Population structure associated with bioregion and seasonal prey distribution for Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in South Africa

O. Alejandra Vargas-Fonseca1,2 | Paige Yates3 | Stephan P. Kirkman4,2 | Pierre A. Pistorius1,2 | Daniel M. Moore3 | Ada Natoli5,6 | Victor Cockcroft1 | A. Rus Hoelzel3

1Department of Zoology, Nelson Mandela University, Port Elizabeth, South Africa
2Marine Apex Predator Research Unit (MAPRU), Institute for Coastal and Marine Research, Nelson Mandela University, Port Elizabeth, South Africa
3Department of Biosciences, University of Durham, Durham, UK
4Branch: Oceans and Coasts, Department of Environment, Forestry and Fisheries (DEFF), Cape Town, South Africa
5Department of Life and Environmental Sciences, College of Natural and Health Sciences, Zayed University, United Arab Emirates
6UAE Dolphin Project, Dubai, United Arab Emirates

**Correspondence**
A. Rus Hoelzel, Department of Biosciences, University of Durham, Durham, UK.
Email: a.r.hoelzel@durham.ac.uk

**Present address**
O. Alejandra Vargas-Fonseca, Department of Oceanography, Institute for Coastal and Marine Research, Nelson Mandela University, Port Elizabeth, South Africa

**Abstract**

Many marine species exhibit fine-scale population structure despite high mobility and a lack of physical barriers to dispersal, but the evolutionary drivers of differentiation in these systems are generally poorly understood. Here we investigate the potential role of habitat transitions and seasonal prey distributions on the evolution of population structure in the Indo-Pacific bottlenose dolphin, *Tursiops aduncus*, off South Africa’s coast, using double-digest restriction-site associated DNA sequencing. Population structure was identified between the eastern and southern coasts and correlated with the habitat transition between the temperate Agulhas (southern) and subtropical Natal (eastern) Bioregions, suggesting differentiation driven by resource specializations. Differentiation along the Natal coast was comparatively weak, but was evident in some analyses and varied depending on whether the samples were collected during or outside the seasonal sardine (*Sardinops sagax*) run. This local abundance of prey could influence the ranging patterns and apparent genetic structure of *T. aduncus*. These findings have significant and transferable management implications, most importantly in terms of differentiating populations inhabiting distinct bioregions and seasonal structural patterns within a region associated with the movement of prey resources.

**KEYWORDS**
Agulhas Bioregion, ddRADseq, genetics, Natal Bioregion, sardine run

**1 | INTRODUCTION**

Potential drivers of population structure in the marine environment include physical distance (isolation by distance) or boundaries (e.g., land masses), environmental transitions that define habitat regions (such as thermal boundaries; e.g., Teske et al., 2019) and oceanic currents (especially for organisms with larval drift; e.g., Knutsen et al., 2007). Historical climate change influencing the connectivity and
distribution of marine organisms via changes in sea level, sea surface temperature and ocean currents can be important in generating vicariance effects (e.g., Brierley & Kingsford, 2009; Gray et al., 2018; Hewitt, 2000). For mobile organisms without a larval life cycle, complex social structure may cause reproductive isolation by encouraging philopatry and limiting dispersal between local populations (Lowther-Thieleking et al., 2015; Parsons et al., 2006; Rosel et al., 2009; Sellas et al., 2005; Van Cise et al., 2017). Resource specializations based on habitat characteristics may increase fitness for philopatric individuals for both marine and terrestrial species (reviewed by Sargeant, 2007), and behavioural or morphological adaptations associated with these specializations could cause assortative mating or physical isolation of populations via philopatry (e.g., Smith & Skulason, 1996).

Highly mobile marine species with strong dispersal potential, such as cetaceans, may be expected to show panmixia, or low levels of genetic differentiation (e.g., Garber et al., 2005; Horne & van Herwerden, 2013; Moura et al., 2013; Reece et al., 2010). Factors including geographical isolation, local genetic drift or isolation by distance can promote population differentiation over large distances (see Palumbi, 2003). However, differentiation has also been found between populations across relatively small geographical distances for various cetacean species (e.g., Kershaw et al., 2017), especially delphinids (see Hoelzel, 2009) and bottlenose dolphins (Tursiops spp.) in particular (Moura et al., 2020; Natoli et al., 2005, 2008; Pratt et al., 2018). This has sometimes been associated with specialization on prey type (Hoelzel et al., 1998; Lowther-Thieleking et al., 2015; Möller et al., 2007; Natoli et al., 2005; Rosel et al., 2009; Segura-Garcia et al., 2018; Sellas et al., 2005), possibly driven by assortative mating or philopatry associated with ecotype (Hoelzel, 1991; Lowther-Thieleking et al., 2015). Killer whales (Orcinus orca) provide a striking example where regional sympatric ecotypes reflecting prey specializations are genetically differentiated (e.g., Hoelzel et al., 2007).

Bottlenose dolphins (Tursiops sp.) provide a useful model for studying the evolution of population structure in marine environments, because a wealth of information exists surrounding the ecology of this genus, providing insight into the mechanisms that might cause reproductive isolation among local populations (e.g., Barros & Wells, 1998; Connor et al., 2000; Fury & Harrison, 2008). The Indo-Pacific bottlenose dolphin (Tursiops aduncus), confirmed as a distinct species following support from genetic data (Moura et al., 2020; Wang et al., 1999), is distributed across the Indian Ocean and southwestern Pacific, including along the southern and eastern coasts of South Africa. We focus on the populations in South Africa, because here the distribution runs through distinct bioregions and is impacted by a seasonal migration of prey species, allowing us to test the possible role of each factor in the evolution of population structure.

Putative population structure along the East Coast of South Africa has previously been studied using allozyme, mitochondrial DNA (mtDNA) and microsatellite DNA markers (e.g., Goodwin et al., 1996; Natoli et al., 2008), though potential inference from the allozyme study was limited to two somatic loci and one sex-linked locus. Natoli et al. (2008), using mtDNA control region sequence and nine microsatellite DNA loci, found small but significant differentiation between samples from subpopulations located north and south of Ifafa (see Figure 1). Weaker evidence was found for differentiation between resident animals located south of Ifafa and animals thought to migrate northwards into KwaZulu Natal (KZN) waters during winter (June–August), coinciding with the annual migration of sardine (Sardinops sagax) into the area (the “sardine run”; van der Lingen et al., 2010; Natoli et al., 2008). On this basis, the existence of two resident and one migratory subpopulation was proposed. The putative migratory animals, estimated to number over 2000 individuals (see Natoli et al., 2008), were proposed to originate from as far south as Plettenberg Bay, and to not travel further north than Ifafa (Natoli et al., 2008). Differentiation between the north and south KZN subpopulations either side of Ifafa was later confirmed by Gopal (2013) based on mtDNA control region sequencing and 14 microsatellite DNA loci, though not clearly supported based on 14 microsatellite DNA loci in Gray et al. (2021). Furthermore, 12 samples from Plettenberg Bay were compared with mtDNA data from Natoli et al. (2008) and haplotype frequency differences were evident between Plettenberg Bay and the Natal Bioregion (e.g., KZN North, KZN South and migratory population on the North Eastern Cape; Gridley, 2011).

Reliable identification of genetically distinct subpopulations is essential for defining effective management units, choosing management interventions, refining future assessments, and monitoring conservation status and trends. Broader conservation inference can be gained when the study system permits testing of transferable questions about the drivers of population structure, as is the case for the T. aduncus populations off South Africa. Genome sampling can greatly enhance the resolution of population structure compared to microsatellite and mtDNA data, and sometimes identify patterns not evident at lower resolution. Therefore, we use high-resolution genome sampling to assess if the transition between the Natal and Agulhas Bioregions provides a barrier to gene flow among T. aduncus populations, even though there is no physical barrier restricting movement. This hypothesis is based on Tursiops sp. population structure identified elsewhere that was putatively associated with habitat boundaries (e.g., Natoli et al., 2005) and is more generally relevant to the impact of the transition between marine bioregions on the distribution and isolation of predatory species. We also test the hypothesis that temporal patterns of prey abundance can affect the dispersion behaviour and population genetic structure of highly mobile species such as T. aduncus, dependent on linear coastal habitat, potentially explaining the earlier reporting of population structure along the Natal coast. These potential drivers of evolutionary differentiation are relevant to a broad range of mobile predatory species, and have the potential to explain cryptic patterns of structure important to effective conservation and management.
2 | METHODS

2.1 | Study area and specimen collection

The study area was along the Southern Coastal and Shelf Waters of the South Africa IMMA (Important Marine Mammal Area; IUCN-Marine Mammal Protected Areas Task Force, 2020; IUCN-MMPATF, 2020), established under the IUCN as an important habitat for *Sousa plumbea*, *Balaenoptera edeni*, *T. aduncus* and *Delphinus delphis* (https://www.marinemammalhabitat.org/portfolio-item/southern-coastal-shelf-waters-south-africa/). The area comprised the Agulhas and Natal Bioregions (see Figure 1 for both place-name landmarks and sampling locations). The Natal Bioregion is strongly influenced by the warm Agulhas southward-flowing current and has a narrow continental shelf (~5 km wide). The Agulhas Bioregion (from Cape Point to the Mbashe River; Sink et al., 2004) is characterized by a broad continental shelf (up to 240 km) and here the Agulhas current moves offshore. It is home to the largest number of South African marine endemic species (Sink et al., 2012). In the Agulhas Bioregion, 45 samples were collected from 2013 to 2016, including 42 that were collected using biopsy darts and three that were collected from stranded dolphins. Of the 45 samples, 24 were collected from Plettenberg Bay and 21 were collected from Knysna, and 15 of the 45 were collected during the sardine run (May–August). We consider Plettenberg Bay and Knysna as two putative populations and test them in that context, based on data from elsewhere in the distribution of *T. aduncus* where differentiation has been found among similar embayments and habitat transitions (e.g., Wiszniewski et al., 2009).

In the Natal Bioregion (from Mbashe River to Cape Vidal; Sink et al., 2004), the sampling area spanned from south of Port Edwards (Natoli et al., 2008) to Richards Bay in KZN, covering 350 km of coastline. From 1994 to 2000, 107 tissue samples were collected from the Natal Bioregion and have previously been analysed using mtDNA and microsatellite DNA markers (Natoli et al., 2004, 2008). Of these 107 samples, 96 provided useful data and among those 48 samples were collected via biopsy sampling of dolphins moving north in large groups of over 500 dolphins along the North Eastern Cape on the Natal Bioregion (20 km or more south of the KZN–Eastern Cape border; Natoli et al., 2008). Because of the sampling time and group characteristics, these samples were considered by Natoli et al. (2008) as belonging to the putative migratory population. The remaining 48 samples were collected by the Natal Shark Board from dolphins caught in shark nets along the North KZN at 10 locations between Richards Bay and Ifafa and South KZN at 13 locations between Ifafa and Port Edward (20 and 28 samples respectively).
total of 59 samples were collected from May to August (during the sardine run period) along the Natal Bioregion. In this Bioregion a division between three putative populations had been suggested, North KZN, South KZN and the sample from the migratory group, and so for this study we consider a total of five putative populations in the sample set (North KZN, South KZN, Eastern Cape, Plettenberg Bay and Knysna). This sampling strategy is in support of testing our two main hypotheses: that a point of genetic division will be found at the transition between bioregions, and that the seasonal change in prey distribution and abundance associated with the sardine run will influence population structure.

Skin samples collected from 2013 to 2016 were frozen and preserved in 90% ethanol. Genomic DNA was extracted using an E.Z.N.A. Tissue DNA Kit (OMEGA biotech). Samples from previous studies (Natoli et al., 2004, 2008) were already available as DNA extracted by the phenol/chloroform method (after Hoelzel, 1998) and stored at −20°C.

### 2.2 RAD library preparation

The library was prepared using the double-digest restriction site associated DNA sequencing (ddRADseq) protocol (Peterson et al., 2012). Extracted DNA was examined for quality on agarose gels and quantified using a Qubit 2.0 Fluorometer (Invitrogen) and a NanoDrop Lite Spectrophotometer (Thermo Scientific). The DNA was double digested at 37°C overnight using the restriction endonucleases MspI and HindIII (according to the manufacturer’s protocols). The resulting DNA fragments were then ligated (according to the manufacturer’s protocols) to P1 and P2 adapters, so that each sample was uniquely barcoded, and divided into pools of 12 samples each. Size selection of DNA fragments between 460 and 560 bp was performed using a Pippin Prep (Sage Science). The pools were combined to produce final DNA libraries containing 10 ng DNA from each pool. Concentrations were confirmed by quantitative polymerase chain reaction (qPCR) and 125-bp paired-end sequencing was performed on an Illumina HiSeq 2500 in two lanes.

### 2.3 Genotype calling

Sequencing data were processed using the STACKS version 1.35 pipeline (Catchen et al., 2011, 2013). Reads were trimmed to 110 bp in length, filtered for quality and demultiplexed using the “process_radtags” program. Reads were aligned to a T. truncatus reference genome (Tur_tru_Illumina_hap_v1 [GenBank accession GCA_003314715.1]) obtained from the Ensembl Genome Browser, release 91 (Zerbino et al., 2018). Alignment was performed using the samtools version 1.2 (Li, 2011; Li et al., 2009), with reads that aligned more than once, aligned nonconcordantly or had low mapping quality (defined as a MAPQ value below 20) being filtered out of the data set. Single-nucleotide polymorphisms (SNPs) were detected using the STACKS program “ref_map” (Catchen et al., 2011, 2013). During short-read alignment the minimum depth of coverage to create a stack was 3, and the maximum distance (in bp) between stacks was 2. The number of mismatches allowed when building a catalogue was 2. Loci were filtered using the STACKS program “populations” (Catchen et al., 2011, 2013), removing those that had a stack depth below 8, or were not present in at least 70% of individuals or in all five groups according to sampling areas (see above for a description of the areas). The lowest allele frequency was 2.5% and loci were not filtered for minor allele frequency. Individuals with greater than 40% missing data were removed from the data set. Data were output from STACKS in the Genepop format and converted to other formats using PGD SPIDER (Lischer & Excoffier, 2012) or PLINK 1.9 (Purcell et al., 2007). Potential duplicate samples were screened for in CERVUS 3.0.7 (Kalinowski et al., 2007) allowing up to 100 mismatches.

### 2.4 Selection analysis

Loci under positive selection were identified by the Fdist method using LOSITAN (Antao et al., 2008) with an infinite alleles model for 50,000 simulations, a subsample size of 30, a confidence interval of 0.95 and a false discovery rate of 0.05. Loci with excessively high $F_{ST}$ values (under positive selection) were removed from the data set for all analyses relying on an assumption of neutrality and were analysed separately to neutral loci in all remaining analyses. By the Fdist method some putative outliers may be due to strong drift, but we chose it because loci under selection are less likely to be included among “neutral” loci than for some alternative methods. Both outlier and neutral loci were investigated for summary statistics and ordination analyses, but only neutral loci were used for assessments of structure or gene flow based on neutral assumptions (e.g., Admixture, Geneland, Correlograms, GeneClass and BayesAss).

### 2.5 Summary statistics

As indicated in the section on study area, we sampled from five putative populations defined based on geography and past studies: North KZN, South KZN, North Eastern Cape (the putative migratory population), Plettenberg Bay and Knysna (Figure 1). Summary statistics were estimated using ARLEQUIN version 3.5.2.2 (Excoffier & Lischer, 2010). Genetic differentiation between populations, indicated by the fixation index $F_{ST}$ was estimated using 10,000 permutations and a significance level of 0.05 after Bonferroni correction. Deviation from Hardy–Weinberg equilibrium (HWE) was evaluated with observed ($H_e$) and expected ($H_e$) heterozygosity was tested for each neutral locus using the “perform exact test of Hardy–Weinberg equilibrium” option with 1,000,000 steps in the Markov chain and 100,000 dememorization steps. A Bonferroni correction was again applied to the significance
level to account for multiple comparisons. A hierarchical assessment of diversity was implemented by an analysis of molecular variance (AMOVA) for various putative groupings using the "locus by locus AMOVA" option with 20,000 permutations, the "include individual level" option selected and a distance matrix being computed based on the pairwise difference method. Multiple putative subdivisions were assessed, for both neutral and outlier loci, as a way to help determine the most likely pattern of structure. The extent of inbreeding, indicated by the inbreeding coefficient $F_{IS}$ was calculated for each population for neutral and outlier loci.

### 2.6 | Population structure discovery

Population structure was investigated using multiple complementary methods, each with different assumptions and limitations, in order to reinforce inference about revealed patterns. We estimated individual ancestries for neutral loci using **admixture** version 1.3 (Alexander et al., 2009). The number of hypothetical ancestral populations ($K$) was varied between runs allowing the optimum value of $K$ to be identified. Ten independent runs were performed for each value of $K$ between 1 and 6, using a different random seed for each run and with the termination criterion for loglikelihood increase between two consecutive iterations set to $10^{-4}$. A 10-fold cross-validation was performed for each run to identify the optimum value of $K$. The run with the lowest cross-validation error was selected for interpretation. Principal component analysis (PCA; Hotelling, 1933a, 1933b; Pearson, 1901) and discriminant analysis of principal components (DAPC; Jombart et al., 2010) were performed on neutral and outlier loci using the package **adegenet** version 2.1.1 in R version 3.4.1 (Jombart, 2008; Jombart & Ahmed, 2011; RCoreTeam, 2017). When prior clustering was not provided for DAPC, successive $K$-means clustering with an increasing number of clusters ($K$) was used to identify groups by maximizing the variation between them. The optimum value of $K$ was inferred using the Bayesian information criterion (BIC) corresponding to each value of $K$ tested (Jombart & Collins, 2015; Jombart et al., 2010). DAPC was performed using the function "dapc" (Jombart et al., 2010), either with prior clusters corresponding to the five putative populations, or with clusters produced by "find.clusters" for values of $K$ between 2 and 4. The number of retained principal components for each analysis was chosen to optimize the a-score in each case (see Table S1). Given that $K$ was low for all analyses, the number of retained discriminant functions in each case was chosen to be one fewer than the value of $K$, as suggested by Jombart and Collins (2015). Assignment of individuals to clusters was visualized using the function "complot" (Jombart et al., 2010). Three-dimensional factorial correspondence analysis (FCA; Benzécri, 1973) was performed on neutral and outlier loci using the program **geneclass2** version 2.0 (Piry et al., 2004). Three analyses were performed: one investigating migration among all five putative populations, one investigating migration among North KZN, South KZN and North Eastern Cape during the sardine run, and one investigating migration among these putative populations outside the sardine run. For each analysis, the program was run with the $L_{home}/L_{max}$ likelihood computation. The Bayesian method (Rannala & Mountain, 1997) was used for likelihood computations. Individual assignment was tested for one putative population at a time, with individuals being assigned to one of two reference populations. For all assignment computations, an assignment threshold of 0.01 was used, along with the same Bayesian criterion (Rannala & Mountain, 1997) used for migrant detection. Recent immigration rates among putative populations and individual immigrant ancestries were examined for neutral loci using a Bayesian method in **bayesass** version 3.04 (Wilson & Rannala, 2003). The same three analyses were performed as for migrant detection in **geneclass2**. Ten independent runs were performed for each analysis, using a different random seed for each. The proposal step lengths of the mixing parameters were adjusted for each analysis (Table S2) to achieve optimal acceptance rates of between 20% and 60% (Rannala, 2007) for each parameter. The program was run each time for 1,000,000 iterations, with a burn-in length of 100,000 and a sampling interval of 100, using the "genotypes" option to output individual ancestry and the "trace" option to output a trace file. As suggested by Faubet et al. (2007) and Meirmans (2014), the trace file was used to calculate the deviance for each run, and for each

### 2.7 | Spatial analyses of population structure

The location of genetic discontinuities between individuals was inferred for neutral loci using the package **geneland** version 4.0.8 in R version 3.4.1 (Guillot et al., 2005, 2008; RCoreTeam, 2017). Samples from the KZN region were from net bycatch, and the precise location of the nets was used. Locations further southwest were located only to the specific bay or launch site (Knysna, Plettenberg Bay and the North Eastern Cape). Five independent runs were performed, each for 100,000 iterations, with a burn-in length of 100 and a sampling interval of 200. The run with the highest average posterior probability was selected for interpretation.

The relationship between genetic and geographical distance along the KZN coast was investigated for neutral loci using spatial autocorrelation in **genalex** version 6.503 (Peakall & Smouse, 2006). Two analyses were performed using samples collected outside and during the sardine run. For each analysis, the program was run for 9999 permutations, with 10,000 bootstraps and 10 distance classes of size 0.2. Considering the proximity of most of the KZN sampling locations, Richards Bay was excluded given the large distance from there to the nearest location in North KZN (after Natoli et al., 2008).
analysis the run with the lowest deviance was selected for interpretation. Circos plots of migration rates from these runs were generated using circos table viewer version 0.69 (Krzywinski et al., 2009).

3 | RESULTS

The "populations" program in STACKS identified 4985 SNPs (Vargas et al., 2021). Twelve individuals (comprising one from each of Zinkwazi, Thompson's Bay, Umlahta, Southport, Ramsgate, Port Edward and Knysna, two from Durban and three from Tongaat), had >40% missing data and therefore were removed from the data set, leaving 140 individuals for analyses. CERVUS identified three pairs in the Eastern Cape sample set that had 10–17 mismatches between the pairs. It is not known if these represent similar genotypes or duplicates (due to sequencing errors). We repeated admixture and summary statistical analyses with one individual removed from each pair, and there was no difference to the outcome (\(K = 2\) is still best supported for admixture, and the \(F_{ST}\) table was essentially the same, with no change in significance; data not shown). Selection analysis using LOSITAN identified 180 outlier loci under positive selection, with the remaining 4805 loci being classified as neutral or under balancing selection. Mean values for \(H_D\) and \(H_E\) were similar among putative populations and mean \(F_{IS}\) values were close to zero (Table 1). Only between 0.03% and 0.86% of neutral loci in each putative population showed significant deviation from HWE following Bonferroni correction (\(\alpha = 2.08 \times 10^{-6}\)), and the mean \(p\)-value was >.7 for each putative population. Thus, the results convey no deviation from HWE for the majority of loci. However, we tested this by repeating the admixture analysis after removing the 37 loci that were out of HWE in one or more populations, and there was no detectable difference (data not shown).

3.1 | Population structure

Pairwise \(F_{ST}\) values were greatest, and significantly different, between each of the Natal putative populations on the East Coast (North KZN, South KZN and North Eastern Cape) and the Agulhas putative populations on the South Coast (Plettenberg Bay and Knysna; Table S3; Figure 2) consistent with our hypothesis that there would be differentiation between populations in the different bioregions. The \(F_{ST}\) value between North KZN and North Eastern Cape (referred to as the "migratory" population; Natoli et al., 2008) was also significant but low. Pairwise \(F_{ST}\) values between adjacent putative populations within the Natal and Agulhas Bioregions were not significant (Table S3). The two-way \(F_{ST}\) comparing Plettenberg Bay and Knysna combined (South Coast) against North KZN, South KZN and North Eastern Cape combined (East Coast) was 0.038. The optimum value of \(K\), corresponding to the run with the lowest cross-validation error in admixture analyses (Figure S1), was 2 (Table S4), and larger values of \(K\) did not show a clear pattern of structure within the Natal Bioregion. When the South Coast and East Coast were analysed separately, the lowest cross-validation error in each case was for \(K = 1\) (Table S4).

Clustering in GENELAND indicated that four populations exist in South Africa, corresponding to North KZN, South KZN, North Eastern Cape (during the sardine run) and the combined South Coast locations (between which there was no differentiation; Figure 3). Although one individual sampled in South KZN clustered with the North KZN putative population, this individual was sampled at Sunwich Port, the most northerly sampling location in South KZN.

In the FCA plots (Figure 4) and PCA plots (Figure S2), the putative populations on the East Coast clustered separately from the putative populations on the South Coast for both neutral and outlier loci. In all six plots, the East and South Coast clusters were separated along the first axis or first principal component (which explains the greatest percentage of the total variation in the data). For the FCAs, when the "sur populations" option was not selected, limited differentiation was evident within the East Coast cluster for both neutral loci (Figure 4a) and outlier loci (Figure 4b). When the "sur populations" option was selected, however, differentiation among East Coast populations became apparent, but was more pronounced for neutral loci (Figure 4c) than for outlier loci (Figure 4d), possibly due to power associated with the reduced number of outlier loci.

<table>
<thead>
<tr>
<th>Population</th>
<th>(N)</th>
<th>Locus type</th>
<th>(H_D)</th>
<th>(H_E)</th>
<th>(F_{IS})</th>
</tr>
</thead>
<tbody>
<tr>
<td>North KZN</td>
<td>20</td>
<td>Neutral</td>
<td>0.265 (±0.182)</td>
<td>0.265 (±0.161)</td>
<td>-0.006</td>
</tr>
<tr>
<td>South KZN</td>
<td>28</td>
<td>Neutral</td>
<td>0.323 (±0.188)</td>
<td>0.324 (±0.159)</td>
<td>0.028</td>
</tr>
<tr>
<td>North Eastern Cape</td>
<td>48</td>
<td>Neutral</td>
<td>0.259 (±0.178)</td>
<td>0.259 (±0.163)</td>
<td>-0.002</td>
</tr>
<tr>
<td>Plettenberg Bay</td>
<td>24</td>
<td>Neutral</td>
<td>0.340 (±0.176)</td>
<td>0.316 (±0.151)</td>
<td>-0.053</td>
</tr>
<tr>
<td>Knysna</td>
<td>20</td>
<td>Neutral</td>
<td>0.242 (±0.179)</td>
<td>0.242 (±0.167)</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outlier</td>
<td>0.300 (±0.149)</td>
<td>0.308 (±0.142)</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neutral</td>
<td>0.291 (±0.180)</td>
<td>0.293 (±0.157)</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outlier</td>
<td>0.322 (±0.177)</td>
<td>0.331 (±0.175)</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neutral</td>
<td>0.302 (±0.187)</td>
<td>0.296 (±0.155)</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outlier</td>
<td>0.310 (±0.175)</td>
<td>0.327 (±0.160)</td>
<td>0.002</td>
</tr>
</tbody>
</table>
Coast and South Coast samples were analysed separately, essentially the same pattern was seen but the separation among putative populations in the East Coast was somewhat clearer (Figure S4). The DAPC plots when prior putative populations were provided for both neutral and outlier loci also showed some differentiation between the populations on the East Coast (Figure S3). When East Coast and South Coast samples were analysed separately, essentially the same pattern was seen but the separation among putative populations in the East Coast was somewhat clearer (Figure S4). The number of principal components retained based on the a-scores for different runs is provided in Table S5. When prior putative populations were not provided for DAPC analyses, the optimum value of K was 2 for both neutral and outlier loci (Figure S5). These two clusters corresponded to the three putative populations on the East Coast (Figure S6a,b). For outlier loci, increasing the value of K beyond 2 revealed further differentiation between each of the three East Coast putative populations, but not between the South Coast putative populations (Figures S6d,f and S7). For neutral loci, increasing the value of K beyond 2 did not reveal any clear structure beyond the two main clusters corresponding to the majority of individuals from the East Coast and all individuals from the South Coast (Figures S6c,e and S7c,e).

For the AMOVA test, the strongest differentiation was again between the South and East Coast groupings, but significant differentiation within the Natal Bioregion was also found (see Table 2). Geneclass2 assigned individuals from the North Eastern Cape to North KZN and South KZN, with 87.5% of individuals assigned to South KZN, and 12.5% to North KZN. Similar variation in assignment was found amongst individuals from North KZN, with 70% of individuals assigned to the North Eastern Cape, and 30% to South KZN. No variation in assignment was found for South KZN, with all individuals assigned to the North Eastern Cape over North KZN. These results suggest greater genetic differentiation between North KZN and the North Eastern Cape than between South KZN and the North Eastern Cape, consistent with the FST data (Table S3; Figure 2).

3.2 | Recent gene flow

Migrant detection in Geneclass2 (Table 3) identified a large number of first-generation migrants from the North Eastern Cape in both KZN putative populations. A small number of migrants from South KZN were also identified in North KZN and the North Eastern Cape, but no migrants from North KZN were identified in any of the other putative populations. In contrast, migrant detection in Bayesass (Table 3; Figure 5) identified recent migrants from North KZN in both South KZN and the North Eastern Cape, and a small number of migrants from the North Eastern Cape in South KZN. Both Geneclass2 and Bayesass identified bidirectional migration between Plettenberg Bay and Knysna, although the migration rate was greater from Plettenberg Bay to Knysna than in the opposite direction.

Analysis of individual ancestry in Bayesass (Table 4) produced similar results to Geneclass2 (Table 3), but identified additional migrants from the North Eastern Cape in North KZN and a single migrant from South KZN in Plettenberg Bay that were not identified by Geneclass2 migrant detection. All migrants identified in North KZN, North Eastern Cape and Plettenberg Bay, along with the majority of those identified in South KZN and Knysna, were second-generation. Apart from the single South KZN migrant in Plettenberg Bay, no other migration was identified between the putative populations on the East Coast and those on the South Coast.

3.3 | Influence of the sardine run on population structure and migration

FST values among some putative populations differed according to the season in which samples were collected (Table S6; Figure 2), consistent with our hypothesis that the annual change in prey distribution and abundance associated with the sardine run could influence population structure. While samples from North KZN collected during the sardine run were significantly differentiated from all South KZN and North Eastern Cape samples, those collected outside the sardine run were not. Additionally, samples from South KZN
collected during the sardine run were not significantly differentiated from North Eastern Cape samples, whilst those collected outside the sardine run were. $F_{IS}$ values were nonsignificant, and positive only for northern KZN outside of the sardine run period (Table S7).

$F_{ST}$ values comparing the East Coast and South Coast putative populations remained significant both during and outside the sardine run (Figure 2; Table S6). Temporal sample comparisons for each location (e.g., South KZN during compared to outside the sardine run) were only significantly different for the North Eastern Cape. Note, however, that most sample sizes are relatively small, and the sample from Knysna during the sardine run (bold in Table S6) is probably too small for useful inference.

Migrant detection in *geneclass2* revealed different patterns of migration between North KZN, South KZN and North Eastern Cape during and outside the sardine run (Table 5). For North KZN, the majority of individuals sampled during the sardine run were identified as nonmigrants, whereas the majority of individuals sampled outside the sardine run were identified as migrants from South KZN or North Eastern Cape. For South KZN, all individuals sampled during the sardine run were identified as migrants from the North Eastern Cape, whereas the majority of those sampled outside the sardine run were identified as nonmigrants. For North Eastern Cape, all individuals sampled during the sardine run were identified as nonmigrants, while half of those sampled outside the sardine run were
identified as migrants from South KZN, with some migrants from North KZN also being found.

The results of migrant detection in *bayesass* also conveyed different migration patterns during and outside the sardine run (Table 5; Figure 6). During the sardine run, migration rates were highest from the North Eastern Cape to North KZN and South KZN, with little other migration taking place. However, outside the sardine run, migration rates were highest from South KZN to North KZN and North Eastern Cape, with some migration also occurring from North Eastern Cape to North KZN and to a lesser extent South KZN. No significant migration from North KZN to South KZN or North Eastern Cape was identified during or outside the sardine run.

Spatial autocorrelation traces were mostly within the 95% confidence limits, but there is some indication of greater oscillation between positive and negative autocorrelation for samples collected during the sardine run (Figure S8). This pattern could suggest greater spatial clustering of genotypes for these samples (Neville et al., 2006; Peakall et al., 2003; Smouse & Peakall, 1999), but in this case these analyses do not provide strong inference.

4 | DISCUSSION

Biodiversity conservation is challenged by a lack of understanding of the mechanisms that generate evolutionary units within a species. In this study we investigated these processes by applying high-resolution population genetic analyses to a species distribution affected by complex physical and temporal environmental factors. Summary statistics and ordination analyses showed strong differentiation between identified populations in the Agulhas and Natal
Bioregions as hypothesized, with low migration rates being found throughout the year, and no recent gene flow. There are no barriers to movement along this stretch of open coastline, but the level of differentiation can be low (in the context of the marker system used) for Indo-Pacific bottlenose dolphin populations elsewhere over a similar geographical range (e.g., between Taiwan and Japan where $F_{ST} = 0.013$ for 20 microsatellite DNA loci; Chen et al., 2017). At the same time, the distance between the Agulhas and Natal study sites is relatively large, and it will be useful in the future to sample more contiguously along the coastline through this range of transition.

### Table 2: AMOVA results for different population groupings for neutral and outlier loci

<table>
<thead>
<tr>
<th>Groups</th>
<th>Locus type</th>
<th>Source of variation</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage variation</th>
<th>Fixation index</th>
</tr>
</thead>
<tbody>
<tr>
<td>• North KZNSouth KZN</td>
<td>Neutral</td>
<td>Among groups</td>
<td>1,147.758</td>
<td>7.293</td>
<td>2.672</td>
<td>0.027***</td>
</tr>
<tr>
<td>North Eastern Cape</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>903.975</td>
<td>0.708</td>
<td>0.259</td>
<td>0.003***</td>
</tr>
<tr>
<td>• Plettenberg Bay Knysna</td>
<td>Neutral</td>
<td>Among groups</td>
<td>417.450</td>
<td>3.442</td>
<td>19.427</td>
<td>0.194***</td>
</tr>
<tr>
<td>• North KZNSouth KZN</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>66.367</td>
<td>0.148</td>
<td>0.837</td>
<td>0.010***</td>
</tr>
<tr>
<td>• North Eastern Cape</td>
<td>Neutral</td>
<td>Among groups</td>
<td>1795.445</td>
<td>5.357</td>
<td>1.983</td>
<td>0.020***</td>
</tr>
<tr>
<td>• Plettenberg Bay Knysna</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>256.288</td>
<td>-0.221</td>
<td>-0.082</td>
<td>-0.001</td>
</tr>
<tr>
<td>• North KZNSouth KZN</td>
<td>Neutral</td>
<td>Among groups</td>
<td>467.643</td>
<td>2.165</td>
<td>13.256</td>
<td>0.133***</td>
</tr>
<tr>
<td>North Eastern Cape</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>16.174</td>
<td>0.042</td>
<td>0.259</td>
<td>0.003</td>
</tr>
<tr>
<td>• Plettenberg Bay Knysna</td>
<td>Neutral</td>
<td>Among groups</td>
<td>1738.983</td>
<td>4.023</td>
<td>1.490</td>
<td>0.015***</td>
</tr>
<tr>
<td>North Eastern Cape</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>312.751</td>
<td>1.065</td>
<td>0.394</td>
<td>0.004***</td>
</tr>
<tr>
<td>• Plettenberg Bay Knysna</td>
<td>Neutral</td>
<td>Among groups</td>
<td>453.985</td>
<td>1.866</td>
<td>11.426</td>
<td>0.114***</td>
</tr>
<tr>
<td>• North KZNSouth KZN</td>
<td>Neutral</td>
<td>Among groups</td>
<td>29.832</td>
<td>0.343</td>
<td>2.100</td>
<td>0.024***</td>
</tr>
<tr>
<td>• North Eastern Cape</td>
<td>Neutral</td>
<td>Among groups</td>
<td>1,720.636</td>
<td>4.577</td>
<td>1.692</td>
<td>0.017***</td>
</tr>
<tr>
<td>• Plettenberg Bay Knysna</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>331.098</td>
<td>0.998</td>
<td>0.369</td>
<td>0.004***</td>
</tr>
<tr>
<td>• North KZNSouth KZN</td>
<td>Neutral</td>
<td>Among groups</td>
<td>464.567</td>
<td>2.414</td>
<td>14.532</td>
<td>0.145***</td>
</tr>
<tr>
<td>• North Eastern Cape</td>
<td>Neutral</td>
<td>Among populations within groups</td>
<td>19.250</td>
<td>0.072</td>
<td>0.435</td>
<td>0.005*</td>
</tr>
</tbody>
</table>

Note: Each group is denoted by a bullet-pointed list. *p < .05, ***p < .001.

### Table 3: Percentage of individuals in each sample population that were identified as nonmigrants, first-generation migrants from another source population in *geneclass2* (top values) and recent migrants from another source population in *bayesass* (bottom values) for neutral loci

<table>
<thead>
<tr>
<th>Sample population</th>
<th>Nonmigrant</th>
<th>Source population</th>
<th>North KZN</th>
<th>South KZN</th>
<th>North Eastern Cape</th>
<th>Plettenberg Bay</th>
<th>Knysna</th>
</tr>
</thead>
<tbody>
<tr>
<td>North KZN</td>
<td>40.00</td>
<td>North KZN</td>
<td>5.00</td>
<td>1.33</td>
<td>1.33</td>
<td>1.30</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>91.39*</td>
<td>South KZN</td>
<td>0.00</td>
<td>6.25</td>
<td>0.63</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>78.01*</td>
<td>North Eastern Cape</td>
<td>0.00</td>
<td>0.62</td>
<td>0.62</td>
<td>37.50</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>62.50</td>
<td>Plettenberg Bay</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>70.00</td>
</tr>
<tr>
<td></td>
<td>30.00</td>
<td>Knysna</td>
<td>0.00</td>
<td>1.13</td>
<td>1.34</td>
<td>1.34</td>
<td>26.69*</td>
</tr>
</tbody>
</table>

*Lower bound of 95% credible set (mean − 1.96 × SD) > 0.00, for *bayesass* results only.
between habitats. Although not extensively studied, there are other reports of genetic differentiation between marine populations in the Agulhas and Natal Bioregions, for example for the catface grouper (*Epinephelus andersoni*; Coppinger et al., 2019).

No differentiation was found between the individuals from the two study areas in the Agulhas Bioregion, which may be expected since the distance between them is relatively small (Figure 1). However, *T. aduncus* from Australian populations showed differentiation across a similar range and are divided by similar geographical features, such as promontories and embayments (e.g., Bilgmann et al., 2007). Reasons for the difference are unknown, but may be associated with historical or environmental factors (e.g., displacement and mixing during the sardine run in South Africa). The degree of differentiation between the Agulhas and Natal regions was stronger for our outlier than for our neutral markers (e.g., Table 2). However, the patterns of structure were consistent with that seen for neutral markers, and it is therefore difficult to determine what proportion of this may be due to selection or loci affected by strong drift.

Weak but significant genetic differentiation for some comparisons was identified between North KZN, South KZN and North Eastern Cape, consistent with data presented earlier (e.g., Natoli et al., 2008). However, given the potential for differential levels or patterns of mobility across seasons (e.g., in association with the annual sardine run; see Peddemors, 1999), it is possible that apparent patterns of structure are distorted by sampling true populations away from their core range, or sampling mixed populations. The sardine run occurs from approximately May to August when billions of sardines spawn over the Agulhas Bank and then migrate north-east along the coast, running closer to shore off KZN than along the North Eastern Cape (Connell, 2010; van der Lingen et al., 2010;
Roberts et al., 2010). We tested the potential for differing mobility across seasons by comparing subsets sampled during and outside of the sardine run. For the Agulhas region on the South Coast (Knysna and Plettenberg), there was no apparent difference between seasons, either between these two putative populations or between them and the KZN region. However, the potential for comparisons with Knysna during the sardine run was compromised by small sample size (Table S6).

For the East Coast samples (North Eastern Cape, KZN South and KZN North, Figure 1) seasonal comparisons showed different patterns, consistent with our hypothesis. When all samples were pooled together, gene flow was apparently from north to south (Figure 5; Table 3), and this was also the case between South KZN and the Eastern Cape when data were restricted to samples collected outside the sardine run (Figure 6; Table 5). The strongest signal for ancestral (2nd generation) migrants was from North KZN to South KZN and the North Eastern Cape (Table 4). During the sardine run, however, there is a strong signal for gene flow northward from the North Eastern Cape into KZN (Figure 6; Table 5). This is consistent with the direction of the sardine run, and may reflect the temporary movement of predating dolphins northward from their home range. The boundary at Ifafa within the KZN region reported on earlier (Natoli et al., 2008) is evident during the sardine run, but not for comparison of samples collected outside the sardine run. Outside the sardine run, the boundary is instead between the North Eastern Cape and KZN, a barrier not evident during the sardine run (Figure 6). One possible interpretation is that the real boundary is between the North Eastern Cape and KZN, a barrier not evident during the sardine run (Figure 6). One possible interpretation is that the real boundary is between the North Eastern Cape and KZN, which then appears to seasonally move northeast with the movement of dolphins following the sardine run (and their capture in the stationary shark nets). Previous studies (Peddemors, 1995) suggested that a seasonal migratory group of dolphins travels no further north than Ifafa, though we did detect some putative migrants from the North Eastern Cape in northern KZN (see Table 3). Another possibility is that both boundaries are real and reflect resident dolphin distributions obscured seasonally by migrating dolphins. This would be consistent with earlier reports of a boundary at Ifafa based on pollutant levels and differential sighting data (with coherent groups sighted in “preferred areas” either side of Ifafa over a period of several years; Cockcroft et al., 1989, 1990). More intensive sampling through this region and across seasons may help resolve this question.

This type of temporal variation in population structure has rarely been reported, but could have important consequences for effective conservation and management. For example, among populations of an African cichlid fish (Pseudocrenilabrus multicolor victoriae) in Uganda, Crispo and Chapman (2010) they found a strong isolation by distance pattern ($r^2 = .73$) that disappeared ($r^2 = .0$) 2 years later. They suggested that flooding in the intervening years promoting gene flow was a possible explanation, but drew no firm conclusions. In migratory species there is often seasonal displacement between summer and winter ranges, typically associated with breeding and foraging habitat. In this case, there may be direct population genetic continuity between summer and winter ranges, as has been seen between Eurasia and southern Africa vs Iberia and North Africa for the European bee-eater (Merops apiaster), although in that example the pattern is disrupted by a founder population established in Iberia from the Eurasian population (Ramos et al., 2016). Among migratory mysticete cetacean species, the pattern may reflect this type of direct continuity between breeding and feeding grounds, or sometimes a mixed assemblage of breeding populations on shared feeding grounds (reviewed by Hoelzel, 1998).
Our results reflect seasonal movement on a smaller geographical scale, probably influenced by the seasonal migration of an important prey resource. An implication may be that for a given geographical region, conservation efforts may need to recognize that different conservation units (populations) may occupy that area at different times of the year. At the same time, the full data set revealed some evidence for two possible boundaries: one between the North Eastern Cape and KZN, and the other north and south of Ifafa within...
KZN (see Figure 3). The important determination for effective management would be whether seasonal movement is either obscuring or artificially generating genetic boundaries. If resident populations were seasonally combined with a migratory population and sampled as one population, then a Wahlund effect and consequently higher values of $F_{IS}$ may be expected. There is only a very subtle (nonsignificant) indication of this in the north KZN region, where $F_{IS}$ is positive in one season, but this occurs outside the sardine run season, not during it when mixing with a migratory population may have generated that effect (Table S7). Another possibility may be that within the KZN region, populations keep within their core range when prey are abundant during the sardine run, and range more broadly searching for alternative prey outside the sardine run. Although this could explain the appearance and loss of the boundary at Ifafa during and outside of the sardine run, expected elevated $F_{IS}$ throughout KZN indicating mixing outside of the sardine run was not observed, and this hypothesis does not easily explain the strengthening of the boundary between KZN and the North Eastern Cape outside of the sardine run.

From a management perspective, recognition of seasonal genetic discontinuities both at Ifafa and between the North Eastern Cape and KZN would therefore be prudent. From an evolutionary perspective, however, the results suggest that genetic differentiation between populations is only partially determined by geography. Assortative mating via direct mate choice could be facilitated by identification of individuals or population members using learnt cultural traits, as has been suggested for other delphinids (Danchin & Wagner, 2010; Riesch et al., 2012). For at least part of the year these populations disperse or migrate, obscuring underlying biogeographical clusters. While this is expected of known migratory species such as humpback (Megaptera novaeangliae) or gray (Eschrichtius robustus) whales, it is less well-established for dolphin species, and especially for a high site fidelity species such as T. aduncus (Möller et al., 2007). The sardine run delivers a large annual injection of nutrients into this otherwise nutrient-poor region (Carter & D'Aubrey, 1988; Hutchings et al., 2010; Meyer et al., 2002), and therefore an impact on distribution might be expected, and consequently the observed effect on population genetic structure.

Potential changes to the timing or extent of the sardine run due to climate change might change this, and should be considered in future management planning. Continued and more extensive temporal sampling throughout this range would facilitate effective conservation management of this species in this region. On a broader scale in South African waters, the divergence between the Agulhas and Natal Bioregions was strong with little evidence for ongoing migration. Habitat conditions are distinct for these two regions (e.g., Sink et al., 2012), though stable isotope analyses for carbon and nitrogen comparing T. aduncus in KZN and Plettenberg Bay showed no clear indication of a difference in diet (Browning et al., 2014).

Our data contribute to ongoing regional conservation efforts, such as through the Marine Mammal Protected Area Task Force (see https://www.marinemammalhabitat.org/portfolio-item/southern-coastal-shelf-waters-south-africa/). The identification of the potential drivers of population differentiation provides broad inference for conservation across species in similar habitat. Understanding the potential for substructure among populations in species with high dispersal potential is challenging, especially in the marine environment when physical barriers to movement are uncommon. However, it is important in support of effective conservation and management. Here we add to the literature on barriers associated with habitat transitions, but also contribute data for a relatively rarely reported effect—the alteration of population structure temporally in association with annual changes in the environment (in this case associated with prey distribution and abundance).

**ACKNOWLEDGEMENTS**

Special thanks to numerous individuals from Ocean Odyssey, Enrico’s Fishing Charters, Orca Foundation, Ocean Blue Adventures and South African National Parks (Tsitsikamma) for providing research vessels, skippers and contributions of their time and fuel costs, and especially to all the skippers and volunteers that helped during this study. Special thanks to Vic Peddemors, the director and staff of the Natal Sharks Board for their support in collecting the samples from the KZN and North Eastern Cape (biopsies collected under permit to Natal Shark Board). This research was conducted under a Memorandum of Understanding for research between Nelson Mandela University and the former Department of Environmental Affairs (DEA), which is now the Department of Environment, Forestry and Fisheries. The study would not have been possible without DEA support. The project was permitted in terms of research permits RES 2013–67 and RES 2015–79, issued by the DEA, and animal ethics clearance A13-SCI-ZOO-001, issued by Nelson Mandela University. The field work was also supported by the Rufford Foundation (UK) and the Society for Marine Mammalogy (South Africa). The Nelson Mandela University Postgraduate Research Scholarship provided a fellowship to O.A.V.F. (2015–2018). Daniel Moore was supported through the NERC IAPETUS doctoral funding program (grant reference NE/L002590/1).

**CONFLICT OF INTEREST**

The authors declare no conflict of interest.

**AUTHOR CONTRIBUTIONS**


**DATA AVAILABILITY STATEMENT**

Sequences associated with RAD analyses are deposited at GenBank under BioProject accession PRJNA746068. RADseq genotype files
REFERENCES


**SUPPORTING INFORMATION**

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