ICES Journal of Marine Science

Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (*Thunnus albacares*)

Rachel B. Mullins[1],[†], Niall J. McKeown[2],[*],[†], Warwick H. H. Sauer[1], and Paul W. Shaw[1],[2]

[1]*Department of Ichthyology & Fisheries Science, Rhodes University, Grahamstown, South Africa*
[2]*Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, UK*

*Corresponding author: tel: 0044 1970622323; e-mail: njm2@aber.ac.uk.
[†]Equal contribution.

The South African (SAF) yellowfin tuna (*Thunnus albacares*) fishery represents a potential example of misalignment between management units and biological processes. The SAF fishery spans an operational stock with a boundary at 20°E, either side of which fish are considered part of Atlantic or Indian Ocean regional stocks. However, the actual recruitment of fish from Atlantic and Indian Ocean spawning populations into SAF waters is unknown. To address this knowledge gap, genomic analysis (11 101 SNPs) was performed on samples from Atlantic and Indian Ocean spawning sites, including SAF sites spanning the current stock boundary. Outlier loci conferred high discriminatory power to assignment tests and revealed that all SAF fish were assigned to the Indian Ocean population and that no Atlantic Ocean fish appeared in the SAF samples. Additionally, several Indian Ocean migrants were detected at the Atlantic spawning site demonstrating asymmetric dispersal and the occurrence of a mixed-stock fishery in Atlantic waters. This study highlights both the spatial inaccuracy of current stock designations and a misunderstanding of interactions between the underlying biological units, which must be addressed in light of local and global declines of the species. Specifically, the entire SAF fishery must be managed as part of the Indian Ocean stock.

Keywords: adaptation, assignment, conservation, dispersal, genetic differentiation, stock, sustainability

Introduction

The worldwide depletion of fish communities (Myers and Worm, 2005) with evidence of fishery induced economic (Botsford et al., 1997) and biological extinctions (Jackson *et al.*, 2001) highlights the importance of identifying biologically distinct units within marine fishes for both sustainable management and conservation of marine biodiversity (Ruzzante, 2006). The ability to monitor the dynamics of such components within systems involving seasonal migration and potential spatial overlap is beneficial, as more easily exploited and/or less productive units may be susceptible to overharvesting, contributing to loss of diversity and adaptive potential (Iles and Sinclair, 1982) and producing negative

effects on recruitment potential and population/fishery viability (Ryman *et al.*, 1995). Recent advances in population genomic methods offer considerable potential as tools to meet the many challenges associated with sustainable fishery management (Allendorf *et al.*, 2010). However, this potential is yet to be fully harnessed for a variety of reasons. The integration of genomics data and fishery management is particularly hampered in developing countries where threats to fishery sustainability may be most concentrated (Waples *et al.*, 2008; Willette *et al.*, 2014; Bernatchez *et al.*, 2017).

Yellowfin tuna, *Thunnus albacares* is globally distributed in tropical waters and supports fishery stocks that extend across

boundaries of national exclusive economic zones (EEZs) and into the high seas. The species accounts for the second largest world-wide catch of tuna and tuna-like species (after skipjack tuna *Katsuwonus pelamis*) in terms of catch weight and volume (Miyake *et al.*, 2010; Juan-Jordá *et al.*, 2013; FAO, 2017), with an average annual catch of ∼1.25 million metric tons over the past decade (Pecoraro *et al.*, 2017). At present, four regional tuna management units are described, Atlantic, Indian, Eastern Pacific, and West Central Pacific, and each is managed as a single stock by the respective tuna Regional Fisheries Management Organisations (tRFMOs) (Pecoraro *et al.*, 2017). Recent genomic studies have confirmed genetic differences and restricted inter-breeding among regional groups (Pecoraro *et al.*, 2016; Barth *et al.*, 2017). As previous studies have been macrogeographic in scale, there is still considerable uncertainty over the fine-scale boundaries and interactions among genetic groups and their congruence with current operational stock boundaries. Mismatch between biological and management units is recognized as a major threat to global fishery sustainability (Reiss *et al.*, 2009), and clarifying any such inconsistencies for yellowfin tuna is an urgent consideration as the management units are currently described as fully exploited and may even be overexploited (Majkowski, 2007), with the species being described as "near threatened" by the IUCN (Collette *et al.*, 2011).

The South African (SAF) yellowfin tuna fishery provides a striking case where management units may be incongruent with biological population processes of dispersal, interbreeding and adaptation. The present management boundary between the Atlantic and Indian Ocean operational stocks lies off South Africa at 20°E (FAO, 2017), and the two stocks are assessed and managed by the International Commission for the Conservation of Atlantic Tunas (ICCAT) and the Indian Ocean Tuna Commission (IOTC), respectively (Figure 1). Yellowfin tuna catches west of the 20°E are reported to ICCAT for inclusion in the Atlantic Ocean stock assessment, whereas catches east of this line are reported to the IOTC. However, this boundary is not based on any recognized biogeographic boundary or species-specific biological information but is based on a geographic
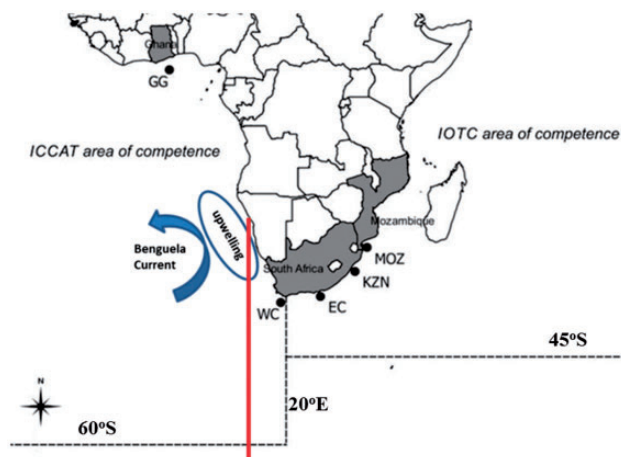
feature, the southernmost tip of Africa at Cape Agulhas at the confluence of the Atlantic and Indian Oceans. Yellowfin tuna do not spawn off the southern coast of South Africa, as the local environment does not provide conditions for optimal survival of early life stages (Pecoraro *et al.*, 2016), and thus adults occurring there represent allochthonous feeding migrants. Barth *et al.* (2017) suggested that the Benguela current system (BCS) along southwest African coast may be a barrier to dispersal of individuals from Atlantic spawning areas north of the BCS into SAF waters. It could therefore be hypothesized that the SAF fishery is being sustained solely by individuals from Indian Ocean spawning population(s). If so, the accuracy of current stock monitoring methods is fundamentally compromised. These inaccuracies need to be urgently resolved, because the Standing Committee on Research and Statistics of ICCAT has reported the Atlantic stock as over-fished (ICCAT, 2017), and the Scientific Committee of the IOTC has declared the Indian Ocean stock as overfished and has proposed an interim rebuilding plan (IOTC, 2016).

The present research aimed to build on and extend previous studies of yellowfin tuna by using RAD-Seq genotyping to assess the membership of yellowfin tuna adults sampled in SAF waters, from sites spanning the operational stock boundary, to Atlantic and Indian Ocean genetic stocks. A specific hypothesis being tested was that, in line with current management unit delineation, Western Cape and Southern/Eastern Cape yellowfin are derived from distinct Atlantic and Indian Ocean spawning populations, respectively. This study reports robust genetic patterns that both corroborate previous studies and reveal new aspects of yellowfin biocomplexity that collectively highlight both the spatial inaccuracy of current stock designations and a misunderstanding of dispersal patterns and population boundaries. These discrepancies prevent the accurate assessment of population productivity and dynamics, and so undermine the effectiveness of management actions that may compromise resilience and sustainability of the resource.

Material and methods
Sample collection, DNA isolation, and species verification
Our sampling strategy (Table 1 and Figure 1) was devised to include fish off the Western (Atlantic Ocean waters) and Eastern (Indian Ocean waters) Cape Provinces, as well as fish from putative spawning areas of regional stocks in the Atlantic Ocean [Gulf of Guinea (CG), ICCAT, 2016 and Indian Ocean (KwaZulu-Natal), IOTC, 2014]. Fin clips from individual yellowfin tuna were fixed in 95% ethanol. DNA was extracted from fin clips using a phenol/chloroform/isoamyl alcohol (PCIA) method following Winnepenninckx et al. (1993). A segment of the mitochondrial Control Region was amplified with the polymerase chain reaction (PCR) and sequenced in both directions using the species-specific primers (5′-TCCTACCCCTAACTCCCAAAG-3′; and reverse primer: 5′-AAACTGTGGGGATTCTCAC-3′). Sequences were used to confirm species identity using the BLAST function in GenBank. MtDNA summary statistics and estimates of inter-sample differentiation were calculated following McKeown et al. (2015).

SNP genotyping
SNP genotyping by sequencing was performed using tunable Genotyping By Sequencing (tGBS), a modified version of RAD-

Figure 1. Geographical location of sample sites (codes correspond to Table 1) with approximate location of the Benguela Current System (BCS), and 20°E operational stock boundary (dashed line). Solid line denotes proposed location of new management boundary at ∼13.35°E based on genetic results from this study.

Table 1. Sample information and summary indices of variation for mtDNA (N = number of individuals sequenced; NHap = number of distinct haplotypes resolved; h = haplotype diversity; π = nucleotide diversity) and nuclear SNP data including numbers of individuals genotyped (N) and polymorphic SNPs (NpolySNP), observed and expected heterozygosities (H_O and H_E, respectively) and their standard deviations, and F_{IS} (F_{IS} values all non-significant).

Sample (code)	Sample coordinates and date	mtDNA				Nuclear SNPs				
		N	N Hap	h	π	N	Npoly SNP	H_O	H_E	F_{IS}
Gulf of Guinea (GG)	2.5N/0.30W August 2015	20	20	1	0.025	20	10 710	0.228 (0.160)	0.256 (0.148)	0.13
Western Cape, South Africa (WC)	35.61S/18.70E October 2014	16	16	1	0.026	13	9 831	0.241 (0.163)	0.276 (0.145)	0.11
Eastern Cape, South Africa (EC)	34.72S/25.18E June 2015	17	17	1	0.024	16	10 413	0.235 (0.164)	0.264 (0.148)	0.11
KwaZulu-Natal, South Africa (KZN)	30.83S/30.48E July 2015	20	20	1	0.022	20	10 766	0.224 (0.225)	0.254 (0.149)	0.11
Mozambique (MOZ)	36.53S/34.50E August 2015	20	19	0.99	0.026	20	10 755	0.227 (0.160)	0.255 (0.148)	0.11

Seq that incorporates digestion with two enzymes for genome reduction and results in an increased number of reads per site (Ott *et al.*, 2017). The tGBS libraries were sequenced on a Life Technologies' Ion Proton System. Sequenced reads were analysed using a custom Perl script (available at https://github.com/orgs/schnablelab; accessed 20 July 2018), which assigned each read to a sample and removed barcode sequences. Seqclean (sourceforge.net/projects/seqclean) was then used to remove proton adaptor sequences and chimaeric reads harbouring internal restriction enzyme sites. Retained reads were subjected to quality trimming in two phases using the software Lucy2 (Li and Chou, 2004) in which bases with PHRED scores <15 (of 40) were removed. In the first phase, sequences were scanned at each end; whereas in the second phase, sequences were scanned using overlapping 10 bp windows. Quality trimmed sequence reads were aligned to the reference genome of Pacific bluefin tuna, *Thunnus orientalis* (Nakamura *et al.*, 2013), using GSNAP (Wu and Nacu, 2010). Sequence alignments were then scanned for polymorphisms. A SNP was called homozygous if at least 15 reads supported the genotype at the site and at least 90% of all reads covering that site shared the same nucleotide. A SNP was considered heterozygous if each of the two nucleotide variants were reported at least 10 times, and each allele was represented in >30% of the total reads.

SNP summary statistics and outlier analysis

Allele frequencies and observed (H_O) and expected (H_E) heterozygosities were estimated using ARLEQUIN 3.4.2.2 (Excoffier and Lischer, 2010). ARLEQUIN was also used to test for departures from expectations of Hardy–Weinberg Equilibrium (HWE). There are numerous challenges to detect SNP "outlier" loci that are putatively under selection (Narum and Hess, 2011). Even though the concept of balancing selection is well established, there remain methodological limitations for its identification in hitchhiking mapping (Hansen *et al.*, 2010). Therefore, we restricted our analysis to the detection of signals of directional selection using two conceptually different approaches. First, we used the FDIST2 outlier detection method (Beaumont and Nichols, 1996) implemented in LOSITAN (Antao et al., 2008). We used the "force mean F_{ST}" option to obtain the genome-wide mean (neutral) F_{ST} value, to reduce the rate of false-positive detection of outliers. We used a null model distribution of the mean

F_{ST} as a function of H_E under an island model of population subdivision using 50 000 simulations. Positive outlier SNPs were identified as those falling above the 99% confidence intervals of the null distribution. Second, the Bayesian approach implemented in BAYESCAN 2.1 (Foll and Gaggiotti, 2008) was performed using default Markov Chain Monte Carlo (MCMC) parameters. Following the suggestion of Foll and Gaggiotti, (2008) to minimize false positives, we used the "decisive" criterion under Jeffreys' scale of evidence (Jeffreys, 1961) to identify outliers. Loci were considered to be under directional selection if identified as outliers using both methods, as the detection of outliers through multiple methods increases the confidence that these loci are non-neutral. The results between global and pairwise sample tests were compared to confirm consistency of outlier designations. Functional significance of loci was investigated using BLAST in GenBank following Milano *et al.* (2014).

Geographical structuring of genetic variation

Genetic differentiation among samples was quantified by global and pairwise F_{ST} (Weir and Cockerham, 1984) with statistical significances evaluated in ARELQUIN with 10 000 permutations. Genetic structuring was also investigated using the Bayesian clustering method in STRUCTURE 2.3.4 (Pritchard *et al.*, 2000), which identifies the most probable number of genetically distinct groups (K) represented by the data and estimates assignment probabilities (Q) for each individual (specifically their genomic components) to these groups. The analysis can be run without prior information; however, incorporating prior information using the LOCPRIOR model allows the clustering algorithm to assume that the probability of assignment varies among samples thereby increasing power while not biasing results (Hubisz *et al.*, 2009). The analysis was performed with and without the LOCPRIOR model, in both cases assuming admixture. Simulations were run 10 times for each proposed value of K (1–5) to assess convergence. Each run had a burn-in of 100 000 MCMC samples followed by 1 000 000 MCMC repetitions. Models were assessed using ΔK (Evanno *et al.*, 2005). To complement the STRUCTURE analysis, classical assignment tests were performed in GENECLASS 2.0 (Piry *et al.*, 2004). The "detect migrants" function was used to identify first generation migrants born in a population other than the one in which they were sampled (Paetkau *et al.*, 2004; Piry *et al.*, 2004). The test

statistic L_h was used following Paetkau *et al.* (2004) and employing Nei's D_A (Nei *et al.*, 1983), which does not require conformance to HWE. The simulation method in GENECLASS (10 000 simulations) was also used to exclude individuals from a candidate population if that individual's probability of assignment to that population fell below a 0.05 threshold. Genetic structuring was also assessed using the discriminant analysis of principal components (DAPC) implemented in ADEGENET (Jombart *et al.*, 2010). Whereas STRUCTURE assigns cluster membership by minimizing Hardy–Weinberg and linkage disequilibria within clusters, DAPC employs fewer assumptions and simply maximizes differences between groups while minimizing differences within groups with the optimal number of groups identified using the Bayesian Information Criterion (BIC).

Results

MTDNA variation

Alignment of 633 bp sequences across 93 individuals revealed 137 variable sites that defined 91 haplotypes, of which 89 were found in a single individual. BLAST confirmed that all sequences were yellowfin tuna. Haplotype diversity was high ($h = 0.999$) across samples (Table 1). Global and pairwise tests of differentiation among samples yielded non-significant values (global $\Phi_{ST} = 0.001$; $p = 0.3$).

Nuclear variation

A total of 179 521 855 raw sequenced reads across 96 yellowfin tuna were obtained with between 56 871 and 10 442 486 raw reads (average = 1 870 018) per individual. Seven individuals from the WC sample with >60% of identified SNPs missing were removed from the dataset, resulting in a dataset of ~30 000 SNP loci for 89 individuals. The number of genotypes was further reduced by removing SNP loci that were missing in >10% of individuals; this resulted in a final SNP genotype dataset of 11 101 SNP loci for 89 fish with a mean of 10 495 (SD = 398) SNPs per individual. Levels of genetic variability were similar across samples, and all samples conformed to HWE (Table 1). Global genetic structuring was weak ($F_{ST} = 0.002$) but statistically significant ($p < 0.0001$). Pairwise tests revealed this significant structuring largely resulted from the divergent GG sample, which showed significant F_{ST} values with all other samples (Table 2). All other pairwise F_{ST} values were not significant, and excluding GG global structuring was not significant ($F_{ST} = -0.0002$).

The LOSITAN outlier analyses consistently reported a greater number of positive outlier loci than the BAYESCAN analysis in the various global and pairwise tests. For example, in the initial global outlier test LOSITAN identified 756 positive outliers compared with 11 by BAYESCAN. In all cases, outliers identified by BAYESCAN were also among the outliers identified by LOSITAN. Therefore, we restricted our discussion to the smaller number of outliers identified using BAYESCAN, as the consensus

outliers were ultimately defined by the BAYESCAN results. The results of outlier tests revealed a clear pattern that aligned with the differentiation of GG from the remaining samples. Firstly, no outliers were identified in global tests excluding the GG sample compared with the detection of 11 in global test including GG. Second, in tests between pairs of samples outliers were only identified in comparisons involving GG. All outliers identified in pairwise tests were among the 11 identified initially in the global test and certain outlier were common across different pairwise comparisons. BLAST analysis did not provide insight into potential functional associations of any of the outlier loci.

When outlier loci were excluded, the pattern of significant differentiation of the GG sample from all other (non-differentiated) samples was still evident in F_{ST} values (Table 2; Figure 2a) and PCA clustering (Figure 2b). Analysis of the 11 outlier loci reported considerably higher levels of differentiation of the GG sample while retaining the pattern of homogeneity among the remaining samples (Table 2; Figure 3a and b). Bayesian clustering analysis of outlier genotypes unanimously supported a model of $K = 2$, with identical results for the analyses with and without LOCPRIOR. Under $K = 2$, all non-GG fish strongly assigned to group B with high individual Q values, whereas the majority of GG fish had higher or intermediate (overlapping confidence
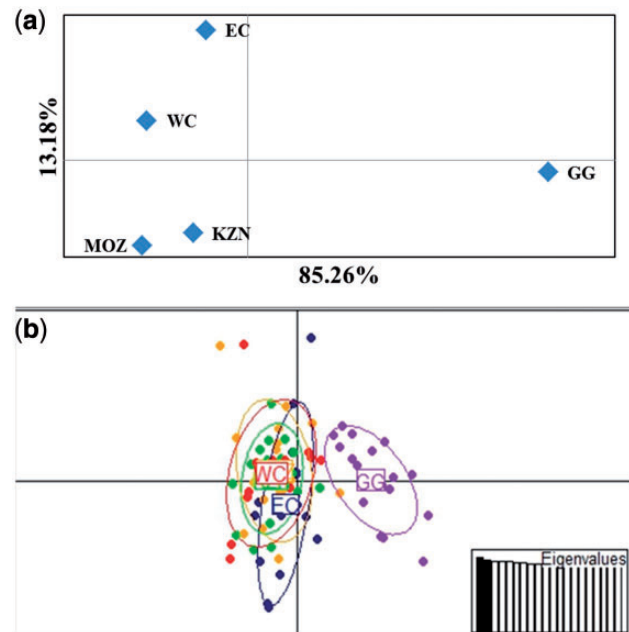
Figure 2. (a) Principal co-ordinate analysis of FST based on neutral loci. (b) Principal component analysis (PCA) clustering analyses of individuals based on neutral loci. For b ellipses summarize variation per site using ADEGENET's default "cellipse" settings.

Table 2. Pairwise F_{ST} values between samples with statistically significant values in bold.
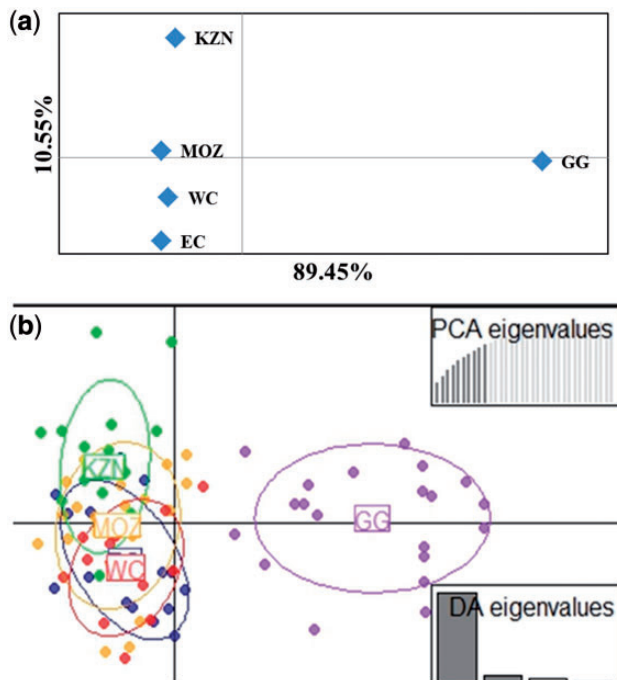
	GG	WC	EC	KZN	MOZ
Gulf of Guinea	–	**0.330**	**0.351**	**0.341**	**0.344**
Western Cape	**0.005/0.005**	–	−0.013	0.028	−0.011
Eastern Cape	**0.004/0.003**	−0.0016/−0.0017	–	0.056	−0.005
KwaZulu-Natal	**0.005/0.003**	−0.0008/−0.0001	−0.0004/−0.0006	–	0.005
Mozambique	**0.005/0.004**	0.0003/0.0002	0.0009/0.0007	0.0001/−0.0001	–

Below diagonal: Estimates across all SNP loci/neutral SNP loci. Above diagonal: estimates based on outlier loci only.
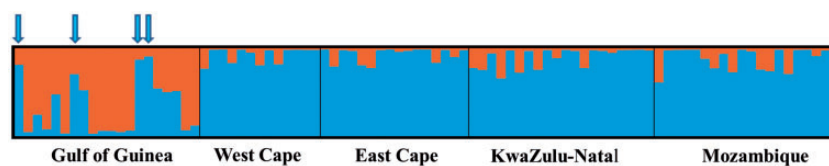
intervals for Q values) assignment probabilities of belonging to group A (Figure 4). Four individuals in the GG sample had larger Q membership values for group B. On the basis of the lack of differentiation among the SAF/Indian Ocean samples, the samples were pooled for the classical assignment tests between GG and the remaining samples. Assignment tests identified the same four GG individuals that had higher group B Q values as being first generation migrants from the pooled SAF/Indian Ocean group. No other migrants were detected with all individuals correctly assigned to the CG or Pooled group. Exclusion analysis using the strict 0.05 threshold further highlighted the genetic differentiation between both groups; first, of 16 non-migrants in the GG sample, 13 could be excluded from the Atlantic group, second, 42 of 69 non-migrants in the Atlantic/Indian ocean group could be excluded from GG.

Discussion

Mismatches between fishery management units and biological population processes of dispersal, interbreeding and adaptation are recognized as fundamental threats to fishery sustainability

Figure 3. (a) Principal co-ordinate analysis of FST based on outlier (non-neutral) loci. (b) Discriminant analysis of principal components (DAPC) describing variation among groups of individuals from different samples estimated from outlier loci. For b ellipses summarize variation per site using ADEGENET's default "cellipse" settings.

and to long-term resilience to climate change (Reiss *et al.*, 2009). In light of global declines in yellowfin tuna numbers and continued intensive harvesting, a primary objective of our study was to assess the biological validity of the current Atlantic Ocean and Indian Ocean operational stocks through genomic analysis of fish spanning the current stock boundary off the south coast of South Africa. The salient findings were (i) a lack of population structuring among the SAF samples, and (ii) genetic differentiation of the GG population from the SAF population. These patterns were evident in F_{ST} estimates from the entire SNP and neutral SNP data sets. A major hindrance to the integration of genetics into fishery management has been the inability of traditional F_{ST}-based measures to infer short-term demographics because these measures reflect connectivity over multiple generations (Whitlock and McCauley, 1999; Hellberg, 2009; Berry *et al.*, 2012). Additionally, divergence is expected to be slow for large populations experiencing little genetic drift. Increasingly, individual-based kinship and assignment tests are being used to directly assess dispersal (Christie *et al.*, 2010). Combining assignment tests with outlier loci can provide high discriminatory power even with low levels of background differentiation among populations (Helyar *et al.*, 2011; Bekkevold *et al.*, 2015). This approach revealed two genetic clusters in which all SAF specimens were assigned to one cluster with the other comprising only GG fish. Collectively the neutral and non-neutral genetic data concurrently resolved two genetic groups.

On the basis of our sampling scheme, it can be assumed that the two genetic groups (GG vs. the rest) correspond respectively to the distinct Atlantic and Indian ocean genetic stocks previously suggested by allozyme variation (Ward *et al.*, 1997) and confirmed by genomic studies (Pecoraro *et al.*, 2016; Barth *et al.*, 2017). The results therefore demonstrate that Western Cape yellowfin tuna, currently assigned to the Atlantic operational stock, are derived from an Indian Ocean genetic stock. Tagging data corroborate the movement of yellowfin tuna individuals from the western equatorial Indian Ocean to the southern Benguela region off western South Africa (Eveson *et al.*, 2015; Murua *et al.*, 2015). The concordant results among the present and previous genetic studies and the temporal sampling range represented point to a temporally stable pattern in which the Atlantic population is not contributing to recruitment in SAF waters. Nevertheless, it cannot be ruled out that there may be seasonal variation in stock mixing or overlap.

The identification of isolating mechanisms is a crucial facet for inferring demographic independence with a central topic in fisheries genetics being the relative roles of environmental forcing and behaviour (Heath *et al.*, 2008; McKeown *et al.*, 2017). The spatial genetic structure reported here and in Barth *et al.* (2017) is compatible with a role for the cool upwelled waters of the BCS as a potential barrier to dispersal of Atlantic yellowfin tuna into SAF waters. Cold water barriers have been reported for several marine

Figure 4. Bar graph illustrating average patterns of Bayesian clustering of individual under the optimum model of $K = 2$. Arrows above bar graph denote Q values from those individuals identified as migrants in the GENECLASS tests.

fishes (Gwilliam et al., 2018). Big eye tuna, *T. obesus*, is an interesting comparison because genetically distinct Atlantic and Indian Ocean fish mix in SAF waters. Yellowfin and bigeye tuna have similar geographical distributions but occupy different ecological niches and vertical habitats (Schaefer et al., 2009); bigeye tuna are more tolerant of cooler water temperatures than are yellowfin tuna (Brill and Lutcavage, 2001). Hence, the cold water of the BCS may have less influence on the movement of bigeye tuna from Atlantic Ocean waters north of the BCS into SAF waters, than on yellowfin tuna.

Barth et al. (2017) reported evidence of asymmetric Indo-Pacific-to-Atlantic gene flow across the BCS in yellowfin tuna, and Reid et al. (2016) described a similar pattern in the cosmopolitan bluefish (*Pomatous saltatrix*). In the present study, the Bayesian clustering analysis revealed patterns compatible with unidirectional gene flow from South Africa to GG. However, such gene flow is typically regarded as being historical and no studies have yet revealed recurrent dispersal of adults across the BCS in species for which the BCS is a barrier to contemporary gene flow. The multilocus identification of four Indian Ocean yellowfin tuna among GG fish is striking as it shows that Indian Ocean spawned yellowfin tuna are crossing the BCS and are mixing with members of the Atlantic genetic stock. These dispersals may be linked to Agulhas warm-core rings (Schouten et al., 2000) that can facilitate the transfer of larvae/adults across the BCS and/or to active swimming. In the latter case, unidirectional dispersals may be facilitated by distinct genetic/plastic thermal tolerances of Indian Ocean spawned fish. This suggests that a cryptic mixed-stock fishery exists north of the BCS. Temporal studies are needed to assess the consistency with which Indian Ocean migrants cross the BCS.

The maintenance of genetic integrity of the Atlantic Ocean stock in the face of potentially high rates of dispersal of Indian Ocean genotypes implicates additional mechanisms that may be isolating both stocks. The capacity for active dispersal of adults highlights the potential importance of natal homing (Gaggiotti et al., 2009; McKeown et al., 2017). Indian Ocean fish may migrate into Atlantic spawning areas but return to the Indian Ocean to spawn. If this is occurring samples collected during spawning times should contain fewer migrants than at other times. In this study the GG sample was collected outside of the spawning period and so comparative analysis of a spawning sample from this area could be highly informative regarding the role of homing. The two genetic stocks may also be isolated by adaptation (Nosil et al., 2009) with migrant offspring being selected against (Hendry, 2004). The increased level of genetic divergence between groups at outlier loci might be explained by directional selection, with the resolved groups corresponding to differentially adapted units (Milano et al., 2014). Alternatively, outlier loci may reflect endogenous forces stemming from pre- and/or post-zygotic incompatibilities between populations with different genetic backgrounds to produce intrinsic incompatibilities, rather than reflecting adaptive environmental selection through genotype-environment associations (Bierne et al., 2011). However, adaptation and intrinsic incompatibilities are not mutually exclusive. Along with natal homing, the identification of slowly changing isolating factors is important, as these factors may lead to resistance to hybridization, colonization of new habitats, and by extension, recovery of stocks (e.g. Svedäng et al., 2007).

Applications and management implications

The use of management units that include only a portion of a larger population can present problems with understanding population stock dynamics and environmental linkages (Frisk et al., 2008). Conversely, management units containing multiple biological populations can lead to inaccurate descriptions of population-specific abundances and productivity, and may conceal the declines of vulnerable populations (Kell et al., 2009; Ying et al., 2011). Our study reveals that the management of the yellowfin tuna fishery around southern Africa presents both types of problems. At a local level, it is recommended that the South Africa Department of Agriculture, Forestry and Fisheries report all future yellowfin tuna catches in SAF waters to the IOTC for inclusion in Indian Ocean stock assessments. Our results indicate that the SAF catch is derived entirely from Indian Ocean spawned fish. It may be more accurate to move the boundary between the Atlantic and Indian Ocean management stocks to the western extent of South Africa's EEZ at 13.35°E (Figure 1). At a regional level, it is essential to understand the proportions of individuals derived from the Atlantic and Indian Ocean genetic stocks over space and time in the mixed-stock fishery north of the BCS. The proportion of "consensus" outliers identified here (0.9%) was generally smaller than the proportions of outlier loci reported by other studies (5.2%, Bradbury et al., 2013; 4.5%, Milano et al., 2014; 3.65%, Hess et al., 2013; 0.99%, Guo et al., 2015). These outlier loci represent a resource that can be used to identify migrants and to characterize mixed-stock dynamics. At present the indiscriminate harvesting of both stocks in the region means that the Atlantic spawning stock biomass is overestimated.

Data accessibility

Data will be deposited on Dryad.

References

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. 2010. Genomics and the future of conservation genetics. Nature Reviews Genetics, 11: 697–709.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., and Luikart, G. 2008. LOSITAN: a workbench to detect molecular adaptation based on a F(st)-outlier method. BMC Bioinformatics, 9: 323.

Barth, J. M. I., Damerau, M., Matschiner, M., Jentoft, S., and Hanel, R. 2017. Genomic differentiation and demographic histories of Atlantic and Indo-Pacific Yellowfin Tuna (*Thunnus albacares*) populations. Genome Biology and Evolution, 9: 1084–1098.

Beaumont, M. A., and Nichols, R. A. 1996. Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society B-Biological Sciences, 263: 1619–1626.

Bekkevold, D., Helyar, S. J., Limborg, M. T., Nielsen, E. E., Hemmer-Hansen, J., Clausen, L. A. W., and Carvalho, G. R. 2015. Gene-associated markers can assign origin in a weakly structured fish, Atlantic herring. ICES Journal of Marine Science, 72: 1790–1801.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., and Maes, G. E. 2017. Harnessing the power of genomics to secure the future of seafood. Trends in Ecology & Evolution, 32: 665–680.

Berry, O., England, P., Marriott, R. J., Burridge, C. P., and Newman, S. J. 2012. Understanding age-specific dispersal in fishes through hydrodynamic modelling, genetic simulations and microsatellite DNA analysis. Molecular Ecology, 21: 2145–2159.

Bierne, N., Welch, J., Loire, E., Bonhomme, F., and David, P. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. Molecular Ecology, 20: 2044–2072.

Botsford, L. W., Castilla, J. C., and Peterson, C. H. 1997. The management of fisheries and marine ecosystems. Science, 277: 509–515.

Bradbury, I. R., Hubert, S., Higgins, B., Bowman, S., Borza, T., Paterson, I. G., Snelgrove, P. V. R., et al. 2013. Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. Evolutionary Applications, 6: 450–461.

Brill, R. W., and Lutcavage, M. E. 2001. Understanding environmental influences on movements and depth distributions of tunas and billfishes can significantly improve population assessments. American Fisheries Society Symposium, 25: 179–198.

Christie, M. R., Tissot, B. N., Albins, M. A., Beets, J. P., Jia, Y., Ortiz, D. M., Thompson, S. E., et al. 2010. Larval connectivity in an effective network of marine protected areas. PLoS ONE, 5: e15715.

Collette, B. B., Carpenter, K. E., Polidoro, B. A., Juan-Jorda, M. J., Boustany, A., Die, D. J., Elfes, C., et al. 2011. High value and long life-double jeopardy for tunas and billfishes. Science, 333: 291–292.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14: 2611–2620.

Eveson, J. P., Hobday, A. J., Hartog, J. R., Spillman, C. M., and Rough, K. M. 2015. Seasonal forecasting of tuna habitat in the Great Australian Bight. Fisheries Research, 170: 39–49.

Excoffier, L., and Lischer, H. E. L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10: 564–567.

FAO. 2017. Fishery Statistical Collections: Global Tuna Catches by Stock. Available at: http://www.fao.org/fishery/statistics/tuna-catches/en.

Foll, M., and Gaggiotti, O. 2008. A genome-Scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics, 180: 977–993.

Frisk, M. G., Miller, T. J., Martell, S. J. D., and Sosebee, K. 2008. New hypothesis helps explain elasmobranch "OutBurst" on Georges bank in the 1980s. Ecological Applications, 18: 234–245.

Gaggiotti, O. E., Bekkevold, D., Jorgensen, H. B. H., Foll, M., Carvalho, G. R., Andre, C., and Ruzzante, D. E. 2009. Disentangling the effects of evolutionary, demographic and environmental factors influencing genetic structure of natural populations: atlantic herring as a case study. Evolution, 63: 2939–2951.

Guo, B. C., DeFaveri, J., Sotelo, G., Nair, A., and Merila, J. 2015. Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. BMC Biology, 13: 19.

Gwilliam, M. P., Winkler, A. C., Potts, W. M., Santos, C. V., Sauer, W. H. H., Shaw, P. W., and McKeown, N. J. 2018. Integrated genetic and morphological data support eco-evolutionary divergence of Angolan and South African populations of *Diplodus hottentotus*. Journal of Fish Biology, 92: 1163–1176.

Hansen, M. M., Meier, K., and Mensberg, K. L. D. 2010. Identifying footprints of selection in stocked brown trout populations: a spatio-temporal approach. Molecular Ecology, 19: 1787–1800.

Heath, M. R., Kunzlik, P. A., Gallego, A., Holmes, S., and Wright, P. J. 2008. A model of meta-population dynamics for North Sea and West of Scotland cod – the dynamic consequences of natal fidelity. Fisheries Research, 93: 92–116.

Hellberg, M. E. 2009. Gene flow and isolation among populations of marine animals. In Annual Review of Ecology Evolution and Systematics, 40: 291–310.

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Cariani, A., et al. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources, 11: 123–136.

Hendry, A. P. 2004. Selection against migrants contributes to the rapid evolution of ecologically dependent reproductive isolation. Evolutionary Ecology Research, 6: 1219–1236.

Hess, J. E., Campbell, N. R., Close, D. A., Docker, M. F., and Narum, S. R. 2013. Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. Molecular Ecology, 22: 2898–2916.

Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources, 9: 1322–1332.

ICCAT. 2016. Report for the biennial period 2014–2015. Part II. 2. (SCRS). Madrid.

ICCAT. 2017. Report of the Standing Committee on Research and Statistics (SCRS). Madrid.

IOTC. 2014. Report of the 17th Session of the Scientific Committee. Victoria, Seychelles, 8–12 December 2014.

IOTC. 2016. Report of the 19th Session of the IOTC Scientific committee. REF: IOTC-2016-SC19-R.

Iles, T. D., and Sinclair, M. 1982. Atlantic herring – stock discreteness and abundance. Science, 215: 627–633.

Jackson, J. B. C., Kirby, M. X., Berger, W. H., Bjorndal, K. A., Botsford, L. W., Bourque, B. J., Bradbury, R. H., et al. 2001. Historical overfishing and the recent collapse of coastal ecosystems. Science, 293: 629–638.

Jeffreys, H. 1961. Theory of Probability. Oxford University Press, Oxford, UK.

Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics, 11: 94.

Juan-Jordá, M. J., Mosqueira, I., Freire, J., and Dulvy, N. K. 2013. The conservation and management of tunas and their relatives: setting life history research priorities. PLoS ONE, 8: e70405.

Kell, L. T., Dickey-Collas, M., Hintzen, N. T., Nash, R. D. M., Pilling, G. M., and Roel, B. A. 2009. Lumpers or splitters? Evaluating recovery and management plans for metapopulations of herring. ICES Journal of Marine Science, 66: 1776–1783.

Li, S., and Chou, H. H. 2004. Lucy 2: an interactive DNA sequence quality trimming and vector removal tool. Bioinformatics, 20: 2865–2866.

Majkowski, J. 2007. Global fishery resources of tuna and tuna-like species. FAO Fisheries Technical Paper No. 483. FAO, Rome.

McKeown, N. J., Arkhipkin, A. I., and Shaw, P. W. 2015. Integrating genetic and otolith microchemistry data to understand population structure in the Patagonian Hoki (*Macruronus magellanicus*). Fisheries Research, 164: 1–7.

McKeown, N. J., Arkhipkin, A. I., and Shaw, P. W. 2017. Regional genetic population structure and fine scale genetic cohesion in the Southern blue whiting *Micromesistius australis*. Fisheries Research, 185: 176–184.

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G. R., Espineira, M., et al. 2014. Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). Molecular Ecology, 23: 118–135.

Miyake, M., Guillotreau, P., Sun, C., and Ishimura, G. 2010. Recent developments in the tuna industry: stocks, fisheries, management, processing, trade and markets. FAO Fisheries and Aquaculture Technical Paper No. 543. FAO, Rome.

Murua, H., Eveson, J. P., and Marsac, F. 2015. The Indian Ocean tuna tagging programme: building better science for more sustainability Preface. Fisheries Research, 163: 1–6.

Myers, R. A., and Worm, B. 2005. Extinction, survival or recovery of large predatory fishes. Philosophical Transactions of the Royal Society B-Biological Sciences, 360: 13–20.

Nakamura, Y., Mori, K., Saitoh, K., Oshima, K., Mekuchi, M., Sugaya, T., Shigenobu, Y., et al. 2013. Evolutionary changes of multiple visual pigment genes in the complete genome of Pacific bluefin tuna. Proceedings of the National Academy of Sciences, USA, 110: 11061–11066.

Narum, S. R., and Hess, J. E. 2011. Comparison of F-ST outlier tests for SNP loci under selection. Molecular Ecology Resources, 11: 184–194.

Nei, M., Tajima, F., and Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. Journal of Molecular Evolution, 19: 153–170.

Nosil, P., Funk, D. J., and Ortiz-Barrientos, D. 2009. Divergent selection and heterogeneous genomic divergence. Molecular Ecology, 18: 375–402.

Ott, A., Liu, S. Z., Schnable, J. C., Yeh, C. T., Wang, K. S., and Schnable, P. S. 2017. tGBS (R) genotyping-by-sequencing enables reliable genotyping of heterozygous loci. Nucleic Acids Research, 45: e178.

Paetkau, D., Slade, R., Burden, M., and Estoup, A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology, 13: 55–65.

Pecoraro, C., Babbucci, M., Villamor, A., Franch, R., Papetti, C., Leroy, B., Ortega-Garcia, S., et al. 2016. Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (Thunnus albacares). Marine Genomics, 25: 43–48.

Pecoraro, C., Zudaire, I., Bodin, N., Murua, H., Taconet, P., Diaz-Jaimes, P., Cariani, A., et al. 2017. Putting all the pieces together: integrating current knowledge of the biology, ecology, fisheries status, stock structure and management of yellowfin tuna (Thunnus albacares). Reviews in Fish Biology and Fisheries, 27: 811–841.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., and Estoup, A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity, 95: 536–539.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155: 945–959.

Reid, K., Hoareau, T. B., Graves, J. E., Potts, W. M., dos Santos, S. M. R., Klopper, A. W., and Bloomer, P. 2016. Secondary contact and asymmetrical gene flow in a cosmopolitan marine fish across the Benguela upwelling zone. Heredity, 117: 307–315.

Reiss, H., Hoarau, G., Dickey-Collas, M., and Wolff, W. J. 2009. Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish and Fisheries, 10: 361–395.

Ruzzante, D. E., Mariani, S., Bekkevold, D., Andre, C., Mosegaard, H., Clausen, L. A. W., Dahlgren, T. G., et al. 2006. Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. Proceedings of the Royal Society B-Biological Sciences, 273: 1459–1464.

Ryman, N., Utter, F., and Laikre, L. 1995. Protection of intraspecific biodiversity of exploited fishes. Reviews in Fish Biology and Fisheries, 5: 417–446.

Schaefer, K. M., Fuller, D. W., and Block, B. A. 2009. Vertical movements and habitat utilization of Skipjack (Katsuwonus pelamis), Yellowfin (Thunnus albacares), and Bigeye (Thunnus obesus) tunas in the Equatorial Eastern Pacific Ocean, ascertained through archival tag data. In Reviews: Methods and Technologies in Fish Biology and Fisheries: Tagging and Tracking of Marine Animals with Electronic Devices, 9, pp. 121–144. Ed. by J. L. Nielsen, H. Arrizabalaga, H. Fragoso, A. Hobday, M. Lutcavage and J. Sibert. Springer, Berlin.

Schouten, M. W., de Ruijter, W. P. M., van Leeuwen, P. J., and Lutjeharms, J. R. E. 2000. Translation, decay and splitting of Agulhas rings in the southeastern Atlantic Ocean. Journal of Geophysical Research-Oceans, 105: 21913–21925.

Svedäng, H., Righton, D., and Jonsson, P. 2007. Migratory behaviour of Atlantic cod Gadus morhua: natal homing is the prime stock-separating mechanism. Marine Ecology Progress Series, 345: 1–12.

Waples, R. S., Punt, A. E., and Cope, J. M. 2008. Integrating genetic data into management of marine resources: how can we do it better? Fish and Fisheries, 9: 423–449.

Ward, R. D., Elliott, N. G., Innes, B. H., Smolenski, A. J., and Grewe, P. M. 1997. Global population structure of yellowfin tuna, Thunnus albacares, inferred from allozyme and mitochondrial DNA variation. Fishery Bulletin, 95: 566–575.

Weir, B. S., and Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38: 1358–1370.

Whitlock, M. C., and McCauley, D. E. 1999. Indirect measures of gene flow and migration: f-ST not equal 1/(4Nm+1). Heredity, 82: 117–125.

Willette, D. A., Allendorf, F. W., Barber, P. H., Barshis, D. J., Carpenter, K. E., Crandall, E. D., Cresko, W. A., et al. 2014. So, you want to use next-generation sequencing in marine systems? Insight from the Pan-Pacific Advanced Studies Institute. Bulletin of Marine Science, 90: 79–122.

Winnepenninckx, B., Backeljau, T., and De Wachter, R. 1993. Extraction of high-molecular-weight DNA from mollusks. Trends in Genetics, 9: 407.

Wu, T. D., and Nacu, S. 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics, 26: 873–881.

Ying, Y. P., Chen, Y., Lin, L. S., and Gao, T. X. 2011. Risks of ignoring fish population spatial structure in fisheries management. Canadian Journal of Fisheries and Aquatic Sciences, 68: 2101–2120.

Handling editor: W. Stewart Grant