SCIENTIFIC COMMITTEE
EIGHTEENTH REGULAR SESSION
ELECTRONIC MEETING
10 – 18 August 2022

No Population Structure of Bigeye Tunas \textit{(Thunnus obesus)} in the
Western and Central Pacific Ocean Indicated by Single Nucleotide Polymorphisms

WCPFC-SC18-2022/SA-IP-11

J. Natasha\textsuperscript{1}, B. L. Stockwell\textsuperscript{1}, A. D. Marie\textsuperscript{2,3}, J. Hampton\textsuperscript{4}, N. Smith\textsuperscript{4}, S. Nicol\textsuperscript{4} and C. Rico\textsuperscript{5}

\textsuperscript{1}Molecular Analytics Laboratory (MOANA-LAB), School of Agriculture, Geography, Environment, Ocean and Natural Sciences, The University of the South Pacific, Suva, Fiji
\textsuperscript{2}ESE, Ecology and Ecosystems Health, Institute Agro, INRAE, Rennes, France
\textsuperscript{3}Agence de Recherche Pour la Biodiversité à la Réunion, Saint-Gilles, France
\textsuperscript{4}Oceanic Fisheries Programme, Pacific Community (SPC), Nouméa, New Caledonia
\textsuperscript{5}Instituto de Ciencias Marinas de Andalucía (ICMAN), Consejo Superior de Investigaciones Científicas, Campus Universitario Río San Pedro, Puerto Real, Spain
No Population Structure of Bigeye Tunas (Thunnus obesus) in the Western and Central Pacific Ocean Indicated by Single Nucleotide Polymorphisms

Janice Natasha1*, Brian L. Stockwell1, Amandine D. Marie2,3, John Hampton4, Neville Smith4, Simon Nicol4 and Ciro Rico5*

1 Molecular Analytics Laboratory (MOANA-LAB), School of Agriculture, Geography, Environment, Ocean and Natural Sciences, The University of the South Pacific, Suva, Fiji; 2 ESE, Ecology and Ecosystems Health, Institute Agro, INRAE, Rennes, France; 3 Agence de Recherche Pour la Biodiversité à la Réunion, Saint-Gilles, France; 4 Oceanic Fisheries Programme, Pacific Community (SPC), Nouméa, New Caledonia; 5 Instituto de Ciencias Marinas de Andalucía (ICMAN), Consejo Superior de Investigaciones Científicas, Campus Universitario Río San Pedro, Puerto Real, Spain

The current management of the bigeye tuna (Thunnus obesus) fishery in the Western and Central Pacific Ocean (WCPO) is based on a “single-stock” paradigm. We investigated the spatial and temporal distribution of the neutral and potentially adaptive genetic diversity in samples collected from waters near Micronesia, Marshall, Phoenix and Gilbert Islands and French Polynesia using genotyping by sequence and single nucleotide polymorphisms (SNP). Analysis of 144 individuals and 2,682 SNP did not provide any substantial evidence to challenge the current assumption of panmixia in the WCPO using putatively neutral loci. However, there was a signal of potentially adaptive divergence (F_{ST} range 0.11–0.24) of samples collected in 2013 in the exclusive economic zone of the Phoenix Islands, possibly because of exposure to differing selective constraints. The geographic origin of these tuna remains to be determined, as the samples originate from tagging cruises that sample large schools of tuna, and some of those tuna have subsequently showed substantial movement. We highly recommend further studies involving large sample sizes of larvae and young-of-the-year across the region and over multiple years to uncover the underlying processes that are potentially driving adaptive divergence and to infer any biological significance it may have on conservation and management of tropical bigeye tuna fishery stock.

Keywords: bigeye tuna (Thunnus obesus), population genetic structure, panmixia, Western and Central Pacific Ocean, adaptive divergence

INTRODUCTION

The current management strategy for the Western and Central Pacific Ocean (WCPO) bigeye tuna (Thunnus obesus, Lowe 1839) fishery is based on a “single-stock” paradigm. However, the recent discovery of population structure using genomic analyses of yellowfin tuna (T. albacares) within the WCPO (Grewe et al., 2015; Pecoraro et al., 2018; Anderson et al., 2019a), the Atlantic
Ocean (Pecoraro et al., 2018) and waters around South Africa (Mullins et al., 2018) suggest that investigations to understand the distribution in time and space of the genetic diversity in other fish with different life history traits is warranted. Factors such as larval behavior and dispersal due to ocean currents (Levin, 2006), oceanographic fronts (Galarza et al., 2009), retention of pelagic eggs and larvae by eddies and gyres in the ocean (Myksvoll et al., 2014), poor adjustment of larvae to their place of destination leading to isolation by adaptation (Nosil et al., 2009), geographic distance (Slatkin, 1993), seascape (Riginos et al., 2016), environmental variables (Riccioni et al., 2013; Benestan et al., 2016) and habitat geomorphology (Lal et al., 2017) could potentially limit gene flow in marine environments and lead to population structure. Of particular interest, is the use of loci under potential selection to assess the geographic and temporal distribution of adaptive genetic diversity to uncover levels of population structure and local adaptation in previously presumed panmictic or nearly genetically homogenous species (Gagnaire et al., 2015). For example, the evidence of within-ocean adaptive genetic divergence in yellowfin (Grewe et al., 2015; Mullins et al., 2018; Pecoraro et al., 2018) and albacore (T. alalunga) (Anderson et al., 2019b) underpins the need to understand the population structure over evolutionary and ecological scales for improved conservation and management of commercially important tunas.

The Pacific bigeye tuna is mainly abundant in the central and eastern Pacific Ocean (Leroy et al., 2015). Pacific bigeye tuna populations were previously considered to be over-exploited (Harley et al., 2015) and listed in the IUCN Red List of Threatened Species as vulnerable (Collette et al., 2011); however, more recent stock assessments based on a range of new information, including a better understanding of age and growth, report that the probability that the WCPO stock is currently in an over-exploited condition is low (McKechnie et al., 2017; Ducharme-Barth et al., 2020). At present, the management of bigeye tuna in the Pacific Ocean is conducted by the relevant Regional Management Fisheries Organizations (RMFO), as two separate stocks. In the WCPO, the fishery is managed by the Western and Central Pacific Fisheries Commission (WCPFC), while in the Eastern Pacific Ocean (EPO), it is under the administrative responsibility of the Inter-American Tropical Tuna Commission (IATTC) (Kumar and Kocour, 2015). The split into two stocks has no biological basis and is purely a matter of governance. Extensive tagging data suggests three putative stock boundaries in the equatorial Pacific Ocean; western, central and eastern with higher levels of abundance and fidelity in eastern and western regions and six additional putative stocks in areas beyond 10°N and three in areas beyond 10°S (Schaefer et al., 2015; Moore et al., 2020a). The data also suggest that the equatorial Central Pacific region may behave as a transition zone between these stocks resulting in a “fluid” rather than hard delineation of stock boundaries (Schaefer et al., 2015; Moore et al., 2020a).

The current understanding of population genetic structure for the Pacific bigeye tuna is based on scant and outdated genetic data. Results of previous studies on population genetics of bigeye tunas using Tg blood group system (Suzuki et al., 1962), serum esterase (Fujino and Kang, 1968), PCR-RFLP of mtDNA and microsatellite generated distances (Grewe and Hampton, 1998) and polymorphism of sequence variations in mitochondrial Cyt b genes (Wu et al., 2014) are consistent with panmixia or indicate very weak population differentiation in the Pacific Ocean. A study of global genetic structure of bigeye tuna using mtDNA haplotype diversity indicated that samples from the Atlantic were genetically distinct from those found in Indian and Pacific Oceans (Bremer et al., 1998). This finding was also confirmed by Durand et al. (2005) and Gonzalez et al. (2008) who reported highly significant mtDNA haplotype frequency differences between Atlantic and Indian or Pacific Ocean samples. However, no significant differences were found among fish captured within oceans, or between Indo-Pacific stocks.

Genomic approaches have been introduced to uncover the distribution of the neutral and potentially adaptive genetic diversity of marine organisms using Next Generation Sequencing (NGS) for underlying environmentally constrained variations that might be under selective pressure (Allendorf et al., 2008). The so called “outlier loci” can unravel genetic differentiation patterns where neutral markers often remain uninformative. In addition, large datasets of neutral markers such as those obtained using genomic approaches can be more informative than small datasets for high gene flow species (Gagnaire et al., 2015). Outlier loci have been used to identify locally adapted populations and redefine conservation and management units in fish such as those depicted in FishPopTrace, a Pan-European project using state-of-the-science technology for sustainable fisheries (Nielsen et al., 2012). This approach is gaining popularity due to the efficiency of selection in opposing the homogenizing effect of migration in comparison to drift (Gagnaire et al., 2015).

In this study, we investigated the genetic structure and the possibility of adaptive divergence in bigeye tunas among five geographically distinct sampling areas across the WCPO and over a span of 3 years and including replicated temporal samples from the Marshall Islands. Using Single Nucleotide Polymorphisms (SNPs), we determined if the temporal and spatial distribution of the neutral and potentially adaptive genetic diversity supports the widely accepted paradigm of panmixia within ocean basins in bigeye tuna.

MATERIALS AND METHODS

All methods and experimental protocols were approved by The University of the South Pacific’s (USP) Faculty of Science, Technology and Environment (FSTE) Research Committee and were carried out in accordance with their guidelines and regulations.

Sample Collection

Tissue samples were collected from muscle plug biopsies of bigeye tuna from five geographical locations across the exclusive economic zones (EEZs) of four island nations and territories, namely, Phoenix and Gilbert Islands of Kiribati, Republic of the Marshall Islands, French Polynesia, and Federated States of Micronesia in the WCPO. Samples were collected between 2011 and 2014 by observers on fishing vessels and by scientists.
Supplementary Table 1

We also selected two temporal samples collected in 2011 and three samples within 6 weeks of each other at different locations. Captured at different sites, months or years apart, by collecting tissue samples archived in the Tissue Bank managed under the auspices of the Western and Central Fisheries Commission (WCPFC), in Noumea, New Caledonia. The samples were taken during research tagging cruises (Figure 1 and Supplementary Table 1). The fish specimens were obtained for analysis from tissue samples archived in the Tissue Bank managed under the auspices of the Western and Central Fisheries Commission (WCPFC), in Noumea, New Caledonia. The samples were taken from an overall area spanning about 60 degrees of longitude and 20 degrees of latitude.

Given the migratory nature of tunas we aimed to minimize the chances of analyzing individuals of the same population captured at different sites, months or years apart, by collecting three samples within 6 weeks of each other at different locations. We also selected two temporal samples collected in 2011 and 2014 in roughly the same area and season. Subsampling from the muscle plug biopsies was performed using sterilized surgical procedures to avoid cross contamination of tissue between samples. A summary of sample locations, numbers, average fork length and year caught is provided in Supplementary Table 1.

DArTseq Genotyping and Single Nucleotide Polymorphisms Quality Control

DNA extraction and sequencing was conducted by Diversity Arrays Technology PL (DArT). The genotyping protocols used for the current study are described in Anderson et al. (2019a,b). The returned dataset was further filtered for locus quality by removing all but one SNP per sequenced DNA fragment, any duplicate SNPs possessing identical Clone IDs and SNPs with a call rate below 95%; a minimum read depth of 7 and minor allele frequency (MAF) below 5%. Hardy Weinberg Equilibrium (HWE) for each locus in the primary dataset was assessed across all populations using the Arlequin v.3.5.1.3 software package using locus by locus test type, 10,000,000 steps in Markov chain and 100,000 dememorization steps (Excoffier et al., 2005). No single locus was found to deviate from HWE ($P < 0.0001$) across all populations and thus, no loci were removed after this filtering step. The raw and filtered dataset in Genepop file format analyzed for this study can be found in the Open Science Framework (OSF) repository (DOI: 10.17605/OSF.IO/6ZRA3).

Outlier Detection

The polymorphic SNPs of all sample collections were screened together to detect outlier loci using three different methods. The first method involved the use of fixation index ($F_{ST}$) outlier approach developed by Beaumont and Nichols (1996) and implemented in LOSITAN (Antao et al., 2008). The Force mean $F_{ST}$ option was selected to approximate desired mean $F_{ST}$ by running a bisection algorithm over repeated 1,000,000 simulations, $P$-value of 0.95 and False Discovery Rate (FDR) of 10%. The detailed method of identification of outliers using this approach is outlined in Grewe et al. (2015). The second method involved the use of the Bayesian simulation-based test of Beaumont and Balding (2004) that has been further refined and implemented in the software BayeScan 2.1 (Foll, 2012) with default settings (burn-in = 50,000, sample size = 5,000, total iterations = 100,000, pilot runs = 20 × 5,000) and a FDR of 10%. The third method involved the use of software HacDivSel which employs the Extreme Outlier Set (EOS) test using a two-step heuristic algorithm to make pairwise population comparisons (Carvajal-Rodríguez, 2017). HacDivSel made 15 pairwise comparisons and generated outlier loci for each paired comparison.

Outlier SNPs generated from LOSITAN were used for the study. These outlier SNPs were removed from the primary dataset to generate a selectively neutral loci dataset (NL). Only candidates for directional selection (known as Loci Under Potential Selection, LUPS, hereafter) were retained for the study.

Population Genetics Structure: Neutral vs. Potentially Adaptive Divergence

Population structure analyses were carried out in parallel on putatively neutral and outlier loci panels to compare different patterns of population. Linkage disequilibrium (LD) for each pair of SNPs in selectively neutral dataset was tested in all populations using the Arlequin v.3.5.1.3 software package all populations and thus, no loci were removed after this filtering step. The raw and filtered dataset in Genepop file format analyzed for this study can be found in the Open Science Framework (OSF) repository (DOI: 10.17605/OSF.IO/6ZRA3).
1.01.7) (Wang, 2011). The filtered genotypes were run initially to estimate the degree of relatedness. The cut-off r-values for categorizing relatedness coefficients into full-sib (FS) and half-sib (HS) were determined by simulating 200 dyads in each group based on the allele frequencies generated from the initial run. The r cut-off values were used to identify the FS and HS in the empirical dataset. Individuals with the highest multilocus heterozygosity from each of the FS and HS pairs were excluded as a signal of potential contamination (Petrou et al., 2019) from the primary dataset. Since Petrou et al. (2019) found that contaminated samples of herring tend to have higher multilocus individual heterozygosity (H_I > 0.4; median = 0.45), which in turn inflated the population structure, bigeye individuals with exceptionally high multilocus heterozygosity (H_I > 0.4), regardless of its association with kin groups, were also removed to reduce potential biasness in estimating population structure. The final primary dataset was filtered for a second time using the same aforementioned SNP quality filtering steps and outlier loci detection but without potential bias introduced by potentially cross-contaminated individuals.

Filtered neutral loci datasets, obtained from the second round of dataset filtering, were subjected to descriptive population genetics analyses including observed, expected, and sample size-corrected multilocus heterozygosity averaged across all loci in Genetix v.4.05.2 (Belkhir et al., 2004). We tested deviations from HWE of the observed value of FIS using 1,000 permutations in Genetix v.4.05.2 (Belkhir et al., 2004). Through this procedure, FIS estimates from the real data are then compared with the distribution obtained by permutation of alleles within populations making it possible to obtain an estimate of the probability of obtaining a value > or = to the value observed under the null hypothesis (i.e., panmixia). Pairwise FST values were calculated for both neutral and outlier datasets at a significance level of 0.05, performed across 10,000 permutations to test for genetic differentiation among populations using the Arlequin v.3.5.1.3 software package (Excoffier et al., 2005). The sequential Bonferroni method (Rice, 1989) for multiple comparisons was applied to correct the significance level. Discriminant analysis of principle components (DAPC) (Jombart et al., 2010), a multivariate approach, was conducted on both datasets (NL and LUPS) and tested for alpha optimization (maximized inter-population variation given a minimum number of principle components). The main genetic clusters (k) were inferred using two methods. First, a k-means clustering was employed using the number of principle components retained by implementing the Bayesian Information Criterion (BIC) method with the use of adegenet function “find.clusters” using the Adegenet package v2.0.1 in R-studio (Jombart and Collins, 2015). Secondly, a model-based estimation of ancestry in unrelated individuals implemented in ADMIXTURE v. 1.3.0 software package was used (Alexander et al., 2009) which adopts the likelihood model embedded in structure and compares the cross-validation error (CV) to recommend k. Successive k-values from one to six were run on ADMIXTURE and the lowest associated CV error was selected as the best k. The k-value obtained was used for Monte Carlo assignment tests using functions “assign.MC” and “assign.kfold” performed in assignPOP (Chen et al., 2018) using a support vector machine (svm) predictive model. Since population genetic analyses suggest that temporal samples of Marshall Islands were genetically similar, they were pooled together into a single population for migration analysis. Gene flow among populations was estimated by calculating the directional relative migration using the neutral dataset on the web-based software application divMigrate-online (Sundqvist et al., 2016) based on the Nm (effective number of migrants) statistic (Alcala et al., 2014) and with alpha value of 0.05. Plots with only statistically significant asymmetric links between population pairs were generated by the software application.

RESULTS

Single Nucleotide Polymorphisms Data Validation

The raw dataset contained 39,661 SNPs genotyped for 175 individuals across the six population samples of bigeye tuna. Upon strict quality control filtering, the final dataset contained 144 individuals with 2,682 putatively neutral loci across the sample collections (Table 1). A small number of individuals with high average observed heterozygosity were found to be kin with other individuals by TrioML algorithm for genetic relatedness (Supplementary Table 3). With no means to validate the exceptionally high heterozygosity in individuals of interest and the potential for the presence of kin groups to misrepresent population structure (Selwyn et al., 2016) or indicate contamination or duplication of DNA samples (Flickinger et al., 2015), all three individuals with H_o > 0.4 (Supplementary Table 2) and 7 other individuals with the highest multilocus heterozygosity and appearing more than once in the potential sibling-pairs were removed from the dataset. A total of 10 individuals were removed.

Descriptive Genetic Diversity

The average number of alleles ranged from 1.97 to 1.99 in individuals caught from the five geographic locations. The observed (H_o) and expected (H_e) multilocus heterozygosity

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Number of neutral loci remaining after each quality control filtering step.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filtering step</td>
<td>Number of neutral loci remaining</td>
</tr>
<tr>
<td>Initial</td>
<td>39,661</td>
</tr>
<tr>
<td>Duplicates on contig</td>
<td>23,445</td>
</tr>
<tr>
<td>Call rate (95%)</td>
<td>14,619</td>
</tr>
<tr>
<td>Read depth</td>
<td>10,555</td>
</tr>
<tr>
<td>MAF (5%)</td>
<td>3,274</td>
</tr>
<tr>
<td>LD (70%)</td>
<td>3,262</td>
</tr>
<tr>
<td>FST Outliers from 1st run on LOSITAN (FDR = 10%)</td>
<td>2,798</td>
</tr>
<tr>
<td>FST Outliers from 2nd run on LOSITAN (FDR = 10%)</td>
<td>2,740</td>
</tr>
<tr>
<td>Loci with P-values between 0.05 and 0.95</td>
<td>2,682</td>
</tr>
<tr>
<td>HWE (P &lt; 0.0001)</td>
<td>2,682</td>
</tr>
</tbody>
</table>
TABLE 2 | Bigeye tuna sampling information and genetic diversity indices.

<table>
<thead>
<tr>
<th>Site</th>
<th>Abbreviation and year sampled</th>
<th>$F_{IS}$</th>
<th>$H_o$ (SD)</th>
<th>$H_{exp}$ (SD)</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phoenix Island (Kiribati)</td>
<td>PX (2013)</td>
<td>-0.004</td>
<td>0.27 (0.17)</td>
<td>0.27 (0.14)</td>
<td>1.98</td>
</tr>
<tr>
<td>Federated States of Micronesia</td>
<td>FM (2011)</td>
<td>0.1</td>
<td>0.24 (0.16)</td>
<td>0.25 (0.15)</td>
<td>1.97</td>
</tr>
<tr>
<td>Gilbert Islands (Kiribati)</td>
<td>GL (2014)</td>
<td>0.035</td>
<td>0.26 (0.17)</td>
<td>0.26 (0.15)</td>
<td>1.98</td>
</tr>
<tr>
<td>Marshal Islands</td>
<td>MH (2011)</td>
<td>0.122</td>
<td>0.23 (0.15)</td>
<td>0.25 (0.15)</td>
<td>1.98</td>
</tr>
<tr>
<td>Marshal Islands</td>
<td>MH (2014)</td>
<td>0.066</td>
<td>0.25 (0.16)</td>
<td>0.26 (0.14)</td>
<td>1.99</td>
</tr>
<tr>
<td>French Polynesia</td>
<td>PF (2013)</td>
<td>0.109</td>
<td>0.23 (0.15)</td>
<td>0.25 (0.15)</td>
<td>1.98</td>
</tr>
</tbody>
</table>

* $F_{IS}$ inbreeding coefficient; $H_o$, observed heterozygosity; $H_{exp}$ expected heterozygosity; NA average number of alleles; SD standard deviation.

TABLE 3 | Pairwise $F_{ST}$ values (above diagonal) and associated $P$-values (below diagonal) between geographical population samples based on neutral loci (NL).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PX (2013)</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>FM (2011)</td>
<td>0.050</td>
<td>0.001</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>GL (2014)</td>
<td>0.001</td>
<td>0.297</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>MH (2011)</td>
<td>0.153</td>
<td>0.378</td>
<td>0.123</td>
<td>0.088</td>
<td>0.088</td>
<td>0.171</td>
</tr>
<tr>
<td>MH (2014)</td>
<td>0.009</td>
<td>0.229</td>
<td>0.081</td>
<td>0.080</td>
<td>0.171</td>
<td></td>
</tr>
<tr>
<td>PF (2013)</td>
<td>0.006</td>
<td>0.489</td>
<td>0.081</td>
<td>0.080</td>
<td>0.171</td>
<td></td>
</tr>
</tbody>
</table>

**Bold indicates pairwise comparisons are statistically significant ($P < 0.007$; with Bonferroni Correction).**

ranged from 0.23 to 0.27 and from 0.25 to 0.26, respectively. While the Phoenix Island (2013) sample showed a heterozygote excess, the $F_{IS}$ value is not significantly different from zero. In contrast, inbreeding coefficients ($F_{IS}$) for all other samples were significantly different from zero and ranged from 0.035 to 0.122 (Table 2).

**Population Structure, Probability of Membership and Migration Patterns Based on Neutral Loci**

Pairwise fixation index ($F_{ST}$) provided low estimates of genetic differentiation between the six sample collections ($F_{ST}$ ranged from 0.001 to 0.004). Significant but marginal genetic differentiation ($P < 0.007$ after Bonferroni Correction, Table 3) was detected between samples collected from the Phoenix (2013) and Gilbert (2014) Islands in the EEZ of Kiribati ($F_{ST} = 0.003$, $P = 0.001$), Phoenix Island (2013) and French Polynesia (2013) ($F_{ST} = 0.003$, $P = 0.006$) and between the Gilbert (2014) and Marshall Islands (2011) ($F_{ST} = 0.004$, $P = 0.002$).

The analysis of individual clustering using the Bayesian Information Criterion (BIC) method and a comparison of cross-validation error values with successive values of $k$ on ADMIXTURE (Table 4) detected the presence of one putative genetic cluster ($k = 1$) of bigeye tuna within the WCPO as also shown by the Discriminant Analysis of Principal Components (DAPC) ($\alpha$-score optimized to retain 48 principal components (Figure 2A). Likewise, bigeye individuals from the six sample collections were poorly allocated as Monte-Carlo assignment accuracy ranged between 20 and 25% only (results not shown). Migration patterns based on neutral loci (Figure 3) showed high significant unidirectional gene flow in collections from same year (Figures 3A–C) and asymmetrical gene flow of bigeye individuals between the sample collection sites with migration coefficients ranging from 0.56 to 1 (Figure 3D).

**Outlier Detection and Population Structure Analysis Based on Potentially Adaptive Loci**

LOSITAN generated 56 outlier loci out of which 26 loci were candidates for directional selection (LUPS). BayeScan generated a panel of 11 LUPS under directional selection, which were also
FIGURE 2 | Discriminant analysis of principal components (DAPC) depicts the extent to which bigeye tuna individuals overlap between population clusters using (A) neutral and (B) outlier SNP loci.

identified by LOSITAN. The 11 loci detected by both methods are the same. HacDivSel generated 492 LUPS from the 15 pairwise sample comparisons made resulting in varying numbers of LUPS for each comparison. However, none of the LUPS identified by HacDivSel were common in all the paired comparisons.

The pairwise $F_{ST}$ values based on the 26-LUPS panel from LOSITAN ranged from 0.113 to 0.237, with significant population differentiation between PX (2013) and rest of population samples, PF (2013) and the rest and between FM (2011) and MH (2014) (Table 5). The pairwise $F_{ST}$ values based on LUPS from HacDivSel ranged from 0.098 to 0.171 and 11-LUPS panel from BayeScan ranged from 0.173 to 0.349 (Supplementary Tables 3, 4). The LUPS panel from HacDivSel showed significant population differentiation between all population samples except between FM (2011) and GL (2014) while the LUPS panel from BayeScan showed significant population differentiation between PX (2013) and the rest. The three algorithms generated varying numbers of LUPS and magnitude of differentiation but all indicated population structure, particularly between the Phoenix Islands sample and the rest. The analysis of individual clustering using the Bayesian Information Criterion (BIC) method and a comparison of cross-validation error values with successive values of $k$ on ADMIXTURE detected two putative genetic clusters ($k = 2$), suggesting that indeed, the Phoenix Islands are different from the rest as ADMIXTURE seems to indicate a gradient between these populations and the other ones, although not a strict separation (Figures 2B, 4). However, scatterplots generated by the Discriminant Analysis of the Principal Components ($\alpha$-score optimized to retain 31 principal components) did not show any distinct cluster, rather an overlapping of the sample collections (Figures 2B, 4). Likewise, the ADMIXTURE graph also fails to show any distinct geographically non-admixed population. Similarly, the Monte-Carlo assignment accuracy for the six sample collections ranged from zero to 50% (results not shown) despite the large and significant $F_{ST}$ values between Phoenix Island and other sample collections.

DISCUSSION

This study examined broad scale population genetic structure in bigeye tunas across the WCPO. It is also the most geographically comprehensive population genetic analysis of bigeye tunas carried out to date in the region with a novel panel of neutral SNP markers ($N = 2,682$) using 144 fish collected across five locations. Most population genetic studies of tunas and other oceanodromous species include samples from different locations within ocean basins collected in different times of the year and often in different years (e.g., Grewe et al., 2015).

The significant inbreeding coefficients ($F_{IS}$) observed in all population samples with the exception of the Phoenix Island can be due to genotyping artifacts, selection, non-random mating, inbreeding, small effective population sizes, migration and admixture. Fisheries overexploitation has also been associated with heterozygous deficits. For example, Hoarau et al. (2005) found that populations of plaice ($Pleuronectes platessa$) examined between 1924 and 1960 were in HWE, whereas populations examined after a fisheries peak in the 1970’s were not. The authors were able to attribute the significant heterozygote deficiencies found from 1970 onward to inbreeding and overexploitation. Unfortunately, we do not have access to archived samples to verify if the analyzed tuna populations were in HWE before the major expansion of tuna fisheries after WWII. Genotyping artifacts are unlikely to be a major cause because the quality filtering steps used in this study are quite stringent and no
FIGURE 3 | Directional relative migration network for sampled bigeye tuna. Directional relative migration estimated by divMigrate-online using Nm migration statistic. Arrows indicate the direction of gene flow and numbers show relative migration coefficient. Arrows with higher numbers appear thicker and stronger in color. Spatial relative migration estimated between (A) FM and MH in 2011 (B) PX and PF in 2013. (C) GL and MH in 2014. (D) Illustrates the calculated migration values, with no filter threshold, for all populations sampled. Temporal samples from Marshall Islands were pooled into a single population to estimate migration network for all populations.

TABLE 5 | Pairwise $F_{ST}$ values (above diagonal) and associated $P$-values (below diagonal) between geographical population samples based on a panel of 26 loci under potential selection (LUPS) obtained from LOSITAN.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PX (2013)</td>
<td>0.190</td>
<td>0.223</td>
<td>0.113</td>
<td>0.155</td>
<td>0.237</td>
<td></td>
</tr>
<tr>
<td>FM (2011)</td>
<td>0.000</td>
<td>0.029</td>
<td>0.021</td>
<td>0.041</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>GL (2014)</td>
<td>0.000</td>
<td>0.030</td>
<td>0.044</td>
<td>0.039</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>MH (2011)</td>
<td>0.001</td>
<td>0.100</td>
<td>0.117</td>
<td>–0.001</td>
<td>0.057</td>
<td>0.043</td>
</tr>
<tr>
<td>MH (2014)</td>
<td>0.000</td>
<td>0.006</td>
<td>0.010</td>
<td>0.537</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>PF (2013)</td>
<td>0.000</td>
<td>0.002</td>
<td>0.004</td>
<td>0.001</td>
<td>0.003</td>
<td></td>
</tr>
</tbody>
</table>

Bold indicates pairwise comparisons are statistically significant ($P < 0.007$; with Bonferroni).

single locus deviated from HWE in all populations (e.g., Kilian et al., 2012; Lal et al., 2017). Selection is unlikely to explain these results because the loci identified as potentially under selection per dataset were removed before the estimation of population fixation indices. Regardless of whether observed heterozygosity may be a consequence of inbreeding, genetic drift, migration, or admixture of genetically differentiated samples, all mechanisms are associated with the presence of population structure. Non-random mating and inbreeding possibly occur, supported by evidence of the presence of kin aggregations. However, these individuals with very high heterozygosity were removed before the estimation of $F_{IS}$ and it is more likely that they are the result of cross contamination. In addition, the presence of kin aggregations in a population sample is predicted
to have a stronger effect on $F_{ST}$ than on $F_{IS}$ (Yearsley et al., 2013), contrary to what we observe. Therefore, admixture is the most likely explanation for the positive inbreeding coefficient observed (Wahlund, 1928). DAPC scatter and ADMIXTURE bar plots, especially for LUPS show that some individuals within sampled locations have a mixed multilocus genotype, suggesting some genetic sub-structuring in fish sampled at individual sites (Figures 2, 4).

Assessments of the population structure at putatively neutral loci identified very small levels of structure to infer any biological significance and they are suggestive of panmixia. The low, yet significant, $F_{ST}$ values calculated are characteristic of highly mobile species and similar to the ranges reported for the Atlantic bluefin (Puncher et al., 2018), South Pacific albacore (Anderson et al., 2019b) and yellowfin tunas (Mullins et al., 2018; Pecoraro et al., 2018). The large population sizes, high potential for gene flow and therefore, a negligible effect of genetic drift explain the minuscule population heterogeneity observed at neutral loci (Palumbi, 1994). Tagging of bigeye tuna in the region (Schafer et al., 2015) shows extensive mixing of individuals, suggesting genetic exchange between the collection sites and corroborates well with our analyses on population structure [pairwise fixation indices ($F_{ST}$), DAPC and assignPOP]. Because our results also suggest extensive gene flow observed at neutral loci in the WCPO region, there is no evidence to disagree with results of previous studies on bigeye tuna using PCR-RFLP (Bremer et al., 1998; Chow et al., 2000) and mitochondrial DNA (Durand et al., 2005; Chiang et al., 2006) that are consistent with panmixia in the Pacific Ocean. Accordingly, we suggest no need to revise the current management policy.

However, in contrast to the panmictic structure suggested by the neutral loci, all panels of LUPS suggest that bigeye tuna, especially those collected from the EEZ of Phoenix Islands, may be subject to complex physical, ecological and evolutionary processes that are signaling an elevated level of differentiation and must be further researched. Furthermore, the comparisons between samples collected in Marshall (2011 and 2014) and Phoenix Islands indicate that the potential adaptive divergence observed may be stochastic due to population dynamics, such as rate of reproduction and spawning dynamics, larval survival and recruitment, adult mortality, and migration, which change between years. Temporal genetic variability has also been reported among bluefin tuna populations in the western and eastern Atlantic Ocean (Puncher et al., 2018), which argues for the necessity of multi-year sampling for investigating population structure of migratory marine fishes. Since at least 80 SNPs are recommended to detect low levels of differentiation among highly mobile species (Ryman et al., 2006; Hauser et al., 2011), this preliminary result must be interpreted with caution as the sample size and the number of LUPS are both small.

The stock structure of tunas may be more complex than previously thought. There are uncertainties in spawning dynamics, the degree of spawning site fidelity and localized residency, provenance of individuals, spatial variation in length-to-age, linkages with adjacent "stocks" and effects of climate change on these processes (Moore et al., 2020a). There is also recent evidence of population differentiation found using LUPS in yellowfin (Grewe et al., 2015) and South Pacific albacore tunas (Anderson et al., 2019b), and the first genetic evidence of natal homing in Atlantic bluefin tuna (Rodriguez-Ezpeleta et al., 2019).

In view of these complexities and uncertainties Moore et al. (2020a,b) identified several general sampling considerations and potential sampling designs to help overcome uncertainties in future studies on stock structure and provenance of tuna. For example, the use of temporally repeated sampling preferably of all age classes including larvae, young of the year (YoY) and actively spawning adults captured in or close to spawning areas is highly recommended. Long-lived migratory fishes with ocean-wide distributions caught from feeding aggregations are not suitable for population genetic studies because they are likely to represent stock mixtures that may generate inconsistent allele frequencies (Puncher et al., 2018). Multiple-year sampling of larvae and YoY from spawning areas helps to reduce the risk of false positives for divergence, which may occur due to sampling bias, temporal variation in the survival of larvae and the sweepstake effect (Larson and Julian, 1999; Maes et al., 2006; Puncher et al., 2018). Furthermore, to alleviate bias such as the Allendorf-Phelps effect, where the samples may represent the progeny of a small number of breeding adults rather than entire population (Allendorf and Phelps, 1981), it was proposed to
use about 40–60 fish samples per spatio-temporal stratum with equal proportions of both sexes. Furthermore, sampling over consecutive years across different ENSO phases, and multiple sampling in equatorial regions within the same year to test for intra-annual patterns, would also be appropriate.

In addition to the suggestions and key research questions proposed by Moore et al. (2020a,b) we recommend sampling in the western part of the EEZ of the Phoenix Island (Sibert et al., 2012) and the Phoenix Island Protected Area (PIPA) (Hernández et al., 2019), which are already known spawning areas for the bigeye tuna. In a study on natal homing in Atlantic Bluefin tuna, Rodríguez-Ezpeleta et al. (2019) found genetic differentiation between samples from two main spawning grounds in the Northwest Atlantic and Mediterranean Sea using thousands of genome-wide SNP markers obtained from more than 200 larvae and small YoY. The sampling design involving the use of larvae, small YoY and/or actively spawning adults captured in or close to known spawning areas will help investigate if bigeye tunas also display homing behavior (i.e., sexually mature individuals return to the same spawning areas where their progenitors spawned). As reported by Rodríguez-Ezpeleta et al. (2019), this will enable the development of an origin traceability tool which effectively assigns tuna catch to a given stock based on where they were spawned rather than where they were caught for more effective tuna fisheries management.

While the analyses carried in the current study are robust, it is not possible with the available data to ascertain the exact reasons for the observed signals of differentiation of the Phoenix Island sample or to establish its biological relevance. As emphasized previously, the results should be interpreted with caution as there are a number of uncertainties surrounding the sampling strategy for this study, for example, the use of smaller sample size increases the likelihood for the Allendorf-Phelps effect. This is particularly important for the sample from the Phoenix Island that indicates family effects in the form of heterozygote excess. In addition, the mixed age composition of population samples used in the study and the lack of temporally repeated sampling further add uncertainties to the patterns observed. Nevertheless, we argue that these results are sufficiently informative to encourage further research in this important area.

CONCLUSION
The current study represents the most geographically comprehensive population genetic assessment of bigeye tuna in the WCPO to date. It provides no substantial evidence to challenge panmixia using potentially neutral loci. However, there are signals of differentiation of samples captured from the EEZ of the Phoenix Island when using LUPS to estimate population partitioning. However, with the available data we are unable to discern the underlying processes leading to divergence in potentially adaptive loci among samples of putatively different geographic origin. Due to the uncertainties surrounding the use of fish caught from a possibly mixed-stock fishery and the opportunistic sampling used in this study, we recommend further high-resolution genomic assessments involving larger samples sizes of spatio-temporal samples of larvae, YoY and actively spawning individuals from throughout bigeye spawning habitat in the Pacific to strengthen future population genetic studies of tropical tunas.

DATA AVAILABILITY STATEMENT
The raw and filtered genotype datasets in Genepop file format analyzed for this study can be found in the Open Science Framework (OSF) repository (DOI: 10.17605/OSF.IO/6ZRA3).

ETHICS STATEMENT
The Animal Ethics Committee of the University of the South Pacific reviewed and approved this study.

AUTHOR CONTRIBUTIONS
CR conceived the project and wrote the grant proposal that funded the study. JN designed and conducted analyses with significant input from BS and AM, collected data, and wrote the manuscript with assistance from CR. JH, SN, and NS were representatives of the Secretariat of the Pacific Community (SPC) and provided fisheries science and management expertise to the study. CR critically revised the manuscript ensuring accuracy and integrity of the data and analyses. SPC was also responsible for tissue sample collection, facilitated sample access and preparation of Figure 1. All authors contributed to the final manuscript.

FUNDING
The genomic analyses in this work were funded by the University of the South Pacific Strategic Research Themes Intramural Program project no. F1006-R1001-71502-622. Funding for the tuna tissue bank was provided by the WCPFC as part of Project 35b.

ACKNOWLEDGMENTS
We gratefully acknowledge all the observers, research staff, and port samplers who collected tissue samples; Valerie Allain for establishing the tissue bank; Jed Macdonald of SPC for preparing Figure 1; Maria Ines Rico and Francois Roupsard for technical assistance with sample preparation; and Anthony Beeching and Feleti P. Teo from the Western and Central Pacific Fisheries Commission (WCPFC) for facilitating and granting access to the Tissue Bank.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.799684/full#supplementary-material
REFERENCES


**Conflict of Interest**: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note**: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Natasha, Stockwell, Marie, Hampton, Smith, Nicol and Rico.** This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.