Chromosomal inversion polymorphisms are widespread across the species ranges of rough periwinkles (*Littorina saxatilis* and *L. arcana*)

Running title: Widespread inversions in *Littorina*

James Reeve[1*], Roger K. Butlin[1,2], Eva L. Koch[3], Sean Stankowski[4], Rui Faria[2,5]

1. University of Gothenburg, Tjärnö Marine Laboratory, Laboratorievägen 10, 452 96 Strömstad, Sweden

2. University of Sheffield, School of Bioscience, Western Bank, Sheffield S10 2TN, United Kingdom

3. University of Cambridge, Department of Zoology, Downing Street, Cambridge CB2 3EJ, United Kingdom

4. Institute of Science and Technology – Austria, Am Campus 1, 3400 Klosterneuburg, Austria

5. University of Porto, Centro de Investigação em Biodiversidade e Recursos Genéticos, Rua Padre Armando Quintas n°7, 4485-661 Vairão, Portugal

*Corresponding author: james.reeve@gu.se

Abstract:

Inversions are thought to play a key role in adaptation and speciation, suppressing recombination between diverging populations. Genes influencing adaptive traits cluster in inversions, and changes in inversion frequencies associate with environmental differences. However, in many organisms it is unclear if inversions are geographically and taxonomically widespread. The intertidal snail, *Littorina saxatilis*, is one such example. Strong associations between putative polymorphic inversions and phenotypic differences have been demonstrated between two ecotypes of *L. saxatilis* in Sweden and inferred elsewhere, but no direct evidence for inversion polymorphism currently exists across the species range. Using whole genome data from 107 snails, most inversion polymorphisms were found to be widespread across the species range. Frequencies of some inversion arrangements were significantly different among ecotypes, suggesting a parallel adaptive role. Many inversions were also polymorphic in the sister species *L. arcana*, hinting at an ancient origin.

Introduction:

Inversions suppress recombination allowing combinations of alleles to be maintained despite gene flow, which potentially plays a key role in local adaptation and speciation (Butlin, 2005; Faria, Johannesson, et al., 2019; Faria & Navarro, 2010; Hoffmann & Rieseberg, 2008; Jackson, 2011; Kirkpatrick & Barton, 2006; Wellenreuther & Bernatchez, 2018). Although inversions have been identified in numerous systems spanning the speciation continuum (Wellenreuther & Bernatchez, 2018), they are often polymorphic in one or both diverging populations, suggesting that they are under balancing selection (Durmaz et al., 2020; Faria, Johannesson, et al., 2019; Wellenreuther & Bernatchez, 2018). To understand the interplay between balancing and divergent selection within inversions, we need to track how they evolve, determining their origin and spread, and the balance of evolutionary forces affecting them over time.

Empirical evidence for the role of inversions in divergence is often limited to small geographical areas, which raises the question: are the same inverted regions driving local adaptation across a species range? Some studies have explored the adaptive role of inversions across global distributions. In some cases, inversion frequencies change across broad biogeographic clines, such as with latitude in *Drosophila melanogaster* (Kapun & Flatt, 2019) or precipitation in several malaria-harbouring mosquito species (Ayala et al., 2014, 2017). In other (although non-mutually exclusive) cases, inversions are involved in local adaptation leading to parallel phenotypic evolution across sites with similar environmental contrasts (Westram et al., 2022). Examples include three inversions in the threespine stickleback that differentiate freshwater and marine populations (Jones et al., 2012), four inversions linked with migratory behaviour in Atlantic cod populations (Matschiner et al., 2022), and 13 inversions associated with changes between forest and prairie habitats in deer mice (Harringmeyer & Hoekstra, 2022). These parallel patterns strongly support inversions'

role in local adaptation and emphasise how characterizing the global distribution of inversions helps to understand the genetic basis of adaptation.

The rough periwinkle (*Littorina saxatilis* (Olivi, 1792)) is a useful study system for understanding the role of inversions in adaptation and speciation. *Littorina saxatilis* is a phenotypically diverse intertidal snail that primarily inhabits rocky seashores across the North Atlantic (Reid, 1996, pp. 324–331). Recently, 18 clusters of loci in linkage disequilibrium have been found within the species, which are indications of polymorphic chromosomal inversions (Faria, Chaube, et al., 2019; Westram et al., 2021). Some of these putative inversions contain loci influencing adaptive traits that differentiate two ecotypes (Koch et al., 2021, 2022): a *crab* ecotype resistant to predation by shore crabs (Boulding et al., 2017; Janson, 1982; Johannesson, 1986) and a *wave* ecotype resistant to dislodgment by waves (Larsson, 2021; Le Pennec et al., 2017). These polymorphic inversions and their associations with ecotypes are repeated across multiple nearby sites in Sweden (Westram et al., 2021). Strong genetic differentiation between ecotypes occurs at genomic regions corresponding to some Swedish inversions suggests that similar associations exist in the United Kingdom, France and Spain (Kess & Boulding, 2019; Morales et al., 2019). However, while a signal of *crab*-*wave* divergence has been inferred across Europe, there is currently no direct evidence that the inversions detected in Sweden are polymorphic across the species range.

The species range of *L. saxatilis* covers many habitats and overlaps substantially with two closely-related species (*L. arcana* and *L. compressa*: all three species ranges overlap from Brittany to the Barents Sea). Morphological studies over the last two centuries have proposed numerous species names and taxonomic subgroupings for *L. saxatilis* (Reid, 1996, pp. 278–292). Reid (1996, pp. 305–318) summarised this variation as four ecotypes: *crab* (*moderate* sensu Reid), *wave*, *brackish* and *barnacle*, that occur within ovoviviparous *L.*

saxatilis and also its egg-laying relatives, *L. arcana* and *L. compressa* (pp. 248-278). *L. saxatilis* and *L. arcana* are considered sister species, with near complete reproductive isolation (Stankowski et al., 2020). There is also increasing evidence of a strong phylogeographic break in *L. saxatilis* around the Bay of Biscay, separating populations in the Iberian Peninsula from those in the North (Doellman, Trussell, Grahame, & Vollmer, 2011; Morales et al., 2019; Tirado, Saura, Rolán-Alvarez, & Quesada, 2016, & Panova et al. 2011). By identifying inversions in other ecotypes, species and geographic regions we can better contextualise their adaptive role in *L. saxatilis* and their taxonomic spread.

Our current knowledge of the *crab* and *wave* ecotypes can help us to infer the adaptive role of inversions in other ecotypes facing similar environmental pressures. The *barnacle* ecotype is found on very exposed shores, thus any inversion arrangement common in *wave* snails is also expected to be common in *barnacle* snails. Similarly, *L. arcana* typically inhabits moderately exposed shores (Reid, 1996, p. 274), therefore, if they share inversions with *L. saxatilis* it is more likely they carry *wave* arrangements. However, this prediction assumes that adaptation is only directed by wave exposure and predation, an assumption that may not always hold (Morales et al., 2019).

To clarify the adaptive role of *L. saxatilis* inversions, we investigated the distribution of each inversion polymorphism across the entire species range. Specifically, we aimed to: i) find whether inversions that were identified in Sweden are polymorphic across the species range, ii) determine if the inversions previously associated with *crab-wave* divergence also differentiate ecotypes consistently throughout the species range, and iii) investigate if these inversion polymorphisms are shared with a sister species (*L. arcana*), which would suggest an ancient origin. Studying the geographic and taxonomic distribution of inversions should provide a broader perspective on their roles in ecotype formation and speciation.

Methods:

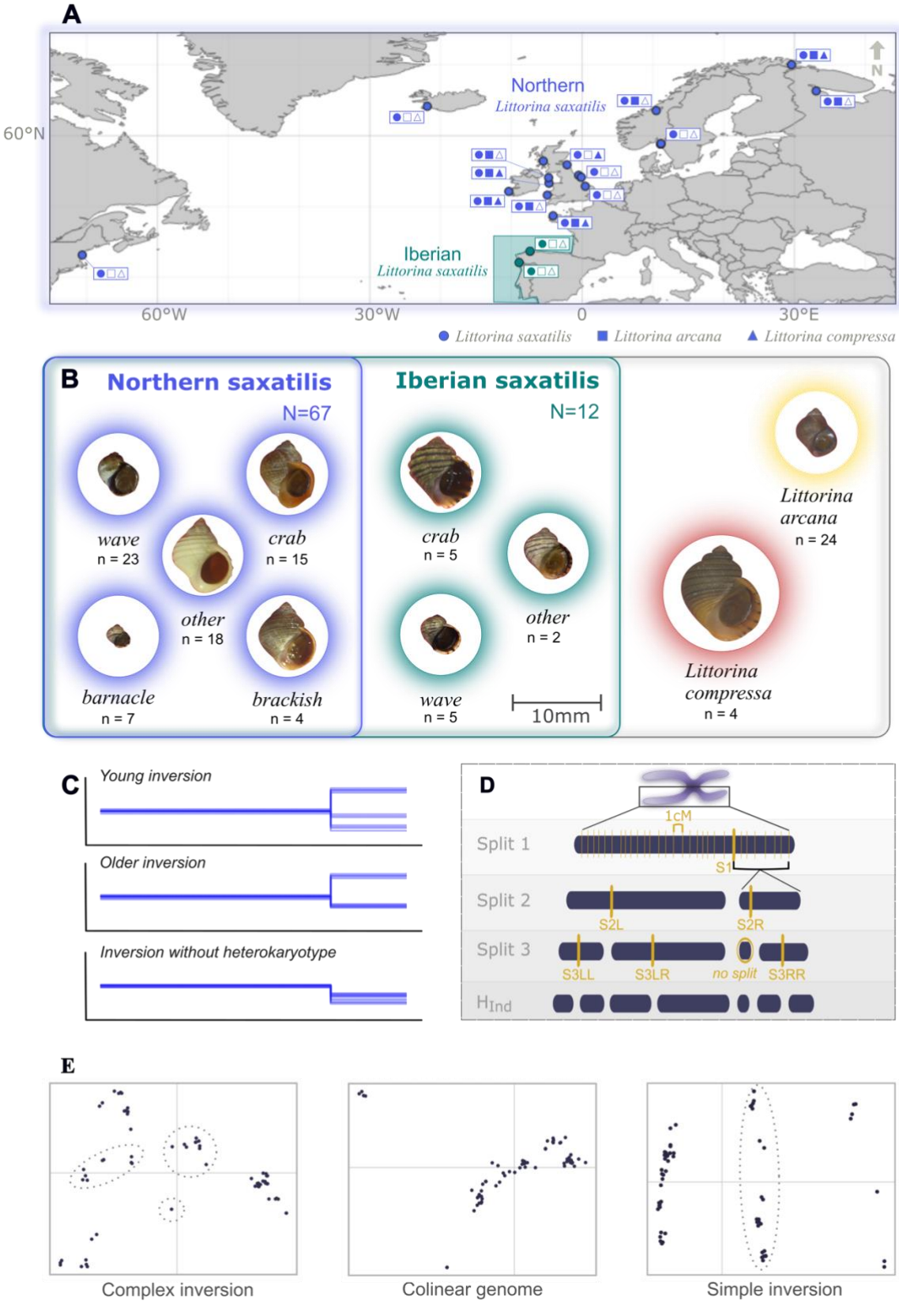
A
Northern
Littorina saxatilis
Iberian
Littorina saxatilis
60°N
60°W
30°W
0
30°E
N
Littorina saxatilis
Littorina arcana
Littorina compressa
B
Northern saxatilis
N=67
wave
n = 23
crab
n = 15
other
n = 18
barnacle
n = 7
brackish
n = 4
Iberian saxatilis
N=12
crab
n = 5
other
n = 2
wave
n = 5
10mm
Littorina
arcana
n = 24
Littorina
compressa
n = 4
C
Young inversion
Older inversion
Inversion without heterokaryotype
D
1cM
Split 1
S1
Split 2
S2L
S2R
Split 3
S3LL
S3LR
no split
S3RR
HInd
E
Complex inversion
Colinear genome
Simple inversion

Figure 1: Distribution of samples included in this study and schematics of detection methods. **A)** map of sampling locations. Solid symbols represent the presence of *Littorina saxatilis* (circle), *L. arcana* (square) and *L. compressa* (triangle) at each site according to Reid (1996). Hollow symbols represent absence. The base map was produced in QGIS using Natural Earth countries data with 1:50m resolution (crs = WGS 84; https://www.naturalearthdata.com/downloads/). **B)** Shell photos of each species, genetic group, and ecotype. All four *L. saxatilis* ecotypes (Reid, 1996, pp. 305–318) are shown, with only *crab* and *wave* existing in Iberia. An additional photo is included for both the Iberian and Northern saxatilis groups, of a snail which does not fit an established ecotype. For Northern saxatilis, the *crab*, *wave* and *brackish* photos were taken from snails collected in Sweden, the *barnacle* and *L. arcana* snails were from the North-East coast of England, *L. compressa* was from Northern Wales, and the 'other' snail was from Iceland. For the Iberian saxatilis, the *crab* and *wave* photos were taken from snails collected from Centinela, while the 'other' snail was from near Burela. **C-D)** Schematics of heterozygosity split approach. **C)** shows hypothetical results for three different patterns that could indicate inversions. **D)** Diagram of how split positions are determined. Each split is represented by a thick yellow line. Each linkage group is split three times for an individual. The final row shows the segments that are used to calculate H_{Ind}. **E)** Example PCA plots showing patterns of a complex double inversion (LGC6.1/2), a colinear segment of the genome (LG8) and simple inversion (LGC1.2) . Heterokaryotype clusters have been circled.

Sample collection:

The dataset used for this study was gathered for a phylogenetic study of the *L. saxatilis* species complex (Stankowski, Zagrodzka, Galindo, et al., 2023; Stankowski, Zagrodzka, Garlovsky, et al., 2023). Snails were collected between 2014 and 2020, by several different collectors, from 18 locations across the North Atlantic (Figure 1a; Table S1). Where possible, all ecotypes and species present were sampled. However, not all locations had habitats or shell characteristics that were typical of the recognised ecotypes described by Reid (1996, pp. 305–318). In such cases these individuals were classified as '*other'*. The collection site details and full list of collectors are available in the supplementary materials (Table S1).

Only reproductively mature individuals were sequenced. Maturity was determined by examining the reproductive anatomy. *L. saxatilis* and *L. arcana* were also distinguished by female reproductive anatomy. Males could not be distinguished with morphology, hence only females were analysed in locations where these species overlapped.

Samples were split into three genetic groups for analysis: Northern saxatilis, Iberian saxatilis and *L. arcana* (Figure 1). These were chosen based on known genetic differences to reduce the impact of geographic isolation (Northern vs Iberian; Doellman et al., 2011; Morales et al., 2019; Panova et al., 2014; Stankowski, Zagrodzka, Galindo, et al., 2023;

Tirado et al., 2016) and reproductive isolation (*L. saxatilis* vs *L. arcana*; Stankowski et al., 2020) while still maintaining a sufficient sample size to identify polymorphic inversions. *L. compressa* was used as an outgroup and not included in inversion-detection analyses due to low sample size (four). One *L. saxatilis* (IMI_6_2) was excluded outright due to missing collection information. Sample sizes are presented in Table S1.

Sequencing and SNP calling

DNA was extracted from a small piece of foot tissue from each snail using a CTAB protocol (Panova et al., 2016). These DNA samples were sent to Edinburgh Genomics (University of Edinburgh) for library preparation using a 350 bp insert TrueSeq DNA Nano gel-free approach and sequenced on an Illumina HiSeqX machine with 150bp pair-end reads and 15X target coverage. Reads were aligned and variants were called following the pipeline in Stankowski et al. (2020). Briefly, reads were trimmed and then mapped to the *L. saxatilis* genome assembly (Westram et al., 2018) using BWAmem (Li & Durbin, 2009), followed by variant calling using GATK4 (McKenna et al., 2010). Called SNPs were then filtered to remove those with multiple alleles (> 2), low mapping quality (Q < 30), minor allele frequencies (maf) below 0.05, and extreme depth of coverage (5 > AD > 35). Only SNPs mapped onto a contig on the *Littorina saxatilis* linkage map were retained.

Updating inversion positions on the linkage map

SNP genotypes were assigned by contig onto a new version of the *L. saxatilis* linkage map, which accounts for map compression within inversions (see supplementary materials). This combines a map from a *crab* ecotype family (Westram et al., 2018, Supp. mat.) with an F2 *crab-wave* map (Koch et al., 2021). This consensus map was filtered to remove any SNP markers within a contig that mapped to a different linkage group or were located > 2cM from the average map position of that contig. A new average map position was then assigned to

each contig. Linkage groups (LG) 10 and 14 retained positions from the *crab* map because few informative markers were found on the *crab-wave* map. Our WGS SNPs did not match the markers used for map construction, thus the average map position of each contig was assigned to SNPs in this dataset. Published positions of the inversion boundaries previously detected in this species (Hearn et al., 2022 for LG12; Westram et al., 2021 for others) were transferred to the new map using contigs that previously mapped to inversion boundaries. Conservatively, the widest distance between inversion boundary positions was adopted.

Inversion detection approaches

Inversions were detected in each genetic group with two complementary approaches: variation in average observed heterozygosity ('heterozygosity splits') and principal component analysis (PCA). Approaches used elsewhere, such as 'LDna' (Kemppainen et al., 2015), 'local_PCA' (Li & Ralph, 2019) or 'asaph' (Nowling et al., 2020), could not be used for our data due to the lack of contiguity of the reference genome and the strong population structure across the species range. For consistency between our approaches, we used contigs in 1cM non-overlapping windows (see below), chosen to give sufficient SNPs for informative PCA results. Visual inspection was used to determine if the positions of inversions were consistent between approaches and genetic groups.

Inversion detection approach 1: Heterozygosity splits

Inversions were identified by scanning each linkage group for significant shifts ('splits') in average individual heterozygosity (i.e. the proportion of SNPs in a genomic window that are heterozygous) among snails in each genetic group (i.e. Northern saxatilis, Iberian saxatilis and *L. arcana*). Heterozygosity is expected to differ among inversion karyotypes and to differ from regions outside the inversion (colinear genome). Young inverted homokaryotypes usually have reduced heterozygosity as they recently expanded

from a single mutated haplotype. As they get older, inverted homokaryotypes accumulate genetic diversity from mutation and gene flux. Heterokaryotypes for older inversions have increased heterozygosity as they contain two isolated haplotypes that are likely to be differentiated at multiple sites. Finally, non-inverted homokaryotypes are expected to have levels of heterozygosity similar to or slightly lower than the genetic background as the inversion reduces effective population size (in proportion to inversion frequency) by acting as a localized reproductive barrier for part of the genome. However, gene flux and selection can distort these expectations. By comparing heterozygosity along the genome among individuals, an inversion can be identified by a cluster of splits in average individual heterozygosity that groups snails into two or three sets, at least one of which differs from the background (Figure 1c).

Splits in average individual heterozygosity (H_{Ind}) were identified by dividing each linkage group into blocks of similar heterozygosity. Specifically, we designed a hierarchical split-function (Figure 1d). For each individual, H_{Ind} was calculated on either side of potential splits, distributed every 1cM along each linkage group. A model was fitted to the data with a single mean H_{Ind} and beta-binomial distribution, with a fitted dispersion parameter ρ. Models were then compared for all possible splits of the LG into two contiguous segments with different H_{Ind} means (but the same dispersion), to find the split with the highest likelihood. The best split model was retained if the difference in likelihood compared to the no-split model was significant (using a chi-squared test with $\chi^2 = -2(LL0 - LL1)$ and $p<0.01$). The split test was then applied to each resulting segment of the linkage group, followed by a third level of splitting. If, at any level, there was no significant split, the splitting function was stopped for that segment. Thus, the splitting function yielded between 1 and 8 segments with different H_{Ind} means. This process was repeated for each LG and individual.

A permutation test was run to look for clusters of splits for each linkage group and genetic group. Counts of significant splits were permuted over the possible boundaries between 1cM windows to determine whether splits were clustered in certain parts of a linkage group. The observed variance in counts across a linkage group was calculated from 3cM sliding windows of split counts, where each 3cM window was the sum of the three 1cM windows. Splits were then shuffled randomly among 1cM windows by drawing counts from a multinomial distribution, ignoring any windows that had no markers on the linkage map. Permuted variance across a linkage group was calculated among 3cM windows, following the approach for observed variance. Empirical *p*-values were calculated for the observed variance compared to permuted variances from 10,000 replicates. Once all genetic groups were tested for each linkage group, the empirical *p*-values were adjusted for multiple-comparisons with a Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Split clusters gave some indication of possible inversion boundaries, however clusters could also represent changes in heterozygosity for other reasons, such as near centromeres or repeat regions. Visual inspection was used to infer if split clusters matched the patterns of inversions. H_{Ind} values were inspected using a plot of H_{Ind} for all individuals in a genetic group across each linkage group. Any bifurcations or trifurcations in H_{Ind} associated with split clusters indicated the presence of an inversion (Figure 1c).

Note that other types of chromosomal rearrangements and gaps in the linkage map can also cause heterozygosity differences among chromosomal regions and individuals (Mérot et al., 2020), which may falsely be called as inversions using this type of analysis alone.

Inversion detection approach 2: Window-based PCA

A second approach used a 1cM window-based PCA to identify inversions by looking for clusters on the first principal component of SNP genotypes within each window. Inversions are expected to separate individuals into three distinct clusters representing the separate karyotypes (Hanlon et al., 2022; Kemppainen et al., 2015; Nowling et al., 2020). The expected pattern is more complex for overlapping inversions on LG6 and LG14 of *L. saxatilis*, where three arrangements may be present forming six clusters on a PC1 vs PC2 plot in a triangular pattern (Faria, Chaube, et al., 2019). Only PC1 was used in detection but further analysis used PC1 and PC2 (see below). PCAs were conducted on all SNPs inside each 1cM window using the *dudi.pca* command from the *adegenet* R package (Jombart, 2008; Jombart & Ahmed, 2011). Missing genotypes were imputed based on the mean score of the window with the *scaleGen()* function. PC1 scores for successive 1cM windows were reorientated by switching the sign of any window where scores were negatively correlated with the preceding window. Adjusted PC1 scores were then visually inspected on the linkage map to detect regions with three clusters.

This approach is very similar to the detection approaches used in 'local_PCA' (Li & Ralph, 2019) and 'asaph' (Nowling et al., 2020). However, we developed a custom script so that local PCA could be run across the linkage map, as the reference genome is too fragmented to run existing tools.

Inversion boundaries

Inversion signals from the heterozygosity splits and window based PCA were used to determine approximate coordinates of inversion boundaries on the linkage map. However, these approaches were primarily meant to indicate the presence of polymorphic inversions, not specific breakpoint positions, thus the edges of inversion signals can vary by a few cM

among genetic groups. To attain a clearer resolution of inversion positions, the inferred boundaries were compared to the coordinates of published inversions (Hearn et al., 2022; Westram et al., 2021) on the new linkage map. If the published boundary positions corresponded to inversion patterns, the published positions were used to filter SNPs for genotyping the inversions with PCA; otherwise we used our inferred boundaries of each genetic group for SNP filtering. Modified boundaries were used for the two missing inversions (see below) to avoid overlaps with neighbouring inversions; for LGC14.3 ([L]inkage [G]roup [C]luster 14.3; notation following Faria et al., 2019) SNPs between 12 and 34.66cM were used, and for LGC12.3 SNPs between 46 and 50.09cM were used. For new putative inversions inferred boundaries were used.

Inspecting and genotyping inversions

Putative inversions were inspected by genetic group through a PCA of all SNP genotypes within each inversion. These PCAs were run using the same command as the window-based PCA. Additional PCAs were run for the full dataset (global PCA) and using only samples from locations where *L. arcana* was collected (*arcana-saxatilis* PCA). Overlap between the coordinates of some inversions in LG12 and LG14 (other than the complex inversions) was observed likely due to low resolution of the genetic maps (Figures S12 and S14, respectively). In such cases, the inversions were genotyped excluding SNPs within those overlapping regions (see above).

As a control, the same analysis was run for the non-inverted (colinear) regions of all linkage groups except for LG10 and LG12, which are mostly covered by inversions (Westram et al. 2021; Hearn et al., 2022). Colinear PCAs included all sites outside of an inversion, merging any potential left and right segments of the respective linkage group (e.g.

LG9), excluding a 2cM buffer around inversions to account for the imprecise positions of boundaries (see Faria, Chaube, et al., 2019 & Westram et al., 2021).

The presence of an inversion was supported by inspecting a scatterplot of PC1 vs PC2 for three clusters of points, which should be divided on PC1 (Figure 1e). For the complex inversions on LG6 and LG14, six clusters were expected forming *a triforce* shape (i.e. triangle within a triangle, from the Zelda video game series) with each vertex-cluster representing one of three possible homokaryotypes and the heterokaryotype clusters found on each vertex of the inner triangle (Figure 1e).

Clusters were assigned with a K-means clustering algorithm in R. One hundred random starting positions were tested for each inversion. The starting positions that had the highest variation between groups, measured as the sum of squares between groups, were retained. Two clustering counts ($2 \leq K \leq 3$) were tested for most inversions, additional clusters were considered ($2 \leq K \leq 9$) for the complex inversions. Different K values were compared using the silhouette method, in the R package *cluster*, retaining the K with the highest mean silhouette value. For simple inversions, the expected K = 2 or 3 fits were found using PC1 for clustering. For complex inversions both PC1 and PC2 were considered in the K-means clustering. In the case of LGC14.1/2, 6 clusters were found only after separating the inversion into two parts (i.e. the simple section of the inversions, LGC14.1 with three clusters; and the complex section where the two inversions overlap, LGC14.2 with six clusters).

Cluster scores were converted into genotypes by adjusting the labels for a consistent order among inversions, whereby the homokaryotypic cluster containing the most *crab* samples was labelled 'RR', the reference arrangement, while the other homokaryotypic cluster was labelled alternative 'AA'. LGC6.1/2 and LGC14.2 had three arrangements. For these cases, the homokaryoptypic cluster furthest from 'RR' on PC1 was labelled 'A_1A_1' and

the final homokaryotypic cluster was labelled 'A_2A_2'. Some inversion labels were manually adjusted to correct for noise from geographic or species diversity (see supplementary material). Inversion PCAs with only two clusters could represent 'RR' and 'AA' or 'RR' and 'RA'. In such cases the global PCA, heterozygosity split plots, and average heterozygosity scores were used to determine the presence or absence of the heterokaryotypic cluster. The frequency of each inversion arrangement was calculated from the counts assigned to clusters. An inversion was considered polymorphic at a sampling site if individuals from that site were heterokaryotypic, or more than one arrangement homokaryotype was present.

Inferring ancestral arrangement using L. compressa

The ancestral arrangements for each inversion were inferred by projecting genotypes from four *L. compressa* onto the PCA plot using the R function *suprow* from the *adegenet* package. Projection improved the resolution of inversion genotypes by preventing the PCA from being dominated by interspecific differences. The arrangement shared with *L. compressa* was considered ancestral. However, in a few cases inversions were also polymorphic in *L. compressa* (Table 2). Deeper sampling of *L. compressa* may reveal additional polymorphic inversions and, therefore, our inference of ancestry should be seen as preliminary.

Association of inversion frequencies with ecotypes

To identify the inversions contributing to divergence we compared arrangement frequencies among ecotypes while controlling for variation among sampling locations. Three logistic regression models were run for each inversion:

Null: N_{inv} ~ location
Eco: N_{inv} ~ ecotype + location
Int: N_{inv} ~ ecotype * location

Inversion frequencies (N_{inv}) were considered as a count of two possible states either the 'R' or 'A' arrangement. For the complex inversions, LGC6.1/2 and LGC14.2, the 'A_1' and 'A_2' arrangements were summed to get the 'A' state count. Subsequent models were run for these inversions to test for differences between 'A_1' and 'A_2'. Models were evaluated through a hierarchical comparison of AIC. First, the null model was compared to the ecotype model to establish if there was an ecotype effect (i.e. $AIC_{Null} - AIC_{Eco}$). If $\Delta AIC > 2$, then *Int* and *Eco* were compared to establish the significance of the ecotype-location interaction. If $AIC_{Null} < AIC_{Eco}$, then *Null* and *Int* were compared instead. Only when $AIC_{Int} - AIC_{Eco} > 2$ was *Eco* the best model. The best model was further evaluated to determine the proportion of variance explained using Cohen's pseudo-R^2 (1 – null deviance / fitted model deviance) and a p-value estimated from the deviance associated with each term (Tables S4-6). P-values were adjusted using a Benjamini-Hochberg correction.

Logistic regressions were run for two ecotype and two species contrasts per inversion. Before running the logistic regression, the data were filtered to keep samples only from locations where both focal ecotypes (or species) were collected, as ecotypes have different distributions (with the *crab* and *wave* being the most widely distributed; Reid, 1996, pp. 305–318). The ecotype contrasts were between *crab* and *wave* or *wave* and *barnacle.* Other contrasts were not possible due to insufficient location overlaps. For the species contrasts, we compared the arrangement frequencies in *L. arcana* separately with the *wave* and *crab* ecotypes of *L. saxatilis*, for sites where both species were sampled. As *crab* and *L. arcana* were only co-sampled in a single location, the interaction model was not fitted, and the null model was simplified to $N_{inv} \sim 1$. Arrangement counts were expected to be more similar between *L. arcana* and the *wave* ecotype because *L. arcana* was found mainly in wave-exposed habitats at the sampled locations.

Results:

Table 1: Summary of inversion detection for all genetic groups and linkage groups. New putative inversions are marked with '!'. '✓' represents a clear pattern of a polymorphic inversion. Smaller '✓' are cases where the heterokaryotype was not detected, '?' represent uncertain patterns, and '×' represents no pattern. Inversion positions were inferred from both identification approaches, each separated by a '|'. When both approaches aligned only a single set of boundaries is presented. Missing inversions are represented by '-'. '*' indicate matches with the published inversion boundaries (Westram et al., 2021; and Hearn et al., 2022 for LGC12.2 and LGC12.3), converted onto the new linkage map (in bold). PCA were run per genetic group for all SNPs within these positions to validate the presence of inversions: 'K' is the optimal number of clusters found via K-means using PC1, or PC1 and PC2 for complex inversions. 'Fig ref' indexes the corresponding plots in the supplementary material.

LG	Inversion	Group	Hetero. splits	PCA per cM	Inv. pos. (cM) Het | PCA	K	Fig. Ref.
LG1	LGC1.1	Northern saxatilis	✓	?	0-14	2	S1
		Iberian saxatilis	?	?	0-16 | 0-15	3	
		Littorina arcana	✓	✓	0-17 | 0-15	3	
					0-19		
	LGC1.2	Northern saxatilis	✓	✓	77-95.6* | 76-95.6*	3	S1
		Iberian saxatilis	✓	✓	76-95.6*	3	
		Littorina arcana	✓	?	77-95.6*	3	
					76.1-95.6		
LG2	LGC2.1	Northern saxatilis	✓	✓	0-14 | 0-11*	3	S2
		Iberian saxatilis	✓	✓	0-12	2	
		Littorina arcana	✓	✓	0-12	3	
					0.3-10.9		
LG3	No inversion						S3
LG4	LGC4.1	Northern saxatilis	✓	?	0-21 | 7-22	3	S4
		Iberian saxatilis	✓	✓	0-23	2	
		Littorina arcana	✓	×	0-24 | -	2	
					5.1-22.8		
LG5	LGC5.1!	Northern saxatilis	✓	?	15-46 | -	3	S5
		Iberian saxatilis	✓	?	15-47 | 16-47	2	
		Littorina arcana	?	?	16-47 | -	3	
LG6	LGC6.1/2	Northern saxatilis	✓	✓	0-19	6	S6
		Iberian saxatilis	✓	✓	0-19	3	
		Littorina arcana	?	?	0-19 | 0-18	4	
					0-24.3		
LG7	LGC7.1	Northern saxatilis	✓	✓	37-43*	3	S7
		Iberian saxatilis	?	?	37-43*	2	
		Littorina arcana	✓	✓	37-42*	3	
					36.6-42.6		
	LGC7.2	Northern saxatilis	✓	✓	45-59.5 | 46-55	3	S7
		Iberian saxatilis	?	✓	45-59.5 | 46-55	3	
		Littorina arcana	✓	✓	44-59.5 | 47-55	3	
					46.9-58.2		

LG8	No inversion						S8
LG9	LGC9.1	Northern saxatilis	✓	✓	18-42*	3	S9
		Iberian saxatilis	✓	✓	18-42* | 18-41*	3	
		Littorina arcana	?	✗	18-42* | -	2	
					18.7-42.1		
	LGC9.2[!]	Northern saxatilis	?	✗	52-59.5 | -	3	S9
		Iberian saxatilis	✓	?	52-59.5 | 52-55	3	
		Littorina arcana	✓	?	53-59.5 | 52-55	3	
LG10	LGC10.1	Northern saxatilis	✓	✓	0-3*	3	S10
		Iberian saxatilis	✓	✓	0-3*	3	
		Littorina arcana	✓	✓	0-3*	3	
					0.9-2.8		
	LGC10.2	Northern saxatilis	✓	✓	4-45.5*	3	S10
		Iberian saxatilis	✓	✓	14-45.5	2	
		Littorina arcana	✓	✓	14-45.5	3	
					3.1-44.1		
LG11	LGC11.1	Northern saxatilis	✓	✓	56-70 | 58-69	3	S11
		Iberian saxatilis	✓	✓	56-70 | 58-69	2	
		Littorina arcana	✓	✓	55-70 | 58-69	3	
					50.9-69.2		
LG12	LGC12.1	Northern saxatilis	✓	✓	0-38 | 16-34	3	S12
		Iberian saxatilis	✗	✗	-	3	
		Littorina arcana	✓	✓	0-39 | 16-38	3	
					9.3-40.5		
	LGC12.2	Northern saxatilis	✓	✓	39-46 | 35-46	3	S12
		Iberian saxatilis	✓	✓	39-49 | 39-46	3	
		Littorina arcana	✗	✗	-	3	
					39.3-45.6		
	LGC12.3	Northern saxatilis	?	?	39-46	2	S12
		Iberian saxatilis	?	?	39-46	3	
		Littorina arcana	✗	✗	-	2	
					40.5-50.1		
	LGC12.4	Northern saxatilis	✓	✓	52-69.9 | 51-69.9	3	S12
		Iberian saxatilis	✓	✓	52-69.9 | 51-69.9	3	
		Littorina arcana	✓	✓	53–69.9 | 52-69.9	3	
					48.3-68.2		
LG13	No inversion						S13
LG14	LGC14.1/2	Northern saxatilis	✓	✓	0-12*	7	S14
		Iberian saxatilis	✓	✓	0-12*	5	
		Littorina arcana	✓	✓	0-12*	9	
					0.7-11.4		
	LGC14.3	Northern saxatilis	✗	✗	-	2	S14
		Iberian saxatilis	✗	✗	-	2	
		Littorina arcana	✗	✗	-	2	
					10.2-34.7		

LG15	No inversion						S15
LG16	No inversion						S16
LG17	LGC17.1	Northern saxatilis	✓	✓	47-62.6*	3	S17
		Iberian saxatilis	✓	✓	47-62.6*	2	
		Littorina arcana	✗	✗	-	3	
					47.5-62.0		

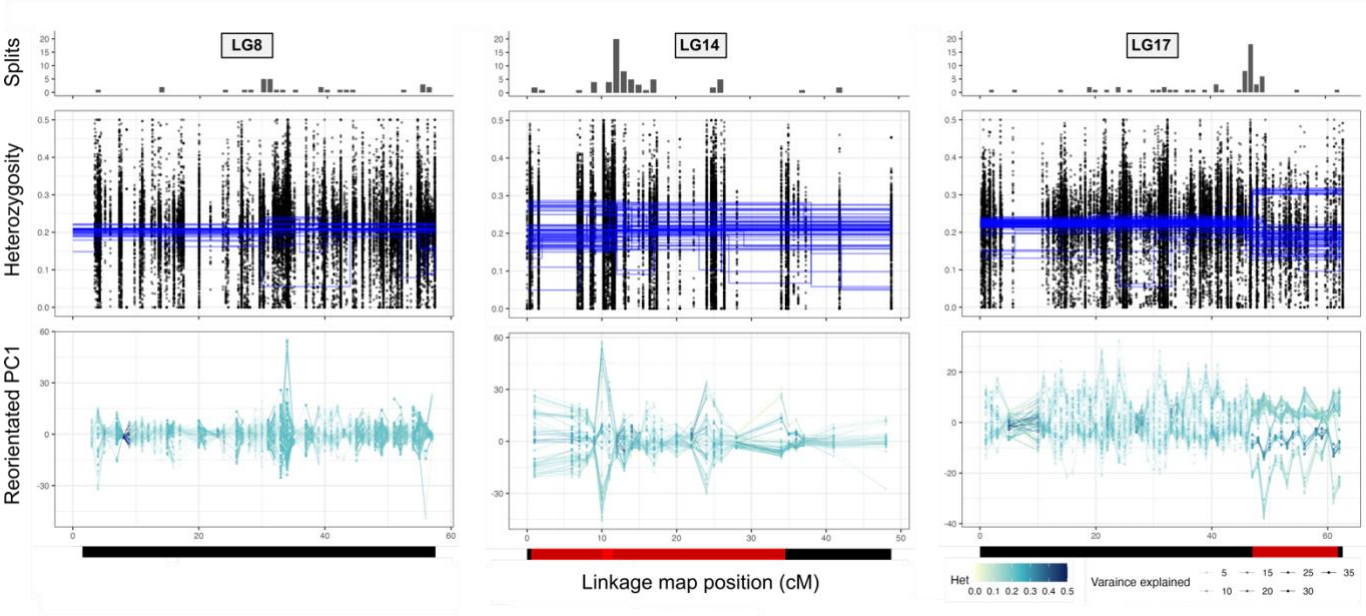

Figure 2: Three examples of the inversion detection results using Northern saxatilis data. From left to right, an example of a linkage group without an inversion (LG8), an example of an opaque pattern of multiple inversions (LG14), and a clear example of a single inversion (LG17). *Top panel*: shows the number of significant splits of heterozygosity between 1cM windows of the linkage map. *2nd panel*: shows the results of the heterozygosity splits approach, where each dot represents the proportion of heterozygous SNPs per contig for each snail, and the blue lines represents the average heterozygosity between two significant splits for an individual (H_{ind}). The y-axis was limited to 0.5 and only contigs on the linkage map are shown. *3rd panel*: PC1 scores from a PCA for each 1cM window. These scores were reorientated to positively correlate with the preceding window's scores. Lines and dots are coloured by average heterozygosity across the window and shaded by the percentage variation on PC1. Heterozygosity exceeding 0.5 was coloured dark blue. *Bottom panel*: red bars show the position of published inversions (Westram et al., 2021) on each linkage group. Overlapping boundaries between LGC14.1/2 (left) and LGC14.3 (right) are in lighter red. Plots for the other groups and inversions are in the supplementary material (Fig S1-S17).

Detecting inversions:

The heterozygosity split approach and window-based PCA provided evidence of polymorphic chromosomal inversions on most linkage groups (Table 1; Figures 2 and S1-S17), most of which were identified in previous studies (Faria, Chaube, et al., 2019; Hearn et al., 2022; Westram et al., 2021). Heterozygosity splits clustered significantly on all linkage

groups, with the ratio of observed versus permuted variance being higher on those containing published inversions (Table S2). In most cases these clusters matched published positions of inversion boundaries (Hearn et al., 2022; Westram et al., 2021), with patterns of H_{Ind} and window-based PCA consistent with inversions (Table 1). Typically, the detection methods were supported by PCA of the whole inverted region, which showed either 3 or 6 clusters, with some variation within clusters driven by geography, and clearly different from PCAs for the collinear regions (Figures 3 and S18).

A clear example is LGC17.1 (Figure 2): this inversion was detected in all *L. saxatilis* populations, with consistent boundaries corresponding to those previously reported from Sweden (Westram et al., 2018, 2021), but it was not polymorphic in our *L. arcana* sample. However, not all published inversions were as easily identified, or their boundaries did not clearly match expectations, especially on LG14 (Figure 2) and LG10 (Figure S10). Heterozygosity varied widely among individuals across the length of both linkage groups, making inversion patterns hard to distinguish. Iberian snails had three inversions (LGC2.1, LGC4.1 & LGC17.1) where only putative homokaryotypes for each arrangement were found (Figures S2, S4 and S17), possibly reflecting strong genetic differentiation between ecotypes (Kess & Boulding, 2019). Finally, patterns of inversions at the expected genomic locations of LGC12.3 and LGC14.3 were absent across the entire data set.

Table 2: Polymorphic status and ancestry of inversions. The PCA of each genetic group was used to determine the polymorphic status at each location. If samples from a single location were present in two or more karyotypes, or just the heterokaryotype, the inversion was considered polymorphic. The heterokaryotype cluster was determined by inspecting the global PCA plot. Polymorphic inversions are marked with a 'Y' on a green background. Non-polymorphic inversions are marked with a 'N' on a red background. Ancestry was determined by the position of *L. compressa* with either the 'A' or 'R' arrangement cluster in the global PCA. Any 'P' represents a case where *L. compressa* overlapped with the heterokaryotype or occurred in two clusters suggesting polymorphism. Any plots without clear clusters are represented by '-'. **N** represents the sample size. The two novel inversions from this study are listed separately on the far right. Site coordinates are in Table S1.

Location	**N**	LGC1.1	LGC1.2	LGC2.1	LGC4.1	LGC6.1/2	LGC7.1	LGC7.2	LGC9.1	LGC10.1	LGC10.2	LGC11.1	LGC12.1	LGC12.2	LGC12.3	LGC12.4	LGC14.1	LGC14.2	LGC14.3	LGC17.1	LGC5.1	LGC9.2
Northern *Littorina saxatilis*																						
Arsklåvet	4	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	Y	N	-
Broad Haven	1	N	N	N	Y	Y	Y	N	Y	Y	N	N	N	Y	-	N	Y	Y	-	N	N	-
Ceann Tra	2	N	N	Y	N	N	Y	Y	N	Y	N	N	N	N	-	N	N	N	-	N	Y	-
Dersingham	2	N	N	N	N	N	N	N	N	N	N	N	N	N	-	N	N	N	-	N	N	-
Holyhead	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	Y	Y	-
Laugarnes	2	N	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	N	N	-
Oban	1	N	N	N	Y	N	Y	N	N	N	N	N	N	N	-	Y	Y	Y	-	Y	N	-
Port Saint Mary	1	N	N	N	N	N	Y	N	N	N	N	N	N	Y	-	Y	N	N	-	N	N	-
Ramsö	4	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	Y	N	-
Ravenscar	7	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	Y	N	-
Roscoff	6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	Y	N	-
Saint Abbs	2	Y	N	Y	N	Y	Y	N	N	Y	N	Y	N	N	-	Y	N	N	-	Y	Y	-
Thornwick	8	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	N	Y	-
Tjärnö	2	N	N	N	N	N	N	N	N	N	N	N	Y	Y	-	N	N	N	-	Y	N	-
Trondheim Fjord	3	N	N	Y	N	N	Y	Y	Y	N	N	Y	N	Y	-	Y	Y	Y	-	N	Y	-
Varanger Fjord	4	N	Y	Y	Y	Y	N	Y	Y	N	N	N	Y	Y	-	N	N	N	-	N	Y	-
White Sea	2	N	N	N	N	N	N	N	N	N	N	N	N	Y	-	N	Y	Y	-	Y	N	-
York ME	4	N	Y	Y	N	Y	N	Y	Y	N	Y	N	N	Y	-	Y	Y	Y	-	N	N	-
Littorina arcana																						
Amble	1	N	N	Y	-	-	Y	N	-	N	Y	N	N	-	-	N	-	Y	-	-	Y	N
Broad Haven	1	Y	Y	N	-	-	N	N	-	N	N	N	N	-	-	N	-	Y	-	-	N	N
Holyhead	4	Y	Y	Y	-	-	N	N	-	Y	Y	Y	Y	-	-	Y	-	N	-	-	N	Y
Ravenscar	7	N	Y	Y	-	-	Y	Y	-	Y	Y	Y	Y	-	-	Y	-	Y	-	-	N	N
Roscoff	4	N	N	N	-	-	N	Y	-	Y	N	Y	Y	-	-	Y	-	Y	-	-	N	N
Saint Abbs	2	Y	N	N	-	-	Y	Y	-	N	N	Y	Y	-	-	Y	-	Y	-	-	Y	N
Trondheim Fjord	1	N	Y	N	-	-	N	N	-	N	N	Y	N	-	-	N	-	Y	-	-	N	N
Varanger Fjord	4	Y	Y	N	-	-	Y	Y	-	Y	Y	Y	N	-	-	Y	-	Y	-	-	N	N
Iberian *Littorina saxatilis*																						
Burela	2	-	N	N	N	Y	-	Y	N	N	Y	Y	-	N	-	N	N	-	-	N	N	N
Centinela	10	-	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	-	Y	-	Y	Y	-	-	Y	Y	Y
Ancestry		R	R	A	P	R	A	A	R	A	A	R	R	P	P	A	P	A_1	-	R	-	R

Distribution of inversions within and between species:

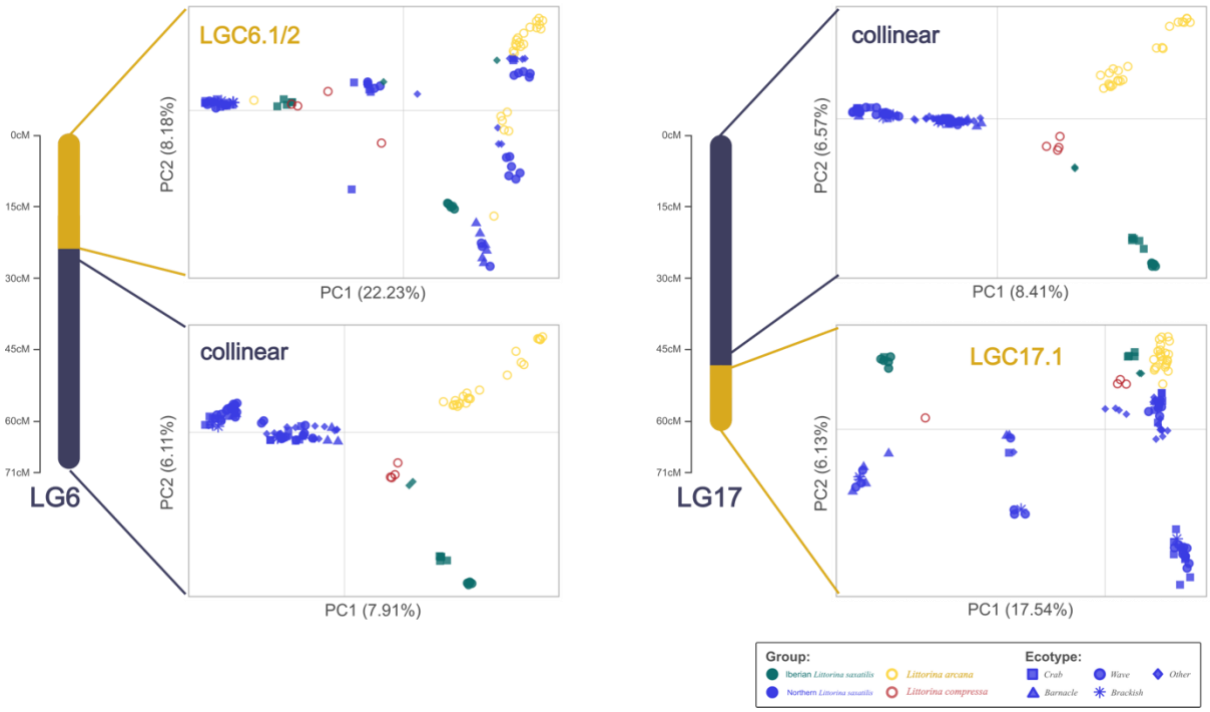

Figure 3: PCA plots for the colinear and inverted segments of two linkage groups using the combined data from all snails. Points are coloured by genetic group and shaped by ecotype. PC1 is on the x-axes and PC2 on the y-axes. Gridlines mark the origin. Bars left of the plots represent linkage groups, with the inverted regions coloured in yellow. Plots for other linkage groups are in the supplementary material (Figure S18).

Most published inversions (89%) were identified as widespread polymorphisms across the species range of *L. saxatilis*. Indeed, all inversions were polymorphic in Northern saxatilis, excluding the two aforementioned exceptions of LGC12.3 and LGC14.3 (Table 2; Figure S18). Four additional inversions were not clearly identified in the PCA of Iberian saxatilis (LGC1.1, LGC7.1, LGC12.1 and LGC14.2), although LGC1.1 and LGC7.1 were identifiable from the global PCA of all snails (Figure S18). LGC12.1 may also be present in the global PCA, however its polymorphic status rests on a single snail from a genetically distinct Iberian location (Burela). Lastly, no heterokaryotype clusters were identified in Iberia for three inversions (LGC2.1, LGC4.1 and LGC17.1) from either the Iberian (Figure S19) or global PCA (Figures 3 and S18).

For most inversions, the PCAs of inversion areas clustered snails by karyotype on PC1, combined with a signal of geographic structure on PC2 (Figures 3 and S18-19). The

strongest geographic signal was the separation of Iberian from the Northern saxatilis, which typically was stronger in one of the two arrangements. Iberian saxatilis also formed a distinct cluster in the colinear genome (Figure S18). A finer scale geographic separation also existed within Northern saxatilis (Figure S19-20). North American samples often were outliers on PC2 and snails sampled from the North Sea (excluding Dersingham, UK) clustered together, within and outside inversions. All inversion polymorphisms were identified in the most densely sampled locations in Sweden (Ramsö), France (Roscoff), Wales (Holyhead), and Spain (Centinela), with a diffuse presence at other sites (Table 2). Of note, two thirds of inversions were polymorphic in the single North American site (York, ME, USA) and no polymorphisms were identifiable in a *brackish* ecotype site (Dersingham, UK). However, the lack of observed polymorphism at a site does not mean inversions are absent or fixed, as many sites had too few samples (≤ 2) to detect rare arrangements.

Most of the previously identified inversion polymorphisms observed in *L. saxatilis* were also polymorphic in *L. arcana* (Table 2). On the PCA plots of inversion regions, *L. arcana* typically clustered separately from *L. saxatilis* within each karyotype (Figures 3 and S18). For the colinear genome, *L. arcana*, Northern saxatilis and Iberian saxatilis formed three groups, with similar separation on PCA axes (Figure S18). Within inversions, PC2 explained most of the difference between species, with the variation on PC1 in *L. arcana* being consistent with the inversion karyotypes in *L. saxatilis*, suggesting the presence of the same arrangements (Figure S21). There was also finer scale geographic separation within *L. arcana*, consistent with the North Sea clustering seen in *L. saxatilis* (Figure S19).

Only two inversion polymorphisms (LGC12.2 and LGC17.1) found in *L. saxatilis* were not identified in *L. arcana* (Table 1). Both inversions were polymorphic in the outgroup, *L. compressa*, suggesting that one arrangement was either lost or fixed in *L. arcana* (Figure S18). The status of LGC4.1 and LGC14.1/2 in *L. arcana* were uncertain. PCA

showed diffuse clusters of samples for the same position on the linkage map, but they were poorly aligned with the *L. saxatilis* clusters (Figures 2, S18 and S21). Further investigation identified weak patterns in *L. arcana* consistent with polymorphism for LGC4.1, but only the first part of LGC14.1/2 (LGC14.1, which appears to be absent in *L. arcana*) could be aligned (Figure S22).

Identification and distribution of new inversion polymorphisms

In addition to identifying previously described inversions from Sweden (Hearn et al., 2022; Westram et al., 2021), new patterns were found on LG5 and LG9 that are consistent with polymorphic inversions (Figures S5 and S9). Following established convention (Faria, Chaube, et al., 2019), these were labelled LGC5.1 & LGC9.2 (Table 1). However, these patterns were not found in all genetic groups and their patterns were inconsistent between the heterozygosity split and PCA-based approaches. PCA for LGC5.1 was similar to the collinear region of LG5 casting doubt on the inversion status of this genomic region (Figure S18). While PCA for LGC9.2 there was a signal the looked like an inversion polymorphism at low frequency, but upon further inspection this was seen to be geographic variation (see supplementary method; Figure S22). The patterns for LGC9.2 for other genetic groups and the global PCA was clear (Table 2; Figure S18). Both putative inversions appear fixed for the more common arrangement in Sweden (Table 2), explaining why they could have been missed in previous studies (Faria, Chaube, et al., 2019; Westram et al., 2021).

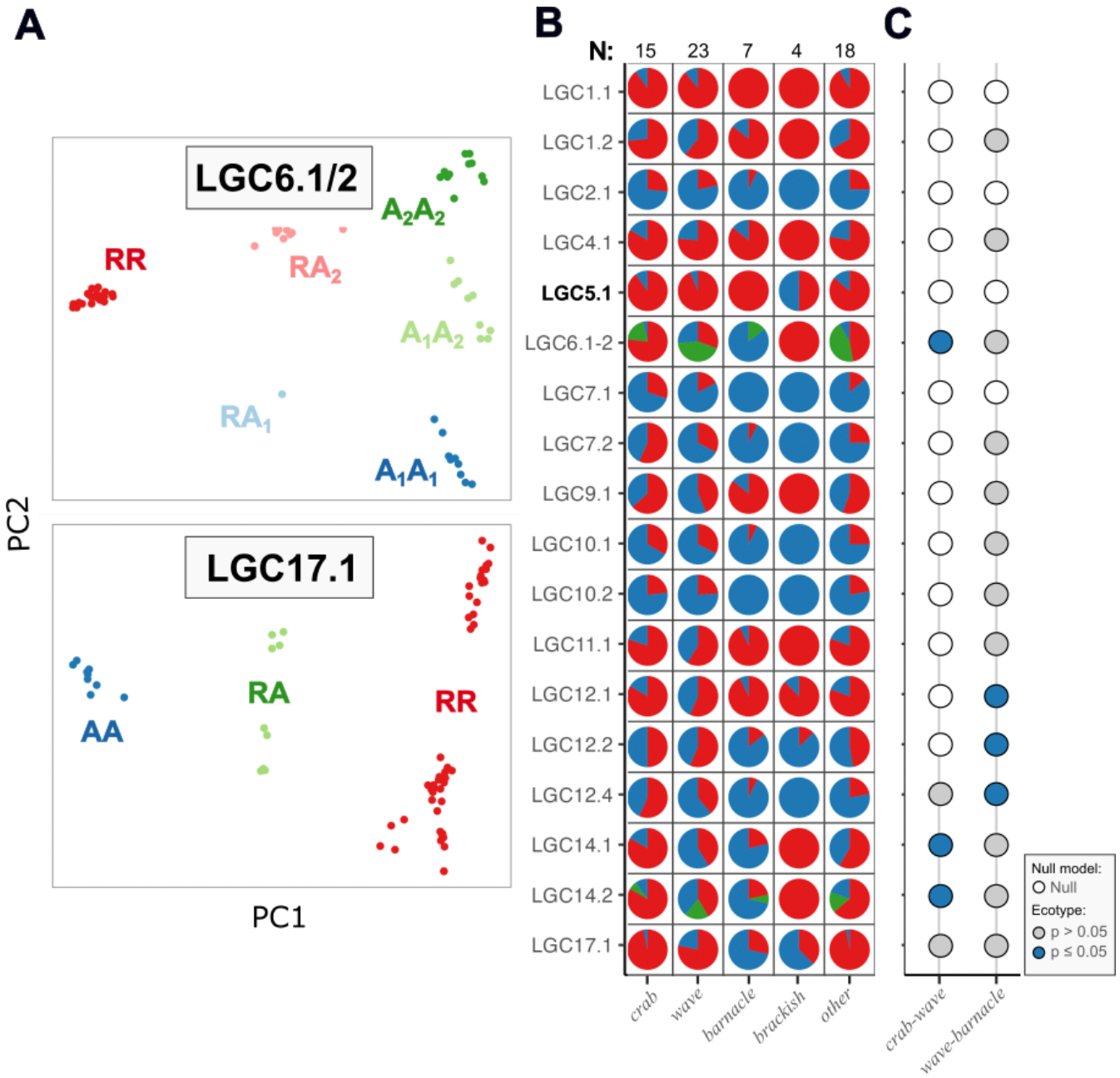

Figure 4: Inversion arrangement frequencies in Northern Littorina saxatilis. **A**) Two examples of PCA plots, coloured by genotype. LGC6.1/2 with six genotypes (complex inversion) and LGC17.1 with three genotypes. 'R' = reference, 'A'= alternative; with 'A' subdivided into 'A_1' = alternative 1 and 'A_2' = alternative 2, the latter corresponding to the third arrangement in LGC6.1/2 (Faria, Chaube, et al., 2019). Plots for other inversions in supplementary material (Figure S19). **B**) Arrangement frequencies for all inversions (new inversion in bold font) with either two or three arrangements. Colours represent the arrangements (R = red; A or A_1 = blue; A_2 = green). Separate pie charts were made for each ecotype. Sample sizes for each ecotype are shown above the panel. **C**) Logistic regression results grouped by ecotype contrast. Fill of circles represent the best model evaluated by AIC: empty = null model and filled = ecotype model. Blue circles were significant for the ecotype effect ($p \leq 0.05$). For these tests A_1 and A_2 were grouped together. Ecotype contrasts subset individuals to locations where both focal ecotypes were sampled. Detailed results are in Table S4 and S6.

Associating inversion frequencies with ecotypes and species

Northern saxatilis ecotype contrasts found that different inversions contribute to divergence between the *crab-wave* and *wave-barnacle* ecotypes. Considering all sites together, 11 of 19 inversions showed clear arrangement frequency differences between the *crab-wave* ecotypes, and 15 of 19 inversions for the *wave* and *barnacle* ecotypes (Figure 4b). However, when a logistic regression analysis was used to account for frequency differences driven by location, only LGC6.1/2, LGC14.1 and LGC14.2 had significant ecotype-related differences for the *crab-wave* contrasts (Figure 4c; Tables S3, S4 and S6). LGC14.1 and LGC14.2 frequencies were better explained by a model containing an interaction term between ecotype and location (Table S4 and Table S6), indicating that the role of these inversions may vary with location. For example, the arrangement frequency differences in LGC14.1 were strong at only two of five sampling locations (Table S3). Meanwhile the three tested inversions on LG12 were significant for *wave-barnacle* contrasts (Figure 4c; Tables S3, S4 and S6), with marginal signals in several other inversions that were lost after correcting for multiple tests.

The species contrasts between *L. saxatilis* ecotypes and *L. arcana* had more signatures of divergence than the *L. saxatilis* ecotype contrasts (Figure 5). Apart from LGC12.1, arrangement frequencies were different between species for all inversions. Breaking this down to *L. saxatilis* ecotypes, 15 of 16 tested inversions were different between *L. arcana* and *wave*, while 13 of 16 inversions were different between *L. arcana* and *crab* (Figure 5b). Logistic regression found five inversions with significant differences in arrangement frequencies among *L. arcana* and the *crab* ecotype, and 11 significant differences among *L. arcana* and the *wave* ecotype (Figure 5c; Tables S5-S6). Three of the significant inversions (LGC1.2, LGC7.1 and LGC10.2) in the *L. arcana-wave* contrasts were better explained by a

model considering the interaction between species and location (Table S5), suggesting that some interspecific differences may be location specific.

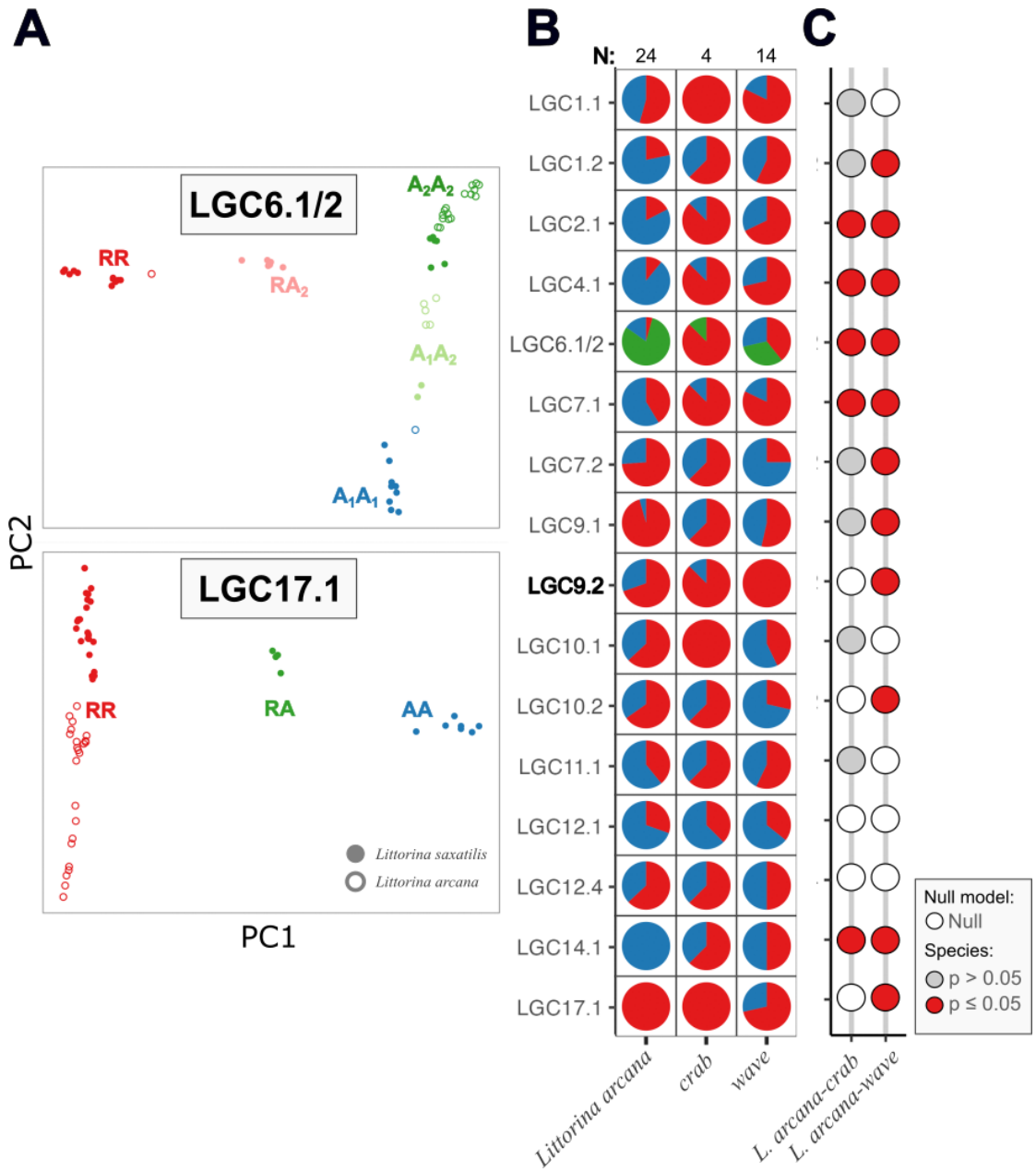

Figure 5: Inversion arrangement frequency comparisons between *Littorina saxatilis* and *L. arcana*. Only locations where both species were collected were included. **A)** Example PCA plots for two inversions coloured by genotype. 'R' = reference, 'A'= alternative; with 'A' subdivided into 'A_1' = alternative 1 and 'A_2' = alternative 2, the latter corresponding to the third arrangement in LGC6.1/2 (Faria, Chaube, et al., 2019). Plots for other inversions in supplementary material (Figure S20). **B)** Arrangement frequencies for all inversions (new inversion in bold font) with either two or three arrangements. Colours represent the arrangements (R = red; A or A_1 = blue; A_2 = green). Separate pie charts have been made for *L. arcana* and two *L. saxatilis* ecotypes. Sample sizes for each group are shown above the panel. **C**) Logistic regression results grouped by species contrast. Fill of circles represent the best model evaluated by AIC: empty = null model and filled = species model. Red circles were significant for the species effect ($p \leq 0.05$). For these tests A_1 and A_2 were grouped together. Since *L. arcana* and *crab* ecotypes were only found together in one location, the interaction model left out these contrasts. Detailed results are in Table S5-6.

Discussion:

We have shown that several inversions identified in Swedish populations of *L. saxatilis* are distributed across the species range (Figure 2; Tables 1-2). Arrangement frequencies for some of these inversions were different among ecotypes (Figure 4). Since inversions contribute to divergent adaptive phenotypes in Sweden, this suggests that they have a widespread role in parallel ecotype formation (Koch et al., 2021, 2022; Morales et al., 2019; Westram et al., 2021). These inversions are possibly ancient polymorphisms, as several were found to be polymorphic in the sister species *L. arcana* (Figure 5).

By aggregating samples over broad areas, our inversion detection approaches clearly showed that all previously published inversion polymorphisms detected in Sweden for *L. saxatilis* (Hearn et al., 2022; Westram et al., 2018, 2021) are widespread across the species' native range, with two exceptions (Figures 2 and S1-S17; Table 1). One exception was LGC14.3 (Figures 2 and S14; Table 1), which was only weakly supported in the original description (Faria, Chaube, et al., 2019). The other was LGC12.3, which was only detected when males and females were analysed separately in the original description (Hearn et al., 2022). The widespread distribution of most inversions across the species range is surprising considering strong geographic structure within *L. saxatilis*, especially the separation between Iberian and North Atlantic populations that is believed to result from long-term isolation on either side of the Bay of Biscay (Doellman et al., 2011; Morales et al., 2019; Panova et al., 2014; Tirado et al., 2016). Moreover, geographic differences in the PCA plots were smaller than the differences among inversion karyotypes (Figures 3 and S18), consistent with an ancient origin of the inversion polymorphisms and low gene flux between arrangements. Interestingly, geographic structure was typically stronger in one of the arrangements (e.g. LGC17.1; Figure 3) which might be explained by the recent geographic spread of one

arrangement (or a derived haplotype within one arrangement) across the species range, perhaps driven by selection.

We also identified patterns consistent with two additional inversion polymorphisms (LGC5.1 & LGC9.2). However, our detection approaches did not always align. One reason is that the success of the heterozygosity split approach is impacted by the age and frequency of an inversion. Young inversions may be missed if the inverted homokaryotype is rare in a population, since the rise in heterozygosity in heterokaryotypes will be hard to perceive until the inverted arrangement accumulates substitutions and becomes common in a population. Old inversions may be missed when the heterokaryotype is rare, since genetic diversity accumulates over time, restoring heterozygosity in the inverted homokaryotype. On the other hand, uncommon inversions may be easier to detect using the heterozygosity split approach than with PCA because a few individuals can stand out on the "split plot", whereas they would not contribute enough variance to group individuals on PC1. Analyses with simulations and inversions of known age are needed to evaluate the heterozygosity split approach more fully relative to other approaches. Identification of inversions was harder in Iberia, due to a small sample size (n = 12) and a phylogeographic barrier between Iberian sampling sites (Tirado et al., 2016). In addition, inversion frequency differences are stronger between ecotypes in Iberia (Morales et al., 2019) and hybrids are less common (Kess & Boulding, 2019), resulting in fewer heterokaryotypes and making our detection methods less effective. Because our approaches can only infer patterns of variation that are typical of inversions, breakpoint sequencing using long-read sequencing in combination with LD analyses, could be used to validate our observations.

Genotyping of the inversions showed differences in arrangement frequency that could be explained if some of the loci contributing to ecotype formation were present within inversions. Around 35% of the tested inversions (7 of 19) had significant differences in

arrangement frequency among *crab-wave* or *wave-barnacle* ecotypes in Northern saxatilis (Figure 4; Table S4 and S6), indicating that several inversions relate to ecotype formation. The general pattern was that differences between the former ecotype contrast were only associated with a limited set of inversions, LGC6.1/2 and LGC14.1/2, which matched some of the previously published *crab-wave* candidate regions (Koch et al., 2022; Westram et al., 2018, 2021), while *wave-barnacle* differences were located on LG12. These *wave-barnacle* differences may relate to shore height, as LG12 has previously been associated with shore height gradients (Morales et al., 2019). Alternatively, this might relate to sex, since LGC12.2 is likely to contain a sex-determining locus (Hearn et al., 2022), and all *barnacle* samples were taken from females while *wave* samples were composed of both males and females. However, the sex-determining locus was only associated with LGC12.2 in *crab* snails, and sex does not explain why LGC12.1 and LGC12.4 showed *wave-barnacle* differences.

These results should be interpreted with some caution. To reduce noise from geographic variation, samples were filtered to only sites where both ecotypes (or species) were collected. These filters mean that the numbers of both snails and sampling locations varied among the different tests (Table S3). However, the handful of investigated locations still covered a large geographic area around North-West Europe, which is compatible with the role of inversions in parallel ecotype formation and the inferences of a previous study (Morales et al., 2019). Despite this, LGC14.1/2 was better supported by the model containing an interaction term between the *crab* and *wave* ecotypes, and location (Table S6), suggesting that the same inversion arrangement may carry different sets of adaptive alleles for different habitats in different populations. Morales et al. (2019) showed that clusters of F_{ST} outliers for several inversions were restricted to Spain and Sweden. Even within a small part of the Swedish coast, associations between inversions, traits and habitat variables are not fully consistent, supporting the hypothesis that an inversion's allelic content can vary among

locations (Koch et al., 2022). Overall, this suggests that, although two inversions contribute to parallel phenotypic divergence between *crab* and *wave* ecotypes, their genetic basis may differ across locations (Faria, Johannesson, et al., 2019).

An alternative but non-mutually exclusive hypothesis is that the selective pressures or habitats differ among locations, reducing the parallelism among environmental contrasts (Bolnick et al., 2018). This is likely if *crab-wave* divergence is also influenced by other environmental features. Shore-height gradients are a prime candidate (Morales et al., 2019). In Spain and the UK, the *crab* and *wave* ecotypes are differentiated by their height on shore (Butlin et al., 2008) leading to an additional adaptive gradient that is conflated with the *crab-wave* gradient, in different directions in the two regions (Morales et al., 2019). Hypothetically, the *wave-barnacle* ecotype contrast should also be affected by a shore-height gradient, as *barnacle* snails typically inhabit lower tidal zones (Reid, 1996, p. 315). In fact, one inversion (LGC12.2) that shows significant differences between *wave* and *barnacle* ecotypes in our study (Figure 4) overlaps with shore-height candidates from Morales et al. (2019). However, other inversions involved in shore-height detected by Morales et al. (2019) were not involved in *wave* and *barnacle* divergence in this study, possibly because differences between the *wave* and *barnacle* ecotypes are more multifaceted than shore-height alone.

Most *L. saxatilis* inversions were also polymorphic in its sister species *L. arcana* (Tables 1-2; Figure S18). Arrangements typically diverged between the species, with one arrangement splitting off from the *L. saxatilis* cluster on the second PC axis. We also saw suggestive evidence of inversion polymorphisms in *L. compressa* (Table 2; Figures 3 and S18) but did not follow up on these observations because of the limited sample size making it difficult to implement our other analyses on this species. The existence of shared inversion polymorphisms among species suggests that the arrangements originated before the species

split (1.7-0.06 My; Reid et al., 2012). Alternative explanations are that arrangements introgressed between species, or inversions evolved repeatedly at the same positions. The repeated evolution scenario has been suggested in humans (Carvalho & Lupski, 2016; Flores et al., 2007), *Drosophila* (Ranz et al., 2007) and deer mice (Harringmeyer & Hoekstra, 2022). However repeated evolution of inversions is unlikely to apply to these *Littorina* species, since the separation of arrangements across species on the same PC axis suggests that they share part of their divergence history, as expected for ancestral polymorphism. It is more challenging to disentangle the patterns of introgression from co-inherited ancestral variation. Hypothetically, an introgressed arrangement should be less diverged than a shared ancestral arrangement, as the arrangement was more recently shared between species (Fuller et al., 2018; Jay et al., 2018). However, in practice, PCA is a poor guide to seeing divergence differences among arrangements. Furthermore, gene flow has been estimated to be extremely infrequent between the two *Littorina* species (Stankowski et al., 2020). Thus, until introgression and shared ancestry can be properly investigated with tree-based and demographic analyses, the most parsimonious scenario is that the inversions were co-inherited from a common ancestor. Assuming common ancestry, these inversions likely appeared before the expansion of *L. saxatilis* after the last glacial period (estimated as 0.37 My; Panova et al., 2014). Many inversions in other species appear to be much older than this (Barrón et al., 2019; Coughlan & Willis, 2019; Fuller et al., 2018; Jay et al., 2018; MacGuigan et al., 2022; Todesco et al., 2020), likely maintained under balancing selection preventing fixation or loss in any specific location (Durmaz et al., 2020; Wellenreuther & Bernatchez, 2018). Whether due to shared ancestry or introgression, polymorphic inversions may have provided a ready source of genetic diversity when the snails expanded into new habitats after glacial retreat (Faria, Johannesson, et al., 2019).

Contrary to our predictions, we detected that *L. arcana* arrangement frequencies were often closer to *crab* ecotype *L. saxatilis* than *wave* ecotype, despite the lower sample sizes (Figure 5; Tables S5 and S6). *L. arcana* typically inhabits the more wave-exposed parts of *L saxatilis'* shore distribution (Reid, 1996). Thus rare gene flow events are expected to be more common in the wave habitat, and selection more similar for *wave* ecotype adapted alleles or inversion arrangements. Therefore, the arrangement frequency differences we observed suggest either that additional habitat variables influence inversion frequencies or that the sets of adaptive alleles carried by each arrangement vary among species.

Parallel ecotype formation may often be underpinned by polymorphic inversions. With our inversion detection approaches we have shown that Swedish inversions in *Littorina saxatilis* are widespread, with some consistently differentiating ecotypes. The majority were also detected in *L. arcana*, suggesting that they are ancient polymorphisms that could be maintained by balancing selection. These approaches could be applied to the vast majority of other species which have fragmented reference genomes. Overall, our detection and genotyping demonstrate the important role that inversions play in diversification of *Littorina saxatilis* and other closely related species.

Acknowledgments:

We would like to thank members of the *Littorina* team for their advice and feedback during this project. In particular, we thank Alan Le Moan who inspired us to look at heterozygosity differences to identify inversions and Katherine Hearn for help with the PCA scripts. We thank Edinburgh Genomics for library preparation and sequencing. Sample collections, sequencing and data preparation were supported by the European Research Council (ERC-2015-AdG-693030- BARRIERS) and Natural Environment Research Council (NE/P001610/1). The analysis was supported by the Swedish Research Council (*vetenskaprådet;* 2018-03695_VR) and the Portuguese Foundation for Science and Technology (Fundación para a Ciência e Tecnologia) through a research project (PTDC/BIA-EVL/1614/2021) and CEEC contract (2020.00275.CEECIND).

Author contributions:

JR, RF & RB conceived this project. SS collated and processed the WGS data. EK constructed the new version of the linkage map. JR conducted the analysis and wrote this paper with the support of RB & RF. All authors reviewed and approved the draft of this paper before submission.

Data Accessibility

All raw sequence data were uploaded to the NCBI short read archive (PRJNA626520). All scripts used for this project are uploaded to GitHub (https://github.com/ja-Reeve/Littorina_inversion_identification.git).

References:

Ayala, D., Acevedo, P., Pombi, M., Dia, I., Boccolini, D., Costantini, C., Simard, F., & Fontenille, D. (2017). Chromosome inversions and ecological plasticity in the main African malaria mosquitoes. *Evolution*, *71*(3), 686–701. https://doi.org/10.1111/evo.13176

Ayala, D., Ullastres, A., & González, J. (2014). Adaptation through chromosomal inversions in Anopheles. *Frontiers in Genetics*, *5*(May), 1–10. https://doi.org/10.3389/fgene.2014.00129

Barrón, M. G., Paupy, C., Rahola, N., Akone-ella, O., Ngangue, M. F., Wilson-bahun, T. A., Pombi, M., Kengne, P., Costantini, C., Simard, F., González, J., & Ayala, D. (2019). A new species in the major malaria vector complex sheds light on reticulated species evolution. *Scienctific Reports*, *9*, 14753. https://doi.org/10.1038/s41598-019-49065-5

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bolnick, D. I., Barrett, R. D. H., Oke, K. B., Rennison, D. J., & Stuart, Y. E. (2018). (Non)parallel evolution. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 303–330. https://doi.org/10.1146/annurev-ecolsys-110617-062240

Boulding, E. G., Rivas, M. J., González-Lavín, N., Rolán-Alvarez, E., & Galindo, J. (2017). Size selection by a gape-limited predator of a marine snail: Insights into magic traits for speciation. *Ecology and Evolution*, *7*(2), 674–688. https://doi.org/10.1002/ece3.2659

Butlin, R. K. (2005). Recombination and speciation. *Molecular Ecology*, *14*, 2621–2635. https://doi.org/10.1111/j.1365-294X.2005.02617.x

Butlin, R. K., Galindo, J., & Grahame, J. W. (2008). Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 2997–3007. https://doi.org/10.1098/rstb.2008.0076

Carvalho, C. M. B., & Lupski, J. R. (2016). Mechanisms underlying structural variant formation in genomic disorders. *Nature Reviews Genetics*, *17*(4), 224–238. https://doi.org/10.1038/nrg.2015.25

Coughlan, J. M., & Willis, J. H. (2019). Dissecting the role of a large chromosomal inversion in life history divergence throughout the Mimulus guttatus species complex. *Molecular Ecology*, *28*, 1343–1357. https://doi.org/10.1111/mec.14804

Coyne, J. A., Aulard, S., & Berry, A. (1991). Lack of underdominance in a naturally occurring pericentric inversion in Drosophila melanogaster and its implications for chromosome evolution. Genetics, 129, 791-802. https://doi.org/10.1093/genetics/129.3.791

Endelman, J. B., & Plomion, C. (2014). LPmerge: an R package for merging genetic maps by linear programming. Bioinformatics, 30(11), 1623-1624. https://doi.org/10.1093/bioinformatics/btu091

Doellman, M. M., Trussell, G. C., Grahame, J. W., & Vollmer, S. V. (2011). Phylogeographic

analysis reveals a deep lineage split within North Atlantic Littorina saxatilis. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1722), 3175–3183. https://doi.org/10.1098/rspb.2011.0346

Durmaz, E., Kerdaffrec, E., Katsianis, G., Kapun, M., & Flatt, T. (2020). How selection acts on chromosomal inversions. *ELS*, *1*, 307–315. https://doi.org/10.1002/9780470015902.a0028745

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

Faria, R., Johannesson, K., Butlin, R. K., & Westram, A. M. (2019). Evolving inversions. *Trends in Ecology and Evolution*, *34*(3), 239–248. https://doi.org/10.1016/j.tree.2018.12.005

Faria, R., & Navarro, A. (2010). Chromosomal speciation revisited: Rearranging theory with pieces of evidence. *Trends in Ecology and Evolution*, *25*(11), 660–669. https://doi.org/10.1016/j.tree.2010.07.008

Flores, M., Morales, L., Gonzaga-Jauregui, C., Domínguez-Vidaña, R., Zepeda, C., Yañez, O., Gutiérrez, M., Lemus, T., Valle, D., Avila, M. C., Blanco, D., Medina-Ruiz, S., Meza, K., Ayala, E., García, D., Bustos, P., González, V., Girard, L., Tusie-Luna, T., … Palacios, R. (2007). Recurrent DNA inversion rearrangements in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(15), 6099–6106. https://doi.org/10.1073/pnas.0701631104

Fuller, Z. L., Leonard, C. J., Young, R. E., Schaeffer, S. W., & Phadnis, N. (2018). Ancestral polymorphisms explain the role of chromosomal inversions in speciation. *PLoS Genetics*, *14*(7), e1007526. https://doi.org/10.1371/journal.pgen.1007526

Hanlon, V. C. T., Lansdorp, P. M., & Guryev, V. (2022). A survey of current methods to detect and genotype inversions. *Human Mutation*, *43*(11), 1576–1589. https://doi.org/10.1002/humu.24458

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

Hearn, K. E., Johannesson, K., Koch, E. L., Stankowski, S., Butlin, R. K., Faria, R., & Westram, A. M. (2022). Differing associations between sex determination and sex-linked inversions in two ecotypes of Littorina saxatilis. *Evolution Letters*. https://doi.org/10.1002/evl3.295

Hoffmann, A. A., & Rieseberg, L. H. (2008). Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics*, *39*, 21–42.

Jackson, B. C. (2011). Recombination-suppression: how many mechanisms for chromosomal speciation? *Genetica*, *139*(3), 393–402. https://doi.org/10.1007/s10709-011-9558-0

Janson, K. (1982). Phenotypic differentiation in littorina saxatilis olivi (mollusca, prosobranchia) in a small area on the swedish west coast. *Journal of Molluscan Studies*, *48*(2), 167–173.

https://doi.org/10.1093/oxfordjournals.mollus.a065633

Jay, P., Whibley, A., Frézal, L., Rodríguez de Cara, M. Á., Nowell, R. W., Mallet, J., Dasmahapatra, K. K., & Joron, M. (2018). Supergene evolution triggered by the introgression of a chromosomal inversion. *Current Biology*, *28*, 1839–1845. https://doi.org/10.1016/j.cub.2018.04.072

Johannesson, B. (1986). Shell morphology of Littorina saxatilis Olivi: The relative importance of physical factors and predation. *Journal of Experimental Marine Biology and Ecology*, *102*(2–3), 183–195. https://doi.org/10.1016/0022-0981(86)90175-9

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., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), 55–61. https://doi.org/10.1038/nature10944

Kapun, M., & Flatt, T. (2019). The adaptive significance of chromosomal inversion polymorphisms in Drosophila melanogaster. *Molecular Ecology*, *28*(6), 1263–1282. https://doi.org/10.1111/mec.14871

Kemppainen, P., Knight, C. G., Sarma, D. K., Hlaing, T., Prakash, A., Maung Maung, Y. N., Somboon, P., Mahanta, J., & Walton, C. (2015). Linkage disequilibrium network analysis (LDna) gives a global view of chromosomal inversions, local adaptation and geographic structure. *Molecular Ecology Resources*, *15*(5), 1031–1045. https://doi.org/10.1111/1755-0998.12369

Kess, T., & Boulding, E. G. (2019). Genome-wide association analyses reveal polygenic genomic architecture underlying divergent shell morphology in Spanish Littorina saxatilis ecotypes. *Ecology and Evolution*, *9*(17), 9427–9441. https://doi.org/10.1002/ece3.5378

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*. https://doi.org/10.1111/evo.14602

Korunes, K. L., & Noor, M. A. F., (2017). Gene conversion and linkage: effects on genome evolution and speciation. Molecular Ecology, 26, 351-364. https://doi.org/10.1111/mec.13736

Larsson, J. (2021). *Understanding the evolution of shell shape in snails*. University of Sheffield.

Le Pennec, G., Butlin, R. K., Jonsson, P. R., Larsson, A. I., Lindborg, J., Bergström, E., Westram, A. M., & Johannesson, K. (2017). Adaptation to dislodgement risk on wave- swept rocky shores in the snail Littorina saxatilis. *PLoS ONE*, *12*(10), e0186901. https://doi.org/10.1371/journal.pone.0186901

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

Li, H., & Ralph, P. (2019). Local PCA shows how the effect of population structure differs along the genome. *Genetics*, *211*(January), 289–304.

MacGuigan, D. J., Krabbenhoft, T. J., Harrington, R. C., Wainwright, D. K., Backenstose, N. J. C., & Near, T. J. (2022). Lacustrine speciation associated with chromosomal inversion in a lineage of riverine fishes. *BioRxiv*. https://doi.org/10.1101/2022.12.12.519811

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 and 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*, 1297–1303. https://doi.org/10.1101/gr.107524.110.20

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

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*, eaav9963. https://doi.org/10.1126/sciadv.aav9963

Nowling, R. J., Manked, K. R., & Emrich, S. J. (2020). Detecting inversions with PCA in the presence of population structure. *PLoS ONE*, *15*(10), e0240429. https://doi.org/10.1371/journal.pone.0240429

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. In *Methods in Molecular Biology* (Vol. 1452). https://doi.org/10.1007/978-1-4939-3774-5_2

Panova, M., Johansson, T., Canbäck, B., Bentzer, J., Rosenblad, M. A., Johannesson, K., Tunlid, A., & André, C. (2014). Species and gene divergence in Littorina snails detected by array comparative genomic hybridization. *BMC Genomics*, *15*, 687.

Ranz, J. M., Maurin, D., Chan, Y. S., von Grotthuss, M., Hillier, L. D. W., Roote, J., Ashburner, M., & Bergman, C. M. (2007). Principles of genome evolution in the Drosophila melanogaster

species group. *PLoS Biology*, *5*(6), 1366–1381. https://doi.org/10.1371/journal.pbio.0050152

Rastas, P., Calboli, F. C. F., Guo, B., Shikano, T., & Merilä, J. (2016). Construction of ultradense linkage maps with Lep-MAP2: stickleback F2 recombinant crosses as an example. Genome Biology and Evolution, 8(1), 8-93. https://doi.org/10.1093/gbe/evv250

Rastas, P. (2017). Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. Bioinformatics, 33(23), 3726-3732. https://doi.org/10.1093/bioinformatics/btx494

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

Reid, D. G., Dyal, P., & Williams, S. T. (2012). A global molecular phylogeny of 147 periwinkle species (Gastropoda, Littorininae). *Zoologica Scripta*, *41*(2), 125–136. https://doi.org/10.1111/j.1463-6409.2011.00505.x

Stankowski, S., Westram, A. M., Zagrodzka, Z. B., Eyres, I., Broquet, T., Johannesson, K., & Butlin, R. K. (2020). The evolution of strong reproductive isolation between sympatric intertidal snails: Strong RI in Littorinids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1806). https://doi.org/10.1098/rstb.2019.0545

Stankowski, S., Zagrodzka, Z. B., Galindo, J., Montano-Rendón, M., Faria, R., Mikhailova, N., Blakeslee, A. M. H., Arnason, E., Broquet, T., Morales, H. E., Grahame, J. W., Westram, A. M., Johannesson, K., & Butlin, R. K. (2023). Whole-genome phylogeography of the intertidal snail Littorina saxatilis. *Evolutionary Journal of the Linnean Society, Online Early Access*. https://doi.org/10.1093/evolinnean/kzad002

Stankowski, S., Zagrodzka, Z. B., Garlovsky, M. D., Pal, A., Castillo, D. G., Moan, A. Le, Leder, E., Reeve, J., Westram, A. M., Butlin, R. K., Biology, E., & Roscoff, S. B. De. (2023). Selection on many loci drove the origin and spread of a key innovation. *BioRxiv*, 1–12. https://doi.org/10.1101/2023.02.13.528213

Tirado, T., Saura, M., Rolán-Alvarez, E., & Quesada, H. (2016). Historical biogeography of the marine snail littorina saxatilis inferred from haplotype and shell morphology evolution in NW Spain. *PLoS ONE*, *11*(8), 1–13. https://doi.org/10.1371/journal.pone.0161287

Todesco, M., Owens, G. L., Bercovich, N., Légaré, J., Soudi, S., Burge, D. O., Huang, K., Ostevik, K. L., Drummond, E. B. M., Imerovski, I., Lande, K., Pascual-Robles, M. A., Nanavati, M., Jahani, M., Cheung, W., Staton, S. E., Muños, S., Nielsen, R., Donovan, L. A., … Rieseberg, L. H. (2020). Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature*, *584*, 602–607. https://doi.org/10.1038/s41586-020-2467-6

Wellenreuther, M., & Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology and Evolution*, *33*(6), 427–440. https://doi.org/10.1016/j.tree.2018.04.002

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

Westram, A. M., Faria, R., Johannesson, K., Butlin, R., & Barton, N. (2022). Inversions and parallel

evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1856). https://doi.org/10.1098/rstb.2021.0203

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

Supplementary method 1: consensus map construction

Two versions of a *Littorina saxatilis* linkage map currently exist. The older version was created from a cross of two *crab* ecotype snails and sequenced with 40,000 capture probes (Westram et al. 2018, Supplementary methods). Henceforth, it will be referred to as the 'crab map'. The younger version was assembled from 13 F2 families from two crosses of *wave* x *crab* parents (Koch et al. 2021). This later map will be called the 'crab x wave map'. Both maps were assembled with Lep-MAP (*crab map* = Lep-MAP2 (Rastas et al., 2016); *crab x wave map* = Lep-MAP3 (Rastas 2017)).

Both maps have a problem with ordering markers in some inverted regions (defined in Westram et al. 2021). The parental genotypes were unknown in both maps prior to crossing. If one of the parents was heterozygous for an inversion, map distances were compressed within this inversion, as recombination in inverted regions is supressed when a parent was a heterkaryotype (Coyne et al., 1991). Map compression was very noticeable on LG10 and LG14, as they are almost entirely covered by inversions (Faria et al., 2019; Westram et al., 2021). To improve the resolution of distances in inversions we merged information across maps to generate a more informative 'consensus map'.

The *consensus map* was built by merging the *crab map* and *crab x wave map* using the R package 'LPmerge' (Endelman & Plomion 2014). Before merging markers in inverted regions were filtered to retain those which had the greatest maps distances, and thus the least recombination suppression. LG10 and LG14 had lower marker density and large gaps in the consensus, therefore only *crab map* markers were used for these linkage groups. The average marker position was calculated for each contig. Any markers >2cM from their contig average were removed and a new average was calculated from the remaining markers. These average positions were defined as the *consensus map* positions.

Supplementary method 2: manual adjustments to inversion genotyping

K-means clustering could not always identify a number of clusters consistent with the patterns of simple (K = 3) or complex (K = 6) inversions. Often this was because some aspect of geographic distance influenced the variance explained on the first principal component. Each inversion in the Northern saxatilis group and the *arcana-saxatilis* comparison, was visually inspected after K-means clustering. Any inversion where the best K ≠ 3 or 6, or where the clustering seemed to be distorted by geographic variation, was manually adjusted. The problems and manual adjustments for each inversion are described below. For step-by-step details see the code at https://github.com/ja-Reeve/Littorina_inversion_identification/tree/main/R_scripts/6.5_manual_edits_of_inversion_genotype.R

Adjustments for Northern saxatilis:

LGC1.1: K-means found only two clusters because clustering was slightly diagonal on the PCA plot (Figure S19). Since K-means only used PC1 scores for most inversions, the heterokaryotypic cluster had to be demarcated visually. Arrangements were manually adjusted with thresholds 'AA' = PC1 < -40, 'RA' = -40 < PC1 < -10, and 'RR' = PC1 > -10.

LGC2.1: a small clump of four American samples falls between the 'AA' and 'RA' clusters down the bottom of the PCA plot (Figure S19). One sample, *York_B-1_Ls*, seemed to belong to the 'RA' cluster while the others belong to the 'AA' cluster. Removing these samples then projecting them back onto the PC axes confirmed this visual assessment.

LGC9.2: the gap between the 'RA' and 'AA' clusters is unclear (Figure S22). Most of spread on PC2 was from snails sampled from the North Sea. Looking at this plot by region, only samples collected from the Celtic Sea appear to have a pattern consistent with an inversion polymorphism. To verify, the largest group, North Sea samples, was removed resulting in four clusters. PC2 explained the variation in the Celtic Sea while PC1 distinguished the American samples (Figure S22). Considering that the variance explained along PC1 was low and reflects geographic separation, we deemed that the support for this inversion was too insubstantial to include it in the association tests.

Adjustments for the arcana-saxatilis *comparisons:*

Note: *L. arcana* and *L. saxatilis* were filtered to keep only samples from locations where both species were collected. For brevity, we will just use the species name in the following section.

LGC4.1: *L. saxatilis* forms three clear clusters, which did not align well with *L. arcana* (Figure S21). To improve the alignment, *L. arcana* samples were projected onto a PCA of *L. saxatilis*. Projected *L. arcana* clustered with the right arrangement in *L. saxatilis* (Figure S22). Interesting, there were two clusters within *L. arcana*, with one cluster falling between the right and central *L. saxatilis* clusters. These centralized points were not associated with any sampling site or specific country, and the average heterozygosity was higher than other *L. arcana*. Based on this information, *L. arcana* samples were labelled using the following thresholds: 'RA' = 15 < PC1 < 30 & PC2 > -5, all others were 'AA'.

LGC5.1: *L. arcana* clusters separate on PC2 while *L. saxatilis* clusters are unclear (Figure S22). PC1 only explains a small proportion of the variance among samples in the Northern saxatilis PCA (Figure S20), suggesting that this inversion may have a weak pattern relative to the differences among species. *L. arcana* samples were projected onto a PCA of *L. saxatilis*, and vice versa (Figure S22). When *L. saxatilis* was projected onto *L. arcana*'s PC axes, *L. saxatilis* becomes a small cluster. The reverse occurs for the inverse projection. Since *L. arcana* aligned poorly with both the 'RR' and 'AA' clusters, it is hard to assign it. Thus, this inversion was not assessed in the species comparisons.

LGC6.1/2: five of the six clusters diagnostic of a complex double inversion are present in the PCA plot of both species (Figure S21), but the clusters in *L. arcana* fall between *L. saxatilis* clusters. Projecting *L. arcana* onto the *L. saxatilis* PC axes, aligned the cluster better between species (Figure S22). One *L. arcana* cluster, which comprised a single outlying sample, still falls between *L. saxatilis* clusters. After looking at average heterozygosity, this sample was assigned to the 'RR' cluster. Arrangements for both species were assigned with the following thresholds: 'RR' = PC1 < -30; 'RA_1' not found; 'RA_2' = -30 < PC1 < 0; 'A_1A_1' = PC2 < -30; 'A_1A_2' = -30 < PC2 < 0; 'A_2A_2' = PC1 > 12 & PC2 > 5.

LGC14.1: similar to LGC4.1, *L. saxatilis* had three clear clusters, while *L. arcana* formed a cloud of points (Figure S21). Projecting *L. arcana* onto *L. saxatilis* PCA axes showed that there was only one cluster for *L. arcana*, which aligned with 'AA' (Figure S22).

LGC14.2: *L. saxatilis* had six clusters and *L. arcana* formed a separate group along the right edge of the PCA plot (Figure S22). This pattern suggests the double inversion is fixed for one arrangement in *L. arcana*. There were no clear clusters in a PCA plot of just *L. arcana*. Instead, samples seemed to be grouped by geographic region. Projecting *L. arcana* onto *L. saxatilis* PCA axes did not improve the clustering, as *L. arcana* still formed a continuous group between two *L. saxatilis* clusters (Figure S22). Further testing of SNPs in the boundary regions also showed similar patterns. Given this complexity and the lack of a clear resolution with these data, LGC14.2 was not assessed in the species comparisons.

Supplementary tables:

Table S1: Snail sampling details. N_{sax} , N_{arc}, and N_{comp} are the number of *Littorina saxatilis*, *L. arcana* and *L. compressa* collected. '-' indicates locations outside the species range of *L. arcana* or *L. compressa*. Ramsö and Arsklåvet were called CZA and CZD in Westram et al. 2021. Collector names are abbreviated in the table.

MMR = Mauricio Montaño-Rendón; Sheffield Chlidren's NHS Foundation Trust, UK
AMW = Anja Marie Westram; Nord University, Norway
KJ = Kerstin Johanesson; University of Gothenburg, Sweden
RB = Roger Butlin; University of Sheffield, UK & University of Gothenburg, Sweden
SS = Sean Stankowski; ISTA, Austria
JL = Jenny Larsson; University of Gothenburg & Chalmers University, Sweden
ZZ = Zuzanna Zagrodzka; University of Sheffield, UK
EA = Einar Árnason; University of Iceland
TB = Tomas Broquet; Station Biologique de Roscoff, France
PK = Petri Kemppainen; University of Helsinki, Finland
NM = Natalia Mikhailova; St. Petersburg State University, Russia
AB = April Blakeslee; East Carolina University, USA
JG = Juan Galindo; University of Vigo, Spain

Location	Country	Latitude	Longitude	N_{sax}	N_{arc}	N_{comp}	Collector
*Northern samples (*L. saxatilis *= 67;* L.arcana *= 24;* L. compressa *= 4):*							
Amble	England (UK)	55.33578	-1.56953	0	1	0	MMR
Arsklåvet	Sweden	58.83091	11.13305	4	-	-	AMW, KJ, RB
Broad Haven	Wales (UK)	51.60891	-4.91878	1	1	0	MMR
Ceann Tra	Ireland	52.13205	-10.36071	2	0	1	MMR
Dersingham	England (UK)	52.86750	0.44738	2	-	-	MMR
Holyhead	Wales (UK)	53.29981	-4.67967	12	4	1	SS, JL, ZZ
Laugarnes	Iceland	64.15250	-21.88383	2	-	-	EA
Oban	Scotland (UK)	56.42207	-5.48392	1	0	0	MMR
Port Saint Mary	Isle of Mann	54.07602	-4.73618	1	0	0	MMR
Ramsö	Sweden	58.82438	11.06258	4	-	-	AMW, KJ, RB
Ravenscar	England (UK)	54.41036	-0.49196	7	7	-	SS, AMW, ZZ
Roscoff	Brittany (France)	48.69481	-4.10734	6	4	1	TB
Saint Abbs	Scotland (UK)	55.89968	-2.13004	2	2	-	MMR
Thornwick	England (UK)	54.13267	-0.11503	8	0	-	RB
Tjärnö	Sweden	58.88994	11.13866	2	-	-	KJ
Trondheim Fjord	Norway	63.55228	10.46486	3	1	0	PK
Varanger Fjord	Norway	70.04039	29.58401	4	4	1	NM
White Sea	Russia	66.33082	33.06251	2	-	-	NM
York	Maine (USA)	43.15093	-70.62182	4	-	-	AB
*Iberian samples (*L. saxatilis *= 12):*							
Burella	Spain	43.66556	-7.35782	2	-	-	JG
Centinela	Spain	42.07786	-8.89555	10	-	-	JG

Table S2: Tests for clustering of heterozygosity splits along each linkage group for each genetic group. V_{obs} = observed variation among 3cM overlapping windows. V_{Perm} = average variance of 10,000 random draws of variance (± standard deviation). Obs/Perm = ratio of observed to permuted variance. P = empirical P-values, adjusted by a Bejamini-Hochberg correction. Significant results are coloured in green, given an $\alpha \leq 0.05$.

LG	Genetic Group	V_{Obs}	V_{Perm} ± SD	Obs/ Perm	P-value	LG	V_{Obs}	V_{Perm} ± SD	Obs/ Perm	P-value
LG1	Northern saxatilis	42.99	4.35±1.00	9.88	0.0001	LG10	19.08	4.50±1.69	4.24	0.0001
	Iberian saxatilis	0.88	0.77±0.17	1.14	0.3061		5.50	2.02±0.76	2.72	0.0031
	Littorina arcana	8.01	1.92±0.44	4.18	0.0003		26.67	5.66±2.18	4.71	0.0003
LG2	Northern saxatilis	12.18	3.84±0.93	3.17	0.0001	LG11	17.91	3.66±1.17	4.89	0.0001
	Iberian saxatilis	5.06	1.07±0.25	4.72	0.0004		2.22	0.62±0.19	3.55	0.0004
	Littorina arcana	3.17	1.70±0.40	1.86	0.0077		8.51	1.87±0.59	4.55	0.0003
LG3	Northern saxatilis	5.68	1.98±0.51	2.86	0.0001	LG12	123.10	3.66±1.17	4.89	0.0001
	Iberian saxatilis	0.25	0.39±0.10	0.64	0.9698		9.25	1.96±0.62	4.72	0.0004
	Littorina arcana	0.84	0.78±0.20	1.07	0.4138		36.48	4.68±1.49	7.80	0.0003
LG4	Northern saxatilis	6.58	2.51±0.62	2.62	0.0001	LG13	3.37	1.49±0.43	2.26	0.0010
	Iberian saxatilis	0.42	0.28±0.07	1.49	0.0910		0.11	0.11±0.03	0.99	0.9394
	Littorina arcana	1.79	1.05±0.26	1.70	0.0272		0.60	0.59±0.17	1.02	0.4422
LG5	Northern saxatilis	37.54	4.67±1.32	8.04	0.0001	LG14	34.67	7.43±3.10	4.67	0.0001
	Iberian saxatilis	1.14	0.87±0.24	1.32	0.1784		1.92	1.18±0.49	1.62	0.1193
	Littorina arcana	1.42	1.12±0.31	1.27	0.2028		4.21	2.26±0.94	1.86	0.0797
LG6	Northern saxatilis	74.99	6.29±1.73	11.93	0.0001	LG15	2.23	1.24±0.35	1.80	0.0125
	Iberian saxatilis	9.13	1.33±0.36	6.86	0.0004		0.37	0.41±0.11	0.89	0.7421
	Littorina arcana	2.11	1.56±0.42	1.35	0.1583		1.71	1.28±0.36	1.33	0.1711
LG7	Northern saxatilis	25.33	4.10±1.16	6.18	0.0001	LG16	1.31	1.05±0.29	1.25	0.1682
	Iberian saxatilis	1.18	0.69±0.19	1.70	0.0408		0.21	0.10±0.02	2.12	0.0300
	Littorina arcana	10.52	2.74±0.76	3.84	0.0003		0.70	0.50±0.14	1.39	0.1583
LG8	Northern saxatilis	7.43	2.49±0.74	2.98	0.0001	LG17	21.12	3.19±0.87	6.63	0.0001
	Iberian saxatilis	0.06	0.05±0.01	1.06	0.9624		2.05	0.74±0.21	2.78	0.0007
	Littorina arcana	0.38	0.49±0.14	0.77	0.7982		0.83	0.78±0.22	1.05	0.4201
LG9	Northern saxatilis	98.78	8.42±2.35	11.73	0.0001					
	Iberian saxatilis	2.68	1.24±0.34	2.17	0.0036					
	Littorina arcana	1.96	1.35±0.37	1.45	0.1264					

Table S3: Inversion arrangement counts for different ecotype and species contrasts. Total arrangement counts are written next to the titles.

Crab vs wave (n = 56)

	LGC1.1				LGC1.2				LGC2.1				LGC4.1				LGC5.1				LGC7.1			
Location	*crab*		*wave*		*crab*		*wave*		*crab*		*wave*		*crab*		*wave*		*crab*		*wave*		*crab*		*wave*	
	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A
Arsklåvet	4	0	4	0	2	2	2	2	3	1	0	4	3	1	4	0	4	0	4	0	2	2	0	4
Ceann Tra	2	0	2	0	2	0	2	0	0	2	1	1	2	0	2	0	1	1	1	1	1	1	1	1
Holyhead	8	0	8	4	5	3	8	4	1	7	3	9	7	1	8	4	7	1	11	1	1	7	3	9
Ramsö	2	2	4	0	3	1	3	1	2	2	0	4	2	2	3	1	4	0	4	0	3	1	1	3
Thornwick	7	1	8	0	6	2	5	3	0	8	0	8	8	0	6	2	7	1	8	0	0	8	1	7

	LGC7.2				LGC9.1				LGC10.1				LGC10.2				LGC11.1				LGC12.1			
Arsklåvet	2	2	0	4	3	4	0	2	2	3	0	4	1	3	0	4	4	0	3	1	4	0	2	2
Ceann Tra	1	1	1	1	2	2	0	2	0	2	1	1	0	2	0	2	2	0	2	0	2	0	2	0
Holyhead	5	3	3	9	3	5	3	10	2	0	6	6	5	3	5	7	5	3	7	5	5	3	10	2
Ramsö	3	1	2	2	3	4	0	2	2	3	0	4	1	3	0	4	2	2	2	2	4	0	2	2
Thornwick	4	4	5	3	4	6	2	2	6	8	2	6	0	8	3	5	7	1	4	4	6	2	2	6

	LGC12.2				LGC12.4				LGC14.1				LGC17.1			
Arsklåvet	2	2	2	2	3	1	2	2	4	0	0	4	4	0	3	1
Ceann Tra	0	2	0	2	0	2	0	2	2	0	2	0	2	0	2	0
Holyhead	8	0	6	6	5	3	5	7	5	3	5	7	8	0	9	3
Ramsö	1	3	2	2	4	0	0	4	3	1	2	2	4	0	3	1
Thornwick	3	5	6	2	3	5	2	6	8	0	1	7	8	0	8	0

	LGC6.1/2						LGC14.2					
	R	A_1	A_2	R	A_1	A_2	R	A_1	A_2	R	A_1	A_2
Arsklåvet	3	0	1	0	1	3	4	0	0	0	4	0
Ceann Tra	2	0	0	2	0	0	2	0	0	2	0	0
Holyhead	7	0	1	6	4	2	5	2	1	5	4	3
Ramsö	3	0	1	0	1	3	3	0	1	2	2	0
Thornwick	6	1	1	1	2	5	8	0	0	1	5	2

Wave vs barnacle (n = 30)

	LGC1.1				LGC1.2				LGC2.1				LGC4.1				LGC5.1				LGC7.1			
Location	*barn*		*wave*		*barn*		*wave*		*barn*		*wave*		*barn*		*wave*		*barn*		*wave*		*barn*		*wave*	
	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A
Holyhead	4	0	8	4	4	0	8	4	0	4	3	9	4	0	8	4	4	0	11	1	0	4	3	9
Ravenscar	4	0	10	0	4	0	4	6	0	4	2	8	4	0	7	3	4	0	10	0	0	4	0	10

	LGC7.2				LGC9.1				LGC10.1				LGC10.2				LGC11.1				LGC12.1			
Holyhead	0	4	3	9	4	4	0	10	2	4	6	6	0	4	5	7	4	0	7	5	4	0	10	2
Ravenscar	0	4	4	6	3	4	0	2	8	4	4	6	0	4	3	7	4	0	6	4	4	0	2	8

	LGC12.2				LGC12.4				LGC14.1				LGC17.1			
Holyhead	0	4	6	6	0	4	5	7	0	4	5	7	2	2	9	3
Ravenscar	0	4	9	1	0	4	6	4	0	4	4	6	0	4	8	2

	LGC6.1/2						LGC14.2					
	R	A_1	A_2	R	A_1	A_2	R	A_1	A_2	R	A_1	A_2
Holyhead	0	4	0	6	4	2	0	4	0	5	4	3
Ravenscar	0	4	0	2	4	4	0	4	0	4	3	3

Crab vs *L. arcana* (n = 16)

	LGC1.1				LGC1.2				LGC2.1				LGC4.1				LGC7.1				LGC7.2			
Location	*crab*		*arcana*		*crab*		*arcana*		*crab*		*arcana*		*crab*		*arcana*		*crab*		*arcana*		*crab*		*arcana*	
	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A
Holyhead	8	0	5	3	5	3	1	7	7	1	1	7	7	1	0	8	7	1	0	8	5	3	8	0

	LGC9.1				LGC9.2				LGC10.1				LGC10.2				LGC11.1				LGC12.1			
Holyhead	5	3	8	0	7	1	6	2	8	0	5	3	5	3	7	1	5	3	1	7	3	5	3	5

	LGC12.4				LGC14.1				LGC17.1				LGC6.1/2					
													R	A_1	A_2	R	A_1	A_2
Holyhead	5	3	7	1	5	3	0	8	8	0	8	0	7	0	1	0	1	7

Wave vs *L. arcana* (n = 56)

	LGC1.1				LGC1.2				LGC2.1				LGC4.1				LGC7.1				LGC7.2			
Location	*wave*		*arcana*		*wave*		*arcana*		*wave*		*arcana*		*wave*		*arcana*		*wave*		*arcana*		*wave*		*arcana*	
	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A	R	A
Broad Haven	2	0	1	1	0	2	1	1	0	2	0	2	1	1	0	2	1	1	2	0	0	2	0	2
Holyhead	8	4	5	3	8	4	1	7	9	3	1	7	8	4	0	8	9	3	0	8	3	9	8	0
Ravenscar	10	0	14	0	4	6	3	11	8	2	7	7	7	3	3	11	10	0	11	3	4	6	11	3
St. Abbs	3	1	2	2	4	0	0	4	2	2	0	4	4	0	1	3	3	1	1	3	0	4	3	1

	LGC9.1				LGC9.2				LGC10.1				LGC10.2				LGC11.1				LGC12.1			
Broad Haven	1	1	2	0	2	0	2	0	1	1	2	0	0	2	2	0	0	2	0	2	0	2	0	2
Holyhead	9	3	8	0	12	0	6	2	6	6	5	3	5	7	7	1	7	5	1	7	2	10	3	5
Ravenscar	5	5	13	1	10	0	14	0	4	6	5	9	3	7	3	11	6	4	9	5	8	2	4	10
St. Abbs	0	4	3	1	4	0	0	4	1	3	4	0	0	4	4	0	3	1	1	3	0	4	2	2

	LGC12.4				LGC14.1				LGC17.1				LGC6.1/2					
													R	A_1	A_2	R	A_1	A_2
Broad Haven	2	0	0	2	1	1	0	2	2	0	2	0	1	0	1	0	1	1
Holyhead	5	7	7	1	5	7	0	8	9	3	8	0	6	4	2	0	1	7
Ravenscar	6	4	7	7	4	6	0	14	8	2	14	0	2	4	4	0	2	12
St. Abbs	1	3	2	2	4	0	0	4	1	3	4	0	2	0	2	0	1	3

Table S4: Logistic regression results for the *crab-wave* and *wave-barnacle* contrasts of the Northern saxatilis genetic group. Three models were run per inversion; Null: the location effect, Eco: location + ecotype effects, and Int: the interaction effect. AIC = Akaike information criteria. R^2 = Cohen's pseudo R^2. Dev = deviance. df = degrees of freedom. P-value = P-value with a Benjamini-Hochberg adjustment. Bold AIC indicates the best model and significant P-values are highlighted in green.

		crab-wave (n=52)					*wave-barnacle* (n=30)				
Inv	Model	AIC	R^2	Dev	df	P-value	AIC	R^2	Dev	df	P-value
LGC1.1	*Eco*	28.34	$7.54e^{-4}$	$7.27e^{-3}$	1	1.00	8.87	1.00	2.72	1	0.30
	Int	26.70	1.00	9.64	4	0.70	10.87	1.00	$2.81e^{-10}$	1	1.00
	Null	**26.34**					**9.59**				
LGC1.2	*Eco*	29.94	0.10	0.03	1	1.00	**11.63**	1.00	8.38	1	0.07
	Int	37.64	1.00	0.30	4	1.00	13.63	1.00	$4.20e^{-11}$	1	1.00
	Null	**27.97**					18.01				
LGC2.1	*Eco*	32.32	0.09	1.09	1	1.00	11.10	1.00	3.42	1	0.26
	Int	29.65	1.00	10.67	4	0.55	13.10	1.00	$2.66e^{-12}$	1	1.00
	Null	**31.41**					**12.53**				
LGC4.1	*Eco*	30.30	0.08	0.50	1	1.00	**11.51**	1.00	5.05	1	0.13
	Int	32.48	1.00	5.82	4	1.00	13.51	1.00	$8.08e^{-13}$	1	1.00
	Null	**28.80**					14.56				
LGC5.1	*Eco*	21.44	0.34	0.53	1	1.00	7.91	1.00	0.60	1	0.44
	Int	28.43	1.00	1.01	4	1.00	9.91	1.00	$1.66e^{-10}$	1	1.00
	Null	**19.97**					**6.51**				
LGC6.1/2	See Table S6										
LGC7.1	*Eco*	33.74	0.05	0.38	1	1.00	8.71	1.00	1.95	1	0.33
	Int	34.63	1.00	7.11	4	1.00	10.71	1.00	$2.44e^{-10}$	1	1.00
	Null	**31.12**					**8.66**				
LGC7.2	*Eco*	35.44	0.34	2.44	1	1.00	**11.47**	1.00	5.24	1	0.13
	Int	38.79	1.00	4.65	4	1.00	13.47	1.00	$1.64e^{-11}$	1	1.00
	Null	**35.88**					14.71				
LGC9.1	*Eco*	30.85	0.67	3.22	1	1.00	**15.75**	0.73	6.69	1	0.12
	Int	37.28	1.00	1.57	4	1.00	15.24	1.00	2.51	1	1.00
	Null	**32.07**					20.43				
LGC9.2	Not tested due to uncertain clustering										
LGC10.1	*Eco*	35.99	0.12	1.80	1	1.00	**11.74**	1.00	7.83	1	0.08
	Int	**30.15**	1.00	13.84	4	0.16	13.74	1.00	$5.76e^{-12}$	1	1.00
	Null	35.79					17.57				
LGC10.2	*Eco*	32.12	0.01	0.11	1	1.00	**11.59**	1.00	5.91	1	0.12
	Int	31.47	1.00	8.65	4	0.91	13.59	1.00	$8.97e^{-12}$	1	1.00
	Null	**30.23**					15.50				
LGC11.1	*Eco*	30.12	0.42	1.80	1	1.00	**11.72**	1.00	6.87	1	0.12
	Int	35.60	1.00	2.52	4	1.00	13.72	1.00	$1.68e^{-13}$	1	1.00
	Null	**29.92**					16.58				
LGC12.1	*Eco*	**34.23**	0.29	3.52	1	0.97	**10.83**	1.00	10.36	1	0.02
	Int	33.56	1.00	8.67	4	0.91	12.83	1.00	0.00	1	1.00
	Null	35.74					19.19				
LGC12.2	*Eco*	38.09	$5.06e^{-3}$	0.05	1	1.00	**10.87**	1.00	0.16	1	$1.15e^{-3}$
	Int	35.46	1.00	10.63	4	0.55	12.87	1.00	$1.47e^{-10}$	1	1.00
	Null	**36.14**					25.16				
LGC12.3	Not tested due to uncertain clustering										
LGC12.4	*Eco*	**33.40**	0.42	5.41	1	0.34	**11.72**	1.00	9.24	1	0.04
	Int	34.04	1.00	7.36	4	1.00	13.72	1.00	$1.93e^{-11}$	1	1.00
	Null	36.80					18.95				
LGC14.1	*Eco*	34.85	0.58	16.57	1	$9.37e^{-4}$	**11.72**	1.00	6.87	1	0.12
	Int	**31.04**	1.00	11.80	4	0.38	13.72	1.00	$1.64e^{-13}$	1	1.00
	Null	49.42					16.58				
LGC14.2	See Table S6										
LGC14.3	Not tested due to uncertain clustering										
LGC17.1	*Eco*	**18.16**	1.00	6.47	1	0.20	**16.20**	0.68	6.81	1	0.12
	Int	26.16	1.00	0.00	4	1.00	15.07	1.00	3.13	1	1.00
	Null	22.63					21.01				

Table S5: Logistic regression results for *Littorina saxatilis – L. arcana* contrasts. *L. saxatilis* was separated into crab and wave ecotypes. Three models were run *per* inversion; Null: the location effect, Eco: location + ecotype effects, and Int: the interaction effect. *Crab-arcana* contrasts were limited to a single location. AIC = Akaike information criteria. R^2 = *Cohen*'s pseudo R^2. Dev = deviance. df = degrees of freedom. P-value = P-value with a Benjamini-Hochberg adjustment. Bold AIC indicates the best model and significant P-values are highlighted in green.

		crab-arcana (n=16)					*wave-arcana* (n=56)				
Inv	Model	AIC	R^2	Dev	df	P-value	AIC	R^2	Dev	df	P-value
LGC1.1	*Eco*	**6.53**	0	4.86	1	0.33	21.96	0.36	0.82	1	1.00
	Int	-	-	-	-	-	26.48	1.00	1.48	3	1.00
	Null	9.39					**20.78**				
LGC1.2	*Eco*	**8.40**	0	4.56	1	0.33	32.66	0.45	8.99	1	0.02
	Int	-	-	-	-	-	**27.64**	1.00	11.02	3	0.19
	Null	10.96					39.65				
LGC2.1	*Eco*	**7.74**	0	10.12	1	0.02	**23.94**	0.87	12.11	1	$4.51e^{-3}$
	Int	-	-	-	-	-	28.06	1.00	1.88	3	1.00
	Null	15.86					34.05				
LGC4.1	*Eco*	**5.87**	0	15.90	1	$1.13e^{-3}$	**24.54**	0.87	22.13	1	$4.08e^{-5}$
	Int	-	-	-	-	-	27.38	1.00	3.16	3	1.00
	Null	19.77					44.67				
LGC5.1	Not tested due to uncertain clustering										
LGC6.1/2	See Table S6										
LGC7.1	*Eco*	**5.87**	0	15.90	1	$1.13e^{-3}$	28.78	0.60	12.90	1	$3.28e^{-3}$
	Int	-	-	-	-	-	**26.30**	1.00	8.48	3	0.48
	Null	19.77					39.69				
LGC7.2	*Eco*	**6.53**	0	4.86	1	0.33	**24.52**	0.81	19.29	1	$1.68e^{-4}$
	Int	-	-	-	-	-	25.95	1.00	4.57	3	1.00
	Null	9.39					41.82				
LGC9.1	*Eco*	**6.53**	0	4.86	1	0.33	**21.90**	0.92	15.80	1	$9.15e^{-4}$
	Int	-	-	-	-	-	26.55	1.00	1.35	3	1.00
	Null	9.39					35.70				
LGC9.2	*Eco*	8.20	0	0.42	1	1.00	**12.33**	1.00	15.10	1	$1.23e^{-3}$
	Int	-	-	-	-	-	18.33	1.00	0	3	1.00
	Null	**6.62**					25.43				
LGC10.1	*Eco*	**6.53**	0	4.86	1	0.33	30.86	0.21	1.74	1	0.75
	Int	-	-	-	-	-	30.44	1.00	6.42	3	1.00
	Null	9.39					**30.60**				
LGC10.2	*Eco*	8.40	0	1.38	1	1.00	33.78	0.37	7.89	1	0.03
	Int	-	-	-	-	-	**26.22**	1.00	13.57	3	0.06
	Null	**7.79**					39.67				
LGC11.1	*Eco*	**8.40**	0	4.56	1	0.33	27.86	0.44	2.95	1	0.43
	Int	-	-	-	-	-	30.08	1.00	3.78	3	1.00
	Null	10.96					**28.81**				
LGC12.1	*Eco*	9.07	0	$1.33e^{-15}$	1	1.00	32.95	0.03	0.37	1	1.00
	Int	-	-	-	-	-	**28.26**	1.00	10.69	3	0.20
	Null	**7.07**					31.32				
LGC12.2	Not tested due to sex-chromosome bias										
LGC12.3	Not tested due to uncertain clustering										
LGC12.4	*Eco*	8.40	0	1.38	1	1.00	34.94	0.03	0.38	1	1.00
	Int	-	-	-	-	-	**30.40**	1.00	10.54	3	0.20
	Null	**7.79**					33.31				
LGC14.1	*Eco*	**6.53**	0	9.29	1	0.03	**17.10**	1.00	27.18	1	$3.16e^{-6}$
	Int	-	-	-	-	-	23.10	1.00	0	3	1.00
	Null	13.82					42.28				
LGC14.2	Not tested due to uncertain clustering										
LGC14.3	Not tested due to uncertain clustering										
LGC17.1	*Eco*	4.00	0	0	1	1.00	**16.83**	1.00	13.26	1	$2.98e^{-3}$
	Int	-	-	-	-	-	22.83	1.00	0	3	1.00
	Null	**2.00**					28.09				

Table S6: Logistic regression results of the complex double inversions; LGC6.1/2 and LGC14.2. Analyses were run twice, firstly combining the two alternate arrangements ($A_1 + A_2 = A$) before contrasting with the R arrangement. Secondly, the A_1 and A_2 arrangement were compared to determine if the ecotype effect is present in the alternate arrangements. Three models were run per inversion; *Null*: the location effect, *Eco*: location + ecotype effects, and *Int:* the interaction effect. *Crab-arcana* contrasts were limited to a single location. AIC = Akaike information criteria. R^2 = Cohen's pseudo R^2. Dev = deviance. df = degrees of freedom. P-value = P-value with a Benjamini-Hochberg adjustment. Bold AIC indicates the best model and significant P-values are highlighted in green.

		crab-wave (n=52)					*wave-barnacle* (n=30)				
Inv	Model	AIC	R^2	Dev	df	P-value	AIC	R^2	Dev	df	P-value
LGC6.1/2 (R vs. A)	*Eco*	**26.89**	0.89	19.92	1	$1.70e^{-4}$	**11.37**	1.00	6.01	1	0.12
	Int	32.50	1.00	2.39	4	1.00	13.37	1.00	$6.43e^{-11}$	1	1.00
	Null	44.81					15.38				
LGC6.1/2 (A_1 vs. A_2)	*Eco*	22.06	1.00	6.56	1	1.00	**10.82**	1.00	6.56	1	0.12
	Int	25.35	1.00	$1.49e^{-11}$	1	1.00	12.82	1.00	$1.49e^{-11}$	1	1.00
	Null	**20.59**					15.37				
LGC14.2 (R vs. A)	*Eco*	34.85	0.58	16.57	1	$9.3e^{-4}$	**11.72**	1.00	6.87	1	0.12
	Int	**31.04**	1.00	11.80	4	0.38	13.72	1.00	$1.64e^{-13}$	1	1.00
	Null	49.42					16.58				
LGC14.2 (A_1 vs. A_2)	*Eco*	19.93	1.00	6.87	1	0.57	**10.78**	1.00	7.23	1	0.11
	Int	18.36	1.00	$1.64e^{-13}$	1	0.82	12.78	1.00	$2.57e^{-12}$	1	1.00
	Null	**18.26**					16.01				

		crab-arcana (n=16)					*wave-arcana* (n=56)				
Inv	Model	AIC	R^2	Dev	df	P-value	AIC	R^2	Dev	df	P-value
LGC6.1/2 (R vs. A)	*Eco*	**5.87**	0	15.90	1	$1.13e^{-3}$	**18.72**	1.00	16.74	1	$6.01e^{-4}$
	Int	-	-	-	-	-	24.72	1.00	$1.73e^{-11}$	3	1.00
	Null	19.77					33.46				
LGC6.1/2 (A_1 vs. A_2)	*Eco*	5.87	0	0.25	1	1.00	**27.82**	0.43	4.18	1	0.25
	Int	-	-	-	-	-	28.26	1.00	5.56	3	1.00
	Null	**4.12**					30.00				

Supplementary figures:

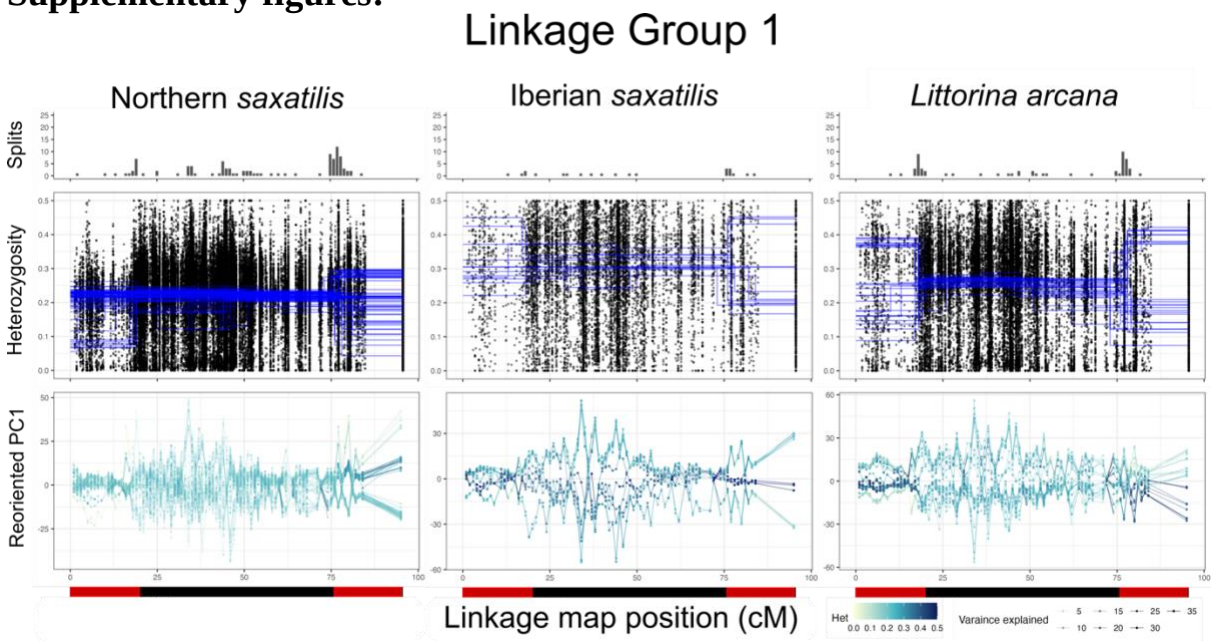

Figure S1: Inversion detection summary for LG1. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

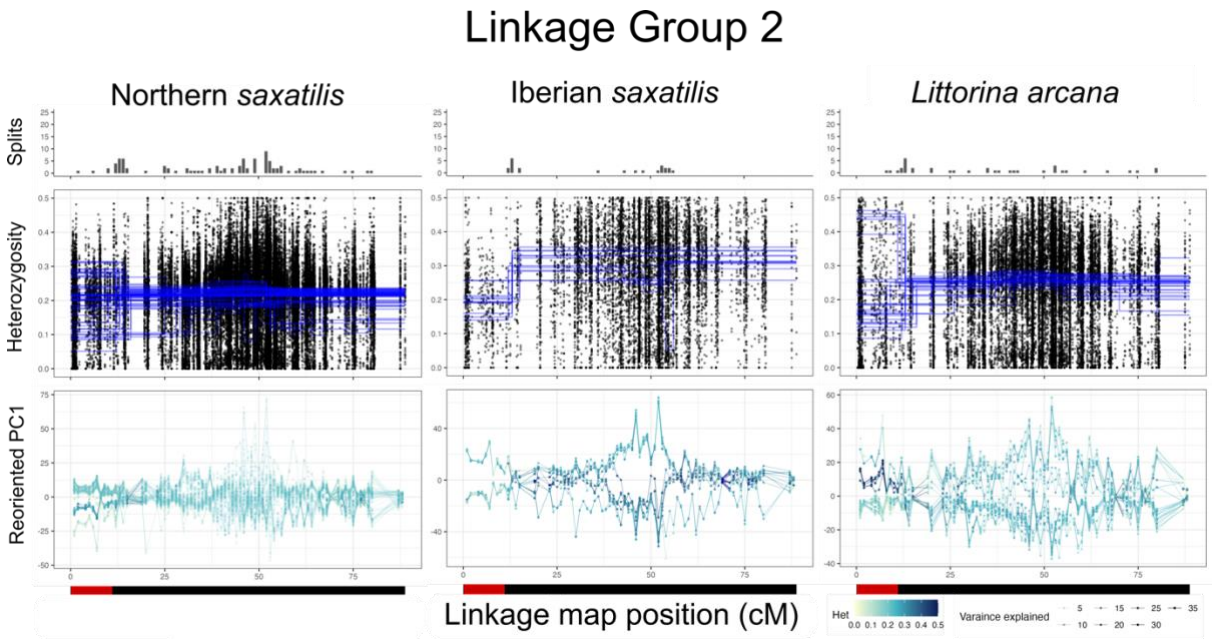

Figure S2: Inversion detection summary for LG2. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

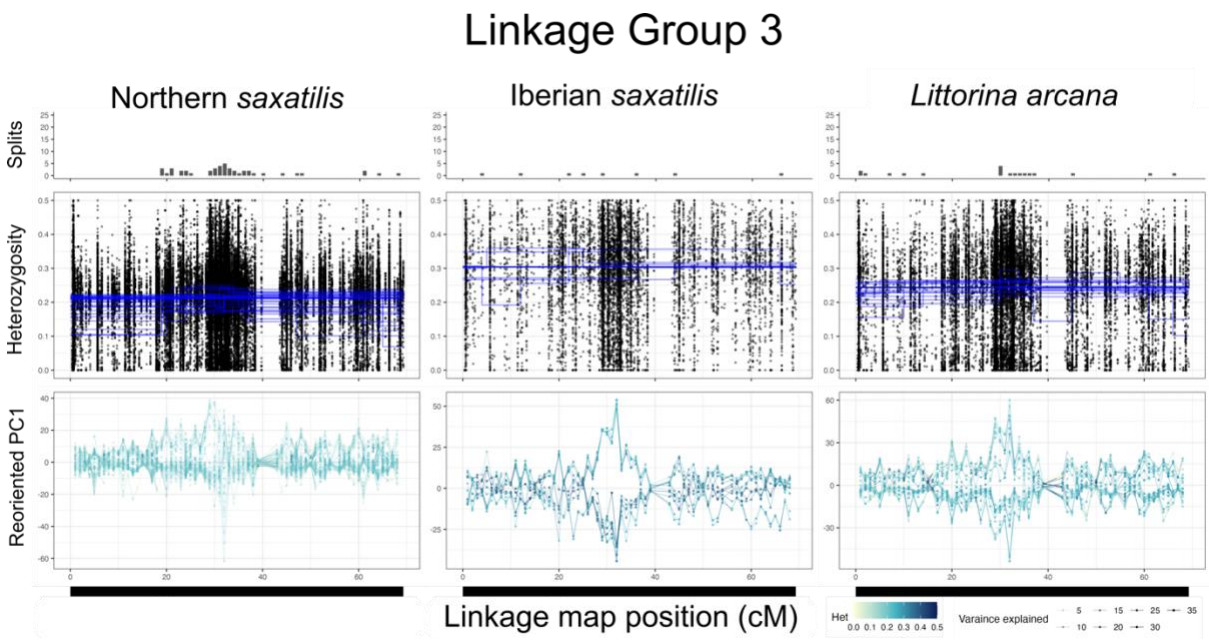

Figure S3: Inversion detection summary for LG3. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

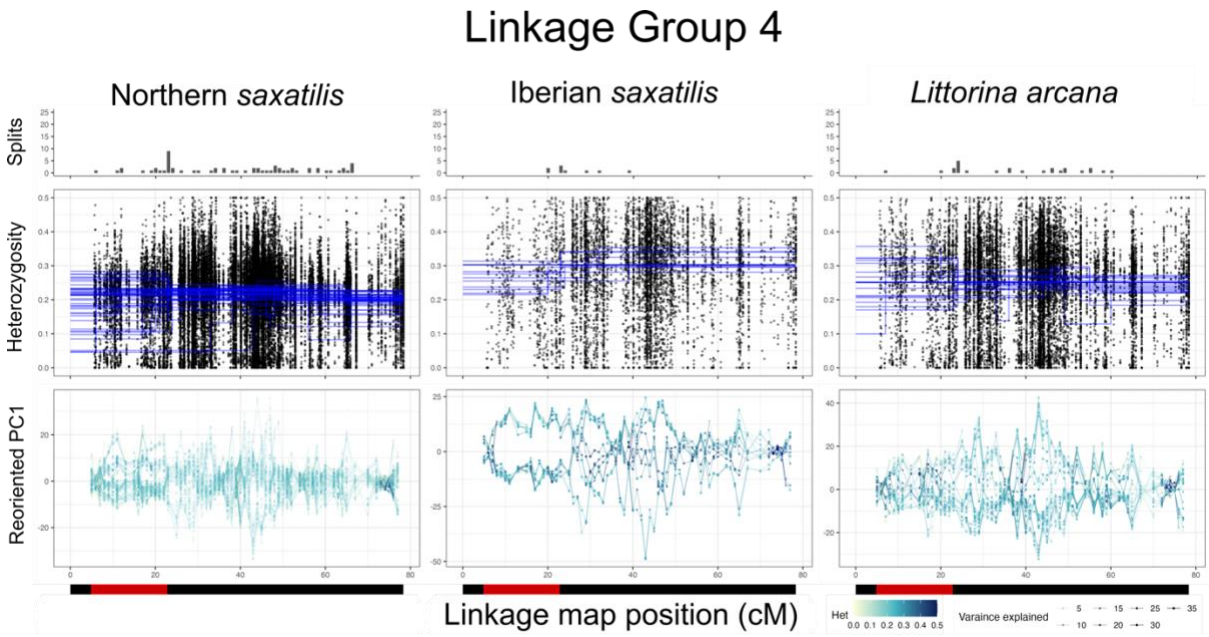

Figure S4: Inversion detection summary for LG4. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

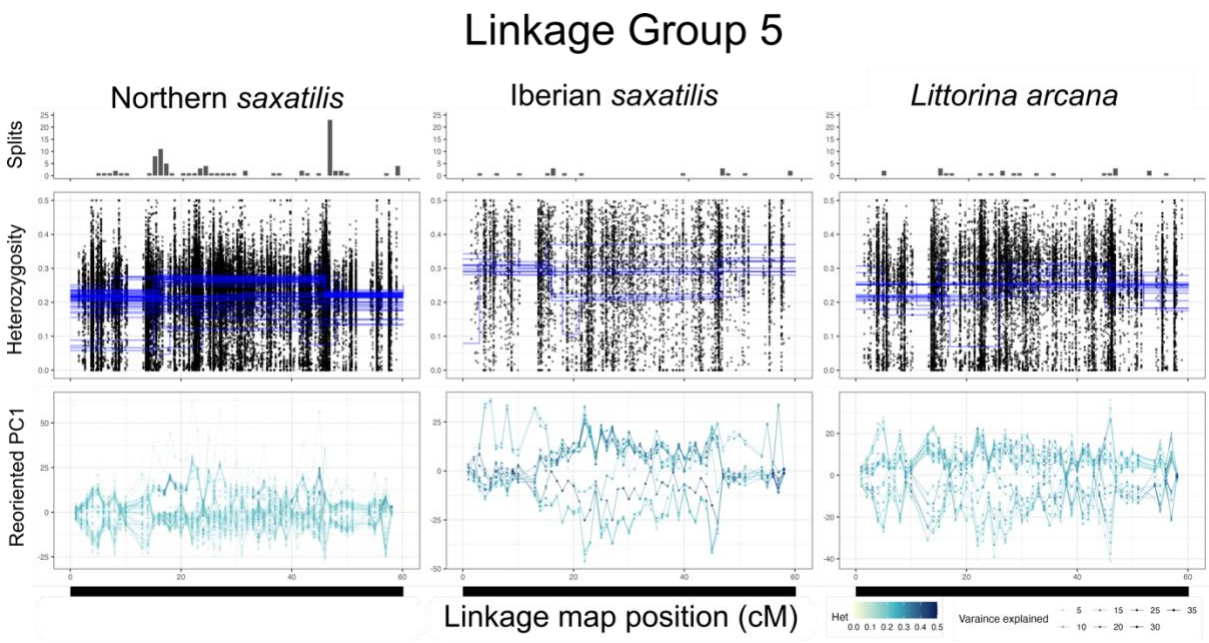

Figure S5: Inversion detection summary for LG5. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

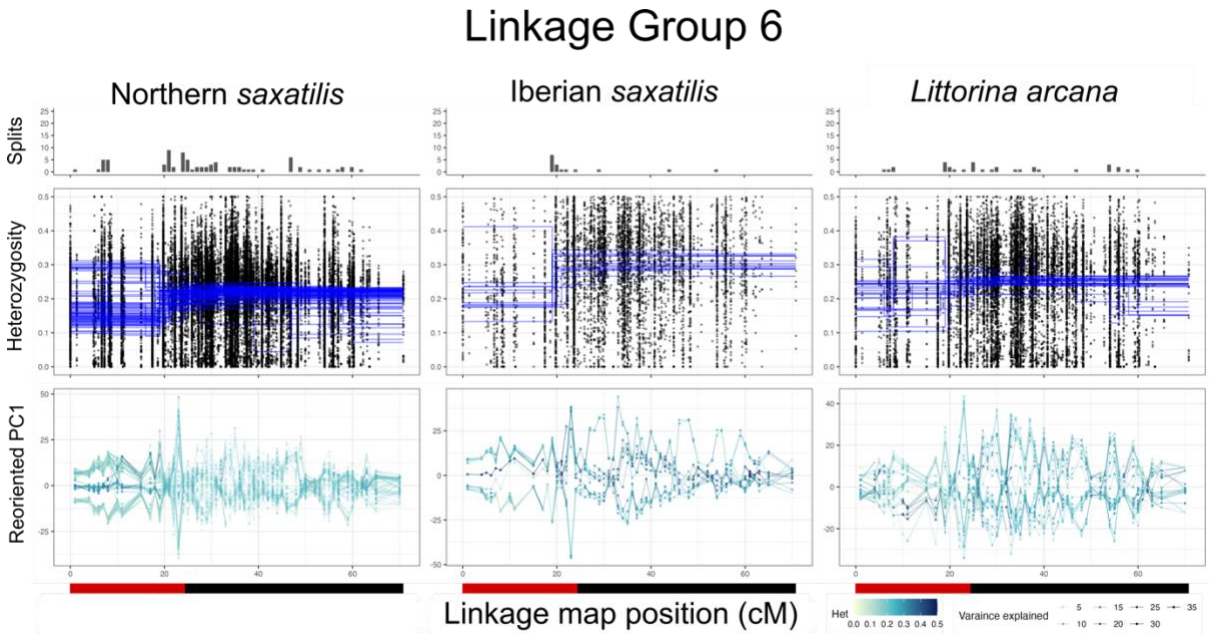

Figure S6: Inversion detection summary for LG6. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

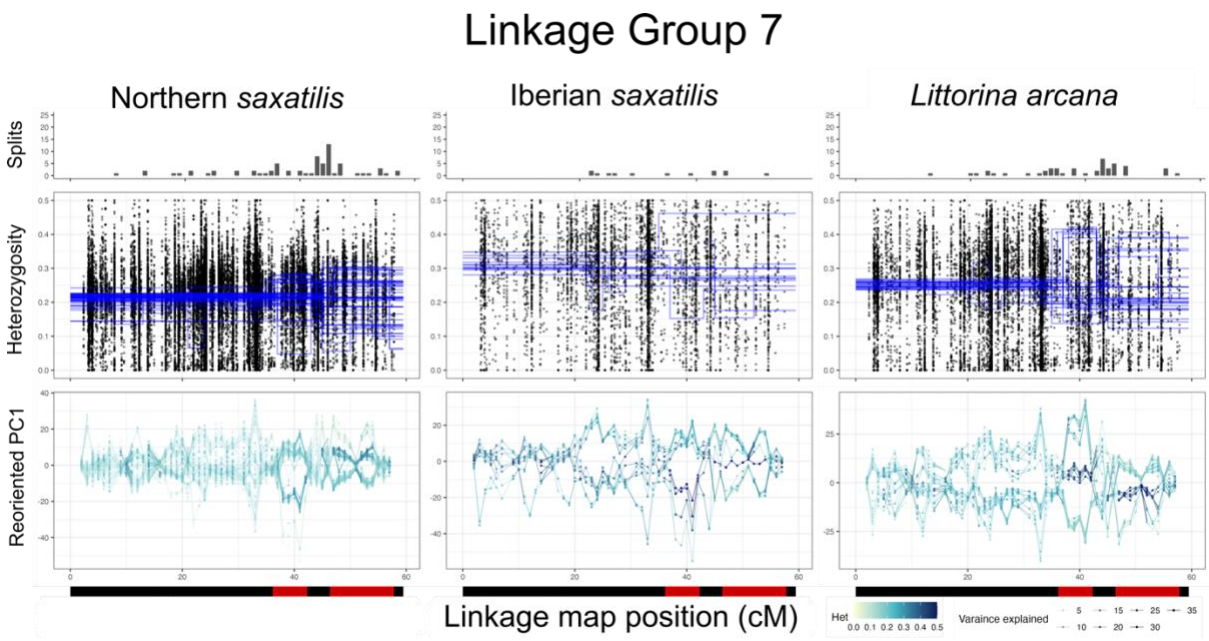

Figure S7: Inversion detection summary for LG7. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

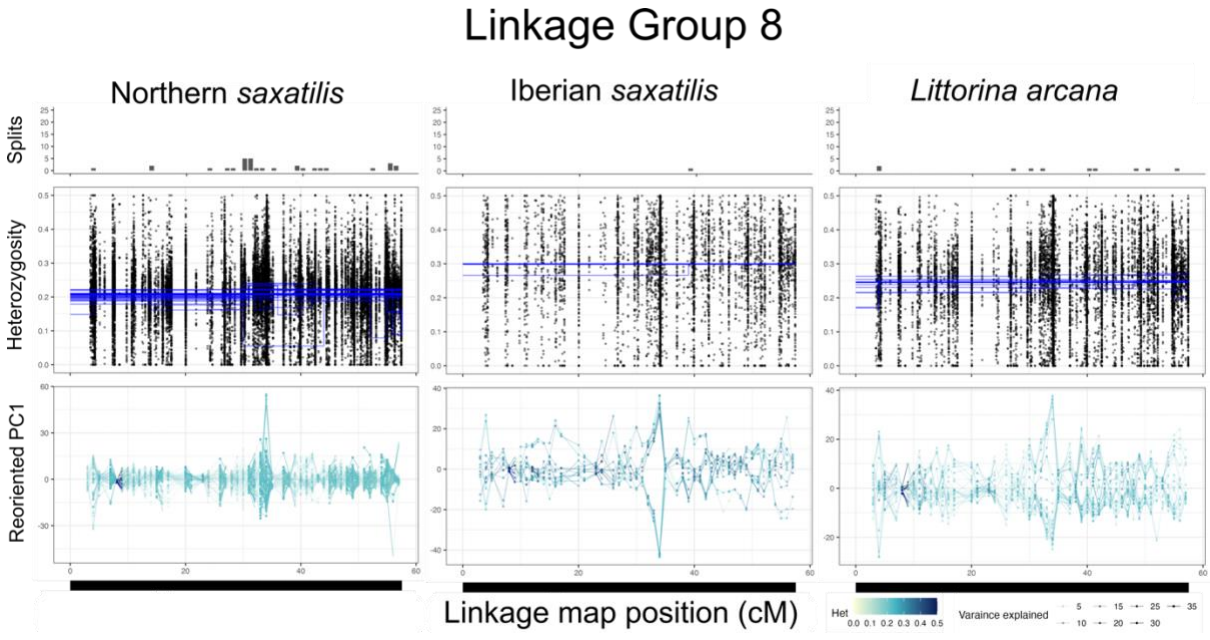

Figure S8: Inversion detection summary for LG8. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

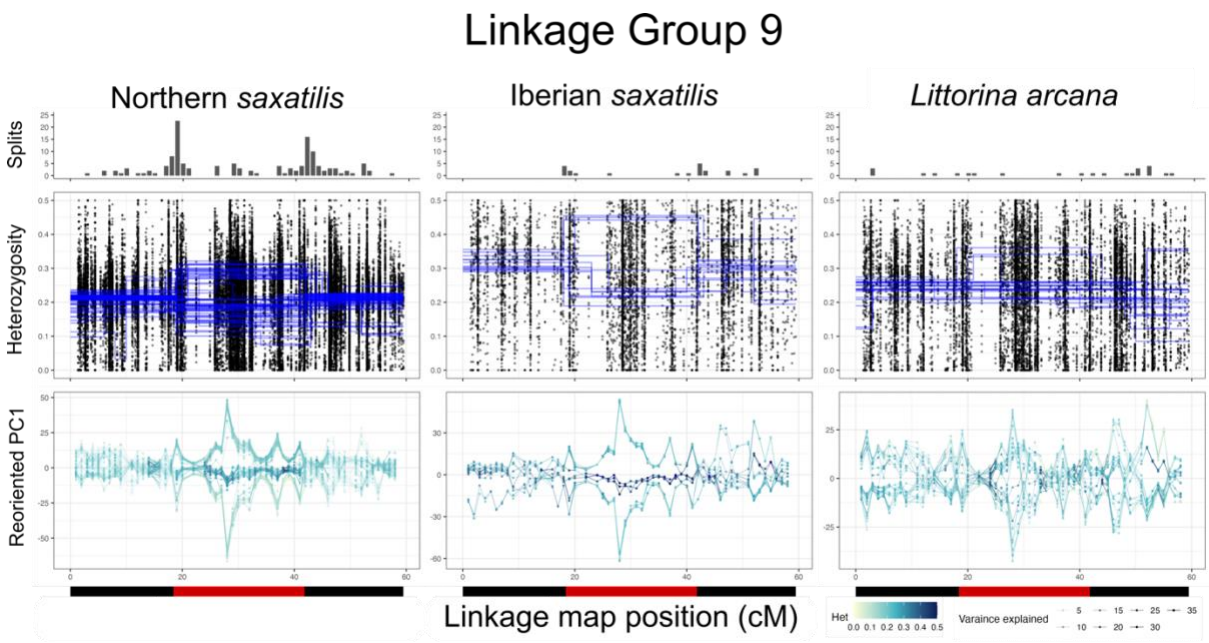

Figure S9: Inversion detection summary for LG9. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

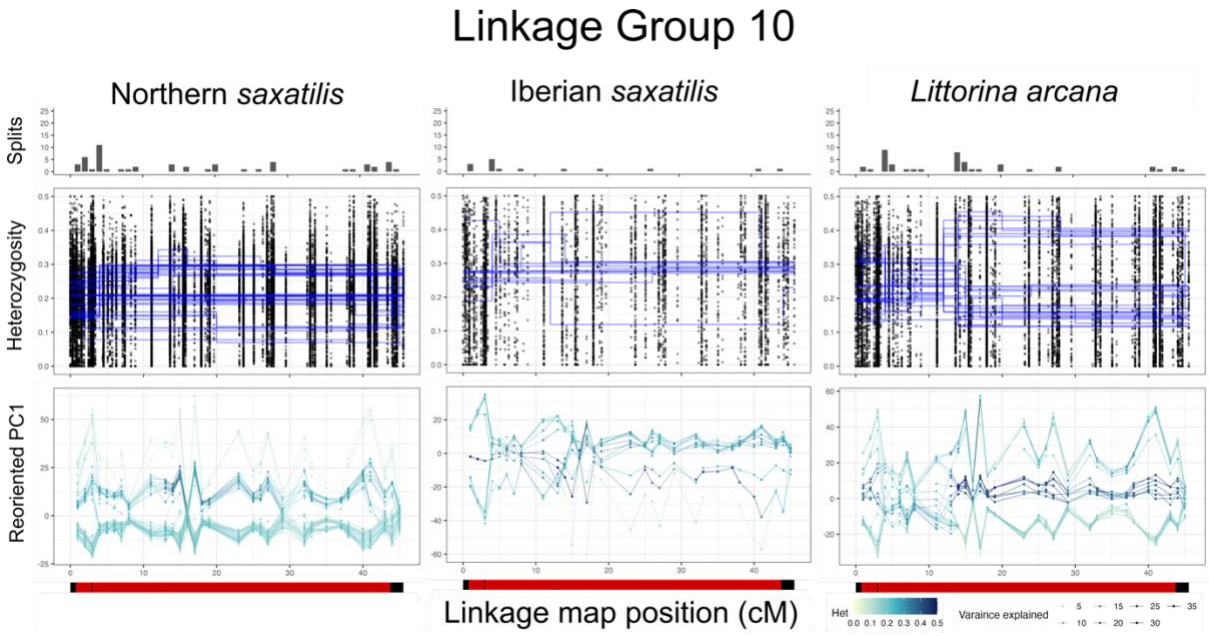

Figure S10: Inversion detection summary for LG10. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

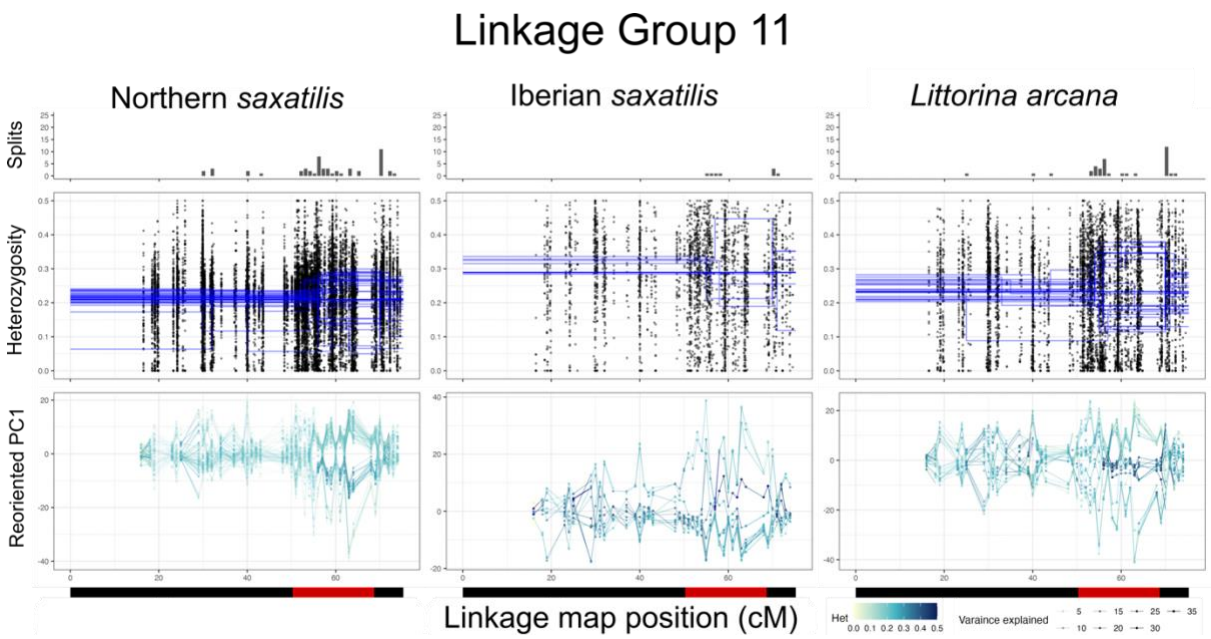

Figure S11: Inversion detection summary for LG11. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

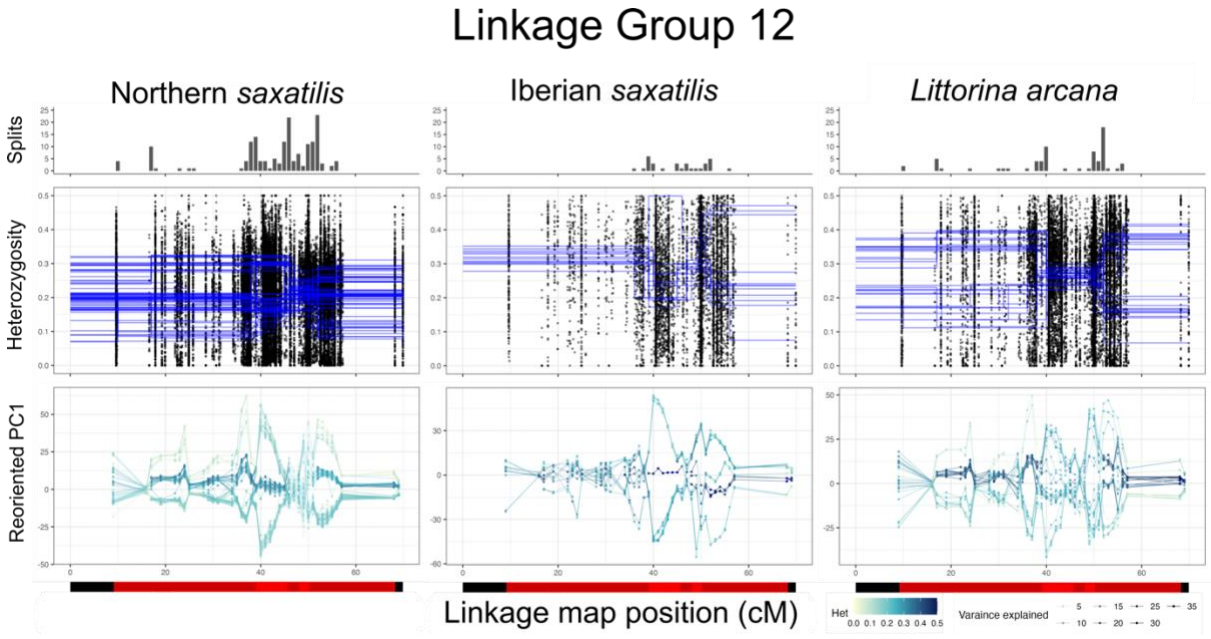

Figure S12: Inversion detection summary for LG12. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

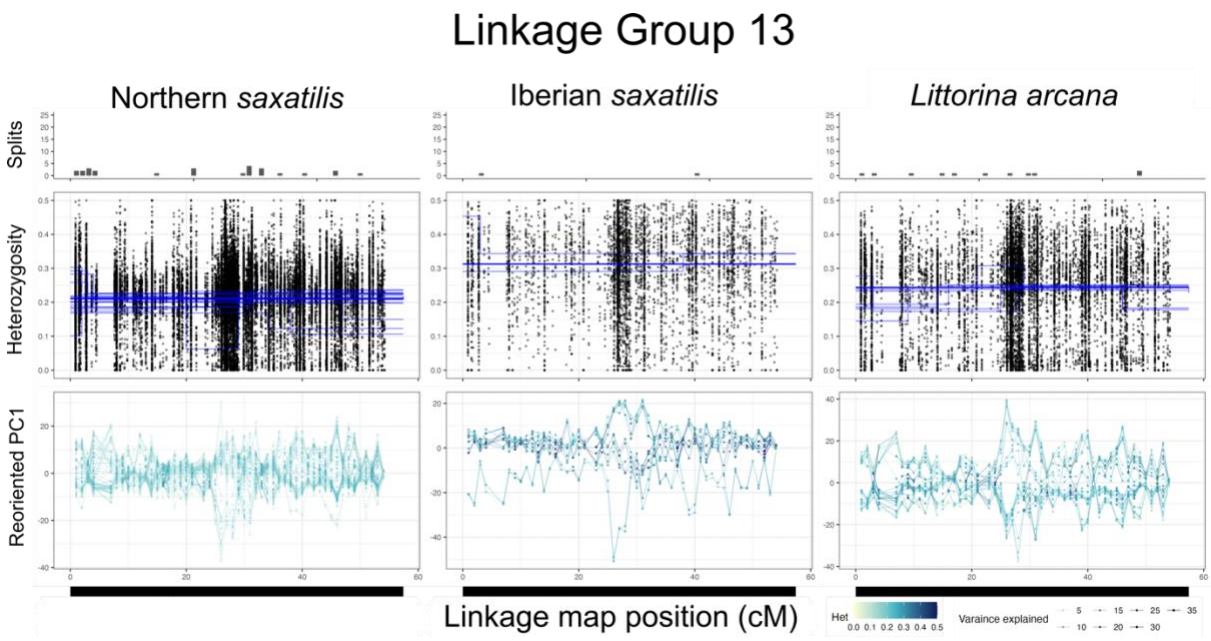

Figure S13: Inversion detection summary for LG13. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

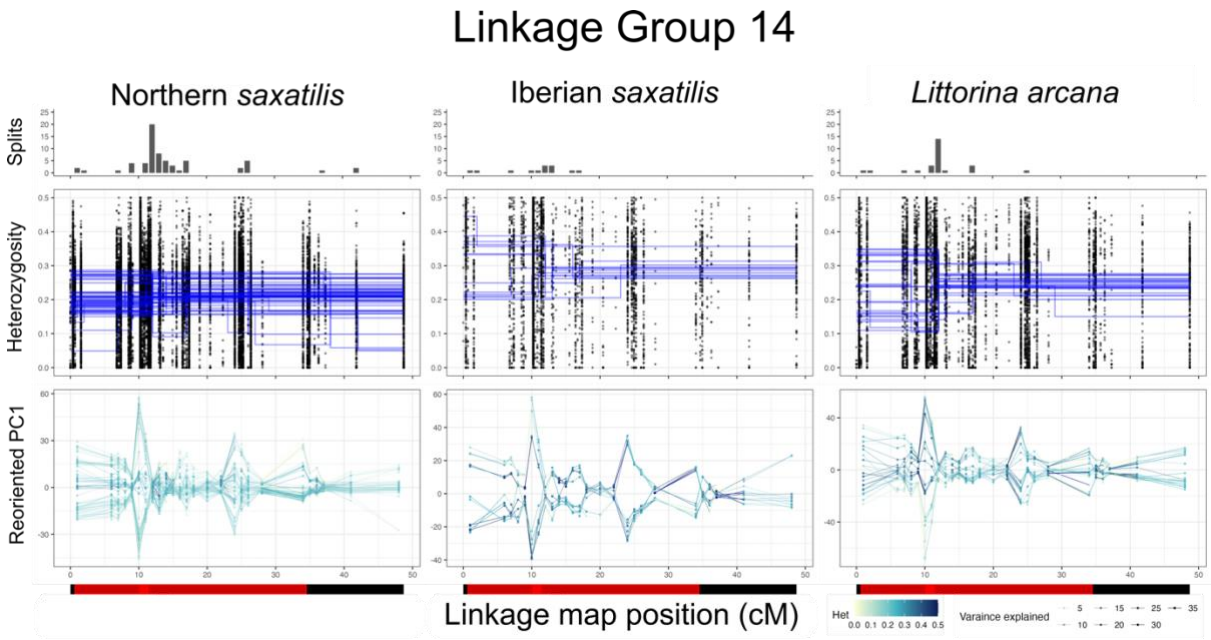

Figure S14: Inversion detection summary for LG14. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

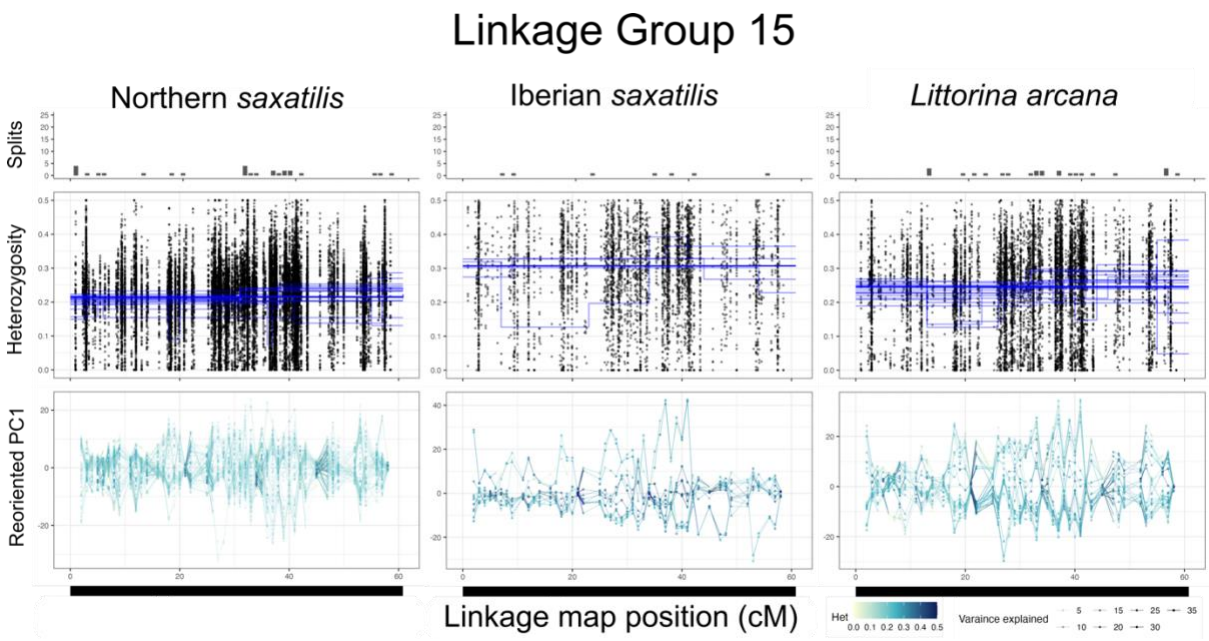

Figure S15: Inversion detection summary for LG15. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

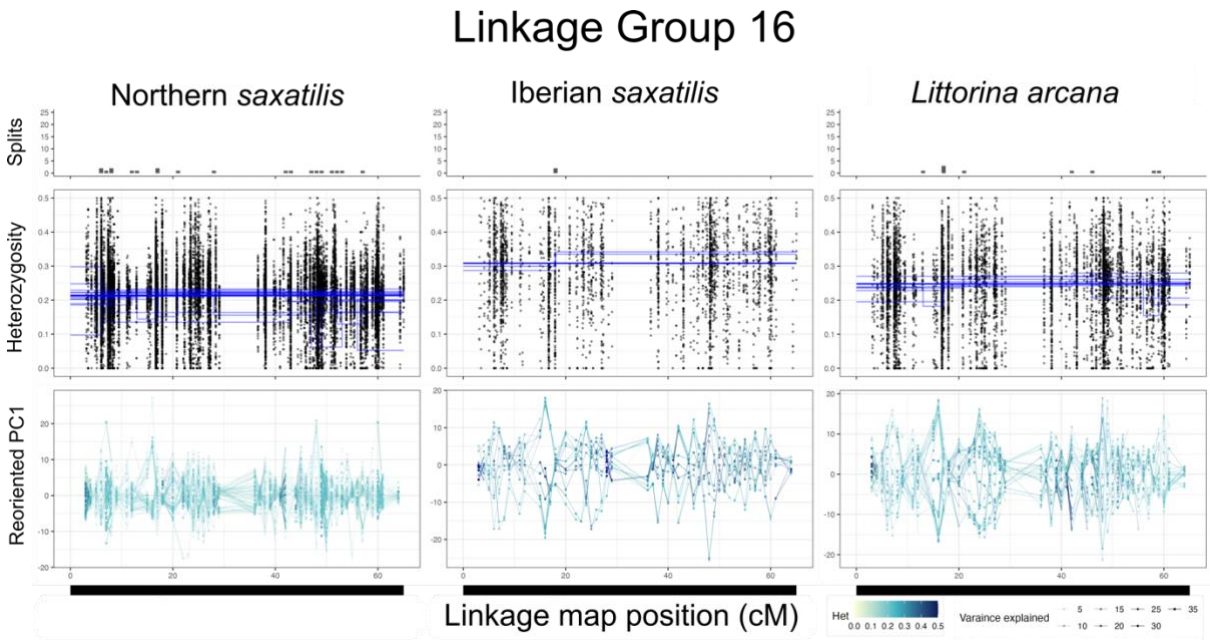

Figure S16: Inversion detection summary for LG16. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

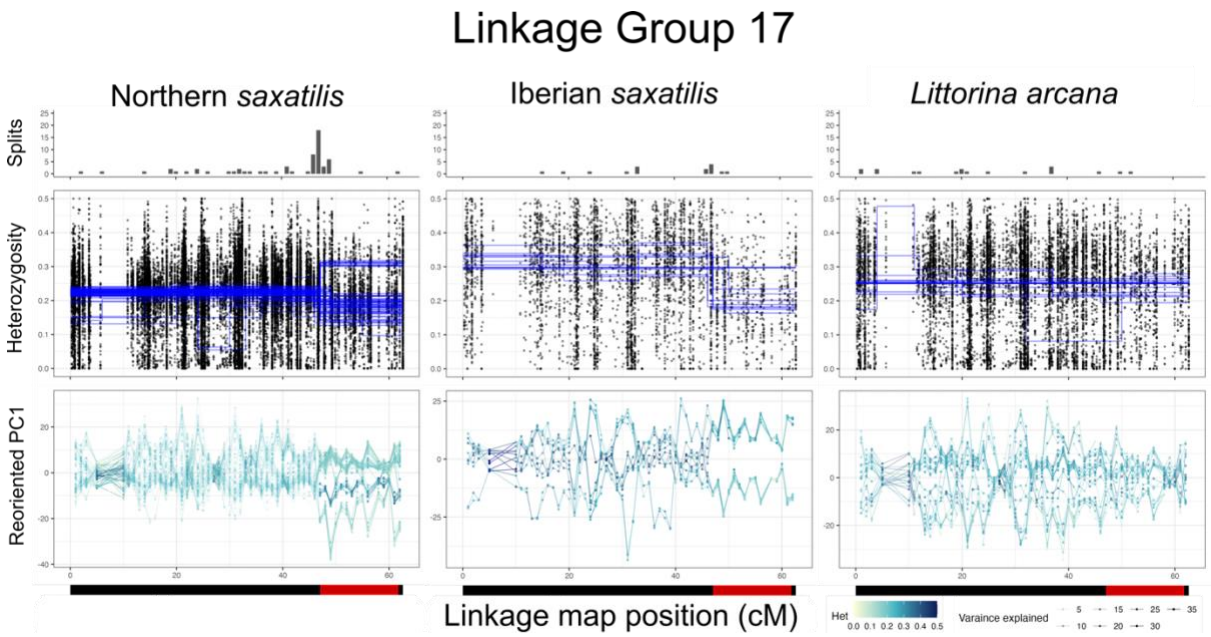

Figure S17: Inversion detection summary for LG17. *Top panel:* number of significant splits between each 1cM window. *2nd panel*: results of heterozygosity split approach. *3rd panel*: PC1 for each window. *Bottom panel*: position of published inversions as red bars. See Figure 2 in the main text for more details.

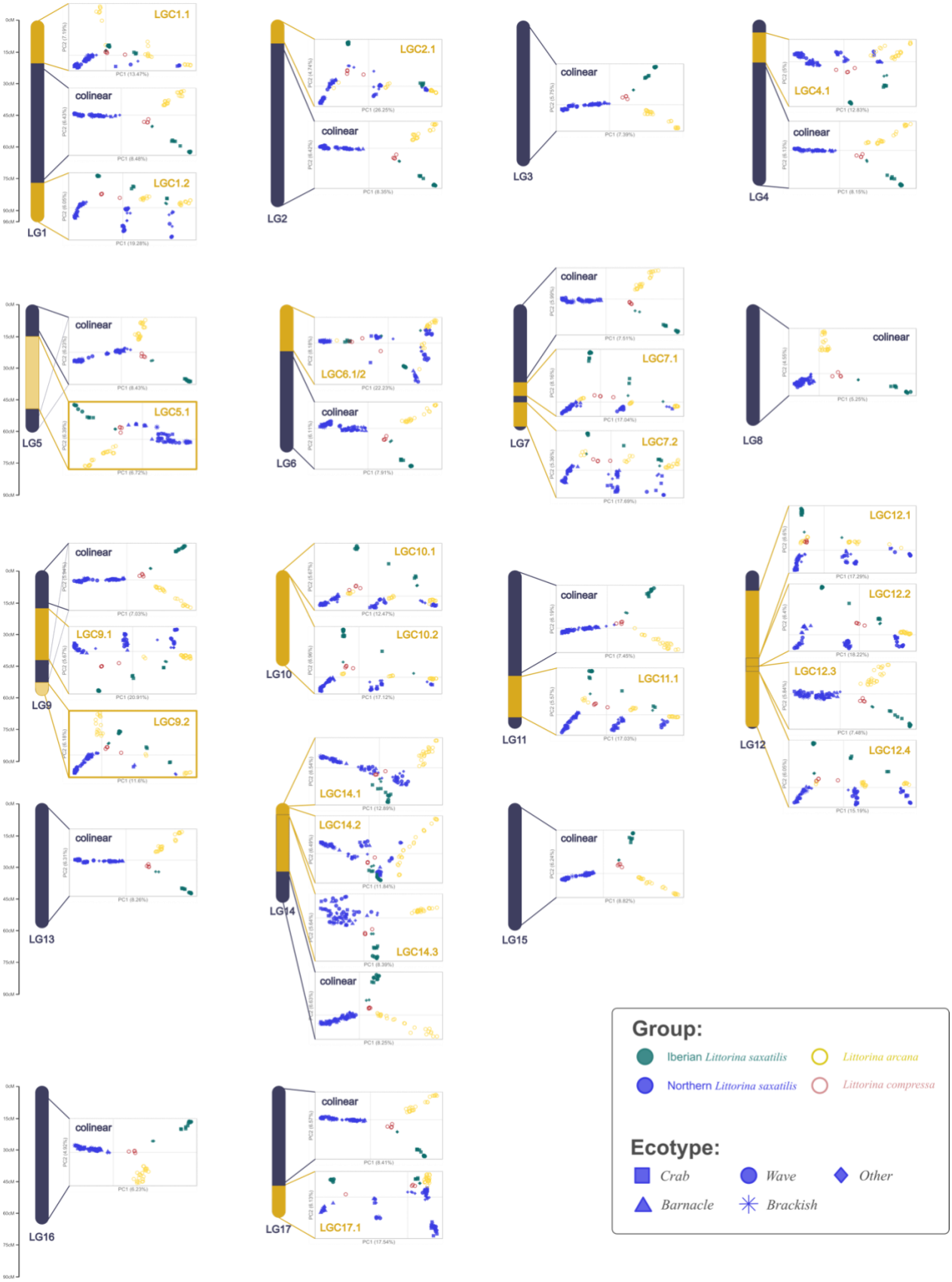

Figure S18: PCA plots for the inverted and colinear regions of each linkage group. Yellow bordered plots indicate the new putative inversions identified in this study. Each plot uses the full data from all snails. Inversion positions are based on Westram et al., 2021 and Hearn et al., 2022 (LG12), updated to fit the new linkage map.

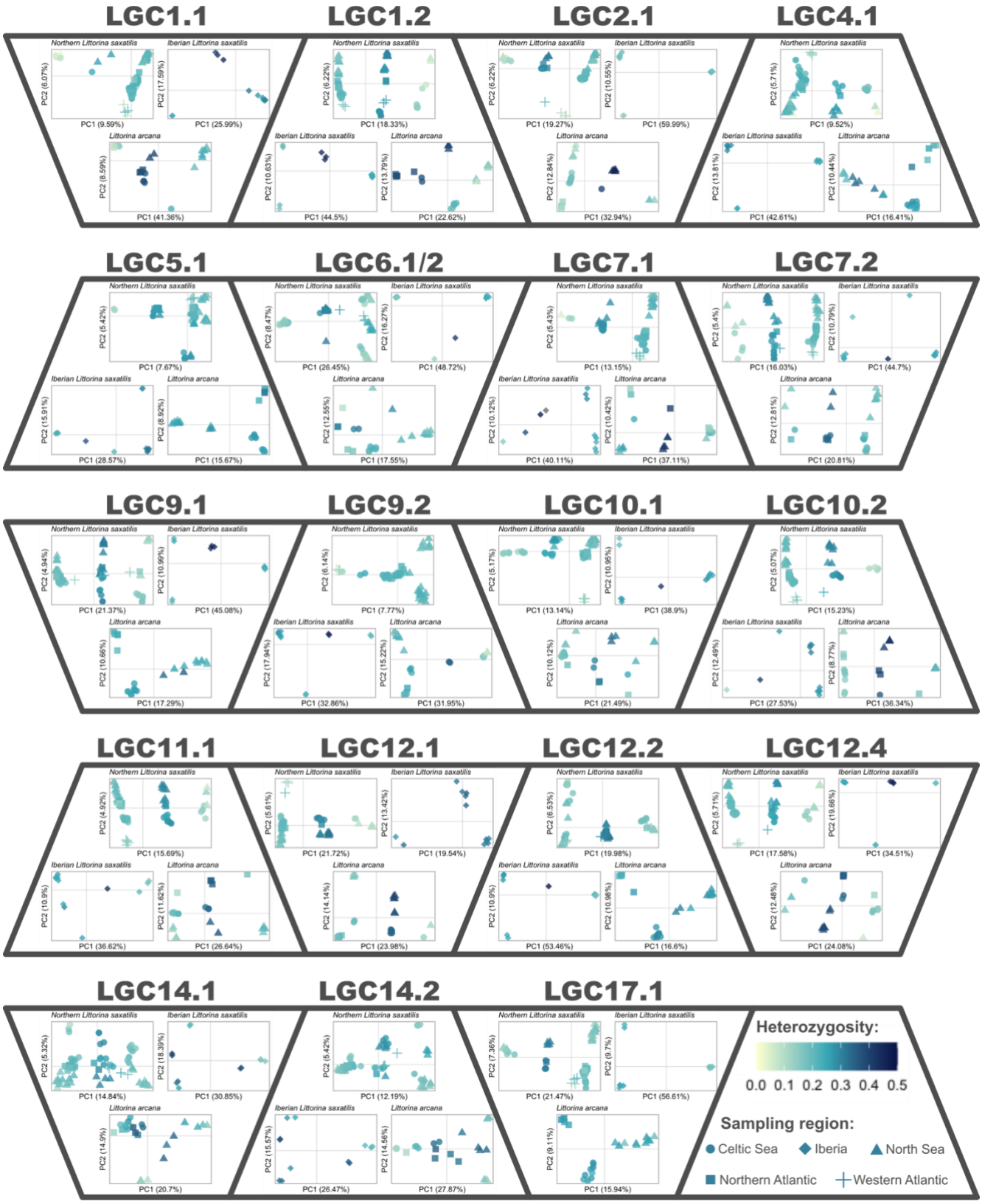

Figure S19: PCA plots of each inversion and genetic group. Points are shaded by the average heterozygosity per sample. Symbols represent the broad scale geographic regions identified in a recent phylogeographic study (Stankowski et al., 2023). Western Atlantic is an additional category we defined based on these plots.

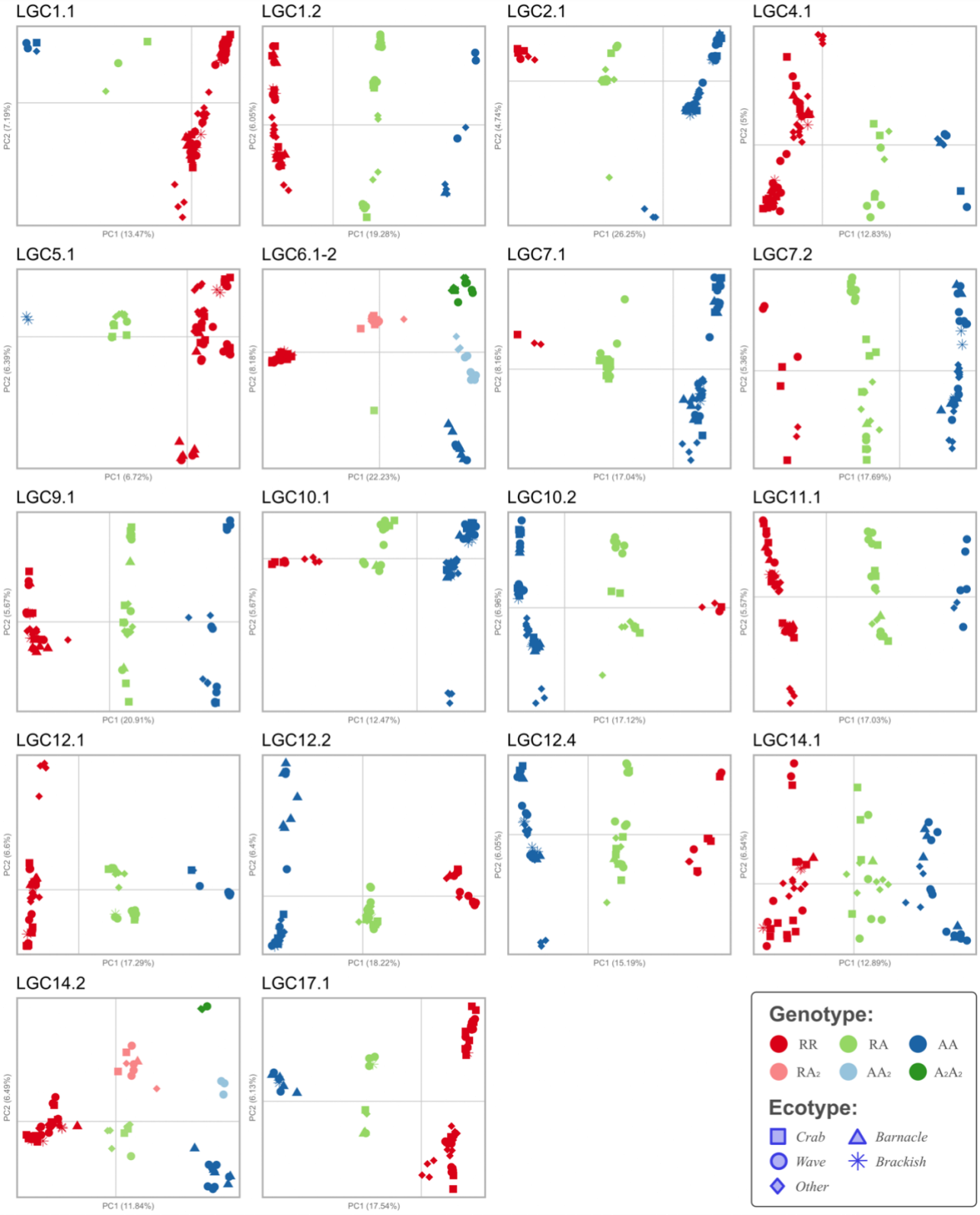

Figure S20: PCA plots of each inversion from Northern *Littorina saxatilis*. Points are coloured by genotype and symbols represent ecotypes. Colour palette matches Figure 4a in the main text. Only inversions that formed 3 or 6 clusters are shown.

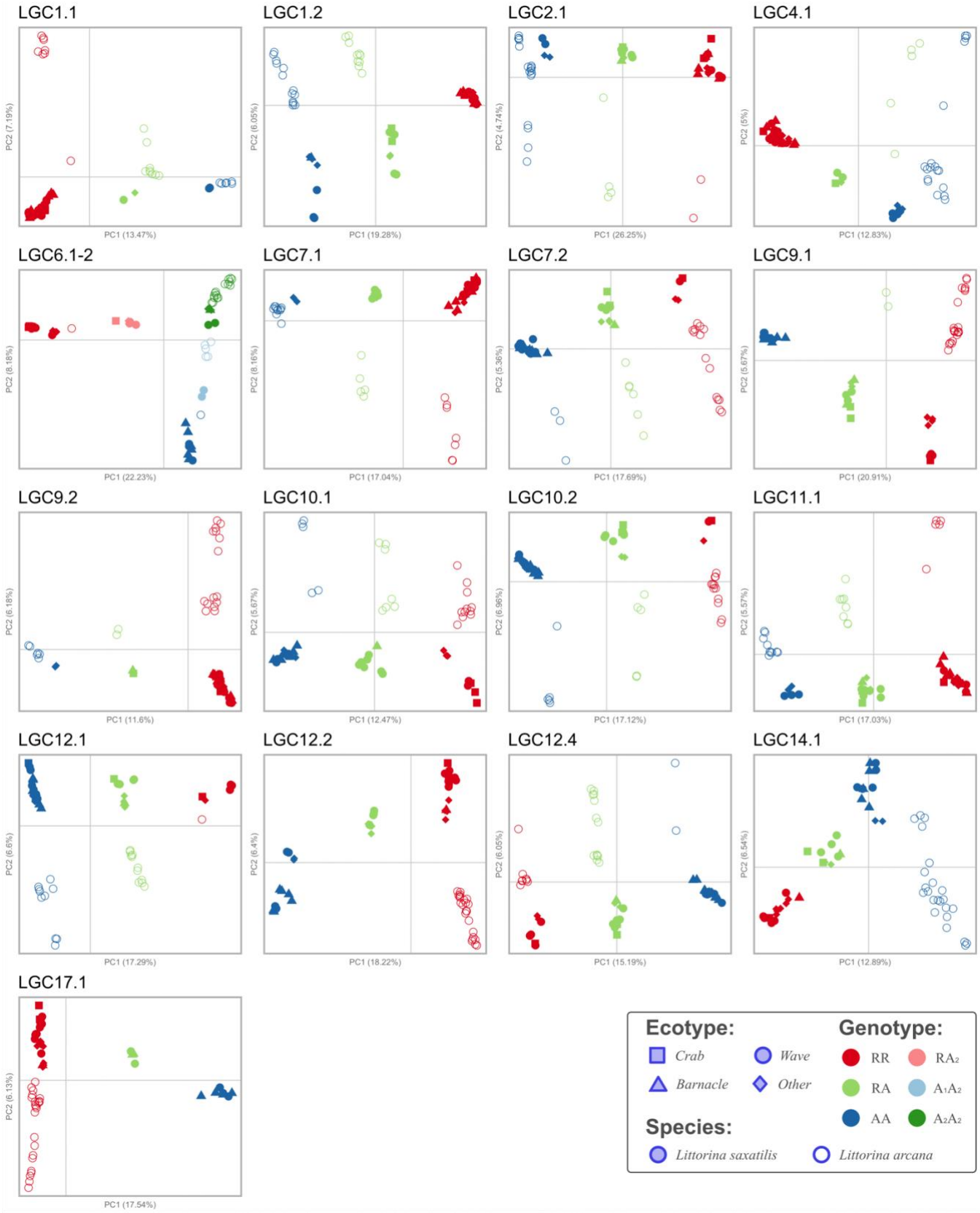

Figure S21: PCA plots of each inversion for sites where both *Littorina saxatilis* and *Littorina arcana* were collected. Points are coloured by genotype and symbols represent ecotypes/species. The colour palette matches Figure 5a in the main text.

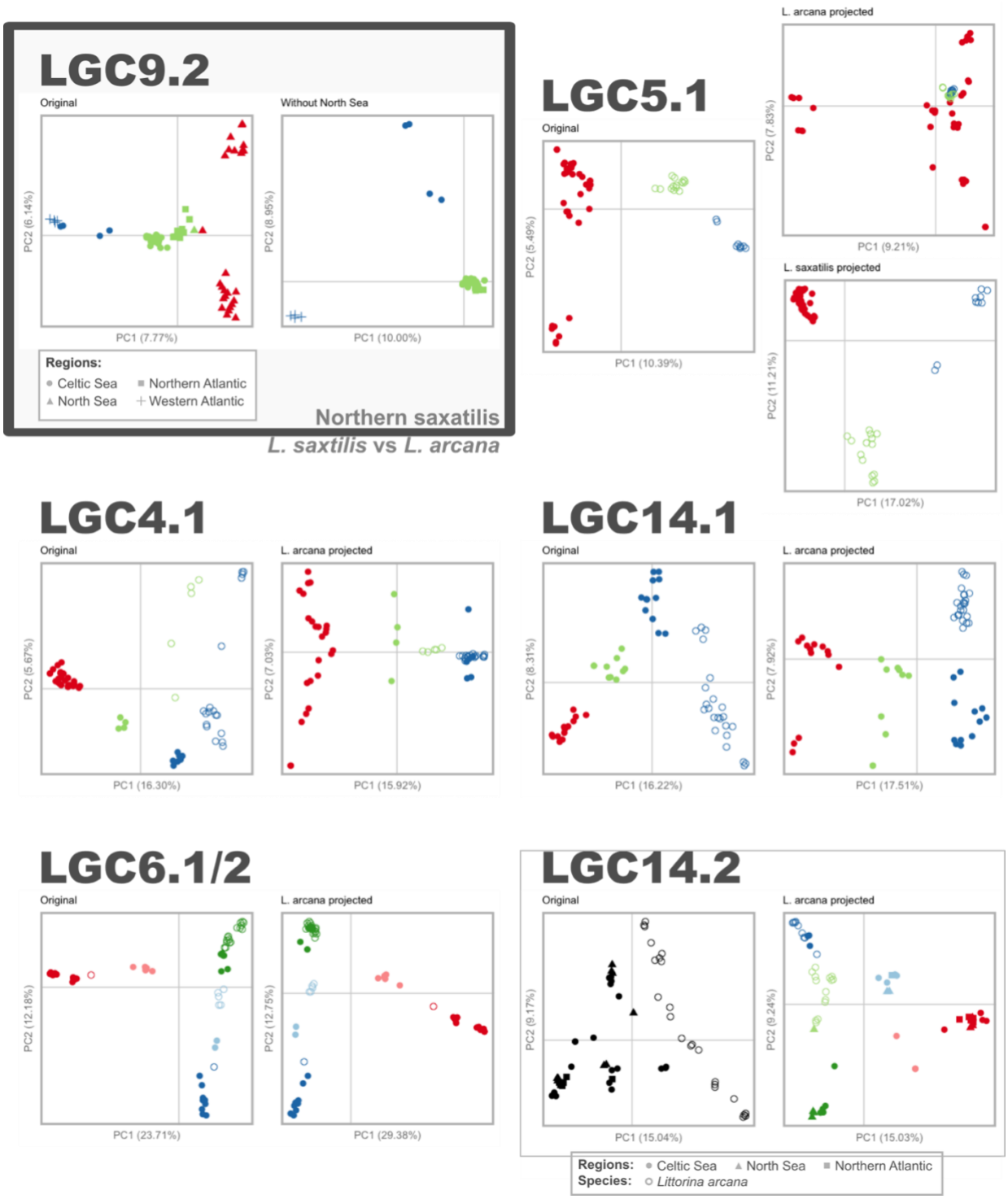

Figure S22: PCA plots demonstrating the manual adjustments made to genotype some inversions. Adjustments are separated into those that were made using the Northern saxatilis data (grey box), and the *Littorina saxatilis* vs. *L. arcana* data. Titles indicate if the panel is the original PCA result or a projection of one species onto the PCA of the other. In cases when geographic signal was pertinent to the manual adjustment, regions are displayed. Geographic regions were defined from a recent phylogeographic study (Stankowski et al., 2023), with USA samples split into a separate Western Atlantic group. Colours and shapes (except for LGC9.2 and LGC14.2) are the same as Figure 5a in the main text. Sampling regions for LGC9.2 and LGC14.2 are assigned different shapes (see legends), with *L. arcana* represented as hollow circles. The original projection of LGC14.2 is coloured black because genotypes could not be determined for this PCA.