nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0581-8

Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation

David A. Marques[1,2,3*], Felicity C. Jones[4,5], Federica Di Palma[6,7], David M. Kingsley[4] and Thomas E. Reimchen[1]

A substantial part of biodiversity is thought to have arisen from adaptive radiations in which one lineage rapidly diversified into multiple lineages specialized to many different niches. However, selection and drift reduce genetic variation during adaptation to new niches and may thus prevent or slow down further niche shifts. We tested whether rapid adaptation is still possible from a highly derived ecotype in the adaptive radiation of threespine stickleback on the Haida Gwaii archipelago, Western Canada. In a 19-year selection experiment, we let giant sticklebacks from a large blackwater lake evolve in a small clearwater pond without vertebrate predators. A total of 56 whole genomes from the experiment and 26 natural populations revealed that adaptive genomic change was rapid in many small genomic regions and encompassed 75% of the change between 12,000-year-old ecotypes. Genomic change was as fast as phenotypic change in defence and trophic morphology, and both were largely parallel between the short-term selection experiment and long-term natural adaptive radiation. Our results show that functionally relevant standing genetic variation can persist in derived radiation members, allowing adaptive radiations to unfold very rapidly.

The colonization of a new habitat or niche requires rapid adaptation to multiple environmental challenges (that is, to 'multifarious' divergent selection). This is most dramatic in adaptive radiations, where rapid successions of niche and habitat shifts occur within a lineage[1–3]. However, most adaptive radiations started thousands of generations ago and we do not know whether major phenotypic and genomic adaptation occurred within the first few generations of colonizing a new habitat, or over longer time scales, and thus how 'rapid' adaptive radiations unfold. Adaptation may be instantaneous when phenotypic plasticity is involved[4,5] or occur over few generations of selection on standing genetic variation[6,7] or admixture variation[8,9]. Alternatively, adaptation may require time for beneficial de novo mutations to arrive, or genomic adaptation may occur slower than phenotypic adaptation if rapid phenotypic plasticity is followed by slower genetic assimilation[4,10]. Furthermore, each new habitat shift will reduce genetic variation through drift and selection and it is unclear whether further adaptation is hampered or slowed down after a first new niche has been colonized in an adaptive radiation.

Evolution experiments and cases of contemporary evolution, such as in biological invasions, may reveal the speed of phenotypic and genomic adaptation[11,12]. However, many 'evolve and resequence' experiments and contemporary evolution studies focused on single selective agents instead of multifarious fitness landscapes[13–21], or phenotypic and genomic adaptation have been studied in isolation[22–25]. Only few examples of phenotypic and genomic contemporary evolution under multifarious divergent selection have been documented, such as in marine threespine stickleback (*Gasterosteus aculeatus*) colonizing freshwater habitats in artificial and natural selection experiments[25–28], showing widespread parallel genomic and phenotypic adaptation compared with thousands-of-generations-older natural populations[7,29].

Here, we quantify the speed of genomic adaptation to multifarious divergent selection in a 19-year selection experiment, starting from a phenotypically highly derived adaptive radiation member, and compare rates of phenotypic and genomic change. We expand on a long-term investigation of the adaptive radiation of threespine stickleback from the Haida Gwaii archipelago off Western Canada[30], where stickleback have colonized multiple watersheds independently and adapted to diverse freshwater habitats, including lakes, ponds and streams with vastly divergent biophysical features, predator and parasite communities, following glacial retreat ~12,000 years ago[31–34]. Phenotypic variation in defensive armour[35–39], such as dorsal and pelvic spines, pelvic girdle and lateral plates, and in trophic morphology[31,39,40], such as body shape, gape and gill rakers, can largely be explained by three main predictors: predation regime, light spectrum and lake size[30].

A selection experiment along these three axes of selection was initiated by T.E.R. in 1993: he transplanted 100 adult stickleback from a large, deep, dystrophic, blackwater lake (Mayer Lake) with vertebrate-dominated predation into a small, shallow, eutrophic, previously unoccupied clearwater pond (Roadside Pond) dominated by invertebrate predators[41]. Mayer Lake contains some of the most derived freshwater stickleback, maximally divergent from the ancestral marine phenotype and occupying the extreme morphospace edge of the Haida Gwaii adaptive radiation[30]. These are 8–10-cm-long, melanistic 'giants' with highly developed predator defence morphology and adaptations to limnetic foraging[42,43], and low levels of phenotypic variance[31], but similar levels of genetic variation to other Haida Gwaii populations[34,44]. After evolving for 16 years in the new selective regime, six predator defence traits, four feeding morphology traits and eye size evolved in the expected direction (Fig. 1a), encompassing ~30% of the morphological distance between natural stickleback populations from large lakes and

[1]Department of Biology, University of Victoria, Victoria, British Columbia, Canada. [2]Aquatic Ecology & Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland. [3]Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland. [4]Department of Developmental Biology, HHMI and Stanford University School of Medicine, Stanford, CA, USA. [5]Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany. [6]Earlham Institute, Norwich Research Park, Norwich, UK. [7]Department of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, UK. *e-mail: david.marques@eawag.ch

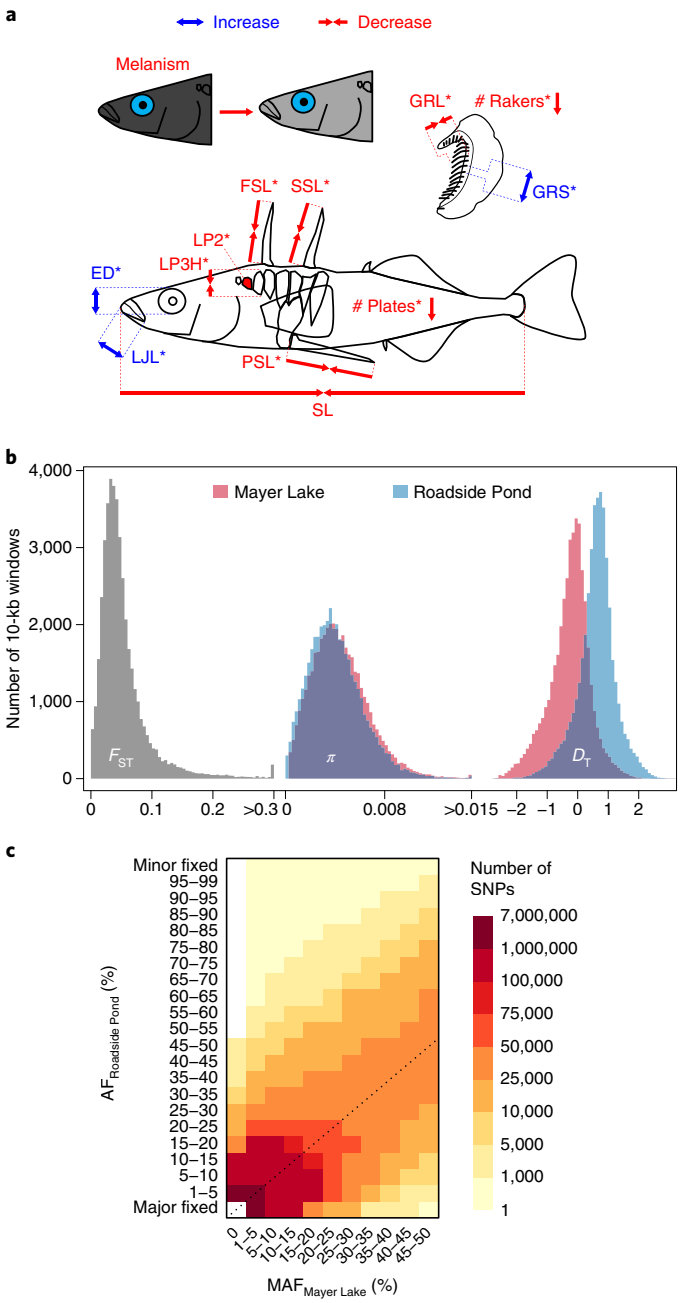


Fig. 1 | Phenotypic and genomic change in the selection experiment. a, Summary of the phenotypic changes observed in the selection experiment, as reported in Leaver and Reimchen[41], with colours indicating a trait increase or decrease, and asterisks indicating significant change. Phenotypic change in six bony predator defence traits (FSL, first dorsal spine length; SSL, second dorsal spine length; PSL, pelvic spine length; # Plates, number of lateral plates; LP3H, lateral plate 3 height; and LP2, lateral plate 2 frequency), four feeding morphology traits (LJL, lower jaw length; # Rakers, number of gill rakers; GRL, gill raker length; and GRS, gill raker spacing) and eye diameter (ED) was in the expected direction (that is, parallel), given the shift from vertebrate- to invertebrate-dominated predation and zooplankton- to invertebrate-dominated diet, and observed phenotypic divergence between large lake and small pond populations in the adaptive radiation on Haida Gwaii[41]. SL, standard length. **b**, Transplant of 100 adult giant threespine stickleback from Mayer Lake into Roadside Pond and evolution for 13 generations led to moderate genomic differentiation (F_{ST}), a minor reduction in nucleotide diversity (π) and a positive shift in the Tajima's D (D_T) distribution. **c** Even though several rare alleles were fixed, allele frequencies (AF) did not change much over 13 generations. The dotted line indicates no change. MAF, minor allele frequency.

small ponds[41] (Fig. 2a). Life history changed from two years to one year for the age of first reproduction, and melanism was reduced[41,45]. Phenotypic evolution was fast with, on average, 0.15 (0–0.25) haldanes over 11 generations, assuming an average generation time of 1.5 years[41]. While strong change in the first generation for four traits suggested phenotypic plasticity, other traits showed slower change, suggesting genetic change. We used whole genomes from 26 natural populations, including the source, Mayer Lake ($n=12$) and the transplant population, Roadside Pond ($n=11$), sampled after evolving for 19 years (or 13 generations) in the new habitat, to identify the speed and targets of genomic adaptation and the extent of genomic parallelism with the Haida Gwaii adaptive radiation.

Results

Moderate genome-wide change, but strong change in many small genomic regions. Giant stickleback evolving for 13 generations in a new habitat showed only moderate genome-wide change (Fig. 1b,c), but strong change in many small genomic regions (Figs. 3 and 4). Allele frequencies (AF) changed on average by 11.4% (weighted mean $|\Delta AF|$; Fig. 1c), leading to a genomic 'background' differentiation between Mayer Lake and Roadside Pond of $F_{ST}=0.057$ for autosomes and $F_{ST}=0.107$ for the female sex chromosome (weighted pairwise F_{ST}; Fig. 1b). Compared with the differentiation observed between natural, postglacial populations, 13 generations of evolution encompassed 41% of the differentiation between lake and stream ecotypes and 22% of the differentiation between stickleback from the large Mayer Lake and three independently colonized small ponds on Haida Gwaii (Fig. 2c). Similarly, mean pairwise divergence (D_{XY}), reflecting the sorting of polymorphic, ancient divergent genomic regions between populations on these short time scales, increased marginally ($D_{XY,\text{within populations}}=0.0037$; $D_{XY,\text{between Mayer–Roadside}}=0.0038$; $t_{183}=-2.88$, $P=0.004$) and encompassed 20% of the divergence between naturally occurring lake and stream ecotypes and 9% of the divergence between Mayer Lake and small ponds (Fig. 2b and Supplementary Results).

The transplant of giant Mayer Lake stickleback to Roadside Pond led to a slight loss of genetic diversity and a prominent, genome-wide distortion of the site frequency spectrum (SFS) (Fig. 1b and Supplementary Fig. 2a). Mean nucleotide diversity was reduced by 7.4% from $\pi_{Mayer}=0.0047$ to $\pi_{Roadside}=0.0043$ (mean 10-kilobase (kb) windows, $t_{85940}=22.56$, $P<0.001$; Fig. 1b), but local diversity across the genome between both populations remained strongly correlated (Pearson's $r=0.92$, linear regression $F_{2,43444}=89{,}485$, $P<0.001$; Supplementary Fig. 3). The distribution of Tajima's D was shifted to a positive mean from $D_{T,Mayer}=-0.27$ to $D_{T,Roadside}=0.60$ (mean 10-kb windows, $t_{85623}=-197.11$, $P<0.001$; Fig. 1b), indicating the loss of rare alleles relative to common alleles, as is evident from the observed one-dimensional SFS (Supplementary Fig. 2a), but Tajima's D remained correlated across the genome (Pearson's $r=0.53$, linear regression $F_{2,43444}=17{,}277$, $P<0.001$; Supplementary Fig. 3).

Genomic footprints of divergent selection are widespread in the genome. Genome-wide average changes, such as diversity loss and a shifted Tajima's D distribution, are probably a product of demographic history, while localized changes in the genome may reflect footprints of selection. To distinguish between the two, we reconstructed the demographic history of Mayer Lake and Roadside Pond stickleback (see Methods). The demographic model best fitting the observed two-dimensional SFS (2D-SFS) features a bottleneck ~8,200 generations ago (translating to ~12,300 years), in line with the postglacial colonization of Mayer Lake (Supplementary Fig. 2). It also recovered the observed population growth of the Roadside Pond population following the transplant[41]. We identified signatures of divergent selection between Mayer Lake and Roadside Pond in the genome from outliers for differentiation (F_{ST}), change in diversity ($\Delta\pi$) or Tajima's D (ΔD_T), as well as haplotype-based

selection statistics integrated haplotype score (iHS) and cross-population extended haplotype homozygosity (XPEHH) against neutral expectations from demographic history by simulating genomic data under the best-fitting demographic model (Supplementary Fig. 3; see Methods). The simulations reproduced both the observed diversity loss and the positive shift in Tajima's *D* (Supplementary Fig. 4).

Traces of divergent selection among the habitats are widespread across the genome: we found 77 outlier regions distributed across 15 chromosomes, covering 15.73 megabases (Mb) or 3.6% of the genome (Figs. 3 and 4, Supplementary Figs. 5–19 and Supplementary Table 1), exceeding expectations from simulated neutral genomic data (0.16–0.25% of the genome). Outlier regions varied in size between 30 and 940 kb (mean = 204 kb, median = 160 kb). Three quarters of outlier regions show patterns consistent with a near-complete, past selective sweep in Mayer Lake, followed by a quick rise of the previously disfavoured allele to high or intermediate frequency in Roadside Pond, indicated by a negative Tajima's *D* in Mayer Lake, a positive Tajima's *D* in Roadside Pond, a negative XPEHH, significant differentiation and exceptional allele frequency shifts between the populations (Fig. 4 and Supplementary Figs. 5–19). The remaining quarter of outlier regions show an opposite pattern, consistent with a selective sweep in Roadside Pond but not in Mayer Lake, indicated by reduced Tajima's *D* and diversity in Roadside Pond, significant iHS and increased H12, another haplotype-based statistic, for Roadside Pond and a positive XPEHH (Supplementary Figs. 6–19). Both patterns are in agreement with divergent selection between the habitats in the experiment.

We computed linkage disequilibrium between outlier regions to test whether divergent selection acted on a single genomic region with others hitchhiking, or whether multiple regions responded independently to divergent selection. Significant interchromosomal linkage disequilibrium was found mainly between outlier regions in the Mayer Lake population (Supplementary Fig. 20), indicating that several regions involved in divergent selection in the experiment are not segregating fully independently in the source population. However, in the transplant population Roadside Pond, we found only seven significant interchromosomal associations (Supplementary Fig. 20), suggesting that while outlier regions on different chromosomes responded largely independently to divergent selection in the experiment, some uncertainty remains about the exact number of independently selected regions.

Adaptive differentiation, defined as the top 5% of single nucleotide polymorphism (SNP) F_{ST} estimates from each genomic outlier region, ranged from $F_{ST} = 0.25$ to $F_{ST} = 0.76$, with a mean of $F_{ST} = 0.44$. Allele frequency change at these SNPs ranged from $|\Delta AF| = 18\%$ to $|\Delta AF| = 81\%$, with a mean of $|\Delta AF| = 51\%$, which under a model of purely selection-driven change would correspond to selection coefficients between $s = 0.24$ and $s = 1$, with mean of $s = 0.62$. Compared with naturally evolved, postglacial ecotypes, genomic adaptive differentiation after only 13 generations of evolution in a new 'ecological theatre' thus encompassed 72% of the degree of adaptive differentiation found between postglacial lake and stream ecotypes[46], and already exceeds adaptive differentiation found between giant Mayer Lake stickleback and their corresponding parapatric stream ecotype (Fig. 2d).

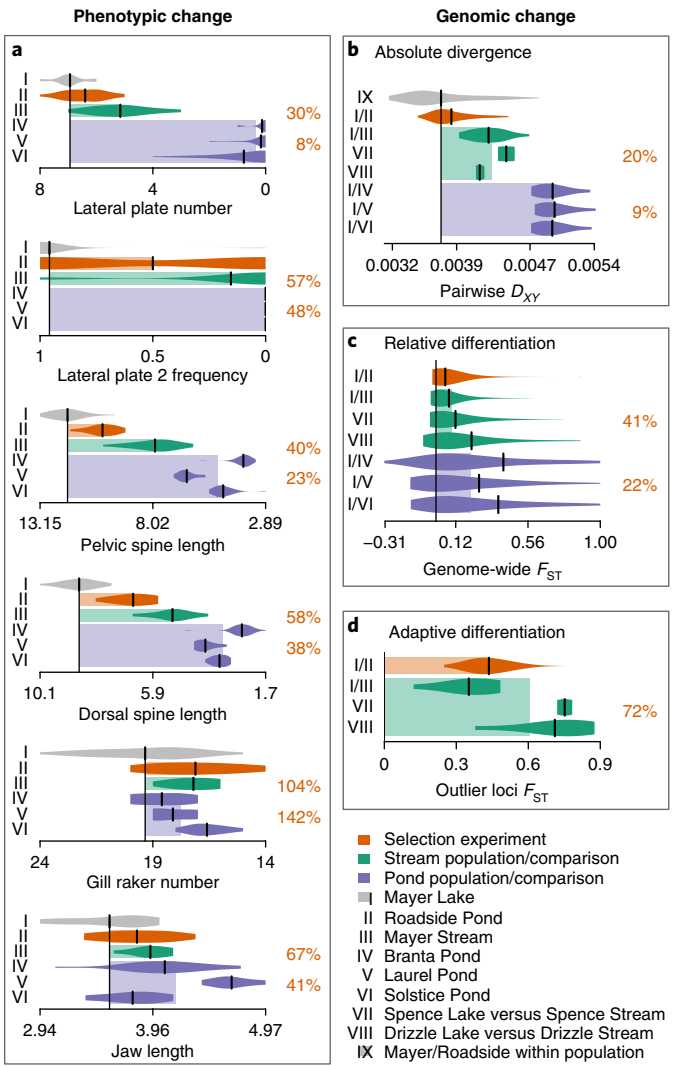


Fig. 2 | Extent of phenotypic and genomic evolution in the 19-year selection experiment compared with the ~12,000-year-old adaptive radiation. Phenotypic divergence and adaptive genomic differentiation arose rapidly in the selection experiment, comparable in extent to postglacial divergence between large lake and pond or stream ecotypes. **a**, Population means and distributions for six phenotypic traits in Mayer Lake (source population, grey), Roadside Pond (transplant population, orange), and postglacial stream (green) and pond (blue) ecotype populations[41]. **b**–**d**, Absolute divergence (**b**), relative differentiation (**c**) and adaptive differentiation (**d**) between the populations listed in **a**, plus those identified in the key. F_{ST} estimates for lake versus stream and large lake versus small pond comparisons are based on SNP chip data from a previous study[46]. Adaptive differentiation SNPs are outlier SNPs from a previous study[46] and the top 5% F_{ST} SNPs in each outlier window for the selection experiment. Values to the right are the mean percentages of phenotypic (**a**) or genomic change (**b**–**d**) in the selection experiment compared with postglacial lake versus stream divergence (upper value) and large lake versus small pond divergence (lower value).

Genomic targets and parallel evolution in experiment and adaptive radiation. We identified potential targets and sources of divergent selection from overlapping genes and quantitative trait loci (QTL)[47] and from genotype–environment and genotype–phenotype (GE/GP) associations in the Haida Gwaii adaptive radiation. In the case of GE/GP associations, we tested whether genomic variation in each outlier region was associated with a change in phenotypic and ecological properties in the selection experiment and across 1 marine and 25 freshwater populations on Haida Gwaii (see Methods and Fig. 5). We found 654 QTL overlapping with outlier regions and 336 candidate genes near the centre of each outlier regions' selective sweep signature, but no Gene Ontology term enrichment (Figs. 4 and 5, Supplementary Figs. 5–19 and Supplementary Tables 2 and 3). Some 36 outlier regions showed parallel GE/GP associations (Fig. 5).

In line with the predation landscape being the most important axis of divergent selection in the adaptive radiation[30], 96 QTL and

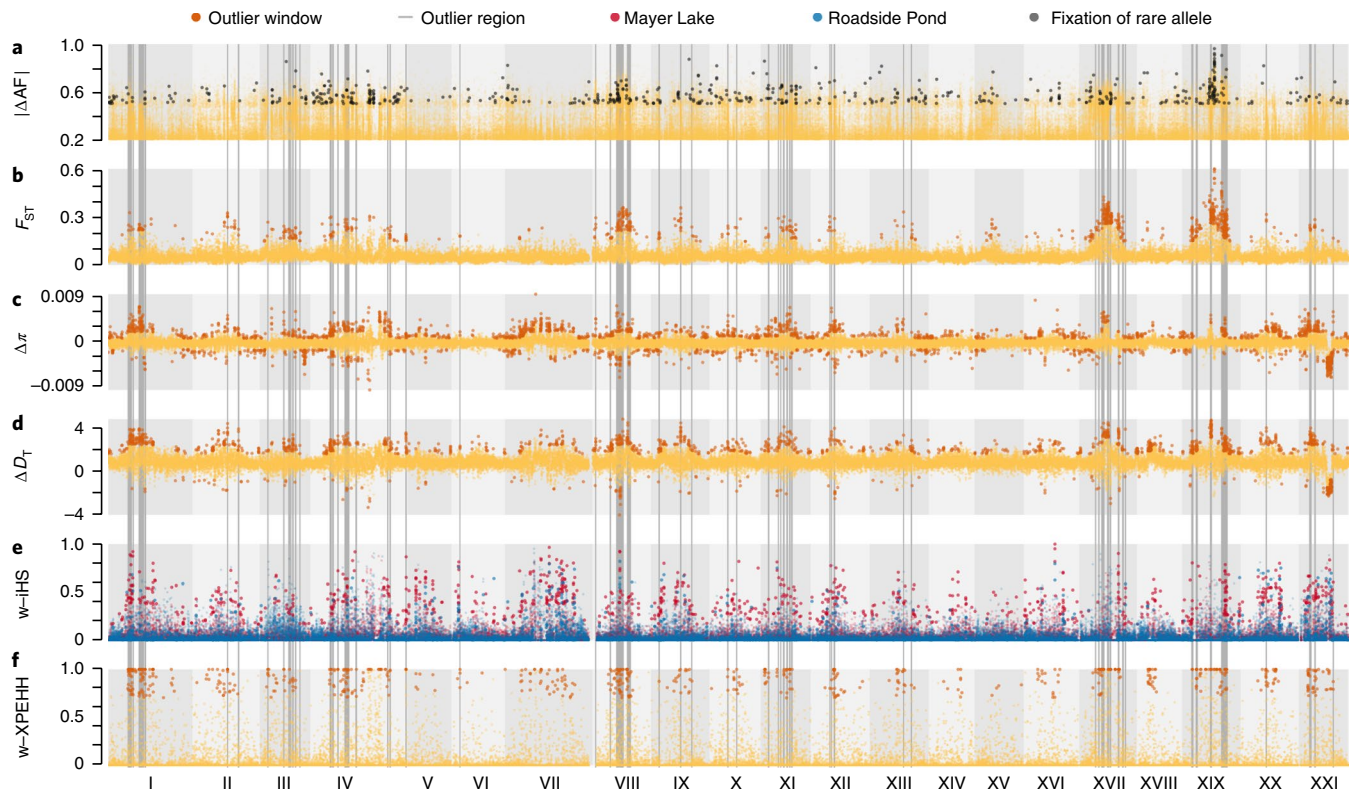


Fig. 3 | Genomic footprints of divergent selection are widespread across the genome. a, Absolute allele frequency change (|ΔAF|) at the top 0.1% strongest |ΔAF| SNPs, with black points highlighting SNPs for which the rarer allele in Mayer Lake went to fixation in Roadside Pond. **b**, Differentiation (F_{ST}). **c**, Change in diversity ($\Delta\pi$). **d**, Tajima's *D* (ΔD_T). **e**,**f**, Haplotype-based selection statistics iHS (**e**) and XPEHH (**f**). The overlapping top 1% outliers among 10-kb windows for statistics **b**–**f** against neutral demographic expectations are highlighted with larger, darker points (see Methods). Grey vertical bars highlight 77 outlier regions with overlapping outliers for statistics **b**, **c**/**d** and **e**/**f** (roman numerals).

one candidate gene controlling predator defence traits overlap with outlier regions. Among these are major-effect QTL for lateral plate number, dorsal spine, pelvic spine and pelvic girdle length, and many intermediate- and minor-effect QTL for these traits on additional chromosomes (Fig. 5 and Supplementary Table 2). Remarkably, phenotypic variation in pelvic and dorsal spine length across the Haida Gwaii radiation is associated with genomic variation in 11 and 6 outlier regions, respectively, and the phenotypic and genomic change observed in the selection experiment paralleled the adaptive radiation in 10 (pelvic spine) and 4 (dorsal spine) of these regions (Fig. 5). Variation in plate number across the radiation is associated with genomic variation in outlier regions IV.i and XVII.f (parallel) and XII.c (non-parallel). While most annotations for candidate genes did not allow us to draw conclusions on defence phenotypes, the *EDA* gene in outlier region IV.e controls lateral plate number[48–50] and several associated traits, such as lateral line pattern and schooling behaviour[51,52] (Supplementary Fig. 9). Genetic variation in the *EDA* region may be responsible for the observed reduction of plate number in the selection experiment[41], while the lack of an association across the adaptive radiation suggests an involvement of different alleles or genes in other populations.

Variation in the light spectrum across the adaptive radiation, the second most important axis of divergent selection[30], and the presence of blackwater show many strong, parallel associations with genetic variation in outlier regions also containing multiple candidate genes involved in (colour) vision (Fig. 5). The six outlier regions most strongly associated with the light spectrum in parallel between adaptive radiation and the selection experiment contain the following genes: *OPN1SW1*, which encodes a photoreceptor sensitive to ultraviolet light[53]; *TRPC7*, which is involved in eye physiology[54]; *CACNA2D3*, which is associated with night blindness in humans[55]; *ATP6V1F*, which is involved in retinal pigmentation[56,57]; *DENND6B* and *CERS2A*, which are nervous system development genes expressed in the eye, lens and retina; and *ADAMTS10*, which is involved in lens development[58]. In addition, outlier region XVII.j shows a parallel association with blackwater habitats and contains the blue-light-sensitive photoreceptor *OPN1SW2*, which we demonstrated previously to be under selection between blackwater and clearwater habitats across both adaptive radiation and the selection experiment[59]. Repeated adaptation to the light spectrum has therefore led to multiple signatures of parallel adaptation on visual perception genes.

Many feeding morphology QTL ($n = 277$) overlap with outlier regions. Jaw length variation in the adaptive radiation is associated with six outlier regions on five chromosomes – four in parallel and two non-parallel (Fig. 5). QTL for gill raker number overlap with many outlier regions, with two parallel associations across the adaptive radiation on chromosomes I and VIII (Fig. 5 and Supplementary Table 2). Similarly, many gill raker length QTL overlap with outlier regions; in particular, intermediate-effect loci on chromosomes IV and VIII. Also for gill raker spacing, intermediate-effect QTL on chromosomes IV and XX, as well as some minor-effect QTL, overlap with outlier regions. The widespread genomic architecture of many feeding-morphology-related traits thus broadly overlaps with genomic regions under divergent selection. Together with several candidate genes involved in craniofacial development or various metabolic processes (Supplementary Table 3), these overlaps might reflect selection on feeding morphology or diet after the shift from a zooplankton-dominated to benthic-invertebrate-dominated habitat.

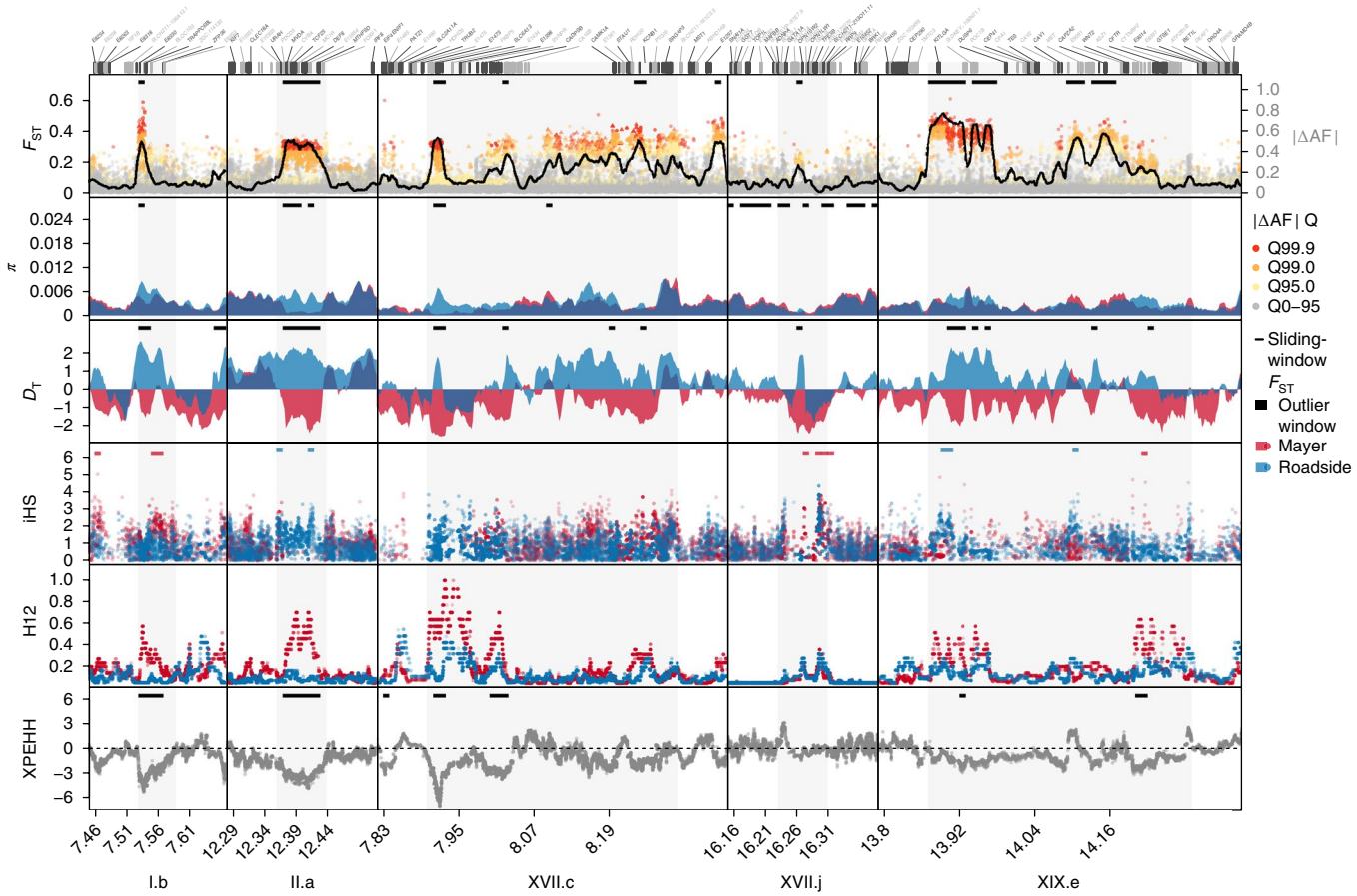


Fig. 4 | Local signatures of divergent selection in the genome. For 5 of the 77 outlier regions in the genome (grey shading), patterns of differentiation (F_{ST}), allele frequency change (|ΔAF|), nucleotide diversity (π), Tajima's D (D_T), iHS, H12 and XPEHH are shown. Outlier region I.b is centred on two genes, *TRAPPC6BL* and *BLOC1S3*, controlling pigmentation in the retina; II.a contains the pigmentation gene *MC1R*; and XVII.j contains the blue-sensitive colour vision gene *OPNSW2*, all of which are probably targets of divergent selection on pigmentation and visual perception. See Supplementary Figs. 5–19 for details of further outlier regions. The top panel colour code indicates quantiles (Q) in the allele frequency change (|ΔAF|) distribution. Black horizontal bars show significant non-overlapping 10-kb outlier windows against neutral expectations. Gene exons are labelled at the top, using shortened gene names (for example, '*ENSG0000000012345*' has been shortened to '*E12345*'). Genomic coordinates (*x* axis) refer to an improved version of the reference genome[80]. Lines are sliding window estimates for 10-kb windows with 2.5-kb step size. Dots represent single SNP estimates (|ΔAF|, iHS and XPEHH) or 81-SNP windows (H12).

Strikingly, we found several candidate genes (*KITLGA*, *MC1R*, *GPR25*, *FOXD3*, *BNC2*, *SYTL2*, *TCTA*, *TRAPPC6BL*, *BLOC1S3*, *MYRIP* and *ATP6V1F*) in outlier regions involved in pigmentation[57,60–65]. These might be associated with the decreased melanism found in Roadside Pond stickleback, the loss of red nuptial coloration or sexual selection for blue signals in blackwater Mayer Lake[45]. *KITLGA* and *MC1R* are well-known targets of divergent selection associated with melanin pigmentation in stickleback, mammals and reptiles[61,65,66]. *KITLGA* overlaps with a large-effect QTL for black pigmentation in other wild stickleback populations[65]. *SYTL2* and *BLOC1S3* overlap with intermediate-effect QTL for red pigmentation[67]. Outlier regions containing these genes do not show associations with melanism across the Haida Gwaii radiation, nor do melanism-associated regions across the adaptive radiation contain known pigmentation genes (Fig. 5). Finally, body size shows a non-parallel association with outlier region XIX.e, containing an intermediate-effect QTL for body size (Supplementary Table 2), while additional, parallel associations do not overlap with known QTL or gene function. Many outlier regions are associated with several ecological and phenotypic traits, other QTL and geography (Fig. 5). This might reflect modular trait architecture[47,68], selection-driven clustering of adaptive variation or indirect associations with correlated ecological and phenotypic variables across the adaptive radiation.

Discussion

Adaptation of a highly derived threespine stickleback ecotype in the Haida Gwaii adaptive radiation to the opposite multifarious selection regime along the 3 major axes of selection resulted in rapid and widespread adaptive genomic change over only 13 generations. We previously demonstrated rapid phenotypic change in the selection experiment, encompassing approximately 30% of the phenotypic divergence of 12,000-year-old natural populations[41] (Figs. 1a and 2a). Here, we found that underlying adaptive genomic variation also responded very quickly to divergent selection, with on average 72% of the expected change occurring in the first 13 generations. The evolutionary rate of genomic adaptation is thus very high and comparable in speed to the contemporary evolution of *Brassica rapa* adapting to drought over 7 generations[18], showing adaptive differentiation of $F_{ST} = 0.17–0.44$, or to Darwin's finches undergoing drought-induced ecological character displacement[69] with a selection coefficient of 0.59 on a major-effect locus. Genomic adaptation in our selection experiment is faster than in marine stickleback adapting to freshwater habitat over approximately 17

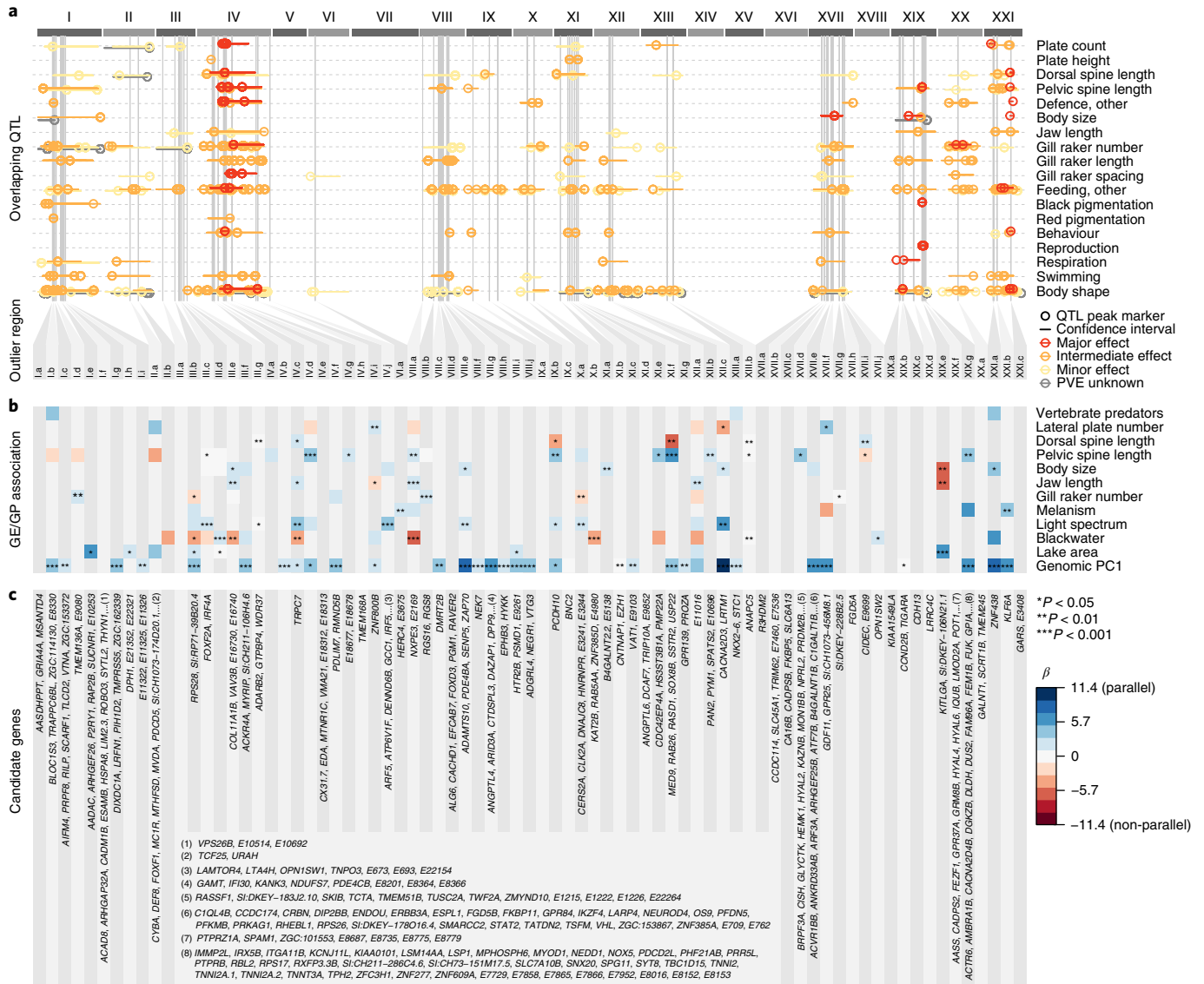


Fig. 5 | Outlier regions and overlapping QTL, candidate genes and genotype-environment and genotype-phenotype (GE/GP) associations across the adaptive radiation. a, Distribution of overlapping QTL. Circles indicate QTL peak markers, horizontal bars confidence intervals and colour codes the effect sizes: major (percentage variance explained (PVE) > 25%); intermediate (25% > PVE > 5%); and minor (PVE < 5%). **b**, GE/GP association and directionality of phenotypic and genomic change between the selection experiment and Haida Gwaii adaptive radiation. Predictors retained in a generalized linear model for each outlier region are shown as coloured squares, with blue boxes representing parallel genomic and phenotypic or genomic and ecological change, and red boxes representing non-parallel change (except for genomic PC1, for which directionality cannot be inferred). The colour code shows the relative effect sizes (β). **c**, Lists of candidate genes centred on divergent selection patterns in the corresponding outlier regions labelled in **a**.

generations in Russia[26], where adaptive alleles increased in frequency by 10–50%. Both the phenotype, with 0.15 haldanes in 12 generations[41], and adaptive genomic variation with a mean F_{ST} of 0.46 in 13 generations thus evolved at a very rapid rate typical of populations colonizing a new adaptive zone in an adaptive radiation[70]. Rapid adaptation to multifarious divergent selection thus seems to occur at a similar speed or faster than adaptation to a single selective force or with a single major locus as in some of these other examples.

Remarkably, the genomic basis of predator defence morphology, colour vision, feeding morphology and pigmentation overlapped with adaptive genomic change in the selection experiment, including some 'master adaptation genes' such as *EDA*, *OPNSW1/2*, *KITLG* and *MC1R*, which are frequently involved in repeated divergent adaptation of body armour, colour vision and pigmentation[48,49,59,61,65,66,71]. Many of these regions showed parallel associations with variation in traits and ecosystem variables across the Haida Gwaii adaptive radiation. This suggests that there was no major gap between phenotypic and genomic change for most of the diverging traits on the contemporary time scale of the selection experiment. Although phenotypic plasticity probably contributed to near-instantaneous phenotypic adaptation for traits such as eye size or gill raker length[41], our genomic findings suggest that selection has operated on genetic variation underlying most diverging phenotypic traits. Much of the genetic variation had to be shared with the Haida Gwaii adaptive radiation as standing genetic variation within Mayer Lake and across the archipelago, as outlier regions in the selection experiment evolved in parallel with the radiation in 36 of the 77 genomic outlier regions (Fig. 5). Determinism of adaptive evolution is therefore not only prevalent in the phenotype[41], but also in the genome, as predicted by the three major axes of natural selection in the radiation[30].

Genomic parallelism is somewhat surprising, given that a highly derived phenotype in the adaptive radiation was the source for the selection experiment. Mayer Lake stickleback are vertebrate-predation-, blackwater- and zooplankton-adapted specialists with little phenotypic variance. It is thus conceivable that such a specialist would have lacked the necessary standing genetic variation for rapid adaptation to the opposite extreme ecological theatre. In addition, a bottleneck during the selection experiment reduced genomic variation by 7%. Indeed, some shared alleles may have been lost during 12,000 years of adaptation to a blackwater lake: strong, but non-parallel associations between blackwater habitation and outlier regions on several chromosomes in the selection experiment suggest that different alleles from the adaptive radiation were favoured once the blackwater population had to re-adapt to the clearwater Roadside Pond (Fig. 5). Nevertheless, the Mayer Lake population still retained shared genetic variation at many loci for parallel adaptation in feeding morphology, defence morphology, pigmentation and vision. Genetic variation in Mayer Lake may have been maintained by disruptive or fluctuating selection[37,38,72] as a result of the large population size at this location (Supplementary Fig. 2), or by occasional introgression of adaptive alleles from adjacent stream ecotypes[46,73] or nearby pond ecotypes. Linkage disequilibrium between physically unlinked genomic regions containing such variation in Mayer Lake suggests that standing genetic variation is correlated in some individuals (Supplementary Fig. 20), compatible with all three hypotheses. Alternatively, bottlenecks during the colonization of Mayer Lake 12,000 years ago and during the experiment may not have been strong enough to remove adaptive genetic variation, such as in biological invasions where adaptive potential is usually not hampered with reductions in genetic diversity of 15–20%[12,74]. Drift during habitat shifts may thus rarely hamper sequential and rapid colonization of new niches in an adaptive radiation.

Our results confirm that natural selection generally overrides historical contingency at the genomic level in the adaptive radiation of threespine stickleback on Haida Gwaii. This is in line with phenotypic patterns[30] and previous genomic results for lake–stream and marine–freshwater divergence[7,26,27,34,46] and in spite of bottlenecks upon colonization and strong selection acting on new colonizers. Similar selection-driven phenotypic and genomic determinism has been found in other adaptive radiations based on adaptive introgression or a hybrid swarm origin rather than standing genetic variation as in the stickleback; for example, in East African cichlids[9,75,76], Darwin's finches[69,77] or *Heliconius* butterflies[78]. However, except for Darwin's finches[79] and our experiment, it remains to be shown whether the colonization of new adaptive zones can occur similarly fast on contemporary time scales. Our findings suggest that multifarious divergent selection acts rapidly on many different genes and regions in the genome, and that large steps in both phenotypic and genomic adaptation in adaptive radiations are taken within the first few generations, even when starting from a highly derived adaptive radiation member. Adaptive radiations may thus rapidly advance on contemporary time scales, given enough standing genetic variation in key functional traits and an ecological theatre offering new niche space and imposing multifarious divergent selection.

Methods

Experimental setup, sampling and ethics statement. In May 1993, 100 adult giant threespine stickleback (approx. 50% males/females) were captured in Mayer Lake and transferred to Roadside Pond (also referred to as 'Mayer Pond' in Leaver and Reimchen[41]). In 2004, 12 females were captured in Mayer Lake and in 2012, 11 females were caught in Roadside Pond corresponding to ~13 generations after release, assuming a population-average generation time of 1.5 years. In addition, stickleback from 25 freshwater populations across the Haida Gwaii archipelago, representing the range of successfully colonized freshwater habitats, were sampled between 1993 and 2012 (see Supplementary Table 1 in Marques, et al.[59]). Stickleback were captured using minnow traps and euthanized with an overdose of tricaine methanesulfonate (MS-222) in agreement with British Columbia's guidelines for scientific fish collection under Ministry of Environment permits SM09-51584 and SM10-62059 and University of Victoria Aquatic Unit facility Standard Operating Procedure OA2003. Collections in Naikoon Provincial Park and Drizzle Lake Ecological Reserve were carried out under park use permits 103171, 103172, 104795 and 104796. Samples were stored in 70% ethanol and the genomes of 58 individuals, including 12 Mayer Lake, 11 Roadside Pond and 1–4 individuals from 25 Haida Gwaii freshwater populations, 2 mainland British Columbia freshwater populations and 1 marine population, were re-sequenced and are listed in Supplementary Table 1 in our previous study focussing on the evolution of colour vision[59]. Alignment, variant and genotype calling, and filtering are described in Marques, et al.[59]. Note that we aligned against an improved ordering of scaffolds of the reference stickleback genome, and all genomic coordinates refer to this improved reference[80]. For the analyses in this study, we used either raw aligned reads with a mapping quality ≥17 and bases with a quality ≥17 for statistics computed on genotype likelihoods, or 1 of 2 subsets from the SNP dataset containing 7,888,602 high-quality SNPs among the 58 sequenced individuals for principal component analysis (unphased SNPs) and haplotype-based statistics (phased and imputed SNPs). The first 'selection experiment' subset contained 4,180,622 SNPs among 12 Mayer Lake and 11 Roadside Pond individuals. The second 'adaptive radiation' dataset contained 6,564,510 SNPs among 1 marine and 25 natural Haida Gwaii freshwater populations (including Mayer Lake) with one randomly picked individual per population. Read-backed phasing and imputation in both adaptive radiation and selection experiment SNP datasets was performed with SHAPEIT version 2.r790 (ref. [81]), with phase-informative reads covering 7.5% of all heterozygote genotypes and 30.9% of all graph segments.

Population genomic analyses. We described genomic change in the selection experiment using the following statistics: absolute allele frequency change (|ΔAF|), differentiation (F_{ST}), nucleotide diversity (π) and SFS (Tajima's D) computed from genotype likelihoods. First, we computed the unfolded 2D-SFS between Mayer Lake and Roadside Pond from aligned autosomal reads with angsd version 0.915, using the reference genome as the ancestral state. We used the 2D-SFS as a prior to estimate F_{ST} at single sites, as well as F_{ST}, Tajima's D and π in windows of 10-kb width, either non-overlapping or sliding with 2-kb step size from raw aligned reads in angsd[82–84] (filters as outlined above). Single-site F_{ST} was calculated from site alphas and betas computed by angsd. We also estimated minor allele frequencies in each population in angsd to calculate |ΔAF|. We calculated a weighted mean |ΔAF| with the weighted.mean function in R, using each SNP's 'starting allele frequency' (that is, the minor allele frequency estimated for the Mayer Lake population) as weights.

We compared the amount of genomic change in the selection experiment with natural populations in the Haida Gwaii radiation for genome-wide differentiation (F_{ST}) and absolute divergence (D_{XY}). For absolute divergence between the populations, we computed the unfolded 2D-SFS for pairs of individuals and calculated mean pairwise D_{XY} from the SFS using custom scripts. We computed pairwise D_{XY} within the populations Mayer Lake and Roadside Pond to obtain a baseline of expected pairwise D_{XY}. Then, we computed pairwise D_{XY} between the Mayer Lake and Roadside Pond populations, along with three lake versus stream ecotype populations (Mayer Lake versus Gold Creek, Drizzle Lake versus Drizzle Lake's inlet and outlet, and Spence Lake versus Spence Lake's outlet) and three large lake versus small pond populations (Mayer Lake versus Branta Pond, Mayer Lake versus Laurel Pond, and Mayer Lake versus Solstice Pond). We also estimated absolute divergence from the proportion of fixed differences among polymorphic sites between pairs of individuals. We annotated SNPs in the 'adaptive radiation' dataset using the Ensembl Variant Effect Predictor[85] and used the Picard Tool LiftoverVCF version 2.7.0 (ref. [86]) to move the SNPs into the original annotation[7]. We partitioned SNPs into missense, synonymous, intron, regulatory and intergenic SNPs using SnpSift version 4.2 (ref. [87]), and computed the proportion of fixed differences from the 012 output format of VCFtools version 0.1.15 (ref. [88]). Genome-wide differentiation between populations (F_{ST}) was calculated from previously published SNP array data[34,46] for the lake versus stream and large lake versus small pond comparisons. We ran a locus-by-locus analysis of molecular variance (AMOVA) for SNPs with at least 3 genotypes per population in Arlequin version 3.5.2.2 (ref. [89]) (Supplementary Table 4), resulting in >400 SNPs per comparison that should give an unbiased genome-wide F_{ST} estimate[90]. For lake versus stream comparisons with multiple stream populations (Drizzle Lake versus Drizzle Lake's inlet and outlet, and Mayer Lake versus Gold, Woodpile and Spam Creeks[46]), we used hierarchical AMOVAs with each population retained as a separate sample but grouped into either the lake or stream group (Supplementary Table 4). Alpha and beta estimates from the AMOVA and the F_{ST} computation in angsd for Mayer Lake versus Roadside Pond were pooled to sums of nominators and denominators to obtain a weighed mean F_{ST} estimate[91].

We identified probable genomic targets of divergent selection between the source and transplant population with a two-step outlier approach. First, we inferred an optimal, neutral demographic model on the 2D-SFS using fastsimcoal2 version 2.6 (ref. [92]). Second, we simulated neutral genomic data under the best demographic model, against which we identified outlying genomic regions in the observed data. We folded the 2D-SFS using custom scripts, fit 12 different demographic models (Supplementary Fig. 2b and Supplementary Data 1) to

the observed 2D-SFS with fastsimcoal2 and compared their likelihoods using the Akaike information criterion following Excoffier et al.[92]. We maximized the likelihood of each model from 100 random starting parameter combinations in 10 to a maximum of 50 ECM cycles, with a stopping criterion of 0.001 (ref. [92]). A total of 100,000 coalescent simulations were used to approximate the expected 2D-SFS. In all simulations, we used a mutation rate of 1.7×10^{-8}, following Feulner et al.[93] and a founding population size of $2N = 200$ individuals for the Roadside Pond population, and generated Mayer Lake samples 5 generations before Roadside Pond to account for different sampling years (Supplementary Data 1). Likelihood and parameter estimates for each model were obtained from the run with the highest likelihood among the 100 optimizations. We simulated neutral genomic data under the best demographic model with fastsimcoal2 for 4 different recombination rates: high = 4–16 cM/Mb; intermediate = 1.5–4.0 cM/Mb; low = 0.5–1.5 cM/Mb; and very low = 0–0.05 cM/Mb. For each recombination range, we generated 1,000 replicate DNA segments of 1 Mb length, with a mutation rate of 1.7×10^{-8} and a random recombination rate from that range, assuming a uniform distribution (very low or low recombination rate) or log-uniform distribution (intermediate or high recombination rate) that reflect the frequency of recombination rate variation in the stickleback genome[80]. We transformed the simulated data into VCF format using custom scripts and computed weighted F_{ST}, Tajima's D and π in non-overlapping 10-kb windows using VCFtools version 0.1.14 (ref. [88]).

A selective sweep caused by divergent selection between habitats is expected to lead to excess differentiation (F_{ST}) between populations at and around the site under selection, as well as reduced diversity in the population experiencing the selective sweep and a shifted SFS, reflected by a strongly negative Tajima's D upon completion of the sweep. In addition, haplotype-based statistics are able to detect soft and incomplete sweeps within a populations (iHS[94] and H12 (ref. [95])) or completed sweeps in one of two populations (XPEHH)[96]. We computed the haplotype-based selection statistics iHS[94], H12[95] and XPEHH[96] for phased and imputed bi-allelic SNPs with a minor allele frequency >5% in the 'selection experiment' dataset and for simulated SNP data. We computed iHS and H12 separately for the Mayer Lake and Roadside Pond populations, using only SNPs with a minor allele frequency of >5% in the respective population. We calculated the proportion of extreme iHS and XPEHH values ('w-iHS', the proportion of $|iHS| > 2$, following Voight, et al.[94] and 'w-XPEHH', the proportion of $|XPEHH| > 2$) in non-overlapping 10-kb windows containing more than 10 iHS or XPEHH estimates, respectively, for both observed and simulated datasets. We used selscan version 1.1.0b[97] with default parameters to compute iHS and XPEHH and the proportion of extreme values in 10-kb windows. We also computed H12 for the observed dataset using scripts published alongside the H12 method[95], with a bin width of 81 SNPs, resulting in, on average, 8.3-kb wide windows (close to the 10-kb windows identified as optimal and robust to various demographic scenarios by Garud et al.[95]).

We identified outliers against neutral expectations for 6 10-kb non-overlapping window statistics: F_{ST}, change in nucleotide diversity ($\Delta\pi = \pi_{Roadside} - \pi_{Mayer}$), change in Tajima's D ($\Delta D_T = D_{T,Roadside} - D_{T,Mayer}$), w-iHS$_{Mayer}$, w-iHS$_{Roadside}$ and w-XPEHH. Our ability to detect signatures of selective sweeps with window-based statistics depends on the local recombination rate, with stronger hitchhiking in low-recombination-rate regions leading to more prominent signals and a greater variation in such statistics (Supplementary Figs. 3 and 4). We therefore identified outlier windows separately in genomic regions with high, intermediate, low and very low recombination rates (see above and Supplementary Fig. 4). We assigned 10-kb windows to recombination rate bins according to local recombination rates estimated in the middle of each 10-kb window, as described previously[59]. For each 10-kb window, we computed the empirical quantile of the observed F_{ST}, $\Delta\pi$, ΔD_T, w-iHS$_{Mayer}$, w-iHS$_{Roadside}$ and w-XPEHH value against the simulated distribution of the statistic in the respective recombination bin with the function 'ecdf' in R version 3.3.1 (ref. [98]). We converted quantiles to two-sided P values for $\Delta\pi$ and ΔD_T and one-sided P values for the other statistics.

We identified genomic regions likely to be under divergent selection between Mayer Lake and Roadside Pond ('outlier regions') based on overlapping outlier signatures in these six selection statistics. To capture the shared signal, we applied Fisher's combined probability test to the four to six P values in each 10-kb window, as implemented in the R package 'metap'. P values from Fisher's combined probability test were corrected for multiple testing using the false discovery rate method[99] implemented in 'p.adjust', converted to q values ($= 1 - p_{adj}$) and z-transformed using the R function 'qnorm'. We used a hidden Markov model (HMM) approach to group adjacent 10-kb windows into outlier regions. The z-transformed q values were used as input to HMMs with two or three normally distributed states. We optimized parameters of both HMMs from 1,000 random starting parameters using the Baum–Welch algorithm implemented in the R package 'HiddenMarkov'. The three-state HMM better fit the data according to the Akaike information criterion and was thus used to assign all 42,996 10-kb windows to the three states using the Viterbi algorithm. Preliminary outlier regions were obtained from joining adjacent windows assigned to the state capturing highly significant Fisher's combined probability test P values. Then, only outlier regions that contained significant outliers with $P < 0.01$ for each of the statistics F_{ST}, $\Delta\pi$ or ΔD_T and w-iHS$_{Mayer}$, w-iHS$_{Roadside}$ or w-XPEHH, as well as outlier regions with strongly aligned signatures for these statistics plus H12$_{Mayer}$ or H12$_{Roadside}$, were retained in the final set of outlier regions reflecting divergent selection between Mayer Lake and Roadside Pond. We did not further analyse signatures of, for example, shared directional or background selection, which should result in reduced diversity and Tajima's D or significant haplotype-based statistics in both populations, but not in differentiation between the populations on such short timescales.

We quantified adaptive differentiation between the source and transplant population by computing single SNP F_{ST} and retaining the top 5% F_{ST} SNPs in each outlier region, thereby probably containing the few SNPs under selection and many more linked, hitchhiking SNPs in each region affected by divergent selection. We compared this distribution of adaptive differentiation in the selection experiment with adaptive differentiation among three postglacial pairs of lake and stream ecotypes on Haida Gwaii, using F_{ST} estimates from only those SNPs previously identified to be under selection in the respective ecotype comparison[46], which also probably reflect hitchhiking SNPs and, to a lesser degree, direct targets of selection. For the top 5% F_{ST} SNPs in each outlier region, we computed the expected selection coefficient under a pure selection model based on the allele frequency changes at these SNPs over 12.7 generations and assuming incomplete dominance $h = 0.5$ following equation (3.2) in Gillespie[100]. These calculations probably overestimate selection coefficients due to unaccounted contributions of drift and should thus be interpreted with caution.

We computed linkage disequilibrium as r^2 between the most divergent 15 SNPs polymorphic in both Mayer Lake and Roadside Pond for each outlier region using VCFtools, both within and between chromosomes. We assessed whether linkage disequilibrium between outlier regions on different chromosomes exceeded neutral expectations of no linkage disequilibrium. We derived the neutral distribution of linkage disequilibrium with the observed sample sizes by randomly choosing SNPs outside outlier regions from each chromosome with a distance of at least 500 kb between SNPs on the same chromosome ($n = 617$) and by computing interchromosomal linkage disequilibrium between these SNPs. Then, we determined whether the mean observed linkage disequilibrium between two outlier regions was greater than the 95% quantile of the neutral observed interchromosomal linkage disequilibrium distribution.

We associated the genomic signatures of divergent selection with potential sources of selection in the experiment and the adaptive radiation by studying their gene content and the gene's functional annotations, from their overlap with previously described stickleback QTL that have been mapped in genetic studies of specific phenotypes[47], and from GE/GP associations across the Haida Gwaii radiation. First, we identified candidate genes by inspecting the patterns of 10-kb sliding-window statistics F_{ST}, π_{Mayer}, $\pi_{Roadside}$, Tajima's D, and the single-locus statistics iHS$_{Mayer}$, iHS$_{Roadside}$, H12$_{Mayer}$, H12$_{Roadside}$, XPEHH and $|\Delta AF|$ visually (Fig. 4 and Supplementary Figs. 5–19). We retained a list of genes centred on, or adjacent to, selective sweep signatures (Supplementary Table 3). Then, we tested this list of candidate genes for enrichment of Gene Ontology terms using the STRING database version 10 (ref. [101]) and retrieved functional and expression information from zebrafish[58] and related mouse, rat and human databases[102,103]. In addition, we identified overlaps between outlier regions and QTL previously identified in other stickleback populations in the Northern Hemisphere using the list and peak marker location and confidence intervals of Peichel and Marques[47]. QTL were grouped into major-, intermediate- or minor-effect size classes, respectively, when they explained >25%, between 5 and 25%, or <5% of the phenotypic variation[47].

Finally, we determined whether outlier regions in the selection experiment evolved in predictable directions given the environmental contrast. In the absence of replicate experimental ponds, we used GE/GP associations across the larger Haida Gwaii stickleback adaptive radiation with many natural replicates to infer whether the same genomic regions evolved in parallel direction. For each outlier region, we identified the SNPs with the strongest allele frequency change between Mayer Lake and Roadside Pond (top 1% $|\Delta AF|$), assigned the alleles as Mayer Lake-like or Roadside Pond-like based on which population they were more frequent in, extracted the genotypes for those SNPs from the adaptive radiation SNP dataset containing single genomes of 1 marine and 25 freshwater populations including Mayer Lake, recoded the alleles as 0 (Mayer-like) or 1 (Roadside-like), combined them into multidimensional scaling (MDS) factors and polarized the MDS factors for Mayer Lake to be represented by low values and Roadside Pond by high values in R. Next, we used the genomic MDS factor of each outlier region as a response variable in a generalized linear model with 12 phenotypic and ecological properties of the 26 Haida Gwaii populations as predictors (see below). For each outlier region's generalized linear model, we performed variable selection by iteratively removing non-significant predictors (χ^2 tests, $P > 0.1$). Parallelism was inferred if the MDS factor of an outlier region was positively associated with a predictor. No parallelism was inferred if the association was negative.

We used the presence of blackwater (transmission at 400 nm < 74%), the presence of vertebrate predators in a population, and whether or not the population consisted of predominantly melanistic phenotypes as binary predictors. Continuous predictors were: lake area (log-transformed), light spectrum, mean body size (standard length), mean lateral plate number (excluding fully plated individuals), mean dorsal spine length, mean pelvic spine length, mean jaw length and mean gill raker number. The linear measurements were size-corrected as

described in Reimchen et al.[30] and all were scaled and standardized to a mean of zero and a standard deviation of one. We polarized all predictors so that the change from Mayer Lake to Roadside Pond would represent a positive shift (that is, a shift from blackwater to clearwater and from melanism to reduced melanism, a decrease in lake area, body size, lateral plate number, pelvic spine length and gill raker number, and an increase in jaw length[41]). As the last predictor, we included geographic structuring using the first principal components axis from genomic variation in the adaptive radiation SNP dataset. We used genotype likelihoods from the adaptive radiation SNP dataset, computed the site allele frequency spectrum to obtain a covariance matrix, as implemented in angsd and ngsCovar[84], and performed the eigenvalue decomposition in R to obtain the first principal component. We visualized phenotypic change in the selection experiment[41] and phenotypic divergence between Mayer Lake, it's stream ecotype (Mayer Stream = Gold Creek[46]), and Laurel, Branta and Solstice ponds[39] from the data of this earlier work using the size-correction of Leaver and Reimchen[41] within each population for all datasets combined. Analyses were performed on Compute Canada's WestGrid computer cluster infrastructure (www.westgrid.ca).

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

Code availability. Custom scripts to compute D_{XY} from the SFS, fold 2D-SFS and fit the HMM are available from https://github.com/marqueda.

Data availability. Aligned sequences can be accessed under accession SRP100209 on the NCBI Sequence Read Archive (www.ncbi.nlm.nih.gov/sra).

Received: 2 May 2017; Accepted: 17 May 2018;
Published online: 25 June 2018

References

1. Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, Oxford, 2000).
2. Grant, P. R. Speciation and the adaptive radiation of Darwin finches. *Am. Sci.* **69**, 653–663 (1981).
3. Losos, J. B., Jackman, T. R., Larson, A., Queiroz, K. & Rodriguez-Schettino, L. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**, 2115–2118 (1998).
4. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, Oxford, 2003).
5. Muschick, M., Barluenga, M., Salzburger, W. & Meyer, A. Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evol. Biol.* **11**, 116 (2011).
6. Barrett, R. D. & Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).
7. Jones, F. C. et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).
8. Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207 (2004).
9. Meier, J. I. et al. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* **8**, 14363 (2017).
10. Price, T. D., Qvarnstrom, A. & Irwin, D. E. The role of phenotypic plasticity in driving genetic evolution. *Proc. Biol. Sci.* **270**, 1433–1440 (2003).
11. Schlotterer, C., Kofler, R., Versace, E., Tobler, R. & Franssen, S. U. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity (Edinb.)* **116**, 248 (2016).
12. Barrett, S. C. H., Colautti, R. I., Dlugosch, K. M. & Rieseberg, L. H. *Invasion Genetics: The Baker and Stebbins Legacy* (Wiley-Blackwell, Hoboken, NJ, 2016).
13. Burke, M. K. et al. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* **467**, 587–590 (2010).
14. Fritz, M. L. et al. Contemporary evolution of a Lepidopteran species, *Heliothis virescens*, in response to modern agricultural practices. *Mol. Ecol.* **27**, 167–181 (2018).
15. Tobler, R. et al. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Mol. Biol. Evol.* **31**, 364–375 (2014).
16. Graves, J. L. et al. Genomics of parallel experimental evolution in *Drosophila*. *Mol. Biol. Evol.* **34**, 831–842 (2017).
17. Huang, Y., Wright, S. I. & Agrawal, A. F. Genome-wide patterns of genetic variation within and among alternative selective regimes. *PLoS Genet.* **10**, e1004527 (2014).
18. Franks, S. J., Kane, N. C., O'Hara, N. B., Tittes, S. & Rest, J. S. Rapid genome-wide evolution in *Brassica rapa* populations following drought revealed by sequencing of ancestral and descendant gene pools. *Mol. Ecol.* **25**, 3622–3631 (2016).
19. van't Hof, A. E., Edmonds, N., Dalikova, M., Marec, F. & Saccheri, I. J. Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science* **332**, 958–960 (2011).
20. Reid, N. M. et al. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* **354**, 1305–1308 (2016).
21. Fraser, B. A., Kunstner, A., Reznick, D. N., Dreyer, C. & Weigel, D. Population genomics of natural and experimental populations of guppies (*Poecilia reticulata*). *Mol. Ecol.* **24**, 389–408 (2015).
22. Hendry, A. P. & Kinnison, M. T. Perspective: the pace of modern life: measuring rates of contemporary microevolution. *Evolution* **53**, 1637–1653 (1999).
23. Reznick, D. N. & Ghalambor, C. K. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112-113**, 183–198 (2001).
24. Stockwell, C. A., Hendry, A. P. & Kinnison, M. T. Contemporary evolution meets conservation biology. *Trends Ecol. Evol.* **18**, 94–101 (2003).
25. Bell, M. A., Aguirre, W. E. & Buck, N. J. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**, 814–824 (2004).
26. Terekhanova, N. V. et al. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genet.* **10**, e1004696 (2014).
27. Lescak, E. A. et al. Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proc. Natl Acad. Sci. USA* **112**, E7204–E7212 (2015).
28. Aguirre, W. E. & Bell, M. A. Twenty years of body shape evolution in a threespine stickleback population adapting to a lake environment. *Biol. J. Linn. Soc.* **105**, 817–831 (2012).
29. Hohenlohe, P. A. et al. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **6**, e1000862 (2010).
30. Reimchen, T. E., Bergstrom, C. & Nosil, P. Natural selection and the adaptive radiation of Haida Gwaii stickleback. *Evol. Ecol. Res.* **15**, 241–269 (2013).
31. Moodie, G. E. E. & Reimchen, T. E. Phenetic variation and habitat differences in *Gasterosteus* populations of the Queen Charlotte Islands. *Syst. Zool.* **25**, 49–61 (1976).
32. Reimchen, T. E. in *The Evolutionary Biology of the Threespine Stickleback* (eds Bell, M. A. & Foster, S. A.) 240–276 (Oxford Univ. Press, Oxford, 1994).
33. Bergstrom, C. A. & Reimchen, T. E. Habitat dependent associations between parasitism and fluctuating asymmetry among endemic stickleback populations. *J. Evol. Biol.* **18**, 939–948 (2005).
34. Deagle, B. E., Jones, F. C., Absher, D. M., Kingsley, D. M. & Reimchen, T. E. Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. *Mol. Ecol.* **22**, 1917–1932 (2013).
35. Reimchen, T. E. Predator handling failures of lateral plate morphs in *Gasterosteus aculeatus*: functional implications for the ancestral plate condition. *Behaviour* **137**, 1081–1096 (2000).
36. Reimchen, T. E. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*—an adaptation to predators. *Can. J. Zool.* **58**, 1232–1244 (1980).
37. Reimchen, T. E. & Nosil, P. Temporal variation in divergent selection on spine number in threespine stickleback. *Evolution* **56**, 2472–2483 (2002).
38. Reimchen, T. E. & Nosil, P. Variable predation regimes predict the evolution of sexual dimorphism in a population of threespine stickleback. *Evolution* **58**, 1274–1281 (2004).
39. Reimchen, T. E., Stinson, E. M. & Nelson, J. S. Multivariate differentiation of parapatric and allopatric populations of threespine stickleback in the Sangan River watershed, Queen Charlotte Islands. *Can. J. Zool.* **63**, 2944–2951 (1985).
40. Spoljaric, M. A. & Reimchen, T. E. 10 000 years later: evolution of body shape in Haida Gwaii three-spined stickleback. *J. Fish. Biol.* **70**, 1484–1503 (2007).
41. Leaver, S. D. & Reimchen, T. E. Abrupt changes in defence and trophic morphology of the giant threespine stickleback (*Gasterosteus* sp.) following colonization of a vacant habitat. *Biol. J. Linn. Soc.* **107**, 494–509 (2012).
42. Moodie, G. E. E. Morphology, life-history, and ecology of an unusual stickleback (*Gasterosteus aculeatus*) in the Queen Charlotte Islands, Canada. *Can. J. Zool.* **50**, 721–732 (1972).
43. Moodie, G. E. E. Predation, natural selection and adaptation in an unusual threespine stickleback. *Heredity* **28**, 155–167 (1972).
44. Oreilly, P., Reimchen, T. E., Beech, R. & Strobeck, C. Mitochondrial DNA in *Gasterosteus* and pleistocene glacial refugium on the Queen Charlotte Islands, British Columbia. *Evolution* **47**, 678–684 (1993).
45. Flamarique, I. N., Bergstrom, C., Cheng, C. L. & Reimchen, T. E. Role of the iridescent eye in stickleback female mate choice. *J. Exp. Biol.* **216**, 2806–2812 (2013).

46. Deagle, B. E. et al. Population genomics of parallel phenotypic evolution in stickleback across stream–lake ecological transitions. *Proc. Biol. Sci.* **279**, 1277–1286 (2012).
47. Peichel, C. L. & Marques, D. A. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Phil. Trans. R. Soc. Lond. B* **372**, 20150486 (2017).
48. Peichel, C. L. et al. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**, 901–905 (2001).
49. Colosimo, P. F. et al. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
50. Colosimo, P. F. et al. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* **2**, E109 (2004).
51. Wark, A. R. et al. Genetic architecture of variation in the lateral line sensory system of threespine sticklebacks. *G3* **2**, 1047–1056 (2012).
52. Greenwood, A. K., Wark, A. R., Yoshida, K. & Peichel, C. L. Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Curr. Biol.* **23**, 1884–1888 (2013).
53. Rennison, D. J., Owens, G. L., Heckman, N., Schluter, D. & Veen, T. Rapid adaptive evolution of colour vision in the threespine stickleback radiation. *Proc. Biol. Sci.* **283**, 20160242 (2016).
54. Perez-Leighton, C. E., Schmidt, T. M., Abramowitz, J., Birnbaumer, L. & Kofuji, P. Intrinsic phototransduction persists in melanopsin-expressing ganglion cells lacking diacylglycerol-sensitive TRPC subunits. *Eur. J. Neurosci.* **33**, 856–867 (2011).
55. Nakajima, Y., Moriyama, M., Hattori, M., Minato, N. & Nakanishi, S. Isolation of ON bipolar cell genes via hrGFP-coupled cell enrichment using the mGluR6 promoter. *J. Biochem.* **145**, 811–818 (2009).
56. Amsterdam, A. et al. Identification of 315 genes essential for early zebrafish development. *Proc. Natl Acad. Sci. USA* **101**, 12792–12797 (2004).
57. Nuckels, R. J., Ng, A., Darland, T. & Gross, J. M. The vacuolar-ATPase complex regulates retinoblast proliferation and survival, photoreceptor morphogenesis, and pigmentation in the zebrafish eye. *Invest. Ophthalmol. Vis. Sci.* **50**, 893–905 (2009).
58. Howe, D. G. et al. ZFIN, the zebrafish model organism database: increased support for mutants and transgenics. *Nucleic Acids Res.* **41**, D854–D860 (2013).
59. Marques, D. A. et al. Convergent evolution of *SWS2* opsin facilitates adaptive radiation of threespine stickleback into different light environments. *PLoS Biol.* **15**, e2001627 (2017).
60. Gwynn, B., Smith, R. S., Rowe, L. B., Taylor, B. A. & Peters, L. L. A mouse TRAPP-related protein is involved in pigmentation. *Genomics* **88**, 196–203 (2006).
61. Hoekstra, H. E., Hirschmann, R. J., Bundey, R. A., Insel, P. A. & Crossland, J. P. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101–104 (2006).
62. Dickinson, M. E. et al. High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).
63. Ignatius, M. S., Moose, H. E., El-Hodiri, H. M. & Henion, P. D. *colgate/hdac1* repression of foxd3 expression is required to permit mitfa-dependent melanogenesis. *Dev. Biol.* **313**, 568–583 (2008).
64. Patterson, L. B. & Parichy, D. M. Interactions with iridophores and the tissue environment required for patterning melanophores and xanthophores during zebrafish adult pigment stripe formation. *PLoS Genet.* **9**, e1003561 (2013).
65. Miller, C. T. et al. *cis*-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* **131**, 1179–1189 (2007).
66. Rosenblum, E. B., Hoekstra, H. E. & Nachman, M. W. Adaptive reptile color variation and the evolution of the *Mc1r* gene. *Evolution* **58**, 1794–1808 (2004).
67. Malek, T. B., Boughman, J. W., Dworkin, I. & Peichel, C. L. Admixture mapping of male nuptial colour and body shape in a recently formed hybrid population of threespine stickleback. *Mol. Ecol.* **21**, 5265–5279 (2012).
68. Miller, C. T. et al. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics* **197**, 405–420 (2014).
69. Lamichhaney, S. et al. A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science* **352**, 470–474 (2016).
70. Gingerich, P. D. Rates of evolution: effects of time and temporal scaling. *Science* **222**, 159–161 (1983).
71. Rennison, D. J., Owens, G. L. & Taylor, J. S. Opsin gene duplication and divergence in ray-finned fish. *Mol. Phylogenet. Evol.* **62**, 986–1008 (2012).
72. Reimchen, T. E. Predator-induced cyclical changes in lateral plate frequencies of *Gasterosteus*. *Behaviour* **132**, 1079–1094 (1995).
73. Stinson, E. M. *Threespine Sticklebacks (Gasterosteus aculeatus) in Drizzle Lake and Its Inlet, Queen Charlotte Islands: Ecological and Behavioural Relationships and Their Relevance to Reproductive Isolation*. MSc thesis, Univ. Alberta (1983).
74. Dlugosch, K. M. & Parker, I. M. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* **17**, 431–449 (2008).
75. Keller, I. et al. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol. Ecol.* **22**, 2848–2863 (2013).
76. McGee, M. D., Neches, R. Y. & Seehausen, O. Evaluating genomic divergence and parallelism in replicate ecomorphs from young and old cichlid adaptive radiations. *Mol. Ecol.* **25**, 260–268 (2016).
77. Lamichhaney, S. et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015).
78. Dasmahapatra, K. K. et al. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).
79. Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711 (2002).
80. Glazer, A. M., Killingbeck, E. E., Mitros, T., Rokhsar, D. S. & Miller, C. T. Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. *G3* **5**, 1463–1472 (2015).
81. Delaneau, O., Howie, B., Cox, A. J., Zagury, J. F. & Marchini, J. Haplotype estimation using sequencing reads. *Am. J. Hum. Genet.* **93**, 687–696 (2013).
82. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
83. Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y. & Wang, J. SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS ONE* **7**, e37558 (2012).
84. Fumagalli, M. et al. Quantifying population genetic differentiation from next-generation sequencing data. *Genetics* **195**, 979–992 (2013).
85. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
86. Picard Tools (Broad Institute, 2017); http://broadinstitute.github.io/picard
87. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strainw1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
88. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
89. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
90. Willing, E. M., Dreyer, C. & van Oosterhout, C. Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE* **7**, e42649 (2012).
91. Bhatia, G., Patterson, N., Sankararaman, S. & Price, A. L. Estimating and interpreting F_{ST}: the impact of rare variants. *Genome Res.* **23**, 1514–1521 (2013).
92. 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).
93. Feulner, P. G. et al. Genomics of divergence along a continuum of parapatric population differentiation. *PLoS Genet.* **11**, e1004966 (2015).
94. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).
95. Garud, N. R., Messer, P. W., Buzbas, E. O. & Petrov, D. A. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genet.* **11**, e1005004 (2015).
96. Sabeti, P. C. et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, U913–U918 (2007).
97. Szpiech, Z. A. & Hernandez, R. D. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. *Mol. Biol. Evol.* **31**, 2824–2827 (2014).
98. R Development Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2016); http://www.r-project.org/
99. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
100. Gillespie, J. H. *Population Genetics: A Concise Guide* 2nd edn (Johns Hopkins Univ. Press, Baltimore, MA, 2004).
101. Szklarczyk, D. et al. STRINGv10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
102. Blake, J. A. et al. Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* **45**, D723–D729 (2017).
103. Shimoyama, M. et al. The Rat Genome Database 2015: genomic, phenotypic and environmental variations and disease. *Nucleic Acids Res.* **43**, D743–D750 (2015).

Acknowledgements

We thank B. Deagle, S. D. Leaver, C. B. Lowe, S. D. Brady, J. Turner, K. Lindblad-Toh and the Broad Institute Genomics Platform for help with sequences, samples and morphometric analysis, and B. Moa for bioinformatics support. This work was funded by the National Research Council Canada grant NRC2354 to T.E.R. and National Institute of Health grants 3P50HG002568-09S1 ARRA and 3P50HG002568 to D.M.K.

Author contributions

T.E.R. conceived the study, ran the experiment, collected fish and ecological data in the field, and acquired morphological data. D.M.K., F.C.J. and F.D.P. generated sequencing data and genotype calls. D.A.M. designed and performed all subsequent analyses and wrote the manuscript with contributions from all co-authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0581-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.A.M.

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

nature research

Corresponding author(s): David A. Marques

Reporting Summary

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

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☒	☐	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	See full documentation in the Methods section and Code Availability statement in the paper.
Data analysis	See full documentation in the Methods section and Code Availability statement in the paper.

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Aligned sequences can be accessed under accession SRP100209 on the NCBI short read archive (www.ncbi.nlm.nih.gov/sra).

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Ecological, evolutionary & environmental sciences study design

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

Study description	Selection experiment: transplant of 100 stickleback from a single population of threespine stickleback (Mayer Lake, Haida Gwaii, Canada) to a single small pond (Roadside Pond), followed by 19 years of evolution in the new habitat and re-sequencing of 12 and 11 individuals from source and transplant populations, respectively. Comparison to 25 additional, natural populations in the Haida Gwaii archipelago, including 1 marine source and 24 freshwater population replicates.
Research sample	Threespine stickleback (Gasterosteus aculeatus), 58 wild-caught individuals (56 females, 2 males) from 26 natural populations on Haida Gwaii and from the transplant population were used for whole genome resequencing.
Sampling strategy	Sample size of 100 fish with similar proportions of males and females for the population transplant was chosen by T.E.R. based on (i) knowledge from population genetics at the time (1993) on what population size would not incur a strong bottleneck and (ii) on what population size a pond of this size should be able to support ecologically with the assumption that all transplanted individuals will survive. Sample sizes for sequencing were determined by research budget and sufficiency of these sample sizes for detecting selection is demonstrated in the paper by using a neutral demographic null model and identifying outliers against this null model.
Data collection	See Methods section and Leaver & Reimchen (2012) Biol J Linn Soc 107: 494-509 for full description.
Timing and spatial scale	Experimental set-up in 1993, collections for genomic samples in 2004 (Mayer Lake) and 2012 (Roadside Pond). Additional collections for phenotypic data are outlined in Leaver & Reimchen (2012) Biol J Linn Soc 107: 494-509. See Methods section for additional detail.
Data exclusions	No data was excluded.
Reproducibility	The 19-year experiment was not repeated. The experimentally evolved population was compared to naturally evolved populations in the Haida Gwaii archipelago (evolution for <12,000 years) as replicates.
Randomization	We subsampled one individual per population for some analyses, using the R-function sample. See Methods section for details.
Blinding	Blinding was not necessary, as all statistics were computed, fully reproducible, directly from genomic data.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Not recorded. See Leaver & Reimchen (2012) Biol J Linn Soc 107: 494-509.
Location	Locations and their properties are given in Tab. 1 in Marques et al. (2017) PLoS Biology 15: e2001627 and lat / long coordinates for these locations are given in Marques et al. (2017) PLoS Biology 15: e2001627, Reimchen et al. (2012) Evol Ecol Res 15: 241-269 and Reimchen & Nosil (2006) Can J Zool 84: 643-654.
Access and import/export	Collections in Naikoon Provincial Park and Drizzle Lake Ecological Reserve were carried out under park use permits: 103171, 103172, 104795 and 104796.
Disturbance	Small population samples were collected on single / few days in a respective year to avoid disturbance. See Leaver & Reimchen (2012) Biol J Linn Soc 107: 494-509.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials: Specimens / materials are stored in the collection of T.E.R at the University of Victoria and can be examined on site on request.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals: Not applicable.

Wild animals: Stickleback were captured using minnow traps and euthanized with an overdose of tricaine methanesulfonate (MS-222) in agreement with British Columbia's guidelines for scientific fish collection, under Ministry of Environment permits SM09-51584 and SM10-62059 and University of Victoria Aquatic Unit facility Standard Operating Procedure OA2003.

Field-collected samples: Stickleback were captured using minnow traps and euthanized with an overdose of tricaine methanesulfonate (MS-222) in agreement with British Columbia's guidelines for scientific fish collection, under Ministry of Environment permits SM09-51584 and SM10-62059 and University of Victoria Aquatic Unit facility Standard Operating Procedure OA2003.

nature ecology & evolution

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41559-018-0581-8

In the format provided by the authors and unedited.

Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation

David A. Marques[1,2,3]*, Felicity C. Jones[4,5], Federica Di Palma[6,7], David M. Kingsley[4] and Thomas E. Reimchen[1]

[1]Department of Biology, University of Victoria, Victoria, British Columbia, Canada. [2]Aquatic Ecology & Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland. [3]Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland. [4]Department of Developmental Biology, HHMI and Stanford University School of Medicine, Stanford, CA, USA. [5]Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany. [6]Earlham Institute, Norwich Research Park, Norwich, UK. [7]Department of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, UK. *e-mail: david.marques@eawag.ch

Supplementary Information

Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation

David A. Marques, Felicity C. Jones, Federica Di Palma, David M. Kingsley & Thomas E. Reimchen

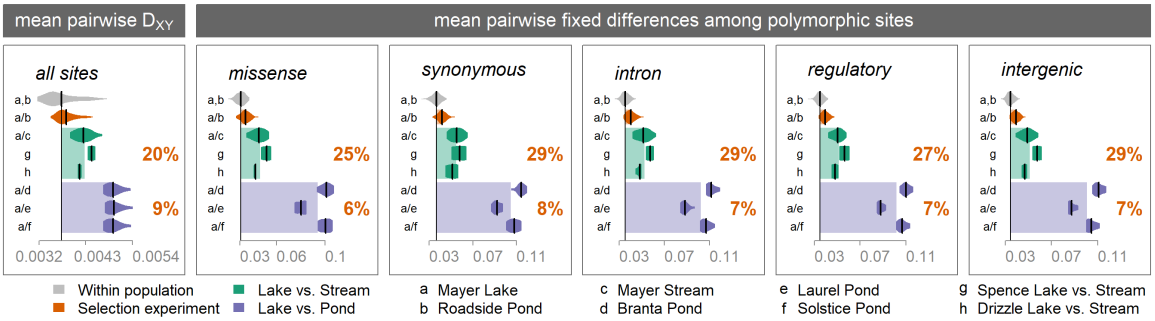
Table of contents

Supplementary Results

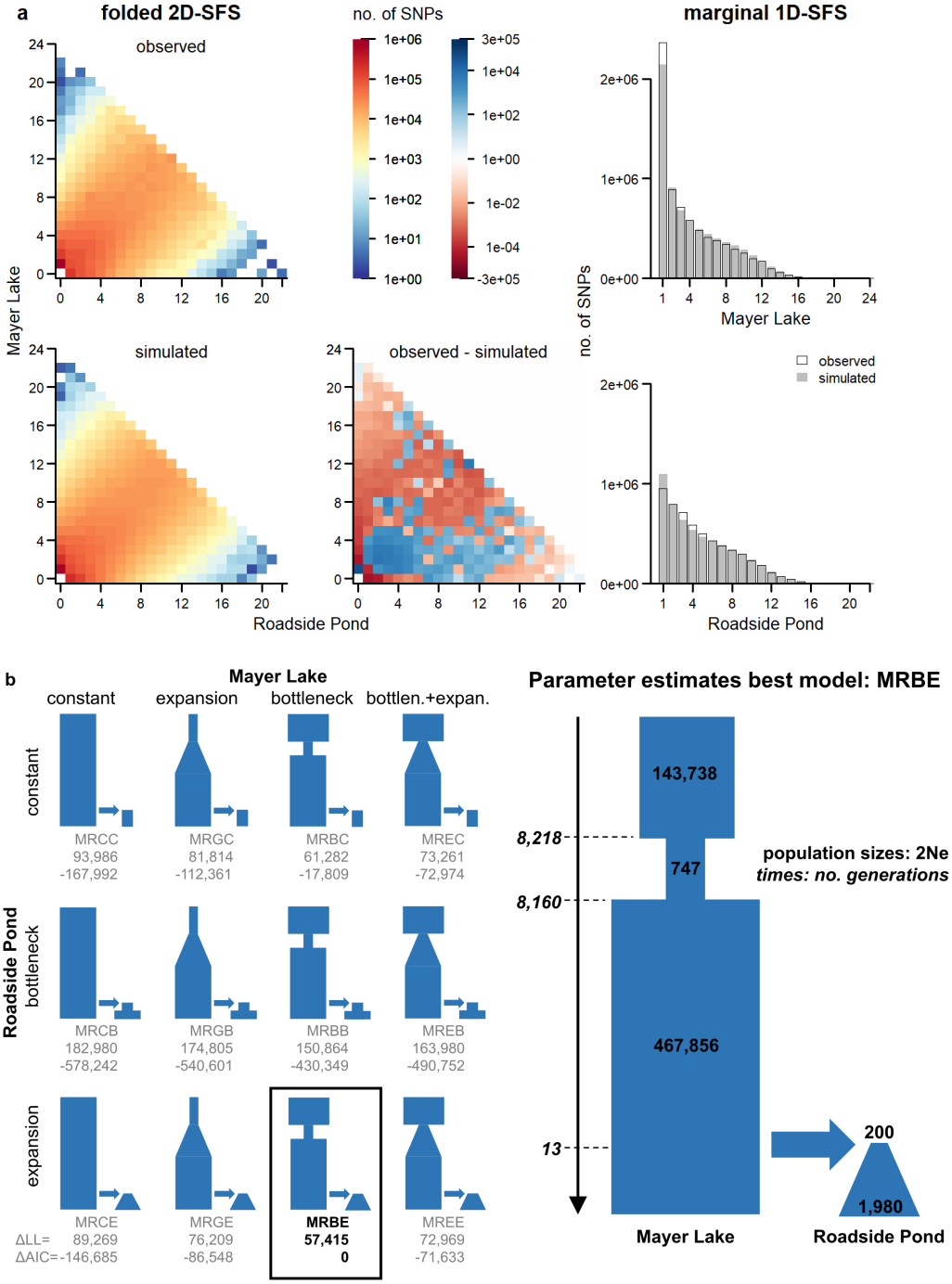
Genomic change at sites of different functional categories. Absolute divergence measured by the proportion of fixed differences among polymorphic sites between pairs of individuals in the selection experiment increased 25-29% compared to lake and stream ecotypes and 6-8% compared to Mayer Lake vs. small pond populations, in similar magnitude for non-synonymous, synonymous, intron, regulatory and intergenic sites (Supplementary Fig. 1).

Supplementary Figures

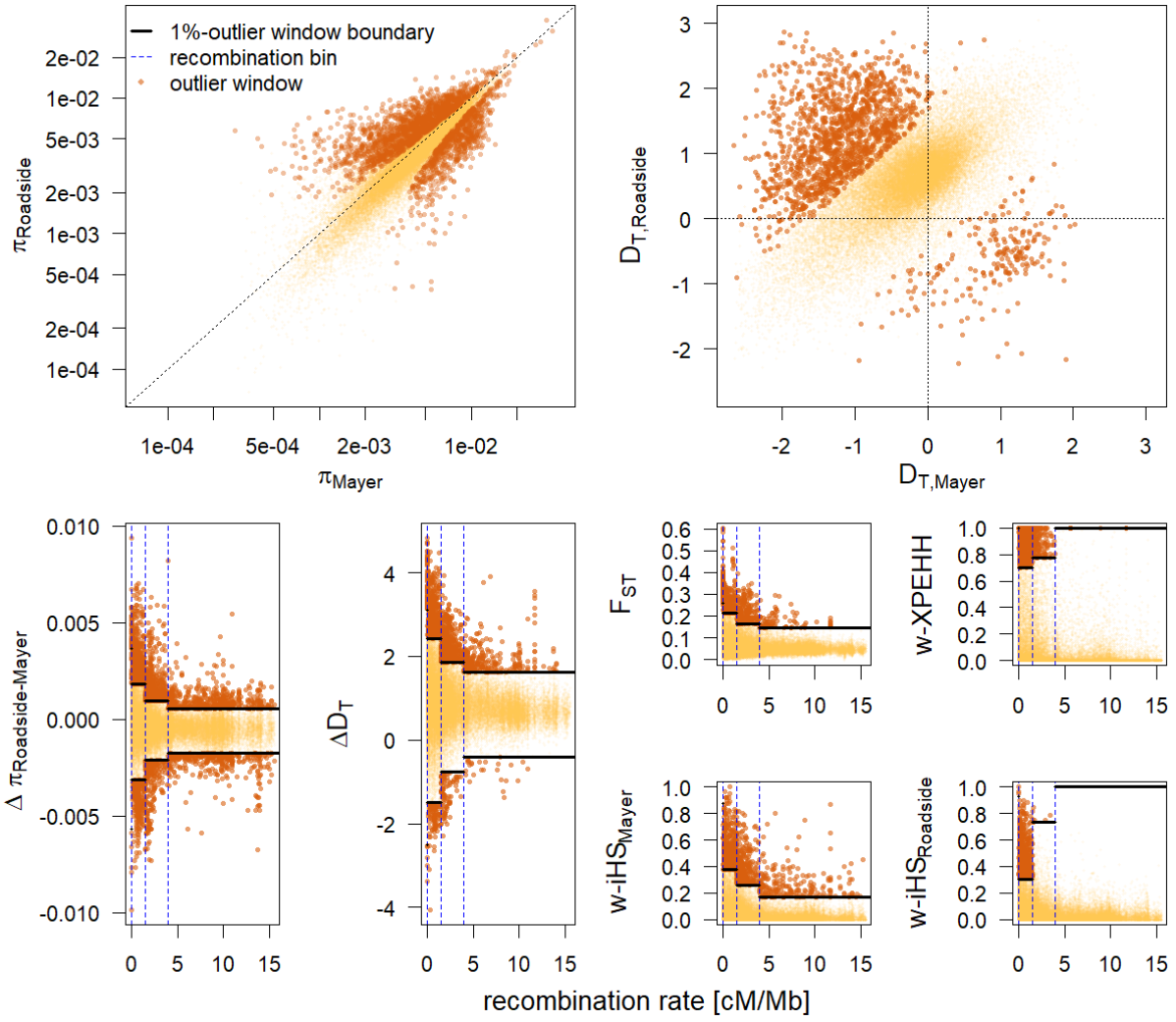
Supplementary Figure 1 | Proportion of fixed differences between two individuals for different SNP categories. The left panel shows the distribution of mean absolute divergence as in Fig. 4 for comparison. The other panels depict the percentage of fixed differences for missense mutations, synonymous mutations, intron, regulatory and intergenic variation.



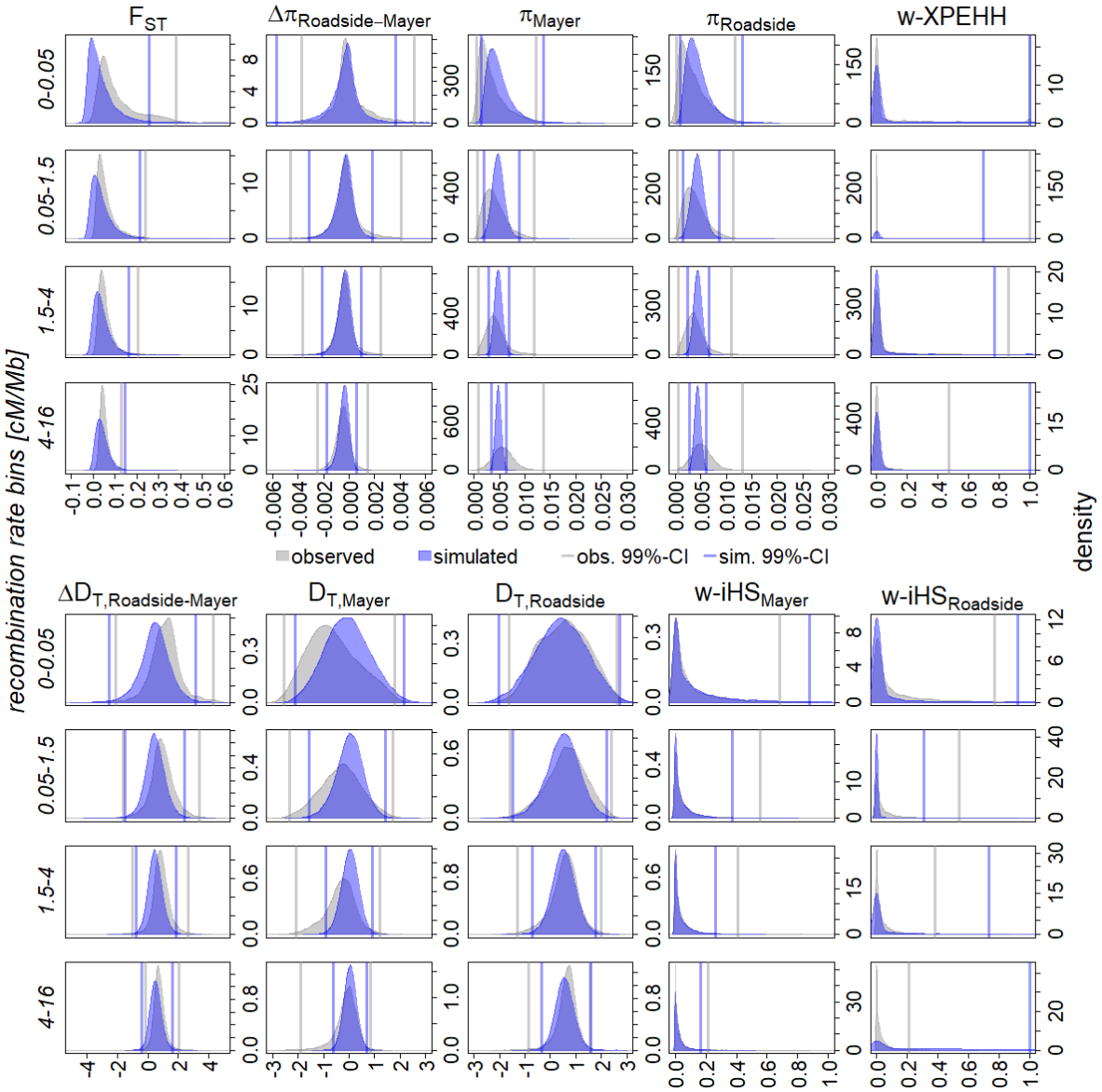
Supplementary Figure 2 | Reconstruction of demographic history based on the 2D-site-frequency spectrum (SFS). a Observed and simulated 2D-SFS under the best demographic model (see below) and marginal 1D-SFS for each population, with entries containing the sum of entries over the folded 2D-SFS for each population. **b** We fit twelve demographic models to the observed folded 2D-SFS between Mayer Lake and Roadside Pond to estimate maximum-likelihood parameters for the simulation of neutral data. The best-fitting demographic model features a bottleneck corresponding to the timing of postglacial colonization and a recent expansion of the transplant population, in line with biogeographic predictions and census data. ΔLL: log-likelihood difference for each model between observed and expected 2D-SFS; ΔAIC: differences between models for the Akaike information criterion.



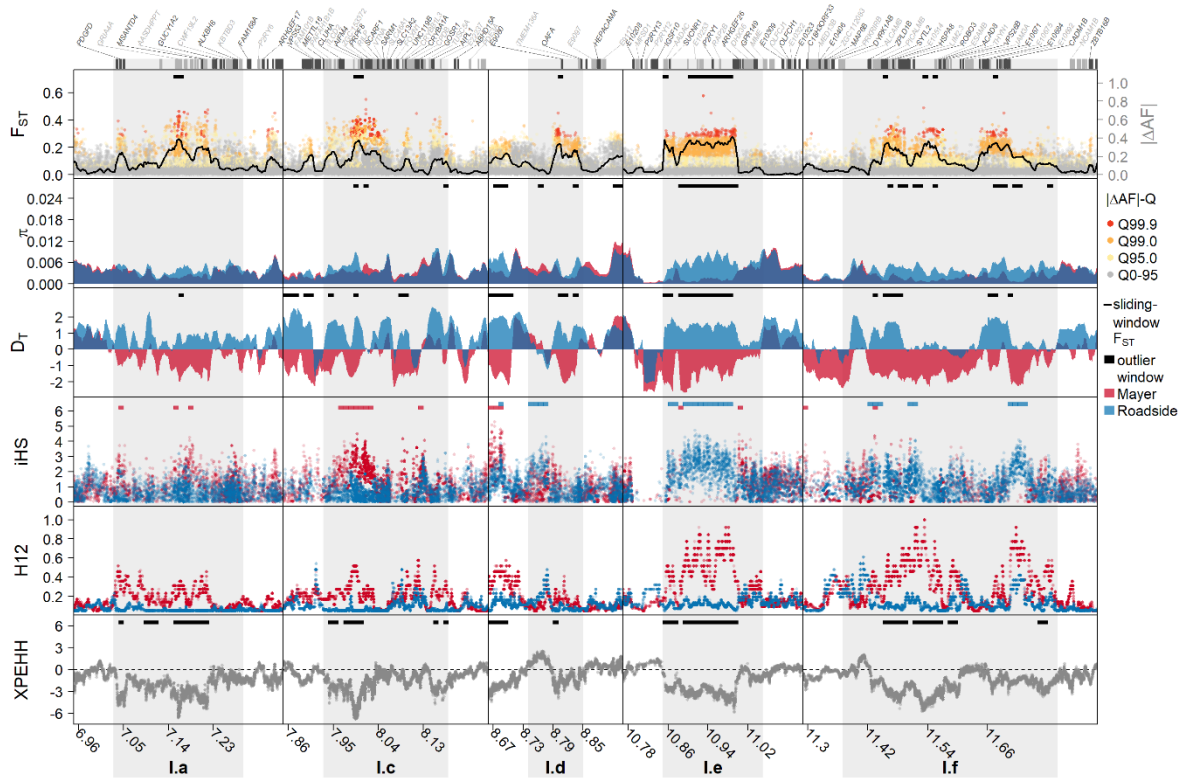
Supplementary Figure 3 | Recombination-rate based cut-offs to identify outlier windows. The lower panels show the distribution of differentiation (F_{ST}), change in diversity ($\Delta\pi$) and Tajima's D (ΔD_T), iHS and XP-EHH and the respective recombination-rate bin based (0–0.05, 0.05–1.5, 1.5–4, >4 cM/Mb) one- or two-sided cut-offs to identify outlier windows (dark brown dots) at the 1%-alpha level against neutral distributions of these statistics (see Supplementary Fig. 4). The upper panels show the correlation of nucleotide diversity and Tajima's D 10 kb window statistics between the two populations with outlier windows highlighted in dark brown. Note the strongly -shifted Tajima's D distribution in Roadside Pond.



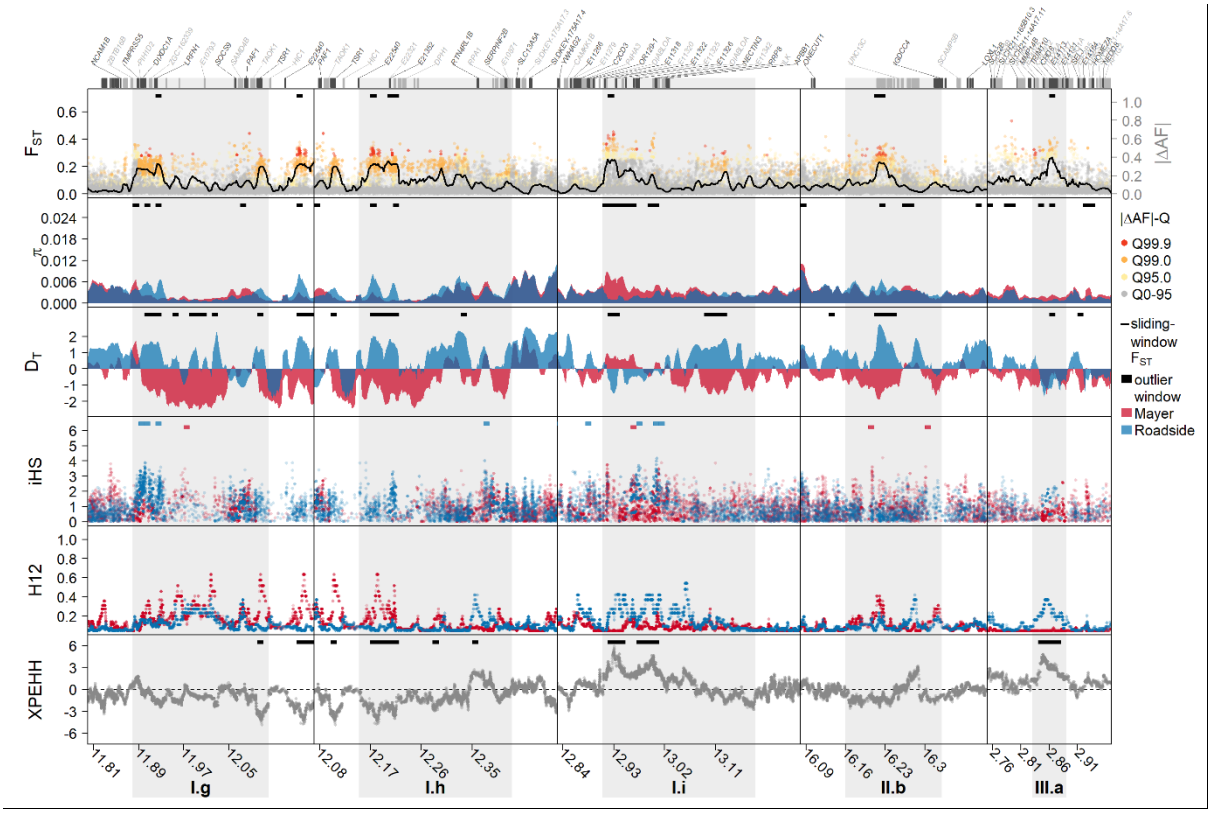
Supplementary Figure 4 | Distribution of observed and simulated summary statistics under the best demographic model. Shown are distributions of 10kb-window statistics in observed and simulated data, for each of four recombination rate bins: differentiation (F_{ST}) between Mayer Lake and Roadside Pond, nucleotide diversity in both populations (π_{Mayer}, $\pi_{Roadside}$) and between populations ($\Delta\pi_{Roadside-Mayer}$), Tajima's D in both populations ($D_{T,Mayer}$, $T_{D,Roadside}$) and between populations ($\Delta D_{T,Roadside-Mayer}$), as well as the haplotype-based selection statistics iHS in each population (w-iHS_{Mayer}, w-$iHS_{Roadside}$) and XP-EHH between populations (w-XP-EHH). Vertical lines depict boundaries for the 1% most extreme windows in the simulated (blue) and observed (grey) data.



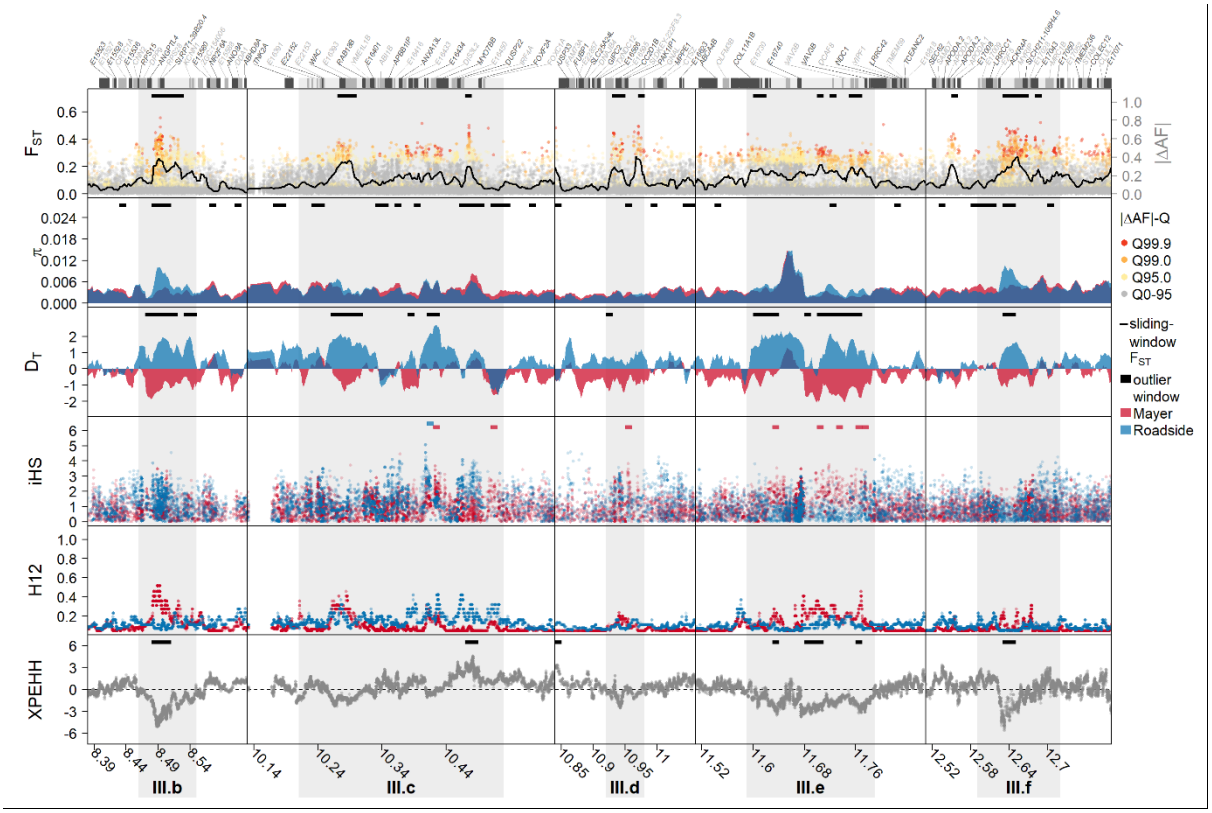
Supplementary Figure 5 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



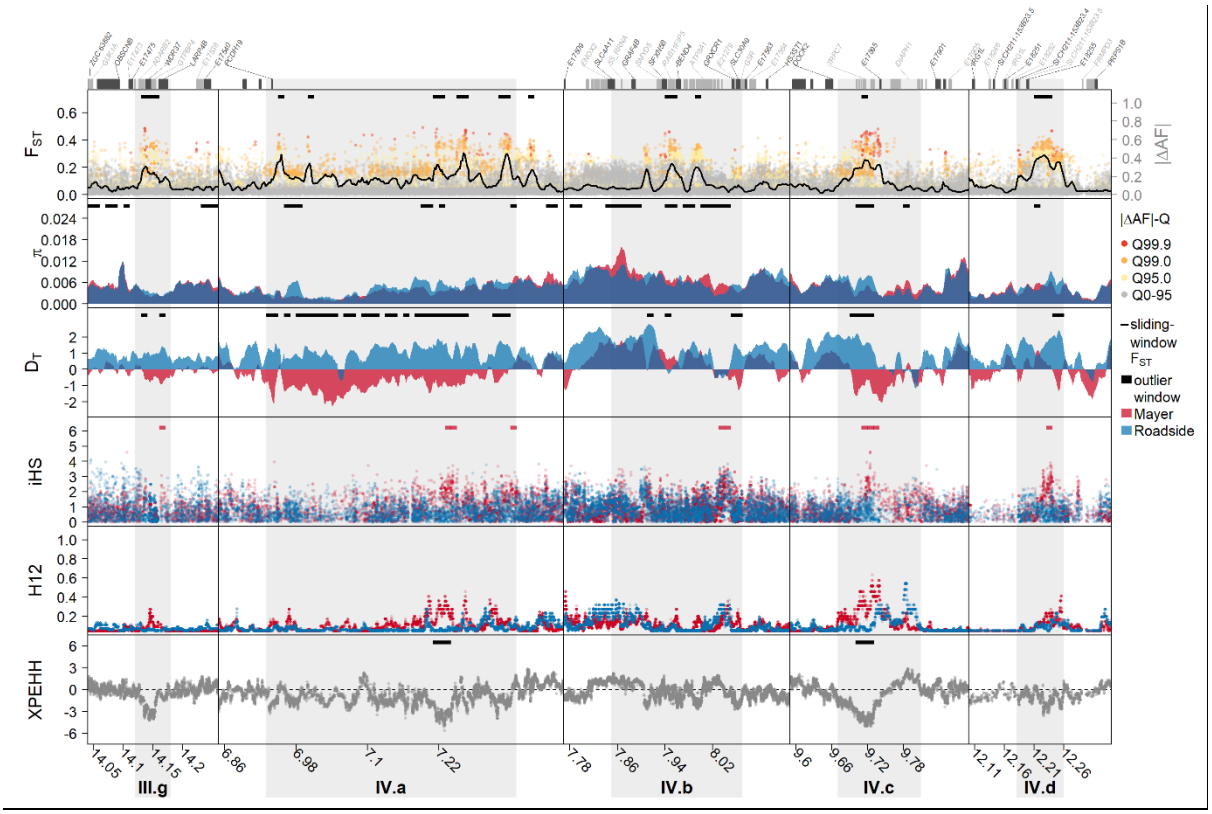
Supplementary Figure 6 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



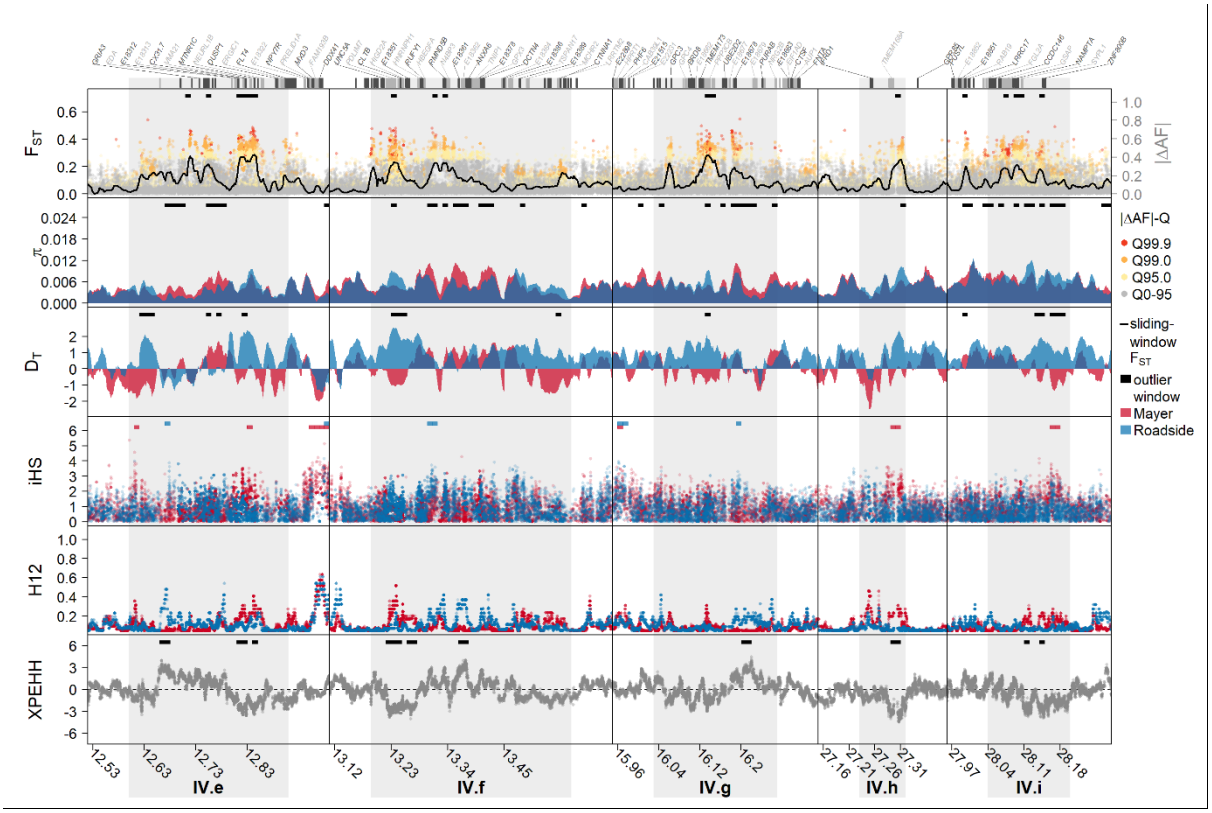
Supplementary Figure 7 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



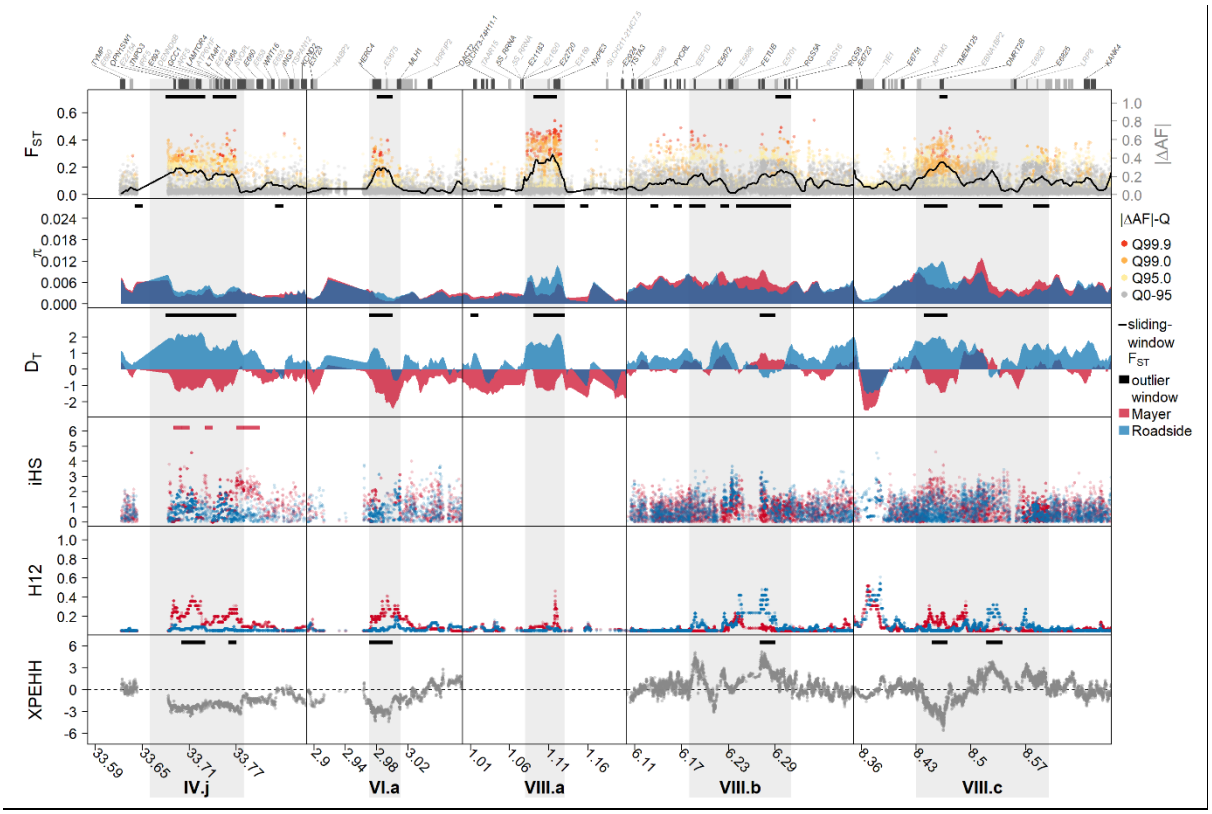
Supplementary Figure 8 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



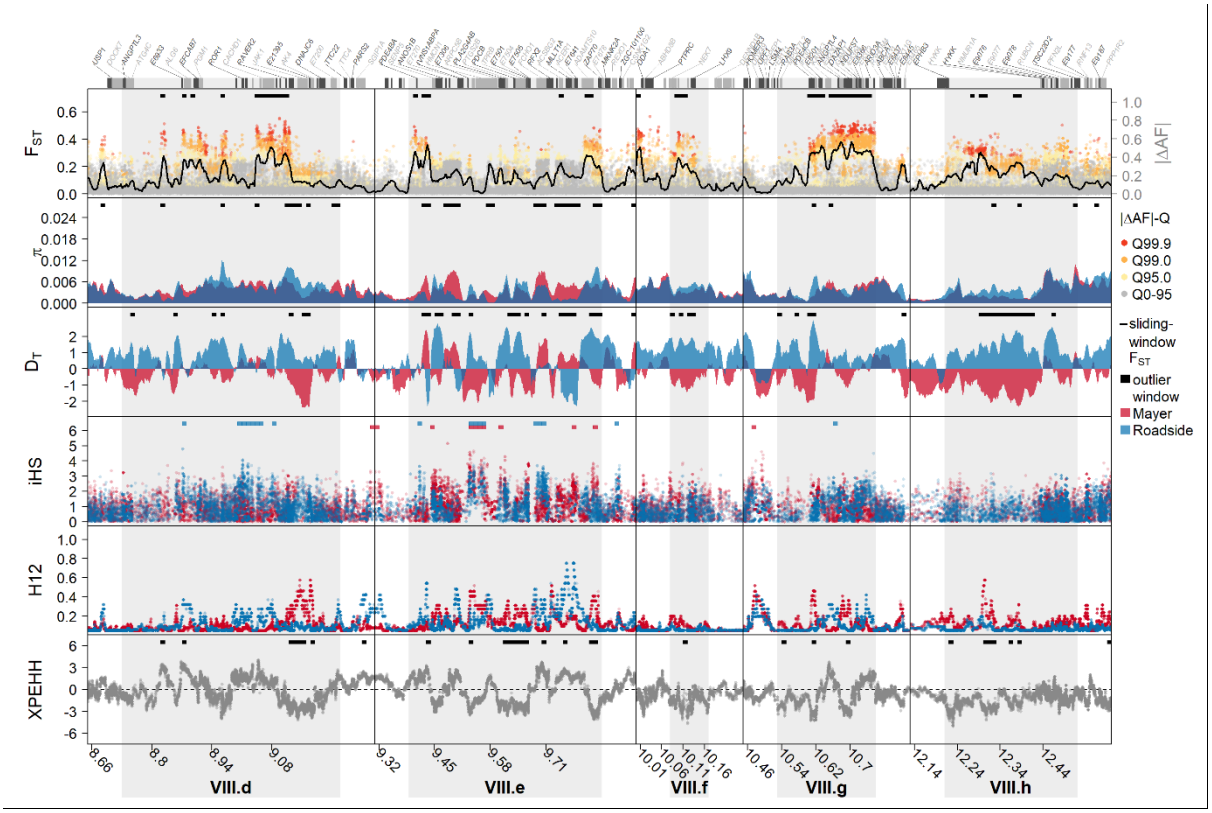
Supplementary Figure 9 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



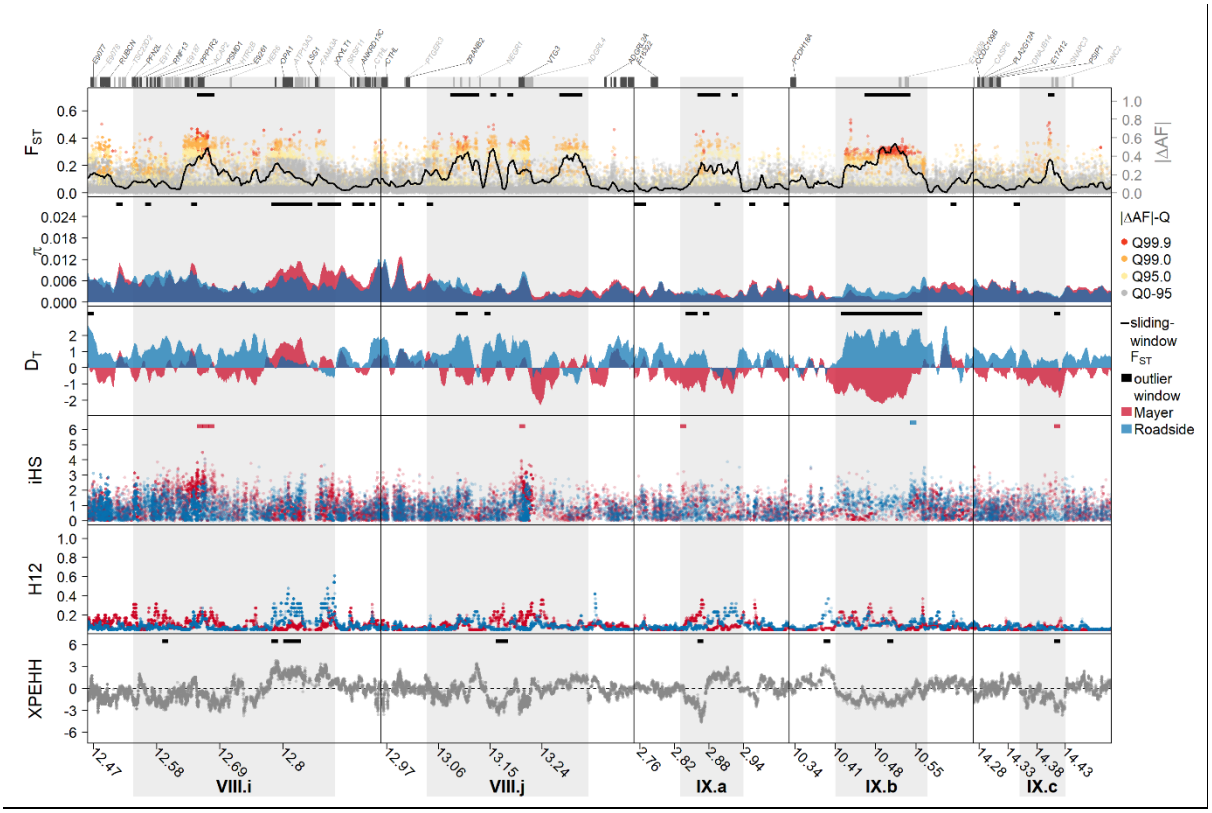
Supplementary Figure 10 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



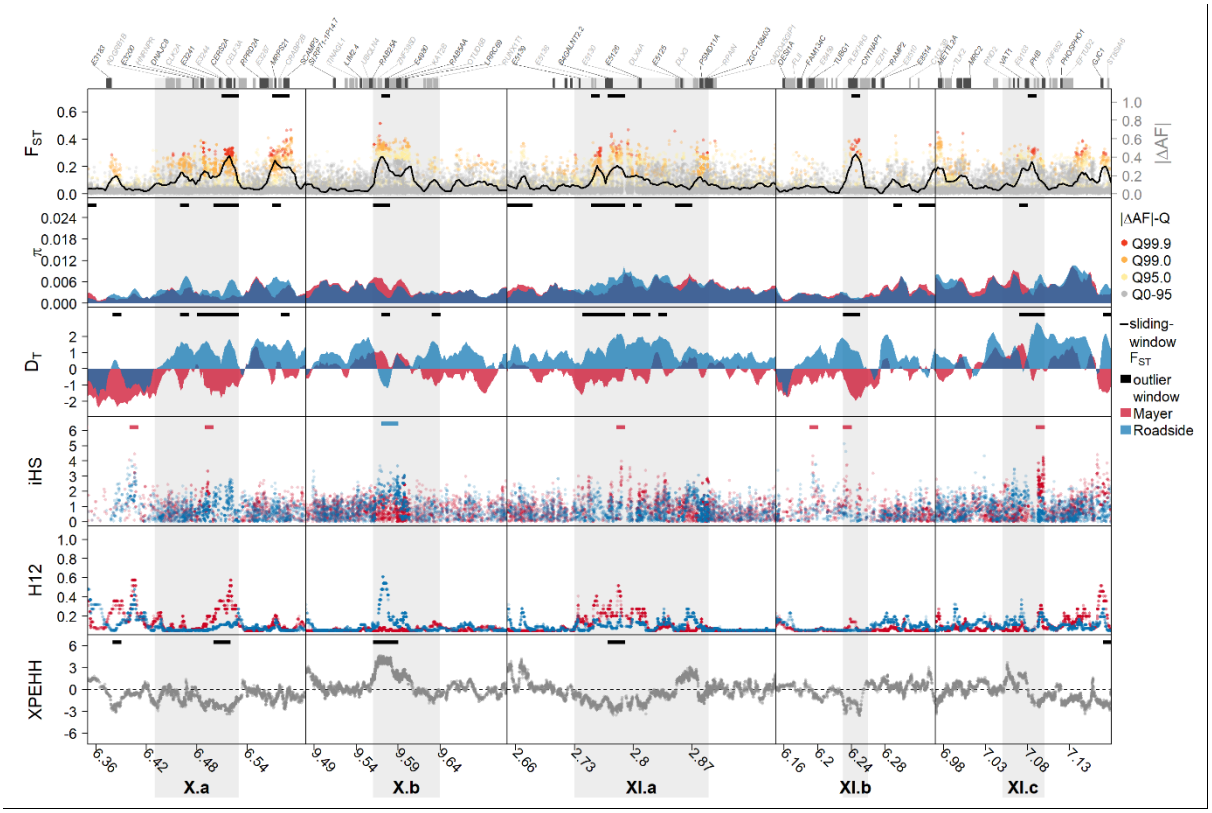
Supplementary Figure 11 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



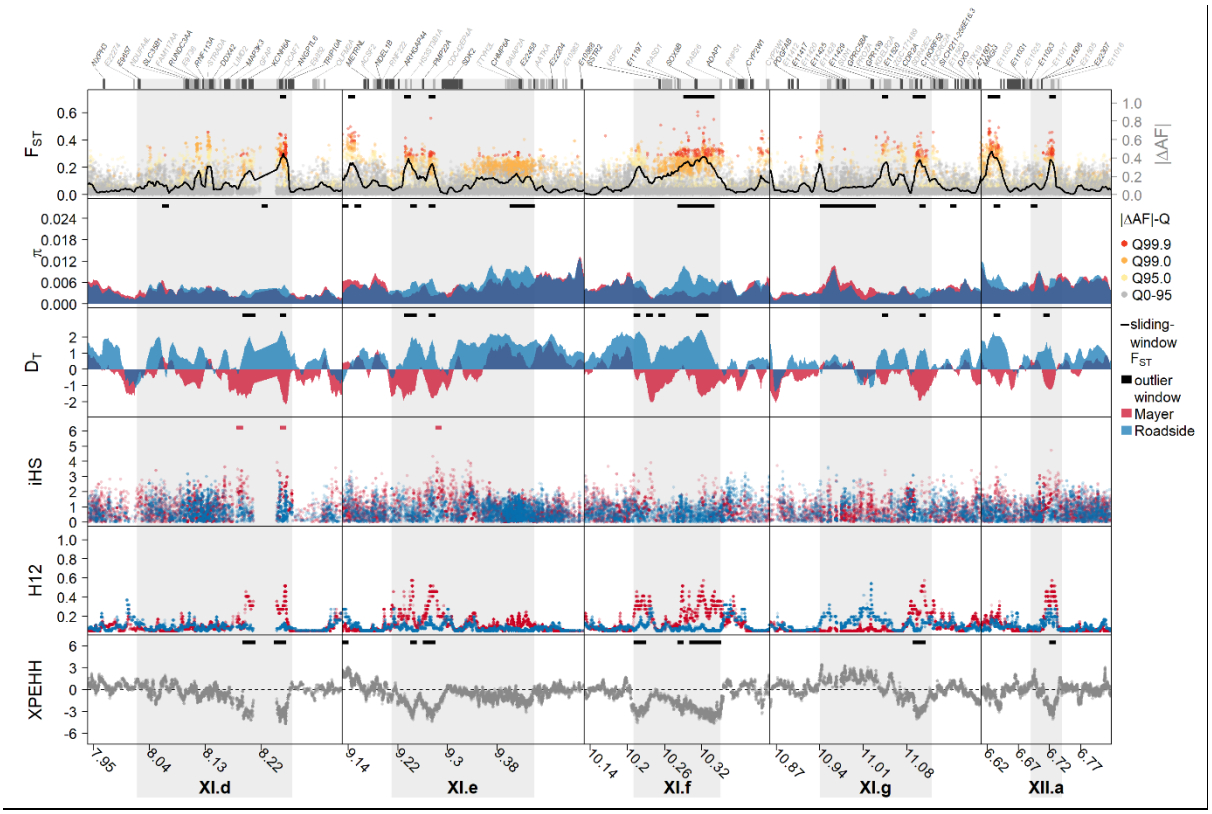
Supplementary Figure 12 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



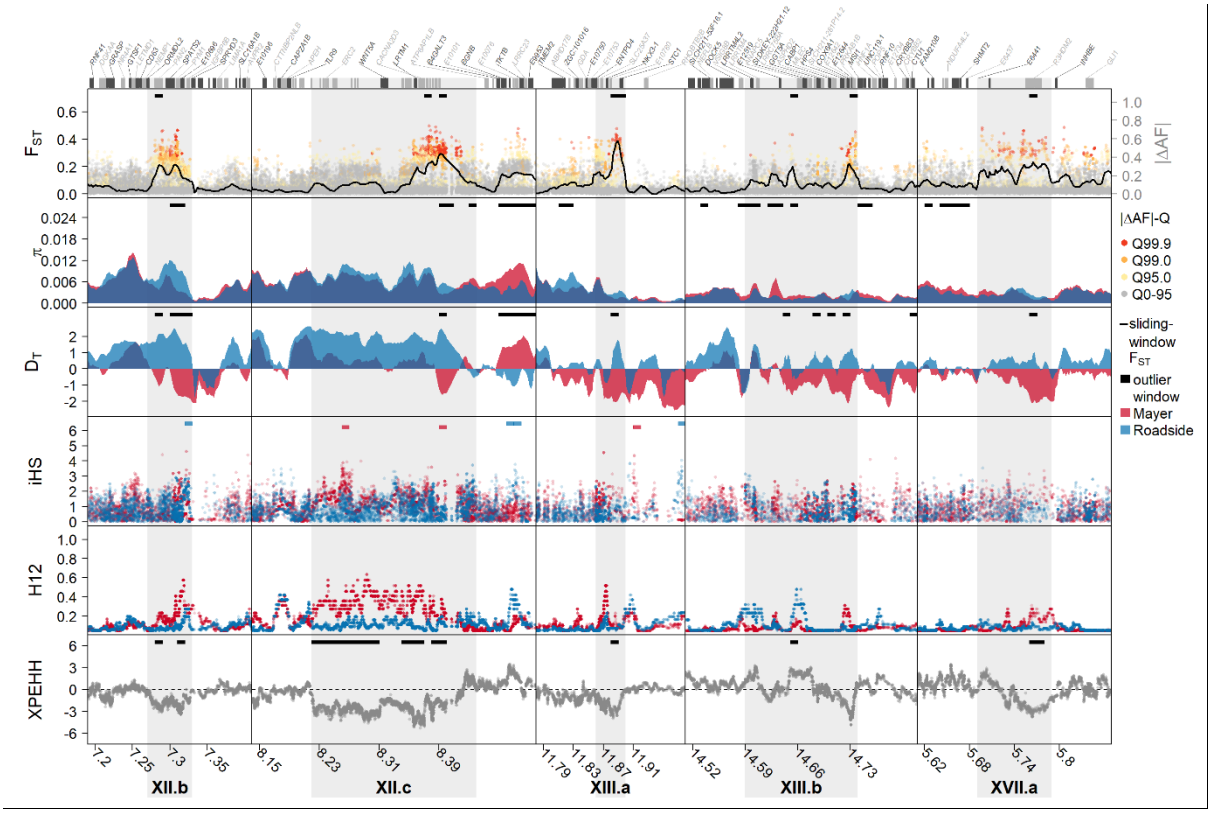
Supplementary Figure 13 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



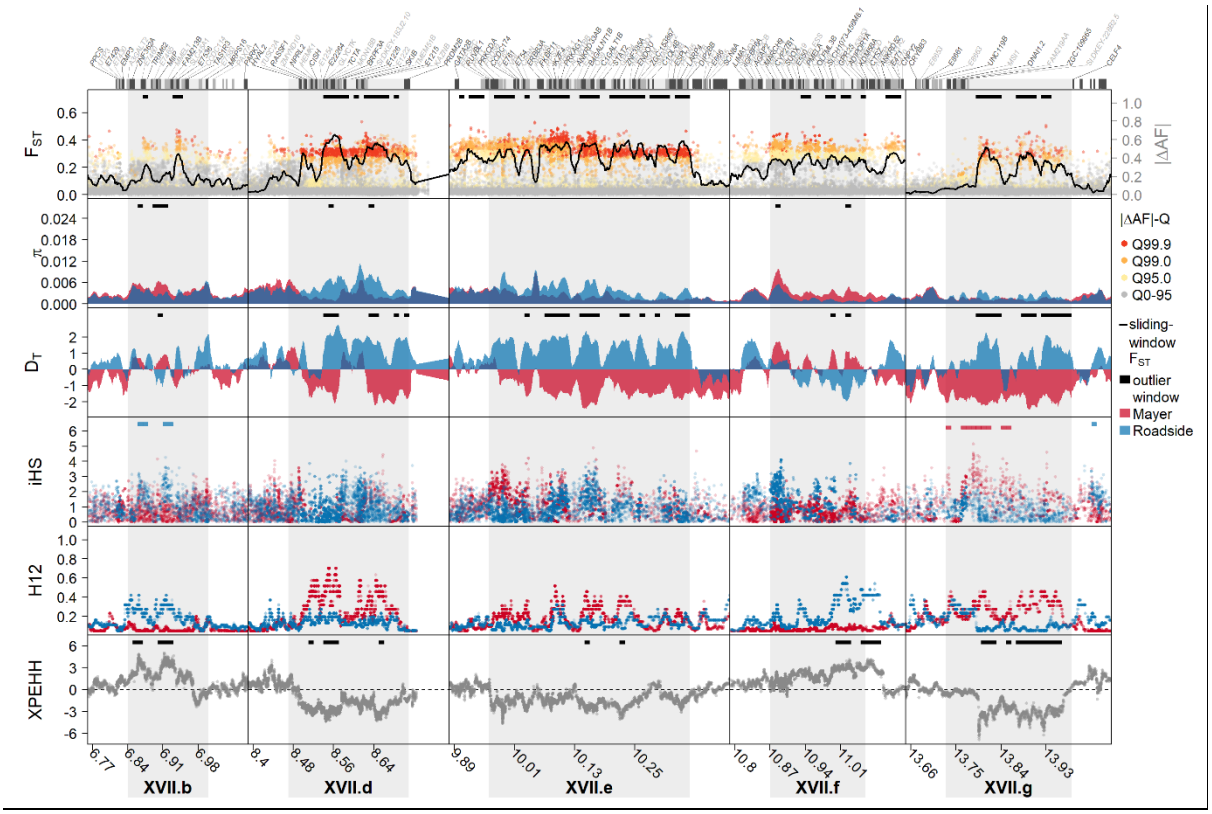
Supplementary Figure 14 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



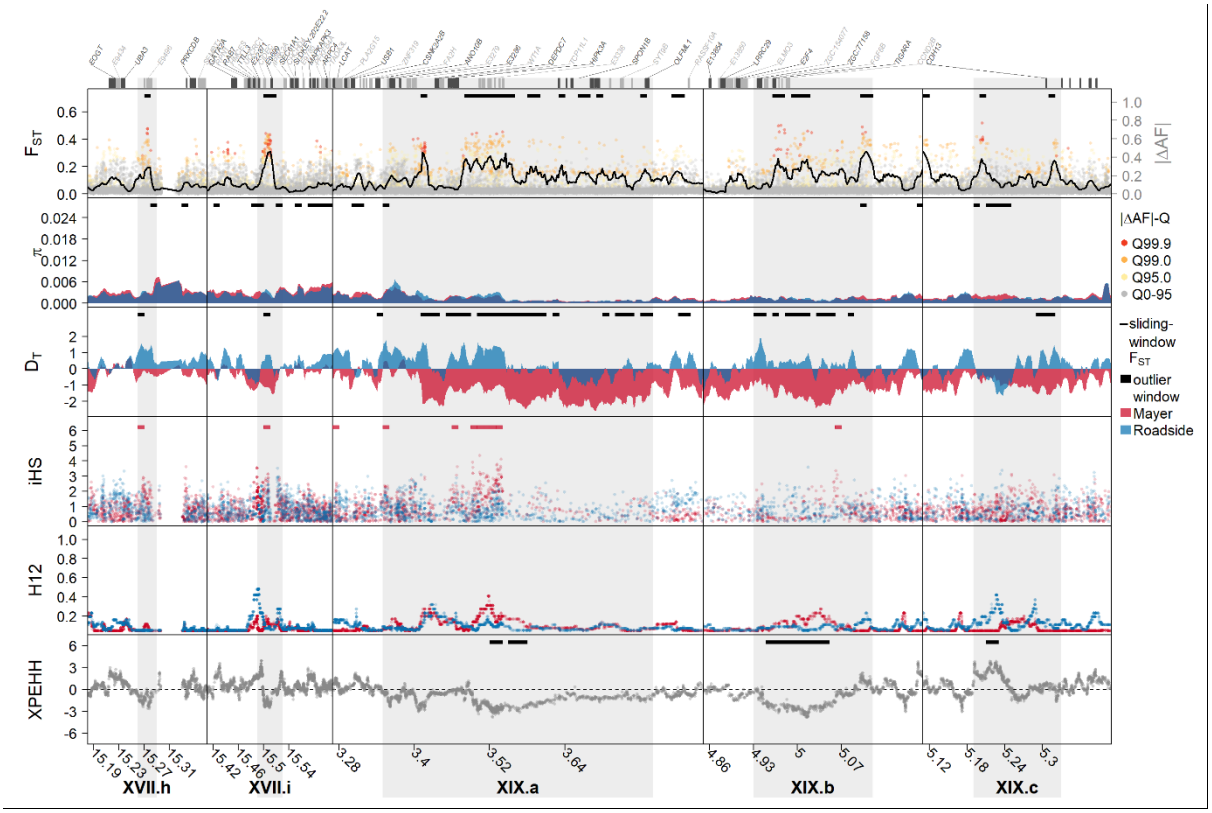
Supplementary Figure 15 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



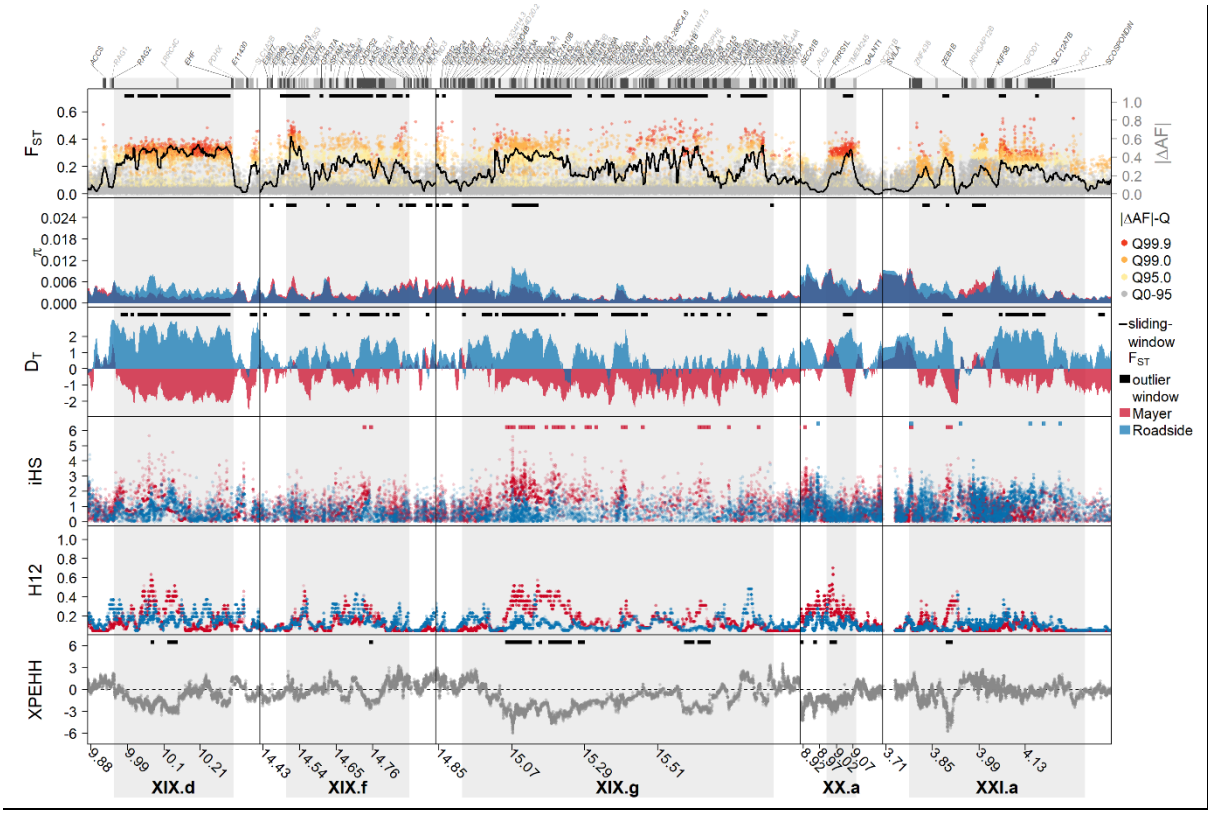
Supplementary Figure 16 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



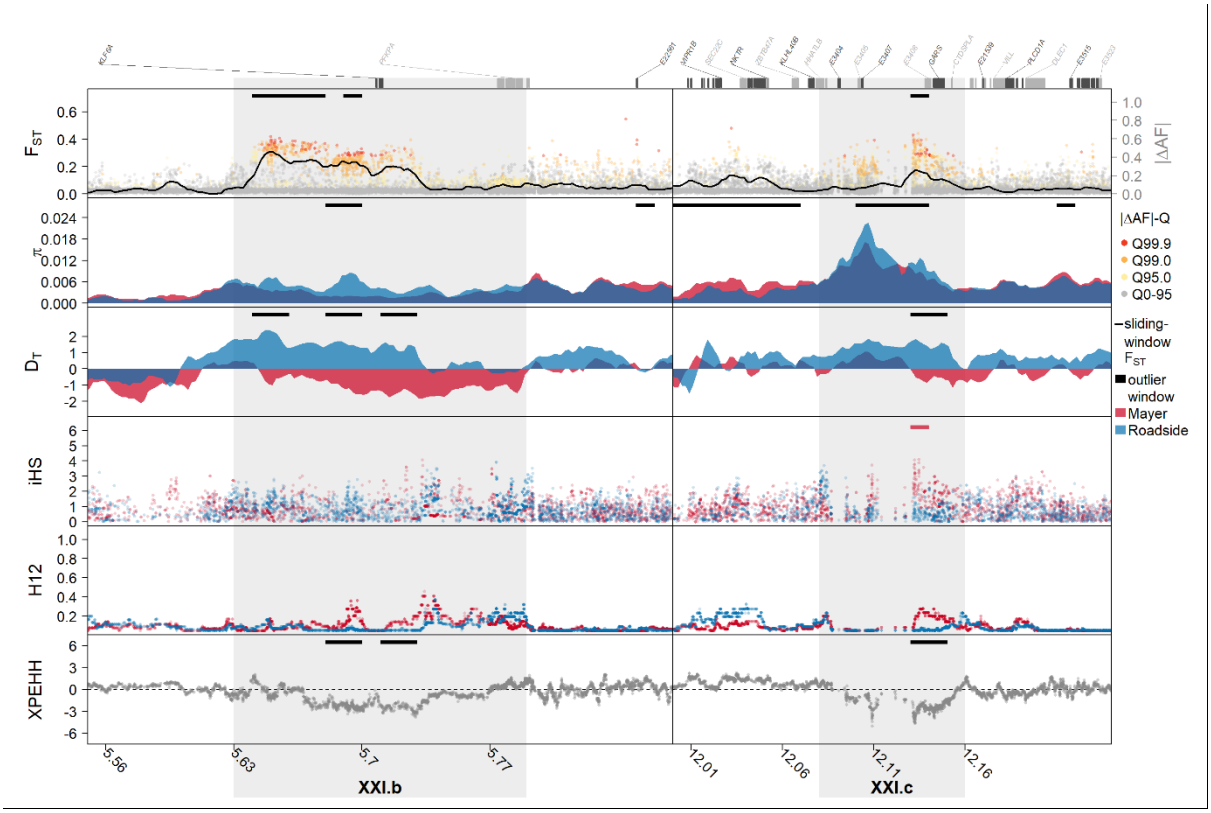
Supplementary Figure 17 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



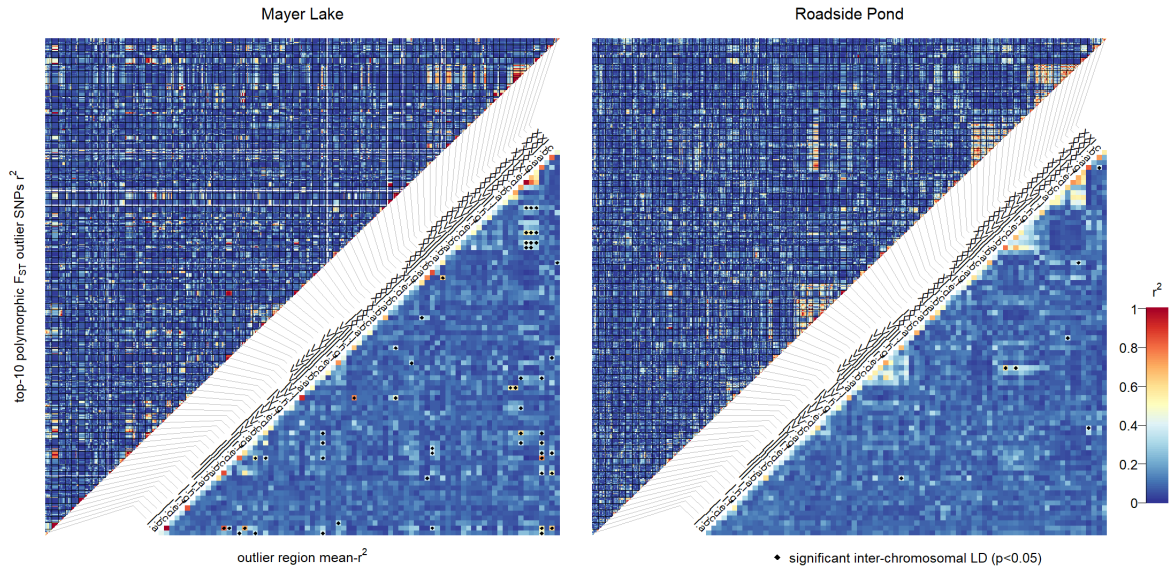
Supplementary Figure 18 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



Supplementary Figure 19 | Local signatures of divergent selection on candidate genes. See Fig. 3 legend.



Supplementary Figure 20 | Linkage disequilibrium (LD) between outlier regions in the selection experiment. Several outlier regions in the source population, Mayer Lake, show higher LD between unlinked outlier regions than expected by chance from inter-chromosomal LD between neutral SNPs. In contrast, only a few outlier regions show inter-chromosomal LD in the transplant population, suggesting selection on many genomic regions instead of selection on one genomic region and a hitchhiking genomic background.



Supplementary Data

Supplementary Data 1 | Twelve demographic models fit to the observed data. Specification of the models (.tpl) and parameter (.est) for each of the 12 tested demographic models used with fastsimcoal2 v2.6 (see Methods section). See Supplementary Fig. 2 for a visual representation of the models.

MRCC.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
200
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
1 historical event
13 1 0 1 1 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRCC.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
[RULES]
[COMPLEX PARAMETERS]
```

MRCB.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
2 historical event
7 1 1 1 RES1 0 0
13 1 0 1 1 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRCB.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
1 NRDSP unif 1000 10000 output
[RULES]
[COMPLEX PARAMETERS]
0 RES1 = 200/NRDSP
```

MRCE.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
RRAT
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
2 historical event
13 1 1 1 1 0 0
13 1 0 1 1 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRCE.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
1 NRDSP unif 1000 10000 output
[RULES]
[COMPLEX PARAMETERS]
0 RATBB = 200/NRDSP hide
0 RCOF = log(RATBB) hide
0 RRAT = RCOF/13 hide
```

MRGC.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
200
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
```

```
3 historical event
13 1 0 1 1 0 0
TMAX 0 0 1 1 GRAT 0
TCOL 0 0 1 1 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRGC.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
1 NANC unif 1000 100000 output
1 TCOL unif 3000 10000 output
0 PROP unif 0.1 0.9 hide
[RULES]
[COMPLEX PARAMETERS]
1 TMAX = TCOL*PROP output
1 TDIV = TCOL-TMAX hide
0 RATIO = NANC/NMAYR hide
0 RTEA = log(RATIO) hide
0 GRAT = RTEA/TDIV hide
```

MRGB.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
4 historical event
7 1 1 1 RES1 0 0
13 1 0 1 1 0 0
TMAX 0 0 1 1 GRAT 0
TCOL 0 0 1 1 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRGB.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
1 NRDSP unif 1000 10000 output
1 NANC unif 1000 100000 output
1 TCOL unif 3000 10000 output
0 PROP unif 0.1 0.9 hide
```

```
[RULES]
[COMPLEX PARAMETERS]
1 TMAX = TCOL*PROP output
1 TDIV = TCOL-TMAX hide
0 RATIO = NANC/NMAYR hide
0 RTEA = log(RATIO) hide
0 GRAT = RTEA/TDIV hide
0 RES1 = 200/NRDSP
```

MRGE.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
RRAT
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
4 historical event
13 1 1 1 1 0 0
13 1 0 1 1 0 0
TMAX 0 0 1 1 GRAT 0
TCOL 0 0 1 1 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRGE.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
1 NRDSP unif 1000 10000 output
1 NANC unif 1000 100000 output
1 TCOL unif 3000 10000 output
0 PROP unif 0.1 0.9 hide
[RULES]
[COMPLEX PARAMETERS]
1 TMAX = TCOL*PROP output
1 TDIV = TCOL-TMAX hide
0 RATIO = NANC/NMAYR hide
0 RTEA = log(RATIO) hide
0 GRAT = RTEA/TDIV hide
0 RATBB = 200/NRDSP hide
0 RCOF = log(RATBB) hide
0 RRAT = RCOF/13 hide
```

MRBC.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
200
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
```

```
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
3 historical event
13 1 0 1 1 0 0
TEND 0 0 1 RES2 0 0
TBOT 0 0 1 BTST 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRBC.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 10000 100000 output
1 NANC unif 10000 1000000 output
1 BTST unif 2 100 output
1 TBOT unif 3000 10000 output
1 LEBO unif 20 200
[RULES]
[COMPLEX PARAMETERS]
1 TEND = TBOT-LEBO output
1 NCOL = NANC/BTST output
0 RES2 = NCOL/NMAYR hide
0 RES1 = NMAYR/NCOL hide
```

MRBB.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
4 historical event
7 1 1 1 RES1 0 0
13 1 0 1 1 0 0
TEND 0 0 1 RES2 0 0
TBOT 0 0 1 BTST 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRBB.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
```

```
//all Ns are in number of haploid individuals
1 NMAYR unif 1000 100000 output
1 NRDSP unif 1000 10000 output
1 NANC unif 1000 100000 output
1 BTST unif 2 100 output
1 TBOT unif 3000 10000 output
1 LEBO unif 20 200
[RULES]
[COMPLEX PARAMETERS]
1 TEND = TBOT-LEBO output
1 NCOL = NANC/BTST output
0 RES2 = NCOL/NMAYR hide
0 RES1 = 200/NRDSP hide
```

MRBE.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
RRAT
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
4 historical event
13 1 1 1 1 0 0
13 1 0 1 1 0 0
TEND 0 0 1 RES2 0 0
TBOT 0 0 1 BTST 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MRBE.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 10000 100000 output
1 NRDSP unif 1000 10000 output
1 NANC unif 10000 1000000 output
1 BTST unif 2 100 output
1 TBOT unif 3000 10000 output
1 LEBO unif 20 200
[RULES]
[COMPLEX PARAMETERS]
1 TEND = TBOT-LEBO output
1 NCOL = NANC/BTST output
0 RES2 = NCOL/NMAYR hide
0 RATBB = 200/NRDSP hide
0 RCOF = log(RATBB) hide
0 RRAT = RCOF/13 hide
```

MREC.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
200
```

```
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
3 historical event
13 1 0 1 1 0 0
TMAX 0 0 1 1 GRAT 0
TBOT 0 0 1 BTST 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MREC.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 10000 100000 output
1 NANC unif 10000 1000000 output
1 BTST unif 2 100 output
1 TBOT unif 3000 10000 output
0 PROP unif 0.1 0.9 hide
[RULES]
[COMPLEX PARAMETERS]
1 TMAX = TBOT*PROP output
1 NCOL = NANC/BTST output
0 RES1 = NMAYR/(NANC/BTST) hide
1 TDIV = TBOT-TMAX hide
0 RATIO = (NANC/BTST)/NMAYR hide
0 RTEA = log(RATIO) hide
0 GRAT = RTEA/TDIV hide
```

MREB.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
0
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
4 historical event
7 1 1 1 RES1 0 0
13 1 0 1 1 0 0
TMAX 0 0 1 1 GRAT 0
TBOT 0 0 1 BTST 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MREB.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 10000 100000 output
1 NRDSP unif 1000 10000 output
1 NANC unif 10000 1000000 output
1 BTST unif 2 100 output
1 TBOT unif 3000 10000 output
0 PROP unif 0.1 0.9 hide
[RULES]
[COMPLEX PARAMETERS]
1 TMAX = TBOT*PROP output
1 NCOL = NANC/BTST output
1 TDIV = TBOT-TMAX hide
0 RATIO = (NANC/BTST)/NMAYR hide
0 RTEA = log(RATIO) hide
0 GRAT = RTEA/TDIV hide
0 RATBB = 200/NRDSP hide
0 RCOF = log(RATBB) hide
0 RRAT = RCOF/13 hide
0 RES1 = 200/NRDSP
```

MREE.tpl

```
//Parameters for the coalescence simulation program : simcoal.exe
2
//Population effective sizes (number of genes)
NMAYR
NRDSP
//Samples sizes and samples age
24 5
22
//Growth rates: negative growth implies population expansion
0
RRAT
//Number of migration matrices : 0 implies no migration between demes
0
//historical event: time, source, sink, migrants, new deme size, new growth rate, migration
matrix index
4 historical event
13 1 1 1 1 0 0
13 1 0 1 1 0 0
TMAX 0 0 1 1 GRAT 0
TBOT 0 0 1 BTST 0 0
//Number of independent loci [chromosome]
1 0
//Per chromosome: Number of contiguous linkage Block: a block is a set of contiguous loci
1
//per Block:data type, number of loci, per generation recombination and mutation rates and
optional parameters
FREQ 1 0 1.7e-8 OUTEXP
```

MREE.est

```
// Priors and rules file
// *********************
[PARAMETERS]
//#isInt? #name #dist.#min #max
//all Ns are in number of haploid individuals
1 NMAYR unif 10000 100000 output
1 NRDSP unif 1000 10000 output
1 NANC unif 10000 1000000 output
1 BTST unif 2 100 output
1 TBOT unif 3000 10000 output
0 PROP unif 0.1 0.9 hide
[RULES]
[COMPLEX PARAMETERS]
1 TMAX = TBOT*PROP output
1 NCOL = NANC/BTST output
```

```
0 RES1 = NMAYR/(NANC/BTST) hide
1 TDIV = TBOT-TMAX hide
0 RATIO = (NANC/BTST)/NMAYR hide
0 RTEA = log(RATIO) hide
0 GRAT = RTEA/TDIV hide
0 RATBB = 200/NRDSP hide
0 RCOF = log(RATBB) hide
0 RRAT = RCOF/13 hide
```