EVOLUTIONARY BIOLOGY

Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast

Hernán E. Morales[1]*†, Rui Faria[2], Kerstin Johannesson[3], Tomas Larsson[1,4], Marina Panova[3], Anja M. Westram[2,5]‡, Roger K. Butlin[2,3]‡

The study of parallel ecological divergence provides important clues to the operation of natural selection. Parallel divergence often occurs in heterogeneous environments with different kinds of environmental gradients in different locations, but the genomic basis underlying this process is unknown. We investigated the genomics of rapid parallel adaptation in the marine snail *Littorina saxatilis* in response to two independent environmental axes (crab-predation versus wave-action and low-shore versus high-shore). Using pooled whole-genome resequencing, we show that sharing of genomic regions of high differentiation between environments is generally low but increases at smaller spatial scales. We identify different shared genomic regions of divergence for each environmental axis and show that most of these regions overlap with candidate chromosomal inversions. Several inversion regions are divergent and polymorphic across many localities. We argue that chromosomal inversions could store shared variation that fuels rapid parallel adaptation to heterogeneous environments, possibly as balanced polymorphism shared by adaptive gene flow.

INTRODUCTION

Uncovering the evolutionary drivers of adaptive divergence is of central importance to understanding how biodiversity is generated and maintained. Cases where phenotypic differentiation has evolved multiple times in response to similar environmental contrasts (i.e., parallel ecological divergence) are considered strong evidence for the consistent operation of natural selection (1). However, environmental variation is often complex, potentially causing replicated populations to evolve along trajectories that are not fully parallel but rather vary in the magnitude or direction of phenotypic divergence (2–4). At the genomic level, it is not clear how often parallelism results from the same underlying genetic changes in the face of environmental heterogeneity (5, 6). Sharing of loci underlying parallel divergence (i.e., genetic sharing) is expected to increase with decreasing evolutionary distance [according to meta-analysis; (6)] and with decreasing geographic distance [according to modeling; (7)]. This is because populations with more recent co-ancestry and higher connectivity can access a common pool of genetic variation, promoting both rapid local adaptation and genetic sharing. However, there are several examples in which parallel ecological divergence evolved rapidly despite exhibiting apparently low genetic sharing (8–10). Multiple factors are likely to modify the amount of genetic sharing, including the congruence in direction for different environmental drivers across locations (e.g., Fig. 1B) and the underlying genomic architecture (2, 3, 5, 11).

Selective pressures that are location specific, or act in different directions in each locality, are expected to lead to nonparallel patterns of phenotypic divergence (3). Thus, although environmental variability could influence the evolution of parallel divergence, this possibility has rarely been addressed directly [but see (12, 13)]. For example, Stuart *et al.* (12) used replicated pairs of lake-stream stickleback populations to show deviations from phenotypic parallel divergence as a function of between-site environmental variation. The magnitude of the deviation was largely explained by the amount of gene flow between sites, but the genetic divergence across the genome in relation to environmental variation was not quantified. Raeymaekers *et al.* (13) found patterns of shared and unique genomic divergence between two related species of sticklebacks (three-spined and nine-spined) in response to a common environmental landscape. The genome-wide variation in both species was partially explained by spatial and environmental drivers, but a dissection of the effects of these drivers across the genomic landscape was not possible. In this matter, the genomic basis of parallel adaptation is known to be largely dependent on the underlying genomic architecture (14), i.e., the number, effect sizes, and additivity of loci and their ordering within the genome into nonrandom arrangements (15, 16). However, it is unclear how genomic architecture interacts with environmental heterogeneity to drive the evolution of parallel divergence.

An important feature of genomic architecture is chromosomal inversions, which suppress recombination in heterozygotes and may facilitate parallel adaptive divergence by maintaining sets of co-adapted alleles across locations (17–19). Inversions can act as reservoirs of adaptive standing variation and as vehicles to distribute adaptive variation via gene flow, thus promoting parallel divergence with a shared genetic basis (20–22). Chromosomal rearrangements have been identified in studies of parallel evolution [e.g., (11, 23)], but the extent to which they contribute to shared genetic differentiation across multiple axes of parallel environmental divergence remains an open question.

Here, we assess how environmental variation and genomic architecture influence the genetic basis of rapid parallel ecological divergence in the intertidal snail *Littorina saxatilis* (5, 24, 25). This species is broadly distributed across the North Atlantic and exhibits a series of life history traits conducive to the maintenance of adaptive

[1]Centre for Marine Evolutionary Biology, Department of Marine Sciences, University of Gothenburg, Göteborg, Sweden. [2]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK. [3]Centre for Marine Evolutionary Biology, Department of Marine Sciences at Tjärnö, University of Gothenburg, Strömstad, Sweden. [4]Department of Cell and Molecular Biology, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Uppsala University, Husargatan 3, SE-752 37 Uppsala, Sweden. [5]IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria.
*Corresponding author. Email: hernanm@sund.ku.dk
†Present address: Evolutionary Genetics Section, Globe Institute, University of Copenhagen, Copenhagen, Denmark.
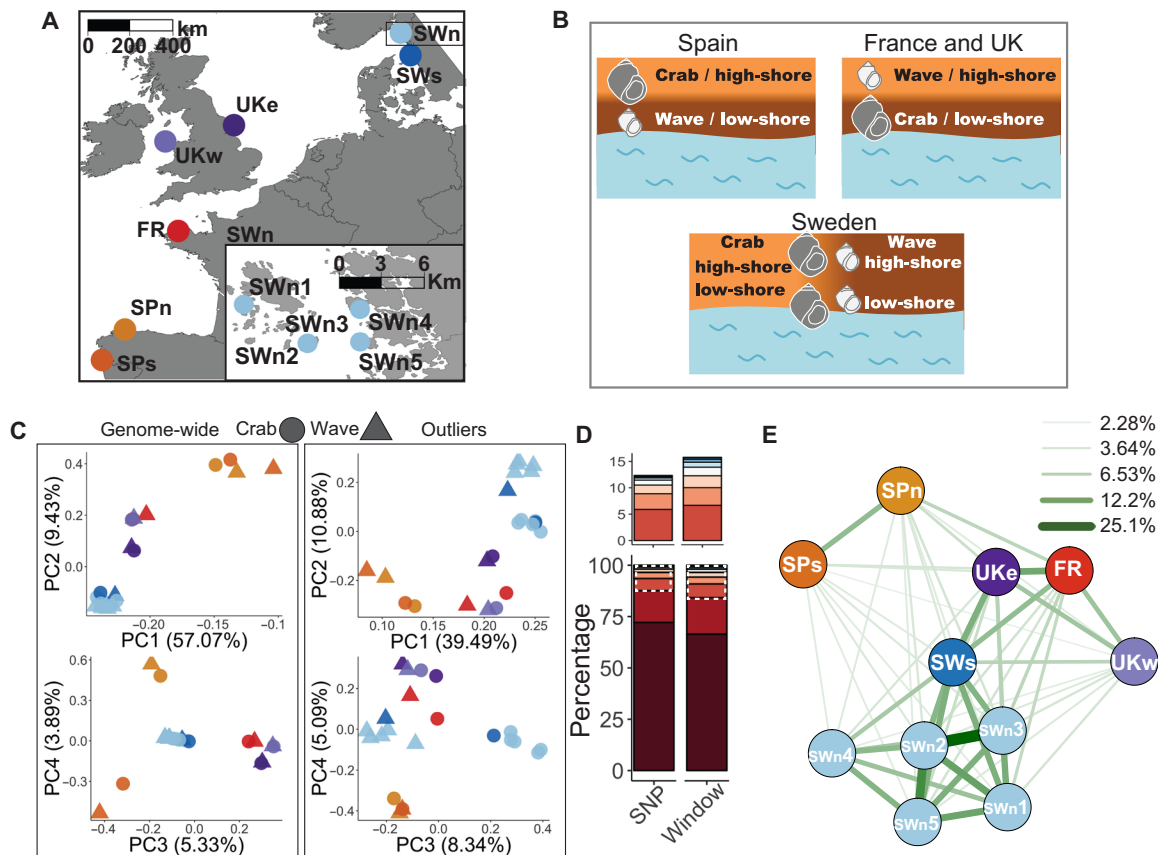‡These authors contributed equally to this work.

Fig. 1. Samples, genetic structure, and Crab-Wave genetic differentiation and sharing. (**A**) The 11 localities sampled at different geographical scales. SWn2 and SWn3 were sampled in different parts of the same island (<1 km). (**B**) Cartoon depicting the relative directions of the two axes of environmental divergence, Crab-Wave and Low-High, which are opposing between Spain and United Kingdom/France and orthogonal within Sweden (details in the Supplementary Materials). (**C**) PCA summarizing genetic variation for 10,263,736 genome-wide SNPs (left) and for 705,786 highly Crab-Wave differentiated SNPs (i.e., outliers) identified across all localities (right). (**D**) Percentage of outlier sharing at SNP (n = 705,786) and 500-bp window (n = 22,315) levels. The color ramp represents the number of localities that share outliers [from 1 (dark red, unique outlier) to 11 (blue, fully shared outlier)], and the dashed white rectangle shows a zoom-in of the top section. (**E**) Network plot showing the percentage of pairwise SNP outlier sharing between localities, with thicker lines representing more sharing. The correlation between pairwise sharing and geographic distance was highly significant. The network plot of window-level sharing shows qualitatively similar results (see fig. S7).

variation [reviewed in (24)]. Traits such as low dispersal due to restricted adult movement (26), internal fertilization and direct development (27), and the evolution of habitat choice [at least in Spain; (28)] might restrict extensive admixture. Moreover, large effective population sizes [high population densities, sperm storage, and multiple paternity; (29)] could lead to high rates of de novo variation and to effective natural selection (26), including the maintenance of balanced polymorphisms observed in *L. saxatilis* (22, 30, 31).

L. saxatilis exhibits a distinctive pattern of replicated parallel phenotypic divergence, replicated at scales from meters to its entire range. Parallel divergence in *L. saxatilis* has evolved in response to at least two independent environmental axes. The main axis of divergence is between habitats dominated by crab-predation and those dominated by wave-action (Fig. 1B and Supplementary Materials). In areas of the shore where crab predation is high, *L. saxatilis* has evolved thick, large shells with a small aperture for the foot and wary behavior. In areas of the shore exposed to strong wave energy, snails have evolved thin, small shells with a large aperture for the foot and bold behavior (5, 24). Across the species range, Crab and Wave forms typically show sharp habitat transitions of a few meters where gene flow occurs in narrow hybrid zones (31). Here, we refer

to this axis of parallel divergence simply as Crab-Wave divergence. Demographic modeling revealed that a scenario of multiple in situ independent Crab-Wave divergences is a better fit to genome-wide neutral data than a scenario of ancestral divergence and secondary contact (25). A second axis of divergence is between low-shore and high-shore habitats (Fig. 1B and Supplementary Materials). The Low and High habitats experience contrasting thermal and desiccation conditions that impose strong selective pressures along a steep vertical gradient (32). Snails in high-shore habitats counteract higher desiccation exposure by lowering metabolic rates, exhibiting higher temperature resistance and lower water loss compared to snails in low-shore habitats (33, 34). Here, we refer to this axis simply as Low-High divergence. Differentiation on this axis has been observed in multiple locations, but its demographic history has not been studied. Crab-Wave and Low-High environmental axes follow different directions across localities (Fig. 1B and Supplementary Materials), allowing us to disentangle their effects. In Spain, the Crab habitat is associated with High-Shore and the Wave habitat with Low-Shore. In France and the United Kingdom, the direction is reversed. In Sweden, the two axes are orthogonal, i.e., each of the crab-predation and wave-action habitats contains an elevation

gradient, with steeper environmental change in the wave-action than in the crab-predation habitat.

Seventeen candidate chromosomal inversions have been described recently in *L. saxatilis* from a single Swedish site [fig. S1; see table S1 for details about all candidate inversions; (*22*)]. The putative inversions were identified using a dense transect sample between Crab and Wave habitats and relying on linkage disequilibrium patterns, population genetic signatures, and recombination suppression from linkage maps. Eleven of them showed a clinal pattern of allelic frequency change between Crab and Wave habitats, and three inversion regions [in linkage groups (LGs) 6, 14, and 17] contained a very large number of non-neutral single-nucleotide polymorphisms (SNPs), suggesting that they are maintained under divergent selection (*31*). Furthermore, 10 of the candidate inversions were polymorphic at both ends of the clines, and 6 of them were polymorphic in one habitat, suggesting a role of balancing selection. However, the contribution of candidate inversions to parallel divergence across the species range and in the face of environmental heterogeneity remains unknown.

Parallel ecological divergence in *L. saxatilis* evolved rapidly after the last glacial maximum [~10,000 years ago; (*25, 35*)], and as late as ~1000 to 2000 years ago in the more recently colonized northern range (*24*). The genomic basis underlying the main axis of divergence between Crab and Wave habitats has been extensively studied using various sources of reduced-representation genomic data [e.g., restriction site–associated DNA sequencing (RAD-seq), capture sequencing, and RNA sequencing (RNA-seq)]. Those studies showed that loci potentially under divergent selection between Crab and Wave habitats (i.e., outlier loci) are rarely shared between regions (Iberia, Northeast Atlantic, and North Sea), countries, or even islands within the same archipelago (*36–38*). However, genetic sharing increases between geographically close populations (*9*), suggesting that shorter geographic and/or evolutionary distance might lead to a more similar genetic basis of parallel divergence (*6, 7*), likely due to stronger effects of gene flow and shared standing variation (*20*). These studies did not consider the potentially confounding effects of the two axes of environmental divergence; likewise, before now, no study has used whole-genome data to ask how genetic sharing is distributed over the genome or to consider the role of chromosomal inversions in the evolution of parallel ecological divergence in *L. saxatilis*.

We used the first whole-genome resequencing dataset for *L. saxatilis* to investigate the genome-wide pattern of genetic differentiation and sharing across multiple instances of parallel ecological divergence. We focus on evaluating the roles of geography, genomic architecture, and environmental variation by testing three general predictions:

1) Genome-wide sharing of genetic differentiation in *L. saxatilis* increases with geographical proximity. Populations that are geographically closer and/or share more recent co-ancestry are expected to have a more similar genetic basis of parallel divergence (*6, 7*) due to stronger effects of gene flow and shared standing variation (*9, 20*). Here, we perform the first evaluation of sharing of genetic differentiation across the entire genome.

2) Chromosomal inversion regions are enriched for shared outlier loci. If inversions contain adaptive standing variation and/or facilitate gene flow of adaptive variation, they should promote similar patterns of genetic divergence across multiple instances of parallel divergence (*19, 21*).

3) Genomic differentiation is influenced by more than a single environmental axis of divergence. Here, we quantify the direction of genetic differentiation across the genome for each of the two axes, Crab-Wave and Low-High, and test how well chromosomal inversions explain differentiation in each case.

Overall, we argue that standing variation can be maintained as balanced polymorphism within chromosomal inversions and, through gene flow, can fuel rapid parallel adaptation to heterogeneous environments.

RESULTS

We performed genome resequencing of 1744 individuals pooled into 26 pool-seq libraries from 11 northern European localities using a hierarchical design covering local and regional scales (Fig. 1A and table S2). In Sweden, the shore-level contrast was sampled within Wave and Crab habitats in two localities (SWn3 and SWn5). Thus, the number of pools used in Crab-Wave analyses was 22, while the number of pools used in Low-High comparisons was 18 (see details in the Supplementary Materials and table S2). Pool-seq is cost efficient and recovers accurate population-level allelic frequencies but can be biased when calculating differentiation metrics, as it is not possible to distinguish the source of each sequencing read (*39*). Our approach was robust because (i) we used a large number of individuals in most pools (16 pools had 100 individuals each and 6 pools had 24 individuals each; table S2); (ii) we sequenced pools at high depth (mean = 68×); (iii) pool-seq allelic frequencies from one locality (SWn4 with 24 individuals per sample pool) were highly correlated with those from individual-based sequencing from the same locality [correlation coefficient (r^2) > 0.88; fig. S2]; (iv) our genetic differentiation estimators were highly correlated with recently developed alternative metrics that control for known biases in pool-seq (*39*) (r^2 = 0.95; fig. S3); and (v) the outlier loci identified by these alternative methods overlapped strongly with ours (92% of overlap in average; fig. S4).

Population structure of Crab-Wave genetic differentiation

Localities varied in their Crab-Wave genome-wide genetic differentiation [ranges of mean F_{ST} in 500–base pair (bp) nonoverlapping windows: Spain = 0.09 to 0.12; France = 0.03; United Kingdom = 0.01 to 0.03; Sweden = 0.05 to 0.07; fig. S5]. We summarized genome-wide genetic variation with a principal components analysis (PCA) of 10,263,736 biallelic SNPs (minor allele frequency > 5% in at least one pool). The first four PCA axes, containing most of the variation (75.7%), depicted a genetic structure consistent with geography (Fig. 1C). This genetic structure agrees with the previously inferred *L. saxatilis* biogeographic history: Populations in the United Kingdom, France, and Sweden originate from a different glacial refuge than that of Spanish snails, and our Spanish sites are separated by a north-south phylogeographic discontinuity (*25, 35, 40*).

Sharing of Crab-Wave genetic differentiation is low but increases with geographical proximity

Initially, we applied a commonly used method to identify (Crab-Wave) outliers as those within the top 1% quantile of the F_{ST} distribution. We performed this F_{ST} scan in each of the localities separately and counted the total number of outliers identified at two different levels: 705,786 outliers at the SNP-level and 22,315 outliers at the 500-bp window level. In a PCA with all outlier SNPs, the first four axes

(63.8% of variation) depicted a hierarchical structure of strong differentiation between countries and weaker ecotype differentiation within countries (Fig. 1C and fig. S6). We quantified the amount of outlier sharing across localities. Overall, higher sharing was observed at the window level than at the SNP level, and most of the SNP- and window-level outlier loci (>66%; SNPs = 529,940; windows = 14,835) were unique to their locality (Fig. 1D). Accordingly, the level of global outlier sharing across all localities was extremely low (<0.016%; 118 SNPs and three windows shared across all localities; Fig. 1D). However, we expected no globally shared outliers at random, assuming independence across this large number of localities (see Methods).

To investigate the geographic context of outlier sharing, we estimated the amount of pairwise SNP outlier sharing, which varied between 2.3 and 25% across comparisons (Fig. 1E), being always higher than the random expectation of sharing between two given localities according to a hypergeometric distribution [1%, 99% confidence interval (CI) = 0.9 to 1.07; see Methods]. The sharing pattern had a strong geographical signal where nearby localities shared a higher number of outliers than localities further apart (Mantel $r = -0.76$, CI = -0.83 to -0.66, $P < 0.001$). Overall, our findings are consistent with previous reports focusing on small portions of the

genome where sharing of Crab-Wave outliers was generally low (36–38, 40) but increased at nearby localities (9). This supports our first prediction: Ancestral variation, common to closely related populations, and/or shared variation via gene flow between spatially proximate populations play an important role in determining the genomic basis of parallel ecological divergence in *L. saxatilis*.

Genomic landscape of Crab-Wave differentiation

Using a recently developed linkage map for *L. saxatilis* (31), we were able to place half of the total genome content into 17 LGs. The map resolution is moderate (~0.5 cM), so multiple scaffolds are associated with the same map position. The patterns of outlier sharing did not vary between scaffolds placed in the linkage map and those not placed (fig. S8). We observed a heterogeneous landscape of genomic differentiation between ecotypes (mean F_{ST} in Fig. 2A and F_{ST} per locality in fig. S9), as commonly observed in other natural systems when divergence proceeds in the face of gene flow [e.g., (8, 41)]. While F_{ST} outlier sharing across the genome was relatively low (mean = 10.6% and SD = 8.5; fig. S10), it was markedly higher in some LGs (e.g., in LG6, mean = 24.5% and SD = 16%; fig. S10).

To further investigate Crab-Wave differentiation across the genome, we used two additional methods to identify outliers that
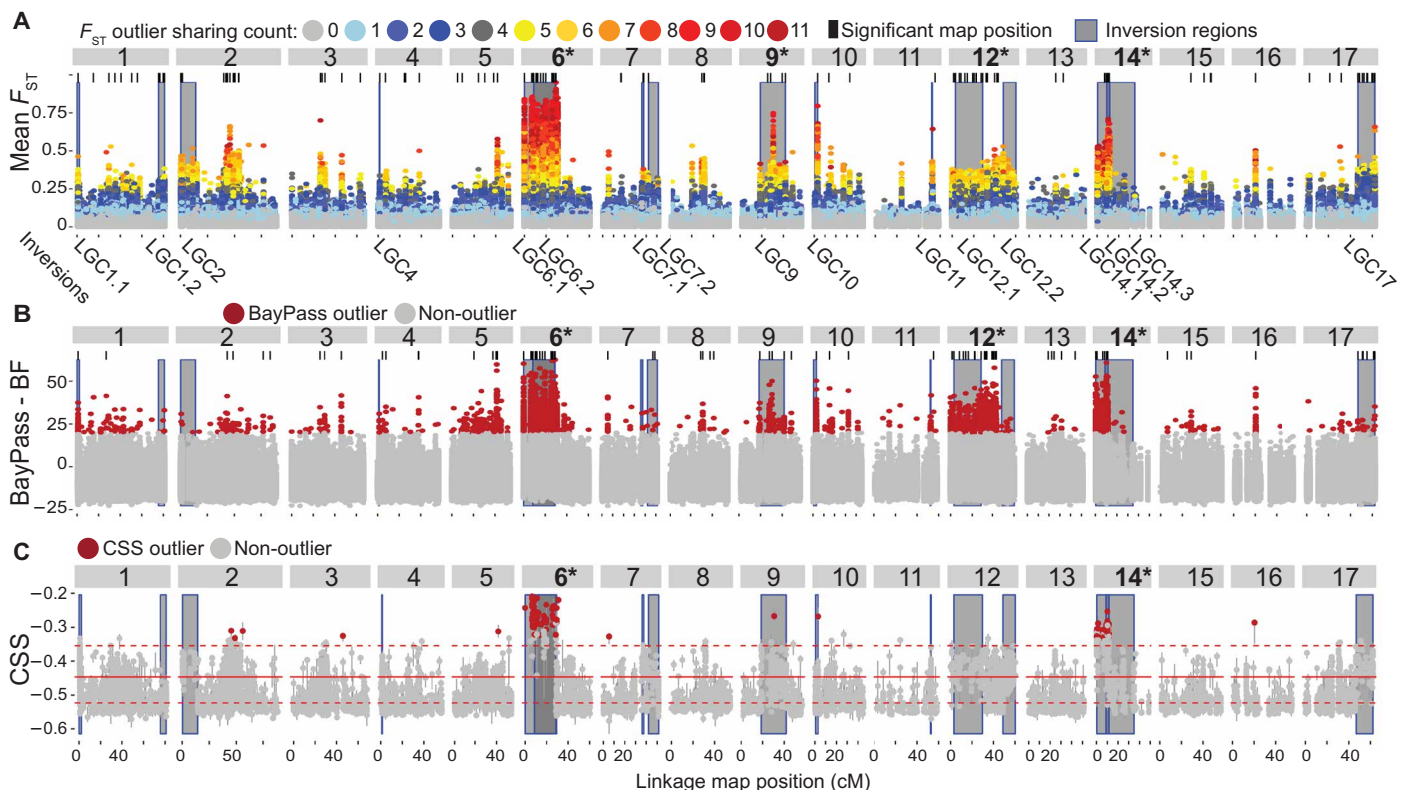

Fig. 2. Genomic landscape of parallel Crab-Wave divergence. (**A**) Average Crab-Wave F_{ST} across all localities, reflecting per SNP genetic differentiation. Dots are colored according to their outlier sharing count. (**B**) BayPass model, reflecting per SNP genetic covariation with Crab versus Wave environments. Outlier SNPs (red) were defined by BF score > 20. (**C**) Cluster separation score (CSS), per map position genetic distance between Crab and Wave groups relative to within-group distance. Gray lines indicate 95% CIs per map position, and red lines indicate the mean and 95% CI for a genome-wide permutation. Outlier map positions (red) are defined by non-overlapping CIs. To identify linkage groups (LGs) and map positions that harbor significantly more divergent shared loci, we performed a permutation test ($P < 0.01$) with values drawn from each of these three tests (F_{ST}, BayPass, and CSS; see the Supplementary Materials for details). At the LG scale, we identified significant LG values as those higher than the 95th percentile of a random genome-wide distribution (highlighted in bold), while at the map position scale, we identified significant map position values as those higher than the 95th percentile of a random distribution within the same LG (highlighted with black bars). Given that CSS operates at the map position level, we did not perform the permutation at this scale.

exploit different types of information (see Methods for details). We used the covariate model from BayPass (42) to measure the per SNP association between allele counts and Crab versus Wave habitat membership across all localities. This approach complements the F_{ST} analysis by accounting for shared demographic history with the covariance matrix Ω [(43); fig. S11] and focusing on covariation rather than on a nondirectional measure of differentiation like F_{ST}. We consider outlier SNPs as those above a Bayes factor (BF) threshold of 20, as recommended in BayPass. Large BF values reflect stronger environmental correlation across multiple (but not necessarily all) localities without relying on the strength of genetic differentiation per se. We also used a cluster separation score (CSS), as in (11), to measure the genetic distance (i.e., Euclidean distance in a PCA) between Crab and Wave populations across localities relative to the genetic distance between localities within each of the Crab and Wave groups. The CSS accounts for direction of differentiation like BayPass, but unlike F_{ST}, it operates at the map position level instead of the SNP level, potentially providing a better signal-to-noise ratio. The higher the CSS, the greater and the more consistent the Crab-Wave genetic distance is across localities. We performed permutations of observed values within map position and a permutation across the genome (i.e., random expectation), and then we considered map positions as outliers when the 95% CIs of observed values and the random expectation did not overlap.

Overall, we found that the three alternative methods showed a high degree of agreement, particularly within some LGs, e.g., LG6, LG9, LG12, and LG14, which showed significantly more outliers than the genome-wide expectation across all three tests (LG numbers highlighted in bold in Fig. 2, A to C). We also observed map positions with a significant excess of outliers scattered throughout the genome (highlighted with black bars in Fig. 2, A and B, and red dots in Fig. 2C). Discrepancy between the F_{ST} and the BayPass tests suggests that SNPs with Crab-Wave habitat associations may not always be accompanied by elevated Crab-Wave genetic differentiation. To confirm this, we compared BayPass BF outliers to the F_{ST} outlier sharing count. Most of the F_{ST} outlier SNPs (98.7%) identified (in any locality) were not identified as BayPass outliers. Moreover, among BayPass outliers, 12% were not identified as F_{ST} outliers at all, and of those that were, 49% had low F_{ST} outlier sharing counts (shared by three localities or fewer; fig. S12). Last, as expected from the pairwise pattern of F_{ST} outlier sharing (Fig. 1E), the genomic landscape of pairwise F_{ST} outlier sharing also had a strong geo-graphical component (see fig. S13 for a genome-wide overview of F_{ST} outlier sharing). For example, sharing in the LG17 cluster (around position 60 cM) was strong in Spain and less prevalent elsewhere, while strong sharing in LG2 (around position 50 cM) was limited to comparisons involving Spanish and Swedish localities.

Strong sharing of Crab-Wave differentiation coincides with putative chromosomal inversions

We were interested in testing whether the presence of polymorphic inversions in *L. saxatilis* could facilitate the evolution of parallel ecological divergence. To investigate this question we used the genomic coordinates of candidate inversions (22) to assess the contribution of genetic differentiation within these regions to the geo-graphical pattern of shared ecotype differentiation across localities. We refer to these genomic regions as "inversion regions" and other parts of the genome as "collinear regions" but stress that polymorphic inversions have been detected only at one Swedish site (SWn4; fig. S1 and table S1). To investigate whether inversion regions contained a higher proportion of outliers than expected at random, we compared the proportion of SNPs contained within collinear and inversion regions to the number of outlier loci contained within collinear and inversion regions. Inversion regions are significantly enriched with outliers in every locality for all three F_{ST}, BayPass, and CSS tests (χ^2 $P < 0.001$; Fig. 3A). This pattern is most easily explained if the inversions present in SWn4 are also polymorphic in many other sites across the species range. To ask whether sharing increases at smaller geographic scales, we correlated the amount of pairwise sharing with pairwise geographic distance between localities for the whole genome, collinear regions, and inversion regions. As expected, sharing was negatively correlated with geographic distance, but sharing was higher and its relationship with distance was stronger in inversions than collinear regions (Fig. 3B; genome-wide: Mantel $r = -0.76$, CI = -0.83 to -0.66, $P < 0.001$; inversion regions: Mantel $r = -0.82$, CI = -0.88 to -0.75, $P < 0.001$; collinear regions: Mantel $r = -0.68$, CI = -0.74 to -0.56, $P < 0.001$). To illustrate sharing behavior across different spatial comparisons, we measured outlier sharing for different pairwise comparisons among groups of localities that are progressively closer to each other, where distance is defined as the average pairwise distances between all localities. Regardless of which genomic regions were considered, sharing decreased with geographical distance (Fig. 3B). Overall, these analyses support our second prediction: Outlier loci for Crab-Wave
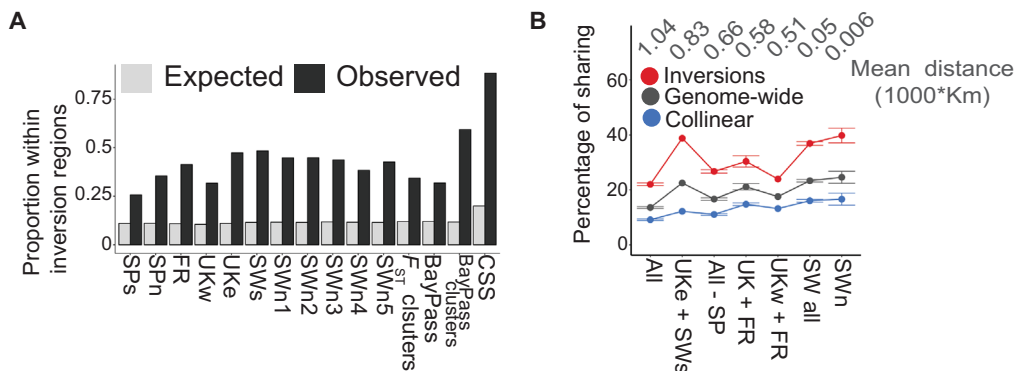

Fig. 3. The influence of chromosomal inversions on outlier loci sharing. (**A**) Expected versus observed count of outliers within inversion regions for F_{ST} outliers for each locality, significant F_{ST} map positions (i.e., black bars in Fig. 2A), BayPass outliers, significant BayPass map positions (i.e., black bars in Fig. 2B), and CSS outliers. (**B**) Average pairwise sharing of F_{ST} outliers and average pairwise geographic distance across different comparisons.

differentiation are clustered across the genome, mostly within putative chromosomal inversions, which show a higher level of sharing among spatially proximate populations. This, in turn, suggests that chromosomal inversions could function as a reservoir of ancestral variation (i.e., more common among populations with more recent co-ancestry) and/or as vehicles for adaptive gene flow (i.e., more common among spatially closer populations).

Genomic differentiation varies according to different axes of environmental variation

Apart from the Crab-Wave axis, *L. saxatilis* experiences a second axis of parallel divergence of Low-High shore-level ecological variation (see Introduction, Fig. 1B, and the Supplementary Materials) (*32*, *34*). To identify genomic regions that are related to the effect of Low-High divergence, we repeated the BayPass and CSS analyses for this environmental axis. Both BayPass and CSS analyses identified several LGs with clusters of differentiation consistent with strong, parallel Low-High divergence, which were sometimes similar to those identified in the Crab-Wave analyses and sometimes different (Fig. 4, A and B). A direct comparison of BayPass BF scores revealed only a modest correlation between the two axes ($r^2 = 0.16$, $P < 0.001$; Fig. 4D), and a comparison of CSS scores revealed an even lower one ($r^2 = 0.04$, $P = 0.07$), suggesting that most SNPs are differentially implicated in each of the axes of parallel divergence.

Leveraging our sampling design, we used the relative direction of the two environmental axes to disentangle their genomic signatures. We developed a metric, the differentiation index (DI), to measure and compare the strength and direction of allele frequency differentiation on each axis of divergence. The DI is a normalized measure of Low-High genetic differentiation at a locality, using the two Spanish pools to confer directionality in assigning reference alleles. We considered the DI value to be positive, negative, or zero if it was higher, lower, or within 1.5 SDs of a random expectation, respectively (positive, green squares; negative, purple squares; or zero, gray squares in Fig. 4C; see Methods and fig. S14). To classify each map position as concordant with Crab-Wave or Low-High divergence, we applied the DI test to all map positions across the genome and relied on the relative directionality of the two axes of environmental divergence, as follows: (i) Crab-Wave divergence: We expect DI to have opposite directions between Spain and the United Kingdom/France and no differentiation in Sweden. Thus, we categorized map positions as showing Crab-Wave divergence if they were positive in at least one Spanish locality, negative in at least one French or U.K. locality, and close to zero in at least one Swedish locality (green rectangles in Fig. 4C). (ii) Low-High divergence: We expect DI to have positive direction in all localities. Thus, we categorized map positions as showing Low-High divergence if they were positive in at least one Spanish locality, positive in at least one French or U.K. locality, and positive in at least one Swedish locality (orange rectangles in Fig. 4C). Applying the DI metric allowed us to highlight strong signals of association with the different axes of parallel environmental variation for genomic regions previously
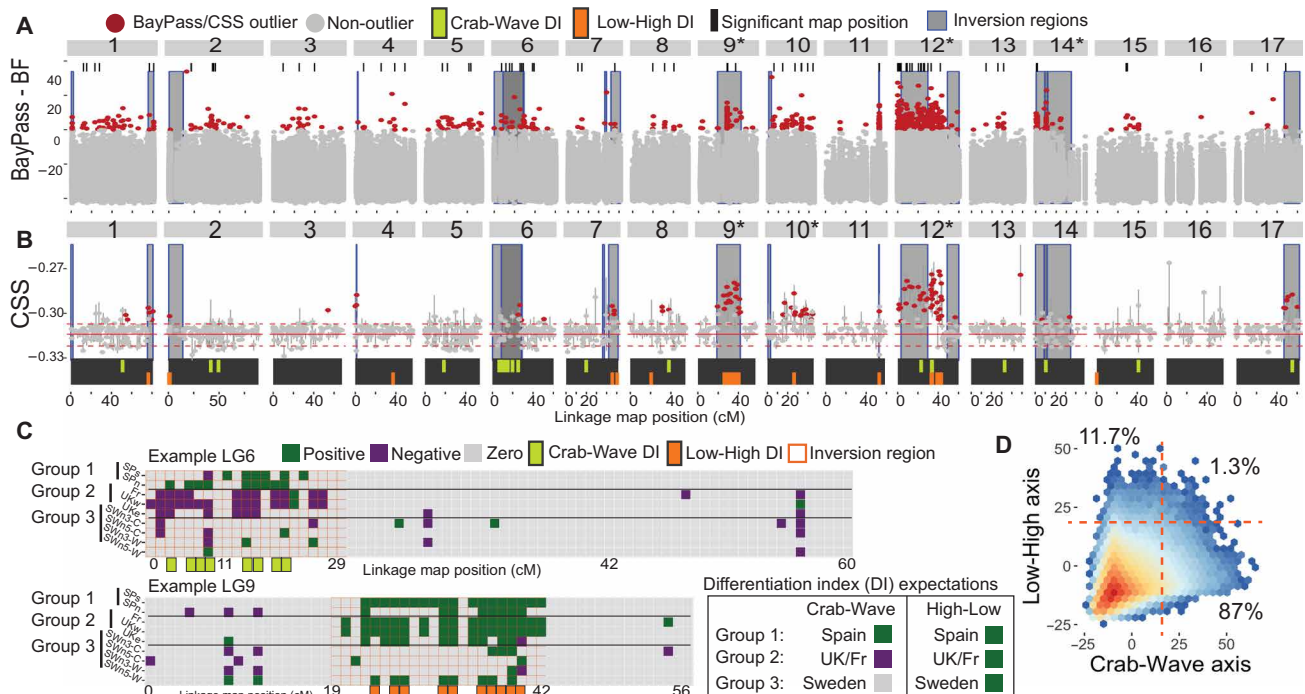

Fig. 4. Two axes of parallel ecological divergence, Crab-Wave versus Low-High habitat contrasts. (**A**) BayPass covariate model for Low-High divergence. (**B**) CSS for Low-High divergence. Green versus orange lines show map positions that were classified as differentiated on Crab-Wave or Low-High axes of divergence, respectively, according to their differentiation index (DI). (**C**) Examples of DI for LG6 and LG9. Each map position was classified as positive, negative, or near zero relative to ±1.5 SDs of the genome-wide DI mean value for a given population. Map positions are likely implicated in Crab-Wave or Low-High divergence [green versus orange lines in (B), see text] if they show the expected directions in at least one locality of each group defined according to the relative direction between the two environmental axes (see inset and main text). (**D**) Two-dimensional density plot comparing BayPass BF scores between Crab-Wave and Low-High tests. The grid colors show the range of count values from blue to red to reflect density of overlapping points. Orange dashed lines represent the BF score = 20 thresholds to define outliers. Among all outliers, 87% were found on the Crab-Wave, 11.7% on Low-High, and 1.3% on both comparisons.

identified by outlier loci tests. To illustrate the DI metric, we selected two LGs that consistently showed shared differentiation in Crab-Wave or Low-High divergence (Fig. 4C). Key map positions of LG6 exhibited DI directions consistent with Crab-Wave divergence, while key map positions of LG9 fitted the expected DI directions for Low-High divergence (see fig. S14 for DI plots for the remaining LGs). Most map positions (96.3%) across the genome were not categorized, as they did not meet the expected combination of directions; however, 21 map positions were consistent with Crab-Wave divergence (green lines in the bottom of Fig. 4B; 10 in collinear and 11 in inversion regions) and 26 with Low-High divergence (orange lines in the bottom of Fig. 4B; 10 in collinear and 16 in inversion regions). As expected, map positions associated with each of the axes of environmental variation were mostly found in genomic regions identified by BayPass and CSS tests (χ^2 $P < 0.01$). Consistent with our third expectation, the two independent axes of parallel divergence are mostly associated with different genomic regions.

Chromosomal inversion regions are differentially involved in two axes of parallel divergence

We wanted to test whether candidate chromosomal inversion regions are likely implicated in parallel ecological divergence to both Crab-Wave and Low-High divergence. We first tested whether genetic differentiation within inversions was significantly higher than a random expectation drawn from the collinear genomic background in each locality (Fig. 5; see details in the Supplementary Materials). While in Spain, United Kingdom, and France genetic differentiation is expected to reflect selection along both environmental axes, in Sweden the axes' effects are separate (y axis in Fig. 5). Our findings indicated that some chromosomal inversion regions are likely repeatedly involved in the Crab-Wave divergence (e.g., in LG6 and LG14) and others in the High-Low divergence (e.g., LG9 and LG12). The patterns observed for other inversions are more complex, as some are predominantly differentiated in Spain and Sweden (group C in Fig. 5), while others show no clear pattern (group D in Fig. 5). Next, following findings from Faria et al. (22) that most inversions are polymorphic in one or both of the Crab and Wave habitats, we tested whether
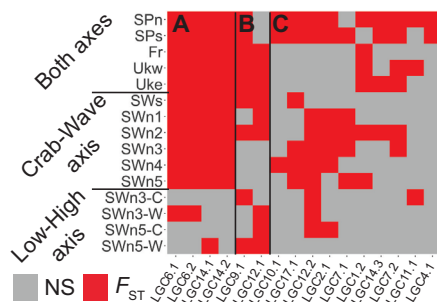

Fig. 5. Genetic differentiation at chromosomal inversion regions. Heatmap showing in red inversion regions that have significantly higher mean genetic differentiation (F_{ST}) compared to the 99% CI of the collinear genomic background. Nonsignificant (NS) tests are indicated in gray. F_{ST} was measured per locality according to their direction along Crab-Wave and Low-High axes as indicated on the y axis (see main text and Fig. 1B). The three groups on the y axis indicate localities that are expected to show a consistent signal depending if they are implicated in the Crab-Wave, Low-High, or both axes of divergence. Groups in the x axis indicate chromosomal inversion regions that have a signal of differentiation consistent with the expectations for A = Crab-Wave divergence, B = Low-High divergence, or C = variable across the species range.

inversion regions have more, the same, or less nucleotide diversity than collinear regions. Despite some variation across localities, we found overall patterns that are highly consistent with Faria et al. (22) (Fig. 6). Most inversion regions harbor significantly more genetic diversity than collinear regions on the same LG in one or both of the Crab and Wave habitats (Fig. 6; $P < 0.01$; see table S3 for full details on linear mixed models). Higher diversity in inverted regions within a pool suggests that inversions are polymorphic. However, given that we used pool-seq, we cannot be certain whether the elevated diversity is between arrangements or at the nucleotide level within an arrangement. At the nucleotide level, arrangements could have accumulated genetic diversity by mutation and gene flux over time. At the arrangement level, the diversity signal could come from the presence of both arrangements in the same habitat, with strong differentiation between them, as seen in the study by Faria et al. (22). Some inversion regions have similar levels of diversity to collinear regions, which suggests that they are not polymorphic, that arrangements have not accumulated differentiation, or, alternatively, that older inversions have accumulated similar levels of diversity to collinear regions. Last, other inversion regions have less diversity than collinear regions, suggesting that these could have swept to high frequency recently. Overall, these results support our third prediction: Accounting for the genomic architecture and the direction of both Crab-Wave and Low-High axes of divergence is essential to understanding the genomic basis of parallel ecological divergence in L. saxatilis.

DISCUSSION

Investigating the genetic basis of parallel ecological divergence is a key step toward understanding whether, and to what extent, evolution is predictable (1, 5). Here, we studied patterns of shared genomic divergence in one of the best-established natural systems of rapid parallel ecological divergence, the marine snail L. saxatilis (5, 25). This species has repeatedly evolved adaptations to local heterogeneous environments within the last 10,000 years, allowing replicated measures of adaptive divergence (Fig. 1, A and B). While previous studies focused on a single axis of ecological divergence in L. saxatilis (36–38), our study disentangles patterns of genome-wide differentiation across two major axes of parallel ecological selection: Crab-Wave and High-Low shore-level divergence (Fig. 1B). Specifically, we carried out the first whole-genome study of genetic differentiation to disentangle these ecological axes of divergence in the context of a common underlying genetic architecture. Overall, the extent of similarity in genomic differentiation can be explained by geographical proximity (Fig. 1E), by the presence of putative chromosomal inversions (Figs. 3A and 5), and by accounting for the effects of multiple environmental axes (Fig. 2, B and C, versus Fig. 4, A and B), supporting our predictions.

Adaptive substrate for parallel divergence via evolutionary and geographic links

Our results indicate that accounting for the geographical context of genetic differentiation is essential to understanding the evolution of parallel ecological divergence. Specifically, we showed that while sharing of outlier loci is generally low, sharing increases with geographical proximity. Ravinet et al. (37), using a low-density genomic scan with RAD-seq, found a range of 8 to 28% of outliers to be shared between pairs of nearby islands in Sweden. Westram et al. (38),
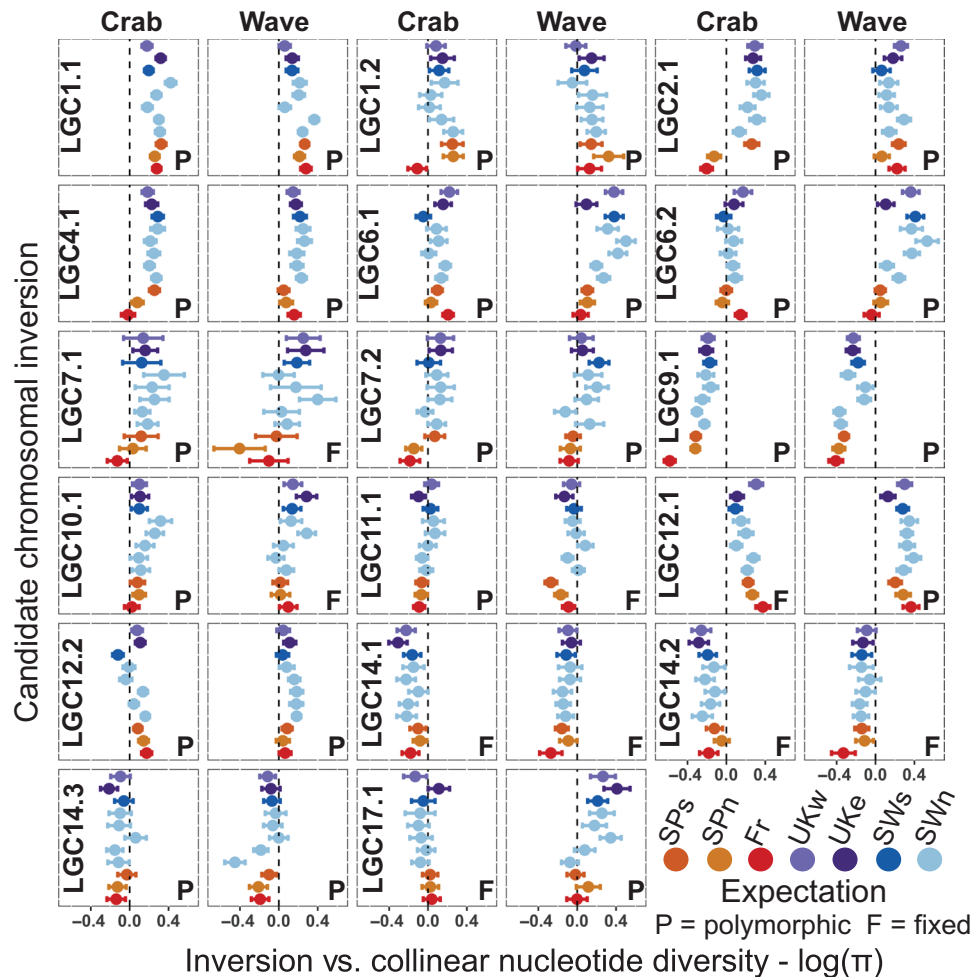
Fig. 6. Nucleotide diversity of candidate inversion regions. Forest plot showing the fixed effect estimates of a linear mixed effect model (LMM) of nucleotide diversity within chromosomal inversion regions compared to collinear regions within the same LG and scaffold number as random effect. Positive values indicate that nucleotide diversity is higher in the inverted than in the collinear region. See table S3 for untransformed mean differences and full LMM output. Expectations for the inversion regions to be polymorphic (P) or fixed (F) are based on a clinal analysis of allelic frequencies in a single Swedish locality (SWn4) as reported by Faria *et al.* (*22*).

using a transcriptome-based genome scan, found pairwise sharing to range from 5% in their most distant comparisons (Spain versus Sweden) to 20% in their closest comparisons (between our Sweden North and Sweden South localities). Here, with a much larger sample of the genome and more reliable estimates of differentiation, we found pairwise sharing to range between 2 and 25% [from the same Spain-Sweden comparison as used by Westram *et al.* (*38*) to islands at similar separation to those by Ravinet *et al.* (*37*)]. The consistency of our results with those of Westram *et al.* (*38*) suggest that their RNA-seq approach provided enough SNP density to estimate genetic sharing accurately.

Our findings, together with those from (*9*), indicate that localities with more recent co-ancestry likely share more standing variation and/or that gene flow can more easily spread adaptive variation between geographically closer localities (*1*, *9*, *20*). Levels of shared genetic differentiation were highly heterogeneous across the genome: We identified a relatively small number of regions that account for most of the sharing, both inside and outside putative inverted regions, and that are differentially involved in the two axes of parallel ecological divergence (Crab-Wave or Low-High). Overall, our find-

ings suggest that *L. saxatilis* was able to use reservoirs of standing genetic variation repeatedly and/or source genetic variation via gene flow to evolve rapid parallel local adaptation to heterogeneous environments.

Chromosomal inversions as reservoirs and vehicles for rapid parallel divergence

Outlier sharing is clearly magnified within putative chromosomal inversions that we defined here on the basis of polymorphisms at a single Swedish site (*22*). We propose that chromosomal inversions might have played an important role during the evolution of rapid postglacial parallel ecological divergence in *L. saxatilis*. We found evidence consistent with widespread chromosomal inversions, often polymorphic, that are differentially involved in two axes of ecological divergence, parallel Crab-Wave divergence (e.g., on LG6 and LG14) and parallel Low-High divergence (e.g., on LG12 and LG9). According to theory and observations from other systems (*17–19*, *21*, *44*), inversion regions might store adaptive loci as ancestral polymorphism that can be used later as a substrate for adaptive divergence and can be distributed via gene flow as "adaptive

cassettes." At the same time, inversions could establish barriers to gene flow by suppressing recombination across large genomic regions (*45*).

The suggestion of chromosomal inversions acting as polymorphism reservoirs and barriers to gene flow is in line with our findings and those of Westram *et al.* (*31*) and Faria *et al.* (*22*), who used a densely sampled transect across a hybrid zone between Wave and Crab habitats in a single Swedish site (SWn4). In those studies, some candidate inversion regions (chiefly those in LG6, LG14, and LG17; table S1) contained clusters of strongly clinal loci (i.e., non-neutral SNPs) but most remained polymorphic in one or both habitats, suggesting that inversion polymorphism might be maintained by a combination of divergent and balancing selection. Here, we investigated the role of putative chromosomal inversions across a much larger geographical scale from Spain to Sweden and observed similar patterns. We found that putative inversion regions behave as large genomic blocks of high divergence but, at the same time, most of them also contain higher nucleotide diversity than collinear regions. Moreover, some SNPs have a strong signal of covariation with habitat contrasts (i.e., BayPass test) despite low F_{ST} values and outlier sharing. We hypothesize that repeated adaptation to heterogeneous environments could be based partly on a shared pool of standing genetic variation maintained within chromosomal inversions as balanced polymorphism that vary in equilibrium frequency according to habitat. It is likely that most of the inversion polymorphisms evolved long before *L. saxatilis*' postglacial colonization, possibly contributing to divergence in the distant past and more recently fuelling rapid parallel divergence (*22*, *46–48*). This hypothesis needs to be tested with demographic modeling including data from sister species. We stress, however, that it is important to be cautious when interpreting the potential role of chromosomal inversions in the evolution of parallel adaptive divergence. Chromosomal inversions are often selected as large haplotype blocks, and thus are expected to have a strong signal of divergence in a genomic scan like ours. Paradoxically, while the presence of inversions allows us to identify large genomic regions implicated in parallel adaptation, it can also prevent us from identifying the causal genes or allelic variants underlying fitness. Nevertheless, narrowing down candidates should be possible. For example, taking advantage of the potential linkage disequilibrium (LD)–disruption effect of gene flux between alternative arrangements in older inversions might allow association mapping of adaptive traits using extreme phenotypes to identify smaller genomic regions [see (*49*)].

The presence of chromosomal inversions can confound the perceived amount of genetic sharing. For instance, we found that sharing increases when measured at different scales from SNPs to 500-bp windows to inversion regions. This finding is consistent with patterns found in other systems where outlier clustering leads to higher sharing at increasingly larger genomic scales (*10*, *50*). This could be explained, in part, by the inherent limitation of defining outlier SNPs as those in the top 1% of the F_{ST} distribution, decreasing sharing estimates for loci that are highly differentiated but just below the stringent threshold. Moreover, inversions might contain many (nonshared) neutral loci of high differentiation that are linked to causal adaptive loci. Particularly, in the case of older inversions, accumulation of mutations and occasional recombination events can drive neutral hitchhikers to high frequencies randomly across localities, leading to overall low sharing at the SNP level. Alternatively, low sharing could be explained by a redundant genomic basis

of local adaptation where different combinations of causal alleles can lead to the same adaptive phenotype (*51*), effectively reducing sharing at the SNP level (*10*).

Parallelism cannot be explained fully by the presence of putative chromosomal inversions, as we also found shared outlier loci scattered across the genome [cf. (*11*)]. This is consistent with polygenic selection of multiple loci of small effect underlying parallel adaptive divergence in *L. saxatilis* (*9*, *30*, *31*, *38*). This further confounds estimates of the contribution of loci within versus outside inversion regions to adaptive divergence if many small-effect loci go undetected. In addition, some clusters of shared genetic differentiation were observed in regions with no known chromosomal inversions (e.g., LG2, LG8, or LG10). There are several probable explanations for these genomic clusters that we cannot distinguish at the moment. On one hand, these clusters could reflect the presence of other currently undetected chromosomal inversions [for example, if some chromosomal inversions have recent origins, or if inversions are variable elsewhere in the species range but fixed for one arrangement in SWn3 and thus undetected by Faria *et al.* (*22*), or if inversions are variable in their allelic content and associated fitness effects across sites]. On the other hand, these clusters might not be related to inversions but to other mechanisms of suppressed recombination [e.g., within centromere regions; (*46*)] that could generate this pattern, in combination with background selection or divergence hitchhiking (*52*).

Parallel divergence at two environmental dimensions

Our results demonstrate that considering environmental heterogeneity across multiple instances of divergence is crucial to understanding the genetic basis of parallel evolution (*3*, *4*, *12*, *13*). When parallel divergence is measured using a single environmental contrast, it is possible that other axes of environmental variation can confound the estimation of genetic differentiation. This is particularly true for those that are associated with more cryptic phenotypic differences, as in the case of physiological differences between low-shore and high-shore habitats in *L. saxatilis*. The contribution of environmental heterogeneity has been discussed previously in other widely studied systems that have focused on a single axis of parallel divergence, but not quantified, e.g., Midas cichlid fish species (*53*), North American lake whitefish species (*54*), *Timema* stick-insect ecomorphs (*8*), and *Pundamilia* cichlid fish species pairs (*55*), but see (*12*, *13*) in stickleback fish. In *L. saxatilis*, measuring genetic differentiation without accounting for its directionality (e.g., with F_{ST}) would have provided an erroneous perception of sharing due to the confounding effects of different axes of parallel divergence (*9*). Explicitly incorporating the directionality of these axes, as we did here with BayPass, CSS, and our DI metric, allow us to clarify their relative contributions, and provide a more accurate representation of genetic differentiation across multiple instances of parallel divergence.

Conclusions

Our findings reveal that considering multiple factors is essential to understanding the genomic basis of parallel ecological divergence. The evolutionary history and geographic context, the congruence in direction between distinct environmental axes, and the maintenance and redistribution via gene flow of alleles contained within chromosomal inversion polymorphisms all contribute to the resulting patterns. In the particular case of *L. saxatilis*, we show that

accounting for an additional axis of environmental heterogeneity allows us to uncover how parallel evolution can proceed in different directions and how this affects different regions of the genome in relation to a common underlying genomic architecture. We show how two independent environmental axes seem to be related to different putative chromosomal inversions, and our results suggest that polymorphic inversions might have played an essential role in maintaining and redistributing the adaptive substrate to fuel rapid divergence in the face of environmental heterogeneity.

METHODS
Sample collection
Samples were collected within typical *L. saxatilis* habitats, far from the contact zones between Crab-Wave and Low-High habitats (see below), in 11 localities: Spain (*n* = 2 pooled samples), France (*n* = 1), United Kingdom (*n* = 2), and Sweden (*n* = 6) (table S2). For three of the six Swedish localities (SWn1, SWn2, and SWn4, *n* = 6 sample pools), pools consisted of 24 individuals from both sexes. For the remaining localities (*n* = 16 pools), pools consisted of 100 female individuals (females were selected to avoid contamination with a sibling species, *Littorina arcana*, which coexists with *L. saxatilis* in the United Kingdom and France; males of the two species cannot be distinguished morphologically). Our sampling design includes two axes of environmental differentiation, Crab-Wave and Low-High habitat contrasts. Crab-Wave divergence consists of two ecotypes locally adapted to crab-predation and wave-action, respectively. Low-High shore-level divergence consists of two local microhabitats of differential temperature and desiccation selective pressures along a sharp vertical gradient. See details in the Supplementary Materials. The two axes of environmental differentiation follow different directions across the different sampling localities (see Fig. 1B). Shore-level samples in Sweden were collected for two locations (Ramsö and Arsklovet; SWn3 and SWn5, respectively), where high- and low-shore individuals within each of the Crab and Wave habitats were kept separate. This means that each of these four pools was divided in two pools of 50 females each for analyses of High-Low divergence and combined for analyses of Crab-Wave divergence into pools of 100 females. Thus, the number of pools we used varied according to which axis of divergence (Crab-Wave or Low-High) was considered. Analyses considering Low-High divergence used a total of 18 pools, i.e., excluding Swedish samples that we did not split into high and low shore (table S2). For the Crab-Wave divergence analyses, we used a total of 22 pools (table S2).

DNA sequencing and processing
Snails were stored in 96% ethanol at −20°C before DNA extraction. For pools of Swedish localities SWn1, SWn2, and SWn4, we used DNA extractions from individual snails used by Ravinet *et al.* (*37*) and pooled them in equimolar amounts after Qubit fluorometer quantification. For the rest of pools, DNA was extracted for batches of five individuals by combining pieces of foot muscle tissue of 1-mm³ size from five snails in one tube. Extractions were done using modified CTAB protocol from (*56*). DNA quality and quantity were assessed by agarose gel electrophoresis and Qubit fluorometer and pooled in equal amount to produce the final pools for sequencing. These samples were further purified using a Zymo Genomic DNA Clean-up and Concentrator kit and subjected to quality and quantity control as above. Genomic libraries were built with the Illumina

TruSeq PCR-Free Kit (Illumina, California, USA), insert sizes between 350 and 670 bp, and were sequenced in an Illumina HiSeq2500 machine with an SBS Kit v4 chemistry at SciLifeLab (Uppsala, Sweden).

Sequencing read processing
Paired-end reads were trimmed with Trimmomatic v.0.36 (*57*) using the following parameters: LEADING:3 TRAILING:3 SLIDING-WINDOW:4:15 MINLEN:30. After trimming, a total of 14.1 million reads were retained, with an average of 706 (SD = 240) thousand reads (lengths = 30 to 126 bp) per pool. Trimmed FASTQ files were mapped against the *L. saxatilis* reference genome from a single Crab ecotype individual [N50, 44,284 bp; NG50 (based on genome size 1.35 Gbp), 55,450 bp; maximum scaffold length, 608,273 bp; total number of contigs, 388,619]. Kofler *et al.* (*58*) reported that pool-seq data tend to show high levels of disagreement in population genetic metrics (e.g., F_{ST}) when different read mappers were used. To attend to these concerns, we initially used two different read mappers: BWA mem v0.7.15 (*59*) and CLC v5.0.3 (www.qiagenbioinformatics.com) with default parameters. Minimal inconsistencies were found between two alternative mappers in depth and F_{ST}. However, CLC proved able to handle repetitive regions better than BWA (i.e., it penalized mapping score appropriately at repeated regions). This is in agreement with Kofler *et al.* (*58*), where CLC was shown to outperform other mappers. Consequently, we decided to use CLC for read mapping. For all samples, more than 97% of the reads were mapped to the reference genome, with a minimal number (<0.7%) mapped as singletons. Given the high level of fragmentation of the reference genome, paired reads were mapped to the same scaffold in only ~70% of the cases. Of the paired reads that mapped to different scaffolds, >50% had high mapping quality scores (Q20 or higher) and were thus retained for downstream analyses. Bam files were processed with SAMtools v1.3.1 (*60*), BEDtools v2.25.0 (*61*), and Picard tools v2.7.1 (http://broadinstitute.github.io/picard). For each set of bam files, we filtered out reads with base quality lower than 30 and mapping quality lower than 20 and those that mapped to very short contigs (<500 bp; 158,060 contigs). We also removed positions of low coverage to avoid uncertain calls and positions of very high coverage to avoid potential repetitive regions. We fitted the depth of coverage to a mixed distribution model to define three classes of coverage: low, medium, and high with the function mixmdl from the R (*62*) package mixtools. We used this model to define cutoff values to remove 100% of the low-coverage distribution and 50% of the high-coverage distribution (low cutoff = 14 and high cutoff = 204; see fig. S15 for details). Population genomic analyses were performed with filtered files with an average coverage of 68× (min = 14×; max = 204×; SD = 18×).

F_{ST} outliers and sharing
To calculate pairwise genetic differentiation (F_{ST}) between samples in each pair of pools, we used the script "fst-sliding.pl" from PoPoolation2 (*63*) for every SNP and for non-overlapping 500-bp windows. PoPoolation2 calculates F_{ST} from the allele frequencies (not allele counts) using the equation from (*64*)

$$F_{ST} = (Pi\_total − Pi\_within)/Pi\_total$$

The formula for Pi is the standard: $Pi = 1 − fA^2 − fT^2 − fC^2 − fG^2$, where *f* is the frequency of each nucleotide. The terms of the equation include Pi_within = (Pi_pool1 + Pi_pool2)/2 and Pi_total = Pi of

the averaged allele frequencies of the two pools. We defined outlier loci as those that were exceptionally differentiated between ecotypes at two different levels, SNPs and 500-bp windows, as those within the top 1% of their F_{ST} distribution.

To measure outlier sharing, we first identified F_{ST} in each locality and then we asked how many times each outlier was found across all localities. To compare the amount of outlier loci sharing observed between localities with a random expectation, we used the formula

$$\frac{\text{Number of outliers}}{\text{Number of SNPs}} * \text{Number of outliers}$$

That is, we expect 0.01% of SNPs (or 1% of outliers) to be shared at random. We also obtained the 99% CI for this random expectation using the hypergeometric distribution function "qhyper" from the base library in R as follows

$$\text{Lower CI} = \text{qhyper}(0.01, \text{Number of outliers}, (\text{Number of SNPs} - \text{Number of outliers}), \text{Number of outliers})$$

$$\text{Higher CI} = \text{qhyper}(0.99, \text{Number of outliers}, (\text{Number of SNPs} - \text{Number of outliers}), \text{Number of outliers})$$

Given this already small sharing probability, the probability for an outlier to be shared across all 11 localities is extremely low: $0.01^{11} \approx 0$. We measured the contribution of geographical distance to outlier sharing as the correlation between pairwise sharing and pairwise geographical distances between localities with a simple Mantel test with the R package ecodist.

BayPass analysis

BayPass (42) is an extension of the Bayesian outlier detection model implemented in BayEnv (43) and has been thoroughly tested with pool-seq data, including simulated data and an empirical RNA pool-seq dataset of *L. saxatilis* (42). We calculated the association of SNP allele counts with sample-specific covariates as BFs. For the ecotype divergence test, we used a binary variable for ecotype membership (Crab ecotype = −1 and Wave ecotype = 1), while for the shore-level test, we used a binary variable for shore-level membership (high shore = −1 and low shore = 1). The covariate test also accounts for the neutral expectation with population covariance matrix Ω. BF is expressed in deciban units [dB = $10\log_{10}$(BF)], and we used Jeffreys' rule to qualify the strength of the SNP-covariate associations: strong (10 < dB < 15), very strong (15 < dB < 20), and decisive (dB > 20), as suggested in the BayPass manual. Only the latter were classified (dB > 20) as outlier SNPs. Because BayPass calculates metrics per SNP and is a very computationally intensive process, we divided our allele count tables into pseudoreplicated datasets of 25,000 random SNPs. This allowed us to test for consistency in the calculation of the Ω matrix and outlier metrics and to reduce the effect of physical linkage in each dataset. We confirmed that the value distance of Ω matrices between replicates and between empirical and simulated datasets was very low (fmd.dist <1 as recommended in BayPass manual), using the R function fmd.dist included in BayPass. We also confirmed that all the obtained BF values and summary statistics across replicates had high correlations using the cor function ($r^2 > 0.98$). For each BayPass run, we used the default parameters of 20 pilot runs, each of 1000 steps, 5000 steps of burn-in, and 1000 steps after burn-in with a thinning sampling of 25. We confirmed convergence of the chains directly from the output of the program.

Cluster separation score analysis

We used the CSS implemented by Jones *et al.* (11) as a measure of between-ecotype genetic distance, weighted by between-localities genetic distance (within each ecotype). However, we used the Euclidean distance between PCA scores as a proxy for genetic distance, as opposed to the multidimensional scaling (MDS) of D_{XY} scores used by Jones *et al.* (11). Using a custom script, we first performed a PCA with all the SNPs contained within a map position. Then, we calculated the Euclidian distance between every pair of Crab versus Wave ecotype, and between every pair of localities within each of the ecotypes (the same procedure was followed for the Low-High axis of divergence). We used the first four axes of the PCA, as they contained most of the genetic variation (>60%). We calculated CSS from the PCA genetic distances following the formula in the Supplementary Materials and Jones *et al.* [(11), p. 5].

To identify map positions with significant CSS values, we compared the bootstrap values for each map position against bootstrap values across the whole genome. First, we estimated a genome-wide random expectation by calculating 10,000 CSS across the genome, each iteration consisting of choosing a random map position, randomly sampling the same number of SNPs contained in that map position, and performing the CSS calculation for that random sample. After 10,000 iterations, we obtained a genome-wide expected distribution of CSS values. Second, we calculated the observed CSS per map position and their CIs with 100 random subsamples with replacement of SNPs within a given map position. Last, to assign significance, we considered that the CSS of a given map position was significant (i.e., strong shared divergence between ecotypes across all localities) if its CSS 95% CI did not overlap with the CSS 95% CI of the genome-wide random expectation.

Clustering patterns of shared differentiation

We used our three metrics of genetic differentiation (F_{ST}), covariation with habitat (BayPass), and genetic distance (CSS) to identify clusters of high "counts" (see below) compared to a random expectation at two genomic scales. At the LG scale, we compared LG counts to those of a genome-wide permutation, and at the map position scale, we compared map position counts to those of a permutation within the same LG. We defined a significant LG or map position if its count was higher than the 95th percentile of the distribution for the respective random permutation. In each of these tests, counts are defined differently: (i) For the F_{ST} analysis, we estimated the mean count of populations that shared a given outlier SNP (outlier loci are those within the top 1% of the F_{ST} distribution). (ii) For the BayPass analysis, we counted the total number of outlier SNPs. (iii) Last, for the CSS analysis, we used the total count of map positions that were classified as outliers.

Chromosomal inversion analyses per locality

We used the exact coordinates of the candidate inversions from Faria *et al.* (22) to define the inverted genomic regions. To define the collinear genomic background, we added a buffer area of 2 cM around inversion regions to allow slight imprecisions on the positions of the breakpoints.

For each locality, we calculated the mean Crab-Wave and/or Low-High genetic differentiation (F_{ST}) from all the 500-bp window values contained in each of the inversion regions (i.e., observed values). Then, we draw a random distribution of mean F_{ST} values from the collinear background for the same locality. We obtained the random

distribution by randomly sampling genomic blocks of the same size as the inversion region and then randomly sampled the same number of 500-bp window values as those contained in the inversion region, repeating 200 iterations. We considered that the observed values were significant if they were higher than the 99th percentile of the distribution of random values.

DI analysis

We developed a metric called the "DI" to test whether differentiation for each map position across the genome was consistent with the directionality expectations of ecotype or shore-level divergence. For each SNP within a map position, we defined the reference allele as the one with higher average frequency over the two Spanish high-shore pools than over the two Spanish low-shore pools. We then averaged the frequencies of this allele over loci in the high-shore pool ($\bar{p}_{Hi}$) and the low shore pool ($\bar{p}_{Li}$) from each locality (i) and took the difference, $\mathrm{DI}_i = \bar{p}_{Hi} - \bar{p}_{Li}$. We also calculated the average DI across all pools ($\overline{\mathrm{DI}} = \bar{\bar{p}}_H - \bar{\bar{p}}_L$), as reference.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/12/eaav9963/DC1

Fig. S1. Distribution of candidate inversions regions.
Fig. S2. Comparison of allelic frequencies between individual sequencing and pool sequencing.
Fig. S3. Comparison of alternative genetic differentiation metrics for pool-seq data.
Fig. S4. Comparison of outlier locus detection for alternative metrics.
Fig. S5. Histograms of Crab-Wave genome-wide genetic differentiation.
Fig. S6. Random iterations of outlier loci PCAs.
Fig. S7. Network plot of outlier sharing at window level.
Fig. S8. Loci included in the linkage map versus unmapped loci.
Fig. S9. Genomic landscape of differentiation between ecotypes.
Fig. S10. Outlier sharing per LG.
Fig. S11. BayPass covariance matrix for the Crab-Wave analysis.
Fig. S12. Comparison between BayPass and F_{ST} Crab-Wave outlier loci.
Fig. S13. Genomic landscape of outlier sharing across demographic groups.
Fig. S14A. DI plots for LGs 1 to 9.
Fig. S14B. DI plots for LGs 10 to 17.
Fig. S15. Mixed distribution model of coverage depth.
Fig. S16. Crab and Wave ecotypes from Spain, United Kingdom, and Sweden.
Fig. S17. Spanish "Crab habitat."
Fig. S18. Spanish Wave/Low habitat exposed to strong wave-action.
Fig. S19. U.K. Crab/Low (yellow arrow) and Wave/High (red arrow) habitats.
Fig. S20. Sweden Crab (yellow arrow) and Wave (red arrow) environments.
Fig. S21. The same site as in Fig. 5 during a storm showing wave splash and breaking waves.
Table S1. Evidence to support the presence of chromosomal inversions.
Table S2. Pool-seq samples.
Table S3. LMM summary.
Supplementary Text: Description of the habitat characteristics
References (65–70)

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. D. Schluter, L. M. Nagel, Parallel speciation by natural selection. *Am. Nat.* **146**, 292–301 (1995).
2. D. I. Bolnick, R. D. H. Barrett, K. B. Oke, D. J. Rennison, Y. E. Stuart, (Non) parallel evolution. *Annu. Rev. Ecol. Evol. Syst.* **49**, 303–330 (2018).
3. K. B. Oke, C. Rolshausen, C. LeBlond, A. P. Hendry, How parallel is parallel evolution? A comparative analysis in fishes. *Am. Nat.* **190**, 1–16 (2017).
4. R. B. Langerhans, Predictability and parallelism of multitrait adaptation. *J. Hered.* **109**, 59–70 (2017).
5. K. Johannesson, M. Panova, P. Kemppainen, C. André, E. Rolán-Alvarez, R. K. Butlin, Repeated evolution of reproductive isolation in a marine snail: Unveiling mechanisms of speciation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 1735–1747 (2010).
6. G. L. Conte, M. E. Arnegard, C. L. Peichel, D. Schluter, The probability of genetic parallelism and convergence in natural populations. *Proc. Biol. Sci.* **279**, 5039–5047 (2012).
7. P. L. Ralph, G. Coop, Convergent evolution during local adaptation to patchy landscapes. *PLOS Genet.* **11**, e1005630 (2015).
8. V. Soria-Carrasco, Z. Gompert, A. A. Comeault, T. E. Farkas, T. L. Parchman, J. S. Johnston, C. A. Buerkle, J. L. Feder, J. Bast, T. Schwander, S. P. Egan, B. J. Crespi, P. Nosil, Stick insect genomes reveal natural selection's role in parallel speciation. *Science* **344**, 738–742 (2014).
9. A. M. Westram, M. Panova, J. Galindo, R. K. Butlin, Targeted resequencing reveals geographical patterns of differentiation for loci implicated in parallel evolution. *Mol. Ecol.* **25**, 3169–3186 (2016).
10. F. Roda, G. M. Walter, R. Nipper, D. Ortiz-Barrientos, Genomic clustering of adaptive loci during parallel evolution of an Australian wildflower. *Mol. Ecol.* **26**, 3687–3699 (2017).
11. F. C. Jones, M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya; Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team, E. S. Lander, F. Di Palma, K. Lindblad-Toh, D. M. Kingsley, The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).
12. Y. E. Stuart, T. Veen, J. N. Weber, D. Hanson, M. Ravinet, B. K. Lohman, C. J. Thompson, T. Tasneem, A. Doggett, R. Izen, N. Ahmed, R. D. H. Barrett, A. P. Hendry, C. L. Peichel, D. I. Bolnick, Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nat. Ecol. Evol.* **1**, 0158 (2017).
13. J. A. Raeymaekers, A. Chaturvedi, P. I. Hablützel, I. Verdonck, B. Hellemans, G. E. Maes, L. De Meester, F. A. M. Volckaert, Adaptive and non-adaptive divergence in a common landscape. *Nat. Commun.* **8**, 267 (2017).
14. S. Yeaman, Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E1743–E1751 (2013).
15. A. D. Foote, Sympatric speciation in the genomic era. *Trends Ecol. Evol.* **33**, 85–95 (2017).
16. E. V. Koonin, Evolution of genome architecture. *Int. J. Biochem. Cell Biol.* **41**, 298–306 (2009).
17. A. A. Hoffmann, L. H. Rieseberg, Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* **39**, 21–42 (2008).
18. R. Faria, A. Navarro, Chromosomal speciation revisited: Rearranging theory with pieces of evidence. *Trends Ecol. Evol.* **25**, 660–669 (2010).
19. M. Kirkpatrick, How and why chromosome inversions evolve. *PLOS Biol.* **8**, e1000501 (2010).
20. R. D. H. Barrett, D. Schluter, Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).
21. M. Kirkpatrick, B. Barrett, Chromosome inversions, adaptive cassettes and the evolution of species' ranges. *Mol. Ecol.* **24**, 2046–2055 (2015).
22. R. Faria, P. Chaube, H. E. Morales, T. Larsson, A. R. Lemmon, E. M. Lemmon, M. Rafajlović, M. Panova, M. Ravinet, K. Johannesson, A. M. Westram, R. K. Butlin, Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes. *Mol. Ecol.* **28**, 1375–1393 (2018).
23. J. M. I. Barth, P. R. Berg, P. R. Jonsson, S. Bonanomi, H. Corell, J. Hemmer-Hansen, K. S. Jakobsen, K. Johannesson, P. E. Jorde, H. Knutsen, P.-O. Moksnes, B. Star, N. C. Stenseth, H. Svedäng, S. Jentoft, C. André, Genome architecture enables local adaptation of Atlantic cod despite high connectivity. *Mol. Ecol.* **26**, 4452–4466 (2017).
24. K. Johannesson, What can be learnt from a snail? *Evol. Appl.* **9**, 153–165 (2016).
25. R. K. Butlin, M. Saura, G. Charrier, B. Jackson, C. André, A. Caballero, J. A. Coyne, J. Galindo, J. W. Grahame, J. Hollander, P. Kemppainen, M. Martínez-Fernández, M. Panova, H. Quesada, K. Johannesson, E. Rolán-Alvarez, Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution* **68**, 935–949 (2014).
26. K. Janson, Selection and migration in two distinct phenotypes of *Littorina saxatilis* in Sweden. *Oecologia* **59**, 58–61 (1983).
27. D. G. Reid, *Systematics and Evolution of Littorina* (Ray Society, 1996).
28. E. Rolán-Alvarez, J. Erlandsson, K. Johannesson, R. Cruz, Mechanisms of incomplete prezygotic reproductive isolation in an intertidal snail: Testing behavioural models in wild populations. *J. Evol. Biol.* **12**, 879–890 (1999).
29. M. Panova, J. Boström, T. Hofving, T. Areskoug, A. Eriksson, B. Mehlig, T. Mäkinen, C. André, K. Johannesson, Extreme female promiscuity in a non-social invertebrate species. *PLOS ONE* **5**, e9640 (2010).
30. K. Johannesson, R. Butlin, What explains rare and conspicuous colours in a snail? A test of time-series data against models of drift, migration or selection. *Heredity* **118**, 21 (2017).
31. A. M. Westram, M. Rafajlović, P. Chaube, R. Faria, T. Larsson, M. Panova, M. Ravinet, A. Blomberg, B. Mehlig, K. Johannesson, R. Butlin, Clines on the seashore: The genomic architecture underlying rapid divergence in the face of gene flow. *Evol. Lett.* **2**, 297–309 (2018).

32. K. Johannesson, B. Johannesson, U. Lundgren, Strong natural selection causes microscale allozyme variation in a marine snail. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2602–2606 (1995).

33. I. M. Sokolova, A. I. Granovitch, V. J. Berger, K. Johannesson, Intraspecific physiological variability of the gastropod *Littorina saxatilis* related to the vertical shore gradient in the White and North Seas. *Mar. Biol.* **137**, 297–308 (2000).

34. M. Panova, K. Johannesson, Microscale variation in Aat (aspartate aminotransferase) is supported by activity differences between upper and lower shore allozymes of *Littorina saxatilis*. *Mar. Biol.* **144**, 1157–1164 (2004).

35. M. Panova, A. M. H. Blakeslee, A. W. Miller, T. Mäkinen, G. M. Ruiz, K. Johannesson, C. André, Glacial history of the North Atlantic marine snail, *Littorina saxatilis*, inferred from distribution of mitochondrial DNA lineages. *PLOS ONE* **6**, e17511 (2011).

36. J. Galindo, J. Grahame, R. Butlin, An EST-based genome scan using 454 sequencing in the marine snail *Littorina saxatilis*. *J. Evol. Biol.* **23**, 2004–2016 (2010).

37. M. Ravinet, A. Westram, K. Johannesson, R. Butlin, C. André, M. Panova, Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Mol. Ecol.* **25**, 287–305 (2016).

38. A. M. Westram, J. Galindo, M. A. Rosenblad, J. W. Grahame, M. Panova, R. K. Butlin, Do the same genes underlie parallel phenotypic divergence in different *Littorina saxatilis* populations? *Mol. Ecol.* **23**, 4603–4616 (2014).

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

40. T. Kess, J. Galindo, E. G. Boulding, Genomic divergence between Spanish *Littorina saxatilis* ecotypes unravels limited admixture and extensive parallelism associated with population history. *Ecol. Evol.* **8**, 8311–8327 (2018).

41. H. E. Morales, A. Pavlova, N. Amos, R. Major, A. Kilian, C. Greening, P. Sunnucks, Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nat. Ecol. Evol.* **2**, 1258–1267 (2018).

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

43. G. Coop, D. Witonsky, A. Di Rienzo, J. K. Pritchard, Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**, 1411–1423 (2010).

44. P. Jay, A. Whibley, L. Frézal, M. Á. Rodríguez de Cara, R. W. Nowell, J. Mallet, K. K. Dasmahapatra, M. Joron, Supergene evolution triggered by the introgression of a chromosomal inversion. *Curr. Biol.* **28**, 1839–1845.e3 (2018).

45. R. Faria, K. Johannesson, R. K. Butlin, A. M. Westram, Evolving inversions. *Trends Ecol. Evol.* **34**, 239–248 (2019).

46. R. K. Butlin, Recombination and speciation. *Mol. Ecol.* **14**, 2621–2635 (2005).

47. S. Bassham, J. Catchen, E. Lescak, F. A. von Hippel, W. A. Cresko, Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. *Genetics* **209**, 921–939 (2018).

48. M. Wellenreuther, L. Bernatchez, Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evol.* **33**, 427–444 (2018).

49. D. Ayala, S. Zhang, M. Chateau, C. Fouet, I. Morlais, C. Costantini, M. W. Hahn, N. J. Besansky, Association mapping desiccation resistance within chromosomal inversions in the African malaria vector *Anopheles gambiae*. *Mol. Ecol.* **28**, 1333–1342 (2018).

50. S. Renaut, G. L. Owens, L. H. Rieseberg, Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Mol. Ecol.* **23**, 311–324 (2014).

51. S. Yeaman, A. C. Gerstein, K. A. Hodgins, M. C. Whitlock, Quantifying how constraints limit the diversity of viable routes to adaptation. *PLOS Genet.* **14**, e1007717 (2018).

52. M. Ravinet, R. Faria, R. K. Butlin, J. Galindo, N. Bierne, M. Rafajlović, M. A. F. Noor, B. Mehlig, A. M. Westram, Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *J. Evol. Biol.* **30**, 1450–1477 (2017).

53. K. R. Elmer, S. Fan, H. Kusche, M. Luise Spreitzer, A. F. Kautt, P. Franchini, A. Meyer, Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nat. Commun.* **5**, 5168 (2014).

54. P.-A. Gagnaire, S. A. Pavey, E. Normandeau, L. Bernatchez, The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution* **67**, 2483–2497 (2013).

55. J. I. Meier, D. A. Marques, C. E. Wagner, L. Excoffier, O. Seehausen, Genomics of parallel ecological speciation in Lake Victoria cichlids. *Mol. Biol. Evol.* **35**, 1489–1506 (2018).

56. M. Panova, H. Aronsson, R. A. Cameron, P. Dahl, A. Godhe, U. Lind, O. Ortega-Martinez, R. Pereyra, V. M. Tesson, A.-L. Wrange, A. Blomberg, K. Johannesson, DNA extraction protocols for whole-genome sequencing in marine organisms, in *Marine Genomics: Methods and Protocols*. *Methods in Molecular Biology*, S. J. Bourlat, Ed. (Springer, 2016), pp. 13–44.

57. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

58. R. Kofler, A. M. Langmüller, P. Nouhaud, K. A. Otte, C. Schlötterer, Suitability of different mapping algorithms for genome-wide polymorphism scans with pool-seq data. *G3* **6**, 3507–3515 (2016).

59. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

60. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/ map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

61. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

62. R Development Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).

63. R. Kofler, R. V. Pandey, C. Schlötterer, PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**, 3435–3436 (2011).

64. D. L. Hartl, A. G. Clark, *Principles of Population Genetics* (Taylor & Francis, 2007), pp. 544–545.

65. J. T. Carlton, A. N. Cohen, Periwinkle's progress: The Atlantic snail *Littorina saxatilis* (Mollusca: Gastropoda) establishes a colony on a Pacific shore. *Veliger* **41**, 333–338 (1998).

66. B. Johannesson, Shell morphology of *Littorina saxatilis Olivi*: The relative importance of physical factors and predation. *J. Exp. Mar. Biol. Ecol.* **102**, 183–195 (1986).

67. I. Sokolova, H. Pörtner, Temperature effects on key metabolic enzymes in *Littorina saxatilis* and *L. obtusata* from different latitudes and shore levels. *Mar. Biol.* **139**, 113–126 (2001).

68. K. Johannesson, B. Johannesson, E. Rolán-Alvarez, Morphological differentiation and genetic cohesiveness over a microenvironmental gradient in the marine snail *Littorina saxatilis*. *Evolution* **47**, 1770–1787 (1993).

69. E. Rolán-Alavarez, M. Carballo, J. Galindo, P. Morán, B. Fernández, A. Caballero, R. Cruz, E. G. Boulding, K. Johannesson, Nonallopatric and parallel origin of local reproductive barriers between two snail ecotypes. *Mol. Ecol.* **13**, 3415–3424 (2004).

70. J. W. Grahame, C. S. Wilding, R. K. Butlin, Adaptation to a steep environmental gradient and an associated barrier to gene exchange in *Littorina saxatilis*. *Evolution* **60**, 268–278 (2006).