

# Reproductive isolation in a three-way contact zone

Laura L. Dean<sup>1</sup>  | James R. Whiting<sup>1,2</sup>  | Felicity C. Jones<sup>3,4</sup>  | Andrew D. C. MacColl<sup>1</sup> 

<sup>1</sup>School of Life Sciences, University of Nottingham, Nottingham, UK

<sup>2</sup>Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada

<sup>3</sup>Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany

<sup>4</sup>Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, The Netherlands

## Correspondence

Laura L. Dean and Andrew D. C. MacColl, School of Life Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK.

Email: [laura.dean@nottingham.ac.uk](mailto:laura.dean@nottingham.ac.uk) and [andrew.maccoll@nottingham.ac.uk](mailto:andrew.maccoll@nottingham.ac.uk)

## Funding information

Natural Environment Research Council, Grant/Award Number: NE/R00935X/1; Deutsche Forschungsgemeinschaft, Grant/Award Number: SPP1819

Handling Editor: Maren Wellenreuther

## Abstract

Contact zones between divergent forms within a species provide insight into the role of gene flow in adaptation and speciation. Previous work has focused on contact zones involving only two divergent forms, but in nature, many more than two populations may overlap simultaneously and experience gene flow. Patterns of introgression in wild populations are, therefore, likely much more complicated than is often assumed. We begin to address this gap in current knowledge by investigating patterns of divergence and introgression across a complex natural contact zone. We use phenotypic and genomic data to confirm the existence of a three-way contact zone among divergent freshwater resident, saltwater resident and saltwater migratory three-spined stickleback (*Gasterosteus aculeatus*) on the island of North Uist, Scottish Western Isles. We find evidence for hybridization, mostly between saltwater resident and saltwater migratory forms. Despite hybridization, genomic analyses reveal pairwise islands of divergence between all forms that are maintained across the contact zone. Genomic cline analyses also provide evidence for selection and/or hybrid incompatibilities in divergent regions. Divergent genomic regions occur across multiple chromosomes and involve many known adaptive loci and several chromosomal inversions. We also identify distinct immune gene expression profiles between forms, but no evidence for transgressive expression in hybrids. Our results suggest that reproductive isolation is maintained in this three-way contact zone, despite some hybridization, and that reduced recombination in chromosomal inversions may play an important role in maintaining this isolation.

## KEYWORDS

adaptation, contact zone, hybrid, introgression, reproductive isolation, speciation

## 1 | INTRODUCTION

Contact zones between divergent forms of the same species provide an excellent opportunity to study how isolating barriers are formed and disintegrated during the process of speciation. Contact zones are typically found in areas of ecological transition, where an environmental gradient exists between habitats, leading to overlap between divergent populations that are adapted to

different environments (Abbott, 2017; Johannesson et al., 2020; Keller & Seehausen, 2012). If speciation is advanced, hybridization across contact zones may be very limited or even non-existent (Wu et al., 2023). If divergence between overlapping populations is low, hybridization may involve multiple generations of hybrid offspring that can backcross to one or both parental populations, allowing genetic exchange (introgression) between divergent forms (Cahan et al., 2023).

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

There are multiple possible outcomes of introgression across contact zones, with varying consequences for speciation. With little or no selection against hybrid genotypes, genomes may become homogenized and differences between divergent populations can become reduced or completely eliminated (Taylor et al., 2006; Xiong & Mallet, 2022). Alternatively, selection against hybrid genotypes that perform sub-optimally in both parental environments may lead to the maintenance or even exaggeration (via reinforcement) of pre-existing differences (Hoarau et al., 2015; Servedio & Noor, 2003). Depending on the strength of selection, this can lead to a stable transition zone with low levels of hybridization (O'Connell et al., 2021). In this case, introgression across the contact zone can introduce novel genomic material to both parental populations, on which selection may then act, potentially driving new environmental adaptations (Gaczorek et al., 2023). It can even contribute to speciation via the formation of new hybrid taxa (Abbott et al., 2013).

Epistatic interactions between genes are critical for the performance of hybrid genotypes. The Bateson–Dobzhansky–Muller model proposes that deleterious interactions between alleles that are harmless or even advantageous in the parental genomic background can result in hybrid disadvantage (Coyne & Orr, 2004). In plants, deleterious epistatic interactions are commonly linked to autoimmune responses in hybrids, resulting in hybrid necrosis (Bomblies et al., 2007; Chae et al., 2014; Freh et al., 2022). In animals, the importance of immune genes and autoimmunity in causing hybrid dysfunction is not yet known, but incompatibilities are more likely to arise in rapidly evolving genes with a high mutation rate (Presgraves, 2010). Several studies have identified substantial differences in gene expression between hybrids and their parental species (Maheshwari & Barbash, 2011), such that gene expression in hybrids is not necessarily additive. In *Drosophila*, multiple hybrid incompatibilities result in gene expression profiles in the two parent species being more similar to each other than those of the hybrids (Ranz et al., 2004). Gene expression profiles can, therefore, reveal potential areas of hybrid dysfunction.

The dynamics of contact zones are usually studied in the context of two interacting populations (Arntzen et al., 2016; Chambers et al., 2022; Jiggins & Mallet, 2000). However, in nature, many of these relationships involve more complex interactions between multiple divergent forms simultaneously (Manuel Penaloza-Ramirez et al., 2010; McGinnity et al., 2015; Natola et al., 2022; Sa-Pinto et al., 2010). For example, if three divergent forms coexist, one form may function as a “bridge species” between two others, hybridizing with both and facilitating indirect allele sharing between populations that would otherwise not exchange genes (Grant & Grant, 2020; Kronforst et al., 2006). Studying reproductive isolation among three incompletely isolated sibling species of the *Anopheles gambiae* complex revealed a complex relationship between ecological divergence, phylogenetic distance and the importance of pre- and post-mating isolation that may otherwise have gone undetected (Pombi et al., 2017). Likewise, studying a sympatric trio of nine-spined stickleback (*Pungitius tymensis* and the *Pungitius pungitius*–*Pungitius sinensis* species complex) allowed for the detection of

a non-parallel genomic basis for the evolution of bony armour plate loss (A. Ishikawa et al., 2013). Thus, these complex contact zones are particularly useful for understanding how secondary contact, hybridization and introgression shape reproductive isolation in natural populations, and how simple, binary, divergence may develop towards adaptive radiation of multiple forms.

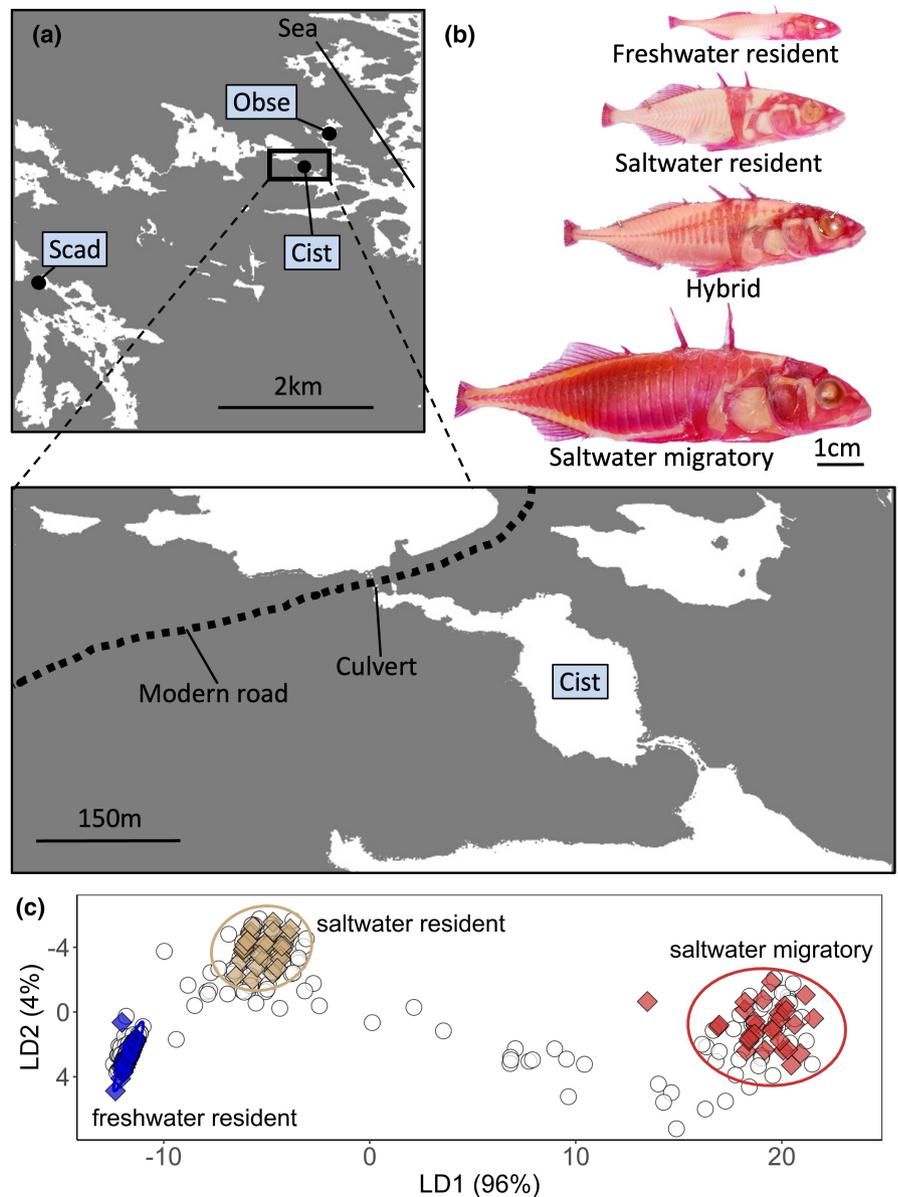
Binary contact zones between divergent forms of three-spined stickleback (*Gasterosteus aculeatus*) are common throughout the Holarctic range of the species. Species pairs of anadromous–freshwater, benthic–limnetic and lake–stream ecotypes are well described (McKinnon & Rundle, 2002). However, contact zones involving more than two forms are almost completely unstudied (but see, Ravinet et al. (2015) and Campbell (1985)). Anadromous or saltwater migratory stickleback spend the majority of their lives in open sea but migrate into freshwater streams, lakes or coastal saltwater lagoons during the spring breeding season. Contact zones between these migratory forms and freshwater resident stickleback are by far the most common (Hendry et al., 2009), with varying levels of hybridization and reproductive isolation reported across locations (Higuchi et al., 1996; Jones et al., 2006; Vines et al., 2016). In multiple locations on the Scottish island of North Uist, migratory stickleback also breed alongside an unusual saltwater resident ecotype that inhabits saline coastal lagoons year-round. In most of these locations, hybridization is rare and reproductive isolation is high (Dean et al., 2019), but in one location where a distinctive freshwater resident form co-occurs, reproductive isolation appears to break down. We investigate divergence in this putative three-way contact zone among freshwater resident, saltwater resident and saltwater migratory stickleback. We use phenotypic and genomic data to show that all three forms co-exist in the contact zone. Our data suggest that hybridization occurs, but small islands of genomic divergence, many of which encompass known chromosomal inversions, are maintained. We also show that the three forms differ in immune gene expression profiles in the contact zone, but find no evidence for transgressive expression of immune genes in hybrids.

## 2 | MATERIALS AND METHODS

### 2.1 | Sample collection

We investigated a potential three-way contact zone on the island of North Uist in the Scottish Western Isles. The contact zone was first discovered by Campbell (1985) and is a small, tidally influenced loch (Loch na Ciste, “Cist”), connected to the sea by a narrow outlet at one end and to a large freshwater drainage catchment by a culvert at the opposite end (Figure 1a). Adult stickleback were collected from the contact zone ( $n=321$ ) and nearby freshwater resident (Loch Scadavay, Scad,  $n=34$ ), saltwater resident (Ob nan Stearnain, Obse,  $n=30$ ) and saltwater migratory (Obsm,  $n=29$ ) populations where hybridization is absent or rare. We used un-baited minnow traps set overnight between April and May (the breeding season for stickleback) in 2015 and 2016. Examples of fish (fixed in

**FIGURE 1** (a) Map of sampling locations. Areas shaded in grey indicate land and white shows the locations of waterbodies. (b) Photographs of examples of each of the three forms from pure populations (Scad for freshwater resident and Obse for saltwater resident and saltwater migratory) and an example hybrid individual from the contact zone in Cist. Fish are stained to highlight areas of external bone. (c) Phenotypic linear discriminant analysis. Diamonds show phenotypic differentiation between reproductively isolated single-form populations (labelled), and circles show predictions for contact zone individuals in the same phenotypic space. Increasing LD1 was associated with more, longer lateral plates and decreasing LD2 was associated with a longer pelvis (this axis is reversed in the plot for clarity). Ellipses represent 99% confidence intervals for single-form populations, which were used to classify contact zone individuals.



formalin and stained with Alizarin Red) from each location are shown in [Figure 1b](#). Fish were euthanized using an overdose of anaesthetic (MS222), followed by destruction of the brain in accordance with UK Home Office regulations. Caudal and pectoral fins were removed and stored in 96% Proof molecular grade ethanol for preservation of genomic DNA. Spleen tissue for 92 individuals was dissected and stored in RNAlater (Life Technologies) at 4°C for 24h and then at -20°C prior to RNA extraction for gene expression analysis. Fish were then preserved in 70% ethanol.

## 2.2 | Phenotypic analyses

Fish were bleached and stained with alizarin red to highlight external bone following standard procedures (Peichel et al., 2001) and measurements of lateral plate number, length of the first and second dorsal spine, length of the pelvic spine, length of the longest lateral

plate, height and length of the pelvis and standard length were taken following Dean et al. (2019). The residuals from regressions of each continuous measurement of body armour (thus excluding lateral plate number) against standard length were used in phenotypic analyses to account for the correlation between body size and continuous body armour measurements.

To identify whether single-form populations could be distinguished phenotypically, and to identify the phenotypic traits most important in distinguishing the three forms, we used a linear discriminant analysis (LDA), which identifies the linear combination of features which best separates two or more pre-determined groups in  $n-1$  dimensional space. LDA was performed using all measured phenotypic traits (lateral plate number, size standardized first and second dorsal spine and pelvic spine lengths, size standardized height and length of the pelvis, size standardized length of the longest plate and standard length) for all fish from single-form populations using the MASS package in R version 4.2.1. All phenotypic traits were

centred and scaled prior to LDA analysis. The resulting discriminant function was used to predict where the contact zone fish fell within the same phenotypic space. Contact zone individuals were classified into one of the following groups: freshwater resident, saltwater resident, saltwater migratory or phenotypic intermediate, based on their phenotype relative to phenotypic 99% confidence ellipses of single-form populations using the SIBER package in R.

## 2.3 | Genomic analyses

For a subset of individuals from the contact zone ( $n=156$ ), DNA was extracted from fin tissue using Qiagen DNeasy blood and tissue kits following the published protocol. Nextera library preparation was performed with assembled Tn5 bound to magnetic beads as previously described (Kirch et al., 2021) and libraries were normalized and split across three pools for sequencing. Sequencing was carried out across three lanes of an Illumina HiSeq 3000 instrument with  $2 \times 150$ bp chemistry. Raw sequencing reads were trimmed using Trimgalore version 0.6.5 and visually assessed for quality using FastQC. Trimmed reads were then aligned to version 5 of the *Gasterosteus aculeatus* reference assembly (Nath et al., 2021) using BWA version 0.7.17, yielding a mean read depth of  $1.82 \pm 0.03 \times$  coverage per individual. Duplicate reads were marked and bam files were indexed using SAMtools version 1.10.2. Base quality score recalibration was calculated and applied using gatk version 4.1.5.0 using the file 'stickleback\_21genome\_snp\_v5.1.bed', a list of known SNPs in the stickleback genome (Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012), lifted over to the *Gasterosteus aculeatus* v5 genome assembly coordinates. Genotype likelihoods were estimated and genotypes were called using bcftools version 1.10.2 with the following filters: Max read depth: 10,000, minimum mapping quality: 20, Minimum base quality: 20, and using the multiallelic caller with a prior expected substitution rate of 0.000001. This resulted in a total of 460,779,110 unfiltered sites (inclusive of invariant sites), which were used for analyses which require invariant sites to be retained in the dataset. Variant quality score recalibration and filtration were applied in gatk using a training set of high ( $\sim 10 \times$ ) coverage individuals of each of the three pure forms from Scad and Obse. This produced a genome-wide SNP set of 7,340,599 SNPs for 156 individuals from the contact zone, which was further filtered for some analysis using VCFtools version 0.1.17 and Plink2 as below.

To identify genomic structure in the contact zone, principal component analysis (PCA) was conducted on all sequenced contact zone individuals ( $n=156$ ) using the adegenet package (Jombart, 2008) in R. A set of 13,639 SNPs, filtered to exclude SNPs with: a minor allele frequency  $< 0.02$  (SNP must be present in at least six individuals),  $> 5$  missing genotypes across all individuals, linkage disequilibrium  $r^2 > 0.2$  (calculated in 100 kb sliding windows using Plink version 1.9b6.21) and all SNPs on the sex and mitochondrial chromosomes, was used for PCA. We also investigated the ancestry of contact zone individuals using the same filtered dataset in ADMIXTURE version 1.3.0 (Alexander et al., 2009). ADMIXTURE was run for  $K=1-10$ , with cross-validation and bootstrap standard error estimations (100

bootstrap replicates, flags  $-CV$  and  $-B100$  respectively). The VCF file was converted to bed format for input to the ADMIXTURE software using Plink version 2.00a2.3. Individuals with  $\geq 95\%$  ancestry from each of the three clusters (which broadly corresponded to the three phenotypic forms) were classified as putatively genomically 'pure' individuals for further analysis. This gave pure sample sizes of: freshwater resident,  $n=51$ , saltwater resident,  $n=31$  and saltwater migratory,  $n=29$ . To test whether inversion polymorphisms affected the admixture classifications, we repeated the admixture analysis with all sites within known and putative chromosomal inversions removed, leaving 13,533 SNPs. In version 5 of the stickleback genome assembly, these inversion coordinates correspond to: ChrI:26,100,000–26,400,000, ChrXI:6,200,000–6,500,000, ChrXXI:10,000,000–11,400,000 identified by Jones, Chan, et al. (2012a), Jones, Grabherr, et al. (2012b) and ChrXI:5,700,000–7,900,000 identified by Liu et al. (2018). We also calculated observed heterozygosity ( $H_o$ ) for each individual using the  $--het$  flag in VCFtools, and tested for differences between admixture classified groups using a linear model, followed by a *post-hoc* Tukey's test in R.  $p$ -values from the Tukey's test were adjusted for multiple testing using the FDR method.

To identify genomic regions with distinct clustering patterns across the genome, we performed a local PCA analysis using the Lostruct version 0.0.0.9000 package in R (Li & Ralph, 2019). Lostruct divides the genome into non-overlapping windows and performs a principal component analysis (PCA) on the SNPs in each window. It then compares the PCAs derived from each window and calculates a similarity score, which is visualized using a multidimensional scaling (MDS) transformation. Lostruct was run using 20 SNP windows and separately for each chromosome, using the genome-wide SNP set, filtered to remove sites with more than 10% missing data using VCFtools, leaving a total of 442,946 SNPs. The first MDS axis was then plotted against the position of each window across each chromosome to visualize genomic regions with distinct clustering patterns. Localized regions of extreme MDS score were selected manually following visual inspection of genome-wide MDS plots. To identify whether regions of the genome showing distinct clustering patterns reflected differential patterns of similarity between genomic groups, PCA was plotted for all SNPs in each discrete genomic region with extreme MDS score using the adegenet R package version 2.1.10 (Jombart, 2008) in R. Individuals were coloured by genomic group (as defined by admixture analysis).

To identify genomic regions of differentiation between putatively 'genomically pure' forms in the contact zone (as identified by admixture analysis), pairwise relative ( $F_{ST}$ ) and absolute ( $d_{xy}$ ) divergence were calculated for each possible comparison of the three genomically pure groups in 5-kb windows.  $F_{ST}$  calculations were made in VCFtools using the genome-wide SNP set of 7,340,599 SNPs, without further filtering.  $d_{xy}$  calculations were made using the unfiltered VCF file inclusive of invariant sites (460,779,110 sites) with custom Python scripts written by Simon Martin ([https://github.com/simonmartin/genomics\\_general](https://github.com/simonmartin/genomics_general)) in Python version 3.6.

Nucleotide diversity ( $\pi$ ) was calculated for each of the three 'genomically pure' groups on a per-site basis and then averaged across sites to compare groups.  $\pi$  calculations were also made in 5-kb

windows across the genome for all contact zone individuals. All  $\pi$  calculations were made using the unfiltered VCF file, including invariant sites (460,779,110 total sites) in VCFtools. To identify potential regions affected by selection, Tajimas D was calculated across all 156 contact zone individuals, in 5-kb windows, using the genome-wide SNP set of 7,340,599 SNPs, without further filtering in VCFtools. All Manhattan plots were generated in R. All individuals with <95% ancestry from any one of the three putatively 'pure' genomic clusters were excluded from  $F_{ST}$ ,  $d_{xy}$  and per-site  $\pi$  calculations.

To investigate patterns of introgression across the contact zone, we estimated hybrid index and carried out a genomic cline analyses using a Bayesian MCMC approach, implemented in the gghybrid package in R. gghybrid uses a maximum-likelihood-based method (Buerkle, 2005) to estimate genome-wide hybrid index and then infers extreme, restricted or biased gene flow using a log-logistic, locus-specific genomic cline function (Fitzpatrick, 2013). Genomic cline analyses relate the hybrid index of an individual to the probability that each locus exhibits ancestry from one parental population, with a null model in which the probability of a locus displaying ancestry from parental population 1, is directly predicted by the hybrid index of the individual. If the probability of ancestry from parental population 1 deviates significantly from the probability predicted by the hybrid index then this will be reflected in excess values of  $u$  (cline centre) and/or  $v$  (cline steepness) parameters. These loci are inferred to exhibit patterns of introgression that diverge from neutral expectations, and as such may be under strong selection and/or be involved in maintaining reproductive isolation across the contact zone (Bailey, 2022).

As almost all of the admixed individuals in the contact zone fell in the genomic space between saltwater resident and saltwater migratory fish, we calculated hybrid indices and estimated genomic clines along this axis only. Parental reference individuals were defined as those with >95% ancestry from either saltwater resident ( $n=31$ ) or saltwater migratory ( $n=29$ ) populations, and all individuals that fell along the axis between these populations in the admixture analysis were included as hybrids ( $n=24$ ). Because loci that are more differentiated between the two parental populations should be more informative in a hybrid cline analysis, we used only SNPs with per-site  $F_{ST}>0.7$  between the parental populations (Gompert & Buerkle, 2011). Genome-wide SNPs were also filtered to exclude sites with a minor allele frequency <0.1, because alleles with low variability in the dataset likely do not produce good fits in cline analyses. Finally, we used BCFtools version 1.10.2 to retain only biallelic sites, leaving a final 2416 SNPs for hybrid index and genomic cline estimation. The MCMC chain was run for 7000 iterations with a 2000 iteration burn-in, both for hybrid index and cline estimation. The fit of the genomic cline model was confirmed via comparison of AIC with models in which cline steepness and/or centre parameters were fixed to null values following the published pipelines for this software (Bailey, 2023).  $p$ -values for genomic clines were corrected for multiple testing using the false discovery rate (FDR) method. The VCF file was converted to structure format for input to R using PGDspider version 2.1.1.5 (Lischer & Excoffier, 2012). We did not consider clines that were significantly shallower than expected in

our analysis because, given the low level of background differentiation, it is difficult to distinguish neutral gene flow from other selective forces, e.g., balancing selection/adaptive introgression/asymmetrical introgression and so in this dataset, there was little to be gained from their inclusion. To determine the distribution of alleles with significant deviations in cline steepness and/or centre in admixed individuals across the three putatively pure populations in the contact zone we calculated the allele frequency of all significant cline alleles in putatively pure (defined by admixture analysis above) freshwater resident,  $n=51$ , saltwater resident,  $n=31$  and saltwater migratory,  $n=29$  individuals using the --freq flag in VCFtools.

## 2.4 | Gene expression

Expression of eight immune genes (*tnfa*, *foxp3a*, *cmip*, *stat4*, *stat6*, *tbet*, *rorc* and *muc2*, see Table S1 for gene information and primers) and two control 'housekeeping' genes (*b2m* and *rpl13a*) was quantified for a subset of contact zone individuals ( $n=89$ ) using real-time quantitative PCR (qPCR) assays. Immune genes were chosen to represent the innate response, Th1-, Th2- and Treg-type adaptive immune responses (Robertson et al., 2016), with the aim of assessing (1) whether gene expression profiles differed between forms in the contact zone and (2) whether expression of immune genes was transgressive in putatively hybrid individuals. All primers used for amplification were previously published in Robertson et al. (2016). RNA was extracted from whole spleens using Genejet RNA purification kits (Thermo Scientific) following the manufacturer's instructions. RNA purity was assessed using a NanoDrop1000 spectrophotometer (Thermo Scientific) and residual gDNA was removed using Precision DNase (Primer Design), following the manufacturer's protocol. Approximately 1.5  $\mu$ g of template was reverse transcribed using nanoscript2 RT kit (Primerdesign) according to the manufacturer's protocol. qPCR reactions were performed in duplicate in 96-well optical PCR plates with optical seals (StarLab) using PrecisionFAST low ROX mastermix with SYBR green (Primer Design). qPCR was conducted using an ABI 7500 Fast real-time thermocycler (Applied Biosystems) with the following cycling conditions: initial incubation at 95°C for 20s, followed by 45 cycles of 95°C for 3s and 60°C for 30s.

Relative expression values were calculated according to the  $\Delta\Delta Cq$  method (Pfaffl, 2001) and respective of the amplification efficiencies of each primer pair as described in Robertson et al. (2016). Ratios for *rorc* and *muc2* were calculated assuming 100% amplification efficiency. Expression values were standardized against the geometric mean Cq of the two housekeeping genes *b2m* and *rpl13a*. Linear models, performed in R, were used to test whether the mean expression of each gene differed across genomic groups within the contact zone (saltwater migratory  $n=22$ , saltwater resident  $n=14$ , freshwater resident  $n=31$  and putatively admixed  $n=22$ , as determined by admixture analysis). For models showing a significant effect of genomic group, *post-hoc* pairwise Tukey's tests were conducted to determine which groups significantly differed from one another, and  $p$ -values were adjusted for multiple testing using the FDR method.

### 3 | RESULTS

#### 3.1 | Morphological differentiation

Freshwater resident, saltwater migratory and saltwater resident stickleback from single form, or ecotype pair, populations, respectively, in lakes near to, but isolated from, the contact zone (Figure 1a), formed three distinct groups in the phenotypic linear discriminant analysis (LDA, Figure 1c). Linear discriminant 1 (LD1) was strongly positively associated with the number of lateral plates and length of the longest lateral plate (which likely reflects differences in body depth), suggesting that, among the nine traits used in the analysis, lateral plates can be used to discriminate between saltwater migratory stickleback and the two resident forms. Linear discriminant 2 (LD2) was negatively associated with pelvis length, suggesting that this trait also distinguishes the three forms. Predictions based on the discriminant function demonstrated that the contact zone contains stickleback that phenotypically resemble all three different forms, as well as individuals with intermediate phenotypes (Figure 1c). Within the contact zone, 60 fish fell inside the 99% confidence ellipse of freshwater residents, 37 fell inside the ellipse of saltwater residents and 33 fell inside the ellipse of saltwater migratory individuals. Forty-five individuals from the contact zone fell outside all three of the phenotypic ellipses of single-form populations and thus, will be classified as phenotypic intermediates.

#### 3.2 | Genomic differentiation

Principal component analysis (PCA) of 12,033 genome-wide SNPs from 156 contact zone individuals suggested that three main genomic groupings exist in the contact zone, with some individuals that are intermediate (Figure 2a). The three groups in the genomic PCA corresponded to the three phenotypic groups identified in the LDA (Figure 1c), with all of the contact zone individuals from within each phenotypic ellipse forming a distinct cluster in genomic space (Figure 2a). Contact zone individuals classified as phenotypic intermediates in the phenotypic LDA mostly clustered with the saltwater resident genomic group, with some also falling in intermediate genomic space (Figure 2a).

Admixture analysis similarly showed that most contact zone individuals belong to one of three distinct groups (Figure 2b), which correspond to the phenotypic groups identified in the LDA. Cross-validation indicated that  $K=3$  was the optimal number of ancestral populations in the dataset (Figure S1). Most of the phenotypically intermediate contact zone individuals had predominantly saltwater resident ancestry, with only a small proportion of both freshwater resident and saltwater migratory ancestry (Figure 2b). The ternary plot (Figure 2c) shows that overall, admixture among the three forms in the contact zone is low. Individuals that did share ancestry from more than one form were restricted to the linear axes between either freshwater resident and saltwater resident forms or saltwater resident and saltwater migratory forms. No individuals

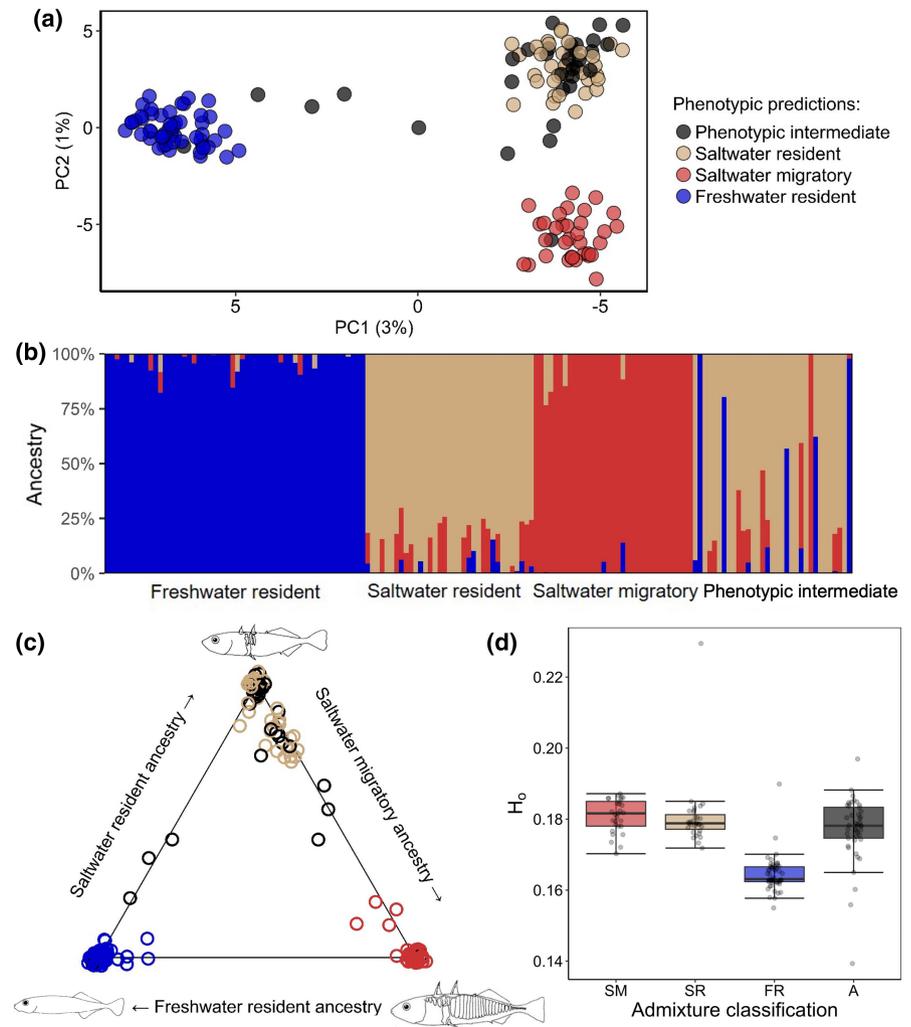
were intermediate on the freshwater resident–saltwater migratory axis. Individuals with  $\geq 95\%$  ancestry from each of the three clusters (which broadly corresponded to the three phenotypic forms) were classified as putatively genomically ‘pure’ individuals for further analysis. This gave pure freshwater resident,  $n=51$ , saltwater resident,  $n=31$  and saltwater migratory,  $n=29$ , and 45 genomically admixed individuals (31% of whom were phenotypically intermediate). The standard error for cluster assignment in admixture was low (mean  $SE=0.032\pm 0.001$ ), with slightly higher standard error for admixed than putatively pure individuals (Figure S2). A number of individuals that were phenotypically saltwater resident fell in the genomic space that would be expected to be occupied by individuals backcrossed from saltwater migratory–resident F1s to pure saltwater resident fish (Figure 2c). To test whether this bimodal pattern could be driven by inversion polymorphism in the saltwater resident population, we repeated the admixture analysis with all sites within known and putative chromosomal inversions removed. The same pattern remained with this reduced dataset (Figure S3), and thus the inversions do not contribute disproportionately to the patterns in the admixture analysis.

Mean observed heterozygosity differed among the four groups classified by admixture analysis: saltwater migratory, saltwater resident, freshwater resident and admixed ( $F=46.95$ ,  $df=3$ ,  $p<.0001$ , Figure 2d). Heterozygosity was lower in freshwater resident fish than in all other groups (*post-hoc* Tukey's test, adjusted  $p<.0001$  for all comparisons). Heterozygosity did not differ between saltwater migratory, saltwater resident and the putatively admixed fish (*post-hoc* Tukey's test, adjusted  $p>.05$  for all comparisons).

$F_{ST}$  calculations between putatively genomically ‘pure’ forms in the contact zone suggested a similar level of divergence between freshwater resident fish and the other two forms (mean  $F_{ST}$ : 0.06 in both cases), and slightly lower divergence between saltwater resident and saltwater migratory forms (mean  $F_{ST}$ : 0.03).  $d_{xy}$  calculations suggested a very similar level of divergence between all forms (mean  $d_{xy}$ : 0.14 for all pairwise comparisons), implying that the elevated mean  $F_{ST}$  in comparisons involving freshwater resident fish likely reflects a reduction in genetic diversity in the freshwater resident population, rather than a true elevation in genomic divergence. Mean nucleotide diversity ( $\pi$ ) for contact zone individuals with  $\geq 95\%$  freshwater resident ancestry ( $n=51$ ) was 0.0028 compared to 0.0030 for both individuals with  $\geq 95\%$  saltwater resident ancestry ( $n=31$ ) and individuals with  $\geq 95\%$  saltwater migratory ancestry ( $n=29$ ).

Local PCA analysis identified 11 genomic regions, across 10 chromosomes, with distinct clustering patterns relative to most of the genome (Figure 3, see Figure S4 for local PCA plots for all chromosomes). Three of these 11 regions corresponded to the locations of known or putative chromosomal inversions (Figure 3). PCAs of these inversion regions all showed three distinct groups separating along PC1, which correspond to the two homokaryotypes and the heterokaryotype (i.e. one copy of each orientation) for the inversion. For the putative inversion on chromosome 9, all genomic groups carried each possible inversion karyotype, suggesting that this inversion is not involved in reproductive isolation between forms.

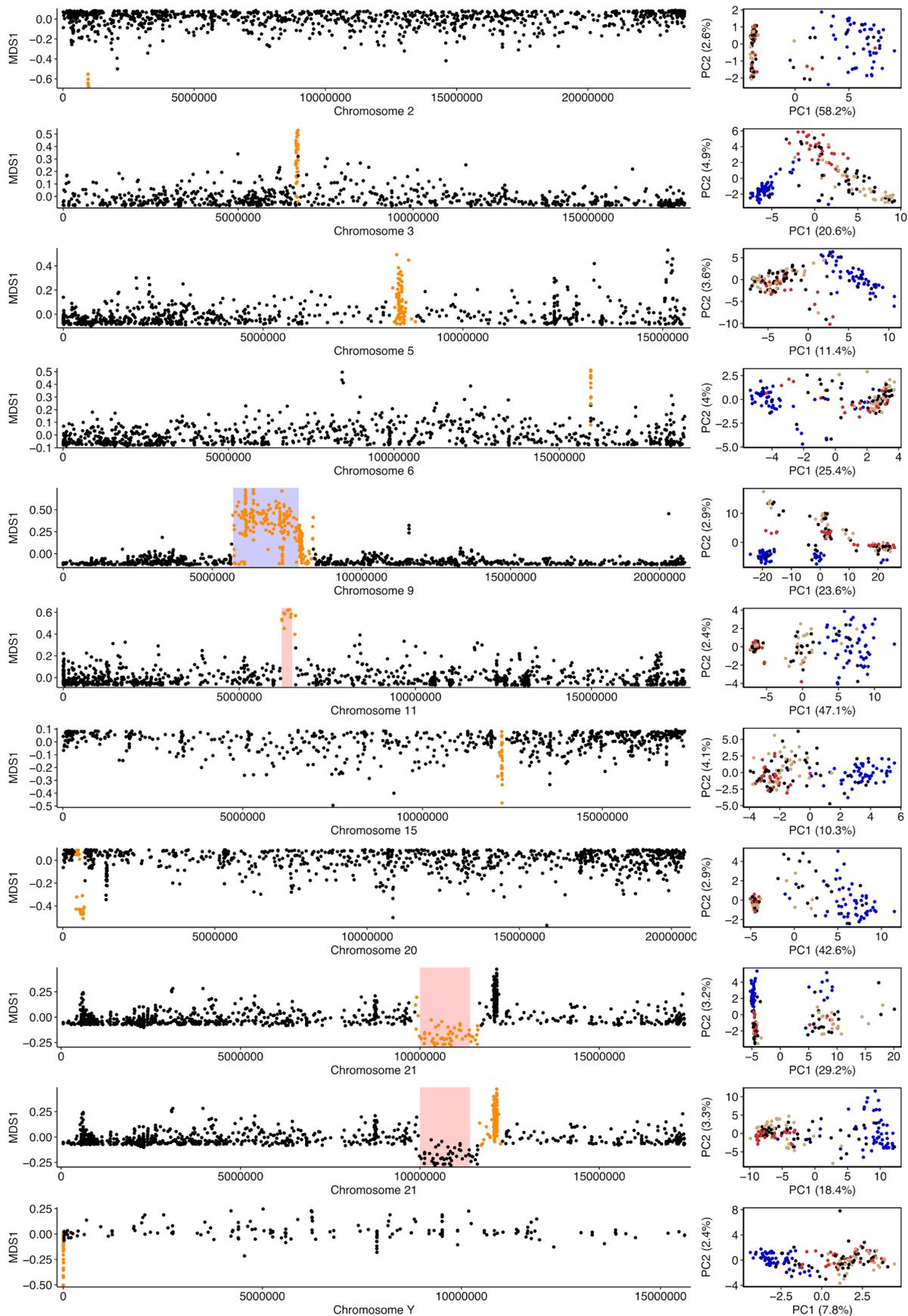
**FIGURE 2** (a) Genomic principal component analysis of 156 contact zone individuals based on 12,033 genome-wide SNPs. Individuals are coloured based on phenotype: individuals within phenotypic 99% confidence ellipses of single-form populations are coloured according to single-form populations (Figure 1c), and individuals falling outside all phenotypic 99% confidence ellipses (Figure 1c) are coloured black. PC1 – principal component 1, PC2 – principal component 2. (b) Admixture plot with  $k=3$  clusters for all sequenced contact zone individuals. Individuals are grouped according to phenotype based on the linear discriminant analysis (LDA, Figure 1c). (c) Ternary plot of admixture in contact zone individuals. Points are coloured based on phenotypic LDA classifications: blue – freshwater resident, beige – saltwater resident, red – saltwater migratory, black – phenotypic intermediate and jittered to show all individuals. (d) Observed heterozygosity ( $H_o$ ) of individuals from each admixture classification. A, admixed; FR, freshwater resident; SM, saltwater migratory; SR, saltwater resident.



The inversion on chromosome 11, which is known to be involved in freshwater adaptation (Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012), separates freshwater resident fish at one extreme, from almost all saltwater migratory fish at the opposite extreme, suggesting that these populations carry alternative homokaryotypes at this inversion. Saltwater resident fish fell mainly within the intermediate group, suggesting that they carry heterokaryotypes. Admixed individuals carried all possible karyotypes of the chromosome 11 inversion. The inversion on chromosome 21 has also been implicated in freshwater adaptation (Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012), but this inversion did not appear to distinguish genomic groups in our dataset, with most individuals carrying a single homokaryotype of the inversion, with some individuals from all groups possessing the heterokaryotype and a few saltwater resident and admixed individuals carrying the alternative homokaryotype (Figure 3). The genomic regions with extreme MDS scores on chromosome 2 (958,716 to 976,888 bp), chromosome 5 (8,243,715 to 8,842,209 bp), chromosome 6 (15,952,181 to 15,959,489 bp), chromosome 15 (12,092,146 to 12,231,595), chromosome 20 (438,004 to 700,797 bp) and chromosome Y (0 to 20,000 bp) all largely separated freshwater resident fish from the three other genomic groups. The region with extreme MDS score on chromosome 3 (6,653,310 to

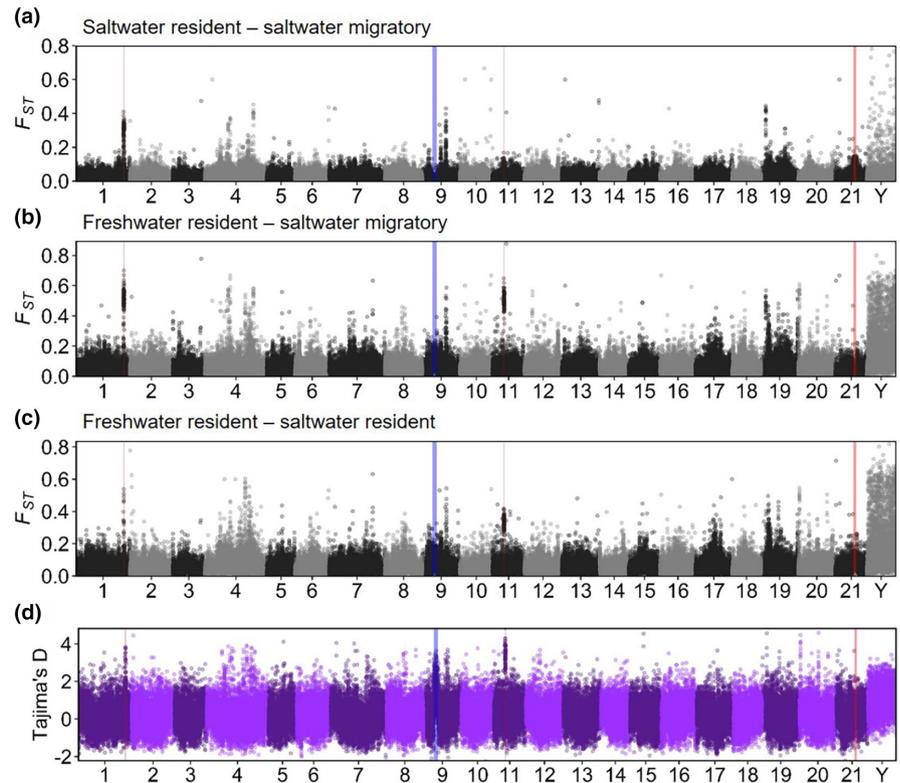
6,713,074 bp) separated freshwater resident and saltwater resident fish along PC1, with saltwater migratory fish falling intermediately, and admixed individuals spanning all three groups.

$F_{ST}$  plots identified multiple peaks of genomic differentiation in each pairwise comparison, (Figure 4a–c). Five clear peaks of divergence were evident between saltwater resident and saltwater migratory forms (Figure 4a). One was located within a known inversion on chromosome 1 (Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012), two peaks were on chromosome 4, containing regions such as the *Eda* locus, which is known to play a key role in determining lateral plate number in most stickleback (Colosimo et al., 2005) and in other populations on North Uist (Dean et al., 2019) and *Ppara* which has been associated with salt tolerance (Wang et al., 2014). There was also a peak on chromosome 9 containing the *Cttna2* gene, which has been associated with temperature adaptation in stickleback (Guo et al., 2015) and a peak on chromosome 19, the X sex determination chromosome. The freshwater resident and saltwater migratory comparison showed a similar pattern of divergence with the same five peaks as above, plus a peak of divergence within a known inversion on chromosome 11 (Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012) and a further peak on chromosome 20 (Figure 4b). Clear peaks of divergence were less evident between



**FIGURE 3** Local PCA analysis. First column shows multidimensional scaling (MDS) plots for each chromosome containing a region of interest (see Figure S4 for MDS plots of all chromosomes). Regions of interest are shown in orange, Red bars show the positions of known inversions identified by Jones, Chan, et al. (2012a), Jones, Grabherr, et al. (2012b) and the blue bar shows the position of a putative inversion identified by Liu et al. (2018). Dots represent 20 SNP windows. Second column shows principal component analyses (PCAs) of SNPs from the corresponding regions of interest (shown in orange in MDS plots), with dots in the PCA plots representing individuals. Points in PCA plots are coloured based on admixture classifications: blue – freshwater resident, beige – saltwater resident, red – saltwater migratory and black – admixed.

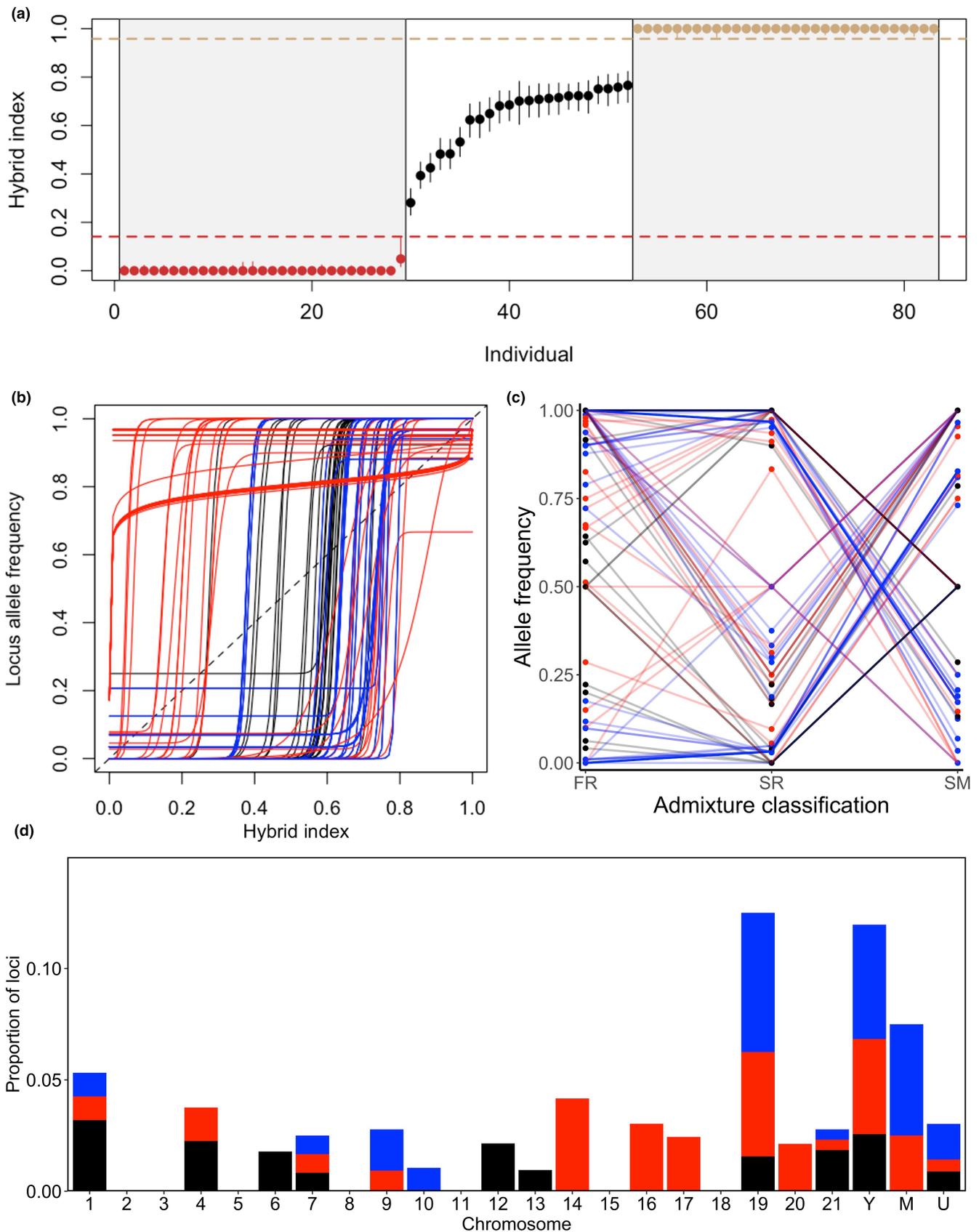
**FIGURE 4** (a–c) Relative genomic differentiation ( $F_{ST}$ ) in pairwise comparisons between putatively ‘pure’ saltwater resident, freshwater resident and saltwater migratory forms from the contact zone. (d) Genome-wide Tajima’s  $D$  for all 156 contact zone individuals. All statistics are calculated in 5-kb windows. Red bars show the positions of known inversions identified by Jones, Chan et al. (2012a), Jones, Grabherr et al. (2012b) and blue bars show the position of a putative inversion identified by Liu et al. (2018).



saltwater resident and freshwater resident forms, but smaller peaks were again present in the same locations on chromosome 4, 9, 11, 19 and 20 (Figure 4c). The Y chromosome shows high divergence in all comparisons, but particularly those involving the freshwater resident form, which is probably a result of drift, given the largely non-recombinant nature of the Y chromosome (Peichel et al., 2020).  $D_{xy}$  plots did not show peaks of divergence in the same genomic regions as  $F_{ST}$ , but instead largely showed elevated divergence only in regions with high nucleotide diversity (Figure S5). This is likely a result of the lower power to detect differences among loci using  $d_{xy}$  than  $F_{ST}$  when divergence times are short (Cruickshank & Hahn, 2014). We also identified a number of genomic regions with elevated Tajima’s  $D$  (Figure 4d), including the known inversions on chromosomes 1, 9 and 11 (but notably not on chromosome 21) and further peaks on chromosomes 4, 9 and 20 that mirror previously identified regions of elevated  $F_{ST}$  (Figure 4a–c). This is consistent with the maintenance of these polymorphisms over time in the contact zone as a result of balancing selection.

Estimations of hybrid index confirmed that all 24 individuals that were putatively admixed between saltwater resident and saltwater migratory fish were likely hybrids, with hybrid indices in these individuals ranging from 0.28 to 0.76 (Figure 5a). Hybrid indices were

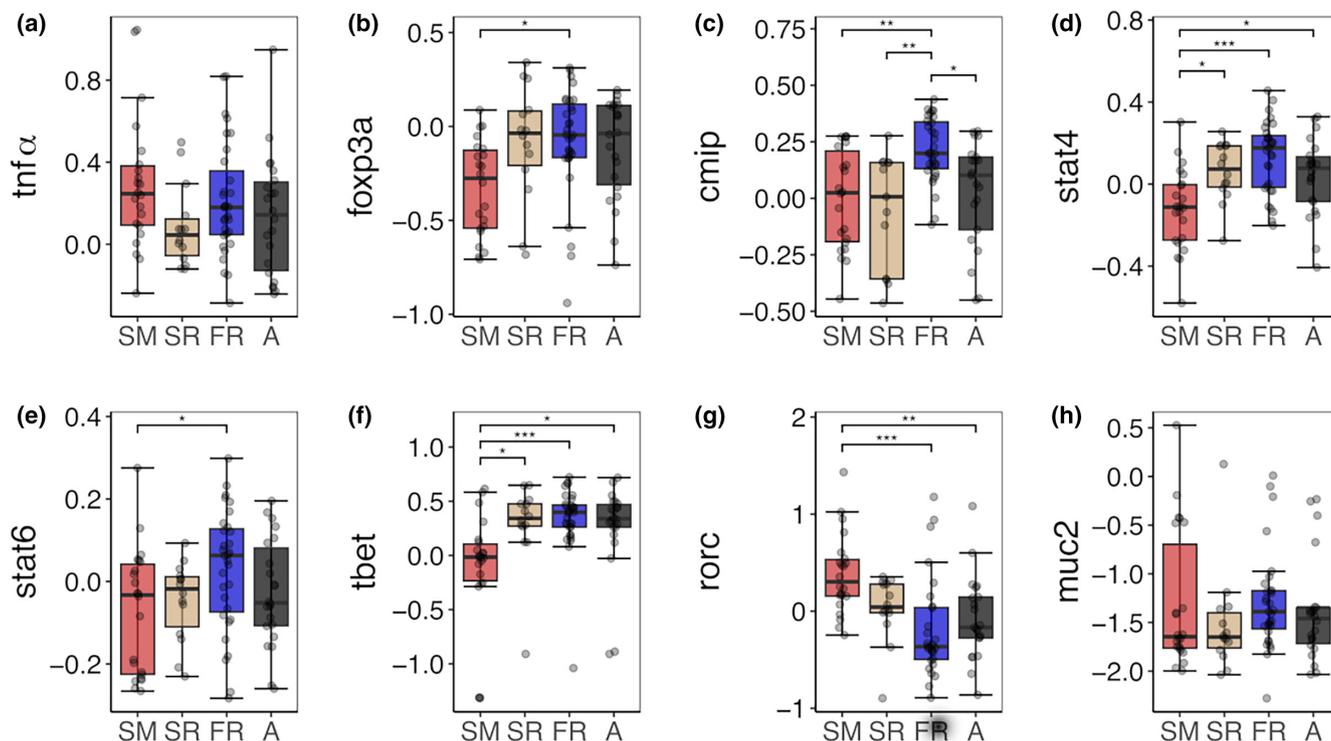
skewed towards saltwater residents, recapitulating the findings of the admixture analysis that most hybrid individuals had predominantly saltwater resident ancestry and likely reflect backcrossing of F1s to the saltwater resident population. Genomic cline analysis detected deviations from the neutral expectation in patterns of introgression across loci. Of the 2416 loci that we analysed, 48 had steeper clines than expected ( $v > 1$ , mean  $v = 24.1 \pm 7.1$ ), which is consistent with hybrid incompatibilities at these loci. Twenty nine of these loci fell within known genes: *myh7bb*, *myrip*, *jph1b*, *jagn1b*, mitochondrial genes *co1* and *nd6* and a number of unnamed genes (Table S2, Figure 5b). Eighty-six loci deviated from the null expectation for the cline centre ( $u$ ) (adjusted  $p < 0.05$  in all cases, Table S2, Figure 5b). Of these 86 loci, 33 showed preferential introgression from the saltwater migratory into the saltwater resident background ( $u > 0.5$ , mean  $u = 0.71 \pm 0.06$ , Figure 5b, Table S2). These loci fell within genes including *pparaa* and the mitochondrial *co1* and *nd6* genes (Table S2). The remaining 53 loci showed preferential introgression from the saltwater resident into the saltwater migratory background ( $u < 0.5$ , mean  $u = 0.07 \pm 0.12$ , Figure 5b, Table S2). These loci fell mostly within the remaining mitochondrial genes (Table S2). Because loci in the cline analysis were selected for high saltwater resident–saltwater migratory  $F_{ST}$ , allele frequencies of the significant cline loci were, by default, opposing in



these two populations (Figure 5c). For some cline loci, the freshwater resident fish had allele frequencies similar to saltwater resident populations, while other loci were similar in allele frequencies in freshwater

resident and saltwater migratory populations. A third subset of alleles that occurred at high frequency in both freshwater resident and saltwater migratory populations occurred at relatively low frequencies in

**FIGURE 5** (a) Hybrid indices for putatively pure saltwater migratory (red), saltwater resident (beige) and admixed (black) individuals calculated using 2416 high  $F_{ST}$  loci in the R package *gghybrid*. (b) Genomic clines for the 48 loci that had significantly steeper than expected clines ( $v$ , black lines), and that deviated from the null expectation in cline centre ( $u$ , red lines) or both (blue lines). Each solid line represents the genomic cline for a single locus. The null prediction that allele frequencies are equal to the hybrid index is shown by the dashed diagonal black line. (c) Allele frequency of loci that had significantly steeper than expected clines ( $v$ , black lines), and that deviated from the null expectation in cline centre ( $u$ , red lines), or both (blue lines) in the three putatively pure populations in the contact zone (determined by admixture analysis). FR, freshwater resident; SR, saltwater resident; SM, saltwater migratory. (d) The proportion of loci included in the cline analysis that had significantly steeper than expected clines ( $v$ , black), and that deviated from the null expectation in cline centre ( $u$ , red) and both (blue). M, Mitochondrial genome; U, unassigned scaffold; Y, Y chromosome.



**FIGURE 6** (a–h) Gene expression of eight immune genes across the three putatively ‘pure’ genomic groups, and remaining, putatively admixed individuals in the contact zone. Y axes show relative gene expression for each gene. Circles represent individuals. A, admixed; FR, freshwater resident; SM, saltwater migratory; SR, saltwater resident. Brackets and asterisks indicate significance thresholds of *post-hoc* Tukey’s tests between groups: \*Indicating  $p < .05$ , \*\*Indicating  $p < .01$  and \*\*\*Indicating  $p < .001$ .

saltwater resident fish (Figure 5c). The proportion of loci with significant deviations from the null expectation in cline steepness and/or cline centre differed across chromosomes (Figure 5d). The two sex chromosomes (Chr 19 and Chr Y) had the highest proportion of significant cline loci (Figure 5d). The locations of loci with significant shifts in cline steepness and/or centre positions relative to genome-wide  $F_{ST}$  are shown in Figure S6.

### 3.3 | Gene expression

Six of the eight immune genes that we investigated showed significant differences in expression between genomic groups within the contact zone ( $F = 2.8–8.6$ ,  $df = 3$ , adjusted  $p < 0.05$  in all cases, Figure 6). Saltwater migratory and freshwater resident groups showed differential expression for all six of these genes (*foxp3a*, *stat4*, *stat6*, *tbet*, *rorc* and *cmip*, Table 1). Saltwater migratory and saltwater resident

groups differed in expression of *stat4* and *tbet* (Table 1). Saltwater migratory and admixed groups differed in expression of *stat4*, *tbet* and *rorc* (Table 1). Saltwater resident and freshwater resident groups differed in *cmip* expression as did freshwater resident and admixed groups (Table 1). Saltwater resident and admixed individuals did not differ in expression of any of the genes we measured (Table 1), likely because most admixed individuals predominantly had saltwater resident ancestry (Figure S7). For all genes, the mean expression in admixed individuals fell between the highest and lowest mean of the three ‘pure’ genomic groups (Figure 6).

## 4 | DISCUSSION

Studies of contact zones between divergent forms of the same species are crucial for understanding how reproductive isolation either develops or collapses following hybridization and introgression.

TABLE 1 Tukey's test results of gene expression differences between forms.

Gene	Comparison	Tukey difference	Adjusted p-value
TNFa	Saltwater resident-saltwater migratory	-0.21	.1529
TNFa	Freshwater resident-saltwater migratory	-0.08	.7638
TNFa	Admixed-saltwater migratory	-0.16	.2733
TNFa	Freshwater resident-saltwater resident	0.13	.4943
TNFa	Admixed-saltwater resident	0.05	.9526
TNFa	Admixed-freshwater resident	-0.08	.7566
FOXP3a	Saltwater resident-saltwater migratory	0.24	.0614
FOXP3a	Freshwater resident-saltwater migratory	0.25	<b>.0118</b>
FOXP3a	Admixed-saltwater migratory	0.21	.0662
FOXP3a	Freshwater resident-saltwater resident	0.00	1.0000
FOXP3a	Admixed-saltwater resident	-0.03	.9878
FOXP3a	Admixed-freshwater resident	-0.03	.9712
CMIP	Saltwater resident-saltwater migratory	-0.05	.9031
CMIP	Freshwater resident-saltwater migratory	0.22	<b>.0018</b>
CMIP	Admixed-saltwater migratory	0.03	.9664
CMIP	freshwater resident-saltwater resident	0.27	<b>.0011</b>
CMIP	Admixed-saltwater resident	0.08	.7029
CMIP	Admixed-freshwater resident	-0.19	<b>.0102</b>
STAT4	Saltwater resident-saltwater migratory	0.18	<b>.0213</b>
STAT4	Freshwater resident-saltwater migratory	0.25	<b>&lt;.0001</b>
STAT4	Admixed-saltwater migratory	0.16	<b>.0263</b>
STAT4	Freshwater resident-saltwater resident	0.07	.6128
STAT4	Admixed-saltwater resident	-0.03	.9757
STAT4	Admixed-freshwater resident	-0.10	.2256
STAT6	Saltwater resident-saltwater migratory	0.03	.9450
STAT6	Freshwater resident-saltwater migratory	0.10	<b>.0444</b>
STAT6	Admixed-saltwater migratory	0.05	.6929
STAT6	Freshwater resident-saltwater resident	0.08	.3195
STAT6	Admixed-saltwater resident	0.02	.9763
STAT6	Admixed-freshwater resident	-0.06	.4526
TBET	Saltwater resident-saltwater migratory	0.41	<b>.0227</b>
TBET	Freshwater resident-saltwater migratory	0.47	<b>.0006</b>

TABLE 1 (Continued)

Gene	Comparison	Tukey difference	Adjusted p-value
TBET	Admixed-saltwater migratory	0.39	<b>.0134</b>
TBET	Freshwater resident-saltwater resident	0.05	.9777
TBET	Admixed-saltwater resident	-0.03	.9980
TBET	Admixed-freshwater resident	-0.08	.9025
RORC	Saltwater resident-saltwater migratory	-0.35	.1018
RORC	Freshwater resident-saltwater migratory	-0.56	<b>.0002</b>
RORC	Admixed-saltwater migratory	-0.45	<b>.0070</b>
RORC	Freshwater resident-saltwater resident	-0.21	.4811
RORC	Admixed-saltwater resident	-0.10	.9162
RORC	Admixed-freshwater resident	0.11	.8283
MUC2	Saltwater resident-saltwater migratory	-0.21	.6982
MUC2	Freshwater resident-saltwater migratory	0.01	.9999
MUC2	Admixed-saltwater migratory	-0.09	.9547
MUC2	Freshwater resident-saltwater resident	0.22	.6278
MUC2	Admixed-saltwater resident	0.12	.9218
MUC2	Admixed-freshwater resident	-0.10	.9275

Bold values indicate Adjusted  $p < 0.05$ .

Our investigation confirms the existence of a three-way stickleback contact zone in Loch na Ciste, Scottish Western Isles with freshwater resident, saltwater resident and saltwater migratory forms co-existing in the same location. Despite considerable morphological variation in the contact zone, genomic and gene expression data are consistent with the existence of three largely distinct genomic groups. This finding is concordant with a similar three-way contact zone in Western Ireland in which morphology alone implies extensive hybridization, but genetic analyses support the existence of three separate ecotypes that are very similar in phenotype to those described here (Ravinet et al., 2015). This suggests that reproductive isolation among these three forms of stickleback is maintained in multiple locations where they occur in sympatric conditions.

We did, however, find evidence for low levels of hybridization, particularly between saltwater resident and saltwater migratory forms in this contact zone. This is particularly interesting because where these two forms occur in the absence of freshwater resident stickleback they form strongly reproductively isolated species pairs, with almost no phenotypically or genomically intermediate individuals (Dean et al., 2019). It is possible that gene flow with the third, freshwater resident form in the contact zone interferes with the processes that prevent hybridization between the two saltwater forms when they exist in isolation, but on balance, this seems unlikely, given how little hybridization is evident between freshwater residents and either of the other two forms.

Most phenotypically intermediate individuals clustered with the saltwater resident fish in genomic analyses. This suggests that most of the individuals that are intermediate in phenotypic space (largely as a result of having an intermediate number of lateral plates) are predominantly of saltwater resident origin, and thus the direction of gene flow is largely into the saltwater resident ecotype. Differences in lateral plate number in *G. aculeatus* are largely determined by a single locus, *Eda* (Colosimo et al., 2005), and fish that are heterozygous at the *Eda* locus often exhibit a partially plated phenotype (Colosimo et al., 2005; Dean et al., 2019; Morris et al., 2019). Introgression of the *Eda* 'complete' allele from the saltwater migratory into the saltwater resident population is, thus, likely responsible for the intermediate phenotypes found in the contact zone. This pattern of unidirectional introgression from saltwater migratory fish into the saltwater resident population could reflect genuine unidirectional gene flow between these forms, e.g., if hybridization only occurs between a specific subset of individuals. Preliminary data suggest that while (small) female residents may be able to spawn in the nests of (large) migratory males, the reverse may be unlikely, because (large) migratory females find it difficult to spawn in the nests of (small) resident males (LLD, personal observations), which could provide a potential mechanism for this. However, it is also possible that gene flow is, in fact, bi-directional, but strong selection for adaptations to living in the open sea means that migratory individuals without the full suite of adaptations to survive in the ocean do not survive to return as breeding adults. If the latter is the case, this strong selection on migratory individuals could play an important role in maintaining the reproductive isolation between saltwater resident and migratory forms.

The genomic PCA and the admixture analysis suggest that freshwater resident and saltwater migratory fish do not hybridize directly in this contact zone. However, indirect gene flow between freshwater resident and saltwater migratory fish may be facilitated via the exchange of genes between both forms and the saltwater resident population. Admixture analysis also highlighted a distinct pattern of bimodality in individuals that were phenotypically like saltwater residents. This is likely a result of an imperfect association between phenotype and genotype, such that some phenotypically saltwater resident individuals are, in fact, admixed backcrosses between saltwater resident by saltwater migratory F1 hybrids and the saltwater resident population.

Our analyses suggested that there are multiple peaks or 'islands' of divergence across the genome which differentiate the three forms in the contact zone. Local PCA identified distinct clustering patterns in genomic regions containing two of the three known chromosomal inversions (on chromosomes 11 and 21), and in the putative inversion on chromosome 9. PCA of these regions showed that the inversions on chromosomes 9 and 21 did not distinguish forms, but the inversion in chromosome 11 had alternate homokaryotypes in saltwater migratory and freshwater resident fish with saltwater resident largely carrying the heterokaryotype or the saltwater migratory homokaryotype. The inversions on chromosomes 1 and 11 also showed clear peaks in  $F_{ST}$ . Chromosomal inversions often underlie ecologically important traits, suggesting that they contribute to adaptation in many species (Berg et al., 2017; Hager et al., 2022; Zhang et al., 2021). As well as

containing important adaptations, inversions can act to maintain reproductive isolation because recombination is suppressed in heterokaryotype individuals (Zhang et al., 2021). Many of the known inversions in stickleback are thought to be related to freshwater-marine divergence (Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012), although our data suggest this may only be true for the chromosome 11 inversion in this contact zone. Nonetheless, our study adds to the growing body of evidence suggesting that inversions are particularly important for adaptation and for maintaining reproductive isolation between divergent forms of stickleback in the face of gene flow. Local PCA also identified a number of outlier regions that are not within known inversions, most of which separated freshwater resident fish from the other two forms. This suggests that these regions may also be important for freshwater adaptation, and likely contribute to the reproductive isolation of freshwater resident fish from the other forms in the contact zone.

Our genomic cline analysis identified multiple loci with significant shifts in cline steepness and centres. Steeper clines are indicative of genomic regions that are under strong selection and/or contain hybrid incompatibilities (Bailey, 2022). Some regions with steeper than expected clines were also within chromosomal inversions, e.g., a region of the chromosome 1 inversion containing the *jagn1b* gene, which is associated with immune function in zebrafish (Doll et al., 2020), and thus could allow different immune responses for the different pathogens that are likely encountered by resident versus migratory stickleback in the contact zone. We also identified loci with significant shifts in cline centres, just over one third of which were shifted towards the saltwater resident fish, indicating preferential introgression of the saltwater migratory allele into the saltwater resident genomic background. One of these loci was located within the *pparaa* gene, which plays a key role in promoting uptake, utilization and catabolism of fatty acids (Desvergne & Wahli, 1999) and is also associated with salt tolerance (Wang et al., 2014). Fatty acid synthesis is a crucial process necessary for the colonization of fatty acid poor freshwater environments (Asano Ishikawa et al., 2019), and so this gene may be particularly important in a contact zone spanning freshwater to seawater. The rest of the cline centre shifts were mostly in the mitochondria and showed preferential introgression of the saltwater resident alleles into the saltwater migratory genomic background. Two major mitochondrial lineages occur in stickleback on North Uist, the trans-Atlantic and European lineages, which likely originated in separate glacial refugia on either side of the Atlantic during the last ice age (Dean et al., 2019; Makinen & Merila, 2008). Previous work has shown that freshwater resident and saltwater resident populations on North Uist are almost entirely of the European mitochondrial lineage, whereas saltwater migratory fish carry approximately an even mix of both mitochondrial lineages (Dean et al., 2019). The preferential introgression of saltwater resident mitochondrial alleles into the saltwater migratory genomic background in our genomic cline analysis has at least two possible explanations. Admixed individuals in the contact zone may benefit from carrying saltwater resident, and therefore probably the European mitochondrial haplotype. Alternatively, admixed individuals may be more likely to have saltwater resident mothers than saltwater migratory mothers since saltwater migratory

females are much larger than resident fish and may struggle to enter the nests and spawn successfully with saltwater resident males (LLD personal observations).

To investigate the wider context and potential origin of the loci that showed significant deviations in cline steepness or centres in admixed individuals, we plotted allele frequencies for these loci in each of the putative genomically pure forms within the contact zone. Some loci that were at high frequency in saltwater residents also occur at high frequencies in freshwater resident fish and vice versa for loci at low frequency. These loci potentially have a shared origin for both resident ecotypes. However, a number of loci at high and low frequencies in saltwater residents have opposing low or high frequencies in both other forms, suggesting that these alleles are unique or mostly unique to the saltwater resident form.

Gene expression profiles also have the potential to reveal areas of hybrid dysfunction, where the hybrid genome results in expression of some genes that are extreme relative to both parental populations. It has been suggested that this phenomenon should be especially common in immune genes, leading to autoimmune phenotypes, a pattern that is relatively common in plants (Bomblies et al., 2007; Maheshwari & Barbash, 2011). However, we did not find any evidence that gene expression in admixed individuals was extreme relative to the three putatively pure parental populations. This suggests that, at least in the handful of immune genes that we studied, hybrid dysfunction as a result of autoimmunity is not likely to contribute to reduced hybrid fitness and reproductive isolation in this contact zone (El Nagar & MacColl, 2016).

Saltwater resident fish were more similar to freshwater residents in expression of some genes (*foxp3a*, *stat4* and *tbet*) and to saltwater migratory fish for others (*cmip*, *stat6* and *muc2*), suggesting that differences between resident and migratory lifestyles are important drivers of differentiation between forms, as well as adaptation to salt- or freshwater. Given the major differences in habitat (and likely selective pressures) between marine and lacustrine/lagoonal environments this seems highly plausible. Much of the current research into contact zones in stickleback has focused on freshwater-marine divergence (DeFaveri et al., 2011; Jones, Chan et al., 2012; Jones, Grabherr, et al., 2012; McCairns & Bernatchez, 2010; Roberts Kingman et al., 2021; Terekhanova et al., 2014; Vines et al., 2016), but migratory vs. resident differences also merit further investigation. Saltwater resident populations of stickleback are rare, relative to freshwater resident and saltwater migratory forms (Campbell, 1985; Dean et al., 2019; Ravinet et al., 2015), but the potential for genomic comparisons among all three forms will likely be the key to separating adaptations to salinity from those relating to migration. This is a crucial next step if we are to fully understand the processes driving adaptive divergence and beginnings of adaptive radiation in the *G. aculeatus* species complex.

Taken together, our results suggest that reproductive isolation is maintained in a rare three-way contact zone, despite some hybridization and introgression among the three forms. We propose that a combination of strong selection for adaptation to different environments, coupled with reduced recombination in key adaptive

genes as a result of their being located on chromosomal inversions, is responsible for maintaining reproductive isolation in this contact zone. Our investigation highlights the importance of studying the complex multi-species interactions that occur in natural populations to fully understand how secondary contact, introgression and selection affect divergence and speciation.

#### AUTHOR CONTRIBUTIONS

L.L.D., A.D.C.M. and F.C.J. conceived the study. L.L.D. wrote the manuscript, collected samples and conducted phenotypic and genomic data analysis. J.R.W. collected samples and conducted phenotypic analysis. F.C.J. carried out genome sequencing. A.D.C.M. supervised all aspects of the study. All authors contributed to editing the final draft of the manuscript.

#### ACKNOWLEDGEMENTS

The authors would like to thank Melanie Kirsch and Kavi Venkataramani for conducting genomic library preparations. This work was funded by NERC grant (NE/R00935X/1), a Leverhulme Travelling Scholarship to ADCM and a German Research Foundation (DFG, SPP1819) grant to FCJ. Analyses of genomic data were conducted using the University of Nottingham's high-performance computer service, Augusta.

#### CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare for this article.

#### OPEN RESEARCH BADGES



This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at doi: [10.17639/nott.7333](https://doi.org/10.17639/nott.7333).

#### DATA AVAILABILITY STATEMENT

Raw phenotypic data, Gene expression data and a VCF file of called genomic variants are available from the University of Nottingham's digital data repository under the doi: [10.17639/nott.7333](https://doi.org/10.17639/nott.7333).

#### BENEFIT SHARING STATEMENT

Benefits generated: A research collaboration was developed between the University of Nottingham (UK) and the Max Planck Institute (Germany), forming institutional and professional relationships and generating training opportunities in library preparation and genomic data processing. Benefits from this research also accrue from the sharing of our data and results on public databases as described above.

#### ORCID

Laura L. Dean <https://orcid.org/0000-0002-4711-0931>

James R. Whiting <https://orcid.org/0000-0001-8936-4991>

Felicity C. Jones <https://orcid.org/0000-0002-5027-1031>

Andrew D. C. MacColl <https://orcid.org/0000-0003-2102-6130>

## REFERENCES

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J. W., 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>
- Abbott, R. J. (2017). Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *Journal of Systematics and Evolution*, 55(4), 238–258. <https://doi.org/10.1111/jse.12267>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Arntzen, J. W., Trujillo, T., Butot, R., Vrieling, K., Schaap, O., Gutierrez-Rodriguez, J., & Martinez-Solano, I. (2016). Concordant morphological and molecular clines in a contact zone of the common and Spined toad (*Bufo bufo* and *B. spinosus*) in the northwest of France. *Frontiers in Zoology*, 13, 52. <https://doi.org/10.1186/s12983-016-0184-7>
- Bailey, R. (2023). Bayesian hybrid index and genomic cline estimation with the R package gghybrid. *Molecular Ecology Resources*, 0(0), 1–15. <https://doi.org/10.1111/1755-0998.13910>
- Bailey, R. I. (2022). *gghybrid R package for Bayesian hybrid index and genomic cline estimation*. Zenodo. <https://doi.org/10.5281/zenodo.3676498>
- Berg, P. R., Star, B., Pampoulie, C., Bradbury, I. R., Bentzen, P., Hutchings, J. A., Jentoft, S., & 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>
- Bombliès, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L., & Weigel, D. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biology*, 5(9), 1962–1972. <https://doi.org/10.1371/journal.pbio.0050236>
- Buerkle, C. A. (2005). Maximum-likelihood estimation of a hybrid index based on molecular markers. *Molecular Ecology Notes*, 5(3), 684–687. <https://doi.org/10.1111/j.1471-8286.2005.01011.x>
- Cahan, S. H., Goodman, P., & Grauer, J. A. (2023). Ecological and genetic distinctiveness of socially hybridogenetic lineages of Pogonomyrmex harvester ants at regional and local scales. *Evolutionary Ecology*, 37, 645–667. <https://doi.org/10.1007/s10682-023-10241-9>
- Campbell, R. N. (1985). Morphological variation in the three-spined stickleback (*Gasterosteus aculeatus*) in Scotland. *Behaviour*, 93(1), 161–168.
- Chae, E., Bombliès, K., Kim, S. T., Karelina, D., Zaidem, M., Ossowski, S., Martin-Pizarro, C., Laitinen, R. A. E., Rowan, B. A., Tenenboim, H., Lechner, S., Demar, M., Habring-Muller, A., Lanz, C., Ratsch, G., & Weigel, D. (2014). Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. *Cell*, 159(6), 1341–1351. <https://doi.org/10.1016/j.cell.2014.10.049>
- Chambers, E. A., Marshall, T. L., & Hillis, D. M. (2022). The importance of contact zones for distinguishing interspecific from intraspecific geographic variation. *Systematic Biology*, 72, 357–371. <https://doi.org/10.1093/sysbio/syac056>
- Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D., & Kingsley, D. M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, 307(5717), 1928–1933. <https://doi.org/10.1126/science.1107239>
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates.
- 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>
- Dean, L. L., Magalhaes, I. S., Foote, A., D'Agostino, D., McGowan, S., & MacColl, A. D. C. (2019). Admixture between ancient lineages, selection, and the formation of sympatric stickleback species-pairs. *Molecular Biology and Evolution*, 36, 2481–2497. <https://doi.org/10.1093/molbev/msz161>
- DeFaveri, J., Shikano, T., Shimada, Y., Goto, A., & Merila, J. (2011). Global analysis of genes involved in freshwater adaptation in threespine stickleback (*Gasterosteus aculeatus*). *Evolution*, 65(6), 1800–1807. <https://doi.org/10.1111/j.1558-5646.2011.01247.x>
- Desvergne, B., & Wahli, W. (1999). Peroxisome proliferator-activated receptors: Nuclear control of metabolism. *Endocrine Reviews*, 20(5), 649–688. <https://doi.org/10.1210/er.20.5.649>
- Doll, L., Aghaallaei, N., Skokowa, J., Welte, K., & Bajoghli, B. (2020). A zebrafish model for severe congenital neutropenia with Jagn1b-deficiency. *Blood*, 136, 14. <https://doi.org/10.1182/blood-2020-141886>
- El Nagar, A., & MacColl, A. D. C. (2016). Parasites contribute to ecologically dependent postmating isolation in the adaptive radiation of three-spined stickleback. *Proceedings of the Royal Society B: Biological Sciences*, 283(1836), 20160691. <https://doi.org/10.1098/rspb.2016.0691>
- Fitzpatrick, B. M. (2013). Alternative forms for genomic clines. *Ecology and Evolution*, 3(7), 1951–1966. <https://doi.org/10.1002/ece3.609>
- Freh, M., Gao, J. L., Petersen, M., & Panstruga, R. (2022). Plant autoimmunity-fresh insights into an old phenomenon. *Plant Physiology*, 188(3), 1419–1434. <https://doi.org/10.1093/plphys/kiab590>
- Gaczorek, T. S., Chechetkin, M., Dudek, K., Caeiro-Dias, G., Crochet, P.-A., Geniez, P., Pinho, C., & Babik, W. (2023). Widespread introgression of MHC genes in Iberian Podarcis lizards. *Molecular Ecology*, 32, 4003–4017. <https://doi.org/10.1111/mec.16974>
- Gompert, Z., & Buerkle, C. A. (2011). Bayesian estimation of genomic clines. *Molecular Ecology*, 20(10), 2111–2127. <https://doi.org/10.1111/j.1365-294X.2011.05074.x>
- Grant, P. R., & Grant, B. R. (2020). Triad hybridization via a conduit species. *Proceedings of the National Academy of Sciences of the United States of America*, 117(14), 7888–7896. <https://doi.org/10.1073/pnas.2000388117>
- Guo, B., DeFaveri, J., Sotelo, G., Nair, A., & Merila, J. (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biology*, 13, 19. <https://doi.org/10.1186/s12915-015-0130-8>
- Hager, E. R., Harringmeyer, O. S., Wooldridge, T. B., Theingi, S., Gable, J. T., McFadden, S., Neugeboren, B., Turner, K. M., Jensen, J. D., & Hoekstra, H. E. (2022). A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science*, 377(6604), 399–405. <https://doi.org/10.1126/science.abg0718>
- Hendry, A. P., Bolnick, D. I., Berner, D., & Peichel, C. L. (2009). Along the speciation continuum in sticklebacks. *Journal of Fish Biology*, 75(8), 2000–2036. <https://doi.org/10.1111/j.1095-8649.2009.02419.x>
- Higuchi, M., Goto, A., & Yamazaki, F. (1996). Genetic structure of threespine stickleback, *Gasterosteus aculeatus*, in Lake Harutori, Japan, with reference to coexisting anadromous and freshwater forms. *Ichthyological Research*, 43(4), 349–358. <https://doi.org/10.1007/bf02347634>
- Hoarau, G., Coyer, J. A., Giesbers, M. C. W. G., Jueterbock, A., & Olsen, J. L. (2015). Pre-zygotic isolation in the macroalgal genus *Fucus* from four contact zones spanning 100–10 000 years: A tale of reinforcement? *Royal Society Open Science*, 2(140538), 140538. <https://doi.org/10.1098/rsos.140538>
- Ishikawa, A., Kabeya, N., Ikeya, K., Kakioka, R., Cech, J. N., Osada, N., Leal, M. C., Inoue, J., Kume, M., Toyoda, A., Tezuka, A., Nagano,

- A. J., Yamasaki, Y. Y., Suzuki, Y., Kokita, T., Takahashi, H., Lucek, K., Marques, D., Takehana, Y., ... Kitano, J. (2019). A key metabolic gene for recurrent freshwater colonization and radiation in fishes. *Science*, 364(6443), 886–+. <https://doi.org/10.1126/science.aau5656>
- Ishikawa, A., Takeuchi, N., Kusakabe, M., Kume, M., Mori, S., Takahashi, H., & Kitano, J. (2013). Speciation in ninespine stickleback: Reproductive isolation and phenotypic divergence among cryptic species of Japanese ninespine stickleback. *Journal of Evolutionary Biology*, 26(7), 1417–1430. <https://doi.org/10.1111/jeb.12146>
- Jiggins, C. D., & Mallet, J. (2000). Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, 15(6), 250–255. [https://doi.org/10.1016/s0169-5347\(00\)01873-5](https://doi.org/10.1016/s0169-5347(00)01873-5)
- Johannesson, K., Le Moan, A., Perini, S., & Andre, C. (2020). A Darwinian laboratory of multiple contact zones. *Trends in Ecology & Evolution*, 35(11), 1021–1036. <https://doi.org/10.1016/j.tree.2020.07.015>
- Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jones, F. C., Brown, C., Pemberton, J. M., & Braithwaite, V. A. (2006). Reproductive isolation in a threespine stickleback hybrid zone. *Journal of Evolutionary Biology*, 19(5), 1531–1544. <https://doi.org/10.1111/j.1420-9101.2006.01122.x>
- Jones, F. C., Chan, Y. F., Schmutz, J., Grimwood, J., Brady, S. D., Southwick, A. M., Absher, D. M., Myers, R. M., Reimchen, T. E., Deagle, B. E., Schluter, D., & Kingsley, D. M. (2012a). A genome-wide SNP genotyping Array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current Biology*, 22(1), 83–90. <https://doi.org/10.1016/j.cub.2011.11.045>
- Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., ... Broad Inst Genome Sequencing, P., & Whole Genome Assembly, T. (2012b). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484(7392), 55–61. <https://doi.org/10.1038/nature10944>
- Keller, I., & Seehausen, O. (2012). Thermal adaptation and ecological speciation. *Molecular Ecology*, 21(4), 782–799. <https://doi.org/10.1111/j.1365-294X.2011.05397.x>
- Kirch, M., Romundset, A., Gilbert, M. T. P., Jones, F. C., & Foote, A. D. (2021). Ancient and modern stickleback genomes reveal the demographic constraints on adaptation. *Current Biology*, 31(9), 2027–+. <https://doi.org/10.1016/j.cub.2021.02.027>
- Kronforst, M. R., Young, L. G., Blume, L. M., & Gilbert, L. E. (2006). Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. *Evolution*, 60(6), 1254–1268. <https://doi.org/10.1554/06-005.1>
- Li, H., & Ralph, P. (2019). Local PCA shows how the effect of population structure differs along the genome. *Genetics*, 211(1), 289–304. <https://doi.org/10.1534/genetics.118.301747>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- Liu, S., Ferchaud, A.-L., Gronkjaer, P., Nygaard, R., & Hansen, M. M. (2018). Genomic parallelism and lack thereof in contrasting systems of three-spined sticklebacks. *Molecular Ecology*, 27(23), 4725–4743. <https://doi.org/10.1111/mec.14782>
- Maheshwari, S., & Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annual Review of Genetics*, 45, 331–355. <https://doi.org/10.1146/annurev-genet-110410-132514>
- Makinen, H. S., & Merila, J. (2008). Mitochondrial DNA phylogeography of the three-spined stickleback (*Gasterosteus aculeatus*) in Europe – evidence for multiple glacial refugia. *Molecular Phylogenetics and Evolution*, 46(1), 167–182. <https://doi.org/10.1016/j.ympev.2007.06.011>
- Manuel Penaloza-Ramirez, J., Gonzalez-Rodriguez, A., Mendoza-Cuenca, L., Caron, H., Kremer, A., & Oyama, K. (2010). Interspecific gene flow in a multispecies oak hybrid zone in the sierra Tarahumara of Mexico. *Annals of Botany*, 105(3), 389–399. <https://doi.org/10.1093/aob/mcp301>
- McCairns, R. J. S., & Bernatchez, L. (2010). Adaptive divergence between freshwater and marine sticklebacks: Insights into the role of phenotypic plasticity from an integrated analysis of candidate gene expression. *Evolution*, 64(4), 1029–1047. <https://doi.org/10.1111/j.1558-5646.2009.00886.x>
- McGinnity, P., Cross, T. F., Hynes, R., Poole, R., Prodohl, P. A., Ravinet, M., & Harrod, C. (2015). *Where the Lake Meets the Sea: Strong Reproductive Isolation Is Associated with Adaptive Divergence between Lake Resident and Anadromous Three-Spined Sticklebacks*. Retrieved from: <Go to ISI>://DRCI:DATA2016064008433000. [http://figshare.com/articles/\\_Where\\_the\\_Lake\\_Meets\\_the\\_Sea\\_Strong\\_Reproductive\\_Isolation\\_Is\\_Associated\\_with\\_Adaptive\\_Divergence\\_between\\_Lake\\_Resident\\_and\\_Anadromous\\_Three\\_Spined\\_Sticklebacks\\_/1379626](http://figshare.com/articles/_Where_the_Lake_Meets_the_Sea_Strong_Reproductive_Isolation_Is_Associated_with_Adaptive_Divergence_between_Lake_Resident_and_Anadromous_Three_Spined_Sticklebacks_/1379626)
- McKinnon, J. S., & Rundle, H. D. (2002). Speciation in nature: The threespine stickleback model systems. *Trends in Ecology & Evolution*, 17(10), 480–488. [https://doi.org/10.1016/s0169-5347\(02\)02579-x](https://doi.org/10.1016/s0169-5347(02)02579-x)
- Morris, M. R. J., Kaufman, R., & Rogers, S. M. (2019). Heterozygosity and asymmetry: Ectodysplasin as a form of genetic stress in marine threespine stickleback. *Evolution*, 73(2), 378–389. <https://doi.org/10.1111/evo.13678>
- Nath, S., Shaw, D. E., & White, M. A. (2021). Improved contiguity of the threespine stickleback genome using long-read sequencing. *G3: Genes, Genomes, Genetics*, 11(2), jkab007. <https://doi.org/10.1093/g3journal/jkab007>
- Natola, L., Seneviratne, S. S., & Irwin, D. (2022). Population genomics of an emergent tri-species hybrid zone. *Molecular Ecology*, 31(20), 5356–5367. <https://doi.org/10.1111/mec.16650>
- O'Connell, K. A., Prates, I., Scheinberg, L. A., Mulder, K. P., & Bell, R. C. (2021). Speciation and secondary contact in a fossorial Island endemic, the Sao Tome caecilian. *Molecular Ecology*, 30(12), 2859–2871. <https://doi.org/10.1111/mec.15928>
- Peichel, C. L., McCann, S. R., Ross, J. A., Naftaly, A. F. S., Urton, J. R., Cech, J. N., Grimwood, J., Schmutz, J., Myers, R. M., Kingsley, D. M., & White, M. A. (2020). Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biology*, 21(1), 177. <https://doi.org/10.1186/s13059-020-02097-x>
- Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L. E., Colosimo, P. F., Buerkle, C. A., Schluter, D., & Kingsley, D. M. (2001). The genetic architecture of divergence between threespine stickleback species. *Nature*, 414(6866), 901–905. <https://doi.org/10.1038/414901a>
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29(9), 45e–445e. <https://doi.org/10.1093/nar/29.9.e45>
- Pombi, M., Kengne, P., Gimonneau, G., Tene-Fossog, B., Ayala, D., Kamdem, C., Santolamazza, F., Guelbeogo, W. M., Sagnon, N. F., Petrarca, V., Fontenille, D., Besansky, N. J., Antonio-Nkondjio, C., Dabire, R. K., Della Torre, A., Simard, F., & Costantini, C. (2017). Dissecting functional components of reproductive isolation among closely related sympatric species of the *Anopheles gambiae* complex. *Evolutionary Applications*, 10(10), 1102–1120. <https://doi.org/10.1111/eva.12517>
- Presgraves, D. C. (2010). The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, 11(3), 175–180. <https://doi.org/10.1038/nrg2718>
- Ranz, J. M., Namgyal, K., Gibson, G., & Hartl, D. L. (2004). Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaster* and *Drosophila simulans*. *Genome Research*, 14(3), 373–379. <https://doi.org/10.1101/gr.2019804>

- Ravinet, M., Hynes, R., Poole, R., Cross, T. F., McGinnity, P., Harrod, C., & Prodoehl, P. A. (2015). Where the lake meets the sea: Strong reproductive isolation is associated with adaptive divergence between lake resident and anadromous three-spined sticklebacks. *PLoS One*, 10(4), 0122825. <https://doi.org/10.1371/journal.pone.0122825>
- Roberts Kingman, G. A., Vyas, D. N., Jones, F. C., Brady, S. D., Chen, H. I., Reid, K., Milhaven, M., Bertino, T. S., Aguirre, W. E., Heins, D. C., von Hippel, F. A., Park, P. J., Kirch, M., Absher, D. M., Myers, R. M., Di Palma, F., Bell, M. A., Kingsley, D. M., & Veeramah, K. R. (2021). Predicting future from past: The genomic basis of recurrent and rapid stickleback evolution. *Science Advances*, 7(25), eabg5285. <https://doi.org/10.1126/sciadv.abg5285>
- Robertson, S., Bradley, J. E., & MacColl, A. D. C. (2016). Measuring the immune system of the three-spined stickleback – investigating natural variation by quantifying immune expression in the laboratory and the wild. *Molecular Ecology Resources*, 16(3), 701–713. <https://doi.org/10.1111/1755-0998.12497>
- Sa-Pinto, A., Baird, S. J. E., Pinho, C., Alexandrino, P., & Branco, M. (2010). A three-way contact zone between forms of *Patella rustica* (Mollusca: Patellidae) in the central Mediterranean Sea. *Biological Journal of the Linnean Society*, 100(1), 154–169.
- Servedio, M. R., & Noor, M. A. F. (2003). The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology, Evolution, and Systematics*, 34, 339–364. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132412>
- Taylor, E. B., Boughman, J. W., Groenenboom, M., Sniatynski, M., Schluter, D., & Gow, J. L. (2006). Speciation in reverse: Morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology*, 15(2), 343–355. <https://doi.org/10.1111/j.1365-294X.2005.02794.x>
- Terekhanova, N. V., Logacheva, M. D., Penin, A. A., Neretina, T. V., Barmintseva, A. E., Bazykin, G. A., Kondrashov, A. S., & Mugue, N. S. (2014). Fast evolution from precast bricks: Genomics of Young freshwater populations of Threespine stickleback *Gasterosteus aculeatus*. *PLoS Genetics*, 10(10), e1004696. <https://doi.org/10.1371/journal.pgen.1004696>
- Vines, T. H., Dalziel, A. C., Albert, A. Y. K., Veen, T., Schulte, P. M., & Schluter, D. (2016). Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution; International Journal of Organic Evolution*, 70(5), 1023–1038. <https://doi.org/10.1111/evo.12917>
- Wang, G., Yang, E., Smith, K. J., Zeng, Y., Ji, G., Connon, R., Fanguie, N. A., & Cai, J. J. (2014). Gene expression responses of threespine stickleback to salinity: Implications for salt-sensitive hypertension. *Frontiers in Genetics*, 5(312), 1–10. <https://doi.org/10.3389/fgene.2014.00312>
- Wu, L., Dang, J., Tang, L., Cheng, Y., Song, G., Sun, Y., Martens, J., Paeckert, M., Alstroem, P., Zhang, D., Jia, C., & Lei, F. (2023). Limited Song mixing without genomic gene flow in a contact zone between two songbird species. *Molecular Biology and Evolution*, 40(3), msad053. <https://doi.org/10.1093/molbev/msad053>
- Xiong, T., & Mallet, J. (2022). On the impermanence of species: The collapse of genetic incompatibilities in hybridizing populations. *Evolution*, 76, 2498–2512. <https://doi.org/10.1111/evo.14626>
- Zhang, L. Y., Reifova, R., Halenkova, Z., & Gompert, Z. (2021). How important are structural variants for speciation? *Genes*, 12(7), 1084. <https://doi.org/10.3390/genes12071084>

### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Dean, L. L., Whiting, J. R., Jones, F. C., & MacColl, A. D. C. (2024). Reproductive isolation in a three-way contact zone. *Molecular Ecology*, 00, e17275. <https://doi.org/10.1111/mec.17275>