



This is a repository copy of *Using replicate hybrid zones to understand the genomic basis of adaptive divergence*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/185726/>

Version: Published Version

---

**Article:**

Westram, A.M., Faria, R., Johannesson, K. et al. (1 more author) (2021) Using replicate hybrid zones to understand the genomic basis of adaptive divergence. *Molecular Ecology*, 30 (15). pp. 3797-3814. ISSN 0962-1083

<https://doi.org/10.1111/mec.15861>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Using replicate hybrid zones to understand the genomic basis of adaptive divergence

Anja M. Westram<sup>1,2,3</sup>  | Rui Faria<sup>2,4,5</sup>  | Kerstin Johannesson<sup>6</sup>  | Roger Butlin<sup>2,6</sup>

<sup>1</sup>IST Austria, Klosterneuburg, Austria

<sup>2</sup>Animal & Plant Sciences, Western Bank, University of Sheffield, Sheffield, UK

<sup>3</sup>Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway

<sup>4</sup>CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO, Laboratório Associado, Universidade do Porto, Vairão, Portugal

<sup>5</sup>CIIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal

<sup>6</sup>Marine Sciences, Tjärnö Marine Laboratory, University of Gothenburg, Strömstad, Sweden

## Correspondence

Anja M. Westram, IST Austria, Klosterneuburg, Austria.  
Email: anja.westram@ist.ac.at

## Funding information

H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 706376, 754411 and 797747; Vetenskapsrådet, Grant/Award Number: 217-2008-1719; Fundação para a Ciência e a Tecnologia, Grant/Award Number: PTDC/BIA-EVL/30628/2017; Natural Environment Research Council, Grant/Award Number: NE/K014021/1; H2020 European Research Council, Grant/Award Number: ERC-2015-AdG-693030-BARRIERS; Svenska Forskningsrådet Formas; European Regional Development Fund, Grant/Award Number: POCI-01-0145-FEDER-030628

## Abstract

Combining hybrid zone analysis with genomic data is a promising approach to understanding the genomic basis of adaptive divergence. It allows for the identification of genomic regions underlying barriers to gene flow. It also provides insights into spatial patterns of allele frequency change, informing about the interplay between environmental factors, dispersal and selection. However, when only a single hybrid zone is analysed, it is difficult to separate patterns generated by selection from those resulting from chance. Therefore, it is beneficial to look for repeatable patterns across replicate hybrid zones in the same system. We applied this approach to the marine snail *Littorina saxatilis*, which contains two ecotypes, adapted to wave-exposed rocks vs. high-predation boulder fields. The existence of numerous hybrid zones between ecotypes offered the opportunity to test for the repeatability of genomic architectures and spatial patterns of divergence. We sampled and phenotyped snails from seven replicate hybrid zones on the Swedish west coast and genotyped them for thousands of single nucleotide polymorphisms. Shell shape and size showed parallel clines across all zones. Many genomic regions showing steep clines and/or high differentiation were shared among hybrid zones, consistent with a common evolutionary history and extensive gene flow between zones, and supporting the importance of these regions for divergence. In particular, we found that several large putative inversions contribute to divergence in all locations. Additionally, we found evidence for consistent displacement of clines from the boulder-rock transition. Our results demonstrate patterns of spatial variation that would not be accessible without continuous spatial sampling, a large genomic data set and replicate hybrid zones.

## KEYWORDS

clines, genomic architecture, inversions, molluscs, speciation

## 1 | INTRODUCTION

Adaptive divergence is ubiquitous, makes a major contribution to biodiversity and might eventually lead to speciation (Hereford, 2009; Nosil, 2012). Yet our understanding of it is incomplete: How do selection,

genomic architecture and demography interact to establish and maintain separate populations, especially when counteracted by gene flow?

Divergence is facilitated by selection on individual loci that is strong relative to gene flow (Haldane, 1930), or by interactions between many loci under (potentially weak) divergent selection

Johannesson and Butlin contributed equally.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

(Barton, 1983; Flaxman et al., 2014). Divergence can be driven by a single divergent selection pressure or a combination of multiple different selection pressures that co-vary in space (Nosil et al., 2009). Balancing selection or positive selection, leading to sweeps shared across populations, might counteract divergence. Identifying the contributions of these various effects is challenging.

The genomic architectures expected to facilitate divergence with gene flow are those that reduce the potential for recombination between divergently selected alleles (Smadja & Butlin, 2011). This includes large-effect loci or pleiotropic loci affecting multiple traits under divergent selection, clustering of loci contributing to adaptation in the genome, and genomic rearrangements reducing recombination within populations, particularly inversions. Empirical work has provided evidence for all of these architectures (Barth et al., 2017; Joron et al., 2011; Pfeifer et al., 2018; Samuk et al., 2017).

Genome scans of divergently adapted populations have been used to detect loci and genomic architectures underlying divergence by identifying genomic regions with high levels of differentiation (Bonin et al., 2006; Nosil et al., 2008; Wood et al., 2008). However, peaks of differentiation are not sufficient to identify divergently selected regions, as various other processes can generate peaks and troughs in the differentiation landscape (Cruickshank & Hahn, 2014; Noor & Bennett, 2009; Ravinet et al., 2017). In addition, genome scans using spatially separate populations are often not suitable for understanding how genetic variation is organized across space, as they exclude areas where the environment changes. Consequently, it is difficult to infer which selection pressures drive genomic divergence, and to understand how selection interacts with other factors at environmental boundaries.

Hybrid zone analysis is an alternative approach that not only enables the identification of divergently selected genomic regions but also has the potential to teach us more about the interplay between different types of selection (e.g., divergent selection and balancing selection), the environment and demographic parameters (Harrison, 1993). It differs from standard genome scan approaches in that it includes samples across the spatial continuum of divergence. The hybrid zone approach has a long history in the study of adaptive divergence and speciation (Barton & Hewitt, 1985), and several recent studies have demonstrated its usefulness in combination with genomic data (Singhal & Bi, 2017; Souissi et al., 2018; Stankowski et al., 2017).

Hybrid zone analysis provides information on cline width (a measure of the strength of selection), symmetry and centre position (Barton & Gale, 1993). Coincidence of cline centres with individual environmental transitions might identify the sources of selection pressures. On the other hand, cline asymmetries or spatial shifts away from environmental boundaries that are consistent across loci could reflect population movement or between-population differences in dispersal or population density.

Independent of whether a standard genome scan or hybrid zone analysis is used, individual outlier loci may be spurious, i.e. not generated by divergent selection but by evolutionary or sampling stochasticity. Some outlier scan studies have therefore used “replicate”

sample pairs and identified strong candidates as those outliers detected in multiple pairs (Perrier et al., 2013; Wilding et al., 2001). With hybrid zone analysis, a similar logic can be applied (Christe et al., 2016; Zieliński et al., 2019). However, in this case one can ask not only whether outlier loci are shared among replicate pairs, but also whether other patterns, including cline widths, cline centres and allele frequencies at cline ends are shared across different zones, strengthening the evidence for nonrandom processes underlying those patterns.

In order to qualify as replicates, pairs of diverged populations must (a) be associated with similar environmental contrasts and show similar phenotypic patterns of divergence; and (b) be connected by ongoing or recently ceased gene flow, so that adaptive variants are likely to be shared among locations. Population pairs separated by large geographical or evolutionary distances are not considered replicates here, as they are expected to have diverged to some extent based on independent genetic variation (e.g., Conte et al., 2012).

Hybrid zone analyses using genome-wide data are still relatively rare (Ryan et al., 2017; but see Singhal & Bi, 2017; Souissi et al., 2018; Stankowski et al., 2017). While several hybrid zone studies have included repeated contacts between the same two taxa, the aim was usually a comparison between the different zones to determine the extent of parallelism, with the zones often deliberately chosen to show differences in age or other features (e.g. Beysard & Heckel, 2014; Schaefer et al., 2016). Very few studies have explicitly included replicate hybrid zones in the sense discussed above (but see Zieliński et al., 2019). Combining replicate hybrid zones and genome-wide data will make it possible to study the genomic basis of adaptation along with its spatial patterns with increased power. Here, we analyse genomic data from replicate hybrid zones in the intertidal snail *Littorina saxatilis*. These hybrid zones connect two partially isolated ecotypes (“Crab” and “Wave”), adapted to different shore habitats. While there is little evidence for intrinsic barriers to gene flow (e.g., Johannesson et al., 2020), there is clear evidence for divergent ecological adaptation. The Crab ecotype is typically associated with boulder fields, where crab predation is intense, while the Wave ecotype inhabits areas of bedrock that are subject to wave exposure (Johannesson et al., 2010). The ecotypes differ in various phenotypes, with the Wave ecotype being much smaller, more globose and bolder than the Crab ecotype (Butlin et al., 2014; Johannesson et al., 2010). Remarkably, individual hybrid zones span distances of only tens of metres (Hollander et al., 2015; Westram et al., 2018), and numerous hybrid zones can be found at small geographical scales (Hollander et al., 2015; Panova et al., 2006; Ravinet et al., 2016). This small-scale structuring is explained by the fact that *L. saxatilis* is a direct developer with internal fertilization and no larval dispersal stage and thus shows very limited lifetime dispersal (~1.5 m per generation; Westram et al., 2018), making this species ideal for studying replicate hybrid zones.

Cline analyses for a single *L. saxatilis* hybrid zone (Westram et al., 2018) revealed hundreds of “non-neutral” single nucleotide

polymorphisms (SNPs) potentially contributing to divergence or, more likely, linked to loci that do. Notably, the majority of these SNPs are located in three large genomic blocks probably representing chromosomal inversions (Faria, Chaube, et al., 2019). We also detected 14 additional putative inversions in the same geographical location, several of which show evidence of frequency differences between ecotypes (Faria, Chaube, et al., 2019). While work on a larger geographical scale indicates increased differentiation in the same genomic regions (Morales et al., 2019), the presence of high-linkage disequilibrium (LD) regions compatible with inversions was only assessed in the first studied hybrid zone (Faria, Chaube, et al., 2019; Westram et al., 2018). Thus, their occurrence in other locations remains to be tested.

It is known from experimental studies that crab predation and wave exposure are strong selection pressures affecting our study system (Johannesson, 1986; Pennec et al., 2017); however, it is difficult to measure these selection pressures in the field. Before our first hybrid zone study, we predicted that clines would centre at the transition from boulder field to bedrock, where we expected selection pressures to change. Surprisingly, we found that most cline centres were shifted several metres from this transition into the rock area in this first analysed location (Westram et al., 2018). We hypothesized that this shift was explained by the presence of a density trough in the rock area, which might attract clines. However, another possible explanation is that the selection pressures driving divergence in this system are in fact not tightly associated with the boulder–rock transition but with other environmental factors that show maximum change at a somewhat different position on the shore.

In this study, we analysed six additional hybrid zones on a small geographical scale to gain insights into how selection, demography and genomic architecture interact in a system of divergence with gene flow. We first ensured that the sampled hybrid zones constitute replicates, by testing whether they show repeatable phenotypic divergence and high levels of background genetic similarity. We then tested for replicated genetic patterns in the following aspects: (a) the occurrence of chromosomal inversions (so far only tested in a single geographical location); (b) the identity of outlier loci (loci putatively affected by divergent selection, here identified by cline analysis); (c) the genomic location of divergent regions (e.g., whether outliers repeatably cluster in inversions); and (d) cline parameters (allele frequencies at cline ends, cline widths and centres), particularly a displacement of cline centres from the boulder–rock transition.

## 2 | MATERIALS AND METHODS

We sampled six hybrid zones and performed phenotyping, genotyping, inversion identification and cline fitting for each zone using similar methods as for the first studied zone (“ANG”; Faria, Chaube, et al., 2019; Westram et al., 2018), which was also included here. We then compared patterns between different zones and performed

further analyses in order to understand causes of common patterns. All analyses were implemented using custom R scripts unless otherwise stated.

### 2.1 | Sampling

We sampled snails from three bays on different islands on the Swedish west coast: Ramsö (58°49′27.8″N, 11°03′45.3″E), Inre Arsklovet (58°50′00.5″N, 11°08′19.6″E) and Yttre Arsklovet (58°49′51.3″N, 11°07′59.0″E), here labelled CZA, CZB, and CZD, respectively (Figure 1b). Distances between islands are between 4.2 and 6.2 km, except for CZB and CZD (0.4 km apart). On each island, we sampled a transect from bedrock via a boulder field to the bedrock on the other side of the bay, so that each transect includes two Crab–Wave hybrid zones (labelled CZA\_left and CZA\_right, etc., as viewed from the sea; Figure 1a). Thus, the Crab ends of the two hybrid zones on the same island are directly adjacent and reflect the same population (Figure 1a,b).

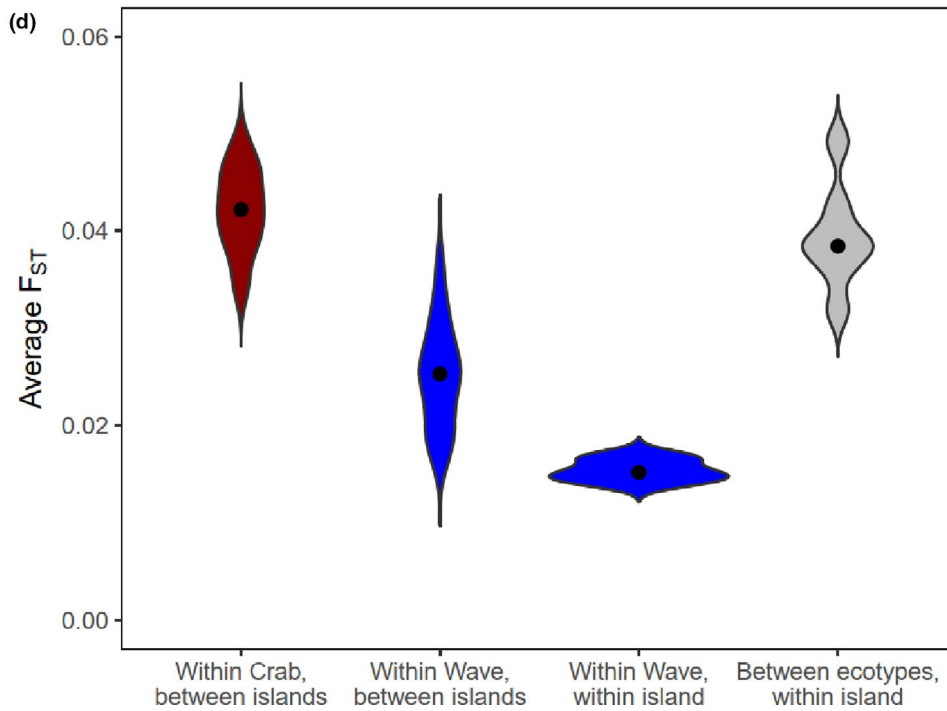
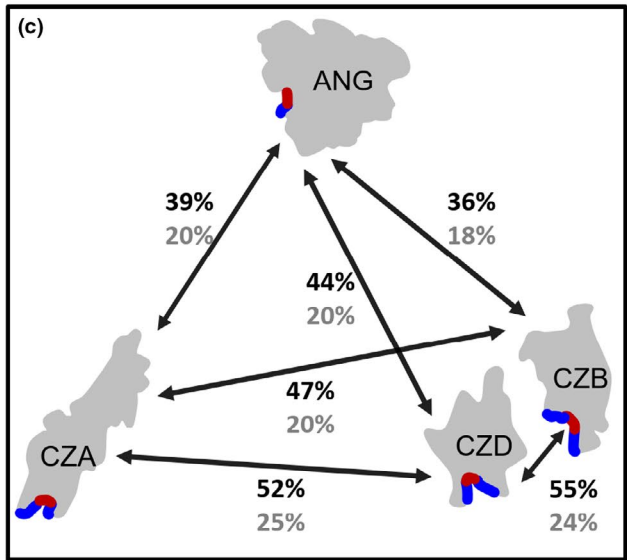
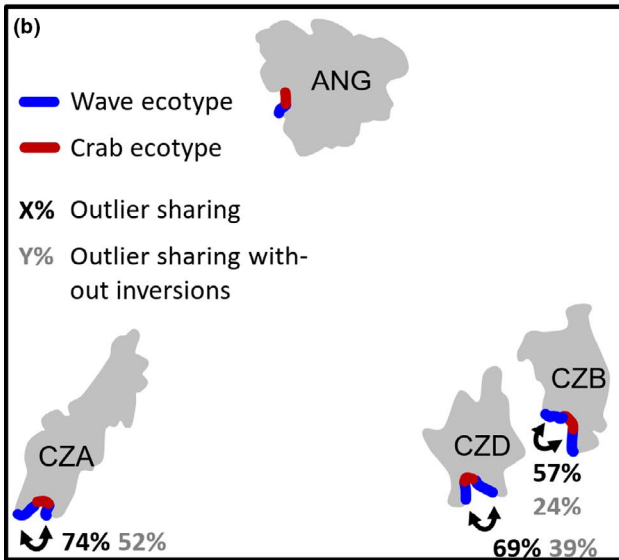
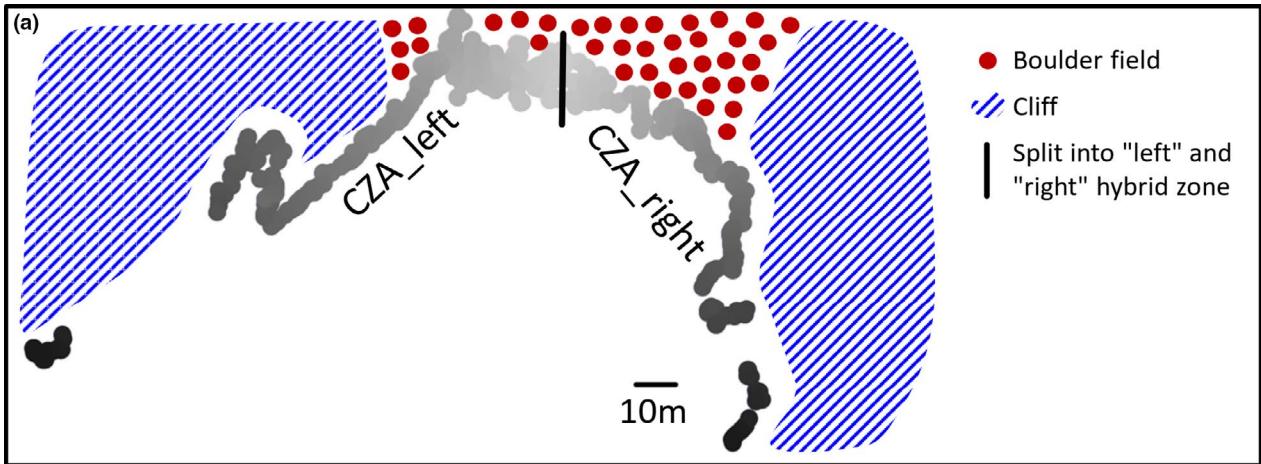
We sampled 600 snails per island, of which 382–384 adults were genotyped. The position of each snail in three dimensions was recorded using a Total Station (Trimble M3; error typically <1 cm). For each snail, shell shape (based on a geometric morphometric analysis) and size (length of the shell) were determined as in Westram et al., (2018). Snails were dissected, and tissue was stored in ethanol.

We recorded three environmental characteristics at ~1,000–2,000 points along each transect: in addition to the substrate (bedrock vs. boulders/stones), we included presence of barnacles (indicative of wave exposure) and furoid seaweed (indicative of weaker wave action and more shelter for shore crabs).

### 2.2 | Definition of snail positions and environmental transitions

Details for the definition of snail positions and environmental transitions along the shore are given in the Supporting Information. In particular, we obtained the position of each snail along a one-dimensional path in order to allow for one-dimensional cline analysis (Figure 1a). For that, we determined the position of each snail on a “least cost path,” minimizing path length while constraining movement mainly to areas of high snail density. The reduction to one dimension is justified as snails occupy only a narrow zone. We did this for each island and each zone (CZA/B/D\_left and CZA/B/D\_right) separately and ensured that each path has a starting position of 0 at the centre of the Crab habitat. For the single hybrid zone on the island of Saltö (ANG), we used the distances from Westram et al., (2018).

We defined the position of different environmental transitions on the path by taking weighted averages within 2 m of each snail for boulder vs. rock, presence of barnacles or furoid seaweed, shore height and a measure of topography, and then fitting step functions to these variables.



**FIGURE 1** (a) Example of the sampling strategy (island CZA). The boulder and rock habitats are shown with red and blue patterns. As we sampled two boulder–rock transitions, we divided the transect into two zones (“left” and “right”) by splitting it at the centre of the boulder habitat (black line). The sampling locations of snails are indicated by grey points, coloured depending on their distance from the centre of the boulder habitat along a one-dimensional least-cost path (from light to dark grey). Snails occur continuously along the shore: Sampling density does not necessarily reflect snail density, and gaps are due to sampling pattern, not distribution gaps. (b,c) Schematic maps (not to scale) of the four islands. For the current study, two hybrid zones were sampled on each of the three islands CZA, CZB and CZD, as shown for CZA in (a). On the island ANG, only a single hybrid zone was sampled for a previous study (Westram et al., 2018). Sharing of cline outliers between hybrid zones on the same island (b) and on different islands (c) is indicated, including SNPs in inversion regions in the analysis (black) or excluding these regions (grey). For the sharing between islands, this was first calculated separately for all possible pairwise combinations of zones and then averaged. (d) Average differentiation between same-ecotype samples from the same and different islands, compared to differentiation between ecotypes on the same island, using only samples distant from the boulder–rock transition, and excluding SNPs in inversion regions. The median is shown as a black point

### 2.3 | Genotyping and bioinformatics

DNA was extracted from foot tissue using a CTAB protocol (Panova et al., 2016). We then used a capture sequencing approach to obtain data for thousands of short (a few hundred bp) randomly distributed genomic regions. Capture sequencing technology allows for the targeting of specific genomic regions (so that the same regions are sequenced in each individual) by performing enrichment before sequencing. We used a total of 40,000 probes of 120 bp in length, identical to those described in Westram et al., (2018).

Library preparation, capture and paired-end sequencing (125 bp) were performed by Edinburgh Genomics. The resulting fastq files were processed as in Westram et al., (2018), with minor modifications, in order to obtain genotype data and allelic read depth data for each individual (see Supporting Information for details).

### 2.4 | Spatial structure and relationships between populations

To represent replicates, the different population pairs must be connected by a recent common history allowing for extensive sharing of adaptive alleles. To test this, we analysed the genetic structure within and between islands and ecotypes, using called genotypes. We calculated  $F_{ST}$  estimates and performed principal component analysis (PCA), excluding SNPs within inversions and outlier SNPs as well as individuals close to boulder–rock transitions (Supporting Information).

### 2.5 | Sharing of chromosomal rearrangements

For each island we identified potential chromosomal rearrangements from patterns of LD generated by reduced recombination in inversion heterokaryotypes (Faria, Chaube, et al., 2019; Kempainen et al., 2015) and also made use of our previously published genetic map of *Littorina saxatilis* (Westram et al., 2018). We followed the procedure of Faria, Chaube, et al., (2019), with minor modifications. Briefly, we used LD network analyses implemented in the LDna R package (Kempainen et al., 2015) to identify SNPs in relatively high LD, within each linkage group and for each island separately (see

Supporting Information for details). The positions of the outermost SNPs of each LD cluster in the *L. saxatilis* genetic map define the boundaries of putative inversions. A PCA was subsequently implemented using all SNPs within each LD cluster for each island separately, using the PCadapt R package (Luu et al., 2017). The aim was to confirm the presence of three separate groups of genotypes, as expected for inverted regions (or six groups for complex regions probably representing overlapping inversions; see below). Finally, we investigated whether the genomic coordinates of inversions detected on the different islands were similar, indicating that the same rearrangements are shared among islands. Strong LD observed between SNPs across large mapping regions generally points to the presence of inversions (Faria, Chaube, et al., 2019; Westram et al., 2018). However, some LD clusters have weaker support in terms of the percentage of variance explained by PC1 or the strength of LD than others. Although we refer to all these regions as inversions, some caution is needed until validation using independent information.

Due to the nonindependence of SNPs within inversions (caused by high LD), we performed some of our analyses without inversions or performed them twice (with and without inversions). All inversions appeared to be shared among all islands (see Section 3); consequently, for analyses without inversions, the same regions were excluded in all zones. When genomic coordinates of inversions varied slightly between islands, we used the outermost coordinates detected on any island (Table 1). When excluding inversions, we also completely removed linkage group 12, which showed unclear patterns suggesting sex linkage and multiple chromosomal inversions spanning a large proportion of the chromosome.

### 2.6 | Cline analyses

We fitted clines for individual SNPs inside and outside rearrangements, based on allelic depth data (read count for each of the two alleles) obtained for each individual. Using a maximum likelihood approach, we fitted a sigmoid cline model to these data, analysing each SNP in each hybrid zone separately, as in Westram et al., (2018) with minor modifications (Supporting Information).

Each cline fit provided five estimates: the position of the cline centre, the cline width, the allele frequency at each cline end,

TABLE 1 All inversions detected in this study

Linkage group	Cluster	ANG Start	ANG End	CZA Start	CZA End	CZB Start	CZB End	CZD Start	CZD End	Consensus start	Consensus end	Type	Comment	No cline
LG1	LGC1.1	0.0	2.1	0.0	2.1	0.0	2.1	0.0	2.1	0.0	2.1	Inversion		
LG1	LGC1.2	75.5	81.0	75.5	81.0	75.5	81.0	75.5	81.0	75.5	81.0	Inversion		ANG
LG2	LGC2.1	0.3	14.2	0.3	14.2	0.3	14.2	0.3	14.2	0.3	14.2	Inversion		
LG4	LGC4.1	1.0	1.5	1.4	1.5	1.0	1.5	1.2	1.5	1.0	1.5	Inversion		CZB_right
LG6	LGC6.1/2	0.0	29.3	0.0	29.9	0.0	28.7	0.0	29.3	0.0	29.9	Complex	Contains three haplotypes (a, b, c) due to two overlapping inversions	
LG7	LGC7.1	36.0	37.7	36.0	37.7	36.0 <sup>a</sup>	37.7 <sup>a</sup>	36.0	37.7	36.0	37.7	Inversion		
LG7	LGC7.2	42.1	51.4	42.1 <sup>a</sup>	51.4 <sup>a</sup>	42.1 <sup>a</sup>	51.4 <sup>a</sup>	42.1	51.4	42.1	51.4	Inversion		
LG9	LGC9.1	18.6	41.8	19.3	41.8	18.6	41.8	18.6	41.8	18.6	41.8	Inversion		CZB_right, CZD_right
LG10	LGC10.1	0.6	3.1	0.9	3.1	0.6	3.1	0.6	3.1	0.6	3.1	Inversion		CZB_left, CZD_left
LG10	LGC10.2	3.1 <sup>a</sup>	44.1 <sup>a</sup>	3.1	41.6	3.1	41.6	3.1	41.6	3.1	44.1	Inversion		CZA_left, CZA_right, CZB_right
LG11	LGC11.1	52.3	52.9	52.3	52.9	52.3	52.9	52.3	52.9	52.3	52.9	Inversion		ANG, CZA_right, CZB_left, CZB_right
LG14	LGC14.1/2	0.4	11.7	0.4	11.7	0.4	11.4	0.4	11.7	0.4	11.7	Complex	Contains three haplotypes (a, b, c) due to two overlapping inversions	
LG14	LGC14.3	11.7	34.9	10.2 <sup>a</sup>	34.7 <sup>a</sup>	10.2 <sup>a</sup>	34.9 <sup>a</sup>	17.4 <sup>a</sup>	34.9 <sup>a</sup>	10.2	34.9	Inversion		
LG17	LGC17.1	47.0	62.3	47.0	62.3	47.0	62.3	47.0	62.3	47.0	62.3	Inversion		

Notes: The start and end coordinates for each inversion on each island are shown in cM. Consensus coordinates are the outermost coordinates detected in any of the four islands. The zones where an inversion did not show a significant cline are indicated in the "No cline" column. Linkage group 12 is likely to contain inversions, but we refrain from listing them as the patterns are highly complex.

<sup>a</sup>Detected only after lowering the LDna or PCA detection thresholds.

and the variance in the data explained by the cline model. We studied the extent to which patterns in each of these estimates were shared among hybrid zones, as described in the following sections.

As SNPs within inversions show high levels of LD, each inversion can be treated as a single locus. We therefore also fitted clines for each inversion, using the inversion genotypes identified for each individual as input. We followed Faria, Chaube, et al., (2019) except that we always used the frequency of the arrangement that was more common in the Wave habitat than the Crab habitat. We recorded which arrangement this was in order to check for any differences in direction among hybrid zones (File S1). To test for cline displacement, we repeated the cline fitting with centres that were constrained to be located at the boulder–rock transition. We considered the unconstrained fit better if the difference in Akaike's information criterion (AIC) was  $>4$ . We considered there to be evidence for clinal variation if either fit had an AIC more than 4 below the AIC for a constant arrangement frequency.

We found two complex inversions (LGC6.1/2 and LGC14.1/2) that each had three haplotypes segregating in our samples (“a,” “b” and “c”). One haplotype was consistently at low frequency in each case and did not show clinal variation. The other two showed reciprocal clines. Therefore, we report clines only for the common haplotype with a higher frequency in the Wave habitat in each case.

In order to test for repeatable phenotypic patterns, clines were also fitted to shell length and shape data (Supporting Information). We compared the best-fitting cline model with an otherwise similar model in which the cline centre was constrained to the boulder–rock transition, providing a test for displacement comparable to the test used for inversion clines.

## 2.7 | Outlier identification and sharing

The presence of a cline alone is not informative regarding whether a locus is affected by divergent selection, as clines can be generated by isolation-by-distance or a genome-wide barrier to gene flow. Here, we use the proportion of the variance in the SNP genotype data that is explained by the best-fitting cline model (*var.ex*) to identify SNPs not consistent with neutrality: in simulations taking into account the general barrier to gene flow, neutral loci that were not linked to selected loci very rarely achieved high values of *var.ex* (Westram et al., 2018). Due to the large number of samples from multiple hybrid zones in this study, detailed simulations of the system were not feasible; however, based on our earlier work, we used the *var.ex* as a measure to infer non-neutrality, and set an arbitrary threshold at the 99% quantile for each hybrid zone. This analysis provided a set of “outlier SNPs” that are likely to be influenced by direct selection or selection at linked loci.

We then investigated to what extent outliers were shared between hybrid zones on the same and on different islands by

calculating the proportion of outliers in zone X that were also outliers in zone Y. We repeated the analysis excluding inversion regions in order to understand to what extent patterns of sharing were driven by the presence of shared inversions.

## 2.8 | Shared genomic architectures of divergence

We investigated the genomic distribution of outlier SNPs, using a previously published linkage map based on the same capture sequencing probes (Westram et al., 2018). First, for each zone separately, we tested whether outliers were significantly enriched in each of the inversions, using permutation. For each inversion and zone, we randomly drew SNPs from the collinear background (i.e., all genomic regions not containing an inversion on any island) until the number of SNPs drawn equalled the number of SNPs in the inversion. The inversion was considered significantly enriched for outliers if it contained a larger number of outliers than found in 95% of 1000 permutations. In cases of complex inversions, a single inversion region encompassing all overlapping inversions was considered.

Second, for each zone separately, we investigated whether outliers outside inversion regions were enriched in other low-recombination regions. Regions of low recombination are likely to contain more SNPs for a given genetic map interval, because their physical length is greater. Therefore, we used the number of non-outliers (on average 99% of all SNPs) in each 1-centimorgan (cM) interval as a proxy for recombination rate and correlated this with the proportions of outliers in these intervals.

## 2.9 | Sharing of cline patterns

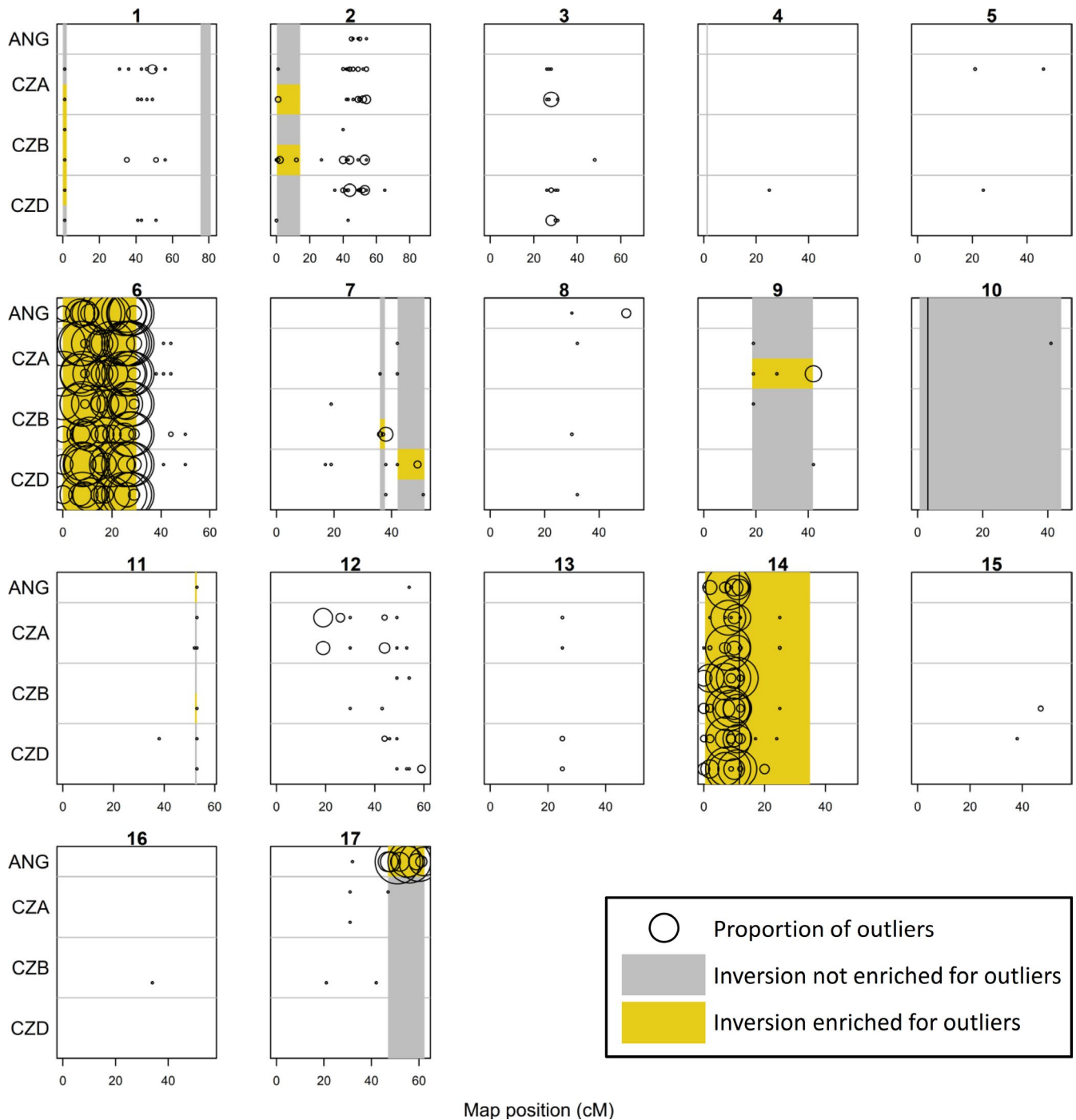
In our previous study (Westram et al., 2018) we found that almost none of the SNPs and inversions putatively under divergent selection showed fixation at cline ends. Here, we investigated whether the same pattern can be found across multiple replicate hybrid zones. For that, we compared the distribution of allele frequencies (for both SNPs and inversions) obtained from the cline fitting for each zone.

We also tested whether cline widths were similar across zones by correlating the values found in different zones.

For cline centres, we investigated whether SNP clines showed shifts away from the boulder–rock transition. We first tested whether SNP cline centres are correlated across hybrid zones, indicating that cline centres of individual loci show consistent shifts in different hybrid zones. We also tested for overall shifts of cline SNP centres by investigating whether the distribution of cline centres was shifted from the boulder–rock transition, using a Wilcoxon signed rank test for each hybrid zone separately.

We then investigated whether cline shifts could be explained by any of the alternative environmental transitions that we defined.





**FIGURE 2** Distribution of outlier SNPs on the genetic map (each subplot shows a linkage group). Each row represents one hybrid zone, labelled on the left; on the islands CZA, CZB and CZD, the “left” hybrid zone is shown above the “right” zone. Each point represents a 1-cM interval where at least one SNP (in a given hybrid zone) is an outlier. The size of the point reflects the proportion of SNPs at a given position that are outliers (the smallest points reflect 2% or less, the largest 40% or more; between these values, circle size scales with the proportion of outliers). Only map intervals containing 10 or more SNPs are shown. Shaded areas represent putative inversion regions. Note that linkage groups 10 and 14 each contain two adjacent inversions; these are separated by a black line. Linkage group 12 is likely to contain inversions but we refrain from highlighting individual inversions as the patterns are highly complex

However, as all other environmental variables were extremely noisy and the variance explained by the step model was often low, instead of presenting a formal analysis we only discuss our results verbally.

As one hypothesis that could explain consistent cline shifts across loci and hybrid zones is a difference between ecotypes in dispersal and/or population density, we tested for the extent of spatial structure within ecotypes by performing an isolation-by-distance analysis (see Supporting Information).

### 3 | RESULTS

#### 3.1 | Analysis of phenotypic clines

All sampled locations showed clear phenotypic clines from large, elongated shells in the boulder field to small, globose shells on the rock (Tables S1 and S2, Figures S1 and S2). Sizes and shapes at cline ends were similar for all contact zones (Tables S1 and S2), as required for replicate zones. The phenotypic cline centres were displaced into the rock habitat in all cases except for the shape cline centre in CZA\_left and CZA\_right (Tables S1 and S2); this displacement was significant in most cases (Tables S1 and S2). The extent of displacement varied from a few metres to more than 20 m for some shell size clines; in all zones the size cline was more displaced than the shape cline.

#### 3.2 | Spatial structure and relationships among populations

After excluding SNPs located in putative inversion regions in any of the hybrid zones, between 8,000 and 20,000 SNPs (depending on the focal population pair) remained for an analysis of population structure.  $F_{ST}$  estimates demonstrated low differentiation between same-ecotype samples from different zones (Figure 1d). Consistent with the patterns revealed by  $F_{ST}$ , the first principal component of a PCA on all putatively neutral SNPs passing filters (excluding inversions, 7,976 SNPs) splits the two ecotypes (Figure S3; however, note that this axis explains only 2.8% of the total variance).

Together these results demonstrate high similarity of same-ecotype samples from different bays (or different sides of the same bay), as required for replicate hybrid zones.

#### 3.3 | Sharing of chromosomal rearrangements

We performed *de novo* detection of inversions for CZA, CZB and CZD, where the presence of inversions has not been tested before. In total, we found evidence for 16 putative chromosomal inversions in 10 out of the 16 linkage groups investigated (excluding LG12), including one (LGC10.2) not described by Faria, Chaube, et al., (2019); Figure 2, Table 1). In LG6 and LG14, we detected two overlapping inversions in each case (Table 1), and report results for the combined inversion regions (LGC6.1/2 and LGC14.1/2).

Despite some slight variation in coordinates, most inversions were consistently detected in an independent manner across all islands. The only three exceptions correspond to three clusters that did not pass one of our detection thresholds in at least one island. However, LD patterns, coordinates or PCA groupings for these clusters strongly suggest the presence of the same inversion across all islands (Table 1), so that we consider all inversions as shared among all zones in the following analyses. All inversions except for LGC11.1 showed significant clinal change in most zones (Table 1; File S1).

#### 3.4 | Sharing of outliers

For the whole genome (i.e., inside and outside inversions) 48,672 SNPs passed filtering and could be used for cline fitting in all locations. In the different zones, between 32% and 59% of SNPs showed nominally significant clines. Cline "outliers," likely to be affected by divergent selection, were defined based on the variance explained by the best cline model. Sharing of outliers between zones was generally high (Figure 1b,c), and higher between zones on the same island, which share the same Crab population. Sharing between islands was more limited but still much larger than the random expectation (five shared outlier SNPs expected based on sampling from a hypergeometric distribution; the probability of sharing 36% or more is essentially 0). Each zone had some private outliers that were not shared with any other zone, but the proportion of these outliers was typically small (CZA\_left: 15%; CZA\_right: 16%; CZB\_left: 15% CZB\_right: 26%; CZD\_left: 9%; CZD\_right: 13%; ANG: 40%; the larger proportion in ANG is partly explained by the fact that no second geographically close zone on the same island was sampled). In total, 87 outliers were shared between all zones.

When inversion regions were excluded from the analysis (leaving 33,226 SNPs in the analysis), outlier sharing was lower (Figure 1b,c; grey numbers), but still much higher than the random expectation (the probability of sharing 18% of outliers, or more, by chance is essentially 0).

#### 3.5 | Shared genomic architectures of divergence

All hybrid zones showed similar genomic distributions of outliers (Figure 2). Most outliers clustered in a few genomic regions, many of them coinciding with inversion regions. Some inversion regions showed a significant enrichment of outliers (yellow shading in Figure 2); most notably, LGC6.1/2 and LGC14.1/2 were enriched for outliers in all zones. An inversion in LG17 showed a clear enrichment of outliers in ANG, but in none of the other zones. While we base our analyses on cline outliers,  $F_{ST}$  estimates were elevated in the same genomic regions (see Figure S4 for Manhattan plots).

Shared outliers were more commonly located in inversion regions than nonshared outliers. For example, 74% of SNPs that were outliers in only a single zone were located in an inversion region ( $n = 646$ ), while 97% of outliers shared among exactly five hybrid zones were located in an inversion region ( $n = 93$ ). All outliers shared among exactly six zones were located in an inversion region ( $n = 76$ ), and all outliers shared among all seven zones ( $n = 87$ ) were located in the LG6 or LG14 inversions. However, some inversion regions (e.g., a large putative inversion on LG10) contained very few outliers in general, and no shared outliers.

A large proportion of outliers were located in inversion regions, but SNPs in an inversion are in high LD and this might distort patterns. Therefore, we repeated the analysis of outlier distribution without inversion regions (Figure S5). While sharing of outliers was considerably lower when inversion regions were

excluded (Figure 1b,c, grey numbers), outlier SNPs outside inversion regions tended to cluster near linkage group centres in all zones (Figure S5), particularly in the centre of LG2. This pattern suggests an association with centromeres (the location of which is not currently known), which are regions of low recombination in many organisms. We therefore hypothesized that outliers tend to be associated with regions of low recombination. We tested this by comparing the numbers of outliers with the numbers of nonoutliers per 1-cM interval. If outliers are randomly distributed across the genome, independent of recombination rate, the slope of this regression is expected to be 0.01 (as 1% of SNPs are outliers by definition). However, if outliers are predominantly located in regions of low recombination where there are many SNPs per cM, the slope is expected to be steeper because outliers are expected to be overrepresented in these intervals. Outliers were indeed overrepresented in intervals with large numbers of non-outlier SNPs (Figure 3).

### 3.6 | Sharing of allele frequency and inversion frequency patterns

For all SNPs with significant clines, allele frequencies at cline ends were estimated. These were strongly correlated between ecotypes in all zones (Figure 4). Outlier SNPs tended to show stronger differences between ecotypes than nonoutliers, as expected. However, differential fixation between ecotypes was rare or absent in all zones; even among outlier SNPs, the great majority were polymorphic at both cline ends. Some inversions showed strong frequency differences between ecotypes, particularly LGC6.1/2 and LG14.1/2, consistent with the observed enrichment in SNP outliers (Figure 4, coloured symbols). However, fixation was rare also for inversions, and differential fixation was never observed (Figure 4).

Allele frequencies for a given ecotype tended to be similar across zones: all pairwise correlations of allele frequencies were positive and statistically significant (all Pearson correlation coefficients ( $r$ )  $\geq .65$ ; all  $p < .001$ ; Tables S3 and S4). Figure 4 shows that inversion-cline end frequencies were also similar between zones.

### 3.7 | Sharing of cline widths

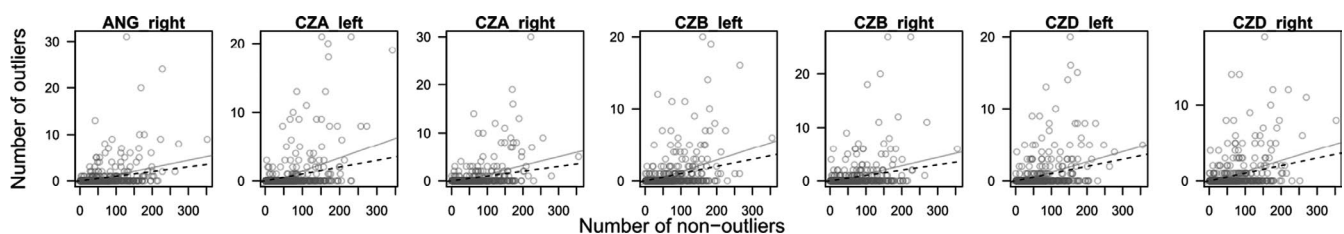
In contrast to allele frequency estimates, the cline width estimates were not (or very weakly) correlated between the two zones on the same island (CZA: Pearson  $r = .02$ ,  $p = .03$ ; CZB:  $r = .02$ ,  $p = .24$ ; CZD:  $r = 0$ ,  $p = .7$ ; Figure S6). A correlation would have been expected at least for shared outlier SNPs if selection was roughly similar in strength. Cline widths were also frequently estimated to be at the lower or upper boundary allowed in the estimation procedure, indicating that many of these estimates are unreliable (Figure S6). We therefore do not discuss cline widths further in this study.

### 3.8 | Sharing of cline centre shifts

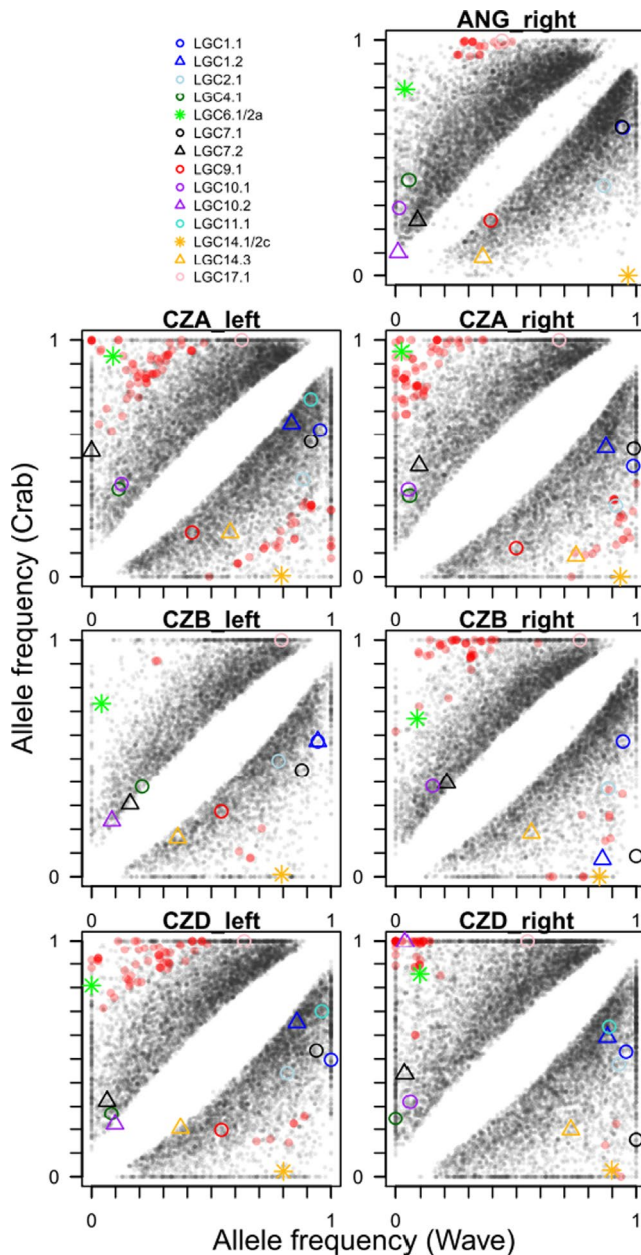
We tested for consistent cline shifts away from the boulder-rock transition across zones based on patterns at individual SNPs. Correlations were either not significant or very weak (CZA: Pearson  $r = .02$ ,  $p = .06$ ; CZB:  $r = .06$ ,  $p < .001$ ; CZD:  $r = .05$ ,  $p < .001$ , Figure S7), indicating that shift patterns differ between zones.

To test for a general pattern of cline shifts across loci, we compared the overall distribution of cline centres (SNPs excluding inversion regions) with the boulder-rock transition and found highly significant displacements for all zones (Wilcoxon signed rank test; all  $p < .001$ ; Figure 5); the median displacement was between 1.7 m (CZA\_left) and 27.0 m (CZB\_right). Similarly, for all inversions with significant clines, a model with displacement provided the best fit. The shift away from the boulder-rock transition was into the rock area for both SNPs and inversions. In many cases there was a long tail of the distribution in the rock area, with many clines displaced for tens of metres (Figure 5).

We investigated whether one of our other environmental variables could provide an explanation for the shifted clines. Fucoïd seaweed did not show a consistent change between the two sides of the transect (Figure S8, Table S5). It is therefore unlikely to explain the consistent cline shifts. The environmental transitions for topography (steepness of the shore) and shore height were sometimes shifted to the right and sometimes to the left of the boulder-rock transition, and both showed very noisy patterns; therefore, they are also unlikely to explain cline shifts (Figure S8, Table S5). In contrast, the



**FIGURE 3** Relationship between the number of outliers and the number of nonoutliers in 1-cM intervals (excluding inversions). The relationship is expected to be steeper if outliers are predominantly located in low-recombination regions (i.e., intervals with many nonoutliers). A linear model fitted to the observed data is shown as a grey line; the expectation (based on 1000 permutations each) is shown as a black dashed line. The difference in slope was significant in all cases, i.e. the slope obtained for the observed data was above the 95% quantile of the distribution of permutation slopes



**FIGURE 4** Fitted allele frequencies at the Wave and Crab ends of the cline for all SNPs with significant clines outside inversion regions (grey points: putatively neutral SNPs; semi-transparent red points: outlier SNPs), and for inversions with significant clines. Cline fitting was not performed for SNPs with an allele frequency difference of less than 0.1 between Crab and Wave; this explains the lack of points around the 1:1 line. For complex inversions (LGC6.1/2 and LGC14.1/2) only the arrangement with the largest change between Crab and Wave is shown

prevalence of barnacles systematically increased on the rock side, and the transition was always on the rock side of the boulder-rock transition (Figure 5; Figure S8, Table S5). However, the variation in the barnacle data was usually considerably higher than the variation in substrate (Figure S8, Table S5); we therefore refrain from any formal test of associations with cline centres.

Another partial explanation for consistent cline shifts could be a difference in population density and/or dispersal between ecotypes. Higher dispersal, higher population density or both in the Crab ecotype would generate overall asymmetric gene flow into the Wave habitat, displacing the clines from the environmental transition where selection changes. We addressed this possibility by performing an isolation-by-distance analysis within ecotypes, using only individuals far from the contact. The results did not show an increase in genetic distances with geographical distances in the Crab populations, but did show evidence of IBD in several of the Wave populations, consistent with higher density or dispersal in Crab snails (Figure S9).

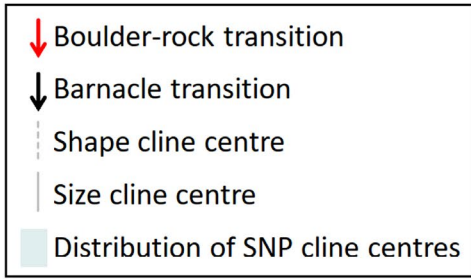
## 4 | DISCUSSION

Hybrid zones are great natural laboratories offering many insights into evolutionary processes (Abbott et al., 2013; Hewitt, 1988). Tracking how allele frequencies change over space may increase the power to detect divergently selected loci and provide information about selection pressures and genomic architectures that is not always accessible when relying on pairs of population samples. However, like other studies of genomic divergence, hybrid zone studies often suffer from an inability to distinguish between chance effects (either due to evolutionary stochasticity [drift] or due to sampling effects) and idiosyncrasies of specific sampling sites on the one hand, and more general patterns on the other hand. Here, we use seven hybrid zones on a small geographical scale as “replicates” of adaptive divergence in order to make this distinction.

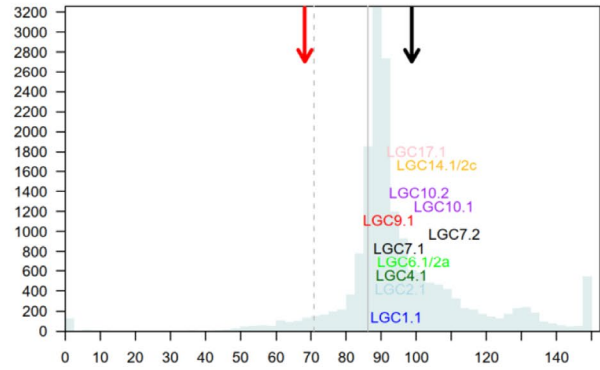
### 4.1 | Replicate hybrid zones

Studies comparing multiple hybrid zones have mostly focused on their differences, for example analysing geographically distant zones, zones of secondary contact of different age or ecologically dissimilar zones. These differences are often associated with differences in the set of genomic regions showing clinal change, the extent of introgression, hybrid zone structure (e.g., mosaic vs. continuous; unimodal vs. bimodal) or patterns of cline displacement (Arntzen et al., 2014; Beysard & Heckel, 2014; Dufresnes et al., 2015; Schaefer et al., 2011). In contrast, here we use multiple hybrid zones that are as similar as possible to serve as replicates, a related approach to that previously employed by Zieliński et al., (2019), who sampled replicate transects of the same newt hybrid zones. To serve as replicates when studying the genomic basis of divergence, the different hybrid zones must show similar patterns of phenotypic divergence and must be connected by recent gene flow, so that adaptive variants can readily be shared. If this is not the case, location-specific adaptations and different genomic solutions to obtaining the same adaptation may be common (Conte et al., 2012).

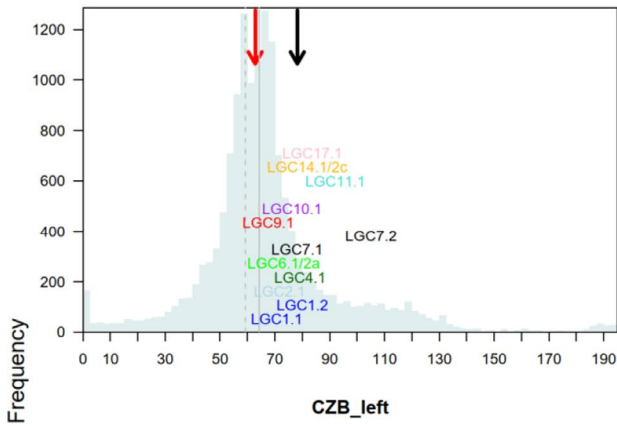
Here, the analysis of phenotypic clines shows that shell shape and size are similarly diverged in all studied zones, suggesting that



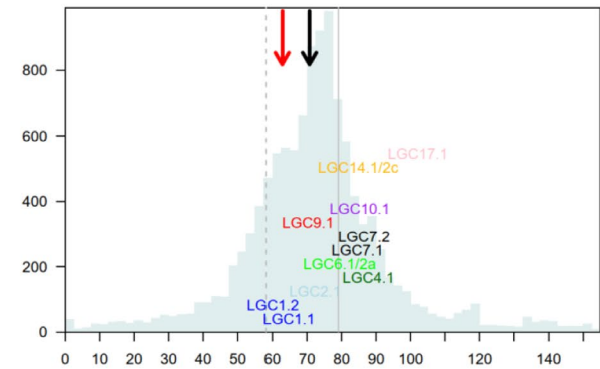
ANG\_right



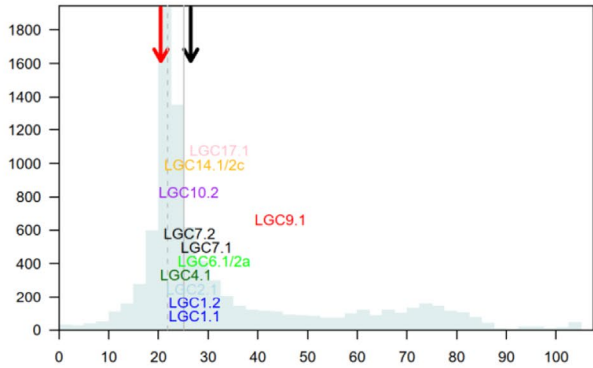
CZA\_left



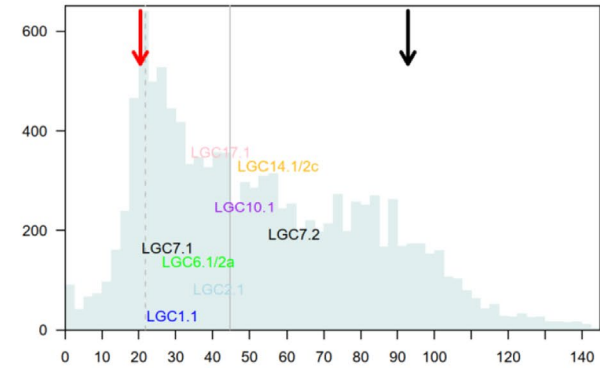
CZA\_right



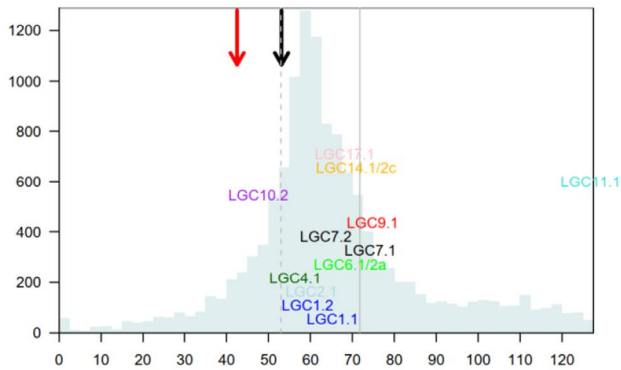
CZB\_left



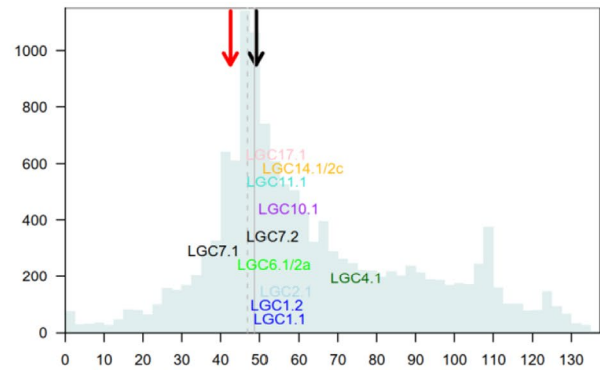
CZB\_right



CZD\_left



CZD\_right



Position along shore (m)

**FIGURE 5** Distribution of cline centres along the shore, from the Crab habitat (left) to the Wave habitat (right). Arrows indicate two environmental transitions (red: boulder–rock; black: barnacle). Phenotypic cline centres are shown as grey solid (size) and dashed (shape) lines. Histograms indicate the position of individual cline centres for SNPs, excluding those in inversion regions. Horizontal positions of inversion names indicate locations of inversion cline centres (left-aligned). Note that a small number of inversion cline centres falling outside the sampling range are not shown (but see File S1 for all centre locations)

selection pressures at least on these traits are similar. Our PCA and  $F_{ST}$  results show that differentiation between ecotypes within hybrid zones and differentiation within ecotypes between hybrid zones is at a similar level, indicating high gene flow within ecotypes between zones and/or a recent common evolutionary history. This system is therefore well suited for identifying repeatable patterns.

Our study builds on the premise that patterns replicated among these geographically close hybrid zones are less likely to be spurious than patterns observed in a single zone. For example, the repeated shifts of both phenotypic and genetic clines into the rock area that we observe are unlikely to be due to chance effects because cline shifts happen independently on each island. One potential criticism of our approach might be that, with geographically close hybrid zones, where gene flow between zones within ecotypes is high, chance differences between ecotypes could be shared among some hybrid zones. This could potentially generate false-positive outliers that are shared among islands. However, we note that levels of sharing are still expected to be higher for selected than neutral SNPs, making the most-shared outliers the strongest candidates. In addition, alleles reaching high-frequency differences due to drift may not show steep clines between ecotypes, so that they are less likely to be detected as outliers with our cline analysis approach in the first place. In other words, the general criticism of genome scans that outliers may not be functionally associated with ecotype divergence still applies to our analysis; however, an analysis including replicates represents an improvement compared to a single-location analysis as confidence is increased for outliers shared among multiple zones.

In the following, we discuss patterns shared between replicates and their implications. We also highlight some nonshared patterns and possible explanations.

## 4.2 | Sharing of inversion polymorphisms

Segregating chromosomal rearrangements are common in many species and have frequently been shown to contribute to divergence (Wellenreuther & Bernatchez, 2018). Inversions could provide an efficient means of transporting sets of adaptive alleles between locations, contributing to rapid adaptation (Faria et al., 2019), for example in postglacial populations such as Swedish *Littorina saxatilis*. If rearrangement polymorphisms are maintained by selection, they should be shared among replicate hybrid zones. There is strong evidence that this is the case in our study system: When analysing the four islands separately, we detected largely the same candidate regions in every case. While the coordinates varied slightly and are relatively imprecise due to the limited resolution of our sequencing data and genetic map, it seems very

likely that the same rearrangements segregate in all studied hybrid zones, and do so at similar frequencies. These findings therefore confirm the presence of rearrangement polymorphism detected in a single location in earlier studies (Faria, Chaube, et al., 2019; Westram et al., 2018). They are also consistent with work demonstrating large blocks of high differentiation at much larger geographical scales in some of these inversion regions (Morales et al., 2019), even if direct evidence for the presence of inversions was not possible due to the use of pool-seq data. Many of the inversions show clinal changes with the same directionality in multiple hybrid zones (File S1). Below, we discuss the role of these shared inversions in divergence.

## 4.3 | Sharing of outliers

While it has been shown that outliers detected in genome scans can be false positives (i.e., not indicative of divergent selection) for various reasons (Ravinet et al., 2017), we here combine two approaches to obtain an especially reliable set of outliers: hybrid zone analysis and the use of replicates. As expected for a study with very closely related replicates and relatively large sample sizes, a large proportion of cline outliers were shared among at least two islands, and we therefore obtain a large set of strong candidate regions.

It is possible that some shared outliers are not related to adaptive divergence between ecotypes, but instead represent old intrinsic incompatibilities that became coupled with extrinsic barriers (Bierne et al., 2011). However, there is little evidence for intrinsic incompatibilities in this system so far (Johannesson et al., 2020), and it is clear that numerous traits change across our hybrid zones. Therefore, while we certainly cannot exclude that some outlier regions represent intrinsic barriers, it seems likely that most shared outliers are associated with genomic regions underlying local adaptation. Further work will focus on annotating these regions and finding associations with phenotypes using mapping studies. However, as discussed below, many of the outliers are located in putative regions of low recombination (inversions or other low-recombination regions); therefore, identifying individual genes underlying divergence will be challenging (Cheng et al., 2012).

While the majority of outliers in each zone are shared with at least one other hybrid zone, there are also a small number of zone-specific outliers. These outliers might be associated with zone-specific selection pressures. The sampled zones differ in some environmental aspects including the height and slope of the cliff and the orientation, which could modify the strength of crab and wave selection pressures and introduce additional local selection pressures. However,

it is currently impossible to distinguish loci under location-specific divergent selection pressures from spurious outliers that are generated by drift or sampling (false positives), or outliers that are under selection in multiple locations but crossed our arbitrary threshold in only one location (false negatives in other zones). The latter might be the case, for example, for outliers in the linkage group 17 inversion (see following section).

It is notable that the proportion of shared outliers clearly decreases when inversions are excluded from the analysis; in particular, the number of outliers shared between all islands decreases substantially—see Morales et al., (2019) for a similar result on a larger geographical scale. This is one piece of evidence for the role of inversions in genomic divergence in this system.

#### 4.4 | Shared genomic architecture of divergence

Numerous theoretical and empirical studies have emphasized the role of recombination reduction in divergence with gene flow. Our findings are consistent with this, as we find clear shared patterns of genomic architectures. First, outliers in all zones cluster in inversion regions, and arrangement frequencies differ strongly between ecotypes for some inversions; second, outliers outside inversions in all zones cluster near chromosome centres and are potentially associated with low recombination rates. We discuss these patterns in turn.

The role of chromosomal rearrangements in adaptive divergence and speciation is becoming increasingly clear. Theory predicts that two populations containing different inversion karyotypes will benefit from reduced recombination, allowing for the maintenance of locally adapted combinations of alleles (Kirkpatrick & Barton, 2006). Inversion polymorphisms associated with adaptive divergence and/or incipient speciation have been found in multiple systems, including lake-stream divergence in sticklebacks (Roesti et al., 2015), divergence between migratory and stationary ecotypes in cod (Berg et al., 2016) and divergence in life-history traits in monkeyflowers (Twyford & Friedman, 2015). We find that multiple inversions show clinal change between ecotypes. However, as for SNPs, clinal change alone is not sufficient evidence for divergent selection, even when shared across replicate zones, as the general barrier to gene flow between ecotypes may generate clinal changes at neutral loci. However, we also find that arrangement frequencies for some inversions show very strong and consistent differences between ecotypes (LGC6.1/2 and LGC14.1/2 are close to differential fixation in some of the zones), a pattern that seems unlikely under neutrality. Consistent with this, we find that outliers are overrepresented in inversion regions, potentially indicating divergent selection on the inversions (but see caveats discussed at the end of this section). The directionality of divergence between the Crab and Wave ecotypes is almost always the same, consistent with one arrangement containing alleles adapted to the Wave environment and the other arrangement containing alleles adapted to the Crab environment. The evidence is particularly strong for the inversions on linkage groups 6

and 14, which we already detected in previous studies in the location ANG (Faria, Chaube, et al., 2019; Westram et al., 2018); remarkably, evidence for high levels of differentiation in these genomic regions has been found across Europe (Morales et al., 2019). It therefore seems likely that these two inversions, and possibly others, make a strong contribution to ecotype divergence in this system. As *L. saxatilis* ecotypes in Sweden have diverged only recently, possibly when crabs recolonized the area a few thousand years ago (Butlin et al., 2014), segregating inversions with sets of co-adapted alleles could have been important drivers of rapid adaptation. These inversions could have been maintained in northern populations during glaciations (e.g., by balancing selection, see below) or they could have introgressed from southern populations (where crabs might have been present throughout glaciations or might have recolonized the area earlier than in the north).

We also find some inversions that seemingly contribute to divergence only locally. Most notably, the inversion on LG17 contains a large number of outliers in ANG, but none in any of the other zones. However, on closer inspection this inversion shows similar frequency patterns in all zones: it is near fixation in the Crab ecotype, but polymorphic on the Wave side, where many heterokaryotypic individuals occur. Frequency differences are slightly more pronounced in ANG, leading to the large number of outliers (as outliers in the inversion are in LD, a sudden increase in outlier number with a small change in arrangement frequency is not surprising). It is therefore feasible that this region is under similar selective pressures in all zones, but that slight chance differences, slight differences in the strength of selection or variation in power among the analyses of the different zones cause the observed pattern.

A general caveat is that inversions generate strong LD and SNPs within inversions do not evolve independently. Indicators of divergent selection ( $F_{ST}$ , cline slope, extent of sharing) at SNPs linked to causal variants are usually weakened by recombination. This weakening effect is diminished substantially in inversions, where strong LD can be maintained across tens of megabases. In contrast, in analyses of reduced-representation data sets, almost all SNPs in collinear regions experience this weakening effect because it is unlikely that directly selected variants are captured. As a result, inversions may stand out regarding indicators of divergent selection even if selection on them is not stronger than selection on many variants in the collinear genome. While it seems clear that inversions are important in this system, it is therefore currently impossible to estimate their relative importance compared to other genomic regions, to pinpoint exactly the location and number of selected loci, or to determine whether selected loci are associated with intrinsic or extrinsic barriers.

As the number of outliers in our analysis was fixed (to 1% of the total number of SNPs), inversions might pocket the majority of outliers, effectively masking patterns in the rest of the genome. We therefore repeated the outlier analysis without inversion regions and found that outliers are also particularly common near chromosome centres. We hypothesized that this represents an accumulation in regions of low recombination, as observed in other systems

(Duranton et al., 2018; Renaut et al., 2013; Roesti et al., 2012). Under divergence with gene flow, outliers are expected to cluster in low-recombination regions, both because low recombination prevents the disruption of favourable allele combinations and because signatures of selection spread further along the chromosome. Empirically, an accumulation of outliers in low-recombination regions has been demonstrated, for example, in sticklebacks, especially when populations are connected by gene flow (Samuk et al., 2017). We used a proxy of the recombination rate based on the *L. saxatilis* genetic map and indeed found a negative relationship between recombination rate and the proportion of outliers. However, we note that we currently cannot distinguish between a larger number of selected loci located in low-recombination regions and an accumulation of outliers simply because signatures of selection are more likely to be detected where recombination is low.

#### 4.5 | Sharing of allele frequency patterns

Considering the geographical proximity, small  $F_{ST}$  values between zones and high levels of outlier sharing, it is not surprising that allele frequencies are correlated between samples of the same ecotype from different islands. Interestingly, fixed differences between ecotypes are rare in all zones (and completely absent in some), for SNPs as well as inversions, including those that show evidence of divergent selection. For SNPs, the pattern might be explained by the fact that we target a relatively small proportion of the genome, and most outlier SNPs are unlikely to be under direct divergent selection but are rather linked to selected loci, as discussed above. As an alternative, but not mutually exclusive explanation, if selection is stabilizing within ecotypes and focused on highly polygenic traits, fixation may not be expected even at causative loci.

The pattern of nonfixation is more surprising for inversions, which are likely to be under direct divergent selection because they contain major-effect loci or multiple loci contributing to polygenic traits. As we previously argued for the location ANG (Westram et al., 2018), patterns of nonfixation for divergently selected inversions could be explained by balancing selection acting on these inversions at the same time, for example due to the presence of recessive deleterious alleles (causing heterokaryotype advantage). Such a combination is feasible given theory that emphasizes both the propensity of inversions to contribute to divergence by generating blocks of locally adapted alleles (Kirkpatrick & Barton, 2006) and the risk for inversions to accumulate recessive deleterious mutations that ultimately lead to overdominance (Faria, Johannesson, et al., 2019; Kirkpatrick, 2010; Ohta, 1971).

#### 4.6 | Sharing of cline centre shifts

In all hybrid zones we found a consistent shift of the median cline centre into the rock area. This shift is consistent with previous work (Johannesson et al., 2020), where we observed that the hybrid

index does not abruptly change at the habitat transition, but rather changes mostly in the rock (Wave) area, often gradually. The centre shift is not only observed for neutral SNPs, outlier SNPs and inversions, but also for one of the two phenotypes we studied (shell size). For individual loci, the extent of the shift is not consistent. While there is some uncertainty associated with each individual cline fit, which may weaken the correlation between cline centre estimates from different zones, the complete lack of a correlation probably indicates that the general shift is explained by genome-wide effects rather than locus-specific effects such as epistasis or dominance. We often found a large number of cline centres clustering on the rock side of the boulder–rock transition, with a long tail of additional cline centres spreading further into the rock habitat.

We suggest two, not mutually exclusive, possible explanations for cline shifts. First, the boulder–rock transition might not correspond to the position where the main selection pressures change. Either the well-established crab and wave selection pressures change to some extent independently of the substrate or other environmental factors contribute to selection. In some locations the bedrock is relatively flat, potentially allowing crabs to forage on it. In others, wave exposure is influenced by shore form and aspect as well as substrate. This is supported by our finding that the presence of barnacles (which indicate wave exposure; Jenkins & Hawkins, 2003) often increases some distance from the boulder–rock transition. If selection pressures change more gradually than the very sudden shift in substrate, this could also explain the trailing-off of the cline centre distribution into the rock habitat.

Second, between-ecotype differences in dispersal or population density could contribute to the shift of the clines. If net dispersal from Crab into Wave is higher than vice versa, this would lead to a general shift of clines. Our isolation-by-distance analysis on “pure” Crab and Wave samples is consistent with this, as Crab individuals show essentially no population structure, while there are patterns of increasing genetic distance with spatial distance in some Wave samples. This could indicate higher dispersal, higher density or both in Crab habitats or in the Crab ecotype, which is conceivable given that dispersal between boulders might be substantially different from that on a cliff, and that Crab snails are much larger than Wave snails. Alternatively, it is possible that the stronger spatial structure in the Wave ecotype is a direct effect of stronger spatial variation in selection pressures. Ultimately, a difference in density or dispersal can only be confirmed with extensive sampling, mark–recapture or experimental studies.

Notably, while cline shifts were observed in all zones, the extent of displacement and the amount of variation among cline centres differs substantially between different zones. This is probably explained by differences among zones regarding the extent and abruptness at which the selective environment is changing. As our alternative environmental measures show, despite the presence of a relatively clear boulder–rock transition the shores are highly complex, with some selection pressures probably varying on a very local scale. Other factors might also contribute, such as differences in the strength of selection between habitats.



Whatever the explanation, it is clear that the cline shifts we observed at ANG are not solely explained by a site-specific density trough, as we hypothesized when studying only a single zone (Westram et al., 2018). Similar density troughs have not been observed at other locations. Our results on cline shifts therefore represent a good example of how replicate hybrid zones can help to disentangle site-specific processes from those that are more general. How multiple environmental transitions, demographic factors and selection strengths interact to generate spatial patterns of allele frequency change, and how this interaction affects the progress of divergence and speciation, is an exciting direction for further theoretical and empirical research.

## ACKNOWLEDGEMENTS

We thank everyone who helped with fieldwork, snail processing and DNA extractions, particularly Laura Brettell, Mårten Duvetorp, Juan Galindo, Anne-Lise Liabot, Mark Ravinet, Irena Senčić and Zuzanna Zagrodzka. We are also grateful to Edinburgh Genomics for library preparation and sequencing, to Stuart Baird and Mark Ravinet for helpful discussions, and to three anonymous reviewers for their constructive comments. This work was supported by the Natural Environment Research Council (NE/K014021/1), the European Research Council (AdG-693030-BARRIERS), Swedish Research Councils Formas and Vetenskapsrådet through a Linnaeus grant to the Centre for Marine Evolutionary Biology (217-2008-1719), the European Regional Development Fund (POCI-01-0145-FEDER-030628), and the Fundação para a Ciência e a Tecnologia, Portugal (PTDC/BIA-EVL/30628/2017). A.M.W. and R.F. were funded by the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreements no. 754411/797747 and no. 706376, respectively.

## AUTHOR CONTRIBUTIONS

A.M.W., K.J. and R.K.B. conceived the study and collected data. A.M.W., R.F. and R.K.B. analysed data. A.M.W. wrote the manuscript with input from all other authors.

## DATA AVAILABILITY STATEMENT

All raw sequencing files in fastq format, as well as individual sampling positions and phenotypes, are available on dryad (<https://doi.org/10.5061/dryad.tb2rbp005>). Scripts for bioinformatic analyses and cline analyses are available on GitHub ([https://github.com/AnjaWestram/Littorina\\_hybrid\\_zone\\_2](https://github.com/AnjaWestram/Littorina_hybrid_zone_2)).

## ORCID

Anja M. Westram  <https://orcid.org/0000-0003-1050-4969>

Rui Faria  <https://orcid.org/0000-0001-6635-685X>

Kerstin Johannesson  <https://orcid.org/0000-0003-0176-7986>

## REFERENCES

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Arntzen, J. W., Wielstra, B., & Wallis, G. P. (2014). The modality of nine *Triturus* newt hybrid zones assessed with nuclear, mitochondrial and morphological data. *Biological Journal of the Linnean Society*, 113(2), 604–622. <https://doi.org/10.1111/bij.12358>
- Barth, J. M. I., Berg, P. R., Jonsson, P. R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., Jakobsen, K. S., Johannesson, K., Jorde, P. E., Knutsen, H., Moksnes, P.-O., Star, B., Stenseth, N. C., Svedäng, H., Jentoft, S., & André, C. (2017). Genome architecture enables local adaptation of Atlantic cod despite high connectivity. *Molecular Ecology*, 26(17), 4452–4466. <https://doi.org/10.1111/mec.14207>
- Barton, N. H. (1983). Multilocus clines. *Evolution*, 37(3), 454–471. <https://doi.org/10.2307/2408260>
- Barton, N. H., & Gale, K. S. (1993). Genetic analysis of hybrid zones. *Hybrid Zones and the Evolutionary Process*, 13–45.
- Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, 16, 113–148.
- Berg, P. R., Star, B., Pampoulie, C., Sodeland, M., Barth, J. M. I., Knutsen, H., Jakobsen, K. S., & Jentoft, S. (2016). Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific Reports*, 6(1), 1–12. <https://doi.org/10.1038/srep23246>
- Beysard, M., & Heckel, G. (2014). Structure and dynamics of hybrid zones at different stages of speciation in the common vole (*Microtus arvalis*). *Molecular Ecology*, 23(3), 673–687. <https://doi.org/10.1111/mec.12613>
- Bierne, N., Welch, J., Loire, E., Bonhomme, F., & David, P. (2011). The coupling hypothesis: Why genome scans may fail to map local adaptation genes. *Molecular Ecology*, 20(10), 2044–2072.
- Bonin, A., Taberlet, P., Miaud, C., & Pompanon, F. (2006). Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, 23(4), 773–783. <https://doi.org/10.1093/molbev/msj087>
- Butlin, R. K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J. A., Galindo, J., Grahame, J. W., Hollander, J., Kempainen, P., Martínez-Fernández, M., Panova, M., Quesada, H., Johannesson, K., & Rolán-Alvarez, E. (2014). Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution*, 68(4), 935–949.
- Cheng, C., White, B. J., Kamdem, C., Mockaitis, K., Costantini, C., Hahn, M. W., & Besansky, N. J. (2012). Ecological genomics of *Anopheles gambiae* along a latitudinal cline: A population-resequencing approach. *Genetics*, 190(4), 1417–1432. <https://doi.org/10.1534/genetics.111.137794>
- Christe, C., Störling, K. N., Bresadola, L., Fussi, B., Heinze, B., Wegmann, D., & Lexer, C. (2016). Selection against recombinant hybrids maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and recurrent gene flow. *Molecular Ecology*, 25(11), 2482–2498. <https://doi.org/10.1111/mec.13587>
- Conte, G. L., Arnegard, M. E., Peichel, C. L., & Schluter, D. (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), 5039–5047.
- Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- Dufresnes, C., Brelsford, A., Crnobrnja-Isailović, J., Tzankov, N., Lymberakis, P., & Perrin, N. (2015). Timeframe of speciation inferred from secondary contact zones in the European tree frog radiation (*Hyla arborea* group). *BMC Evolutionary Biology*, 15(1), 155. <https://doi.org/10.1186/s12862-015-0385-2>

- Duranton, M., Allal, F., Fraïsse, C., Bierne, N., Bonhomme, F., & Gagnaire, P.-A. (2018). The origin and remodelling of genomic islands of differentiation in the European sea bass. *Nature Communications*, 9(1), 1–11. <https://doi.org/10.1038/s41467-018-04963-6>
- Faria, R., Chaube, P., Morales, H. E., Larsson, T., Lemmon, A. R., Lemmon, E. 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 & Evolution*, 34(3), 239–248. <https://doi.org/10.1016/j.tree.2018.12.005>
- Flaxman, S. M., Wacholder, A. C., Feder, J. L., & Nosil, P. (2014). Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Molecular Ecology*, 23(16), 4074–4088. <https://doi.org/10.1111/mec.12750>
- Haldane, J. B. S. (1930). A mathematical theory of natural and artificial selection. (Part VI, isolation). *Mathematical Proceedings of the Cambridge Philosophical Society*, 26(2), 220–230. <https://doi.org/10.1017/S0305004100015450>
- Harrison, R. G. (1993). *Hybrid zones and the evolutionary process*. Oxford University Press.
- Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, 173(5), 579–588. <https://doi.org/10.1086/597611>
- Hewitt, G. M. (1988). Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, 3(7), 158–167. [https://doi.org/10.1016/0169-5347\(88\)90033-X](https://doi.org/10.1016/0169-5347(88)90033-X)
- Hollander, J., Galindo, J., & Butlin, R. K. (2015). Selection on outlier loci and their association with adaptive phenotypes in *Littorina saxatilis* contact zones. *Journal of Evolutionary Biology*, 28(2), 328–337. <https://doi.org/10.1111/jeb.12564>
- Jenkins, S. R., & Hawkins, S. J. (2003). Barnacle larval supply to sheltered rocky shores: A limiting factor? In M. B. Jones, A. Ingólfsson, E. Ólafsson, G. V. Helgason, K. Gunnarsson, & J. Svavarsson (Eds.), *Migrations and dispersal of marine organisms* (pp. 143–151). Springer Netherlands. [https://doi.org/10.1007/978-94-017-2276-6\\_16](https://doi.org/10.1007/978-94-017-2276-6_16)
- Johannesson, B. (1986). Shell morphology of *Littorina saxatilis* Oliv.: The relative importance of physical factors and predation. *Journal of Experimental Marine Biology and Ecology*, 102(2), 183–195. [https://doi.org/10.1016/0022-0981\(86\)90175-9](https://doi.org/10.1016/0022-0981(86)90175-9)
- Johannesson, K., Panova, M., Kempainen, P., André, C., Rolán-Alvarez, E., & Butlin, R. K. (2010). Repeated evolution of reproductive isolation in a marine snail: Unveiling mechanisms of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1547), 1735–1747.
- Johannesson, K., Zagrodzka, Z., Faria, R., Westram, A. M., & Butlin, R. K. (2020). Is embryo abortion a post-zygotic barrier to gene flow between *Littorina* ecotypes? *Journal of Evolutionary Biology*, 33(3), 342–351. <https://doi.org/10.1111/jeb.13570>
- Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., Whibley, A., Becuwe, M., Baxter, S. W., Ferguson, L., Wilkinson, P. A., Salazar, C., Davidson, C., Clark, R., Quail, M. A., Beasley, H., Glithero, R., Lloyd, C., Sims, S., ... French-Constant, R. H. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature*, 477(7363), 203–206. <https://doi.org/10.1038/nature10341>
- Kempainen, 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>
- Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biology*, 8(9), e1000501. <https://doi.org/10.1371/journal.pbio.1000501>
- Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, 173(1), 419–434. <https://doi.org/10.1534/genetics.105.047985>
- Le Penne, 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>
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>
- Morales, H. E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A. M., & Butlin, R. K. (2019). Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast. *Science Advances*, 5(12), eaav9963. <https://doi.org/10.1126/sciadv.aav9963>
- Noor, M. A. F., & Bennett, S. M. (2009). Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, 103(6), 439–444. <https://doi.org/10.1038/hdy.2009.151>
- Nosil, P. (2012). *Ecological speciation*. Oxford University Press.
- Nosil, P., Egan, S. P., & Funk, D. J. (2008). Heterogeneous genomic differentiation between walking-stick ecotypes: “Isolation by adaptation” and multiple roles for divergent selection. *Evolution*, 62(2), 316–336. <https://doi.org/10.1111/j.1558-5646.2007.00299.x>
- Nosil, P., Harmon, L. J., & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. *Trends in Ecology & Evolution*, 24(3), 145–156. <https://doi.org/10.1016/j.tree.2008.10.011>
- Ohta, T. (1971). Associative overdominance caused by linked detrimental mutations\*. *Genetics Research*, 18(3), 277–286. <https://doi.org/10.1017/S0016672300012684>
- Panova, M., Aronsson, H., Cameron, R. A., Dahl, P., Godhe, A., Lind, U., & Johannesson, K. (2016). DNA extraction protocols for whole-genome sequencing in marine organisms. In S. J. Bourlat (Ed.), *Marine genomics* (pp. 13–44). Springer. [https://doi.org/10.1007/978-1-4939-3774-5\\_2](https://doi.org/10.1007/978-1-4939-3774-5_2)
- Panova, M., Hollander, J., & Johannesson, K. (2006). Site-specific genetic divergence in parallel hybrid zones suggests nonallopatric evolution of reproductive barriers. *Molecular Ecology*, 15(13), 4021–4031.
- Perrier, C., Bourret, V., Kent, M. P., & Bernatchez, L. (2013). Parallel and nonparallel genome-wide divergence among replicate population pairs of freshwater and anadromous Atlantic salmon. *Molecular Ecology*, 22(22), 5577–5593. <https://doi.org/10.1111/mec.12500>
- Pfeifer, S. P., Laurent, S., Sousa, V. C., Linnen, C. R., Foll, M., Excoffier, L., Hoekstra, H. E., & Jensen, J. D. (2018). The evolutionary history of Nebraska deer mice: Local adaptation in the face of strong gene flow. *Molecular Biology and Evolution*, 35(4), 792–806. <https://doi.org/10.1093/molbev/msy004>
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation. *Journal of Evolutionary Biology*, 30(8), 1450–1477.
- Ravinet, M., Westram, A., Johannesson, K., Butlin, R., André, C., & Panova, M. (2016). Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Molecular Ecology*, 25(1), 287–305.
- Renaut, S., Grassa, C. J., Yeaman, S., Moyers, B. T., Lai, Z., Kane, N. C., Bowers, J. E., Burke, J. M., & Rieseberg, L. H. (2013). Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications*, 4, 1827. <https://doi.org/10.1038/ncomms2833>
- Roesti, M., Hendry, A. P., Salzburger, W., & Berner, D. (2012). Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. *Molecular Ecology*, 21(12), 2852–2862. <https://doi.org/10.1111/j.1365-294X.2012.05509.x>

- Roesti, M., Kueng, B., Moser, D., & Berner, D. (2015). The genomics of ecological vicariance in threespine stickleback fish. *Nature Communications*, 6(1), 1–16. <https://doi.org/10.1038/ncomms9767>
- Ryan, S. F., Fontaine, M. C., Scriber, J. M., Pfrender, M. E., O'Neil, S. T., & Hellmann, J. J. (2017). Patterns of divergence across the geographic and genomic landscape of a butterfly hybrid zone associated with a climatic gradient. *Molecular Ecology*, 26(18), 4725–4742. <https://doi.org/10.1111/mec.14236>
- Samuk, K., Owens, G. L., Delmore, K. E., Miller, S. E., Rennison, D. J., & Schluter, D. (2017). Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Molecular Ecology*, 26(17), 4378–4390. <https://doi.org/10.1111/mec.14226>
- Schaefer, J., Duvernell, D., & Campbell, D. C. (2016). Hybridization and introgression in two ecologically dissimilar *Fundulus* hybrid zones. *Evolution*, 70(5), 1051–1063. <https://doi.org/10.1111/evo.12920>
- Schaefer, J. F., Duvernell, D. D., & Kreiser, B. R. (2011). Ecological and genetic assessment of spatial structure among replicate contact zones between two topminnow species. *Evolutionary Ecology*, 25(5), 1145–1161.
- Singhal, S., & Bi, K. (2017). History cleans up messes: The impact of time in driving divergence and introgression in a tropical suture zone. *Evolution*, 71(7), 1888–1899. <https://doi.org/10.1111/evo.13278>
- Smadja, C. M., & Butlin, R. K. (2011). A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, 20(24), 5123–5140. <https://doi.org/10.1111/j.1365-294X.2011.05350.x>
- Souissi, A., Bonhomme, F., Machado, M., Bahri-Sfar, L., & Gagnaire, P.-A. (2018). Genomic and geographic footprints of differential introgression between two divergent fish species (*Solea* spp.). *Heredity*, 121(6), 579–593. <https://doi.org/10.1038/s41437-018-0079-9>
- Stankowski, S., Sobel, J. M., & Streisfeld, M. A. (2017). Geographic cline analysis as a tool for studying genome-wide variation: A case study of pollinator-mediated divergence in a monkeyflower. *Molecular Ecology*, 26(1), 107–122. <https://doi.org/10.1111/mec.13645>
- Twyford, A. D., & Friedman, J. (2015). Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution*, 69(6), 1476–1486. <https://doi.org/10.1111/evo.12663>
- Wellenreuther, M., & Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology & Evolution*, 33(6), 427–440. <https://doi.org/10.1016/j.tree.2018.04.002>
- 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>
- Wilding, C. S., Butlin, R. K., & Grahame, J. (2001). Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology*, 14(4), 611–619.
- Wood, H. M., Grahame, J. W., Humphray, S., Rogers, J., & Butlin, R. K. (2008). Sequence differentiation in regions identified by a genome scan for local adaptation. *Molecular Ecology*, 17(13), 3123–3135.
- Zieliński, P., Dudek, K., Arntzen, J. W., Palomar, G., Niedzicka, M., Fijarczyk, A., Liana, M., Cogălniceanu, D., & Babik, W. (2019). Differential introgression across newt hybrid zones: Evidence from replicated transects. *Molecular Ecology*, 28(21), 4811–4824. <https://doi.org/10.1111/mec.15251>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Westram AM, Faria R, Johannesson K, Butlin R. Using replicate hybrid zones to understand the genomic basis of adaptive divergence. *Mol Ecol*. 2021;30:3797–3814. <https://doi.org/10.1111/mec.15861>