




Rapid parallel adaptation despite gene flow in silent crickets

Xiao Zhang ^{1,3}, Jack G. Rayner^{1,3}, Mark Blaxter² & Nathan W. Bailey¹

Gene flow is predicted to impede parallel adaptation via de novo mutation, because it can introduce pre-existing adaptive alleles from population to population. We test this using Hawaiian crickets (*Teleogryllus oceanicus*) in which ‘flatwing’ males that lack sound-producing wing structures recently arose and spread under selection from an acoustically-orienting parasitoid. Morphometric and genetic comparisons identify distinct flatwing phenotypes in populations on three islands, localized to different loci. Nevertheless, we detect strong, recent and ongoing gene flow among the populations. Using genome scans and gene expression analysis we find that parallel evolution of flatwing on different islands is associated with shared genomic hotspots of adaptation that contain the gene *doublesex*, but the form of selection differs among islands and corresponds to known flatwing demographics in the wild. We thus show how parallel adaptation can occur on contemporary timescales despite gene flow, indicating that it could be less constrained than previously appreciated.

¹School of Biology, University of St Andrews, St Andrews, Fife KY16 9TH, UK. ²Tree of Life, Wellcome Sanger Institute, Cambridge CB10 1SA, UK. ³These authors contributed equally: Xiao Zhang, Jack G. Rayner. ✉email: xz42@st-andrews.ac.uk; nwb3@st-andrews.ac.uk

Adaptive evolution can be surprisingly repeatable. Convergent phenotypic evolution occurs when similar selection pressures drive the spread of functionally identical traits at the organismal level but in different populations^{1,2}. One way this can occur is through adaptive introgression, by repeatedly introducing the same pre-existing adaptive allele from one population into others, whereupon they are selected. In this scenario, gene flow is predicted to play an important role in moving the same allelic variants, which share a single mutational origin, among populations³. In contrast, parallel evolution occurs when traits that are functionally identical at the organismal level nonetheless have independent mutational origins at the genetic level^{3,4}. In the latter case, uniquely arisen adaptive genetic variants are independently selected in different populations, and theory strongly predicts that high levels of gene flow will reduce the likelihood of parallel evolution^{5–10}. For that reason, the most compelling examples of parallel evolution are those that have been documented in highly isolated populations, for example, the independent selection of different genetic variations that enable rabbits (*Oryctolagus cuniculus*) on different continents to survive myxoma virus¹¹, and the repeated adaptive deletion of the *Pitx1* gene, which causes stickleback (*Gasterosteus aculeatus*) pelvic reduction in different spatially isolated lake systems^{12,13}. Exceptions to this general pattern are only predicted when selection is extreme or mutation rates are particularly high^{3,14}, but these scenarios are difficult to empirically test in natural systems. Most adaptive novelties arose and spread in the distant past, necessitating the use of sophisticated techniques to distinguish processes such as incomplete lineage sorting and gene flow^{15–17}. Dissecting the mutational origins of novel adaptive traits undergoing convergent phenotypic evolution, while simultaneously inferring gene flow, is complicated by these sorts of demographic developments occurring during and after episodes of adaptive evolution^{18,19}. Thus a question at the centre of arguments about evolvability and adaptive potential remains largely unsolved: Can parallel evolution underpinned by independent mutations occur despite high levels of gene flow, and if so, how?

We addressed this question using a unique system of rapidly evolving field crickets (*Teleogryllus oceanicus*) on the Hawaiian archipelago. Typically, male crickets attract females for mating by rubbing their wings together to produce a conspicuous, long-range acoustic signal. However, in approximately 2003, a variant segregating in the manner of a single locus on the X chromosome—*flatwing*—arose and spread in a population on the island of Kauai²⁰. This adaptive allele erases sound-producing structures on males' wings during development, causing them to resemble the undifferentiated wings of females (Fig. 1a)²¹. Silence protects these 'flatwing' males from fatal attack by an acoustically orienting endoparasitoid fly, *Ormia ochracea*²⁰. The combination of a novel adaptive variant and acoustically orienting parasitoids has driven the recent, rapid evolution of male silence, increasing the abundance of flatwing males from 0% to over 95% in ca. 20–30 generations²⁰. Protective flatwing male morphs were found 2 years later on the neighbouring islands of Oahu, and 5 years later on the Big Island (hereafter Hilo), and evidence from the field shows that the flatwing morph is now fixed at 100% in the Kauai population (Fig. 1a)^{21–25}. The flatwing male morph has never been found outside of Hawaii. Previous research using lab-reared lines provided evidence that both Kauai and Oahu *flatwing* loci (f_w^K and f_w^O , respectively) show the same mode of inheritance on the X chromosome yet are predominately associated with different restriction site associated DNA (RAD) markers, suggesting they are independently generated alleles^{21,23}.

Here, we take advantage of the repeated evolutionary origin and spread of flatwing crickets in multiple Hawaiian island populations to test the expected trade-off between gene flow and rapid parallel adaptation via independent mutational events, which is predicted by standard evolutionary models^{5–10}. We first compare morphometric differences among flatwing male crickets on three islands of the Hawaii archipelago, then resequence 70 cricket genomes at high coverage: normal-wing and flatwing males from all Hawaiian populations, plus a non-parasitized, normal-wing Australian population and a sister species outgroup *T. commodus*. Association and selection analyses localise *flatwing* to different loci in each of the three Hawaiian populations but also highlight a putative genomic hotspot of adaptation, which has been subject to distinct processes that have left different signatures of selection in each population. By reconstructing historical population demographic patterns both autosome-wide and for X-linked loci, we find evidence of extensive, recent, and ongoing genome-wide gene flow among all three populations, but no evidence of *flatwing* introgression. We reconstruct the historical demographic population genetics of these populations as well as signatures of selection on adaptation hotspots in each. We interrogate a candidate region identified by the analyses above, containing the gene *doublesex* (*dsx*), which has been widely implicated in sex-specific adaptive traits in other insects, by quantifying gene expression in developing wingbuds of different male morphs on two islands where flatwing crickets exist in appreciable abundance (Kauai and Oahu). Our results reveal that parallel evolution can occur despite gene flow, indicate how this process occurs, and implicate the *doublesex* pathway in the repeated origins of an adaptive, male-feminising trait in wild insects.

Results

Morphological analysis indicates unique wing adaptations.

Structures formed by specialised veins on the forewings of male crickets normally produce sound when males move their wings back and forth across one another during singing. However, these structures are either absent or severely reduced in flatwing males. Thus feminised, flatwing males are prevented from making sound even though they still move their wings in the motor pattern characteristic of singing, but without producing audible sound (Fig. 1a)^{20,23,26}. Flatwing venation differs between Kauai and Oahu males and appears to be under independent genetic control²³, so to test whether the wing venation of flatwing males is morphologically distinct among all three populations we studied here, we collected right forewings of 114 wild flatwing male crickets over multiple years from populations on the islands of Kauai, Oahu and Hilo (Supplementary Table 1). Landmark-based geometric morphometric analysis^{23,27} revealed that wing venation of flatwing males differed markedly among the three populations (multivariate analyses of variance: all $P < 0.001$, Supplementary Table 2). We then used canonical variate analysis to visualise the main sources of variation among flatwing males from the three islands. All three populations were distinguishable with Hilo flatwings more similar to Kauai flatwings than to Oahu flatwings (Fig. 1b, c), though this must be interpreted cautiously owing to the small number of Hilo flatwing crickets in the wild. However, this pattern is visible to the naked eye (Fig. 1a). Hilo and Kauai flatwing individuals tend to have completely lost their scraper—the thickened ridge of tissue on the wing margin, which engages the other wing during movement—but a vestigial scraper has been retained by most Oahu flatwings (Fig. 1a, c). Despite this variation, all flatwing males are functionally silent²⁶. Although all flatwing males had reduced or absent resonators (enlarged wing

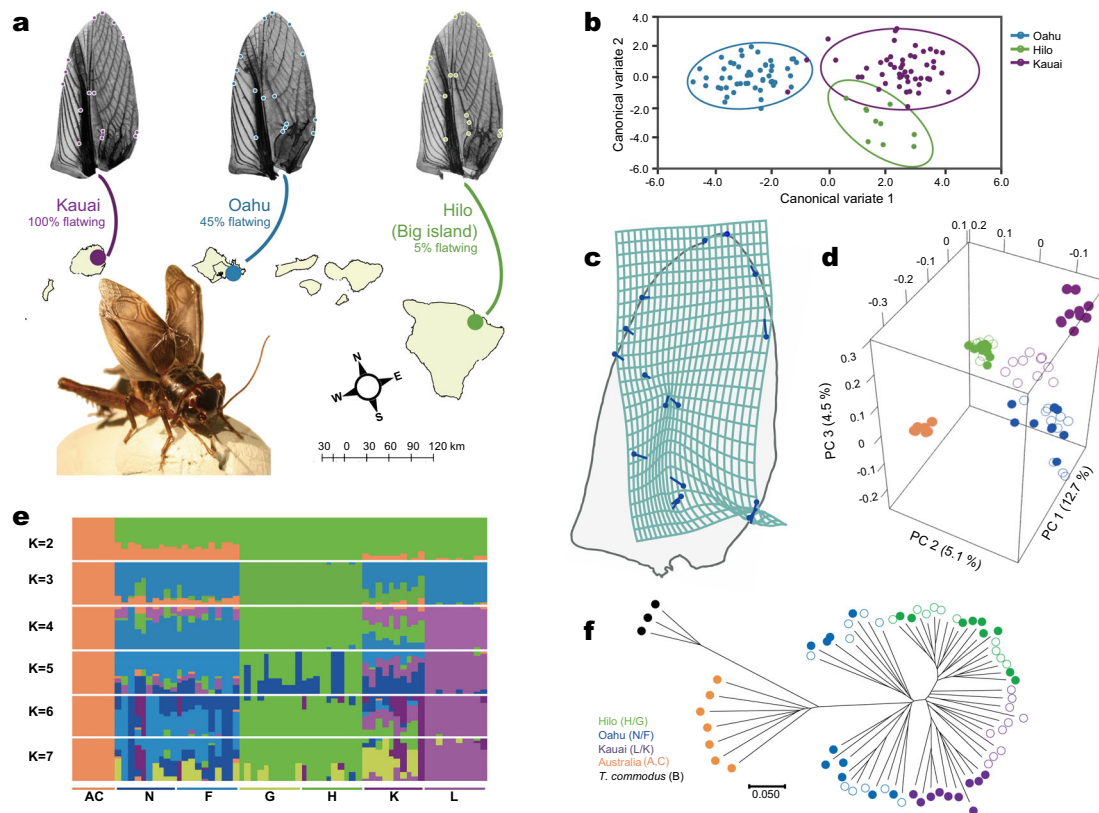


Fig. 1 Geographic, morphological and genetic variation in silent Hawaiian crickets. **a** Representative right forewings of flatwing *T. oceanicus* males from populations on the Hawaiian islands of Kauai, Oahu and the Big Island (Hilo). Map drawn using QGIS v2.18.24 (<http://qgis.org>). Dots on wing veins indicate geometric morphometric landmarks, and colours reflect the scheme used in this manuscript if not stated otherwise (purple = Kauai, blue = Oahu, green = Hilo). Bottom left photo: A normal-wing male for comparison (Photo credit: N.W. Bailey). **b** Flatwing wing venation differs among the three populations. Canonical variates analysis (CVA) showing multivariate differences in wing veins of flatwing males from Kauai, Oahu and Hilo. Source data are provided as a Source Data file. **c** Deformation grid illustrating the major differences in forewing venation between flatwing males from the three populations. The outline indicates a generic male forewing, and arrows show vectors representing landmark variation on CV1 (scaling factor of 10). **d** There is clear population genetic differentiation among the three islands. Genome-wide differentiation among *T. oceanicus* populations shown in a three-dimensional PCA scatter diagram based on autosomal SNPs. Australian *T. oceanicus* outgroup samples are indicated in orange. Solid circles = normal-wing males, open circles = flatwing males. **e** However, population structure analysis indicated an optimal group size of $K = 5$, likely reflecting our inclusion of lab-raised normal-wing Kauai males and lab-raised flatwing Hilo males. Different colours indicate genetically distinct groups. **f** Neighbour-joining phylogenetic tree based on autosomal SNPs. Symbols and colour coding for *T. oceanicus* samples are as above, and black circles = *T. commodus* outgroup. The scale bar represents the level of genetic similarity.

cells that vibrate during singing), the degree of this reduction varied among populations (Fig. 1a).

Background genetic structure of cricket populations under convergent fly selection. We resequenced whole genomes of 70 male crickets to an average depth of 25 \times . Twenty males (10 of each morph; flatwing and normal-wing) were derived from parasitized populations on each of three Hawaiian islands (Kauai, Oahu and Hilo) (Fig. 1a). Seven normal-wing male *T. oceanicus* from Australia plus three males of the sister species *T. commodus* (which also have normal-wing morphology) were also sequenced. Supplementary Table 3 provides further sampling details. Australian populations are not known to be subject to selection pressure from the parasitoid fly (Fig. 1a), making them a useful comparison in selection analyses, and the sibling species *T. commodus* was used as an outgroup (Fig. 1 and Supplementary Table 3, Supplementary Data 1). All crickets were wild-caught with the following exceptions: Because normal-wing crickets now appear to be extinct in the Kauai population^{24,28}, and flatwing crickets only exist in extremely low abundance in the Hilo population²³, we resequenced lab-reared individuals established from wild-caught ancestors for these groups.

More than 96% of the re-sequencing reads successfully mapped to the *T. oceanicus* reference genome²¹, confirming the quality of this re-sequencing dataset. We used a customised pipeline to account for potential biases caused by outgroups or sequencing errors to obtain a total of 107 million high-quality single nucleotide polymorphisms (SNPs) (Supplementary Data 2). To examine background genome-wide population structure excluding any sex-linked or flatwing-associated effects, a neighbour-joining tree based on pairwise genetic distances was constructed using autosomal SNP data (Fig. 1f). In Hawaii, individuals grouped into three population-specific clades, and Oahu was most closely related to the Australian outgroup (Fig. 1f and Supplementary Fig. 2). Flatwing and normal-wing individuals were not split into two subclades except for Kauai individuals (purple circles of Fig. 1d, f and Supplementary Figs. 2, 3), which likely reflects founder effects among these lab-reared samples. Principal component analysis (PCA) and genetic structure analysis are consistent with the NJ tree (Fig. 1d, e and Supplementary Fig. 1). Morphometric and autosomal genetic data yielded notably similar patterns of relatedness among island populations: Kauai and Hilo were more similar to each other than either was to Oahu despite their greater geographic distance.

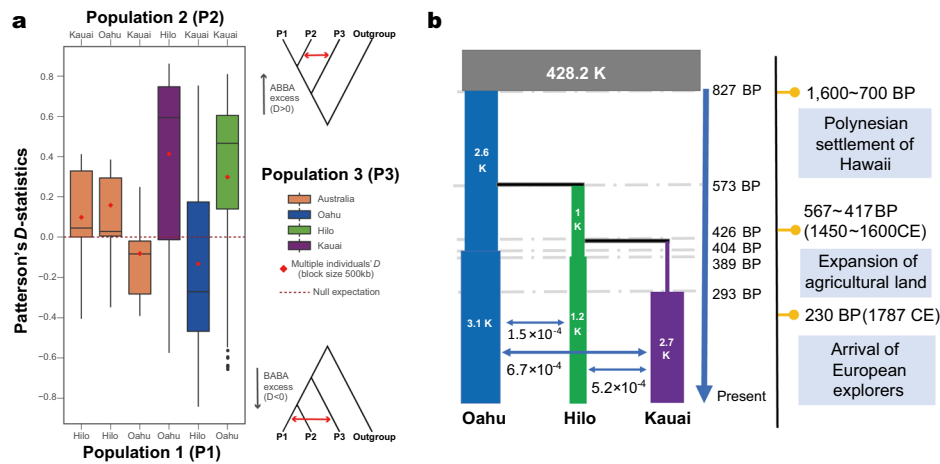


Fig. 2 Extensive gene flow among three Hawaiian island populations of *T. oceanicus*. **a** ABBA–BABA tests illustrate significant gene flow among all three Hawaiian populations on the X chromosome. Sites used in the test are derived in *T. oceanicus* compared to the outgroup (*T. commodus*) but differ among P1 and P2 *T. oceanicus* individuals. Tests with Australian *T. oceanicus* populations are included as control comparisons (left three bars) and the plot shows all possible topological permutations of the three Hawaiian island populations relevant for the ABBA–BABA test (right three bars). The null expectation of no gene flow is indicated by $D = 0$. Positive D indicates an excess of shared derived alleles—evidence for greater gene flow—between populations 2 and 3 in the comparison (ABBA excess), whereas negative D indicates evidence for greater gene flow between populations 1 and 3 in the comparison (BABA excess). The significance of gene flow among island populations was also supported by all comparisons between all Hawaiian scenarios and Australian scenarios (two-sided t -tests: all $P < 0.001$; $n = 700$ independent D -values for each Australian control; $n = 1000$ each for Kauai and Hilo; $n = 2000$ for Oahu. Statistical details are provided in Supplementary Table 4). Box plots show upper quartile, median, and lower quartile values, 1.5 \times interquartile ranges, and outliers. Multiple individuals' D tests using a smaller block size are represented by red diamonds. Source data are provided as a Source Data file. **b** The most likely demographic history of Hawaiian *T. oceanicus* populations reconstructed using Fastimcoal2 indicates recent, extensive gene flow among all islands. The ancestral population is shown in grey. Bar width for each population indicates estimated effective population sizes (numbers indicated in each block) and black horizontal lines represent population divergence. Blue horizontal arrows and the numbers underneath them indicate per-generation migration rates, i.e. the estimated rate of symmetrically exchanged migrants between populations every generation; these episodes of gene flow were inferred to have started approximately within the past 300–400 years and are ongoing. This reconstruction provided evidence that all three populations experienced ancient bottlenecks, plus recent population expansions accompanied by gene flow. An approximate timeframe indicating initial human colonisation of the Hawaiian archipelago, agricultural activity, and European contact is provided on the right of the panel. The population demographic reconstruction for *T. oceanicus* indicates that the crickets are likely to have accompanied the first human arrivals to Hawaii.

These morphometric and genomic patterns implied two possible mechanisms for the repeated, rapid, adaptive evolution of flatwing crickets across the Hawaiian archipelago. Under the first scenario, this similarity could be due to adaptive introgression of one causative flatwing variant from island to island. Alternatively, the phenotypic expression of different independent flatwing variants could be influenced by different genomic background effects, with the similarity of flatwings in Kauai and Hilo caused by their more similar genetic backgrounds.

Recent, extensive and ongoing gene flow among Hawaiian cricket populations. Migration could have facilitated the adaptive introgression of flatwing variants from population to population, whereupon they spread under selection from the fly²⁹. In the wild, populations on different islands are assumed to be isolated from one other because of the distinct geographical barrier of the sea; *T. oceanicus* from Hawaii are not known to be capable of sustained flight such as would be required for long-distance dispersal and would have great difficulty crossing many miles of open ocean. However, in Hawaii, human activities including the transport of agricultural goods, tourists and residents make these barriers surmountable. To evaluate gene flow, we assessed the presence and extent of admixture among the three island populations and reconstructed their demographic history. Only wild individuals were used in these analyses.

We used two approaches to evaluate gene flow. First, we implemented Patterson's D statistics (ABBA–BABA test)³⁰ using X chromosome data because we were specifically interested in understanding gene flow that could affect the movement of

flatwing alleles among populations. This analysis revealed considerable gene flow among the three Hawaiian island populations (right three bars of Fig. 2a), which was considerably and significantly higher than that involving Australian populations (orange bars of Fig. 2a and Supplementary Table 4, Supplementary Data 3). All comparisons yielded D -values which deviated significantly from the null hypothesis ($D = 0$), which is consistent with common observations that modest but significant gene flow can be detected among geographically distinct distant populations within the same species^{31–33}. The key observation is that all Hawaii island scenarios showed significantly more extreme D -values compared to gene flow involving Australian populations (two-sided t -tests: all $P < 0.001$; Supplementary Table 4). In addition, multiple individuals' D ³⁴ test using a much smaller block size showed similar results (Fig. 2a and Supplementary Table 5). Although it would be problematic to directly compare D -values of the X chromosome and autosomes due to differences in effective population size, we nevertheless note that the X (LG1) showed the strongest signals of gene flow, and the large majority of autosomal linkage groups (LGs) showed qualitative patterns consistent with the X (Supplementary Fig. 4).

Second, to test more detailed scenarios of gene flow and its timing we used the joint site frequency spectrum (SFS) of autosomal SNPs implemented in Fastsimcoal2³⁵. The latter analysis is more sensitive but relies on several assumptions that preclude using X-linked markers, so we evaluated 62 models using the SFS of autosomal SNPs to reconstruct the recent demographic history of Hawaiian crickets, incorporating divergence time, changes in effective population size, and gene flow (Supplementary Fig. 5). The models included seven divergence patterns possible for three

extant populations plus one assumed ancestral population, all possible scenarios of island permutation for each divergence pattern, and two gene flow scenarios (i.e. whether or not recent gene flow occurred among the three islands) for each permutation.

The results indicate that gene flow among all three islands started recently, and occurred continuously until the present, rather than transiently. Almost all scenarios assuming recent gene flow were more probable than those without gene flow (Supplementary Data 4). The best-fit model suggested that crickets appeared in Oahu first, more than 800 years ago. Then, the Hilo population split from the Oahu population about 570 years ago. Finally, about 400 years ago, the Kauai population diverged from the Hilo population. All three populations experienced ancient bottlenecks followed by recent population expansions, with symmetrical gene flow starting within the last 300–400 years and continuing today (Fig. 2b and Supplementary Data 4). These dates coincide with the history of the Polynesian settlement of Hawaii, followed by expansion of the human population, and the arrival of European explorers, respectively (Fig. 2b). We also used Pairwise Sequentially Markovian Coalescent (PSMC) models to infer more ancient demographic history (Supplementary Fig. 6)³⁶. The ancestral effective population size of Hawaiian crickets showed a peak about 20,000 years ago followed by a decline involving at least a twofold decrease in the effective population size coinciding with the last glacial maximum³⁷, likely reflecting their putative ‘out-of-Australia’ origins and island-hopping across the Pacific^{38,39}.

Genomic hotspots implicated in rapid parallel adaptation of silent crickets. To dissect the genetic basis of convergent phenotypic evolution in silent Hawaiian crickets and test whether it has been driven by adaptive introgression or parallel evolution, we performed three independent genome-wide association studies (GWAS), one for each island. Because of the extreme scarcity of Kauai normal-wing males and Hilo flatwing males, GWAS involving these two morphs used lab stocks established from the offspring of wild-caught individuals (see ‘Methods’ section). Our previous quantitative trait locus (QTL) mapping of flatwing in Kauai yielded a broad peak occupying approximately one-third of the X chromosome (redrawn from ref. 21 using pseudomolecule coordinates (bp); top panel of Fig. 3a). Here, we substantially narrowed this candidate region to three small SNP datasets varying from 5 to 615 flatwing-associated SNPs, spanning 3–17 scaffolds of the X chromosome (Fig. 3a and Supplementary Table 6, Supplementary Data 5). Among all flatwing-associated SNPs, 83 overlapped with 32 protein-coding genes, most of which have a role in *D. melanogaster* wing development or other relevant developmental processes (Supplementary Data 6).

It was important that inferences from separate GWAS analyses were comparable across Hawaiian populations containing flatwing males; that is, flatwing-associated SNPs needed to be detected with a uniform criterion for false discovery rate (FDR) in all three analyses. We therefore subsetted the Hilo and Kauai flatwing-associated SNPs according to the most conservative criterion based on the Oahu GWAS (Fig. 3a, red dots), for which all samples from both morphs had been obtained from the wild. Both the initial, relaxed set of flatwing-associated markers and the second, conservative set of flatwing-associated markers yielded the same pattern: no flatwing-associated SNP was shared by any of the three populations (Supplementary Table 6, Supplementary Data 5). This pattern was consistent with different genetic causes of flatwing in Kauai, Oahu and Hilo, however, we found that different flatwing-associated SNPs in the populations were sometimes located on the same scaffold. Two of these candidate

flatwing-associated scaffolds were shared by all three islands (Fig. 3b and Supplementary Fig. 7, Supplementary Table 7). Considering that one scaffold is a physically continuous genetic region, one possible explanation is that these regions contain potential ‘genomic hotspots’ of adaptation in which recurrent mutations lead to the evolution of flatwing morphology. Alternatively, this pattern could be caused by adaptive introgression of the same causative flatwing mutation(s) being masked in our analysis by potential biases of performing three separate GWASs, for example, due to population genetic structure. To further evaluate the idea that parallel evolution occurred in genetic hotspots, we established maximum-likelihood phylogenetic trees using SNPs extracted from (1) flatwing-associated scaffolds shared by different populations, (2) GWAS candidate SNP sets from all populations combined, (3) each of the three population-specific GWAS candidate SNP sets separately, and (4) flanking regions of these flatwing-associated SNPs. Given our precise knowledge of how recently flatwing appeared in each population^{20,22,23}, if a single flatwing allele had arisen and introgressed among populations, then trees built using restrictive sets of markers strongly linked to the flatwing phenotype should show elevated relatedness of all flatwing individuals regardless of their population origin compared to normal-wing individuals. None of the unrooted trees using Hawaiian individuals, nor rooted trees with Australian individuals as outgroups, supported such an adaptive introgression scenario: flatwing individuals from different populations did not all cluster together as would be expected if flatwing-associated regions were shared due to recent introgression across populations (Fig. 3c, d and Supplementary Figs. 8–25). Because the sex of crickets is determined by an XX/XO (female/male) system, and only males were used here (except the one reference genome that was represented by triangle), these trees can also be considered to suggest relationships of different haplotypes.

We further tested this idea using samples for which wild-caught flatwing individuals were resequenced, i.e. Kauai and Oahu flatwings, to avoid confounds of using lab-bottlenecked individuals. If flatwing introgressed very recently through these Hawaiian populations, a divergence of flatwing-associated allele (s) between population pairs should be lower than the divergence of other loci on the X chromosome^{40–42}. Instead, we found that net nucleotide divergence (d) between Oahu and Kauai wild flatwing individuals at flatwing-associated regions was significantly greater than that for all other X chromosome regions (Fig. 3e, two-sided t -test: $P < 0.001$; $n = 437$ sliding windows within flatwing region, $n = 66,211$ sliding windows outside flatwing region; additional statistical details are reported in Supplementary Fig. 26a), while absolute divergence (d_{xy}) between flatwing-associated regions versus other X chromosome regions was similar (Supplementary Fig. 26b, see ‘Methods’ section). These results, combined with a lack of shared flatwing-associated SNPs and structural variants and long-range, complete linkage disequilibrium on flatwing-associated scaffolds (see below) supported the recent independent origins of flatwing crickets on different islands.

Structural variants (SVs) can have a strong influence on adaptive evolution in insects^{43,44}. We identified a total of 6971 deletions, 4482 translocations, 3385 inversions, 211 insertions and 17,670 copy number variation regions (CNVRs) located in the X chromosome among Kauai, Oahu and Hilo populations (Supplementary Fig. 27, Supplementary Table 8). If the same structural variants were associated with flatwing morphology on different islands, this would be inconsistent with independent mutational origins of flatwing in those populations. However, we did not uncover any candidate association between structural variants and the flatwing phenotype within islands except for Kauai

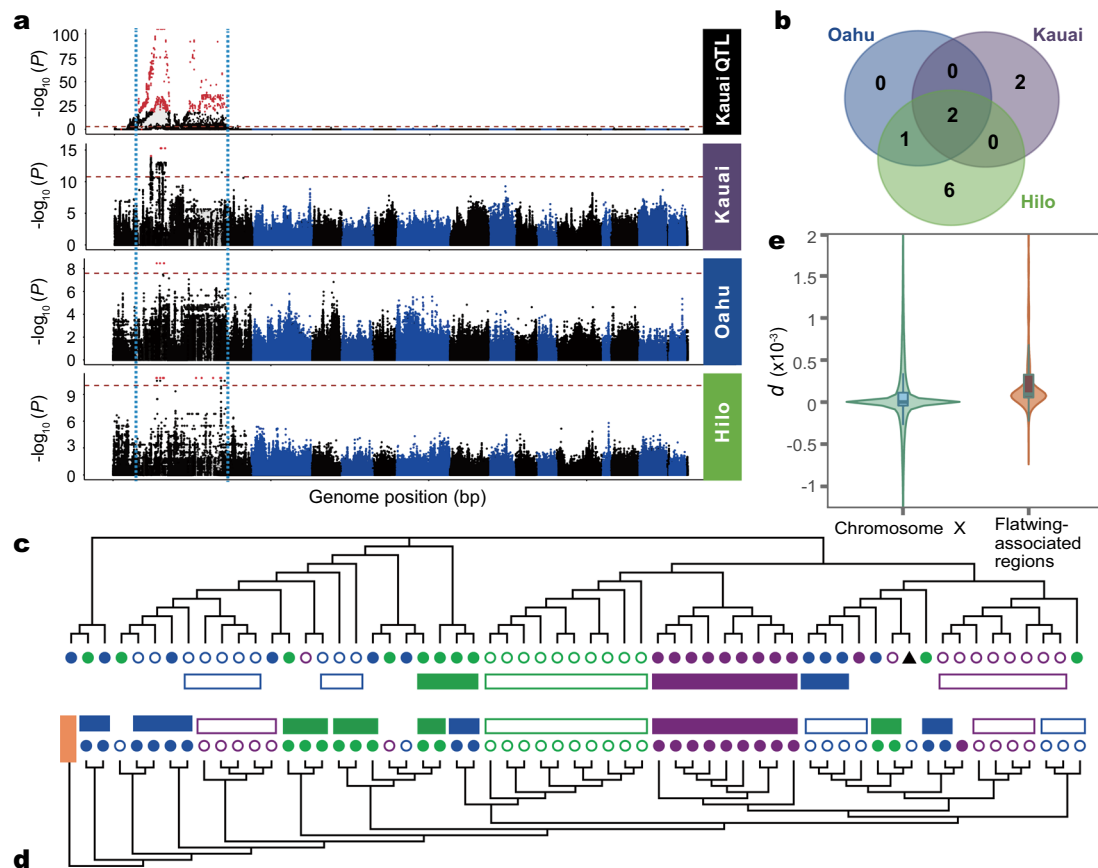


Fig. 3 Putative genomic hotspots containing different flatwing-associated loci in cricket populations on three Hawaiian islands. **a** Flatwing association mapping. Data from a previously published flatwing QTL²¹ are shown on top, with red symbols showing flatwing-associated SNPs and blue vertical lines the extent of the QTL on the X-linkage group (left-most block of black markers). Horizontal dashed lines indicate FDR-corrected significance thresholds, and alternating shades of black and blue indicate different linkage groups. For the lower three panels, the most significant markers are defined as described in the main text using the strictest criterion and plotted in red. **b** Venn diagram showing candidate X chromosome scaffolds associated with the flatwing phenotype for each population, all detected using the most stringent Oahu criterion. **c** Unrooted maximum-likelihood consensus tree for the two flatwing-associated scaffolds shared by all islands. **d** Rooted maximum-likelihood consensus tree for all 1 kb flanking regions of all flatwing-associated SNPs. **c, d** Normal-wing males are represented with solid circles and flatwing males with open circles. Clustered individuals (≥ 2) are highlighted using filled (normal-wing) and open (flatwing) boxes if they were from the same island population and shared the same wing phenotype. Black triangles represent the individual used as the reference genome²¹. Orange rectangle represents seven Australian individuals, which were used as an outgroup. **e** Violin plot of net genetic distance (d) outside and within flatwing-associated regions on chromosome X between flatwing Kauai and flatwing Oahu samples. Distances were calculated in 10 kb sliding windows with 2.5 kb steps ($n = 66,211$ and 437 windows, respectively). Values for upper quartile, median, and lower quartile, plus $1.5\times$ interquartile ranges, are indicated with box plots in the middle of the violin plots. Additional statistical details are provided in Supplementary Fig. 26a.

(Supplementary Fig. 27 and Supplementary Data 7–8), consistent with the hypothesis of parallel evolution.

Signatures of selection on flatwing-associated regions differ across populations. To examine the genomic signatures of selection on flatwing males, we identified selective-sweep regions by comparing π log-ratios vs. population differentiation (F_{ST}) between wild flatwing Hawaiian crickets and normal-wing Australian crickets, which are not under selection pressure from *O. ochracea* (Fig. 4a)^{33,45}. Approximately 1460 putative selective-sweep regions were identified, comprising ca. 1% of the genome. Fifteen flatwing-associated SNPs were located in the selective-sweep regions (Supplementary Table 9). The two flatwing ‘hotspot’ scaffolds identified by GWAS analyses and shared by all islands (Contig18404_pilon and Contig6636_pilon, hereafter ‘scaffold 18404’ and ‘scaffold 6636’; Fig. 3b) were also recovered in this independent selective-sweep analysis (Fig. 4 and Supplementary Figs. 28–30).

Selection is expected to reduce genetic diversity and elevate linkage disequilibrium (LD)^{33,45}. For all wild populations, flatwing individuals repeatedly showed significantly lower nucleotide diversity (π) than normal-wing individuals in the flatwing hotspot scaffolds (Wilcoxon rank-sum tests, all $P < 0.001$, 5% of empirical distribution, Fig. 4c, d and Supplementary Fig. 30a, b). F_{ST} between flatwing and normal-wing individuals was also significantly elevated above genome-wide levels (Fig. 4c, d and Supplementary Fig. 30). Consistent with recent extreme positive selection, we observed unusually long-range complete LD and significantly negative Tajima’s D -values in Kauai, in which rapid evolution completely fixed the flatwing morphotype within ca. 15 years^{22,24,25} (top panels of Fig. 4b and Supplementary Figs. 28, 29). However, the Oahu population showed a much different signature of selection in this region (bottom panel of Fig. 4b, e). For Oahu flatwing individuals, the flanking regions of flatwing-associated SNPs showed only a very modest elevation of LD (Supplementary Fig. 28). In addition, when we analysed all the wild Oahu samples (including flatwing and

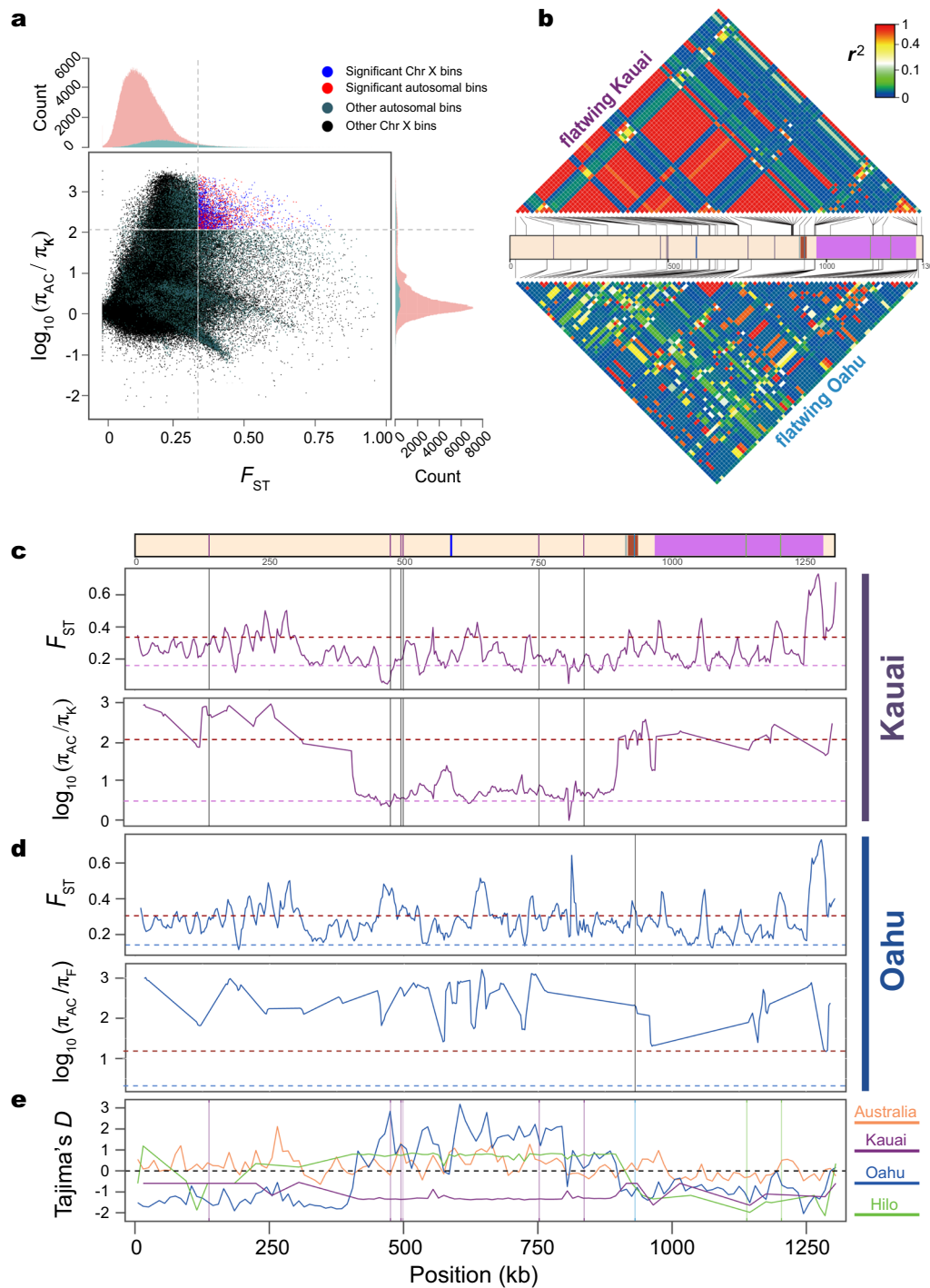


Fig. 4 Parallel adaptation is associated with different signatures of selection on putative flatwing hotspot, scaffold 18404. **a** Loci showing signatures of selective sweeps in wild Kauai flatwing males were identified by examining the distribution of fixation index (F_{ST}) and relative diversity ($\log_{10}[\theta_{\pi, \text{unselected_normal-wing}}/\theta_{\pi, \text{selected_flatwing}}]$) between flatwing Kauai individuals and normal-wing Australian individuals. Datapoints represent 10 kb sliding windows calculated in 2.5 kb increments across the entire genome. Blue points and red points correspond to X chromosome and autosomal bins, respectively, showing the most significant selective-sweep signatures (top 5% of the empirical distribution). **b** Patterns of linkage disequilibrium (LD) along the flatwing-associated scaffold 18404. Top and bottom heatmaps show pairwise LD (r^2) between SNPs (top—Kauai, bottom—Oahu), and SNPs were filtered as detailed in the ‘Method’ section to ensure LD is comparable between flatwing males from Kauai vs. Oahu. The colour scale refers to the value of r^2 . Protein-coding genes and flatwing-associated SNPs are indicated by blocks and bars with different colours respectively (*Prospero* = brown, *Doublesex* = purple) in the middle panel. **c–e**, Signatures of selection on scaffold 18404 associated with parallel evolution in Hawaiian populations. Genes and flatwing-associated SNPs are shown in the top panel using the same scheme as in **b**. F_{ST} and π were calculated using sliding-window analyses with 10 kb windows for **c** Kauai and **d** Oahu. $\pi_{\text{ratio}} = \theta_{\pi, \text{unselected_normal-wing}}/\theta_{\pi, \text{selected_flatwing}}$. Horizontal dashed lines represent mean whole-genome value for the values shown in each panel, and dark red horizontal dashed lines in **c** and **d** represent the top 5% threshold values. **e** Tajima’s D for all wild-caught individuals. Kauai flatwing (purple), Oahu (blue, all samples combined), Hilo normal-wing (green), Australian normal-wing (orange). The dashed horizontal line indicates the Tajima’s D null hypothesis for neutrality.

normal-wing), Tajima's D was significantly positive at this strongly selected region (Fig. 4e). This pattern of Tajima's D is consistent with balancing selection, which in turn is consistent with field observations that the population in Oahu persistently contains around 45% flatwing males²². The discovery of a new morph of silent male, which is known to protect against the fly, 'curly-wing', in Oahu raises the intriguing possibility that balancing selection arises as a result of a competing, adaptive male-silencing variant, which exerts its phenotypic effects not through disrupting male wing venation, as in flatwing, but instead through a disruption of the 3-dimensional structure of the wings²⁵. Finally, in contrast with Kauai and Oahu populations, Tajima's D -values of Australian and Hilo populations fluctuate around 0 in this region (Fig. 4e).

Flatwing gene expression during wing development differs across islands and implicates *doublesex*. We performed a separate gene expression study to examine the developmental genetic basis of the flatwing phenotype and test whether this is linked to the putative hotspot identified above, scaffold 18404. Our GWAS and selection analyses highlighted two important candidate genes, *prospero* (*pros*) and *doublesex* (*dsx*), and three flatwing-associated SNPs supported in the selection analysis were directly located within introns of these genes (Fig. 4c, top panel). *Pros* and *dsx* both have known roles in insect wing development. *Pros* is a transcription factor targeting *Tollo*, which induces cell sorting in *Drosophila* wings^{46,47}. *Dsx* is also a transcription factor widely involved in sex-specific developmental patterning^{48,49}, with intriguing relevance to the evolution of sex-biased wing venation in flatwing crickets. For example, a wing-specific isoform of *dsx* acts as a master regulatory switch in the butterfly *Papilio polytes*, switching an entire polymorphic, mimetic wing pattern on and off in females⁵⁰.

Crickets are hemimetabolous and wing venation is established early during nymphal development⁵¹, and there are widespread gene expression differences between *flatwing* and *normal-wing* genotypes during embryonic development even prior to hatching²¹. Crickets develop wingbuds from thoracic precursors, which eventually become externally articulated during later nymphal stages⁵². As such, we performed gene-level differential expression analyses of developing wingbuds at the earliest stage when they become freely articulated, two instars prior to adult eclosion (Fig. 5a). We collected wingbud samples from pure-breeding *normal-wing* and *flatwing* lines that we previously established from Kauai and Oahu populations (no such lines were available for Hilo), and expression analysis reinforced the importance of the genomic region containing *dsx* to the manifestation of independently derived flatwing phenotypes from different populations.

There was little overall similarity in wingbud gene expression changes associated with *flatwing* genotypes from Kauai vs. Oahu (Fig. 5b). In Kauai, 190 transcripts were uniquely DE between morphs, and in Oahu 449 transcripts were uniquely DE between morphs. Just 10 transcripts showed correlated differential expression in both populations, and all of them were down-regulated in *flatwing* genotypes. Two of these consistently DE transcripts were located on scaffold 18404. One was *dsx*, the other a transcript with no functional annotation from an overlapping region of the scaffold; both were among a cluster of transcripts that showed correlated morph-associated expression differences between islands (Supplementary Fig. 31). The extent of *dsx* downregulation in *flatwing* genotypes mirrored the completeness with which song-producing structures on male wings were feminised in each population; that is, *dsx* was more down-regulated in Kauai flatwing males, where male-specific wing venation is more reduced (Fig. 5c). *Dsx* showed higher overall

expression in females, but was also sensitive to genotype with lower expression in *flatwing* carriers. These sex-specific effects of *dsx* are consistent with sex-specific developmental effects that have been found in other organisms, such as horned beetles (*Onthophagus taurus*) in which *dsx* dysregulation affects the development of male and female horn length in opposing directions^{48,49}. We found no evidence of sex- or morph-segregating *dsx* isoforms. We cannot exclude that these might be detected at a different developmental stage or in different tissues, but note with interest that there are now multiple exceptions to the expectation of sex-specific splicing in arthropods⁵³. The expression patterns of the above two transcripts were further validated by re-analysing an independent, previously collected RNA-seq dataset (see 'Methods' section). *Prospero* was lowly expressed in wingbud tissues (mean count-per-million of 0.35) so did not pass expression filtering of transcripts, nor showed any evidence of differential expression between genotypes after relaxing the filtering threshold (unadjusted $P > 0.88$) (Supplementary Fig. 33). These results implicate a well-known regulator of sex-specific phenotypic development in the rapidly evolved loss of male song in *T. oceanicus*, and provide further evidence that a genomic adaptation hotspot has a vital role in determining multiple independently derived male-silencing phenotypes. Although it is currently not possible to establish whether *dsx* expression is causative or an associated downstream effect of flatwing development, these findings suggest a tentative genetic mechanism of flatwing determination. Mutations in this region may alter the expression level of *dsx*, and contribute to spatio-temporal variation of downstream genes.

Discussion

Understanding the genetic mechanisms of adaptive evolution is a major focus of evolutionary biology⁵⁴. Dissecting the mechanisms underlying convergent phenotypic evolution is extremely useful for gaining a better understanding of adaptation more generally: why, in some cases, do adaptations in different populations and species share a genetic origin, whereas in other cases functionally equivalent adaptations are the result of multiple independent evolutionary origins? Gene flow is generally thought to counteract the generation and maintenance of such parallel adaptations because it can introduce pre-existing adaptive alleles into populations under selection pressure, which can then be directly selected faster than de novo adaptive mutations can arise³⁻¹⁰. However, our results from silent Hawaiian crickets surprisingly show that this need not be the case.

The recent, abrupt, and rapid evolution of flatwing male crickets in multiple Hawaiian populations is one of the fastest rates of convergent phenotypic evolution documented in the wild^{21,23}. We found readily detectable variation in the morphology of adaptive flatwing phenotypes from populations on Kauai, Oahu and Hilo, and these morphotypes were associated with different genetic loci. There were multiple signals of independent origins of the silent flatwing male phenotype in different island populations: (1) distinct morphometric patterns of flatwing venation on different islands, (2) entirely non-overlapping flatwing-associated SNPs among all three populations, (3) all phylogenetic trees for flatwing-associated genetic regions falsified a scenario of adaptive introgression (4) greater net nucleotide divergence at flatwing-associated regions (5) lack of shared flatwing-associated structural variants (6) long-range, complete linkage disequilibrium on flatwing-associated scaffolds in Kauai but not Oahu (7) little overall similarity in wingbud gene expression changes associated with flatwing genotypes from different islands. This parallel adaptation occurred despite considerable recent, ongoing gene flow among all three populations.

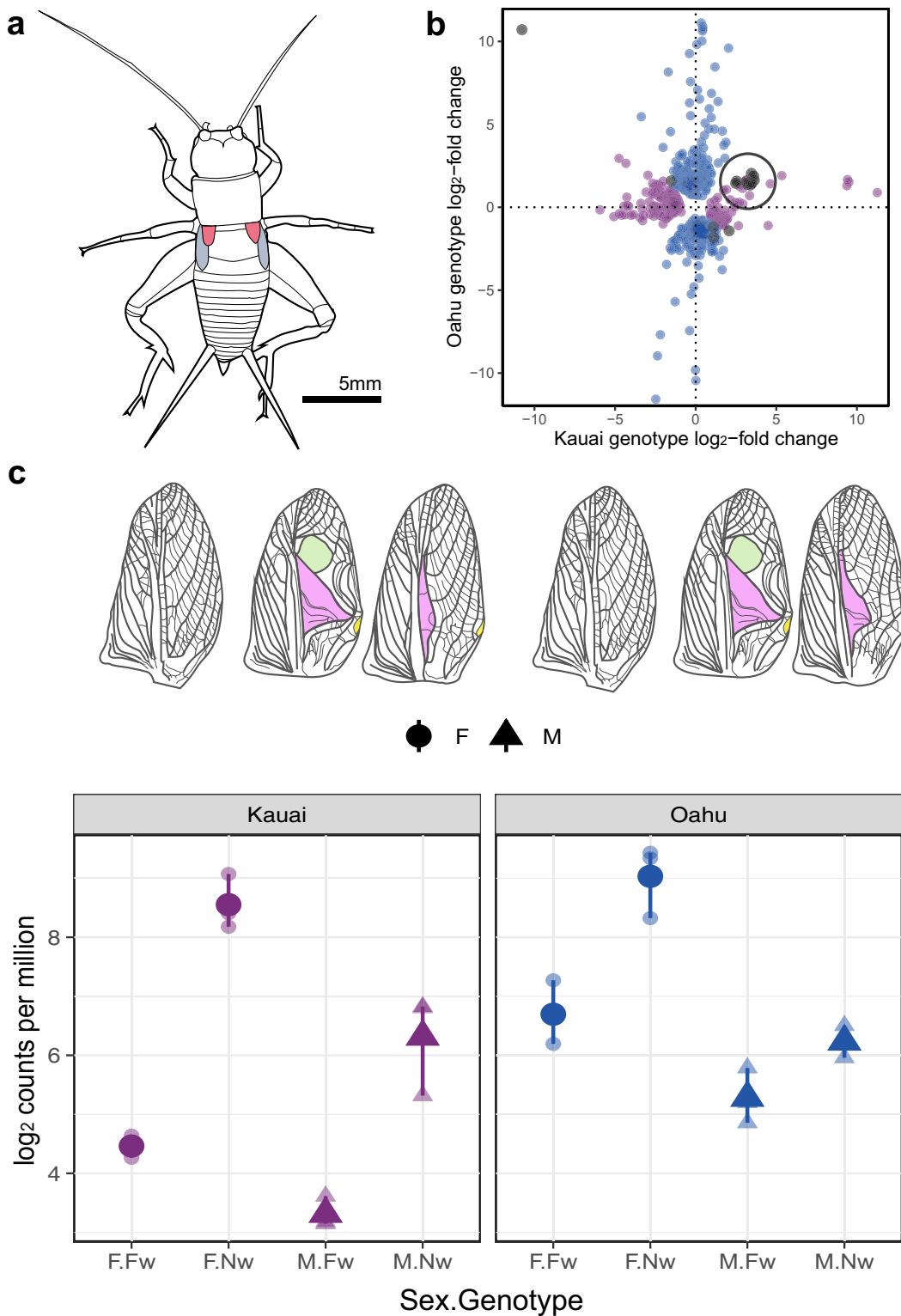


Fig. 5 Differential gene expression analysis of developing male forewings. **a** Diagram of the penultimate nymphal instar from which right dorsal forewing buds were sampled. Forewings are highlighted in red, hindwings in blue. Diagram drawn by JG Rayner. **b** The majority of transcripts identified as differentially expressed between *normal-wing* and *flatwing* genotypes in Kauai (purple) and Oahu (blue) were non-overlapping, with the exception of 15 transcripts identified in black. Ten of the latter showed coordinated changes in expression (highlighted with a circle). Log₂-fold changes > 0 on either axis indicate normal-wing biased expression. **c** Top panel: diagrams of forewing phenotypes for female, flatwing and normal-wing males from Kauai and Oahu²³. Flatwing males' wings are feminised to different degrees in Kauai vs. Oahu, but female forewings do not noticeably differ between genotypes. Bottom panel: corresponding patterns of relative expression of *doublesex* in females and males from each population. Large points show means, small points show individual observations for each replicate, and lines indicate the minimum and maximum values. Source data are provided as a Source Data file. Diagram is drawn by JG Rayner.

T. oceanicus in the Hawaiian islands have, like many island species, lost long-distance flight capability, and migration appears to be maintained by human activities, which became more frequent as human settlement increased 300 years ago (Fig. 2b).

Flatwing phenotypes from different islands were associated with a small number of shared putative adaptation hotspots in the genome, and signatures of selection on these varied from population to population and corresponded to known demographic dynamics in the wild. Follow-up gene expression analyses gave further support for the involvement of a transcription factor well-known for controlling sexually dimorphic morphological traits, *doublesex*, in multiple origins of adaptive male wing feminisation. These results illustrate how strong selection associated with genomic hotspots of adaptation can overwhelm the generally predictable influence of gene flow. It remains an open question what level of gene flow would inhibit the parallel evolution of male-silencing variants in this system. However, it is notable that the gene flow we detected in *T. oceanicus* was particularly strong compared to estimates in other species^{55–57}. The precise dynamics of mutation, migration and drift in shaping adaptation is a longstanding and complex problem, so in the silent flatwing cricket system and in other systems undergoing rapid convergent evolution, it would be illuminating to perform evolve and resequence experiments to define threshold levels of gene flow beyond which parallel mutations localised within distinct populations become swamped by adaptive variants originating from other populations.

The tension between mutation and migration in shaping the dynamics of adaptive evolution is a central idea in evolutionary theory, which predicts that adaptive introgression should reduce the role of de novo mutations when selection is extreme and gene flow is high^{3,5,14}. Exceptions may occur when mutation rates are particularly high, but it has been challenging to evaluate whether these trade-offs shape convergent phenotypic evolution in a predictable fashion. Most well-characterised cases of parallel evolution via independent mutation involve a clear trade-off between mutation and migration. For example, recurrent pelvic reduction in freshwater populations of the stickleback *Gasterosteus aculeatus* occurs due to a hyper-mutable enhancer region of the transcription factor *Pitx1*^{12,13}. Affected populations are isolated due to the repeated invasion of separate freshwater systems, yet high mutation rate of an important genomic hotspot of adaptation has facilitated parallel evolution of adaptive pelvic reduction in each habitat. In contrast, introgressive hybridisation between different species of mice introduced a pesticide resistance allele from *Mus spretus* to *M. musculus domesticus*, which was subsequently selected due to the increased application of anticoagulant pesticides in the mid-1900s⁵⁸. Similarly, there is compelling evidence that adaptive loci associated with mimetic wing patterning in *Heliconius* butterflies have introgressed from species to species, facilitating convergent adaptive radiations^{16,59}. It is important to note that shared derived adaptations introduced via migration as well as uniquely evolved adaptations can co-occur, and compelling examples have been observed in the context of antimalarial resistance in *Plasmodium falciparum*⁶⁰. However, the evolutionary processes underlying the origins of these adaptations overwhelmingly obey the mutation-migration trade-off. Our findings contradict this general tendency and raise the interesting possibility that incompatibilities between unique, derived adaptive alleles could act as constraints on convergent evolution via introgressive hybridisation. A hypothetical example is where one coevolved system is resistant to introgression of alternative alleles because the latter are individually maladaptive in combination with the resident system. There is precedence for such genomic incompatibilities to affect adaptive variants, for example the homozygous lethal inversion polymorphism in the ruff,

*Philomachus pugnax*⁶¹. In the cricket system, it is known that flatwing haplotypes from Kauai (*fw^K*) and from Oahu (*fw^O*) can be expressed in alternative genomic backgrounds^{23,27}, but the fitness consequences of this are not known. Future work testing constraints on convergent evolution via adaptive introgression is necessary to evaluate this scenario.

Empirical examples that violate this predictable trade-off are rare. Part of the reason for this may be the well-known challenges of disentangling demographic signal from introgression or independent mutational origins when convergent evolution is supposed to have occurred in the distant evolutionary past¹⁹, but the silent Hawaiian cricket system is uniquely suited to overcome this issue owing to the recent origin and spread of flatwing morphs in all three populations studied here²². If, as our results suggest, parallelism is less constrained by gene flow than previously appreciated, it is worth considering what factors favour such an outcome. The functional genetic basis of adaptive traits and the phenotypic nature of the adaptation may each have important roles in relaxing this constraint. In crickets and other ensiferan insects, forewing venation is a sexually dimorphic trait critical for sexual reproduction. Signals emitted by resonating membranes on the wings are major components of the mate recognition system, often the first long-range signal detected by receivers, and they have a well-known role in sexual selection, reproductive isolation and diversification⁶². Male forewings and the signals they produce are highly variable across ensiferan taxa^{27,63–65}. Traits showing an obvious signature of extensive evolutionary variation have, by definition, arisen through the action of repeated genetic mutation, and could therefore be particularly disposed to recurrent parallel evolution. In addition, loss of an existing complex trait may be more evolvable than gain of a novel trait⁶⁶, because disrupting the developmental genetic pathways underlying a complex structure such as a cricket wing resonator or a stickleback's pelvic girdle can be achieved through simple genetic change affecting one of many loci. Involvement of the *doublesex* pathway in the evolution and development of flatwing male crickets is consistent with such an evolutionary mode of action. *Dsx* is an essential regulator of sexual dimorphism in insects, for example, controlling the development of sexually dimorphic horns in the dung beetle *Onthophagus taurus*⁶⁷ and the Japanese rhinoceros beetle *Trypoxylus dichotomus*⁶⁸, female-limited wing mimicry in *Papilio* butterflies^{50,69,70}, sexual morphology in *Drosophila melanogaster*^{48,49}, and male scent organs in *Bicyclus* butterflies⁷¹.

The emergence and spread of parallel adaptations in populations subject to gene flow are liable to be biased towards highly evolvable traits such as those functioning in sexual selection and towards genomic hotspots of adaptation where relatively simple changes result in functional losses. Given the increasing research focus on evolutionary rescue and rapid adaptation in wild populations, the speed of such adaptation, its repeatability, and the factors favouring it are important to dissect and translate to applied contexts, to mitigate resistance evolution in agricultural pests and human pathogens, and also to design effective interventions to cope with negative consequences of anthropogenic environmental change⁷². The repeated loss of song in Hawaiian crickets under pressure from eavesdropping parasitoids provides a compelling test case in a wild population undergoing 'real-time' evolutionary adaptation, showing that extremely rapid parallel adaptation can occur despite evolutionary forces traditionally thought to impede it.

Methods

Morphometric analyses of male wings. Right forewings of wild flatwing male crickets were collected from three Hawaiian islands and mounted between two microscope slides (Supplementary Table 1). Fifty-two Kauai flatwings, 52 Oahu

flatwings, and 10 Hilo flatwings were sampled. The unbalanced sample size was caused by the extremely low frequency of flatwing crickets in the wild in Hilo²². Slides were photographed using a Leica DFC295 digital camera mounted on a Leica M60 dissecting microscope. We performed a landmark-based geometric morphometric analysis of these photos²⁷. Specifically, landmark data were generated by manually placing 16 landmarks on each wing photograph using the tpsDig2 (v2.31)⁷³ as described in Fig. 1a. The criteria for placement of each landmark are identical to those used by Pascoal et al.²³ and Bailey et al.²⁷. Morphological comparisons were then conducted using MorphoJ (v1.07a), which extracts shape information by performing a Procrustes superimposition on all the landmark data⁷⁴. A covariance matrix generated from Procrustes distances was then used to perform principal component analysis (PCA) without a priori information of group (i.e. island) identity. The PCA results and corresponding group information were then analysed using a multivariate analysis of variance (MANOVA) in R v. 3.5.2 to test whether flatwings from the three groups were morphologically different. They were, so we used a canonical variates analysis (CVA) to visualise the differences among islands. CVA is designed to maximise the separation of three populations, producing axes, which best explain inter-group variation²³. A scatter diagram and a transformation grid plot were plotted to visualise the CVA results.

Sampling, DNA extraction and whole-genome re-sequencing. Seventy individuals were collected for whole-genome re-sequencing (Supplementary Table 3). We used ten flatwing males and ten normal-wing males from each Hawaiian island (Kauai, Oahu and Hilo), seven Australian *T. oceanicus* males and three males of a closely related, but reproductively isolated sister species from Australia, *T. commodus*. The normal-wing males from Hilo and flatwing males from Kauai were collected from the wild in 2017. However, normal-wing males were non-existent in the Kauai population at that time, and flatwing males were extremely scarce in the Hilo population, so for these groups we used laboratory specimens from stock populations derived from previous collections at exactly the same locations; this enabled us to breed extremely large population sizes to obtain enough of the rare morphs for re-sequencing. For Oahu crickets, both flatwing and normal-wing samples were from wild populations. Australian individuals were collected from the wild in Australia in 2011. Single hind legs of each sampled individual were collected and preserved in 100% ethanol. Genomic DNA was extracted from legs using the CTAB method. The quantity of the extracted DNA was measured using a Qubit dsDNA kit. The quality of the DNA samples was evaluated using Nanodrop, 1% agarose gel electrophoresis, AATI Fragment Analyser and the Standard Sensitivity Genomic DNA Analysis Kit. The Hamilton MicroLab STAR liquid handling system was used to pre-normalise DNA samples before library preparation. DNA was sheared to a 450 bp mean insert size using a Covaris LE220 focused-ultrasonicator. Sequencing libraries were prepared using Illumina SeqLab specific TruSeq Nano High Throughput library preparation kit, Hamilton MicroLab STAR and Clarity LIMS X (4.2) Edition, and then clustered onto a HiSeqX Flow cell v2.5 on cBot2s. The clustered flow cell was transferred to a HiSeqX for sequencing using a HiSeqX Ten Reagent kit v2.5.

Sequence quality checking, read alignment and SNP identification. Demultiplexing was performed using bcl2fastq (v2.20, Illumina), allowing 1 mismatch when assigning reads to barcodes. Adaptors were trimmed during the demultiplexing process. Raw reads that included $\geq 5\%$ unidentified nucleotides (N) or more than 65% low-quality nucleotides (Phred quality value ≤ 7) were discarded³³. Before alignment, index files of the field cricket (*T. oceanicus*) reference were generated using bwa and samtools (v1.3.1)⁷⁵. The filtered pair-end sequence reads were then mapped to the reference genome²¹ using BWA-MEM (v0.7.12) with default parameters⁷⁶. Secondary alignments with flag ≥ 255 were discarded using a customised Perl script³³. Samtools (v1.3.1) was used to assess coverage individually. We used samtools and picard (2.14.1) to sort, format and index the resulting binary alignment map (BAM) files and remove PCR duplicates. To increase the accuracy of alignments around indel regions, we detected indels and re-aligned these regions using RealignerTargetCreator and IndelRealigner functions in GATK (v3.7.0)⁷⁷. All following analyses were conducted based on these re-aligned BAM files. We detected SNPs and indels following GATK best practice workflows⁷⁸. Briefly, raw SNPs and indels were first detected individually using HaplotypeCaller. CombineGVCFs was then applied to merge SNPs of different individuals from the same population. The VCF records of multiple populations were then re-genotyped and aggregated using GenotypeGVCFs. Following the authors' recommendations, hard-filtering criteria (Qual By Depth < 2.0, Fisher Strand > 60.0, RMS Mapping Quality < 40.0, Mapping Quality Rank-Sum Test < -2.5, Read Pos Rank-Sum < -1.0) were decided manually based on density plots of the raw SNP dataset and applied using SelectVariants and VariantFiltration. To further improve the accuracy of genotyping and economise computing resources, SNPs of extremely short scaffolds (≤ 5 kb) were not called. In consideration of the fact that sex of crickets is determined by an XX/XO (female/male) system²¹, all males were fixed to have homozygous genotypes of X-linked loci⁷⁹. Low-quality SNPs potentially caused by sequencing error were also discarded using a customised Perl script. Specifically, all SNPs with Phred-scaled quality score < 30 or with abnormal depth (lower than one-third or higher than three times the average sequencing depth of a particular individual) were filtered³³. The X scaffold was defined according to the previously published linkage map²¹. For SNPs located in X-linked scaffolds, the criteria of abnormal

depth were set as the half that of the autosomes. Heterozygous calls of X-linked SNPs were fixed to homozygous based on the allele depth if the depth of one allele was greater than three times the other one. If none of these alleles were supported by the supermajority, the genotype was treated as unknown. Loci that had unknown genotypes in more than 25% of all individuals were also excluded. To avoid bias caused by outgroups (unrelated populations), multiple SNP datasets were generated independently and utilised accordingly for the following analyses (Supplementary Data 2).

Phylogenetic, principal component (PC) and population structure analyses. We used PLINK (v1.90b6.8)⁸⁰ to calculate identity-by-state based on SNP datasets. The resulting distance matrix was then passed to PHYLIP (v3.696)⁸¹ to construct neighbour-joining phylogenetic trees (Fig. 1f and Supplementary Figs. 2, 3). *T. commodus* individuals were used as an outgroup. We generated consensus trees by performing bootstrap replication 1000 times and merging them. Consensus trees were then labelled and visualised using MEGA (v10.0.5)⁸² and FigTree (v1.4.4). We used the smartpca programme in EIGENSOFT (v7.2.1)⁸³ to perform the principal component analysis (PCA) and visualised the result by plotting a scatter diagram using rgl and ggplot2 R packages⁸⁴. We used ADMIXTURE (v1.3.0) to estimate population structure with the numbers of ancestral populations (*K*) from 2 to 7 after thinning the SNP set for linkage disequilibrium using PLINK with an r^2 threshold of 0.1⁸⁵.

Genetic diversity and linkage disequilibrium analyses. We used VCFtools (v0.1.16)⁸⁶ to calculate genetic diversity with a sliding-window size of 10 kb and a step size of 2.5 kb. To calculate linkage disequilibrium, the original SNP dataset was first divided into subsets of each population using vcfutils. The split VCF files were then converted to haploview-recognised format using PLINK. Haploview (v4.2) calculated correlation coefficients (r^2), which are used to infer pairwise linkage disequilibrium⁸⁷. To visually depict LD patterns on scaffold 18404 (Fig. 4b), a reasonable number of SNPs dispersed across the whole scaffold is necessary. However, different sets of positions were polymorphic in flatwing Oahu and flatwing Kauai individuals, which makes direct visual comparison difficult to interpret. To minimise visual bias caused by using different SNP sets, we used the flatwing Oahu group as a standard and then selected the corresponding SNP in the Kauai group; for sites at which the same SNP was unavailable in Kauai, the nearest one in its flanking regions was selected (Supplementary Data 9).

Gene flow and demographic history inference. To avoid bias caused by inbreeding, only wild-caught, unrelated samples were used to infer gene flow and demographic history of Hawaiian cricket populations. We used the ABBA-BABA (*D*-statistic) test^{30,88} and multiple individuals' *D*-statistic test³⁴ to detect gene flow among different populations. Six different scenarios were statistically tested. To do so, BAM files of individuals from three Hawaiian populations, one Australian population and one outgroup population (*T. commodus*) were passed to ANGSD (v0.930), which counted the number of ABBA and BABA sites and calculated *D*-statistic metrics^{34,88}. We inferred relatively late (within 10,000 generations) demographic parameters including the divergence of three populations, bottlenecks and migrations from the site frequency spectrum (SFS) using Fastsimcoal³⁵. To minimise bias caused by potential inbreeding, linkage and selection, the following criteria were applied to further filter SNP datasets before format conversion: (i) Only autosomal scaffolds were retained, based on previous linkage map, (ii) SNPs with an $r^2 > 0.3$ were excluded using the sliding-window method, (iii) PLINK was used to exclude one of the individual pairs, which had a proportion identity-by-descent outlier, (iv) SNPs located inside genic regions were excluded, (v) Sites with unknown genotype were also filtered. The filtered SNP dataset was converted to Arlequin format using PGDSpider (v2.1.1.5)⁸⁹. We used Arlequin (v3.5)⁹⁰ to produce the minor allele SFS (folded SFS) and joint multidimensional SFS. Using Fastsimcoal (v2.6.0.3), we tested 62 models including effective population size changes, different divergence scenarios, presence or absence of gene flow, and permutations of three island populations. For each model, we performed 50 independent parameter inferences. For each run, 100,000 simulations per likelihood and 40 Expectation Conditional Maximisation (ECM) cycles were performed to estimate parameters with maximum-likelihood. We compared different models using the Akaike Information Criterion (AIC), and estimated the 95% confidence interval of the best model using 30 parametric bootstraps³⁵. We used the Pairwise Sequentially Markovian Coalescent (PSMC) method to reconstruct the most recent common ancestor (TMRCA) distribution, infer the divergence time from now to the TMRCA, and finally infer the history of effective population size³⁶. Since local density of heterozygotes information was used by PSMC, X-linked scaffolds were excluded from this analysis. To further improve the accuracy of this analysis, only wild individuals with the highest sequencing depth (top 2, sequencing coverage > 26 \times) were used. We used samtools to generate diploid consensus sequences based on individual BAM file. Loci with abnormal depth were masked by N. PSMC (v0.6.5-r67) was used to infer historical effective population size with a mutation rate of 3.5×10^{-9} and a generation time of 0.25 year^{91,92}.

Genome-wide association analysis. The flatwing phenotype is known to be inherited as a single locus, X-linked trait. Sex of crickets are determined by XX/XO

(female/male) system, and the expression of *flatwing* is male-limited²¹. We used flatwing and normal-wing males to perform three independent, basic case/control association tests for each Hawaiian population sampled. For Oahu, 10 wild flatwing males and 10 wild normal-wing males were used. Because of the rarity of flatwing males in the Hilo population, and normal-wing males in the Kauai population, it was impossible for us to collect enough samples in the wild. Hence, laboratory stocks were used (see above). Thus for Hilo, we used 10 normal-wing wild males and 10 flatwing lab-reared males. For Kauai, because of the inbreeding of lab-reared normal-wing individuals (caused by the absence of wild normal-wing males), the association analysis was performed between 10 wild Kauai flatwing males and 30 normal-wing males from Kauai lab stock, wild Hilo, and wild Oahu populations. The standard case/control threshold for association significance was generated using Fisher's exact test implemented in PLINK^{79,80}. *P*-values were Bonferroni-adjusted in *R* to account for multiple testing. To ensure results were not affected by different numbers of individuals entering GWAS analyses for different islands, we performed another independent GWAS for the Oahu population using the same strategy as for Kauai (10 wild Oahu flatwing vs. all normal-wing). It generated qualitatively identical results (Supplementary Table 10). This analysis was unavailable for Hilo, because in that population the case group (flatwing) was lab-raised. However, phylogenetic tree analyses (see below, and Fig. 3c, d and Supplementary Figs. 8–25) tested all possibilities, including with trees built from a merged set of all GWAS-significant SNPs and their flanking regions, and none were consistent with *flatwing* introgression.

Structural variant and copy number variation region detection, and comparison between male phenotypes. We detected structural variants (SVs) across the X chromosome. The alignment results of reads mapped to the X chromosome were first extracted from BAM files using customised Perl script and SAMTOOLS⁷⁵. These extracted BAM files were then passed to bam2cig to generate a per-invocation config file⁹³. Based on this config file and the X-chromosomal BAM files, Breakdancer (1.3.6) was used to detect deletion, insertion, inversion, and translocation⁹³. The SVs (i) smaller than 100 bp, or (ii) with a quantity score < 30, or (iii) supported by fewer than 2 reads were filtered⁹⁴. For each island, SVs of flatwing vs. normal-wing morphs were compared. Flatwing-differentiated SVs were defined according to the association significance calculated using Fisher's exact test. We used the read-depth method implemented in CNVnator (v0.3.3), which is a software combined mean-shift approach, multiple-bandwidth partitioning and GC correction, to detect copy number variants (CNVs)⁹⁵. To maximise the accuracy and sensitivity, a bin size of 100 was used. We also filtered the detected CNVs with $p > 0.05$ or $q0 > 50\%$ to avoid potential bias caused by inequality of sequencing depth and the reads mapped to multiple genomic regions⁹⁵. Considering the fact that all individuals used in this research were males, which have only one X chromosome, the copy number was normalised to one. Copy number variation regions (CNVRs) were defined as the genomic regions that showed copy number variations in multiple individuals (>3). The overlapped CNVs of different individuals were merged into CNVRs using the method, which defined the minimum and maximum boundaries of the CNVs as the boundaries of the CNVR^{96,97}. CNVRs supported by fewer than three individuals were filtered to further reduce the false discovery rate⁹⁷. To detect flatwing-differentiated CNVRs, both V_{ST} and student's *t*-tests were used⁹⁷. Here, $V_{ST} = (V_T - V_S)/V_T$. V_T is the total variance in copy number between the two groups. V_S is the average of the variance within each group, weighted for its sample size^{98,99}. A CNVR in the top 5% of all V_{ST} values plus a student's *t*-test *P*-value < 0.05 was considered to be flatwing-differentiated CNVR⁹⁷.

Parallel evolution analysis. Intersections of flatwing-associated regions of different populations were calculated and visualised using a Venn diagram tool (Bioinformatics & Evolutionary Genomics, Ghent University). The SNPs of flatwing-associated regions were extracted and converted to MEGA-format using Perl scripts. We constructed maximum-likelihood (ML) phylogenetic trees for flatwing-associated SNPs and their flanking regions (Fig. 3c, d and Supplementary Figs. 8–25) using the Tamura–Nei model in MEGA (v10.0.5)⁸². Absolute divergence (d_{xy}) and net nucleotide divergence (d) were used to estimate the amount of DNA divergence (nucleotide substitutions). For the latter, $d = d_{xy} - (d_x + d_y)/2$, where d_x and d_y are per-site pairwise nucleotide differences within populations X and Y, respectively, and d_{xy} represents pairwise nucleotide differences between populations X and Y^{100,101}. Estimates of d and d_{xy} were calculated using PopGenome (v2.6.1)¹⁰². The d and d_{xy} values of different genomic regions in different populations were compared using sliding-window approaches, and statistical differences were tested using 1000 random simulations implemented with custom Perl and R scripts.

Detection of candidate selective-sweep regions. We used samtools to calculate nucleotide diversities (π) and Tajima's *D* for different populations⁷⁵. Fixation indices (F_{ST}) between different populations were also calculated. Specifically, VCF files of corresponding populations were analysed using a sliding-window approach with 10 kb windows and a 2.5 kb step size. Windows overstepping boundaries of scaffolds were filtered. Selection pressure reduces genetic diversity of the targeted region and enhances the degree of differentiation between populations under

selection versus those not under selection^{33,45}. Here, we compared nucleotide diversities of different populations and calculated their π log-ratios in Perl and tested the empirical significance of these measures in *R* using empirical percentiles. Genomic regions within the top 5% of F_{ST} values and top 5% of π log-ratios were identified as candidate selective-sweep regions^{33,45}, and any adjacent (overlapped) candidate regions were merged in to a single candidate region.

Gene content. The gene content of candidate selective-sweep regions, significant GWAS loci, flatwing-differentiated SVs, and flatwing-differentiated CNVRs was obtained by comparing their coordinates with the official gene set for *T. oceanicus*²¹. Genes overlapping the corresponding candidate region (loci) were defined as candidate genes. Functional annotations of these genes were assigned based on Swissport, TrEMBL and the NR database¹⁰³.

Sampling developing cricket wingbuds for gene expression experiment. We collected and sequenced 24 RNA samples from the wingbuds of *normal-wing* and *flatwing* genotypes of both sexes, from Kauai and Oahu lab populations. There were 3 replicate pure-breeding lines per group. Populations used in the experiment were derived from eggs laid by wild-caught females from populations of Kauai and Oahu in 2014⁵². Laboratory populations were reared in a growth chamber at 25 °C, under a 12-h L:D cycle. Kauai lines had been recently outcrossed by mixing all lines, and pure-breeding lines were reconstituted prior to the current experiment by performing genetic crosses across two generations to select only those families showing patterns of offspring morph ratios indicating they were derived from homozygous *flatwing* or *normal-wing* dams; males and females from these lines were, therefore, pure-breeding for the respective genotype, with males carrying one copy and females two. Kauai and Oahu lines were reared and maintained under identical conditions, in 20 L boxes with cardboard shelter, and food and water available *ad libitum*. Individuals were removed from stock boxes at their penultimate nymphal instar (Fig. 5a), when developing wings are first externalised, ca. 1.5 months after hatching. Wing differences between *normal-wing* and *flatwing* genotype males arise early during their development and become pronounced at this stage⁵². Individuals were briefly anaesthetised using CO₂, and the dorsal-right forewing bud was removed using micro-dissection scissors and stored in RNAlater. Wingbuds from five individuals were pooled per sample, and samples were frozen at –20 °C after 24 h at 4 °C.

RNA extraction and sequencing for gene expression experiment. RNA extractions were performed using a Trizol protocol. RNA purity was assessed using a NanoDrop ND-1000 spectrophotometer, and quality assessed using an Agilent 2100 Bioanalyzer. Following library preparation with RiboZero, samples were sequenced on an Illumina HiSeq 4000, generating 2 × 150 bp paired-end reads. Due to logistical constraints, library preparation and sequencing were performed separately for samples from the two populations, but using identical protocols; to minimise the effect this could have on our results, we restricted differential expression analyses to within-population comparisons. CASAVA v1.8.2 (Illumina) was used for base calling and demultiplexing indexed reads. Adaptor sequences were trimmed from fastq files using Cutadapt v1.2.1¹⁰⁴ and low-quality bases were removed using Sickle v1.200 with a minimum window quality score of 20. Sequences with high similarity to Eukaryotic ribosomal RNAs were removed from the dataset using sortmeRNA¹⁰⁵.

Differential expression analysis. Genome alignment, transcriptome assembly and quantification of expression were performed following Pertea et al.¹⁰⁶. Reads were aligned to the *T. oceanicus* genome using HiSat2 v2.1.0. A genome-guided transcriptome was then assembled from output files using StringTie v1.3.4, and gene expression values quantified for each of the samples. Differential expression analyses were performed at the 'gene' level. To retain in our transcriptome only transcripts with strong empirical support and which appear to be protein-coding, we filtered any without open reading frames of >100 amino acids, and those which were not expressed at >1 count per million in at least three samples from each population as well as a previously sequenced Kauai wingbud RNA-seq dataset (see below). The expression filter was relaxed to >0.1 count per million when investigating patterns of *prospero* expression (Supplementary Fig. 33). The resulting merged and filtered transcriptome contained 30,299 unigenes. Gene counts produced by StringTie were prepared for input into edgeR v3.20.9¹⁰⁷ using the prepDE.py script made available by the authors¹⁰⁶. In edgeR, counts were normalised by trimmed means of *M*-values (TMM), after which a single negative binomial GLM was fit incorporating all data, using per-transcript normalised expression values as the response variable. Differential expression (DE) analyses were performed using likelihood ratio tests for pairwise comparisons between morph genotypes and sexes. *P*-values were FDR-adjusted using the Benjamini–Hochberg procedure, and transcripts were considered DE between groups if FDR values were < 0.05. Statistical analyses were performed using R v3.4.1. DE transcripts associated with morph genotype in each population were highly correlated between sexes (Supplementary Fig. 31), so males and females with the same genotype were pooled to maximise statistical power: 205 transcripts were identified as DE between genotypes in Kauai, and 464 in Oahu.

Validation of gene expression experiment. As partial validation of our results, we compared the identity of transcripts DE between morph genotypes from Kauai with those found to be DE between male morphs in previously sequenced samples from male Kauai wingbuds⁵². The latter data were collected using methods considerably different to our own: individuals came from different biological lines, were not anaesthetised prior to sampling, and there were fewer individuals per pooled sample and more samples per group ($N = 3$ individuals per pool; $N = 6$ samples per male morph). Samples were also sequenced on a different platform (Illumina HiSeq 2000). Thus, overlap between datasets in transcripts identified as DE would be indicative of biologically meaningful and robust results. We aligned these samples to the transcriptome and performed DE analysis following the same procedure as above. The results showed that of the 30 transcripts identified as DE between Kauai male morphs in the current dataset, 10 (including both those located on scaffold 18404, one of which is *dsx*) were also significantly DE in the previously collected dataset, all in the same direction, with well-correlated \log_2 -fold changes (Spearman's $\rho = 0.648$, $P = 0.049$), supporting the validity of our results.

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

Data availability

Raw re-sequencing reads that support the findings of this study have been deposited in the European Nucleotide Archive with the accession number [PRJEB39125](https://www.ebi.ac.uk/ena/record/PRJEB39125). Raw RNA-seq reads have been deposited in the same archive with the accession number [PRJEB40088](https://www.ebi.ac.uk/ena/record/PRJEB40088). Raw wing morphology data are publicly available at ChirpBase.org [<http://download.chirpbase.org/>]. Previously published RAD-seq²¹ and RNA-seq data⁵² have been deposited in the European Nucleotide Archive under accession numbers [PRJEB29921](https://www.ebi.ac.uk/ena/record/PRJEB29921) and [PRJNA283744](https://www.ebi.ac.uk/ena/record/PRJNA283744), respectively. Source data are provided with this paper.

Code availability

Custom scripts used in this project are archived and publicly available at ChirpBase.org [<http://download.chirpbase.org/>]

Received: 27 May 2020; Accepted: 23 November 2020;

Published online: 04 January 2021

References

- Losos, J. B. Convergence, adaptation, and constraint. *Evolution* **65**, 1827–1840 (2011).
- 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).
- Stern, D. L. The genetic causes of convergent evolution. *Nat. Rev. Genet.* **14**, 751–764 (2013).
- Lee, K. M. & Coop, G. Population genomics perspectives on convergent adaptation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20180236 (2019).
- Ralph, P. L. & Coop, G. Convergent evolution during local adaptation to patchy landscapes. *PLoS Genet.* **11**, e1005630 (2015).
- Ralph, P. & Coop, G. Parallel adaptation: one or many waves of advance of an advantageous allele? *Genetics* **186**, 647–668 (2010).
- Feder, A. F., Pennings, P. S., Hermisson, J. & Petrov, D. A. Evolutionary dynamics in structured populations under strong population genetic forces. *G3* **9**, 3395–3407 (2019).
- Patwa, Z. & Wahl, L. M. The fixation probability of beneficial mutations. *J. R. Soc. Interface* **5**, 1279–1289 (2008).
- Aeschbacher, S. & Burger, R. The effect of linkage on establishment and survival of locally beneficial mutations. *Genetics* **197**, 317–336 (2014).
- Bradburd, G. S. & Ralph, P. L. Spatial population genetics: it's about time. *Annu. Rev. Ecol. Syst.* **50**, 427–449 (2019).
- Alves, J. M. et al. Parallel adaptation of rabbit populations to myxoma virus. *Science* **363**, 1319–1326 (2019).
- Xie, K. T. et al. DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science* **363**, 81–84 (2019).
- Chan, Y. F. et al. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**, 302–305 (2010).
- Olson-Manning, C. F., Wagner, M. R. & Mitchell-Olds, T. Adaptive evolution: evaluating empirical support for theoretical predictions. *Nat. Rev. Genet.* **13**, 867–877 (2012).
- Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **20**, 229–237 (2005).
- Heliconius Genome Consortium. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).
- Skov, L. et al. The nature of Neanderthal introgression revealed by 27,566 Icelandic genomes. *Nature* **582**, 78–83 (2020).
- Whiting, M. F., Bradler, S. & Maxwell, T. Loss and recovery of wings in stick insects. *Nature* **421**, 264–267 (2003).
- Whiting, J. R. & Fraser, B. A. Contingent convergence: the ability to detect convergent genomic evolution is dependent on population size and migration. *G3* **10**, 677–693 (2020).
- Zuk, M., Rotenberg, J. T. & Tinghitella, R. M. Silent night: adaptive disappearance of a sexual signal in a parasitized population of field crickets. *Biol. Lett.* **2**, 521–524 (2006).
- Pascoal, S. et al. Field cricket genome reveals the footprint of recent, abrupt adaptation in the wild. *Evol. Lett.* **4**, 19–33 (2019).
- Zuk, M., Bailey, N. W., Gray, B. & Rotenberg, J. T. Sexual signal loss: The link between behaviour and rapid evolutionary dynamics in a field cricket. *J. Anim. Ecol.* **87**, 623–633 (2018).
- Pascoal, S. et al. Rapid convergent evolution in wild crickets. *Curr. Biol.* **24**, 1369–1374 (2014).
- Tinghitella, R. M., Broder, E. D., Gurule-Small, G. A., Hallagan, C. J. & Wilson, J. D. Purring Crickets: the evolution of a novel sexual signal. *Am. Nat.* **192**, 773–782 (2018).
- Rayner, J. G., Aldridge, S., Montealegre, Z. F. & Bailey, N. W. A silent orchestra: convergent song loss in Hawaiian crickets is repeated, morphologically varied, and widespread. *Ecology* **100**, e02694 (2019).
- Schneider, W. T., Rutz, C., Hedwig, B. & Bailey, N. W. Vestigial singing behaviour persists after the evolutionary loss of song in crickets. *Biol. Lett.* **14**, 20170654 (2018).
- Bailey, N. W., Pascoal, S. & Montealegre, Z. F. Testing the role of trait reversal in evolutionary diversification using song loss in wild crickets. *Proc. Natl Acad. Sci. USA* **116**, 8941–8949 (2019).
- Rayner, J. G., Pascoal, S. & Bailey, N. W. Release from intralocus sexual conflict? Evolved loss of a male sexual trait demasculinizes female gene expression. *Proc. Biol. Sci.* **286**, 20190497 (2019).
- Nouhaid, P., Blanckaert, A., Bank, C. & Kulmuni, J. Understanding admixture: haplodiploidy to the rescue. *Trends Ecol. Evol.* **35**, 34–42 (2020).
- Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
- Anderson, C. J., Tay, W. T., McGaughan, A., Gordon, K. & Walsh, T. K. Population structure and gene flow in the global pest, *Helicoverpa armigera*. *Mol. Ecol.* **25**, 5296–5311 (2016).
- Li, J. et al. Allopatric divergence and hybridization within *Cupressus chengiana* (Cupressaceae), a threatened conifer in the northern Hengduan Mountains of western China. *Mol. Ecol.* **29**, 1250–1266 (2020).
- Qiu, Q. et al. Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nat. Commun.* **6**, 10283 (2015).
- Soraggi, S., Wiuf, C. & Albrechtsen, A. Powerful inference with the D-statistic on low-coverage whole-genome data. *G3* **8**, 551–566 (2018).
- Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V. C. & Foll, M. Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9**, e1003905 (2013).
- Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
- Timmermann, A., Justino, F., Jin, F. F., Krebs, U. & Gosses, H. Surface temperature control in the North and tropical Pacific during the last glacial maximum. *Clim. Dynam.* **23**, 353–370 (2004).
- Tinghitella, R. M., Zuk, M., Beveridge, M. & Simmons, L. W. Island hopping introduces Polynesian field crickets to novel environments, genetic bottlenecks and rapid evolution. *J. Evol. Biol.* **24**, 1199–1211 (2011).
- Pascoal, S. et al. Sexual selection and population divergence I: The influence of socially flexible cuticular hydrocarbon expression in male field crickets (*Teleogryllus oceanicus*). *Evolution* **70**, 82–97 (2016).
- Norris, L. C. et al. Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proc. Natl Acad. Sci. USA* **112**, 815–820 (2015).
- Clarkson, C. S. et al. Adaptive introgression between *Anopheles* sibling species eliminates a major genomic island but not reproductive isolation. *Nat. Commun.* **5**, 4248 (2014).
- Ai, H. et al. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat. Genet.* **47**, 217–225 (2015).
- Zhou, J., Lemos, B., Dopman, E. B. & Hartl, D. L. Copy-number variation: the balance between gene dosage and expression in *Drosophila melanogaster*. *Genome Biol. Evol.* **3**, 1014–1024 (2011).
- Massouras, A. et al. Genomic variation and its impact on gene expression in *Drosophila melanogaster*. *PLoS Genet.* **8**, e1003055 (2012).
- Liu, Z. et al. Genomic mechanisms of physiological and morphological adaptations of limestone Langurs to Karst Habitats. *Mol. Biol. Evol.* **37**, 952–968 (2019).

46. Guenin, L., Raharjaona, M., Houlgatte, R. & Baba-Aissa, F. Expression profiling of prospero in the *Drosophila* larval chemosensory organ: Between growth and outgrowth. *BMC Genomics* **11**, 47 (2010).
47. Kim, S., Chung, S., Yoon, J., Choi, K. W. & Yim, J. Ectopic expression of Toll/ Toll-8 antagonizes Dpp signaling and induces cell sorting in the *Drosophila* wing. *Genesis* **44**, 541–549 (2006).
48. Rideout, E. J., Dornan, A. J., Neville, M. C., Eadie, S. & Goodwin, S. F. Control of sexual differentiation and behavior by the doublesex gene in *Drosophila melanogaster*. *Nat. Neurosci.* **13**, 458–466 (2010).
49. Kijimoto, T., Moczek, A. P. & Andrews, J. Diversification of doublesex function underlies morph-, sex-, and species-specific development of beetle horns. *Proc. Natl Acad. Sci. USA* **109**, 20526–20531 (2012).
50. Kunte, K. et al. doublesex is a mimicry supergene. *Nature* **507**, 229–232 (2014).
51. Snodgrass, R. E. *Principles of Insect Morphology* (Cornell University Press, 1993).
52. Pascoal, S. et al. Rapid evolution and gene expression: a rapidly evolving Mendelian trait that silences field crickets has widespread effects on mRNA and protein expression. *J. Evol. Biol.* **29**, 1234–1246 (2016).
53. Price, D. C., Egizi, A. & Fonseca, D. M. The ubiquity and ancestry of insect doublesex. *Sci. Rep.* **5**, 13068 (2015).
54. Charlesworth, D., Barton, N. H. & Charlesworth, B. The sources of adaptive variation. *Proc. Biol. Sci.* **284**, 20162864 (2017).
55. Corl, A. et al. The genetic basis of adaptation following plastic changes in coloration in a novel environment. *Curr. Biol.* **28**, 2970–2977.e2977 (2018).
56. Sprengelmeyer, Q. D. et al. Recurrent collection of *Drosophila melanogaster* from wild African environments and genomic insights into species history. *Mol. Biol. Evol.* **37**, 627–638 (2020).
57. Lujan, N. K., Weir, J. T., Noonan, B. P., Lovejoy, N. R. & Mandrak, N. E. Is Niagara Falls a barrier to gene flow in riverine fishes? A test using genome-wide SNP data from seven native species. *Mol. Ecol.* **29**, 1235–1249 (2020).
58. Song, Y. et al. Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Curr. Biol.* **21**, 1296–1301 (2011).
59. Edelman, N. B. et al. Genomic architecture and introgression shape a butterfly radiation. *Science* **366**, 594 (2019).
60. Pearce, R. J. et al. Multiple origins and regional dispersal of resistant dhps in African *Plasmodium falciparum* malaria. *PLoS Med.* **6**, e1000055 (2009).
61. Kupper, C. et al. A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83 (2016).
62. Alexander, R. D. Evolutionary change in Cricket acoustical communication. *Evolution* **16**, 443–44 (1962).
63. Ragge, D. R. *The Wing-Venation of the Orthoptera Saltatoria: With Notes on Dictyopteran Wing-venation* (British Museum, Natural History, London, 1955).
64. Gwynne, D. T. Phylogeny of the Ensifera (Orthoptera): a hypothesis supporting multiple origins of acoustical signalling, complex spermatophores and maternal care in Crickets, Katydid, and Weta. *J. Orthop. Res.* **4**, 203–218 (1995).
65. Desutter-Grandcolas, L. & Robillard, T. Phylogeny and the evolution of calling songs in Gryllus (Insecta, Orthoptera, Gryllidae). *Zool. Scr.* **32**, 173–183 (2003).
66. Wiens, J. J. Widespread loss of sexually selected traits: how the peacock lost its spots. *Trends Ecol. Evol.* **16**, 517–523 (2001).
67. Ledon-Rettig, C. C., Zattara, E. E. & Moczek, A. P. Asymmetric interactions between doublesex and tissue- and sex-specific target genes mediate sexual dimorphism in beetles. *Nat. Commun.* **8**, 14593 (2017).
68. Ito, Y. et al. The role of doublesex in the evolution of exaggerated horns in the Japanese rhinoceros beetle. *EMBO Rep.* **14**, 561–567 (2013).
69. Palmer, D. H. & Kronforst, M. R. A shared genetic basis of mimicry across swallowtail butterflies points to ancestral co-option of doublesex. *Nat. Commun.* **11**, 6 (2020).
70. Komata, S., Lin, C. P., Iijima, T., Fujiwara, H. & Sota, T. Identification of doublesex alleles associated with the female-limited Batesian mimicry polymorphism in *Papilio memnon*. *Sci. Rep.* **6**, 34782 (2016).
71. Prakash, A. & Monteiro, A. Doublesex mediates the development of sex-specific pheromone organs in *Bicyclus* butterflies via multiple mechanisms. *Mol. Biol. Evol.* **37**, 1694–1707 (2020).
72. Hoffmann, A. A. & Sgro, C. M. Climate change and evolutionary adaptation. *Nature* **470**, 479–485 (2011).
73. Rohlf, F. J. The tps series of software. *Hystrix* **26**, 9–12 (2015).
74. Klingenberg, C. P. Morpho: an integrated software package for geometric morphometrics. *Mol. Ecol. Resour.* **11**, 353–357 (2011).
75. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
76. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
77. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
78. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11 10 11–11 10 33 (2013).
79. Kim, K. W. et al. Genetics and evidence for balancing selection of a sex-linked colour polymorphism in a songbird. *Nat. Commun.* **10**, 1852 (2019).
80. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
81. Retief, J. D. Phylogenetic analysis using PHYLIP. *Methods Mol. Biol.* **132**, 243–258 (2000).
82. Kumar, S., Stecher, G., Li, M., Nkay, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
83. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
84. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis. Use R*, 1–212 (Springer, 2009).
85. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
86. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
87. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
88. Korneliusen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
89. Lischer, H. E. & Excoffier, L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298–299 (2012).
90. Excoffier, L., Laval, G. & Schneider, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* **1**, 47–50 (2007).
91. Keightley, P. D. et al. Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Res.* **19**, 1195–1201 (2009).
92. Mousseau, T. A. & Roff, D. A. Genetic and environmental contributions to geographic-variation in the ovipositor length of a cricket. *Ecology* **76**, 1473–1482 (1995).
93. Chen, K. et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* **6**, 677–681 (2009).
94. Wang, K. et al. Genome-wide variation within and between wild and domestic yak. *Mol. Ecol. Resour.* **14**, 794–801 (2014).
95. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011).
96. Lou, H. et al. Copy number variations and genetic admixtures in three Xinjiang ethnic minority groups. *Eur. J. Hum. Genet.* **23**, 536–542 (2015).
97. Zhang, X. et al. Genome-wide patterns of copy number variation in the Chinese yak genome. *BMC Genomics* **17**, 379 (2016).
98. Redon, R. et al. Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
99. Pezer, Z., Harr, B., Teschke, M., Babiker, H. & Tautz, D. Divergence patterns of genic copy number variation in natural populations of the house mouse (*Mus musculus domesticus*) reveal three conserved genes with major population-specific expansions. *Genome Res.* **25**, 1114–1124 (2015).
100. Nei, M. & Roychoudhury, A. K. Gene differences between Caucasian, Negro, and Japanese populations. *Science* **177**, 434–436 (1972).
101. Wakeley, J. The variance of pairwise nucleotide differences in two populations with migration. *Theor. Popul. Biol.* **49**, 39–57 (1996).
102. Pfeifer, B., Wittelsburger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936 (2014).
103. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **28**, 45–48 (2000).
104. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
105. Kopylova, E., Noe, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
106. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).
107. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

Acknowledgements

We are grateful to W.V. Bailey, B. Gray, J. Rotenberry, W.T. Schneider and M. Zuk for assistance with cricket sampling. M.G. Ritchie and D.M. Zonana provided helpful comments on an early version of the manuscript. J. Santoyo-Lopez and P. Koldkjær assisted with next-generation sequencing at Edinburgh Genomics and the Centre for Genomic Research, (University of Liverpool), respectively. P. Thorpe provided valuable bioinformatics assistance, and T. Sneddon and N. Cook provided valuable wet lab assistance. The work was funded by Natural Environment Research Council awards to N.W.B. [NE/I027800/1, NE/L011255/1]. Bioinformatics support was provided by a Wellcome Trust ISSF award [105621/Z/14/Z]. X.Z. was supported by a China Scholarship Council PhD studentship [201703780018].

Author contributions

Conceived the project: X.Z. and N.W.B. Designed experiments: X.Z., M.B., J.G.R. and N.W.B. Collected data: X.Z., J.G.R. and N.W.B. Analysed data: X.Z. and J.G.R. Wrote the manuscript: X.Z., J.G.R. and N.W.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-20263-4>.

Correspondence and requests for materials should be addressed to X.Z. or N.W.B.

Peer review information *Nature Communications* thanks Nicola Nadeau, and the other, anonymous, reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021