


Modular chromosome rearrangements reveal parallel and nonparallel adaptation in a marine fish

Tony Kess¹  | Paul Bentzen² | Sarah J. Lehnert¹ | Emma V. A. Sylvester¹ | Sigbjørn Lien³ | Matthew P. Kent³ | Marion Sinclair-Waters⁴ | Corey Morris¹ | Brendan Wrings⁵ | Robert Fairweather² | Ian R. Bradbury^{1,2}

¹Fisheries and Oceans Canada, Northwest Atlantic Fisheries Centre, St. John's, NL, Canada

²Biology Department, Dalhousie University, Halifax, NS, Canada

³Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Centre for Integrative Genetics, Norwegian University of Life Sciences, Ås, Norway

⁴Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki, Finland

⁵Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS, Canada

Correspondence

Tony Kess, Fisheries and Oceans Canada, Northwest Atlantic Fisheries Centre, St. John's, NL, Canada.
Email: tonykess@gmail.com

Funding information

NSERC Discovery Grant

Abstract

Genomic architecture and standing variation can play a key role in ecological adaptation and contribute to the predictability of evolution. In Atlantic cod (*Gadus morhua*), four large chromosomal rearrangements have been associated with ecological gradients and migratory behavior in regional analyses. However, the degree of parallelism, the extent of independent inheritance, and functional distinctiveness of these rearrangements remain poorly understood. Here, we use a 12K single nucleotide polymorphism (SNP) array to demonstrate extensive individual variation in rearrangement genotype within populations across the species range, suggesting that local adaptation to fine-scale ecological variation is enabled by rearrangements with independent inheritance. Our results demonstrate significant association of rearrangements with migration phenotype and environmental gradients across the species range. Individual rearrangements exhibit functional modularity, but also contain loci showing multiple environmental associations. Clustering in genetic distance trees and reduced differentiation within rearrangements across the species range are consistent with shared variation as a source of contemporary adaptive diversity in Atlantic cod. Conversely, we also find that haplotypes in the LG12 and LG1 rearranged region have diverged across the Atlantic, despite consistent environmental associations. Exchange of these structurally variable genomic regions, as well as local selective pressures, has likely facilitated individual diversity within Atlantic cod stocks. Our results highlight the importance of genomic architecture and standing variation in enabling fine-scale adaptation in marine species.

KEYWORDS

Atlantic cod, environmental association, genomic architecture, marine, migration, parallel evolution

1 | INTRODUCTION

Instances of parallel evolution provide the opportunity to quantify the variety of solutions to ecological problems at different scales of organization, from molecular to phenotypic levels. Understanding

the processes underlying parallel genomic adaptation is a fundamental goal in biology and can clarify to what extent evolution is deterministic (Blount, Lenski, & Losos, 2018; Conte, Arnegard, Peichel, & Schluter, 2012; Elmer & Meyer, 2011; Stern, 2013). Genomic organization and history may be integral to genomic parallelism, affecting

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.

© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

the rate and predictability of adaptation (Stern, 2013; Tigano & Friesen, 2016). Recent evidence has demonstrated that the amount of shared variation within a genome and how this variation is organized structurally can facilitate local adaptation and may influence rates of genomic parallelism (Nelson & Cresko, 2018; Pearse, Miller, Abadia-Cardoso, & Garza, 2014). However, empirical support demonstrating the role of shared variation and structural variation across systems is only beginning to be understood.

Genomic architecture, the organization and interactions of genes within the genome, can affect rates of adaptation (Flaxman, Wacholder, Feder, & Nosil, 2014; Rogers, Mee, & Bowles, 2013). Genomic rearrangements, such as inversions, can reduce recombination within rearranged genomic regions, allowing evolution of co-adapted complexes of alleles advantageous for adaptation in heterogeneous environments (Barb et al., 2014; Hoffmann & Rieseberg, 2008; Kirkpatrick & Barton, 2006). Compartmentalization of genes underlying different traits into functionally independent units (i.e., functional modularity) may also affect rates of adaptation (Ragland et al., 2017; Rogers et al., 2013; Stern & Orgogozo, 2007; Wagner, Pavlicev, & Cheverud, 2007) and can enable evolution of phenotypes adapted to different selective pressures without fitness reduction caused by genetic constraints (Wagner et al., 2007).

Standing variation in these genomic architectural features can allow rapid responses to selection and directly influence genomic parallelism (Bolnick, Barrett, Oke, Rennison, & Stuart, 2018). Examples of repeated incorporation of shared rearrangements into the genetic architecture of adaptive traits have been identified across taxa, including loci contributing to crypsis in *Timema cristinae* stick insects (Lindtke et al., 2017), diapause timing in apple maggot flies (*Rhagoletis pomonella*, Feder et al., 2003; Ragland et al., 2017), and rapid colonization of novel habitats in threespine sticklebacks (*Gasterosteus aculeatus*, Bassham, Catchen, Lescak, Hippel, & Cresko, 2018; Jones et al., 2012). However, understanding the eco-evolutionary functions of genomic architectural features has only recently become possible with high-resolution sequencing, genomic mapping, and genome assemblies (Wellenreuther & Bernatchez, 2018).

Atlantic cod (*Gadus morhua*), a heavily exploited, demersal, broadcast spawning marine fish, provides a study system well suited to investigate the extent of modularity and parallelism among genomic architectural features. Atlantic cod exhibit high connectivity and dispersal, but also show divergent adaptation and genetic structuring associated with genomic architectural variation (Barth et al., 2017; Sinclair-Waters, Bentzen, et al., 2018). Past genetic studies of Atlantic cod have suggested parallel environmental adaptation to temperature and salinity in the east and west Atlantic (Berg et al., 2015; Bradbury et al., 2010). Phenotypic parallelism of migratory and stationary ecotypes occurring in Europe and North American coastal waters has also been identified (Robichaud & Rose, 2004). This environmental and behavioral divergence has likely been facilitated by rearrangements consisting of multiple potential inversions on four separate linkage groups (LG 1, 2, 7, and 12, Kirubakaran et al., 2016; Sodeland et al., 2016). However, several outstanding

questions remain regarding the origin and function of diverged, rearranged genomic regions among individual Atlantic cod across the range. The extent of coinheritance and functional modularity of rearrangements in Atlantic cod, as well as the extent that the same rearrangements are shared across the Atlantic remains unclear. Previous range-wide comparisons have revealed interchromosomal linkage between rearranged regions, suggesting a shared evolutionary trajectory for rearranged regions (Berg et al., 2015; Bradbury et al., 2014). However, structuring associated with environmental clines and trans-Atlantic divergence can produce interchromosomal linkage that may mask how frequently chromosomal rearrangements are co-inherited at finer scales.

Here, we build on previous studies and explore the potential for exchange of these rearrangements across the entire Atlantic range, and their role in local adaptation to fine-scale ecological variation in Atlantic cod. First, we quantify coinheritance of chromosomal rearrangements within populations to test whether multiple rearrangements show a shared evolutionary trajectory. We then conduct environmental and migratory associations with genomic markers using multivariate and machine-learning methods that can detect relationships between alleles that would be missed in single-locus models and genome scans (Forester, Lasky, Wagner, & Urban, 2018; Gagnaire & Gaggiotti, 2016; Stephan, Stegle, & Beyer, 2015). Next, we use multivariate clustering to determine whether loci within rearrangements show consistent associations with the same environmental variables, indicating functional modularity (Lotterhos, Yeaman, Degner, Aitken, & Hodgins, 2018). Finally, we test for signatures of adaptation from shared variation to infer whether the same haplotypes in rearranged regions are shared across the Atlantic. Overall, the results from our study highlight how features of genomic architecture and history can enable repeated differentiation and adaptation across a variety of habitats within a highly connected marine species.

2 | METHODS

2.1 | Genotype data

To identify population structure and genome architecture variation across the range of Atlantic cod, we combined five datasets that were previously genotyped using the same 12K SNP array, along with additional samples from the Flemish Cap genotyped using the same array (Barth et al., 2017; Berg et al., 2017; Kess et al., 2019; Sinclair-Waters, Bentzen, et al., 2018; Sinclair-Waters, Bradbury, et al., 2018). Atlantic cod genotypes from North America and Europe were obtained from 44 separate sampling trips from 37 distinct sampling locations (Figure 1a, sampling site and year in Table S1). Detailed genotyping and collection information are available in the studies cited above. These genotypes were combined into a final dataset of $n = 1,230$ ($n = 531$ North America, $n = 699$ Europe) individuals genotyped at 6,669 informative SNPs genotyped in all populations, corrected for strand flips,

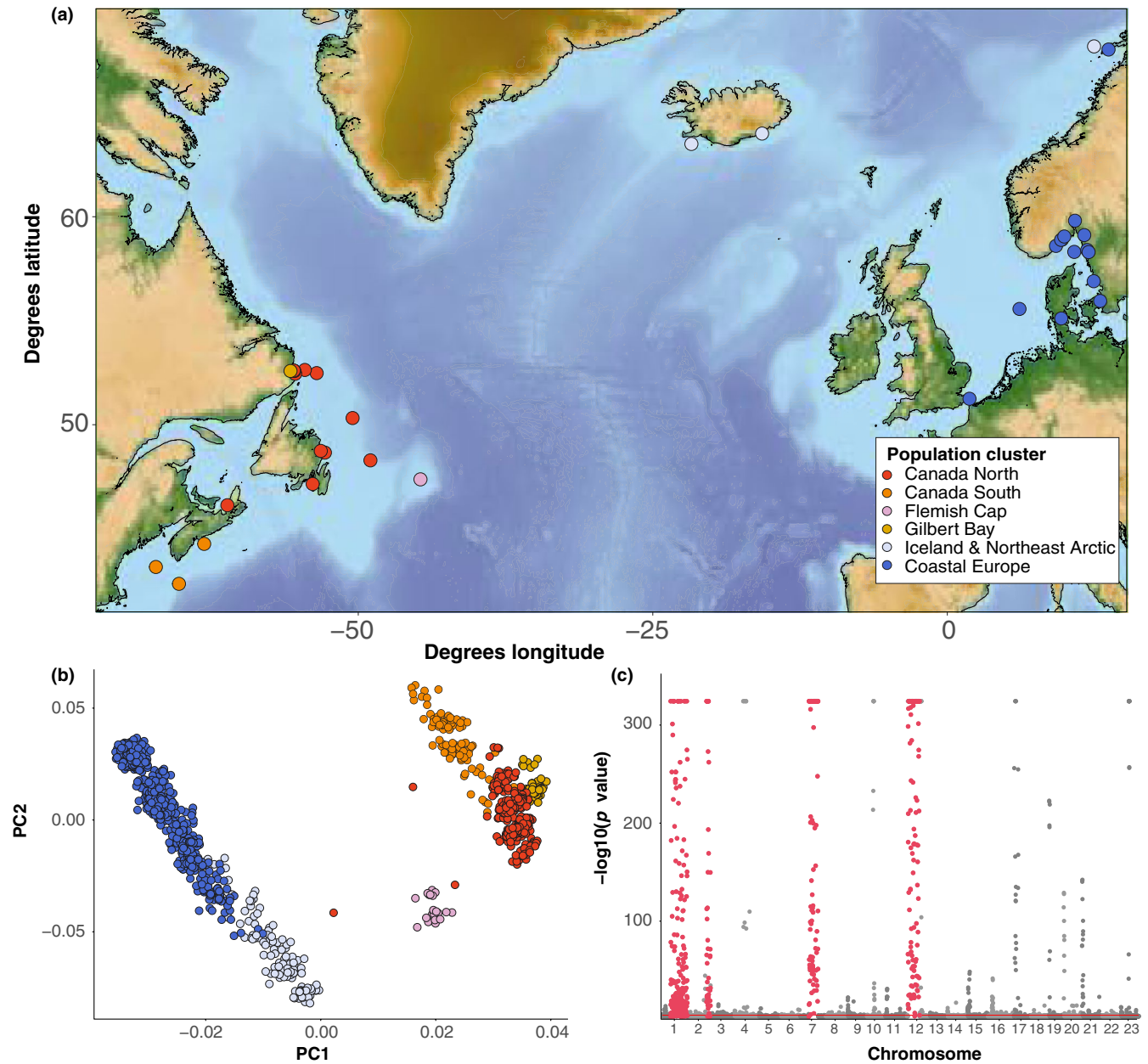


FIGURE 1 Geographic location of sampling sites, population clusters, and single nucleotide polymorphism (SNP) correlations with principal component axes associated with population structure in Atlantic cod. (a) Location and population clusters of sampling sites. (b) Population clustering of samples on the first two principal component axes in *pcadapt*. (c) Significance of correlation of each SNP (p value) with the $K = 25$ retained principal components in *pcadapt*. Chromosome rearrangements are marked in red

and filtered to remove SNPs with ambiguous genotype coding across datasets.

2.2 | Range-wide population structure

We quantified range-wide population structure using principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE), a nonlinear machine-learning clustering algorithm (van der Maaten & Hinton, 2008), with genotypes from all 1,230 individuals from all populations. We detected outliers driving population clustering in PCA by calculating Mahalanobis distances for each SNP from a

vector of z-scores describing its relationship with K tested principal components. Using the R package *pcadapt* (Luu, Bazin, & Blum, 2017), we carried out PCA with $K = 2$ to identify SNPs associated with the first two PC axes explaining most geographic variation, assuming a false discovery rate (FDR, Storey & Tibshirani, 2003) cutoff of $q = 0.1$. We then used $K = 25$ to identify SNPs that were correlated with principal components that explained less variation but captured more subtle genetic differences between populations and individuals. To represent this variation on two axes, we also calculated t-SNE scores for each individual using the *tSNE* R package (Donaldson, 2016). We assigned a perplexity value of 25; this parameter characterizes the number of effective nearest neighbors expected in the data.

2.3 | Individual genetic variation

To quantify individual genetic variation, we separated individuals into geographic groups from Canada North, Canada South, Gilbert Bay, Flemish Cap, Iceland and Northeast Arctic, and Coastal Europe (Figure 1a,b). These groupings are supported by separation of geographic regions in both PCA and t-SNE, although the split with Glibert Bay is more distinct in t-SNE, suggesting multiple axes of genetic divergence (Figures 1b and S1). We then conducted PCA within each of these clusters using *pcadapt*. Outliers differentiated between individuals were identified by significance of Mahalanobis distances for each SNP, using an FDR of $q = 0.1$. We set $K = 25$ to increase the chance of identifying SNPs making marginal contributions to individual structuring. PCA clustering in the Canada North region has been performed previously in Kess et al., 2019, but was limited to $K = 1$ principal components and did not include the population in the Gulf of St. Lawrence (CSG), as the goal of the previous study was to identify the primary driver of individual genetic structuring within the Northern cod stock around Newfoundland and Labrador.

2.4 | Range-wide genetic structure of rearrangements

Separate Bayesian clustering of SNPs within each chromosomal rearrangement was conducted using STRUCTURE v2.3.4 (Pritchard, Stephens, & Donnelly, 2000) to determine individual rearrangement genotype. We carried out separate analyses of the 531 individuals within North America and the 699 individuals within Europe. We conducted separate runs for all SNPs within rearrangements on LG1, LG2, LG7, and LG12 in R using *parallelstructure* (Besnier & Glover, 2013). For each rearrangement, we ran three replicate Markov Chain Monte Carlo runs of 100,000 burn-ins followed by 500,000 iterations, assuming $K = 2$. Bayesian clustering of the LG1 rearranged region was previously carried out by Kess et al., 2019, but did not include the Flemish Cap samples. Individual genetic ancestry proportions were summarized and visualized using CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015). Filtering of rearrangement SNPs based on genome locations used by Berg et al., 2017 was performed using *genepopedit* (Stanley, Jeffery, Wringe, DiBacco, & Bradbury, 2016), and conversion to STRUCTURE format was carried out with PGDSpider (Lischer & Excoffier, 2012). We calculated rearrangement frequencies by grouping individual rearrangement genotypes into homozygous rearranged, homozygous nonrearranged, and heterozygous categories inferred from STRUCTURE admixture proportions (Q -values). Because categorization of chromosomal orientations into rearranged and nonrearranged categories requires phylogenetic comparison across multiple species to test the origin of rearranged regions, we use “rearranged” genotype and “nonrearranged” genotype designations only to characterize alternative chromosomal states, but follow designations used by Berg et al. (2017).

2.5 | Coinheritance and linkage disequilibrium

We tested for coinheritance of all *pcadapt* outlier SNPs identified within each regional cluster exhibiting within-population architectural variation using two linkage-based methods. We calculated an r^2 matrix quantifying pairwise linkage disequilibrium (LD) between each SNP using plink 1.9 (Chang et al., 2015). LD calculations were performed separately within Canada North ($n = 553$ SNPs), Canada South ($n = 490$ SNPs), Iceland and Northeast Arctic ($n = 618$ SNPs), and Coastal Europe population clusters ($n = 730$ SNPs). We visualized linkage between outlier SNPs using an LD heatmap in the *superheat* R package (Barter & Yu, 2017). Using the *LDna* R package (Kemppainen et al., 2015), we generated network plots to visualize each SNP as a node connected by edges exceeding an r^2 threshold of .25.

2.6 | Environmental and migratory association with genomic variation

To identify association of all 6,669 SNPs with environmental gradients, we carried out genome-wide association using two polygenic methods: redundancy analysis (RDA, Rao, 1964) and random forest (Breiman, 2001). We conducted RDA and random forest separately for all individuals in North America and in Europe to limit the possibility of nonparallel genetic architecture and residual population structure diluting signals of local adaptation. We extracted 12 environmental variables for each sampling location from the Bio-ORACLE v2.0 database (Assis et al., 2018) using the *sdmpredictors* R package (Bosch, 2018). Environmental variables were obtained from mean bottom depth layers for monthly (2000–2014) mean, minimum and maximum values for salinity, dissolved oxygen, and temperature. To reduce the number of correlated variables in RDA, we standardized each variable by its standard deviation and then conducted separate PCAs on variables for salinity, temperature, and dissolved oxygen. We used the first PC for each of these variables as individual phenotypes used as predictors in an RDA model to identify environmental association. SNPs with absolute RDA scores exceeding the 95th percentile were categorized as outliers. We conducted redundancy analysis using the *vegan* R package (Oksanen et al., 2011). We did not include a covariate for geographic structure in environmental association, as true signals of association may be lost when environmental adaptation tracks geographic separation (Yeaman et al., 2016), and dimensionality reduction used in redundancy analysis has been shown to effectively control for background structure (Forester, Jones, Joost, Landguth, & Lasky, 2016).

We carried out separate random forest regressions for each environmental variable in North America and Europe. Loci significantly associated with each environmental variable were identified as those with mean decrease in accuracy (MDA) scores exceeding a threshold identified by MDA drop-off inferred from binning of MDA scores. SNPs with MDA values below this threshold exhibited low power in predicting environmental features. Random forest regressions were conducted using the *randomForest* R package (Liaw & Wiener, 2002).

We explored genetic associations with migration phenotype for all individuals in Europe using all 6,669 SNPs, following the association methods used in Kess et al. (2019) in North America. Regional-level migration phenotype was assigned to each individual from migratory classes identified by Robichaud and Rose (2004). As above, we identified SNP associations with migratory phenotype using redundancy analysis without correction for geographic structure and random forest classification.

To determine the direction of association between environmental and migratory variables and rearrangement frequencies, we calculated Pearson's correlation between the frequency of each rearrangement and each environmental variable. We assessed significance at a lenient $p < .05$, as the goal of this test was to quantify direction of association.

2.7 | Clustering of loci by multiple environmental associations

We examined patterns of SNP co-association to uncover groups of loci that exhibited similar patterns of associations across multiple environmental variables. As in Lotterhos et al. (2018), we calculated absolute values of Pearson's correlation coefficient (r^2) between each SNP and all environmental variables: principal component, minimum, maximum and mean values for salinity, temperature, and dissolved oxygen. We calculated separate correlations for each SNP that was identified as significantly associated with environment in RDA in Europe and in North America. We then produced a separate pairwise Euclidean distance matrix for Europe and North America from distances between SNPs and their associations across variables and carried out hierarchical clustering on each matrix using Ward's hierarchical clustering algorithm in R using the *hclust* function. Loci that exhibit shared patterns of association across multiple environmental variables will be found in the same cluster using hierarchical clustering; these clusters represent proxies for functional modules. The R package *superheat* was then used to visualize patterns of shared associations using paired dendrograms and heatmaps.

2.8 | Signatures of adaptation from shared variation

To identify whether environment and migration-associated rearrangements represent adaptation from shared variation, or from de novo mutations, we measured differentiation and separation in trees of genetic distance between European and North American individuals with shared rearrangement genotypes. We compared F_{ST} and d_{XY} across linkage groups LG1, LG2, LG7, and LG12 between individuals with each rearrangement genotype in North America and Europe and calculated nucleotide diversity (π) separately for each rearrangement genotype in North America and Europe. F_{ST} calculations were carried out in the *diveRsity* R package (Keenan, McGinnity, Cross, Crozier, & Prodöhl, 2013). We used Beagle 5.0 (Browning, Zhou, & Browning, 2018) to phase vcf files and then calculated d_{XY} and π using the R package *PopGenome* (Pfeifer, Wittelsbürger, Ramos-Onsins, & Lercher,

2014). Differences in F_{ST} , d_{XY} , and π between rearrangement genotype groups, and within and outside structurally variable regions on each chromosome were tested using a Mann-Whitney U test (Mann & Whitney, 1947). We then calculated range-wide individual genetic distance trees for each rearrangement separately with all 1,230 genotyped individuals. We calculated Cavalli-Sforza and Edwards chord distances (Cavalli-Sforza & Edwards, 1967) with 1,000 loci bootstraps in POPULATIONS v1.2.33 (Langella, 2012) and visualized trees using FigTree v1.74 (Rambaut, 2012).

3 | RESULTS

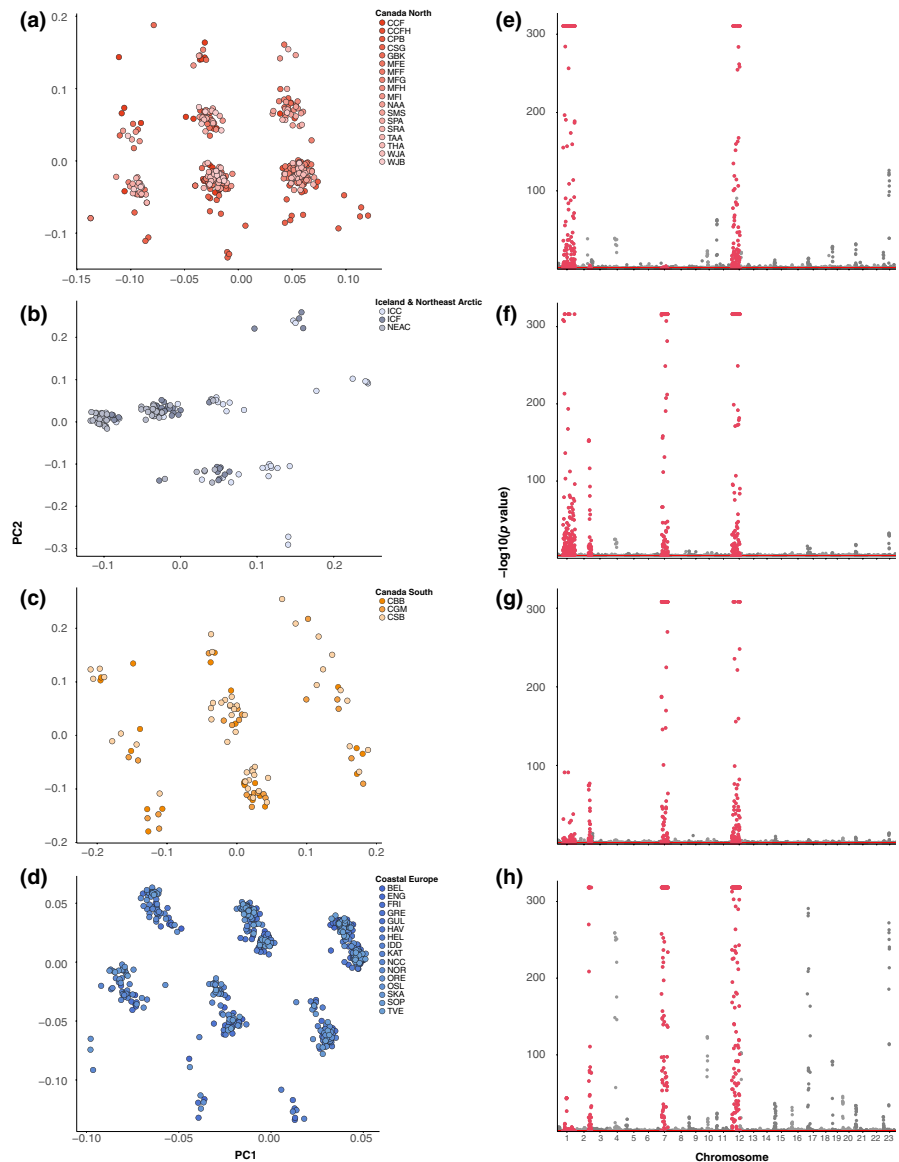
3.1 | Population structure and individual genetic variation

Using dimensionality reduction, we identified population structure associated with differentiation among geographic regions, and structuring by rearrangement genotype within these geographic regions. Structuring within the total dataset uncovered trans-Atlantic and latitudinal separation (Figures 1b and S1), driven by 1,012 SNPs with significant q -values (<0.1), concentrated in four known rearrangements on LG1, LG2, LG7, and LG12 (Figure 1c). Within the Canada North, Canada South, Iceland and Northeast Arctic, and Coastal Europe regions, we found multiple discrete genetic clusters primarily driven by different complements of chromosomal rearrangements (Figure 2a-d), whereas no clear separation of individuals was observed in the Flemish Cap or Gilbert Bay regions. Clustering along PC1 and PC2 axes corresponded to different combinations of linkage groups with known rearrangements contributing to structuring within each geographic region (Figure 2e-h). Bayesian clustering of LG1, LG2, LG7, and LG12 rearrangements within Europe and North America revealed ancestry proportion distributions consistent with homozygous or heterozygous rearrangement genotypes ($Q \sim 0, 0.5, 1$, Figure S4). We identified the presence of all three genotypes for each rearrangement across several sites in North America and Europe (Figure 3), indicating variation in rearrangement genotype is a consistent driver of individual genetic differences across many Atlantic cod populations (Table S2).

3.2 | Coinheritance and linkage disequilibrium

Consistent with independent inheritance of rearrangements, we observed low LD of outlier SNPs on different rearrangements within any geographic regions (Figure 4a-d). However, patterns of interchromosomal LD varied slightly between regions, with small elevation in LD observed in Iceland and Northeast Arctic samples relative to other regions (Figure 3). We identified high intrachromosomal LD within each rearrangement, consistent with reduced recombination between alternative rearrangement orientations (Figure 4a-d). Linkage networks also supported extensive LD within rearrangements, and low LD between different rearrangements within each geographic cluster at an r^2 cutoff of .25, indicating independent inheritance of each rearrangement within each geographic cluster (Figure 4e-h).

FIGURE 2 Clustering of individuals from sampling sites within major population genetic clusters and single nucleotide polymorphism (SNP) correlations with principal component axes associated with individual genetic structure within each major cluster. (a–d) Individual clustering of samples on the first two principal component axes in *pcadapt* in Canada North, Iceland and Northeast Arctic, Canada South, and Coastal Europe population clusters. (e–h) Significance of correlation of each SNP (p value) with the $K = 25$ retained principal components in *pcadapt* in Canada North, Iceland and Northeast Arctic, Canada South, and Coastal Europe population clusters. Chromosome rearrangements are marked in red



3.3 | Environmental and migratory association with genomic variation

Tests of polygenic environmental adaptation revealed associations primarily within chromosomal rearrangements in both North America and Europe, supporting a role for these regions in local adaptation. Redundancy analysis (RDA) indicated that variation in principal components of salinity, oxygen, and temperature was strongly associated in with LG1, LG2, LG7, and LG12 rearrangements in both North America and Europe (Figure 5a,b). Random forest regression of SNPs with environmental variables also revealed consistent association of rearrangements with variation in temperature, oxygen, and salinity (Table 1, Figures S5 and S6). We also identified a shared positive association with at least one temperature variable and LG2, LG7, and LG12 rearrangement haplotype frequency on either side of the Atlantic (Table S3).

Redundancy analysis and random forest classification have been previously used to identify an association between migratory behavior

and the LG1 rearrangement in North America (Kess et al., 2019). Using both of these methods to repeat this analysis in European waters, we found migratory phenotype associations predominantly with SNPs within the LG1 rearrangement, (Figure S7), suggesting parallelism of genetic architecture for this trait. We also found that frequency of the rearranged LG1 haplotype was positively correlated with migration in both North America and Europe (Table S3).

3.4 | Clustering of loci by multiple environmental associations

We found that most SNPs within rearranged regions belonged to the same functional cluster, indicating functional modularity, but in several instances outlier SNPs within multiple rearranged regions were found in the same environmental co-association cluster (Figure 6a,b). In North America, environmental clusters were driven by strength of association with principal components. Although

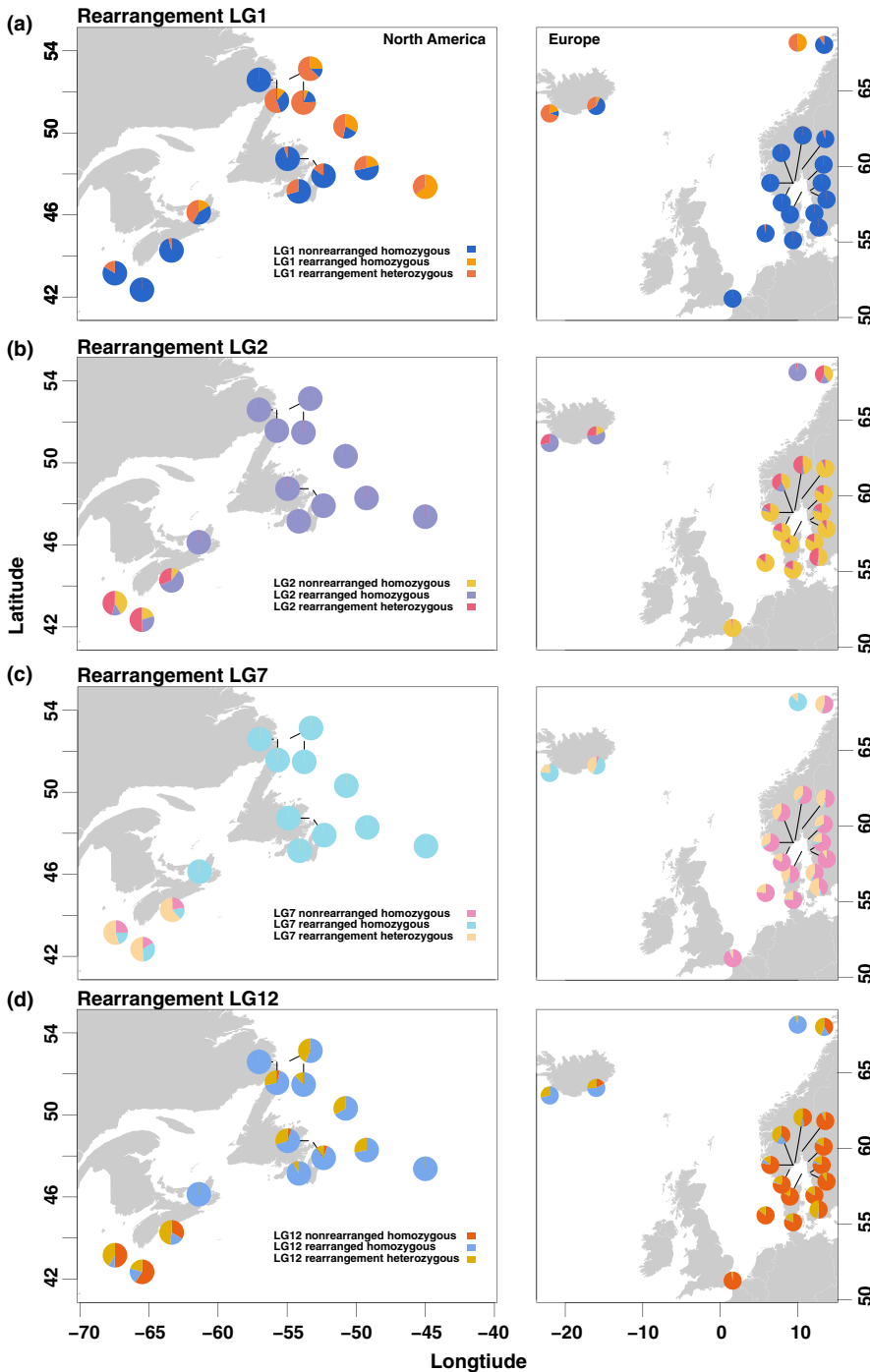


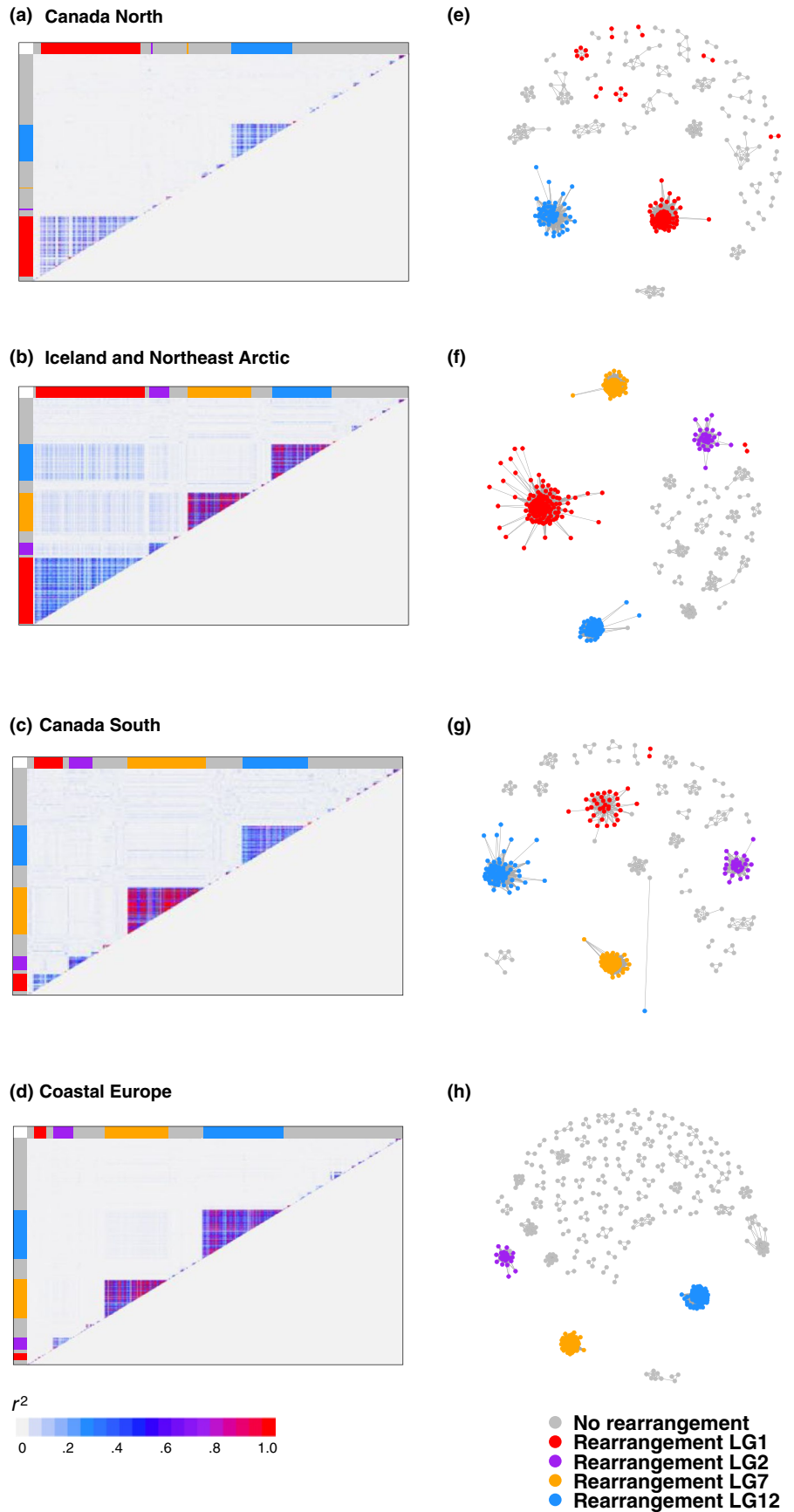
FIGURE 3 Proportion of homozygous rearranged, homozygous non-rearranged, and heterozygous individuals for LG1, LG2, LG7, and LG12 rearrangements in sites in North America and Europe. (a) Proportions of LG1 genotypes in North America and Europe (b) Proportions of LG2 genotypes in North America and Europe (c) Proportions of LG7 genotypes in North America and Europe (d) Proportions of LG12 genotypes in North America and Europe

SNPs within the same rearranged region tended to show similar patterns of co-association, we observed SNPs from each rearrangement within the same cluster across multiple major co-association clusters (Figure 6a). In Europe, environmental clustering was driven by associations between principal components as well as individual environmental variables. The pattern of co-association of SNPs with the same environmental variable across rearrangements was less pronounced in Europe, indicating greater functional modularity. Reduced co-association between rearrangements was greatest in LG12; SNPs in LG12 exhibited almost no clustering with SNPs within other rearrangements (Figure 6b).

3.5 | Signatures of adaptation from shared variation

For each rearrangement, trees of genetic distance showed groupings distinct to rearrangement genotype. For LG1, individual genetic distance trees revealed no separation of LG1 homozygous rearranged genotype individuals by geography, supporting a hypothesis of adaptation from shared variation for these rearrangements (Figure 7a). Distinct clusters were identified for European and North American nonrearranged genotype and heterozygous genotype individuals, indicating high trans-Atlantic differentiation for the nonrearranged LG1 orientation.

FIGURE 4 Linkage disequilibrium between outlier single nucleotide polymorphisms (SNPs) significantly correlated with principal component axes describing individual genetic variation in *pcadapt* in Canada North, Iceland and Northeast Arctic, Canada South, and Coastal Europe population clusters. (a–d) Linkage as measured by r^2 between outlier SNPs visualized using LD heatmaps (e–h) Linkage between outlier SNPs visualized in a linkage network in *LDna* with an r^2 cutoff of .25. Chromosomal rearrangements are represented by separate colors



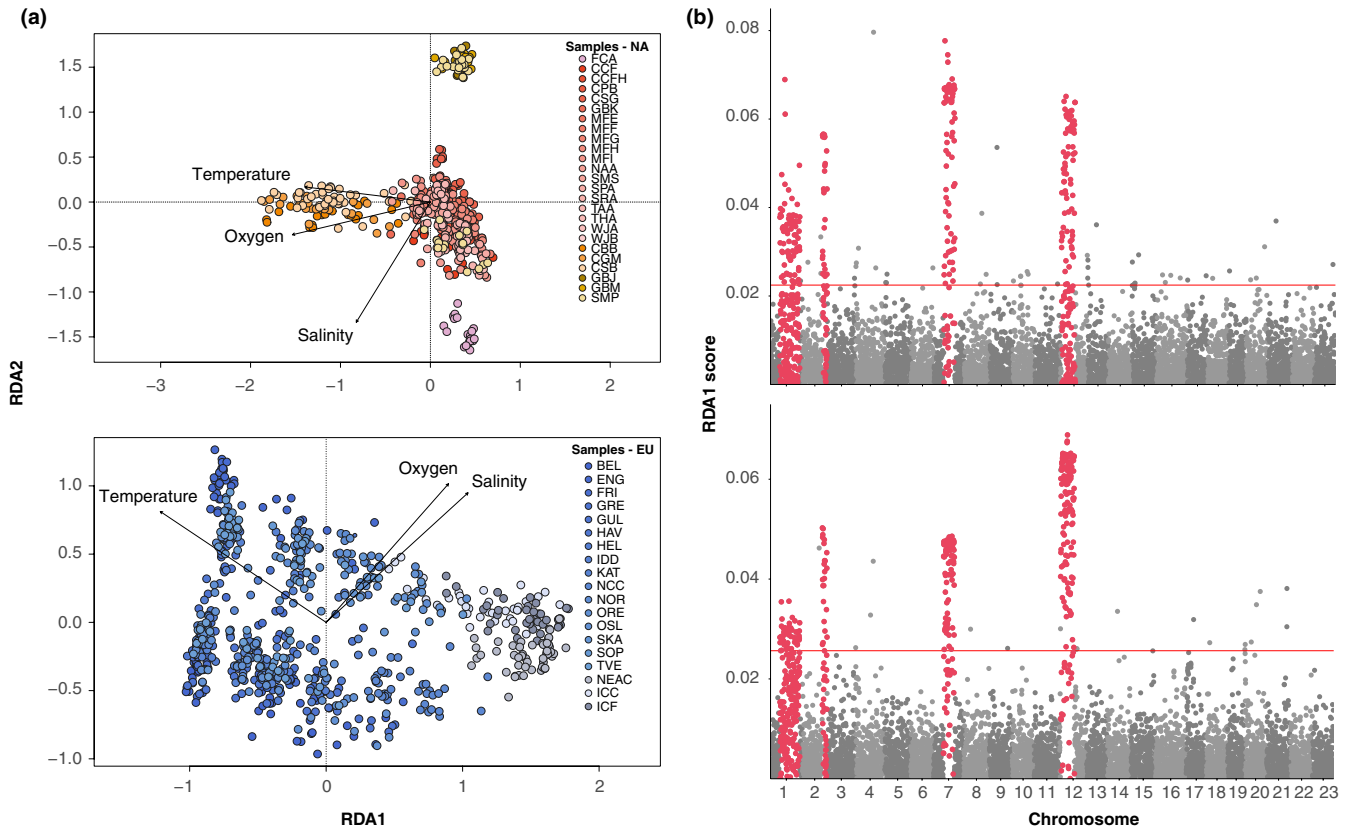


FIGURE 5 Redundancy analysis of single nucleotide polymorphisms (SNPs) with temperature, salinity, and oxygen in North America and Europe. (a) Biplots of RDA1 and RDA2 scores for each individual color coded by population with vectors describing variation due to oxygen, temperature, and salinity gradients in North America and Europe. (b) Absolute values of scores for each SNP on RDA1 within North America and Europe. Significant SNPs are those with scores greater than the 95th percentile. Chromosome rearrangements are marked in red

In LG2 and LG7, we observed separate clusters corresponding to homozygous rearranged, homozygous nonrearranged, and heterozygous genotype individuals (Figure 7b–c), which exhibited low trans-Atlantic differentiation relative to divergence by rearrangement genotype. Genetic distances between these clusters were larger in LG7 relative to LG2, indicating different rates or of differentiation.

In LG12, we observed low divergence among rearranged genotype individuals, suggesting this region is also shared across the North Atlantic (Figure 7d). We found strong separation between North America and Europe in nonrearranged genotype individuals in LG12 and observed distinct heterozygous genotype clusters for these regions, reflecting divergence in nonrearranged LG12 genotype individuals (Figure 7d).

Comparison of differentiation between North American and European individuals at the same rearranged genomic region revealed patterns consistent with adaptation from shared variation for each rearranged region, but also revealed elevated trans-Atlantic divergence within nonrearranged haplotypes for LG1 and LG12 (Figure 7e,g,h). These values were significantly below the chromosome-wide average for rearranged genotype individuals in LG1 ($W = 5,129$, $p < .04$, $F_{ST} = 0.0184$ rearranged region, $F_{ST} = 0.0516$ collinear region) and rearranged genotype individuals in LG12 ($W = 13,332$, $p < 1.3 \times 10^{-7}$,

$F_{ST} = 0.0328$ rearranged region, $F_{ST} = 0.0835$ collinear region), but were not significantly different in LG2. Nonrearranged genotype individuals also exhibited reduced F_{ST} in the structurally variable region of LG7 ($W = 5,975$, $p < .001$, $F_{ST} = 0.0464$ rearranged region, $F_{ST} = 0.0563$ collinear region). In contrast, LG12 nonrearranged genotype individuals ($W = 4,063$, $p < 2.2 \times 10^{-16}$, $F_{ST} = 0.505$ rearranged region, $F_{ST} = 0.084$ collinear region Figure 7d) and LG1 nonrearranged genotype individuals ($W = 8,966.5$, $p < 3.1 \times 10^{-13}$, $F_{ST} = 0.191$ rearranged region, $F_{ST} = 0.063$ collinear region) exhibited significantly elevated differentiation between North America and Europe. F_{ST} was consistently high at rearranged regions relative to regions without rearrangements (Figure S8). We found that d_{XY} also tracked the same pattern as F_{ST} (Figure S9a–d).

Observed nucleotide diversity (π) values differed between rearranged and nonrearranged groups on LG1; rearranged genotype individuals in Europe ($W = 192$, $p < 3.2 \times 10^{-5}$) and North America ($W = 189$, $p < 5.61 \times 10^{-6}$) had lower nucleotide diversity than the remainder of the chromosome (Figure S9e–f). We found that π was reduced in structurally variable regions on LG12 for both rearranged genotype and nonrearranged genotype individuals in Europe (nonrearranged: $W = 163$, p -value = .002078, rearranged: $W = 191$, $p < 9.5 \times 10^{-7}$) and North America (nonrearranged: $W = 153$, p -value < .011, rearranged: $W = 166$, $p < .0012$).

TABLE 1 Number of single nucleotide polymorphisms (SNPs) associated with temperature, dissolved O₂, and salinity in North America and Europe identified using random forest regression, and proportion of environment-associated SNPs in chromosomal rearrangements

Location	Variable	MDA threshold	Associated SNPs	Proportion of environment-associated SNPs in rearrangements
Europe	Maximum temperature	0.05	31	0.81
	Mean temperature	0.02	12	0.92
	Minimum temperature	0.01	22	0.36
	Max dissolved O ₂	1.0	58	0.53
	Mean dissolved O ₂	0.5	27	0.30
	Minimum dissolved O ₂	0.05	81	0.72
	Maximum salinity	0.01	18	0.50
	Mean salinity	0.01	24	0.54
	Minimum salinity	0.01	48	0.57
North America	Maximum temperature	0.05	133	0.67
	Mean temperature	0.01	41	0.95
	Minimum temperature	0.05	43	0.95
	Max dissolved O ₂	10	56	0.95
	Mean dissolved O ₂	10	72	0.97
	Minimum dissolved O ₂	5	83	0.98
	Maximum salinity	0.002	51	0.86
	Mean salinity	0.005	22	0.82
	Minimum salinity	0.01	48	0.79

4 | DISCUSSION

Genomic architecture and standing variation can be crucial in determining evolutionary trajectories, particularly for instances of parallel evolution (Bolnick et al., 2018; Elmer & Meyer, 2011). In fact, genomic features may represent intrinsic drivers of adaptation and speciation, but their roles in this process have been underappreciated until recently (Campbell, Poelstra, & Yoder, 2018). Here, we investigated the extent of functional modularity and genomic parallelism in rearranged genomic regions in Atlantic cod and demonstrate that all four chromosomal rearrangements show low interchromosomal linkage within geographic regions, indicating these rearrangements exhibit independent inheritance. Further, our results support signatures of adaptation from shared variation in at least one chromosomal orientation in LG1, LG7, and LG12 chromosomal rearrangements, emphasizing the importance of shared architectural diversity. Interestingly, one chromosomal orientation for each of LG1 and LG12 has undergone trans-Atlantic differentiation, suggesting nonparallel genomic adaptation for each of these regions. Together, these results demonstrate the importance of modular genomic architecture and both shared and nonparallel components of structural variation in enabling greater differentiation within Atlantic cod than previously recognized.

We demonstrate low interchromosomal linkage of rearranged regions, suggesting independent evolutionary trajectories for each rearrangement. These results are contrary to previous findings of high interchromosomal linkage between rearranged regions (Bradbury et al., 2014) and provide evidence against a genome-wide barrier to reproduction. Our comparison of rearrangement genotypes within

geographic regions revealed that previous observed patterns of interchromosomal linkage likely represent differences in rearrangement frequencies between regions. Widespread individual variation identified in our study also highlights potential genomic mechanisms underlying the presence of sympatric ecotypes occurring across the species range. Our observation of regional and individual levels of genetic structure associated with chromosomal rearrangements is consistent with the finding of physiological, behavioral, and life history variation within multiple discrete cod stocks, and observations of genetic divergence occurring within individual geographic regions (Conroy, Calvert, Sherwood, & Grabowski, 2017; Knutsen et al., 2018; Olsen et al., 2008; Puncher et al., 2019; Roney et al., 2016; Sinclair-Waters, Bentzen, et al., 2018). Our identification of extensive individual-level variation and low linkage in rearrangement genotypes indicates a genetic mechanism underpinning this biocomplexity: Atlantic cod distribution across differing environments may be enabled by locally adapted stocks within geographic regions, each with ecotype and individual variation in chromosomal rearrangement genotype. Independent inheritance of adaptive chromosomal rearrangements could facilitate exchange of these rearrangements as distinct adaptive units, enabling rapid production of novel ecotypes and colonization of a variety of habitats (Kirkpatrick & Barrett, 2015).

We found a shared pattern of migratory association of the LG1 rearrangement in independent polygenic association tests in Europe and North America. Associations of LG1 with migratory phenotype found here are consistent with past genome scan studies using migratory and nonmigratory ecotypes from waters around Europe (Berg et al., 2017) and genome-wide association in North American waters (Kess

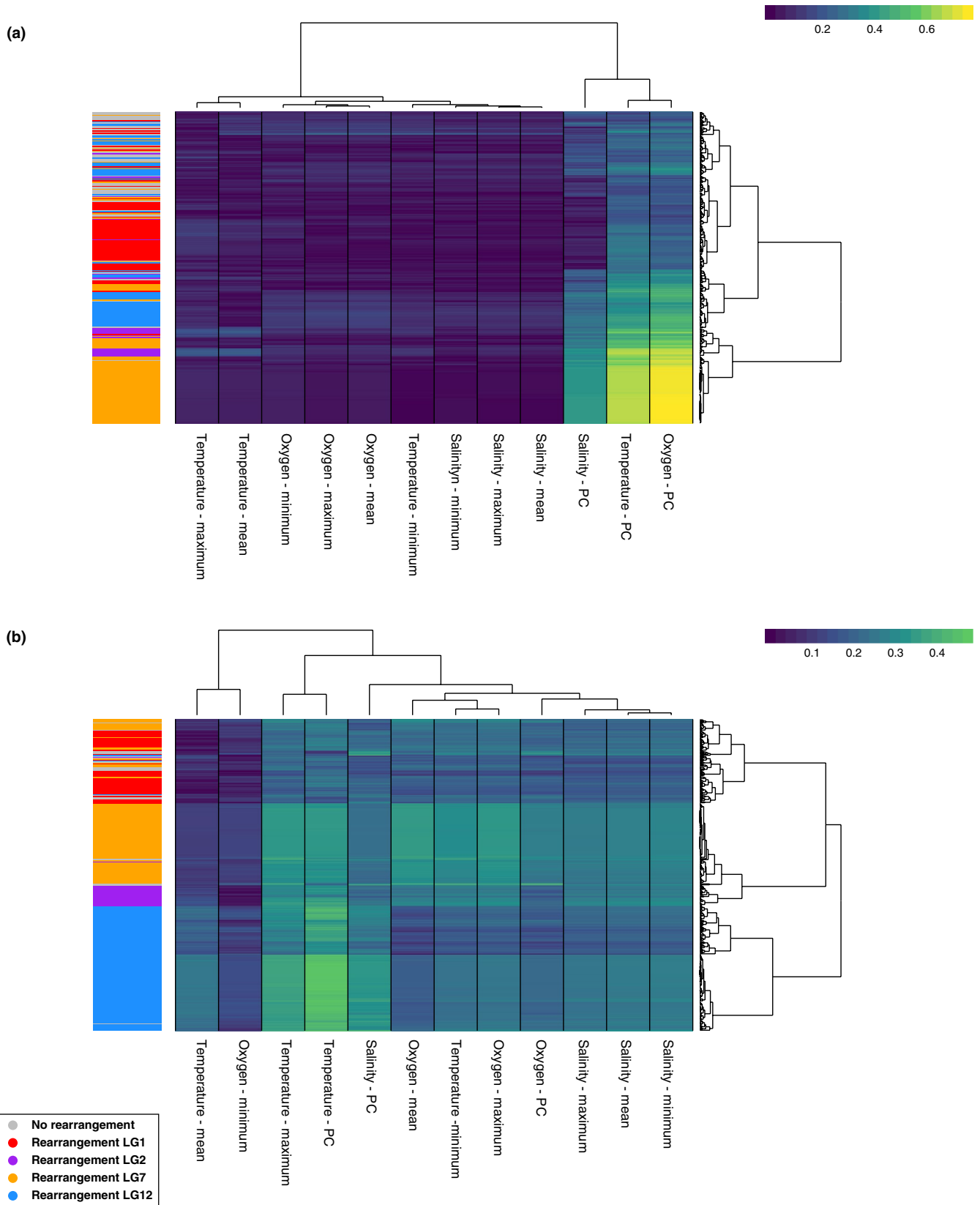


FIGURE 6 Co-association of RDA outlier single nucleotide polymorphisms (SNPs) with environmental variables grouped by hierarchical clustering of SNP association across environmental variables in North America (a) and Europe (b)

- Population cluster**
- Canada North
 - Canada South
 - Flemish Cap
 - Gilbert Bay
 - Iceland & Northeast Arctic
 - Coastal Europe

- Europe rearranged versus nonrearranged
- Europe rearranged versus North America nonrearranged
- North America rearranged versus Europe nonrearranged
- North America rearranged versus nonrearranged
- North America versus Europe rearranged
- North America versus Europe nonrearranged

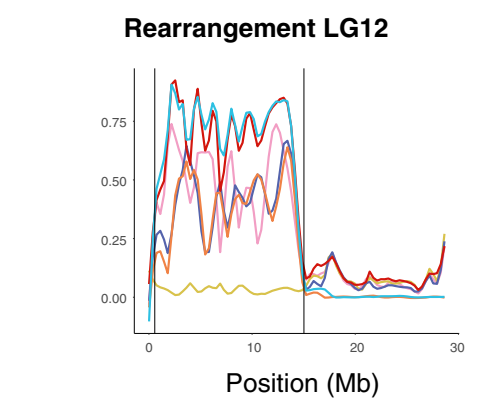
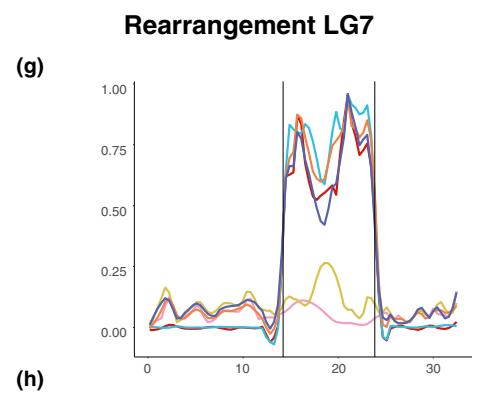
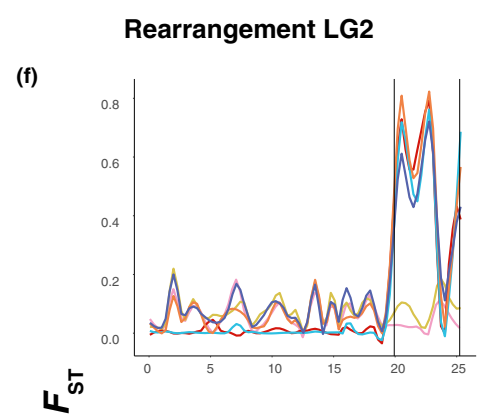
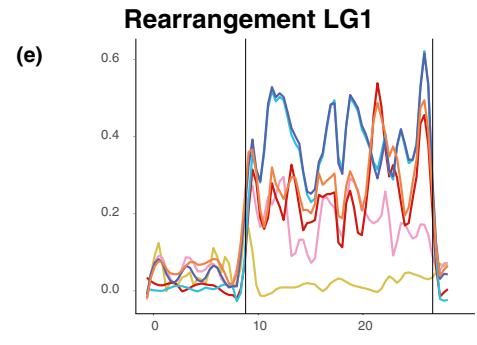
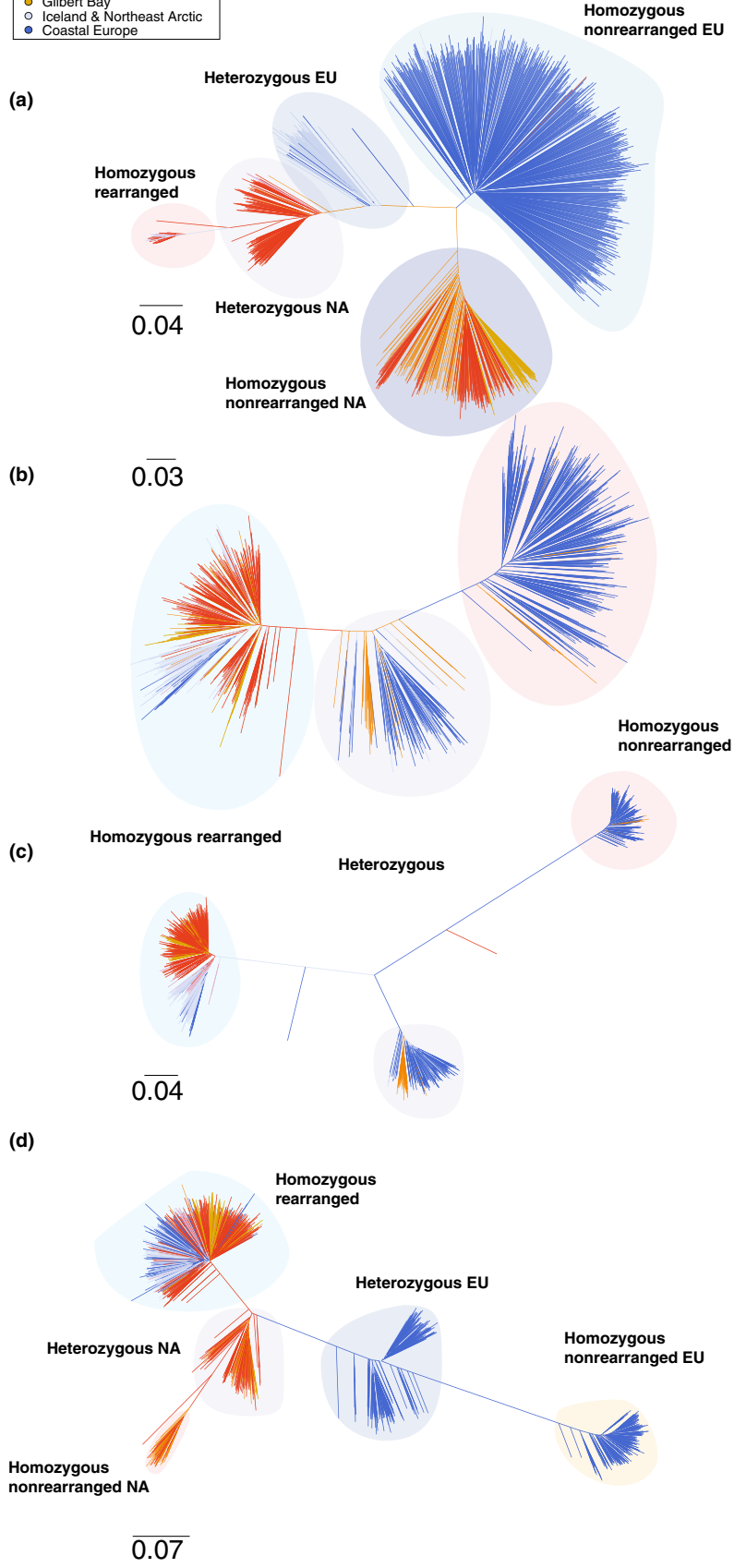


FIGURE 7 Comparison of F_{ST} along LG1, 2, 7, and 12 between individuals with each rearrangement genotype in North America and Europe and genetic distance trees for each rearrangement. (a–d) Comparison of F_{ST} along LG1, 2, 7, and 12 between alternative homozygous rearrangement genotype categories in North America and Europe. Black vertical lines indicate the boundaries of each rearrangement. (e–h) Range-wide individual trees of genetic distance for LG1, 2, 7, and 12 rearrangements calculated using Cavalli-Sforza and Edwards chord distances

et al., 2019). By using available migratory information for each site, we provide further support for the association of the LG1 rearrangement with migratory phenotype in the Eastern Atlantic. Recent molecular dating of this rearrangement indicates that its evolution substantially predates (1.6–2 Mya) postglacial colonization of Northwest Atlantic waters (Kirubakaran et al., 2016) and further supports the hypothesis of ancestral introduction of this rearrangement during colonization of North American waters (Sinclair-Waters, Bentzen, et al., 2018).

Consistent with past studies using single-locus association methods, we also find polygenic association of rearrangements with temperature, salinity, and oxygen availability in North America and Europe (Barney, Munkholm, Walt, & Palumbi, 2017; Berg et al., 2016; Bradbury et al., 2010). Although parallel clinal association is ecologically informative, autocorrelation of environmental variables within each site complicates interpretation of the exact adaptive role of each rearrangement (Hoban et al., 2016). We find associations for rearrangement frequencies with oxygen and salinity differed in direction and significance between North America and Europe, whereas the relationship with temperature and rearrangement frequency was significant and in the same direction on both continents, suggesting that temperature may primarily dictate clinal structuring of rearrangements.

Using multivariate co-association and clustering, we show that SNPs within each rearrangement exhibited similar patterns of environmental co-association. However, some SNPs overlapped in environmental associations between the rearrangements, suggesting a subset of loci are adapted to a shared set of environmental variables across rearrangements. This observation is consistent with these SNPs exhibiting functional modularity despite physical linkage. A similar pattern has also been observed among environment-adapted genomic regions in *Pinus contorta*, indicating physical separation among loci is not always necessary for adaptation to different environmental variables (Lotterhos et al., 2018). Functional modularity was greater in Europe than in North America; our identification of divergence between the same structurally variable chromosome regions in Europe and North America likely contributes to varying patterns of environmental association identified here. Future studies using common garden experiments and crosses with manipulation of individual environmental parameters will be required to more clearly identify the functions of these structural variants (Bierne, Bonhomme, Boudry, Szulkin, & David, 2006; Wang et al., 2018).

We find signatures of adaptation from shared variation across the Atlantic for chromosomal orientations specific to each chromosome. Using differentiation-based and genetic distance comparisons, we show low trans-Atlantic differentiation between the rearranged orientation of LG1, and LG12, and the nonrearranged orientation of LG7, suggesting selection on shared variation in these regions in both Europe and North America. These patterns of localized reduced

differentiation are consistent with incorporation of these genomic regions into the genetic architecture of adaptive traits from shared variation (Barrett & Schluter, 2008). This observation is similar to differentiation observed between marine and freshwater stickleback in comparison with the inverted region housing the *Atp1a1* gene, although the stickleback study focused specifically on multiple derived populations (Roesti, Gavrilets, Hendry, Salzburger, & Berner, 2014). Our results provide further support for the increasingly prominent role of shared structural variation in facilitating local adaptation (Barrett & Schluter, 2008; Kirkpatrick & Barton 2006).

We also uncovered elevated trans-Atlantic differentiation in the nonrearranged orientations of LG1 and LG12, indicating distinct evolutionary histories of these chromosomal regions in Europe and North America. The identification of elevated trans-Atlantic divergence within rearranged LG1 and LG12 genotype individuals suggests gene content in these regions may be amenable to repeated selection during ecological specialization in Atlantic cod. Repeated divergence of the same genomic region has been found to underlie parallel evolution of pelvic spine reduction in threespine stickleback through repeated deletions within *Pitx1* regulatory regions (Chan et al., 2010), and evolution of albinism in *Astyanax* cavefish has also arisen from independent deletions within the *Oca2* locus (Protas et al., 2006). Alternatively, reduced recombination and a long period of trans-Atlantic divergence in ancient rearrangements predating speciation may also drive this elevated divergence relative to the genomic background. Surprisingly, we find that nucleotide diversity is reduced in the LG12 rearranged region for nonrearranged individuals despite high trans-Atlantic F_{ST} and d_{XY} , suggesting diversity-reducing sweeps in these regions. However, low precision due to low density of SNPs, ascertainment bias, and high variance in d_{XY} when applied to SNP data may bias measures of diversity and divergence (Wakeley, 1996). Future characterization of long-read sequence from LG1 and LG12 across Europe and North American Atlantic cod may help clarify factors driving divergence identified here.

Shared, modular, structural variation is hypothesized to be integral to ecological adaptation and may facilitate repeatability of the genomic architecture underlying parallel adaptation. We show that modular, shared chromosomal rearrangements are associated with ecological adaptation within Atlantic cod. Additionally, we show that separate divergence in LG12 and LG1 rearrangements has occurred in North American and European populations, revealing a nonparallel component to adaptation within the same, structurally variable region of the genome. Our findings here also provide new understanding of biocomplexity organized at different spatial scales within this species. Substantial cryptic variation is mediated by these rearrangements, indicating different adaptive potentials among cod stocks. Understanding and conserving this variation may be crucial in

mitigating anthropogenic impacts on Atlantic cod such as those imposed by climate change and overharvest.

ACKNOWLEDGMENTS

We thank the Centre for Integrative Genetics (CIGENE) for genotyping of Atlantic cod used in this study. This study was supported by the Genomics Research and Development Initiative (GRDI) of the Department of Fisheries and Oceans Canada (DFO) and the Natural Sciences Engineering and Research Council Canada (NSERC).

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

TK, SJL, IRB, and PB designed the study. TK, EVAS, and RF carried out statistical analyses. PB, IRB, MS-W, MPK, and SL provided the molecular data. TK drafted the manuscript, and all authors contributed to writing and editing of the final manuscript.

ORCID

Tony Kess  <https://orcid.org/0000-0002-1079-3791>

DATA AVAILABILITY STATEMENT

A plink format file of SNP genotypes for all genotyped individuals, lists of SNPs included in inversions, migratory phenotypes for European individuals, and environmental data for Europe and North America are available on dryad (<https://doi.org/10.5061/dryad.f1vhhmgsg>).

REFERENCES

- Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., & De Clerck, O. (2018). Bio-ORACLE v2. 0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, 27(3), 277–284. <https://doi.org/10.1111/geb.12693>
- Barb, J. G., Bowers, J. E., Renaut, S., Rey, J. I., Knapp, S. J., Rieseberg, L. H., & Burke, J. M. (2014). Chromosomal evolution and patterns of introgression in *Helianthus*. *Genetics*, 197, 969–979.
- Barney, B. T., Munkholm, C., Walt, D. R., & Palumbi, S. R. (2017). Highly localized divergence within supergenes in Atlantic cod (*Gadus morhua*) within the Gulf of Maine. *BMC Genomics*, 18, 271. <https://doi.org/10.1186/s12864-017-3660-3>
- Barrett, R., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, 23, 38–44. <https://doi.org/10.1016/j.tree.2007.09.008>
- Barter, R. L., & Yu, B. (2017). Superheat: An R package for creating beautiful and extendable heatmaps for visualizing complex data. *Journal of Computational and Graphical Statistics*, 27, 910–922.
- Barth, J. M. I., Berg, P. R., Jonsson, P. R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., ... 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>
- Bassham, S., Catchen, J., Lescak, E., von Hippel, F. A., & Cresko, W. A. (2018). Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. *Genetics*, 209(3), 921–939. <https://doi.org/10.1534/genetics.117.300610>
- Berg, P. R., Jentoft, S., Star, B., Ring, K. H., Knutsen, H., Lien, S., ... André, C. (2015). Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biology and Evolution*, 7, 1644–1663. <https://doi.org/10.1093/gbe/evv093>
- Berg, P. R., Star, B., Pampoulie, C., Bradbury, I. R., Bentzen, P., Hutchings, J. A., ... Jakobsen, K. S. (2017). Trans-oceanic genomic divergence of Atlantic Cod ecotypes is associated with large inversions. *Heredity*, 119(6), 418–428. <https://doi.org/10.1038/hdy.2017.54>
- Berg, P. R., Star, B., Pampoulie, C., Sodeland, M., Barth, J. M. I., Knutsen, H., ... Jentoft, S. (2016). Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic Cod. *Scientific Reports*, 6(23246), 1–12. <https://doi.org/10.1038/srep23246>
- Besnier, F., & Glover, K. A. (2013). ParallelStructure: A R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. *PLoS ONE*, 8, e70651. <https://doi.org/10.1371/journal.pone.0070651>
- Bierne, N., Bonhomme, F., Boudry, P., Szulkin, M., & David, P. (2006). Fitness landscapes support the dominance theory of post-zygotic isolation in the mussels *Mytilus edulis* and *M. galloprovincialis*. *Proceedings of the Royal Society B: Biological Sciences*, 273, 1253–1260.
- Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. *Science*, 362, eaam5979.
- Bolnick, D. I., Barrett, R. D. H., Oke, K. B., Rennison, D. J., & Stuart, Y. E. (2018). (Non)Parallel evolution. *The Annual Review of Ecology, Evolution, and Systematics*, 49, 303–330. <https://doi.org/10.1146/annurev-ecolsys-110617-062240>
- Bosch, S. (2018). *sdmpredictors: Species distribution modelling predictor datasets. R package version 0.2.8*. Retrieved from <http://CRAN.R-project.org/package=sdmpredictors>
- Bradbury, I. R., Bowman, S., Borza, T., Snelgrove, P. V. R., Hutchings, J. A., Berg, P. R., ... Bentzen, P. (2014). Long distance linkage disequilibrium and limited hybridization suggest cryptic speciation in Atlantic cod. *PLoS ONE*, 9, e106380. <https://doi.org/10.1371/journal.pone.0106380>
- Bradbury, I. R., Hubert, S., Higgins, B., Borza, T., Bowman, S., Paterson, I. G., ... Bentzen, P. (2010). Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proceedings of the Royal Society B: Biological Sciences*, 277, 3725–3734.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next generation reference panels. *American Journal of Human Genetics*, 103(3), 338–348. <https://doi.org/10.1016/j.ajhg.2018.07.015>
- Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018). What is speciation genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, 124(4), 561–583. <https://doi.org/10.1093/biolinnean/bly063>
- Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics*, 19, 233–257.
- Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Shapiro, M. D., Brady, S. D., ... Kingsley, D. M. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science*, 327(5963), 302–305. <https://doi.org/10.1126/science.1182213>
- Chang, C., Chow, C., Tellier, L., Vattikuti, S., & Purcell, S. M. & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Conroy, C. W., Calvert, J., Sherwood, G. D., & Grabowski, J. H. (2017). Distinct responses of sympatric migrant and resident Atlantic cod phenotypes to substrate and temperature at a remote Gulf of Maine

- seamount. *ICES Journal of Marine Science*, 75(1), 122–134. <https://doi.org/10.1093/icesjms/fsx101>
- 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, 5039–5047.
- Donaldson, J. (2016). *tsne: T-distributed stochastic neighbour embedding for R (t-SNE)*. R package version 0.1-3. Retrieved from <https://CRAN.R-project.org/package=tsne>
- Elmer, K. R., & Meyer, A. (2011). Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology & Evolution*, 26, 298–306. <https://doi.org/10.1016/j.tree.2011.02.008>
- Feder, J. L., Berlocher, S. H., Roethele, J. B., Dambroski, H., Smith, J. J., Perry, W. L., ... Aluja, M. (2003). Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences of the United States of America*, 100(18), 10314–10319. <https://doi.org/10.1073/pnas.1730757100>
- 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, 4074–4088. <https://doi.org/10.1111/mec.12750>
- Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, 25(1), 104–120. <https://doi.org/10.1111/mec.13476>
- Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Molecular Ecology*, 27, 2215–2233. <https://doi.org/10.1111/mec.14584>
- Gagnaire, P.-A., & Gaggiotti, O. E. (2016). Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. *Current Zoology*, 62, 603–616. <https://doi.org/10.1093/cz/zow088>
- Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., ... Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, 188(4), 379–397. <https://doi.org/10.1086/688018>
- 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 and Systematics*, 39, 21–42. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173532>
- Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., ... Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484, 55–61. <https://doi.org/10.1038/nature10944>
- Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). *diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors*. *Methods in Ecology and Evolution*, 4(8), 782–788.
- Kempainen, P., Knight, C. G., Sarma, D. K., Hlaing, T., Prakash, A., Maung Maung, Y. N., ... 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., Bentzen, P., Lehnert, S. J., Sylvester, E. V., Lien, S., Kent, M. P., ... & Bradbury, I. R. (2019). A migration-associated supergene reveals loss of biocomplexity in Atlantic cod. *Science Advances*, 5(6), eaav2461.
- Kirkpatrick, M., & Barrett, B. (2015). Chromosome inversions, adaptive cassettes and the evolution of species' ranges. *Molecular Ecology*, 24(9), 2046–2055. <https://doi.org/10.1111/mec.13074>
- Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, 173, 419–434. <https://doi.org/10.1534/genetics.105.047985>
- Kirubakaran, T. G., Grove, H., Kent, M. P., Sandve, S. R., Baranski, M., Nome, T., ... Andersen, Ø. (2016). Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Molecular Ecology*, 25, 2130–2143. <https://doi.org/10.1111/mec.13592>
- Knutsen, H., Jorde, P. E., Hutchings, J. A., Hemmer-Hansen, J., Grønkvær, P., Jørgensen, K.-E., ... Olsen, E. M. (2018). Stable coexistence of genetically divergent Atlantic cod ecotypes at multiple spatial scales. *Evolutionary Applications*, 11(9), 1527–1539. <https://doi.org/10.1111/eva.12640>
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15, 1179–1191.
- Langella, O. (2012). *POPULATIONS 1.2.28. Population genetic software (individuals or populations distances, phylogenetic trees)*. Retrieved from <https://launchpad.net/~olivier-langella/+archive/ubuntu/ppa>
- Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, 2, 18–22.
- Lindtke, D., Lucek, K., Soria-Carrasco, V., Villoutreix, R., Farkas, T. E., Riesch, R., ... Nosil, P. (2017). Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Molecular Ecology*, 26, 6189–6205. <https://doi.org/10.1111/mec.14280>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28, 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- Lotterhos, K., Yeaman, S., Degner, J., Aitken, S., & Hodgins, K. (2018). Modularity of genes involved in local adaptation to climate despite physical linkage. *Genome Biology*, 19, 157.
- Luu, K., Bazin, E., & Blum, M. G. (2017). *pcadapt: An R package to perform genome scans for selection based on principal component analysis*. *Molecular Ecology Resources*, 17, 67–77.
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18, 59–60. <https://doi.org/10.1214/aoms/1177730491>
- Nelson, T. C., & Cresko, W. A. (2018). Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evolution Letters*, 2, 9–21. <https://doi.org/10.1002/evl3.37>
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G. L., ... Wagner, H. (2011). *vegan: Community Ecology Package*. R package version 1.17-6. Retrieved from <http://CRAN.R-project.org/package=vegan>
- Olsen, E. M., Knutsen, H., Gjøsæter, J., Jorde, P. E., Knutsen, J. A., & Stenseth, N. C. (2008). Small-scale biocomplexity in coastal Atlantic cod supporting a Darwinian perspective on fisheries management. *Evolutionary Applications*, 1, 524–533. <https://doi.org/10.1111/j.1752-4571.2008.00024.x>
- Pearse, D. E., Miller, M. R., Abadia-Cardoso, A., & Garza, J. C. (2014). Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20140012. <https://doi.org/10.1098/rspb.2014.0012>
- Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, 31, 1929–1936. <https://doi.org/10.1093/molbev/msu136>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Protas, M. E., Hersey, C., Kochanek, D., Zhou, Y. I., Wilkens, H., Jeffery, W. R., ... Tabin, C. J. (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics*, 38, 107–111. <https://doi.org/10.1038/ng1700>

- Puncher, G. N., Rowe, S., Rose, G. A., Leblanc, N. M., Parent, G. J., Wang, Y., & Pavey, S. A. (2019). Chromosomal inversions in the Atlantic cod genome: Implications for management of Canada's Northern cod stock. *Fisheries Research*, 216, 29–40. <https://doi.org/10.1016/j.fishres.2019.03.020>
- Ragland, G. J., Doellman, M. M., Meyers, P. J., Hood, G. R., Egan, S. P., Powell, T. H. Q., ... Feder, J. L. (2017). A test of genomic modularity among life-history adaptations promoting speciation with gene flow. *Molecular Ecology*, 26, 3926–3942.
- Rambaut, A. (2012). *FigTree v1.4*. Retrieved from <http://tree.bio.ed.ac.uk/software/figtree/>
- Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhya: The Indian Journal of Statistics, Series A*, 26, 329–358.
- Robichaud, D., & Rose, G. A. (2004). Migratory behaviour and range in Atlantic cod: Inference from a century of tagging. *Fish and Fisheries*, 5, 185–214. <https://doi.org/10.1111/j.1467-2679.2004.00141.x>
- Roesti, M., Gavrillets, S., Hendry, A. P., Salzburger, W., & Berner, D. (2014). The genomic signature of parallel adaptation from shared genetic variation. *Molecular Ecology*, 23(16), 3944–3956. <https://doi.org/10.1111/mec.12720>
- Rogers, S. M., Mee, J. A., & Bowles, E. (2013). The consequences of genomic architecture on ecological speciation in postglacial fishes. *Current Zoology*, 59, 53–71. <https://doi.org/10.1093/czoolo/59.1.53>
- Roney, N. E., Hutchings, J. A., Olsen, E. M., Knutsen, H., Albretson, J., & Kuparinen, A. (2016). Fine-scale life-history structure in a highly mobile marine fish. *Evolutionary Ecology Research*, 17, 95–109.
- Sinclair-Waters, M., Bentzen, P., Morris, C. J., Ruzzante, D. E., Kent, M., Lien, S., & Bradbury, I. R. (2018). Genomic tools for management and conservation of Atlantic Cod in a coastal marine protected area. *Canadian Journal of Fisheries and Aquatic Sciences*, 75(11), 1915–1925. <https://doi.org/10.1139/cjfas-2017-0254>
- Sinclair-Waters, M., Bradbury, I. R., Morris, C. J., Lien, S., Kent, M. P., & Bentzen, P. (2018). Ancient chromosomal rearrangement associated with local adaptation of a post-glacially colonized population of Atlantic cod in the northwest Atlantic. *Molecular Ecology*, 27, 339–351. <https://doi.org/10.1111/mec.14442>
- Sodeland, M., Jorde, P. E., Lien, S., Jentoft, S., Berg, P. R., Grove, H., ... Knutsen, H. (2016). "Islands of Divergence" in the Atlantic Cod genome represent polymorphic chromosomal rearrangements. *Genome Biology and Evolution*, 8(4), 1012–1022. <https://doi.org/10.1093/gbe/evw057>
- Stanley, R. R., Jeffery, N. W., Wringe, B. F., DiBacco, C., & Bradbury, I. R. (2016). genepopedit: A simple and flexible tool for manipulating multilocus molecular data in R. *Molecular Ecology Resources*, 17(1), 12–18.
- Stephan, J., Stegle, O., & Beyer, A. (2015). A random forest approach to capture genetic effects in the presence of population structure. *Nature Communications*, 6, 10.
- Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, 14, 751–764. <https://doi.org/10.1038/nrg3483>
- Stern, D. L., & Orgogozo, V. (2007). The loci of evolution: How predictable is genetic evolution? *Evolution*, 62(9), 2155–2177. <https://doi.org/10.1111/j.1558-5646.2008.00450.x>
- Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genome-wide studies. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 9440–9445. <https://doi.org/10.1073/pnas.1530509100>
- Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, 25, 2144–2164. <https://doi.org/10.1111/mec.13606>
- van der Maaten, L., & Hinton, G. E. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.
- Wagner, G. P., Pavlicev, M., & Cheverud, J. M. (2007). The road to modularity. *Nature Reviews Genetics*, 8, 921–931. <https://doi.org/10.1038/nrg2267>
- Wakeley, J. (1996). The variance of pairwise nucleotide differences in two populations with migration. *Theoretical Population Biology*, 49, 39–57. <https://doi.org/10.1006/tpbi.1996.0002>
- Wang, J., Ding, J., Tan, B., Robinson, K. M., Michelson, I. H., Johansson, A., ... Ingvarsson, P. K. (2018). A major locus controls local adaptation and adaptive life history variation in a perennial plant. *Genome Biology*, 19, 72. <https://doi.org/10.1186/s13059-018-1444-y>
- 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>
- Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., ... Aitken, S. N. (2016). Convergent local adaptation to climate in distantly related conifers. *Science*, 353(6306), 1431–1433. <https://doi.org/10.1126/science.aaf7812>

SUPPORTING INFORMATION

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

How to cite this article: Kess T, Bentzen P, Lehnert SJ, et al. Modular chromosome rearrangements reveal parallel and nonparallel adaptation in a marine fish. *Ecol Evol*. 2020;10:638–653. <https://doi.org/10.1002/ece3.5828>