analyses are archived on DRYAD (https://doi.org/10.5061/dryad.51c59zwbr). Correspondence for materials (data, scripts or samples) should be addressed to S.C. (samridhi.chaturvedi@gmail.com), Z.G. (zach.gompert@usu.edu) or P.N. (patrik.nosil@cefe.crns.fr).
20. Yeaman, S. Local adaptation by alleles of small effect. Am. Nat. 186, 574–589 (2015).
21. Yeaman, S. et al. Convergent local adaptation to climate in distantly related conifers. Science 353, 1431–1433 (2016).
22. Chaturvedi, S. et al. The predictability of genomic changes underlying a recent host shift in Melissa blue butterflies. Mol. Ecol. 27, 2561–2566 (2018).
23. Arendt, J. & Reznick, D. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23, 26–32 (2008).
24. Conte, G. L., Arnegard, M. E., Peichel, C. L. & Schluter, D. The probability of genetic parallelism and convergence in natural populations. Proc. Biol. Sci. 279, 5039–5047 (2012).
25. Bailey, S. F., Rodrigue, N. & Kassen, R. The effect of selection environment on the probability of parallel evolution. Mol. Biol. Evol. 32, 1436–1448 (2015).
26. Lenski, R. E. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J. 11, 2181–2194 (2017).
27. Roda, F. et al. Convergence and divergence during the adaptation to similar environments by an Australian groundsel. Evolution 67, 2515–2529 (2013).
28. Stuart, Y. E. et al. Contrasting effects of environment and genetics generate a continuum of parallel evolution. Nat. Ecol. Evol. 1, 158 (2017).
29. Morales, H. E. et al. Genomic architecture of parallel ecological divergence: beyond a single environmental contrast. Sci. Adv. 5, eaav9963 (2019).
30. Manousaki, T. et al. Parsing parallel evolution: ecological divergence and differential gene expression in the adaptive radiations of thick-tipped Midas cichlid fishes from Nicaragua. Mol. Ecol. 22, 650–669 (2013).
31. Rennison, D. J. Shared patterns of genome-wide differentiation are more strongly predicted by geography than by ecology. Am. Nat. 195, 192–200 (2020).
32. Schluter, D., Cliford, E. A., Nemethy, M. & McKinnon, J. S. Parallel evolution and inheritance of quantitative traits. Am. Nat. 163, 809–822 (2004).
33. Roesti, M., Gavrilets, S., Hendry, A. P., Salzburger, W. & Berner, D. The genomic signature of parallel adaptation from shared genetic variation. Mol. Evol. 23, 3944–3956 (2014).
34. Meyers, P. J. et al. Can the genomics of ecological speciation be predicted across the divergence continuum from host races to species? A case study in Rhaegoletis. Phil. Trans. R. Soc. B 375, 20190534 (2020).
35. Matos, M. et al. History, chance and selection during phenotypic and genomic experimental evolution: replaying the tape of life at different levels. Front. Genet. 6, 71 (2015).
36. Good, B. H. et al. The dynamics of molecular evolution over 60,000 generations. Nature 551, 45–50 (2017).
37. Storz, J. F. Causes of molecular convergence and parallelism in protein evolution. Nat. Rev. Genet. 17, 239–250 (2016).
38. Kohler, A. et al. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat. Genet. 47, 410–415 (2015).
39. Haldane, J. B. The Causes of Evolution (Princeton Univ. Press, 1990).
40. Gompel, N. & Carroll, S. B. Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. Nature 424, 931–935 (2003).
41. Orgogozo, V. Replaying the tape of life in the twenty-first century. Interface Focus 5, 20150057 (2015).
42. Blount, Z. D., Lenski, R. E. & Losos, J. B. Contingency and determinism in evolution: replaying life’s tape. Science 362, eaam5979 (2018).
43. Nosil, P. Divergent plant adaptation and reproductive isolation between ecotypes of Timema cristinae walking sticks. Am. Nat. 169, 151–162 (2007).
44. Comeault, A. A., Carvalho, C. F., Dennis, S., Soria-Carrasco, V. & Nosil, P. Color phenotypes are under similar genetic control in two distantly related species of Timema stick insect. Evolution 70, 1283–1296 (2016).
45. Comeault, A. A. et al. Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. Curr. Biol. 25, 1975–1981 (2015).
46. Lindtke, D. et al. Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. Mol. Ecol. 26, 6189–6205 (2017).
47. Villoutreix, R. et al. Large-scale mutation in the evolution of a gene complex for cryptic coloration. Science 369, 460–466 (2020).
48. Barghi, N., Hermisson, J. & Schlötterer, C. Polygenic adaptation: a unifying framework to understand positive selection. Nat. Rev. Genet. 21, 769–781 (2020).
49. Rockman, M. V. The QTN program and the alleles that matter for natural selection. Science 355, 959–962 (2017).
50. Law, J. H. & Crespi, B. J. The evolution of geographic parthenogenesis in Timema walking-sticks. Mol. Ecol. 11, 1471–1489 (2002).
51. Nosil, P. et al. Natural selection and the predictability of evolution in Timema stick insects. Science https://science.sciencemag.org/content/359/6377/765 (2018).
52. Riesch, R. et al. Transitions between phases of genomic differentiation during stick-insect speciation. Nat. Ecol. Evol. 1, 82 (2017).
53. Harvey, M. G. et al. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. Syst. Biol. 65, 910–924 (2016).
54. Schluter, D. & Conte, G. L. Genetics and ecological speciation. Proc. Natl Acad. Sci. USA 106, 9955–9962 (2009).
55. Liu, S., Ferchaud, A.-L., Grønknjaer, P., Nygaard, R. & Hansen, M. Genomic parallelism and lack thereof in contrasting systems of three-spined sticklebacks. Mol. Ecol. 27, 4725–4743 (2018).
56. Gompert, Z. et al. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common genomic differentiation during stick-insect speciation. Nat. Ecol. Evol. 1, 82 (2017).
57. Harvey, M. G. et al. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. Syst. Biol. 65, 910–924 (2016).
58. Gompert, Z. et al. Experimental evidence for ecological selection on genome variation in the wild. Ecol. Lett. 17, 369–379 (2014).
59. Sprenger, P. P., Lars, H. B., Abou, B., Federle, W. & Menzel, F. Coping with the climate: cuticular hydrocarbon acclimation of ants under constant and fluctuating conditions. J. Exp. Biol. 221, jeb171488 (2018).
60. Botella-Cruz, M., Velasco, J., Millán, A., Hetz, S. & Pallarés, S. Cuticle hydrocarbons show plastic variation under desiccation in saline aquatic beetles. Insects 12, 285 (2021).
61. Walden, N., Lucek, K. & Willi, Y. Lineage-specific adaptation to climate involves flowering time in North American Arabidopsis lyrata. Mol. Ecol. 29, 1436–1451 (2020).
65. Rose, N. H., Bay, R. A., Morikawa, M. K. & Palumbi, S. R. Polygenic evolution drives species divergence and climate adaptation in corals. *Evolution* **72**, 82–94 (2018).

66. Blanco-Pastor, J. L. et al. Annual and perennial Medicago show signatures of parallel adaptation to climate and soil in highly conserved genes. *Mol. Ecol.* https://doi.org/10.1111/mec.16061 (2021).

67. Wang, L. et al. Molecular parallelism underlies convergent highland adaptation of maize landraces. *Mol. Biol. Evol.* **38**, 3567–3580 (2021).

68. Bohutińska, M. et al. Genomic basis of parallel adaptation varies with divergence in *Arabidopsis* and its relatives. *Proc. Natl Acad. Sci. USA* **118**, e2022713118 (2021).

69. Lobréaux, S. & Melodelima, C. Detection of genomic loci associated with environmental variables using generalized linear mixed models. *Genomics* **105**, 69–75 (2015).

70. Frachon, L. et al. A genomic map of climate adaptation in *Arabidopsis thaliana* at a micro-geographic scale. *Front. Plant Sci.* **9**, 967 (2018).

71. Contreras-Moreira, B. et al. Genetic association with high-resolution climate data reveals selection footprints in the genomes of barley landraces across the Iberian Peninsula. *Mol. Ecol.** 28**, 1994–2012 (2019).

72. Heliconius Genome Consortium. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).

73. Henning, F. & Meyer, A. The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. *Annu. Rev. Genomics Hum. Genet.* **15**, 417–441 (2014).

74. Marburger, S. et al. Interspecific introgression mediates adaptation to whole genome duplication. *Nat. Commun.* **10**, 5218 (2019).

75. Giska, I. et al. Introgression drives repeated evolution of winter coat color polymorphism in hares. *Proc. Natl Acad. Sci. USA* **116**, 24150–24156 (2019).

76. Menon, M. et al. Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. *Commun. Biol.* **4**, 160 (2021).

77. Bay, R. A., Taylor, E. B. & Schluter, D. Parallel introgression and selection on introduced alleles in a native species. *Mol. Ecol.* **28**, 2802–2813 (2019).

78. Zhang, X., Ryaner, J. G., Blaxter, M. & Bailey, N. W. Rapid parallelism underlies convergent adaptation despite gene flow in silent crickets. *Nat. Commun.* **10**, 5218 (2019).

79. Marques, D. A., Jones, F. C., Di Palma, F., Kingsley, D. M. & Reimchen, E. Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nat. Ecol. Evol.* **2**, 1128–1138 (2018).

80. Li, H. et al. and 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

81. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* https://doi.org/10.1093/gigascience/gia008 (2021).

82. Gautier, M. Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics* **201**, 1555–1579 (2015).

83. Günther, T. & Coop, G. Robust identification of local adaptation from allele frequencies. *Genetics* **195**, 205–220 (2013).

84. Clarke, R. T., Rothery, P. & Raybould, A. F. Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *J. Agric. Biol. Environ. Stat.* **7**, 361–372 (2002).

85. Soria-Carrasco, V. et al. Stick insect genomes reveal natural selection’s role in parallel speciation. *Science* **344**, 738–742 (2014).

86. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria (2022).
Extended Data Fig. 1 | PCA of climate variation. Ordination of climate variation (22 variables, see Supplementary Table 2 for code descriptions) via principal component analysis (PCA). Points denote the study populations, colour-coded by species.
Extended Data Fig. 2 | Manhattan plot for PC1. Manhattan plots showing the strength of evidence for association (measured here using the Bayes factor from the software BayPass) between a SNP window and climate for PC1. Results are shown along the 13 linkage groups. In each panel title, the two values in parentheses are the number of SNP windows in the top 10% quantile (‘windows’), followed by the number of linkage groups with at least 1 SNP window in the top 10% quantile (‘LG’).
Extended Data Fig. 3 | Manhattan plot for PC2. Manhattan plots showing the strength of evidence for association (measured here using the Bayes factor from the software BayPass) between a SNP window and climate for PC2. Results are shown along the 13 linkage groups. In each panel title, the two values in parentheses are the number of SNP windows in the top 10% quantile (‘windows’), followed by the number of linkage groups with at least 1 SNP window in the top 10% quantile (‘LG’).

(A) T. bartmani (155 windows, 13 LG)
(B) T. podura (215 windows, 13 LG)
(C) T. chumash (157 windows, 13 LG)
(D) T. cristinae (807 windows, 13 LG)
(E) T. knulli (386 windows, 13 LG)
(F) T. poppensis (315 windows, 13 LG)
(G) T. landelsensis (333 windows, 12 LG)
(H) T. californicum (342 windows, 13 LG)
Extended Data Fig. 4 | Parallelism tests for PC1. Tests for parallel climate-associated SNP windows between species of *Timema* stick insects (all plots are for the top 10% empirical quantile) for PC1. Plot shows x-fold enrichments for the number of overlapping climate-associated SNP windows for PC1 for comparisons between multiple species, that is, beyond pairs of species (for example, 2 or more species, 3 or more species, 4 or more species). Gray dots denote x-fold values expected under 1000 randomizations for a null distribution. Black diamond denotes median of the x-fold values expected under 1000 randomizations for a null distribution. Red dot and N value above each group indicates the observed number of overlapping climate-associated SNP windows for each comparison. *P*-value above each group denotes whether the overlap is greater than expected by chance from a one-sided randomization test. * Indicates x-fold enrichments with P-value ≤ 0.05.
Extended Data Fig. 5 | Parallelism tests for PC2. Tests for parallel climate-associated SNP windows between species of *Timema* stick insects (all plots are for the top 10% empirical quantile) for PC2. Plot shows x-fold enrichments for the number of overlapping climate-associated SNP windows for PC2 for comparisons between multiple species, that is, beyond pairs of species (for example, 2 or more species, 3 or more species, 4 or more species). Gray dots denote x-fold values expected under 1000 randomizations for a null distribution. Black diamond denotes median of the x-fold values expected under 1000 randomizations for a null distribution. Red dot and N value above each group indicates the observed number of overlapping climate-associated SNP windows for each comparison. *P*-value above each group denotes whether the overlap is greater than expected by chance from a one-sided randomization test. * Indicates x-fold enrichments with *P*-value ≤ 0.05.
Extended Data Fig. 6 | Tests of the ‘shared ecology’ versus ‘shared genetics’ hypothesis for PC1. Test results of the ‘shared ecology’ versus ‘shared genetics’ hypotheses for PC1. (a) Scatterplot shows the relationship between X-fold enrichment (measure for parallelism) and climatic distance (measured as the distance in PC1 scores) based on a single-factor linear model. (b) Scatterplot shows the relationship between X-fold enrichment (measure for parallelism) and genetic distance (measured as pairwise phylogenetic distance) based on a single-factor linear model. (c) Scatterplot shows the relationship between climatic distance (measured as the distance in PC1 scores and is the distance in climate variables) and genetic distance (calculated as pairwise phylogenetic distance) based on a single-factor linear model. (d) Plot shows parameter estimates with standardized coefficients for the full model for PC1. This test was implemented for all eight species and 56 species pairs. Error bars indicate 95% equal-tail probability intervals (ETPIs). A negative or positive estimate that deviates from zero indicates the effect on parallelism.
Extended Data Fig. 7 | Tests of the ‘shared ecology’ versus ‘shared genetics’ hypothesis for PC2. Test results of the ‘shared ecology’ versus ‘shared genetics’ hypotheses for PC2. (a) Scatterplot shows the relationship between X-fold enrichment (measure for parallelism) and climatic distance (measured as the distance in PC2 scores) based on a single-factor linear model. (b) Scatterplot shows the relationship between X-fold enrichment (measure for parallelism) and genetic distance (measured as pairwise phylogenetic distance) based on a single-factor linear model. (c) Scatterplot shows the relationship between climatic distance (measured as the distance in PC2 scores and is the distance in climate variables) and genetic distance (calculated as pairwise phylogenetic distance) based on a single-factor linear model. (d) Plot shows parameter estimates with standardized coefficients for the full model only for PC2. This test was implemented for all eight species and 56 species pairs. Error bars indicate 95% equal-tail probability intervals (ETPIs). A negative or positive estimate that deviates from zero indicates the effect on parallelism.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  
  *Give P values as exact values whenever suitable.*

- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#) is as follows:

| Data collection | No software was used. |
| Data analysis | We used custom scripts in Perl, Python (version 3), and UNIX/BASH to modify files and run remote jobs on University of Utah High Performance Computing Cluster. We used specific programs such as BAYPASS, TREEMIX, ENTROPY, and R (version 3.4) code for data analyses, statistical testing and visualization. The associated scripts for this study are archived at [https://github.com/karwaan/Timmel_climate_adaptation_genomics](https://github.com/karwaan/Timmel_climate_adaptation_genomics). All associated data files are archived on DRYAD ([https://doi.org/10.5061/dryad.51c59zwbr](https://doi.org/10.5061/dryad.51c59zwbr)). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.
Data

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

DNA sequence data, genome, experimental data, and CHC data used in this study are associated with the previously published studies (Gompert, Comeault, et al. 2014; Riesch et al. 2017). The associated DNA sequence data have been archived on NCBI’s SRA (Accession: PRJNA356405 ID: 356405).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender
This study did not involve human participants.

Population characteristics
This study did not involve human participants.

Recruitment
This study did not involve human participants.

Ethics oversight
This study did not involve human participants.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences  ☐ Behavioural & social sciences  ☑ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description
Evolution can repeat itself, resulting in parallel adaptations in independent lineages occupying similar environments. Moreover, parallel evolution sometimes, but not always, uses the same genes. Two main hypotheses have been put forth to explain the probability and extent of parallel evolution. First, parallel evolution is more likely when shared ecologies result in similar patterns of natural selection in different taxa. Second, parallelism is more likely when genomes are similar, because of shared standing variation and similar mutational effects in closely related genomes. Here we combine ecological, genomic, experimental, and phenotypic data with Bayesian modeling and randomization tests to quantify the degree of parallelism and its relationship with ecology and genetics. Our results show that the extent to which genomic regions associated with climate are parallel among species of Timema stick insects is shaped collectively by shared ecology and genomic background. Specifically, the extent of genomic parallelism decays with divergence in climatic conditions (i.e., habitat or ecological similarity) and genomic similarity. Moreover, we find that climate-associated loci are likely subject to selection in a field experiment, overlap with genetic regions associated with cuticular hydrocarbon traits, and are not strongly shaped by introgression between species. Our findings shed light on when evolution is most expected to repeat itself.

Research sample
Our study system is wingless, univoltine, herbivorous stick insects in the genus Timema, many species of which are endemic to California, USA. These insects are best-studied for their cryptic colours and colour-patterns, which are controlled by the same genetic region (termed Mel-Stripe) in all species studied to date. Timema colouration thus provides a striking example of highly parallel evolution at the level of a single, largely non-recombining gene region that could be considered akin to a major effect locus. However, adaptation often involves many genes, including those with alleles of minor effect, arrayed throughout the genome, where the probability of parallel genetic evolution is less clear. In this context, we study a novel ecological dimension in Timema, namely climate, motivated by the fact that adaptation to varying climatic (abiotic) conditions of the environment can be polygenic, and the genus Timema inhabits variable habitats in California. For example, the occupied habitats of Timema range from sea-level to mountainous regions, and from arid semi-deserts near the Mexican border to wet evergreen forests in northern California [50]. Moreover, there is climatic variation both within and among species, with several species being distributed along elevational gradients (ranging from 10 meters to ~2800 meters). This creates an opportunity to test the role of climatic variables, such as precipitation and temperature, in driving parallel evolution in Timema, which are known to be important determinants of selection in many organisms.

Sampling strategy
For this study, we used 8 species which included ~54 populations across species (N=1412). This sample size was ideal for testing...
### Sampling strategy
Hypothesis for parallel adaptation to climate at the genomic level. Comparisons across species are rare for tests of parallelism and this dataset provided a novel opportunity to identify the causes of decay of parallelism across species.

### Data collection
The data included in this study was generated previously as specified in the manuscript.

### Timing and spatial scale
The data used in this study was published in a study in 2017. Samples were collected from California, USA.

### Data exclusions
No data were excluded from this study.

### Reproducibility
We did not have experimental tests in this study, but all code used for analysis has been archived on Github.

### Randomization
One way in which correlations were controlled in climate variables was by using the first three principal components as environmental variables (representing climate variables) to test the genomic association to climate using BayPass software. We also performed one-way randomizations for all parallelism tests and conducted additional permutation tests to check the robustness of our parallelism and shared-ecology and shared-genetics hypothesis tests.

### Blinding
Blinding is not relevant to this study as we are not dealing with human subjects or making any inference about behavior. We are just using the samples for genomic DNA to make inferences about evolution and population genetics.

### Did the study involve field work?
- Yes [ ]
- No [x]

---

### Reporting for specific materials, systems and methods
We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

#### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| [ ] | Antibodies            |
| [x] | Eukaryotic cell lines |
| [x] | Palaeontology and archaeology |
| [x] | Animals and other organisms |
| [x] | Clinical data         |
| [x] | Dual use research of concern |

---

#### Methods

| n/a | Involved in the study |
|-----|-----------------------|
| [x] | ChIP-seq              |
| [x] | Flow cytometry        |
| [x] | MRI-based neuroimaging |