Evolution Letters, 2024, **8**(4), 575–586

https://doi.org/10.1093/evlett/qrae014
Advance access publication 23 April 2024
Letter

Coupling of twelve putative chromosomal inversions maintains a strong barrier to gene flow between snail ecotypes

Alan Le Moan[1,2], Sean Stankowski[3], Marina Rafajlović[1], Olga Ortega-Martinez[1], Rui Faria[4,5], Roger K. Butlin[1,6], Kerstin Johannesson[1]

[1]Department of Marine Sciences, Tjärnö Marine Laboratory, University of Gothenburg, 452 96 Strömstad, Sweden,
[2]Adaptation et Diversité en Milieu Marin, UMR7144, Station Biologique de Roscoff, Sorbonne Université, 29680 Roscoff, France,
[3]Institute of Science and Technology Austria, 3 21 44 Klosterneuburg, Austria,
[4]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal,
[5]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal,
[6]Ecology and Evolutionary Biology, School of Biosciences, University of Sheffield, Sheffield S10 2TN, United Kingdom

Corresponding authors: Department of Marine Sciences, Tjärnö Marine Laboratory, University of Gothenburg, 452 96 Strömstad, Sweden. Email: alan.le.moan@gmail.com; Department of Marine Sciences, Tjärnö Marine Laboratory, University of Gothenburg, 452 96 Strömstad, Sweden. Email: kerstin.johannesson@gu.se

Abstract

Chromosomal rearrangements can lead to the coupling of reproductive barriers, but whether and how they contribute to the completion of speciation remains unclear. Marine snails of the genus *Littorina* repeatedly form hybrid zones between populations segregating for multiple inversion arrangements, providing opportunities to study their barrier effects. Here, we analyzed 2 adjacent transects across hybrid zones between 2 ecotypes of *Littorina fabalis* ("large" and "dwarf") adapted to different wave exposure conditions on a Swedish island. Applying whole-genome sequencing, we found 12 putative inversions on 9 of 17 chromosomes. Nine of the putative inversions reached near differential fixation between the 2 ecotypes, and all were in strong linkage disequilibrium. These inversions cover 20% of the genome and carry 93% of divergent single nucleotide polymorphisms (SNPs). Bimodal hybrid zones in both transects indicated that the 2 ecotypes of *Littorina fabalis* maintain their genetic and phenotypic integrity following contact. The bimodality reflects the strong coupling between inversion clines and the extension of the barrier effect across the whole genome. Demographic inference suggests that coupling arose during a period of allopatry and has been maintained for > 1,000 generations after secondary contact. Overall, this study shows that the coupling of multiple chromosomal inversions contributes to strong reproductive isolation. Notably, 2 of the putative inversions overlap with inverted genomic regions associated with ecotype differences in a closely related species (*Littorina saxatilis*), suggesting the same regions, with similar structural variants, repeatedly contribute to ecotype evolution in distinct species.

Keywords: speciation, hybrid zones, chromosomal rearrangements, linkage disequilibrium, *Littorina*

Lay Summary

The details of biological speciation remain enigmatic. In particular, it is unclear how initial barriers to gene flow are reinforced to complete reproductive isolation. Local barriers in the genome, established by divergent selection and/or stochastic events, need to be coupled together to form genome-wide barriers. Chromosomal rearrangements, such as inversions, establish the coupling of large numbers of genes and facilitate divergent adaptation despite gene flow. Here we report ecotypes of a marine snail in which 12 putative inversions covering 20% of the genome have established barriers in contact zones between two ecotypes. Despite being distributed over nine different chromosomes, the inversions are all strongly genetically associated and this strengthens the overall barrier to gene flow that extends outside the inversions.

Introduction

Understanding reproductive isolation from genome-wide polymorphism data has become a key goal in speciation research (Ravinet et al., 2017). Reproductive isolation evolves in response to divergent natural and sexual selection, and through mutation order effects, both of which lead to the accumulation of barriers to gene flow along the genome (Abbott et al., 2013; Nosil & Flaxman, 2011). The strength of the barrier experienced by a locus depends on its proximity to a causal barrier locus, the fitness effects of the barrier loci, the genome-wide distribution of these barriers, and their interaction in recombinant genomes (Satokangas et al., 2020). These recombinant genomes can be found in hybrid zones, which represent natural laboratories in which to study reproductive isolation (Barton & Hewitt, 1989). When reproductive isolation

Received January 16, 2024; revisions received March 7, 2024; accepted March 29, 2024

is weak, a high level of admixture between the parental lineages is expected in the center of the hybrid zone, which leads to the formation of a unimodal hybrid zone (Barton & Gale, 1993). In contrast, when reproductive isolation is very strong, hybrids are rare and parental genotypes coexist in the center of the hybrid zone, which leads to a bimodal hybrid zone (Gay et al., 2008; Harrison & Bogdanowicz, 1997). Studying the formation of bimodal hybrid zones can provide an important step toward understanding the completion of speciation (Jiggins & Mallet, 2000).

The formation of a bimodal hybrid zone requires the coupling of many barrier loci, that is the build-up of strong linkage disequilibrium (LD) between pre- and post-zygotic barriers to gene flow (Butlin & Smadja, 2018). This coupling can easily occur during an allopatric phase of divergence prior to secondary contact and hybrid zone formation, or, through reinforcement when divergence occurs in the face of gene flow (Butlin, 1987). In addition, barriers of different geographical origins can overlap in space due to the attraction of allelic clines to one another or to the same ecotone, a process coined "spatial coupling" (Barton, 1979; Bierne et al., 2011). Coupling can also be facilitated when individual barrier loci cluster within genomic regions where recombination is suppressed, for example, within chromosomal inversions (Kirkpatrick, 2010). This genomic coupling effectively combines many individual barrier loci into one large effect locus (Schaal et al., 2022), and also extends local barrier effects across larger segments of chromosomes (Kirkpatrick & Barton, 2006; Rieseberg, 2001). By maintaining LD among barrier loci in the presence of gene flow, such rearrangements facilitate the spread of locally adaptive alleles (Kirkpatrick & Barton, 2006), promote the accumulation of genetic incompatibilities (Navarro & Barton, 2003), and maintain local adaptation. For instance, previous studies have found pronounced local barrier effects due to direct selection acting on chromosomal inversions during the evolution of partially isolated ecotypes (e.g., deer mouse (Harringmeyer & Hoekstra, 2022), Atlantic cod (Matschiner et al., 2022), seahorse (Meyer et al., 2024), stick insects (Nosil et al., 2023), and sunflower (Huang et al., 2020)). Yet, little evidence showed that these barriers generate substantial genome-wide reproductive isolation.

Marine snails of the genus *Littorina* are emerging models for studying the role of chromosomal inversions in speciation (Faria et al., 2019; Johannesson et al., 2024; Le Moan et al., 2023; Reeve et al., 2023; Westram et al., 2018). *Littorina* includes several species that form ecotypes that differ in multiple phenotypic traits (including shell size, shell morphology, and behavior). The ecotypes evolve repeatedly across similar types of meter-scale environmental gradients (Johannesson et al., 2010, 2024). In *L. saxatilis*, the parallel evolution of "wave" and "crab" ecotypes involves more than 10 chromosomal inversions (Faria et al., 2019; Morales et al., 2019). Many of these show signatures of divergent selection, and contribute to phenotypic differences between the ecotypes (Koch et al., 2021, 2022; Westram et al., 2018, 2021). Despite steep frequency clines across ecotype hybrid zones, the coupling of the different inversions is weak in *L. saxatilis*, and hybrid zones appear unimodal, indicating that the overall barrier to gene flow is weak (Janson & Sundberg, 1983; Westram et al., 2018, 2021). In addition, several inversions remain polymorphic within one or both ecotypes, and they only explain roughly half of the phenotypic variation among snails (Koch et al., 2022). Therefore, despite much of the genetic divergence underlying local adaptation being associated with the inversions, it remains unclear whether the establishment of these chromosomal rearrangements also generates genome-wide barriers.

A closely related species, *Littorina fabalis*, living adjacent to *L. saxatilis* but in the seaweeds on European shores, forms a dwarf and a large ecotype associated with different wave exposure (Galindo et al., 2021; Tatarenkov & Johannesson, 1999). Field experiments using transplant and tethering of snails on algae and rocks underneath showed that large snails have better survival in the exposed environment because they better resist crab predation after being dislodged from the algae during heavy wave action (Kemppainen et al., 2005). Sharp allele frequency clines separate the ecotypes at one out of four highly polymorphic allozyme loci (Tatarenkov & Johannesson, 1994, 1998). In contrast to the unimodal hybrid zones of *L. saxatilis*, *L. fabalis* shows some evidence of bimodality in ecotype contact zones, with deficiency of heterozygotes, and a bimodal size distribution (Tatarenkov & Johannesson, 1998). Recently, we found that the divergent allozyme gene is located inside a large putative chromosomal inversion (Le Moan et al., 2023), and preliminary data suggested the presence of additional inversions. Here, we study the genomic landscape of differentiation between the two *L. fabalis* ecotypes across the whole genome using 295 snails sampled along parallel transects from two hybrid zones between the large and the dwarf ecotypes of *L. fabalis*, located on the northern (167 snails sampled) and southern shores (128 snails) of a small island on the Swedish west coast (Figure 1a). Using low-coverage whole genome sequencing (lcWGS), we found that most genetic differences between the ecotypes were localized inside 12 large haplotype blocks that are probably associated with large chromosomal inversions. While we refer to them as "inversions" throughout the text, it should be acknowledged that we currently lack direct molecular or cytogenetic evidence for the nature of the recombination suppression and so we consider them putative inversions at this stage. We map the inversion arrangement frequencies across the hybrid zones to assess their barrier effects and show that the coupling between the differentially fixed inversions results in strong reproductive isolation associated with the formation of bimodal hybrid zones. We furthermore explore the evolutionary origin of the two ecotypes of *L. fabalis* and provide a comparative genomic analysis of ecotypic divergence with the related *L. saxatilis*.

Results

Population structure and the landscape of differentiation

The two ecotypes of *L. fabalis* were clearly genetically distinct on the first axis of a principal component analysis (PC1), which explained a large part of the total variation (PC1 = 13.55%, PC2 < 1%, Supplementary Fig. S1) and showed very similar clines in the two transects (Figure 1b). The average F_{ST} between ecotypes, calculated between 20 snails collected from each transect end, was 0.090 (0.092 and 0.088 across the northern and southern shore transects, respectively). Rather than being randomly distributed across the genome, we found that highly differentiated loci formed clusters on nine of the 17 chromosomes when mapping the sequences to the *L. saxatilis* reference genome (Figure 1c and d).

The signatures of chromosomal rearrangements

The clusters of high differentiation across the genome might be caused by large chromosomal inversions. Therefore, we looked for signatures commonly associated with inversions (Faria et al., 2019; Le Moan et al., 2021; Mérot et al., 2020). First, we

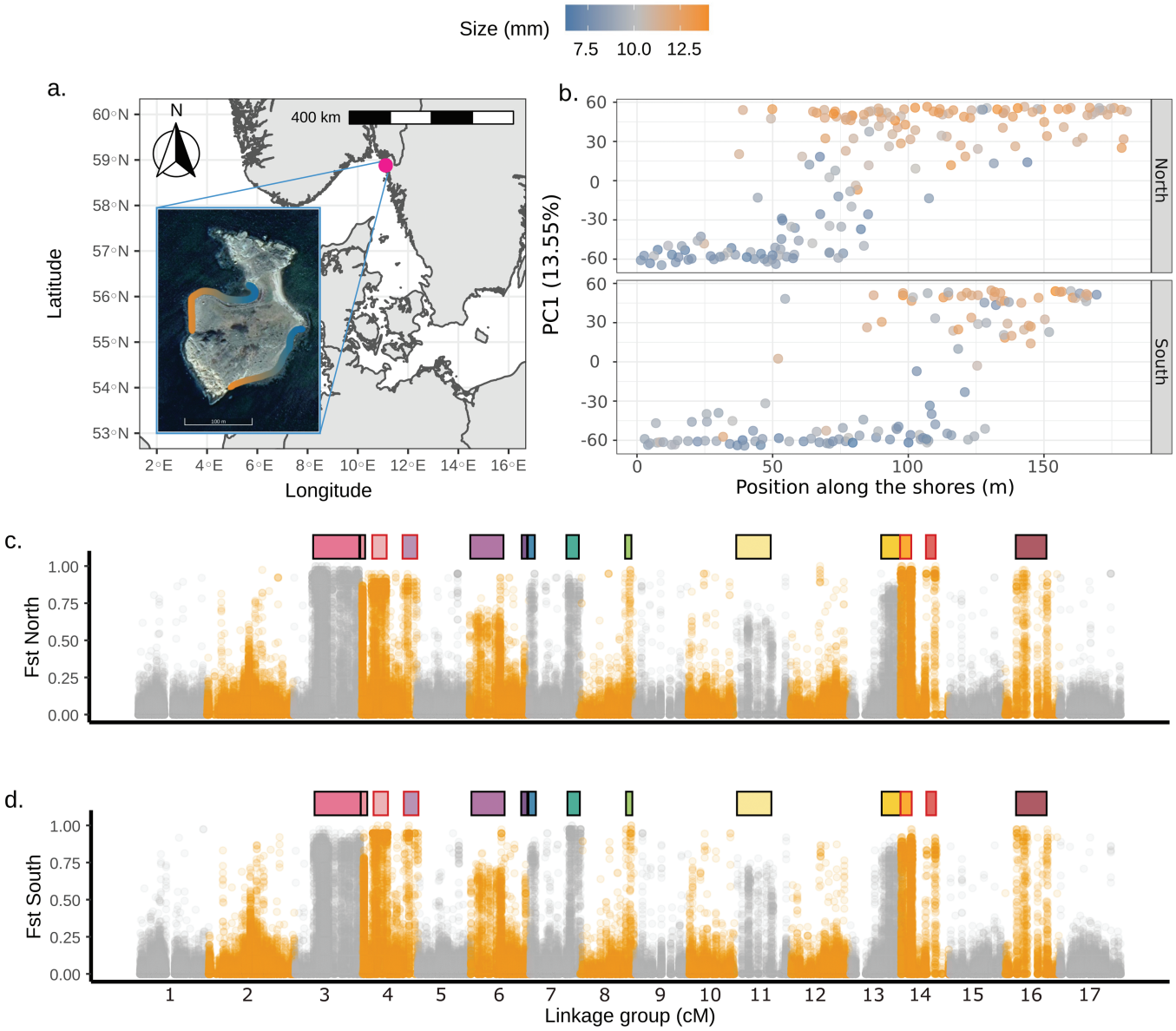

Figure 1. Population structure analyses showing: (a) map of the sampling locations; (b) the PC1 score of a PCA performed on the 295 snails based on genotypes at 110k SNPs (filtered set, see *Methods*) with minor allele frequency (MAF) of > 5% that are placed on the linkage map. The individual PC1 scores were plotted against the positions of the snails on the northern (top) and southern (bottom) transects; (c) and (d) Manhattan plots of pairwise F_{ST} calculated between 20 snails sampled from the opposite ends of each transect (top: northern and bottom: southern transect). In (a) and (b), the color gradient represents the size of individual snails (following the legend above the graph). In (c) and (d), the orange and grey show the limits of different linkage groups (LGs) on which the contigs were ordered by linkage map position, and the horizontal bars on top of the plot show the positions of the 14 blocks (one color per block) corresponding to 12 putative chromosomal inversions (after grouping the two pairs of blocks with LD of 1 from LG4 and LG14 encircled in red, see text and Supplementary Fig. S2).

found strong LD among SNPs located several cM apart, revealing 14 LD blocks ranging from 1 to 33 Mbp (Supplementary Fig. S2 and Supplementary Table S1). These blocks were characterized by plateaus of similarly high LD typical of large inversions (Supplementary Fig. S3). The average LD was much stronger within blocks than the average LD found outside ($r^2 = 0.387$ and 0.006, respectively, Supplementary Fig. S4 and Supplementary Table S1). Second, local PCAs performed on the SNPs within each block revealed three distinct clusters of individuals on PC1 (Supplementary Figs. S5–S7). Observed heterozygosity was approximately twice as high in the central cluster compared with those on either side (Supplementary Figs. S8–S10 and Supplementary Table S2). Moreover, hundreds of SNPs were differentially fixed between the two distant clusters (Supplementary Table S1). All of these observed patterns are hallmarks of chromosomal inversions. Third, as shown earlier, the largest block located on linkage group 3 (LG3) shows a "suspension bridge" pattern (Le Moan et al., 2023), providing further support for this block being an inversion. In some local PCAs (e.g., for LG14_inv2 and to some extent for LG7_inv, Supplementary Fig. S6), more complex patterns of clustering were observed, suggesting the presence of multiple overlapping polymorphic inversions (Faria et al., 2019). In addition, r^2=1 between separate LD blocks on LG4 and LG14 suggests that they are probably part of the same putative inversion, but have been broken up due to differences in synteny between the reference used (*L. saxatilis*) and *L. fabalis* (Supplementary Fig. S11).

LD between other pairs of rearrangements never reaches such high values (maximum r^2 = 0.67), which means that recombinants between pairs of rearrangements occurred and that they are indeed different arrangements. One additional island of high differentiation was visible near the center of LG2 (Figure 1). Here genomic signatures were different from those of the putative inversions (Supplementary Figs. S2–S5), and this may be a region of low recombination associated with the centromere (Mérot et al., 2021).

Consequences of chromosomal rearrangements for population structure

The 12 putative chromosomal inversions (after combining the two blocks on LG4 and the two on LG14) covered around 20% of the reference genome. The differentiation between ecotypes inside these putative inversions was an order of magnitude higher (mean F_{ST} ~ 0.3) than outside of them (mean F_{ST} ~ 0.03). In addition, they carried 94% of the SNPs with F_{ST} values above the 95% upper quantile of differentiation (i.e., $F_{ST} \geq 0.61$ and 0.59 in the northern and southern transect, respectively), and 92% of these SNPs were shared between the two transects. These results showed that 12 putative inversions carry most of the divergence between the two ecotypes in *L. fabalis*. Nevertheless, the ecotypes remained distinct after removing SNPs within the putative inversions (F_{ST} ~ 0.03), as illustrated by the first axis of a PCA performed on the collinear part of the genome (only capturing 2% of the variation; Supplementary Fig. S1b). Linkage disequilibrium between arrangements located on different linkage groups was generally very high ($r^2 = 0.21 < LD_{between} < 0.66$) with the exception of the inversions on LG6 and LG11 (Supplementary Fig. S11). This coupling of multiple large genomic barriers to gene flow is expected to increase the barrier effect of each inversion and to extend the barrier to other parts of the genome.

Arrangements and allelic clines

Sampling was conducted over continuous transects allowing exploration of the strength of the barrier effect provided by the putative inversions, and their coupling, from the steepness of the clines. We inferred arrangement-frequency clines at the 12 putative inversions, using information from clusters inferred in a local DAPC to obtain the karyotypes of each snail (Supplementary Fig. S7). As expected from strong coupling, all rearrangement clines showed similar trends with cline centers in close proximity to the phenotypic shift and with different arrangement frequencies reaching near fixation at transect ends in both shores, with the exception of arrangements on LG6 and LG11 that remained polymorphic in the dwarf ecotype (Figure 2a, Supplementary Table S3). In addition, SNPs found within the inversions showed similar trends to the rearrangement clines (Supplementary Fig. S12). Using an estimate of dispersal of $\sigma = 8.27 m.gen^{-1}$ (± 2.01 SE) directly inferred from the distances between 19 related individuals sampled along the transects (10 half-sib pairs and 1 full-sib pair (Supporting text p.3 and Supplementary Fig. S13)), we estimated that a coefficient of selection ranging from 0.02 to 0.09 was required to maintain the arrangement clines. All the differentially fixed rearrangements showed strong and significant deficiencies of heterokaryotypes near the center of the clines (inferred directly from the cline fitting, Supplementary Table S3). Nevertheless, all inversions were found as heterozygotes in the center of the hybrid zone. By representing the hybrid index against the heterozygosity calculated from the inversion karyotypes, we found that inversions were heterozygous in F1 and backcross individuals between the two ecotypes (i.e., individuals located on the top and along the sides of a triangle plot; Figure 2b), but that recombinant individuals from the two genetic backgrounds were largely missing (F2 or later generations of recombinant hybrids, located in the center of the triangle plot, Figure 2b).

Bimodal hybrid zones between ecotypes

The strong deficiency of heterokaryotypes and the absence of recombinant hybrids suggest that the parental lineages maintain their phenotypic and genetic integrity despite rare hybridization events. We used quantitative cline analyses to test this hypothesis by comparing a unimodal distribution of phenotypes or hybrid indexes in the center of the hybrid zone to a bimodal or trimodal distribution. We found evidence for bimodality in all three parameters explored: shell size, PC1 score based on the genome-wide polymorphism, hybrid index based on the inversion genotypes (histograms in Figure 2c–e). For each transect, all clines followed similar trends, being centered at 64 and 122 m in the northern and southern transects, respectively. In each case, multimodality was maintained in the centers of the clines (red group in histograms—Figure 2c–e and Supplementary Figs. S14–S19), with bimodal clines often providing similar fits to trimodal clines in the southern transect while trimodality was always the favored model in the northern transect (Supplementary Tables S4–S6). The higher number of hybrids found in the northern transect explains why trimodality was more often supported in this transect, but otherwise, the two transects were very similar. This multimodality shows that the *L. fabalis* ecotypes are substantially reproductively isolated, with low levels of hybridization when they co-occur.

Inversion coupling

The high transect-wide LD (Supplementary Fig. S11), the similar cline shapes among inversions (Figure 2a) and bimodal hybrid zone (Figure 2b) suggest that the inversions do no behave independently. This led to the question whether the inversions remain coupled in the center of the hybrid zone. We inferred that the LD between the differentially fixed inversions was six and 13 times higher in the northern and southern transect (Supplementary Table S7), respectively, than what was predicted based on the interaction between selection, dispersal, and recombination (detailed method in Supporting text of p.3). This strong coupling was also supported by the low coefficient of variation (CV) of the cline slopes across inversions ($CV_{north} = 0.21$, $CV_{south} = 0.10$, Supplementary Table S8), which is a good proxy for the coupling coefficient (Firneno et al., 2023). Here, the CV suggests a coupling coefficient close to 1, which means that both the direct effects of selection on each inversion, and the indirect effects of selection on each inversion on the others, contribute to maintaining frequency clines of the inversions and, thus, the bimodal hybrid zone. In addition, bimodality was also inferred in the collinear part of the genome when assessed from the PC1 score computed without the inversions (Supplementary Table S5), or a hybrid index computed with outlier SNPs found outside the inversions (Supplementary Table S6). This suggests that the barrier effect provided by the coupling of the inversions extends into the collinear part of the genome.

Origin of the hybrid zones in *L. fabalis*

We used demographic inference to explore the divergence history of the ecotypes in *L. fabalis* (De Jode et al., 2023). We asked whether the divergence between ecotypes was better explained by a model of divergence in the face of gene flow (primary divergence), or was the result of a demographic history involving a period of isolation

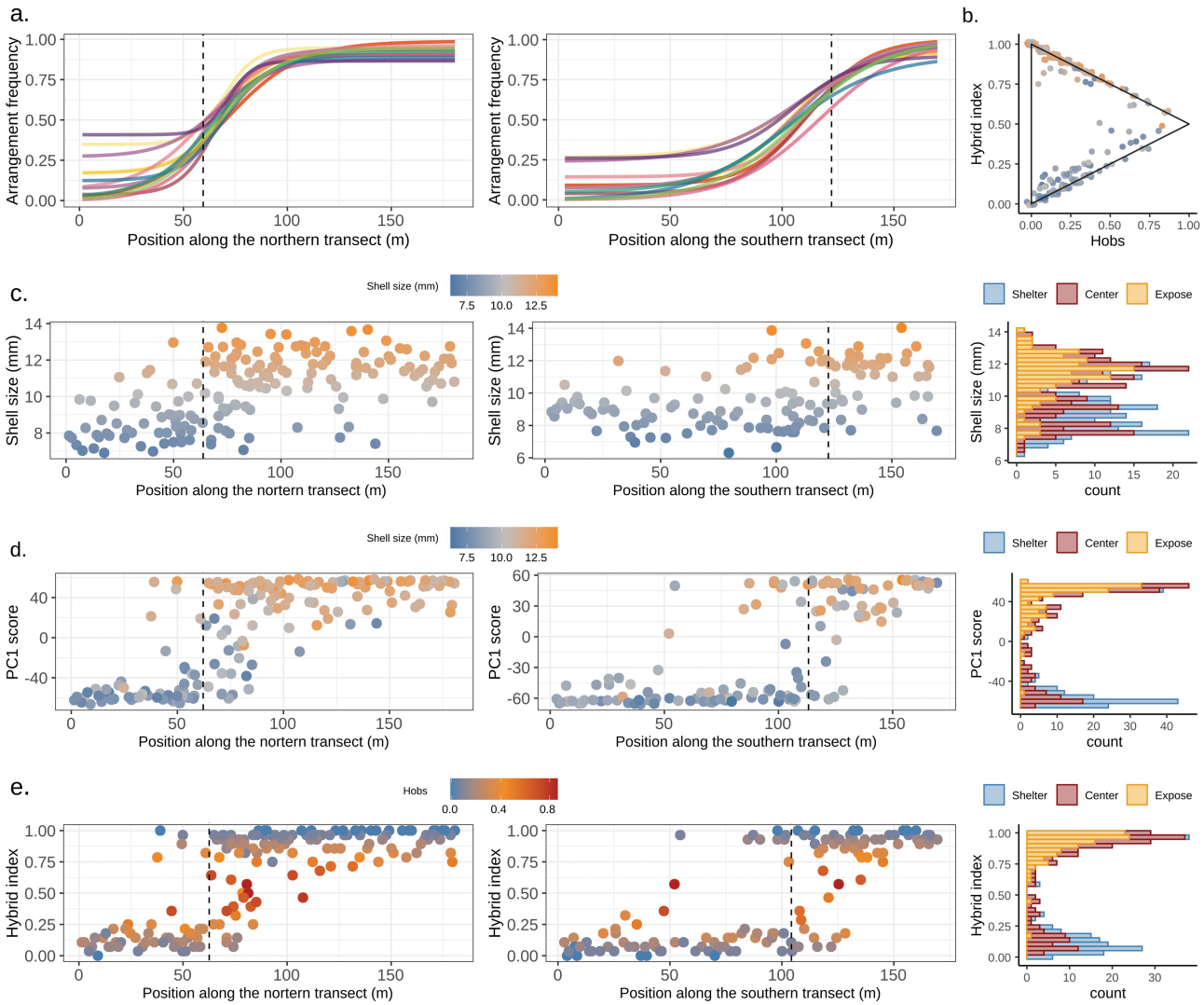

Figure 2. Cline analyses for the northern (left) and southern transect (middle). Graph (a) shows frequency clines inferred from the 12 putative chromosomal inversions, which are colored according to the color rectangles used in Figure 1c and d. The dotted vertical line in each of the top graphs shows the center of the size phenotypic cline. Graph (b) shows the triangle plot for hybrid detection with F1 hybrids expected at the "top right" of the triangle (both transects pooled). The graphs (c)–(e), respectively, show the shell size variation, the PC1 score variation (as in Figure 1b), and, the hybrid index, here calculated from the inversion karyotypes. The dashed line shows the center of the cline (*SI Appendix*, Supplementary Tables S4–S6). All the histograms ((c)–(e) on the right, both transects pooled) separated individuals from the middle of the hybrid zone in red (center of the cline ± the width/2) from individuals located on left of the cline in the sheltered area (in blue) and on the right of the cline in the exposed area (in yellow), and all show a clear bimodality of each parameter. In (c) and (d), dots are colored by the shell size gradient, and in (e), dots are colored by the observed heterozygosity (H_{obs}).

prior to the current contact. We found support for secondary contact in all inferences conducted (using all SNPs, using collinear SNPs, and using inversion SNPs only; Figure 3 and Supplementary Table S9). The inferences made with collinear or inversion SNPs alone supported an ancestral population expansion and a heterogeneous effective population size along the genome (SC_ae_2N). The best models differed mainly by the inferred effective population sizes and effective gene flow, which were an order of magnitude larger in the collinear model than in the inversion model (Supplementary Tables S9 and S10). The relative timing of splits and secondary contacts was similar between collinear and inversion SNP analyses ($T_{SecondaryContact}/T_{Split}$ ~ 0.1, Supplementary Table S9), but the secondary contact period was shorter in the overall dataset ($T_{SecondaryContact}/T_{Split}$ ~ 0.01). Only the inferences made from the overall dataset showed support for heterogeneous gene flow (best-supported model = SC_2N_2M). We inferred that 33% of markers experienced a barrier effect where gene flow was reduced by a factor of 10 compared to the remaining background (Supplementary Table S9).

Correlated landscapes of differentiation across *Littorina* species

The WGS data from *L. fabalis* were analyzed using the same reference genome as used in previous studies of *L. saxatilis* (Morales et al., 2019), and this allowed us to search for similarities in the architecture of reproductive isolation. We found a weak but significant correlation in the mean F_{ST} per contig between ecotypes across the two species (r^2 = 0.211, $p < 0.001$). This correlation was mostly driven by two linkage groups (LG6 and LG14) carrying an excess of contigs that differed between the contrasting ecotypes in both species (mean $F_{ST} > 0.1$) (Figure 4a). Removing those two LGs from the correlation leads to a correlation coefficient close to 0 ($r = -0.004$,

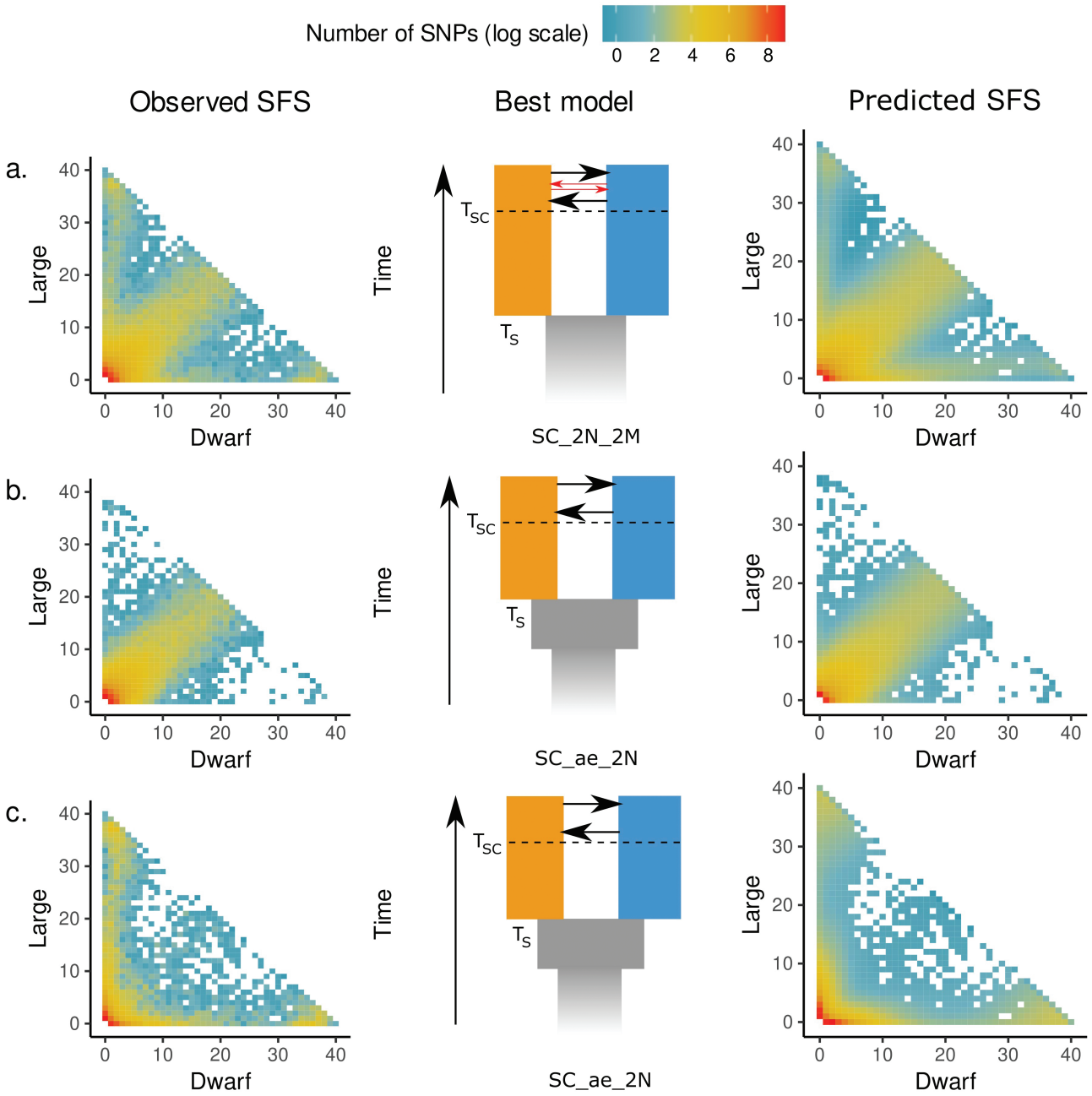

Figure 3. Summary of the demographic inferences performed on the (a) overall dataset, (b) on SNPs located outside the putative inversions, and (c) on SNPs located within the inversions. In each part, the graphs show the observed joint site frequency spectrum (SFS), a schematic representation of the best model inferred based on AIC, and the predicted SFS obtained from the optimization of the best model. All inferences showed support for a secondary contact (SC) scenario, including heterogeneous effective population size along the genome (2N). Only the overall dataset showed support for heterogeneous gene flow (2M), illustrated by the thin red arrows in the model depicted in (a), while the two other inferences also showed support for population expansion in the ancestral population (ae).

$p = 0.001$). The shared differences were clustered at one end of LG6 and LG14 (Figure 4b and c). The genetic signatures in LG14 and LG6 matched those expected from chromosomal inversions (*L. fabalis*: Supplementary Figs. S2–S10, *L. saxatilis* (Faria et al., 2019)), suggesting that two chromosomal rearrangements cover similar genomic regions in differentiating ecotypes in both species. Indeed, these two putative inversions show clines over ecotype contact zones in both species (*L. fabalis*: Figure 2a, *L. saxatilis* (Westram et al., 2018, 2021)). Overall, these results show that the same regions, with similar structural variants, repeatedly contribute to ecotype evolution in distinct, albeit closely related, species.

Discussion

Size is the only major phenotypic difference between the dwarf and the large ecotype of *L. fabalis* (Reimchen, 1981; Tatarenkov & Johannesson, 1998). In addition, they have slightly different associations with wave exposure: the dwarf ecotype is distributed along the more sheltered part of the shore and the large ecotype is in the moderately exposed parts, with shifts from one ecotype to the other sometimes taking place over a few meters only (Tatarenkov & Johannesson, 1999). Here we found that the two ecotypes show strong but heterogeneous genomic differences

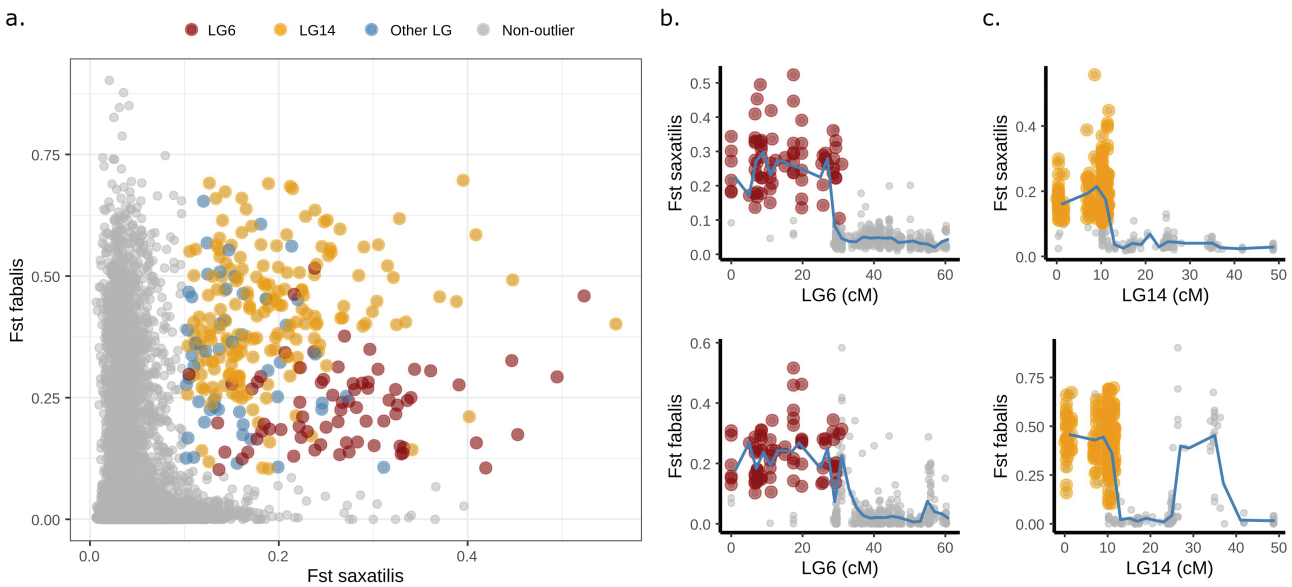

Figure 4. Correlated genomic landscapes of ecotype differentiation in *L. saxatilis* and *L. fabalis*. Panel (a) shows the correlation of the mean F_{ST} value per contig between ecotypes in both species ($r = 0.211$, $p < 0.001$). The colors highlight contigs with $F_{ST} > 0.1$ in both species. Red dots correspond to shared outliers located on LG6, orange on LG14, and blue on other LGs. Panels (c) and (d) show the variation of mean F_{ST} in *L. saxatilis* (top) and *L. fabalis* (bottom) along the two linkage groups carrying a high number of contigs with $F_{ST} > 0.1$ in both species (LG6 and LG14), with dots showing mean F_{ST} per contig, and blue line the average F_{ST} per cM. The shared outlier contigs from LG6 and LG14 highlighted in (a) are highlighted with the same colors in (b) and (c).

concentrated in 12 putative inversions distributed over nine of the 17 *Littorina* chromosomes. The inversions form steep and coupled clines across hybrid zones between the two ecotypes and, at contact, a deficiency of heterokaryotypes for the inversions is associated with bimodality of phenotype and genotype distributions indicating strong barriers to genetic exchange. Our demographic analyses strongly suggest that this hybrid zone is a result of secondary contact following a period of isolation. This isolation may have facilitated both adaptive divergence and the accumulation of genetic incompatibilities that decrease hybrid fitness (Bierne et al., 2011).

The barrier to gene flow is highly heterogeneous across the genome (Figure 3), with strong barriers caused by the arrangements of nine of the 12 inversions that have near-fixed differences between the two ecotypes, but much weaker barriers across the collinear parts of the genome. The strength of the overall barrier depends on the fitness effects of the barrier loci (i.e., the linked effects of the loci inside each inversion) and the distribution of the loci along the genome (Dufresnes et al., 2021; Schaal et al., 2022; Wessinger et al., 2023). We found that strong selection maintains the repeated and highly predictable association between the wave-exposure gradient and the distribution of the two ecotypes, despite the apparently subtle environmental transition. The overall selection involves an extrinsic component: an earlier study using reciprocal transplants and tethering of snails in the seaweed, or on the rock underneath, showed differential survival of the ecotypes with the large ecotype having the highest survival in moderately exposed environments, because it better resisted crab attacks if dislodged from the seaweed during heavy wave action (Kemppainen et al., 2005). Bimodality in the contact zone might also imply assortative mating (Jiggins & Mallet, 2000), although previous field observations suggested close to random mating (Tatarenkov & Johannesson, 1999), arguing against this possibility. Yet, we found evidence that the association between arrangements of different inversions was much stronger than expected from the effect of selection acting independently on each inversion. These results, together with the deficiency in heterokaryotypes, imply coupling between inversions, perhaps due to epistatic effects with extrinsic and/or intrinsic components. It also implies that *L. fabalis* ecotypes are further along the speciation continuum than are the Swedish *L. saxatilis* ecotypes that show unimodal distributions of genotypes and phenotypes in contact zones, and greater variation among clines at contact (Supplementary Table S8).

The structure of the hybrid zones in *L. fabalis* differs from the situation in *L. saxatilis* where the coupling of the different inversions is weak in the centers of the hybrid zones (Johannesson et al., 2020), and almost all inversions remain polymorphic at cline ends (Faria et al., 2019; Westram et al., 2018, 2021). These differences coincide with different evolutionary histories of the ecotypes in the two species. While the unimodal hybrid zones between the crab and wave ecotypes of *L. saxatilis* have evolved by divergent selection with short interruptions to gene flow, at most (Butlin et al., 2014), we found that the *L. fabalis* hybrid zones originated from a secondary contact after a long period of isolation. The inferred period of separation was during the quaternary glacial period (68–150k year ago) while the secondary contact matches the post-glacial period (2–9k year ago). The allopatric divergence of *L. fabalis* ecotypes likely facilitated the coupling of barriers to gene flow, including components of adaptation and genetic incompatibility, which is now involved in the maintenance of the bimodal hybrid zones.

Despite the differences in the hybrid zone structures, and the different demographic backgrounds, the evolution of *L. fabalis* and *L. saxatilis* ecotypes shares some interesting features. For instance, the barriers to gene flow are distributed over many polymorphic inversions, some of which remain highly polymorphic within ecotypes in both species. Furthermore, two inversions, on LG6 and LG14, cover similar genomic regions in the two species. In Swedish populations of *L. saxatilis*, these are the two

most important inversions involved in ecotype differentiation (Morales et al., 2019; Westram et al., 2021) and they host many QTLs associated with phenotypic traits under divergent selection (Koch et al., 2021, 2022). These inversions could have the same origin and be shared by a recent introgression event between *L. saxatilis* and *L. fabalis*, or through repeated selection of an ancestral polymorphism segregating in the standing variation of both species. Alternatively, the repeated and independent origins of new inversions in the same genomic regions in both species, presumably favored by selection because they captured co-adapted alleles, could also have led to this correlated genomic landscape of ecotypic differentiation. Correlated landscapes of differentiation between sister taxa are present in other systems (Burri et al., 2015; Shang et al., 2023Van Doren et al., 2017) and can often be explained by the effect of background selection on shared recombination landscapes (Burri, 2017). The chromosomal inversions detected here represent a low recombining region expected to amplify the hitchhiking effects associated with background selection. However, in the snails, these regions are associated with sharp allelic clines in both species (here and in Westram et al., (2018, 2021)), which suggests that shared architectures of barriers to gene flow can also lead to correlated landscapes of differentiation.

A heterogeneous gene flow between rearranged and collinear genomic regions is expected and commonly found in many systems, such as in migratory vs stationary ecotypes of cod (Matschiner et al., 2022), meadow vs forest ecotypes of deer mice (Harringmeyer & Hoekstra, 2022), and dune vs field ecotypes of sunflower (Huang et al., 2020). In *L. fabalis*, the strong coupling of inversions reinforces the general barrier which extends into the collinear part of the genome. Despite a very small proportion of divergent SNPs being present in the collinear genome, the reduction in gene flow in these collinear parts affects the background differentiation resulting in an F_{ST} = 0.03 which is similar to the level of differentiation between the two ecotypes of the sympatric *L. saxatilis*. Notably, the barrier effect on the collinear genome of *L. fabalis* is present despite the fact that restriction of recombination in inversion heterozygotes is expected to increase recombination in the collinear genome (Stevison et al., 2011). Overall, while our results concur with previous work highlighting the importance of large chromosomal rearrangements for ecotypic divergence and barriers to gene flow, and show that the coupling of several inversions can generate strong reproductive barriers, it remains an open question whether, or not, this type of genomic architecture of reproductive isolation will eventually lead to the completion of speciation.

Materials and methods

Geographic and genomic sampling

Snails were collected over two transects covering a wave exposure gradient located along the southern (128 snails sampled) and northern shores (167 snails) of Lökholmen island on the Swedish west coast (58.89°N, 11.11°E). The level of wave exposure was indicated by the presence or absence of *Ascophyllum nodosum* with this species being intolerant to exposure levels ranging from moderate to strong wave action. The position and elevation of each snail along the transects were recorded with a Trimble total station. The three-dimensional position of each snail was transformed into one-dimensional spatial coordinates using least-cost distance following Westram et al. (2021). We took pictures of each snail with the aperture up alongside a scale, and the size of each snail was measured in ImageJ as the longest distance from the aperture rim to the apex. Snails were thereafter dissected and a piece of foot tissue was stored in 95% ethanol. We extracted DNA from tissue samples using the protocol by Panova et al., (2016). After measuring concentrations and purity of the extractions on Qbit and nanodrop, DNA was shipped to SciLifeLab (Sweden) for individual Nextera WGS library preparations, followed by sequencing on a NovaSeq S6000. Targeted coverage was 5X except for 12 individuals taken from the ends of the transects which were sequenced at ~15X coverage.

Bioinformatics pipeline

The WGS short reads were mapped to the reference genome of the related species *L. saxatilis* using *bwa-mem* (Li & Durbin, 2009). This reference genome is fragmented but is complemented by a dense linkage map (Westram et al., 2018). Mapped reads were processed through the GATK pipeline (McKenna et al., 2010) following good practice guidelines including the removal of PCR duplicates and base-pair recalibration. Further filtration steps were applied using vcftools (Danecek et al., 2011) in order to keep only biallelic SNPs sequenced in at least 90% of the individuals and with an average minimum sequencing depth of 3 and maximum of 15 across samples. This raw dataset contained 1.38M SNPs, which were further filtered depending on the analyses.

Population structure analyses

For these analyses, the raw dataset was thinned by keeping only SNPs with a minor allele frequency above 5%, leading to 451k SNPs. This dataset was further pruned for physical linkage by keeping 1 SNP per 1k base-pair chosen at random (based on LD decay in *Littorina* (Westram et al., 2018)), resulting in a total of 110k SNPs. We computed a PCA on all individuals (both transects included) from the 110k SNPs pruned for linkage using the R package adegenet (Jombart, 2008). Pairwise F_{ST} values between ecotypes were calculated per pruned-SNP using the 20 snails nearest the ends of both transects, analyzed independently for each transect using vcftools (Danecek et al., 2011), and visualized using Manhattan plots in ggplot. Then, we used a sliding window approach over 1cM bins using the 458k SNPs unpruned for linkage to compute local PCA with adegenet (Jombart, 2008) and variation of heterozygosity with the *Hsplit* function (Reeve et al., 2023) to explore the variation in structure and diversity patterns along the *L. saxatilis* linkage map (Westram et al., 2018). Pairwise LD between SNPs was then calculated per linkage group using only markers with minor allele frequency above 25% using the r^2 method in LDheatmap (Shin et al., 2006). Based on variation patterns of LD, *Hsplit*, F_{ST}, and local PCA along the chromosomes, we visually defined 14 linkage blocks carrying high differentiation between ecotypes, which were likely due to chromosomal inversions. Then, we used unlinked SNPs from outside the linkage blocks to compute the relatedness among all pairs of snails using the a_{jk} statistic from (Yang et al., 2010) in vcftools (Danecek et al., 2011). These analyses allowed us to identify multiple pairs of full-sibs and half-sibs, which we used to calculate the standard deviation of the dispersal distance between parent and offspring (σ) from the mean transect distance between sibling individuals using the following formula: $\sigma = \frac{\bar{d}}{2/\sqrt{\pi}}$. More details about this equation and its assumptions are available as Supporting text p.3. We computed DAPC using SNPs from within each of the linkage blocks for a k value of three clusters using adegenet (Jombart et al., 2010) in order to infer the karyotypes of the putative chromosomal rearrangements for each snail. These karyotypes were

finally used to calculate the hybrid index and the observed heterozygosity per snail.

Genetic cline analyses

Chromosomal arrangement frequency cline analyses were performed within a Bayesian framework using the R package Rstan (Stan Development Team, 2023) using the karyotypes inferred from the local DAPC. The cline functions tested included the estimation of F_{IS}, assumed to peak at the cline center and estimated at 5 points, following the formulation in Cfit (Gay et al., 2008) but also rewritten as Rstan functions. In addition, the coefficients of selection required to maintain the cline were calculated from the width of the cline, and the estimate of dispersal, σ, inferred from the distance between pairs of related individuals, using the formula from Barton & Gale, (1993) assuming an abrupt habitat change and no dominance effect: $width = 1.732x^{\sigma}/_{\sqrt{s}}$, where s is the reduction in mean fitness at the cline center. Furthermore, we explored the allelic clines at each SNP from the linked dataset (i.e., not filtered to one SNP per kb) with an MAF above 0.3 and a difference in allele frequency between the extreme parts of a transect above 0.1. In total, we inferred simple sigmoid clines, without assuming fixation at the ends, from 90k SNPs using maximum likelihood in bbmle (which is less computationally intensive than Rstan). Then, following the procedure from Westram et al., (2018, 2021), we compared the AIC of models with fixed frequency and linear frequency variation to clinal variation along the transect. Clinal fits showing an AIC difference > 4, compared with a non-clinal fit, were considered as significant clines. The goodness of each cline fit was then estimated using a generalized linear regression model with binomial error to evaluate how well frequencies estimated from the cline explained the genotype observed for each snail given its position on the transect.

Phenotypic cline analyses

The phenotypic cline analyses were also performed in Rstan (Stan Development Team, 2023). We fitted cline models for continuous traits to the variation observed along the two transects in shell length, the PC1 score computed for the overall dataset, the PC1 score for the dataset from SNPs outside chromosomal rearrangements, and the hybrid index computed from the chromosomal rearrangement genotypes. We compared the ability of five contrasted cline models to fit the data, following the formulation in the Cfit package (Gay et al., 2008) but re-coded in custom Rstan functions: a unimodal cline, two bimodal clines, one with and one without introgression, and two trimodal clines, one with and one without introgression. Unimodal clines are generally used for quantitative traits when the phenotype observed in the center of the hybrid zone follows a normal distribution centered around the mean of the two parental phenotypes while multimodal clines are expected when the parental phenotypes remain more or less discrete even in the center of the hybrid zone. Bimodal clines are expected between taxa overlapping in the center of a hybrid zone that maintain phenotypic differences in the center of the cline similar to those between the extremes. Such hybrid zones are expected with strong reproductive isolation or for traits with a simple genetic basis or threshold behavior. Furthermore, the phenotype can also be inferred as trimodal if hybrids with intermediate phenotypes occur in the center of the cline but have low fitness, avoiding production of a hybrid swarm. The differences between the parental modes near to the center of the cline can be reduced in the case of introgression, which is the reason why the introgression was also modeled in the multimodal clines.

Coupling analyses

Coupling can be considered as a measure of the extent to which selection acts independently on each barrier locus (weak coupling) or depends on the interaction among many barrier loci (strong coupling). We measured the coupling using two different approaches. The first approach is based on LD between barrier loci in the center of the hybrid zone following Barton & Gale, (1993). More specifically, we computed the expected disequilibrium (D_{exp}) when selection acts independently across loci, which depends on the per generation dispersal (estimated in Supporting text p.3), the width of the cline (here based on the hybrid index, Supplementary Table S5), and recombination between barrier loci (here the different chromosomal rearrangements were assumed to recombine freely, details in Supporting text of the p.4). Then, we compared the D_{exp} value to the observed disequilibrium (D_{obs}) in the center of the hybrid zones calculated following Barton & Gale, (1993) (see details in Supporting text of p.4), a $D_{obs} >> D_{exp}$ suggesting a strong coupling. The second approach to measure coupling was based on the coefficient of variation in cline slope (CV) across inversions following the method by (Firneno et al., 2023). CV can be calculated by dividing the standard deviation of the cline slopes by the mean of the cline slopes (here, $slope = \frac{Plarge-Pdwarf}{width}$) and is negatively associated to the coupling coefficient (Firneno et al., 2023). Here, a weak coupling coefficient typically generates CV above one while strong coupling generates a CV close to zero.

Demographic analyses

We used the software *moments* (Jouganous et al., 2017) to infer the demographic history of divergence between the large and the dwarf ecotype of *L. fabalis*. This software uses the information contained in the joint site frequency spectrum to fit contrasted demographic scenarios using an approach based on a diffusion approximation. We calculated the SFS using 20 snails sampled from each end of each transect. To avoid bias in the SFS due to the miss-calling of the heterozygotes with low coverage depth in our data, one allele per polymorphic position was randomly sampled independently in each individual. Individuals from the ends of both transects were then pooled to produce a SFS of 40 chromosomes from the large ecotype (i.e., one chromosome per individual) against 40 chromosomes from the dwarf ecotype and to explore the divergence between ecotypes ("overall SFS"). This SFS was further subsampled to include only the SNPs from within the putative chromosomal inversions ("inversion SFS") or from outside the inversions ("collinear SFS"). Demographic inferences were then made independently for the three SFS. In total, we compared the ability of 28 demographic scenarios to reproduce the observed SFS (Le Moan et al., 2021; Momigliano et al., 2021), including scenarios of allopatric divergence (4 variants of a strict isolation (SI) model), and divergence with gene flow (8 variants of an isolation-with-migration (IM) model, 8 variants of an ancestral migration (AM) model, and 8 variants of a secondary contact (SC) model). All scenarios are detailed in Supplementary Fig. S20, and they were adjusted to the observed data using the optimization procedure described in Portik et al. (2017). The different demographic scenarios were compared using AIC, with models differing in AIC by < 10 considered as equally good. Demographic parameters were transformed using the method described in (Rougeux et al., 2019), using a mutation rate of 1×10^{-8} per bp and generation and a generation time of 1 year (Reid, 1996).

Interspecific comparative analysis of differentiation

As the same reference genome was used between this work and previous work performed on ecotype evolution in *L. saxatilis*,

we could compare the genomic landscape of ecotype differentiation between the two species. We used a simple correlation test on mean F_{ST} per contig between ecotypes in each species using SNPs derived from the whole genome. Those F_{ST} values were calculated independently in both species, using the lcWGS data described in the present study for *L. fabalis*, and previously published F_{ST} values from Morales et al., (2019) computed from pool-seq WGS data for *L. saxatilis*. For *L. saxatilis*, we used only the F_{ST} between ecotypes from two sites (ANG and CZB), which were located on two islands of the same archipelago as the *L. fabalis* hybrid zone studied here. Shared outlier regions were then considered as any contig with mean F_{ST} between ecotypes above 0.1 (arbitrary threshold).

Supplementary material

Supplementary material is available online at *Evolution Letters*.

Data and code availability

The raw fasta files are available on NCBI SRA under the project name "*Littorina fabalis* transect WGS" with accession number PRJNA836378. The filtered vcf file, metadata, as well as the R scripts used to analyse these datasets, are available as a zenodo archive here: doi: 10.5281/zenodo.10790257.

Author contributions

Conceptualization: A.L.M., R.B., and K.J.; Laboratory work: O.O.M.; Methodology: A.L.M. and R.B.; Investigation: A.L.M., R.B., and K.J.; Visualization: A.L.M.; Supervision: S.S., R.F., K.J., and R.B.; Writing—original draft: A.L.M. and K.J.; and Writing—review & editing: M.R., R.F., S.S., K.J., and R.B.

Funding

K.J., M.R., and R.K.B. were supported by grants from the Swedish Research Council (grant numbers: 2021-0419, 2021-05243, and 2018-03695, respectively). R.K.B. was also supported by the Leverhulme Trust (RPG-2021-141), R.F. by the Portuguese Science Foundation (PTDC/BIA-EVL/1614/2021 and 2020.00275.CEECIND), and ALM by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No BIENVENÜE 899546.

Conflict of interest: The authors declare no conflict of interest.

Acknowledgments

The computations and data handling were enabled by resources provided by the Swedish National Infrastructure for Computing at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973. We thank all the member of the *Littorina* team for the stimulating discussions about the manuscripts, James Reeves for his help the implementation of *Hsplit*, and Thomas Broquet for his useful comments in the latter stage of manuscript revisions.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Barton, N. H. (1979). The dynamics of hybrid zones. *Heredity*, *43*(3), 341–359. https://doi.org/10.1038/hdy.1979.87

Barton, N. H., & Gale, K. S. (1993) Genetic analysis of hybrid zones. In R. G. Harrison (Ed.), *Hybrid zones and the evolutionary process*. Oxford University Press.

Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*, *341*(6242), 497–503. https://doi.org/10.1038/341497a0

Bierne, N., Welch, J., Loire, E., Bonhomme, F., & David, P. (2011). The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, *20*(10), 2044–2072. https://doi.org/10.1111/j.1365-294X.2011.05080.x

Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, *1*(3), 118–131. https://doi.org/10.1002/evl3.14

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S. -A., Sætre, G. -P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665. https://doi.org/10.1101/gr.196485.115

Butlin, R. (1987). Speciation by reinforcement. *Trends in Ecology & Evolution*,*2*(1),8–13.https://doi.org/10.1016/0169-5347(87)90193-5

Butlin, R. K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J. A., Galindo, J., Grahame, J. W., Hollander, J., Kemppainen, P., Martínez-Fernández, M., Panova, M., Quesada, H., Johannesson, K., & Rolán-Alvarez, E. (2014). Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution*, *68*(4), 935–949. https://doi.org/10.1111/evo.12329

Butlin, R. K., & Smadja, C. M. (2018). Coupling, reinforcement, and speciation. *The American Naturalist*, *191*(2), 155–172. https://doi.org/10.1086/695136

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R.; 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Jode, A., Le Moan, A., Johannesson, K., Faria, R., Stankowski, S., Westram, A. M., Butlin, R. K., Rafajlović, M., & Fraïsse, C. (2023). Ten years of demographic modelling of divergence and speciation in the sea. *Evolutionary Applications*, *16*(2), 542–559. https://doi.org/10.1111/eva.13428

Dufresnes, C., Brelsford, A., Jeffries, D. L., Mazepa, G., Suchan, T., Canestrelli, D., Nicieza, A., Fumagalli, L., Dubey, S., Martínez-Solano, I., Litvinchuk, S. N., Vences, M., Perrin, N., & Crochet, P. -A. (2021). Mass of genes rather than master genes underlie the genomic architecture of amphibian speciation. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(36), e2103963118. https://doi.org/10.1073/pnas.2103963118

Faria, R., Chaube, P., Morales, H. E., Larsson, T., Lemmon, A. R., Lemmon, E. M., Rafajlović, M., Panova, M., Ravinet, M., Johannesson, K., Westram, A. M., & Butlin, R. K. (2019). Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes. *Molecular Ecology*, *28*(6), 1375–1393. https://doi.org/10.1111/mec.14972

Firneno, T. J., Semenov, G., Dopman, E. B., Taylor, S. A., Larson, E. L., & Gompert, Z. (2023). Quantitative analyses of coupling in hybrid zones. *Cold Spring Harbor Perspectives in Biology*, *15*(12), a041434. https://doi.org/10.1101/cshperspect.a041434

Galindo, J., Carvalho, J., Sotelo, G., Duvetorp, M., Costa, D., Kemppainen, P., Panova, M., Kaliontzopoulou, A., Johannesson, K., & Faria, R. (2021). Genetic and morphological divergence between *Littorina fabalis* ecotypes in Northern Europe. *Journal of Evolutionary Biology*, *34*(1), 97–113. https://doi.org/10.1111/jeb.13705

Gay, L., Crochet, P. -A., Bell, D. A., & Lenormand, T. (2008). Comparing clines on molecular and phenotypic traits in hybrid zones: A window on tension zone models. *Evolution*, *62*(11), 2789–2806. https://doi.org/10.1111/j.1558-5646.2008.00491.x

Harringmeyer, O. S., & Hoekstra, H. E. (2022). Chromosomal inversion polymorphisms shape the genomic landscape of deer mice. *Nature Ecology & Evolution*, *6*(12), 1965–1979. https://doi.org/10.1038/s41559-022-01890-0

Harrison, R. G., & Bogdanowicz, S. M. (1997). Patterns of variation and linkage disequilibrium in a field cricket hybrid zone. *Evolution*, *51*(2), 493–505. https://doi.org/10.1111/j.1558-5646.1997.tb02437.x

Huang, K., Andrew, R. L., Owens, G. L., Ostevik, K. L., & Rieseberg, L. H. (2020). Multiple chromosomal inversions contribute to adaptive divergence of a dune sunflower ecotype. *Molecular Ecology*, *29*(14), 2535–2549. https://doi.org/10.1111/mec.15428

Janson, K., & Sundberg, P. (1983). Multivariate morphometric analysis of two varieties of *Littorina saxatilis* from the Swedish west coast. *Marine Biology*, *74*(1), 49–53. https://doi.org/10.1007/bf00394274

Jiggins, C. D., & Mallet, J. (2000). Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, *15*(6), 250–255. https://doi.org/10.1016/s0169-5347(00)01873-5

Johannesson, K., Faria, R., Moan, A. L., Rafajlović, M., Westram, A. M., Butlin, R. K., & Stankowski, S. (2024). Diverse pathways to speciation revealed by marine snails. *Trends in Genetics*, *40*(4), 337–351. https://doi.org/10.1016/j.tig.2024.01.002

Johannesson, K., Panova, M., Kemppainen, P., André, C., Rolán-Alvarez, E., & Butlin, R. K. (2010). Repeated evolution of reproductive isolation in a marine snail: Unveiling mechanisms of speciation. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *365*(1547), 1735–1747. https://doi.org/10.1098/rstb.2009.0256

Johannesson, K., Zagrodzka, Z., Faria, R., Marie Westram, A., & Butlin, R. K. (2020). Is embryo abortion a post-zygotic barrier to gene flow between *Littorina* ecotypes? *Journal of Evolutionary Biology*, *33*(3), 342–351. https://doi.org/10.1111/jeb.13570

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94. https://doi.org/10.1186/1471-2156-11-94

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. (2017). Inferring the joint demographic history of multiple populations: Beyond the diffusion approximation. *Genetics*, *206*(3), 1549–1567. https://doi.org/10.1534/genetics.117.200493

Kemppainen, P., Nes, S., van Ceder, C., & Johannesson, K. (2005). Refuge function of marine algae complicates selection in an intertidal snail. *Oecologia*, *143*(3), 402–411. https://doi.org/10.1007/s00442-004-1819-5

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biology*, *8*(9), e1000501. https://doi.org/10.1371/journal.pbio.1000501

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, *173*(1), 419–434. https://doi.org/10.1534/genetics.105.047985

Koch, E. L., Morales, H. E., Larsson, J., Westram, A. M., Faria, R., Lemmon, A. R., Lemmon, E. M., Johannesson, K., & Butlin, R. K. (2021). Genetic variation for adaptive traits is associated with polymorphic inversions in *Littorina saxatilis*. *Evolution Letters*, *5*(3), 196–213. https://doi.org/10.1002/evl3.227

Koch, E. L., Ravinet, M., Westram, A. M., Johannesson, K., & Butlin, R. K. (2022). Genetic architecture of repeated phenotypic divergence in *Littorina saxatilis* ecotype evolution. *Evolution*, *76*(10), 2332–2346. https://doi.org/10.1111/evo.14602

Le Moan, A., Bekkevold, D., & Hemmer-Hansen, J. (2021). Evolution at two time frames: Ancient structural variants involved in post-glacial divergence of the European plaice (*Pleuronectes platessa*). *Heredity*, *126*(4), 668–683. https://doi.org/10.1038/s41437-020-00389-3

Le Moan, A., Panova, M., De Jode, A., Ortega-Martinez, O., Duvetorp, M., Faria, R., Butlin, R., & Johannesson, K. (2023). An allozyme polymorphism is associated with a large chromosomal inversion in the marine snail *Littorina fabalis*. *Evolutionary Applications*, *16*(2), 279–292. https://doi.org/10.1111/eva.13427

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Matschiner, M., Barth, J. M. I., Tørresen, O. K., Star, B., Baalsrud, H. T., Brieuc, M. S. O., Pampoulie, C., Bradbury, I., Jakobsen, K. S., & Jentoft, S. (2022). Supergene origin and maintenance in Atlantic cod. *Nature Ecology & Evolution*, *6*(4), 469–481. https://doi.org/10.1038/s41559-022-01661-x

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Mérot, C., Berdan, E. L., Cayuela, H., Djambazian, H., Ferchaud, A. -L., Laporte, M., Normandeau, E., Ragoussis, J., Wellenreuther, M., & Bernatchez, L. (2021). Locally adaptive inversions modulate genetic variation at different geographic scales in a seaweed fly. *Molecular Biology and Evolution*, *38*(9), 3953–3971. https://doi.org/10.1093/molbev/msab143

Mérot, C., Oomen, R. A., Tigano, A., & Wellenreuther, M. (2020). A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends in Ecology & Evolution*, *35*(7), 561–572. https://doi.org/10.1016/j.tree.2020.03.002

Meyer, L., Barry, P., Riquet, F., Foote, A., Der Sarkissian, C., Cunha, R. L., Arbiol, C., Cerqueira, F., Desmarais, E., Bordes, A., Bierne, N., Guinand, B., & Gagnaire, P. A. (2024). Divergence and gene flow history at two large chromosomal inversions underlying ecotype differentiation in the long-snouted seahorse . *Molecular Ecology*, e17277. https://doi.org/10.1111/mec.17277

Momigliano, P., Florin, A. -B., & Merilä, J. (2021). Biases in demographic modeling affect our understanding of recent divergence. *Molecular Biology and Evolution*, *38*(7), 2967–2985. https://doi.org/10.1093/molbev/msab047

Morales, H. E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A. M., & Butlin, R. K. (2019). Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast. *Science Advances*, *5*(12), eaav9963. https://doi.org/10.1126/sciadv.aav9963

Navarro, A., & Barton, N. H. (2003). Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation.

Evolution, 57(3), 447–459. https://doi.org/10.1111/j.0014-3820.2003.tb01537.x

Nosil, P., & Flaxman, S. M. (2011). Conditions for mutation-order speciation. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *278*(1704), 399–407. https://doi.org/10.1098/rspb.2010.1215

Nosil, P., Soria-Carrasco, V., Villoutreix, R., De-la-Mora, M., de Carvalho, C. F., Parchman, T., Feder, J. L., & Gompert, Z. (2023). Complex evolutionary processes maintain an ancient chromosomal inversion. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(25), e2300673120. https://doi.org/10.1073/pnas.2300673120

Panova, M., Aronsson, H., Cameron, R. A., Dahl, P., Godhe, A., Lind, U., Ortega-Martinez, O., Pereyra, R., Tesson, S. V. M., Wrange, A. -L., Blomberg, A., & Johannesson, K. (2016). DNA extraction protocols for whole-genome sequencing in marine organisms. *Methods in Molecular Biology*, *1452*, 13–44. https://doi.org/10.1007/978-1-4939-3774-5_2

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., Rödel, M. -O., Blackburn, D. C., & Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*(19), 5245–5263. https://doi.org/10.1111/mec.14266

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450–1477. https://doi.org/10.1111/jeb.13047

Reeve, J., Butlin, R. K., Koch, E. L., Stankowski, S., & Faria, R. (2023). Chromosomal inversion polymorphisms are widespread across the species ranges of rough periwinkles (*Littorina saxatilis* and *L. arcana*). *Molecular Ecology*, https://doi.org/10.1111/mec.17160

Reid, D. G. (1996). *Systematics and evolution of Littorina*. The Ray Society.

Reimchen, T. (1981). Microgeographical variation in *Littorina mariae* Sacchi & Rastelli and a taxonomic consideration. *Journal of Conchology*, 30, 341–350.

Rieseberg, L. H. (2001). Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, *16*(7), 351–358. https://doi.org/10.1016/s0169-5347(01)02187-5

Rougeux, C., Gagnaire, P. -A., & Bernatchez, L. (2019). Model-based demographic inference of introgression history in European whitefish species pairs'. *Journal of Evolutionary Biology*, *32*(8), 806–817. https://doi.org/10.1111/jeb.13482

Satokangas, I., Martin, S. H., Helanterä, H., Saramäki, J., & Kulmuni, J. (2020). Multi-locus interactions and the build-up of reproductive isolation. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 375(1806), 20190543. https://doi.org/10.1098/rstb.2019.0543

Schaal, S. M., Haller, B. C., & Lotterhos, K. E. (2022). Inversion invasions: When the genetic basis of local adaptation is concentrated within inversions in the face of gene flow. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1856), 20210200. https://doi.org/10.1098/rstb.2021.0200

Shang, H., Field, D. L., Paun, O., Rendón-Anaya, M., Hess, J., Vogl, C., >Liu, J., Ingvarsson, P. K., Lexer, C., & Leroy, T. (2023). Drivers of genomic landscapes of differentiation across a Populus divergence gradient. *Molecular Ecology*, 32(15), 4348–4361. https://doi.org/10.1111/mec.17034

Shin, J. -H., Blay, S., McNeney, B., & Graham, J. (2006). LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *Journal of Statistical Software*, *16*, 1–9. https://doi.org/10.18637/jss.v016.c03

Stan Development Team. (2023). *RStan: The R interface to Stan*. https://mc-stan.org/

Stevison, L. S., Hoehn, K. B., & Noor, M. A. F. (2011). Effects of inversions on within- and between-species recombination and divergence. *Genome Biology and Evolution*, 3, 830–841. https://doi.org/10.1093/gbe/evr081

Tatarenkov, A., & Johannesson, K. (1994). Habitat related allozyme variation on a microgeographic scale in the marine snail *Littorina mariae* (Prosobranchia: Littorinacea). *Biological Journal of the Linnean Society*, 53(2), 105–125. https://doi.org/10.1111/j.1095-8312.1994.tb01004.x

Tatarenkov, A., & Johannesson, K. (1998). Evidence of a reproductive barrier between two forms of the marine periwinkle *Littorina fabalis* (Gastropoda). *Biological Journal of the Linnean Society*, 63(3), 349–365. https://doi.org/10.1111/j.1095-8312.1998.tb01522.x

Tatarenkov, A., & Johannesson, K. (1999). Micro- and macrogeographic allozyme variation in *Littorina fabalis*; do sheltered and exposed forms hybridize? *Biological Journal of the Linnean Society*, *67*(2), 199–212. https://doi.org/10.1111/j.1095-8312.1999.tb01861.x

Van Doren, B. M., Campagna, L., Helm, B., Illera, J. C., Lovette, I. J., & Liedvogel, M. (2017). Correlated patterns of genetic diversity and differentiation across an avian family. *Molecular Ecology*, *26*(15), 3982–3997. https://doi.org/10.1111/mec.14083

Wessinger, C. A., Katzer, A. M., Hime, P. M., Rausher, M. D., Kelly, J. K., & Hileman, L. C. (2023). A few essential genetic loci distinguish Penstemon species with flowers adapted to pollination by bees or hummingbirds. *PLoS Biology*, *21*(9), e3002294. https://doi.org/10.1371/journal.pbio.3002294

Westram, A. M., Faria, R., Johannesson, K., & Butlin, R. (2021). Using replicate hybrid zones to understand the genomic basis of adaptive divergence. *Molecular Ecology*, *30*(15), 3797–3814. https://doi.org/10.1111/mec.15861

Westram, A. M., Rafajlović, M., Chaube, P., Faria, R., Larsson, T., Panova, M., Ravinet, M., Blomberg, A., Mehlig, B., Johannesson, K., & Butlin, R. (2018). Clines on the seashore: The genomic architecture underlying rapid divergence in the face of gene flow. *Evolution Letters*, *2*(4), 297–309. https://doi.org/10.1002/evl3.74

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of heritability for human height. *Nature Genetics*, *42*(7), 565–569. https://doi.org/10.1038/ng.608