frontiers in Marine Science

ORIGINAL RESEARCH
published: 12 July 2019
doi: 10.3389/fmars.2019.00341

RETRACTED: Close Kin Proximity in Yellowfin Tuna (*Thunnus albacares*) as a Driver of Population Genetic Structure in the Tropical Western and Central Pacific Ocean

Giulia Anderson[1]*, Monal Lal[1,2†], John Hampton[3], Neville Smith[3] and Ciro Rico[1,4]

[1] Molecular Analytics Laboratory (MOANA-LAB), School of Marine Studies, Faculty of Science, Technology and Environment, The University of the South Pacific, Suva, Fiji, [2] Australian Centre for Pacific Islands Research and School of Science and Engineering, University of the Sunshine Coast, Maroochydore, QLD, Australia, [3] Oceanic Fisheries Programme (OFP), Pacific Community, Nouméa, New Caledonia, [4] Instituto de Ciencias Marinas de Andalucía (ICMAN), Consejo Superior de Investigaciones Científicas, Puerto Real, Spain

OPEN ACCESS

Edited by:
Sophie von der Heyden,
Stellenbosch University, South Africa

Reviewed by:
Pablo Saenz-Agudelo,
Austral University of Chile, Chile
Stewart Grant,
University of Alaska Fairbanks,
United States

***Correspondence:**
Giulia Anderson
giuanders@gmail.com

†Present address:
Monal Lal,
University of the Sunshine Coast,
Maroochydore, QLD, Australia

Specialty section:
This article was submitted to
Marine Conservation
and Sustainability,
a section of the journal
Frontiers in Marine Science

Received: 16 February 2019
Accepted: 04 June 2019
Published: 12 July 2019

Citation:
Anderson G, Lal M, Hampton J,
Smith N and Rico C (2019) Close Kin
Proximity in Yellowfin Tuna (Thunnus
albacares) as a Driver of Population
Genetic Structure in the Tropical
Western and Central Pacific Ocean.
Front. Mar. Sci. 6:341.
doi: 10.3389/fmars.2019.00341

Recent studies argue for the presence of genetic population structure in yellowfin tuna (*Thunnus albacares*) in all oceans. However, the persistence of family groups has never been considered a viable mechanism of structure, nor has it been measured. We analyzed genetic similarity among 280 yellowfin tunas from seven population samples collected in the Western and Central Pacific Ocean (WCPO) using single nucleotide polymorphisms, and found population structure that was significantly explained by the presence of 96 individuals involved in 332 half or full sib dyads. We found significantly higher mean and median relatedness between individuals from the same sample groups, compared to individuals from different groups; and high relatedness between individuals caught at the same fish-aggregating device (FAD) than between those caught in the wider exclusive economic zone (EEZ). Alternatively, the EEZ of the Federated States of Micronesia may harbor exceptionally large numbers of close kin. We conclude that yellowfin directly school with related individuals through their first year, and at least demonstrate tightly overlapping regional fidelity as adults. These results may explain, to some extent, the patterns of population genetic structure recently observed in yellowfin tuna.

Keywords: ***Thunnus albacares*, co-dispersal, fidelity, population genetic structure, FAD, Federated States of Micronesia**

INTRODUCTION

Yellowfin tuna (*Thunnus albacares*) is a pantropical scombrid of great fishery importance. Yellowfin represents 28% of the global tuna fishery by volume, second only to skipjack (*Katsuwonus pelamis*) (ISSF, 2017), and is highly versatile as a consumer product. Mature individuals can produce sashimi-quality meat, and younger animals are often processed as high quality canned tuna, valued only slightly less than or equal to "white meat" albacore tuna (Macfadyen and Defaux, 2016). The species also exhibits greater fecundity and more rapid growth to maturity than other

RETRACTED

quality-oriented tuna species such as Atlantic, Pacific and Southern bluefin (*Thunnus thynnus, Thunnus orientalis*, and *Thunnus maccoyii*) and bigeye (*Thunnus obesus*) tunas, the former of which are recovering from or currently experiencing over-exploitation in specific oceanic regions (ISSF, 2017). In contrast, yellowfin tuna in the Western and Central Pacific Ocean (WCPO) is identified as merely fully exploited, yet supports a catch volume four times greater than that of bigeye (also "fully exploited") in the same region (ISSF, 2017). With careful management, yellowfin tuna can continue to provide large quantities of high-quality protein for human consumption.

Although tunas have traditionally been treated as panmictic within the WCPO, yellowfin tuna display some of the strongest population structuring of the tropical tuna species (Pecoraro et al., 2016). Numerous types of studies have detected some level of genetic differentiation in yellowfin populations within the Pacific Ocean. Morphometric calculations have been used to propose semi-independent eastern, central, and western populations (Suzuki et al., 1978); allozymes differentiated fish caught in the western and central Pacific from eastern fish (Ward et al., 1997); microsatellites detected structure between the Philippines and Papua New Guinea (Aguila et al., 2015); and genome-wide single nucleotide polymorphisms (SNP) distinguished fish from eastern and central Pacific, and to a lesser degree central from western Pacific regions (Grewe et al., 2015). Some of the same studies did not detect structure in sympatric species, like skipjack, at the same intra-ocean scale and using the same molecular markers (Ely et al., 2005). The genetic observations are backed by morphometric and behavioral studies that have confirmed the presence of a magnetite organ that facilitates electromagnetic navigation and precise homing (Walker et al., 1984). Furthermore, tagging studies describe a median individual range of less than 800 km in the WCPO (Sibert and Hampton, 2003) despite the demonstrated capacity to travel 4000 km or more (Wild and Hampton, 1993). However, there is no consensus to date about the extent to which yellowfin tuna populations are structured, or which of the acknowledged mechanisms drive genetic differentiation.

Given our currently imperfect understanding of yellowfin tuna population structure, it is reasonable to explore previously overlooked mechanisms. For example, the presence of family groups can significantly impact distribution of the neutral genetic diversity and thus the population structure of marine taxa. Selwyn et al. (2016) found spatially clustered family groups in a marine goby species (*Coryphopterus personatus*) associated with small, but significant, population structure over short distances and within-population deviations from Hardy–Weinberg equilibrium (HWE), both of which are consistent with a pattern known as chaotic genetic patchiness (CGP) or fluctuating genetic mosaics (Eldon et al., 2016). Although CGP has never been directly applied to a tuna species, it shares similarities, and possibly driving factors, with the highly complex system of structure proposed by the collective literature for yellowfin tuna. However, genetic relatedness has never been thoroughly explored in tuna because of the traditionally held improbability of finding persistent family organization in an ocean dwelling, broadcast spawning, cosmopolitan species.

To assess this possibility, we established kin relationships using 1340 SNPs and 280 yellowfin tuna specimens showing unexpected patterns of population structure. The individuals were collected opportunistically from seven exclusive economic zones (EEZs) across the WCPO in 30 catch events conducted between 2009 and 2014. We observed the presence of full- and half-sib dyads between and within EEZs and school-specific catch events and compared the overall level of relatedness within and between groups. Apart from defining the impact of kin presence on population structure, we also explore the significance of our findings for the biology and fisheries management of yellowfin tuna. We attempt to separate the impacts of kin co-dispersal from other biological and behavioral mechanisms, such as sweepstakes reproduction and site fidelity, and discuss the implications in the context of genetic population structure and fisheries management.

MATERIALS AND METHODS

Sample Collection

Yellowfin specimens were selected for analysis from tissue samples archived in a tissue bank collection managed by the Pacific Community (SPC) under the auspices of the Western and Central Pacific Fisheries Commission (WCPFC), in Nouméa, New Caledonia. Tissue samples consisted of muscle plug biopsies taken by scientists during research cruises. To counteract any potential contamination that may have occurred during initial field sampling, and avoid cross-contamination during subsampling of each specimen, the exposed surface of each tissue plug was sliced away prior to sample dissection for DNA extraction. Specifically, a 2 mm^3 piece of tissue taken from the core of the sample using a scalpel that was re-sterilized with ethanol over a flame after any cut that contacted the tissue surface.

Tissue samples for this study were originally selected opportunistically from the available collection based on catch location and date, such that all samples in a grouping were collected within 6 weeks of each other, and within an area of 3400 km^2 (radius of 330 km). Population samples from different years were selected from archived tissue collected in the same area and season but in different years. Samples were accompanied by metadata including catch location, date, and catch event number, fish size, and a description of objects associated with the school of interest. A total of 280 individuals from 7 geographic locations were analyzed, representing 4 countries and spanning 6 years (**Figure 1** and **Supplementary Table S1**). Population samples are labeled based on country and year of collection: from the Federated States of Micronesia in 2009 and 2011 (FM09 and FM11, respectively), the Gilbert Islands of Kiribati in 2009 (GL09), New Caledonia in 2014 (NC14), and in Papua New Guinea in 2009, 2011, and 2013 (PG09, PG11, and PG13).

DNA Sequencing and SNP Calling

DNA extraction and sequencing was conducted by Diversity Arrays Technology (DArT PL). Its patented next generation

RETRACTED

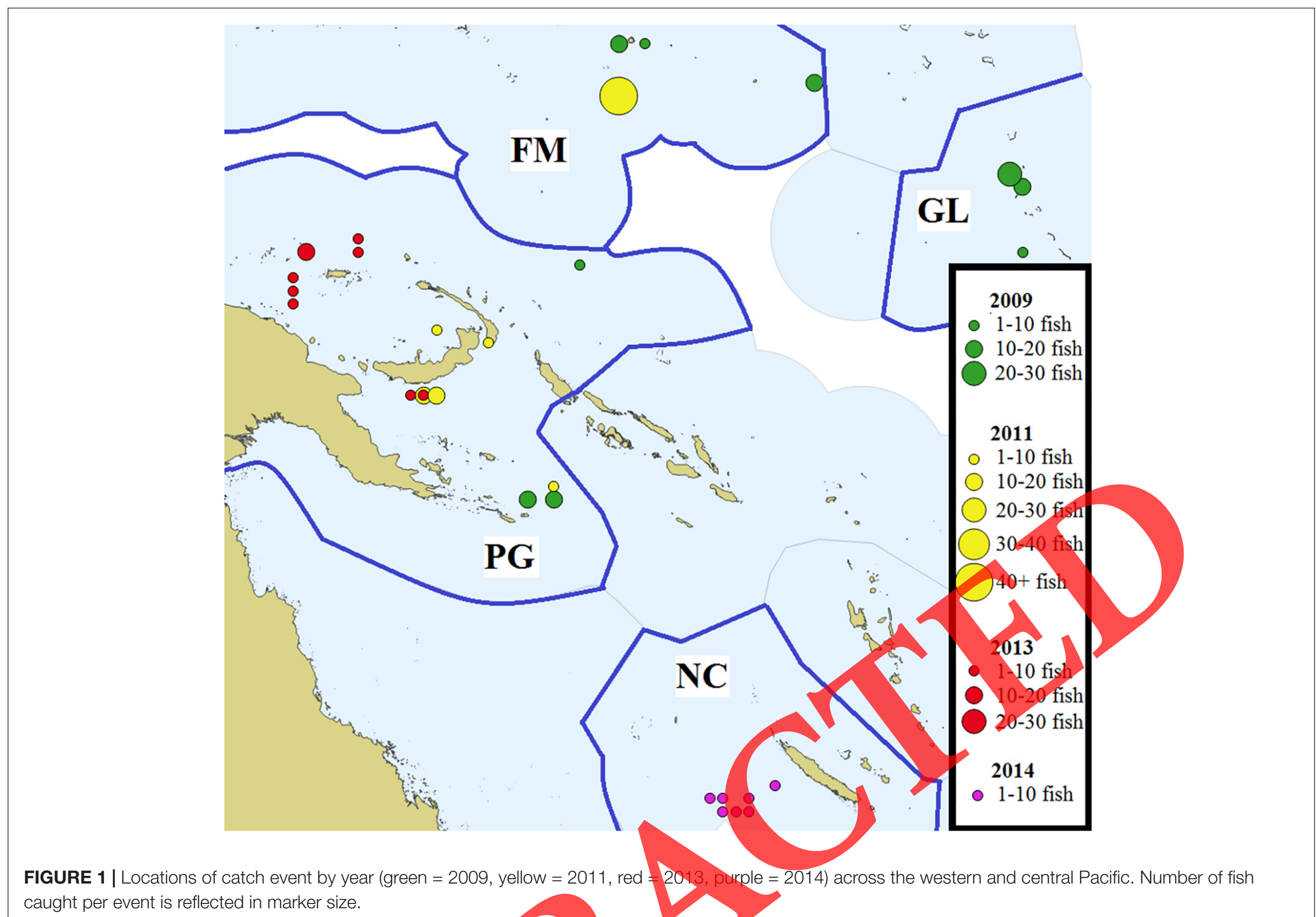

FIGURE 1 | Locations of catch event by year (green = 2009, yellow = 2011, red = 2013, purple = 2014) across the western and central Pacific. Number of fish caught per event is reflected in marker size.

RETRACTED

sequencing protocol, DArTseq, is a cost-effective option for generating high quality, high-throughput SNP datasets for non-model species. Although some steps are proprietary knowledge, a description of the DArTseq protocol is available in Kilian et al. (2012), Sansaloni et al. (2011), and Ren et al. (2015). Following automated DNA extraction, samples were digested using *Pst*I and *Sph*I restriction enzymes. Methylation-sensitive enzymes were chosen to avoid highly repetitive, methylated genomic regions that are minimally informative and tend to carry elevated risk of misinterpreting paralogs as a single locus during marker calling. Specialized adaptors were ligated to digested DNA. Both *Pst*I and *Sph*I adapters included a PCR primer sequence and Illumina flowcell attachment sequence, and the *Pst*I adaptor also included a unique, varying length barcode sequence for sample recognition within pooled libraries. PCR only amplified fragments capped with both adaptors, using the following protocol: 1 min denaturation at 94°C, 30 cycles of 20 s at 94°C, 30 s and 58°C and 45 s at 72°C, and a final extension step of 7 min at 72°C. Libraries were then further amplified using bridge PCR on the Illumina HiSeq 2500 platform and sequenced on the same platform. The resulting data was submitted to an in-house software, DArTsoft, which interprets sequences from images of fluorescence taken during Illumina sequencing and produces FASTQ files. Files were quality controlled for sequences with 90% confidence at 50% of bases, and split by barcode into individual specimens. Sequences were aligned *de novo*. A separate algorithm, DArTsoft14, called SNPs and further quality filtered for singletons and other suspected sequencing errors. The final output produced by DArT was a genotype report of all identified SNPs, their global call rate, polymorphic information content, and their co-dominant status in each sequenced specimen.

The returned dataset was further filtered for locus quality. Loci were first culled by removing all but one SNP per sequenced DNA fragment. Remaining loci were selected based on a 99% call rate, a minimum read depth of 7×, and 5% minor allele frequency. F_{ST} outlier analyses were conducted with BAYESCAN v. 2.1 using the prior odds for neutral model and a 10% false discovery rate. Individuals were submitted to BAYESCAN in their 7 original sample groups. Next, loci were extracted that showed deviation from Hardy–Weinberg equilibrium across all populations. HWE tests if loci occur at frequencies that deviate from selectively-neutral assumptions and was analyzed using Arlequin v. 3.5.2.1 (Excoffier and Lischer, 2010) using the most sensitive possible p-value threshold of 0.0001. HWE results were also filtered for loci with a maximum observed heterozygosity of 0.5, as an independent control for the potential merging of paralogous loci in the DArTseq pipeline. All individuals were kept in their

original sample groups during SNP filtering. Raw datasets are publicly available at: https://doi.org/10.17605/OSF.IO/WTZD8.

Population Structure

To test for population structure, genetic diversity and population structure parameters were measured for each sample group. The inbreeding coefficient (F_{IS}) and adjusted expected and observed heterozygosity ($H_{n.b.}$ and H_o) were obtained in GENETIX v. 4.05 (Belkir et al., 2004), while pairwise F_{ST} (Weir and Cockerham, 1984) was calculated using Arlequin. ADMIXTURE v. 1.3.0 (Alexander et al., 2015) was used to recommend the number of independent genetic clusters (k) among all sampled individuals and then estimate the probability of membership of each individual to the resulting clusters. Analyses were run with hypothetical k values ranging from 1 to 10, and the optimal k value was selected based on a low coefficient of variance. A Discriminant Analysis of Principle Components (DAPC) was conducted using the "dapc" and "optim.a.score" commands in the *adegenet* package in R v. 3.3.1 to validate the results of pairwise F_{ST} and ADMIXTURE analyses.

Relatedness

In response to unexpected heterozygosity and inbreeding assessments, we explored the possibility that observed population genetic structure resulted from the presence of related individuals in our sample groups. First, to more accurately group individuals for relatedness analyses, all samples were submitted to a clustering algorithm using the "find.clusters" command in the R package *adegenet*. All further analyses were carried out in triplicate: once on all sampled individuals treated as a single populations, and once on each of the two recommended clusters. However, due to similarity of results, this report focuses on universal analyses.

To assess relatedness among sampled individuals we used two software programs: RelateAdmix (Moltke and Albrechtsen, 2014) and COANCESTRY v. 1.0.1.7 (Wang, 2011). The algorithm employed in RelateAdmix maximizes its accuracy when working with admixed populations, such as our dataset that incorporates all individuals. COANCESTRY allows for the consideration of inbreeding, which F_{IS} results indicate was a relevant feature for some sample groups. Due to their more conservative statistical trends and appropriateness for admixed tuna populations, RelateAdmix results are given precedence during interpretation. Conveniently, results from the two algorithms rarely contradict each other.

COANCESTRY was first run using simulated genotypes based on empirical allele frequency data, which were constructed in full sib, half sib, and unrelated pairs. The generated dyads were submitted to seven algorithms offered by the COANCESTRY software, and the average difference between empirical results and expected values calculated to determine which algorithm was most accurate for determining the relationship correlation coefficient (r) for this species. TrioML was selected and used in further analyses of empirical data to assess r between all sampled individuals. RelateAdmix requires *a priori* information about general genetic clustering, which was recycled from ADMIXTURE. The output.P and .Q files from the selected ADMIXTURE analysis were submitted to RelateAdmix, and output values were used to calculate $\theta = (k_1/4)+(k_2/2)$, where θ is the expected fraction of two genomes is identical by decent.

RelateAdmix and COANCESTRY both produce relatedness coefficients but no recommendation of how to differentiate full sibling pairs from half siblings and more distantly related individuals. To determine cut-off values for categorizing relatedness coefficients, we again used COANCESTRY to simulate 200 dyads each of full sibling, half siblings, and unrelated individuals based on the same allele frequencies as the original population samples. The simulated genotypes were then submitted to both COANCESTRY and RelateAdmix, and results used to calculate the lower 95% confidence interval of the mean of each group, assuming normal distribution. These values became the cut-offs for delineating full- and half-siblings in empirical datasets in the respective software programs. The pairs of specimens identified as related by either software were then cross-referenced, and those identified by both algorithms consolidated into a consensus dataset.

Upon confirming that our samples include numerous, closely related individuals, we assessed the extent to which the observed relatedness was responsible for the presence of population genetic structure by conducting a second analysis of pairwise F_{ST} after removing one of each sibling from each dyad in the consensus dataset.

Finally, to further ascertain how the identified kin groups are distributed among samples, a Spearman's rho statistic and linear regression was also calculated between the number of full sib dyads and potential dyads per group, to assess the influence of sample size on number of full sibs per bin. We also explored spatial and temporal distribution of kin using the Mann–Whitney test for comparing medians of non-parametric data, and Student's T-test for mean. We organized individuals from the all-sample dataset by location (at the EEZ level), year at a single location, and catch event at a single location. We also explored the importance of environmental context by binning dyads based on the presence and type of floating objects specimens were associated with when caught. All statistical analyses were conducted in the R statistical analysis package (R Core Team 2016) using commands "cor.test" with the Spearman method, "lm," "wilcox.test," and "t.test," respectively.

RESULTS

Sequencing using the DArTseq pipeline identified 49,078 SNP loci on 30,727 DNA fragments. Further, stringent filtering retained 1135 neutral, polymorphic markers in HWE that occur in at least 99% of samples analyzed, and demonstrated a minimum read depth and allele frequency of 7 and 5%, respectively (**Table 1**).

Genetic diversity and population genetics parameters were measured, including multi-locus heterozygosity and fixation indices, using all individuals in their original sample groups. All measurements varied among the sample groups, with FM09 and FM11 always similar and distinct from the other samples (**Table 2**). They had the highest expected heterozygosity values and deviation from Hardy–Weinberg Expectations in the form

RETRACTED

TABLE 1 | Number of loci remaining after each quality filtering step.

Filtering step	Number of loci
Initial	49078
Duplicates on contig	30727
Call rate (99%)	7092
Read depth (7×)	6635
MAF (5%)	1361
FST outliers (FDR 90%)	1354
HWE ($p < 0.0001$)	1354
Heterozygosity (0.5)	1135
LD (70%)	**1135**

TABLE 2 | Population genetic diversity parameters.

Population	Heterozygosity		F_{IS}
	$H_{n.b.}$ (stdev)	H_o (stdev)	
FM09	0.283 (0.12)	0.353 (0.19)	−0.255
FM11	0.278 (0.12)	0.346 (0.19)	−0.255
GL09	0.203 (0.11)	0.207 (0.12)	−0.020
NC14	0.185 (0.12)	0.177 (0.13)	0.041
PG09	0.182 (0.11)	0.171 (0.11)	0.058
PG11	0.180 (0.11)	0.173 (0.11)	0.041
PG13	0.197 (0.11)	0.191 (0.12)	0.036

TABLE 3 | Pairwise F_{ST} values (below diagonal) and *p*-values (above diagonal).

	FM09	FM11	GL09	NC14	PG09	PG11	PG13
FM09	–	**0.0001**	**0**	**0**	**0**	**0**	**0.00129**
FM11	**0.00948**	–	**0**	**0**	**0**	**0**	**0.0002**
GL09	**0.02328**	**0.02317**	–	0.0319	**0**	**0**	0.01673
NC14	**0.02789**	**0.02998**	0.00274	–	0.00535	0.00426	0.04099
PG09	**0.04378**	**0.04363**	**0.00817**	0.00729	–	**0**	**0**
PG11	**0.04689**	**0.04744**	**0.00887**	0.00755	**0.01097**	–	**0**
PG13	**0.02848**	**0.03045**	0.01	0.01131	**0.01651**	**0.01789**	–

Statistically significant values are bold, using a Bonferroni corrected alpha value of 0.00238.

RETRACTED

of significant heterozygosity excess ($H_{n.b} = 0.28$ and $H_o = 0.35$ in both cases), as well as significant outbreeding, as measured by F_{IS} (more extreme than −0.25, $p < 0.0001$). In contrast, GL09, NC14, PG09, PG11, and PG13 had much more moderate heterozygosity values, with $H_{n.b.}$ values ranging from 0.18 to 0.20, and slight-to-negligible heterozygosity deficits (H_o between 0.17 and 0.20; F_{IS} between −0.02 and 0.05, $p < 0.0001$). Pairwise F_{ST} values were significant in all comparisons except five after Bonferroni correction: between NC14 and all three PG groups, and between GL09 both NC14 and PG13. Statistically significant values ranged from 0.008 to 0.047 (**Table 3**). All values greater than 0.02 involved pairwise comparisons with either FM09 or FM11. A DAPC using *adegenet* in R confirmed the separation of FM samples from the remaining sample groups (**Figure 2**). ADMIXTURE recommended the separation of sampled individuals into two populations, one of which accounted for 19 samples from either FM09 or FM11 and two individuals in GL09, and the other that incorporated all other individuals with high confidence (**Figure 3**). Collectively, the tests provided clear evidence of population genetic structure in the WCPO. The heterozygosity and F_{IS} values justify further exploration of collective dispersal of kin (Selwyn et al., 2016).

To ensure that pre-existing population structure did not influence the perceived relatedness of compared individuals in tests for kinship, *adegenet*'s "find.cluster" command in R was used in conjunction with ADMIXTURE results to separate specimens into populations with minimal sub-structure. *Adegenet* recommended the separation of 38 individuals from the main dataset, including all of those highlighted by ADMIXTURE. Specimens were separated as recommended and processed by COANCESTRY and RelatedAdmix in two separate analyses. The cross-validation of results from both software programs for the cluster of 38 FM and GL samples identified between 4 and 19 siblings for every sampled individual, with an average of 8 siblings. Among the 21 individuals that ADMIXTURE recommended as a separate cluster, the consensus of the two software programs labeled on average ten siblings per specimen. Siblings also occurred in the dataset of remaining 240 samples from all locations, but at a much lower rate (36 of 28680 possible dyads).

When all samples were analyzed as a single group, COANCESTRY and RelateAdmix concurred on the presence of 29 full-sib dyads, and 332 dyads that were equally or more similar than half sibs. Of all high-confidence siblings, 283 occurred within the same EEZ, compared to 49 between EEZs; 190 that occurred in the same EEZ were caught in the same year, compared to 93 between years; and 88 occurred within the same catch event, compared to 102 between catch events but in the same EEZ and year (**Table 4**). Despite recommendations to divide the specimens into two clusters, 37 dyads from between the resulting groups were listed as siblings in the consensus dataset.

Given the confirmation that related individuals were present in numbers that might affect population structure, we removed one individual from each sibling pair from the dataset and re-calculated pairwise F_{ST} values (**Table 5**). The number of significant comparisons after Bonferroni correction dropped from 16 to 9, and the largest F_{ST} values dropped from 0.047 to 0.016. Although comparisons with FM11 were compromised due to the inappropriately low number of remaining samples in that group, the reduction was evident in all groups in which kin were observed. ADMIXTURE results likewise changed from recommending $k = 2$ to $k = 1$.

The recommended sibling groups were not evenly distributed between population samples. Of the 190 half- or full-siblings that occurred in the same EEZ and year, 106 were from sample group FM11, 61 from FM09, 12 from PG13, and five or fewer from GL09, NC14, PG11, and PG09. Specific to full siblings, 9 of 23 relevant dyads occurred in FM09, 5 in FM11, 5 in PG13, and two or fewer in the other population samples. Of the 142 cross-location or cross-year dyads, 91 were between FM samples from different years, and another 44 were between either FM sample and the nearby GL sampling sites. Two half-sib dyads were found between years in PG, and only 5 occurred between all other, more disparate locations like PG and FM. Among dyads

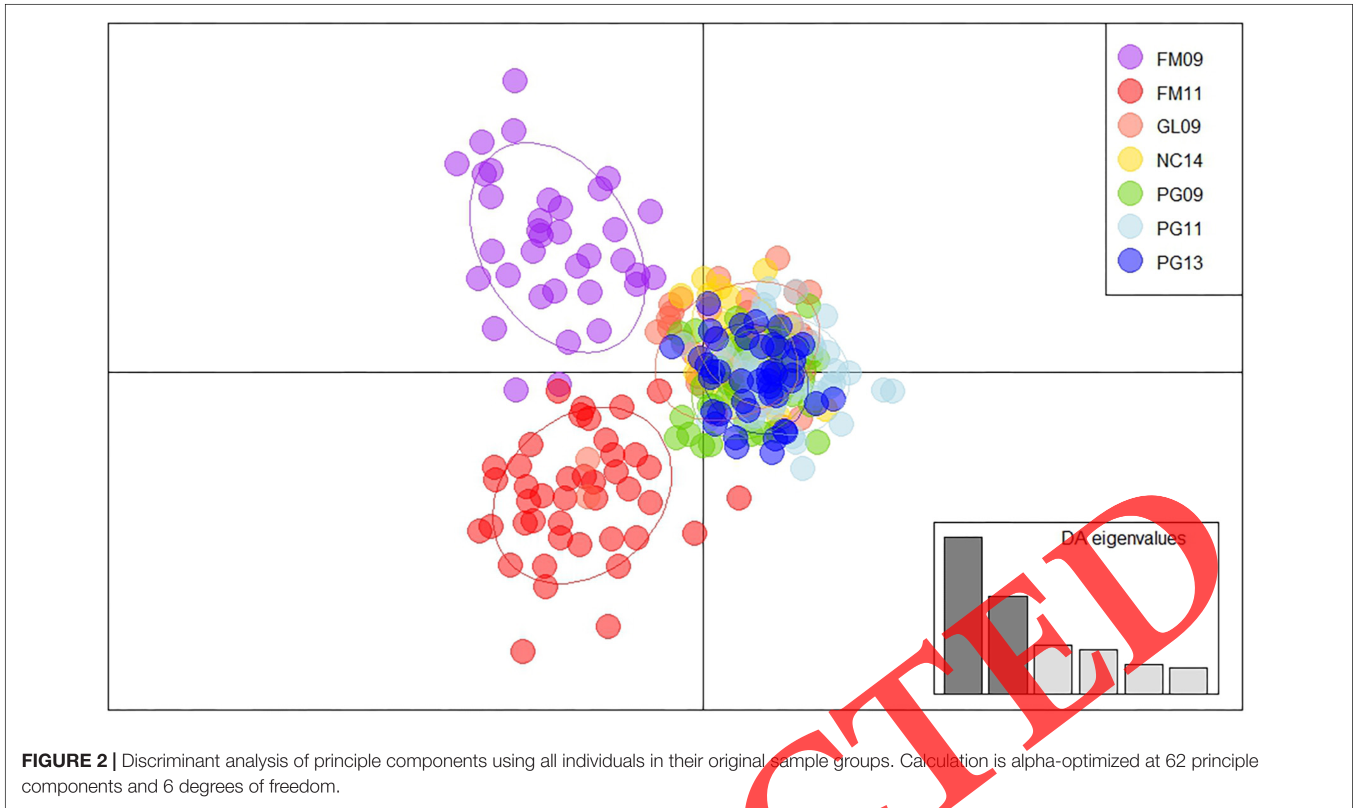

FIGURE 2 | Discriminant analysis of principle components using all individuals in their original sample groups. Calculation is alpha-optimized at 62 principle components and 6 degrees of freedom.

Probability of Membership
0.0 0.2 0.4 0.6 0.8 1.0
FM09 FM11 GL09 NC14 PG09 PG11 PG13
Sample Group

FIGURE 3 | Probability of membership of all sampled individuals to either of two genetic clusters, as recommended by ADMIXTURE.

RETRACTED

TABLE 4 | The distribution of kin dyads, organized by population sample, catch event, EEZ.

	Grouping	Full sibs	% of possible dyads	Half and full sibs	% of possible dyads
Within sample group	FM09	9	1.4	61	9.7
	FM11	5	0.5	106	11.2
	GL09	2	0.2	5	0.5
	NC14	1	0.3	2	0.6
	PG09	0	0.0	2	0.2
	PG11	1	0.1	2	0.2
	PG13	5	0.5	12	1.3
Between sample groups	FM09-FM11	1	0.1	91	5.9
	FM09-GL09	0	0.0	20	1.3
	FM09-PG13	0	0.0	2	0.3
	FM11-GL09	5	0.3	24	1.2
	FM11-PG13	0	0.0	2	0.1
	GL09-PG13	0	0.0	1	0.1
	PG11-PG13	0	0.0	2	0.1
Within catch events	FM09 1	3	3.3	17	18.7
	FM09 2	0	0.0	3	8.3
	FM09 3	3	3.8	17	21.8
	FM09 total	6	2.9	37	18.0
	FM11 1	2	0.9	32	13.9
	FM11 2	1	0.4	28	12.1
	FM11 total	3	0.6	60	13.6
	GL09 1	1	1.8	2	3.6
	GL09 2	1	1.5	2	3.0
	GL09 total	2	0.9	4	1.0
	NC14 3	1	100	1	100.0
	NC14 total	1	3.2	1	3.2
	PG09 total	0	0.0	0	0.0
	PG11 3	1	1.0	2	13.3
	PG11 total	1	0.5	2	0.9
	PG13 3	0	0.0	1	1.3
	PG13 4	4	11.1	6	16.7
	PG13 7	0	0.0	1	0.1
	PG13 total	4	2.5	7	4.3
Between catch events within a sample group	FM09 1–2	0	0	7	5.6
	FM09 1–3	1	0.5	20	11.0
	FM09 2–3	2	1.7	8	6.8
	FM09 total	3	0.7	35	8.2
	FM11 1–2	2	0.4	58	12.0
	FM11 total	2	0.4	58	12.0
	GL09 2–3	0	0.0	1	0.7
	GL09 total	0	0.0	1	0.2
	NC14 2–5	0	0.0	1	100
	NC14 total	0	0.0	1	0.3
	PG09 1–2	0	0.0	1	1.7
	PG09 2–3	0	0.0	1	0.4
	PG09 total	0	0.0	2	0.4
	PG11 total	0	0.0	0	0.0
	PG13 2–3	0	0.0	3	2.6
	PG13 6–7	1	0.2	2	4.4
	PG13 total	1	0.2	5	0.6

(Continued)

TABLE 4 | Continued

	Grouping	Full sibs	% of possible dyads	Half and full sibs	% of possible dyads
Within EEZ	FM	15	0.5	258	8.2
	GL	2	0.2	5	0.6
	NC	1	0.3	2	0.6
	PG	6	0.1	18	0.2
Between EEZs	FM-GL	5	0.1	44	1.3
	FM-PG	0	0.0	4	0.0
	GL-PG	0	0.0	1	0.0

TABLE 5 | Pairwise F_{ST} values (below diagonal) and *p*-values (above diagonal) of sample groups with siblings removed.

	FM09	FM11	GL09	NC14	PG09	PG11	PG13
FM09	–	0.02237	**0**	**0**	**0.0003**	**0.0002**	0.05653
FM11	0.00289	–	0.01406	0.04188	0.10643	0.04198	0.13603
GL09	**0.00767**	0.00364	–	0.16909	**0.0003**	**0.0003**	0.03950
NC14	**0.00694**	0.00393	0.00153	–	0.01693	0.01129	0.05653
PG09	**0.01421**	0.00694	**0.00752**	0.00670	–	**0**	**0**
PG11	**0.01619**	0.00962	**0.00746**	0.00653	**0.01118**	–	**0**
PG13	0.01246	0.00700	0.00858	0.00853	**0.01350**	**0.01454**	–

Statistically significant values are bold, using a Bonferroni corrected alpha value of 0.00238.

identified as full sibs by both algorithms, cross-sample siblingship only occurred between FM11 and either FM09 or GL09.

The same uneven trends were evident in continuous *r* data (**Figures 4**, **5**). Comparisons of RelateAdmix median results using Mann–Whitney *U* tests are reported here, as RelateAdmix tended to be more conservative than COANCESTRY and tests of median more appropriate for the demonstrably non-parametric data than tests of mean. COANCESTRY data and analyses using *T*-tests for mean produced similar statistical conclusions. When samples were organized by location, the median level of relatedness of all FM samples ($r = 0.117$) was significantly larger than that of any other EEZ (median *r* values ranging from 0.097 to 0.102, all *p*-values $< 2.2e\text{-}16$). Median relatedness was also significantly larger within the NC and PG EEZs (median *r* values of 0.102 and 0.097) than between EEZs (between EEZ $r = 0.096$, compared to NC $p = 3.49e\text{-}12$, PG $p = 0.005$), although the absolute difference was at the potentially uninformative scale of 10e-4. Median relatedness in GL was not significantly different from that between EEZs ($p = 0.284$).

On the suspicion that uneven distribution of kin might stem from opportunistic and uneven sample collection across locations, we explored the correlation between number of possible dyads in a group and the number of kin identified. A linear regression between number of siblings and number of possible dyads when data was organized by EEZ, catch event, or associated objects produced a statistically significant but incomplete relationship. Number of dyads only accounts for 34% of the variance in number of siblings ($p = 0.0004$). When incorporated into a multiple regression that also considered

RETRACTED

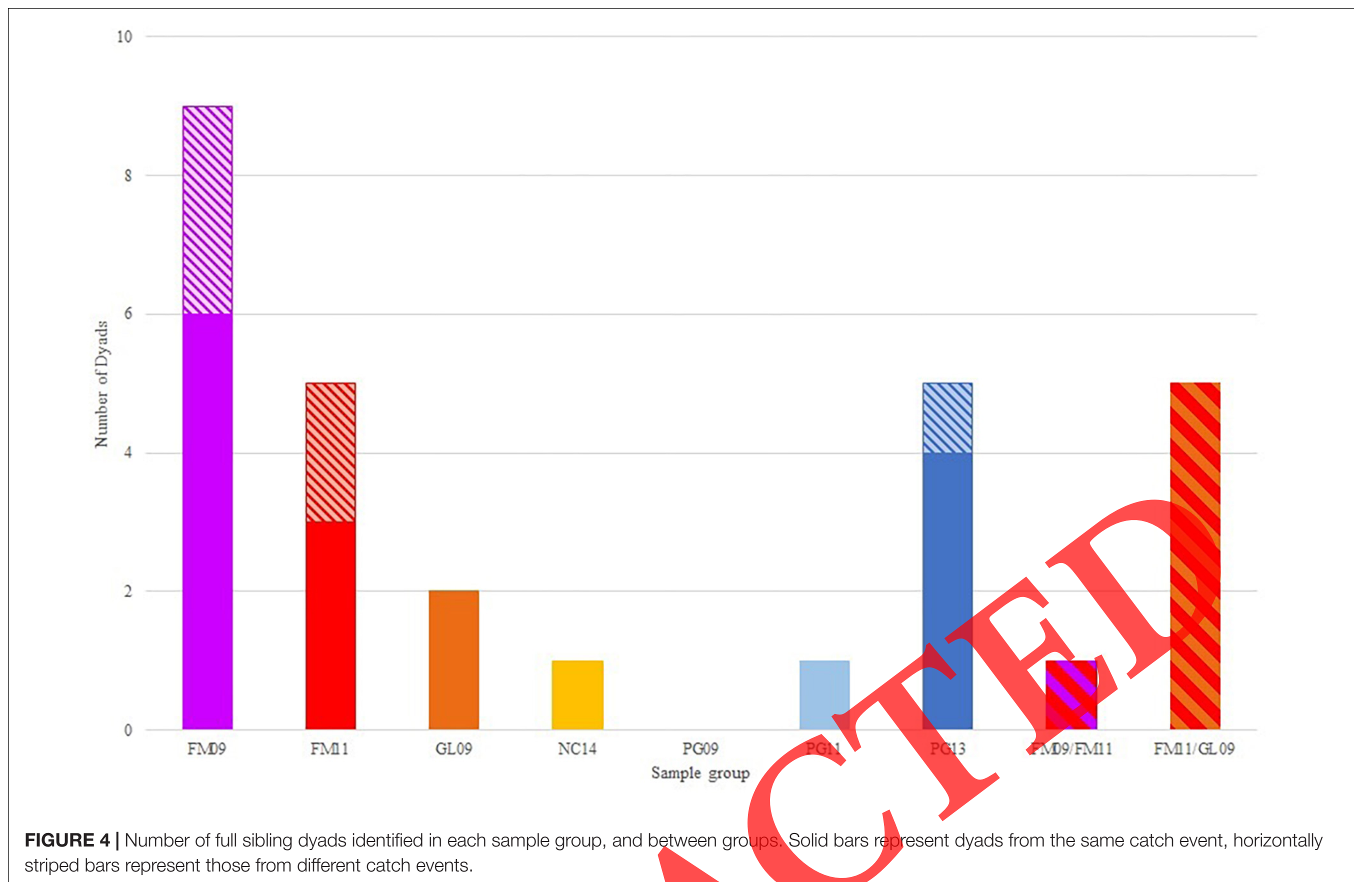

FIGURE 4 | Number of full sibling dyads identified in each sample group, and between groups. Solid bars represent dyads from the same catch event, horizontally striped bars represent those from different catch events.

RETRACTED

sample group, number of dyads became not significant (p = 0.204). Alternatively, a Spearman correlation produced a stronger relationship at 50% correlation and p-value of 0.005. However, this is potentially an overly simplistic representation, and still leaves a great deal of the observed variation unexplained. It was therefore appropriate to further explore the distribution of relatedness for temporal and environmental patterns.

Both FM and PG were sampled over multiple years, and demonstrated similar temporal trends. Median relatedness in both EEZs was larger among specimens caught in the same year than between years (Micronesia median r = 0.121 within years, 0.113 between years, p = 3.94e-8; Papua New Guinea median r = 0.098 within years, 0.097 between years, p = 0.031), but median relatedness between years in the same EEZ was still significantly larger than that between EEZs (FM p < 2.2e-16; PG p = 2.24e-6). Furthermore, when samples from each EEZ were subdivided into catch events, FM and PG demonstrated significantly larger median relatedness within catch events (median r = 0.124 and 0.099, respectively) than between events (median r = 0.116 and 0.097; p = 5.06e-7 and 0.036, respectively). When divided by year, only FM09 retained this trend according to both algorithms. PG13 also showed a weaker but significant trend in results produced by RelateAdmix. Samples from GL, NC, as well as FM11 and PG09 and PG11, did not demonstrate any difference within-versus-between catch events. One detail that distinguished FM09 and PG13 from other sample groups was the environmental context within which samples were collected, specifically the presence and type of floating objects associated with sampled schools. Dyads were therefore re-organized into groups based on associated objects and re-assessed using the Mann–Whitney U test and Student's T-test.

Organization by associated object produced several significant trends. Dyads from within a catch event that lacked data on associated objects had significantly larger median relatedness than any other category (median r = 0.012, p between 0.0001 and 2.2e-16), followed by those in the same sampling event around drifting fish aggregating devices (dFADs) (median r = 0.105). Median relatedness around dFADs was greater than that from individuals found around anchored fish aggregating devices (aFADs), whale sharks, current lines, seamounts, or no object at all (medians range from 0.096 to 0.101, p ranges from 0.009 to 4.10e-6). Although COANCESTRY and RelateAdmix produced some contradictory conclusions about the ranking of remaining bins, the absolute differences under discussion were miniscule. Functionally, there was no significant difference in relatedness between fish caught in free swimming schools and those caught around current lines, seamounts, and natural drifting objects. Alternatively, when comparing means, individuals from around aFADs (mean r = 0.113) showed significantly larger relatedness values than those around seamounts, current lines, natural drifting objects, whale sharks, or in free schools (mean r ranges from 0.099 to 0.101, p-values range from 0.006 to 0.0002). These

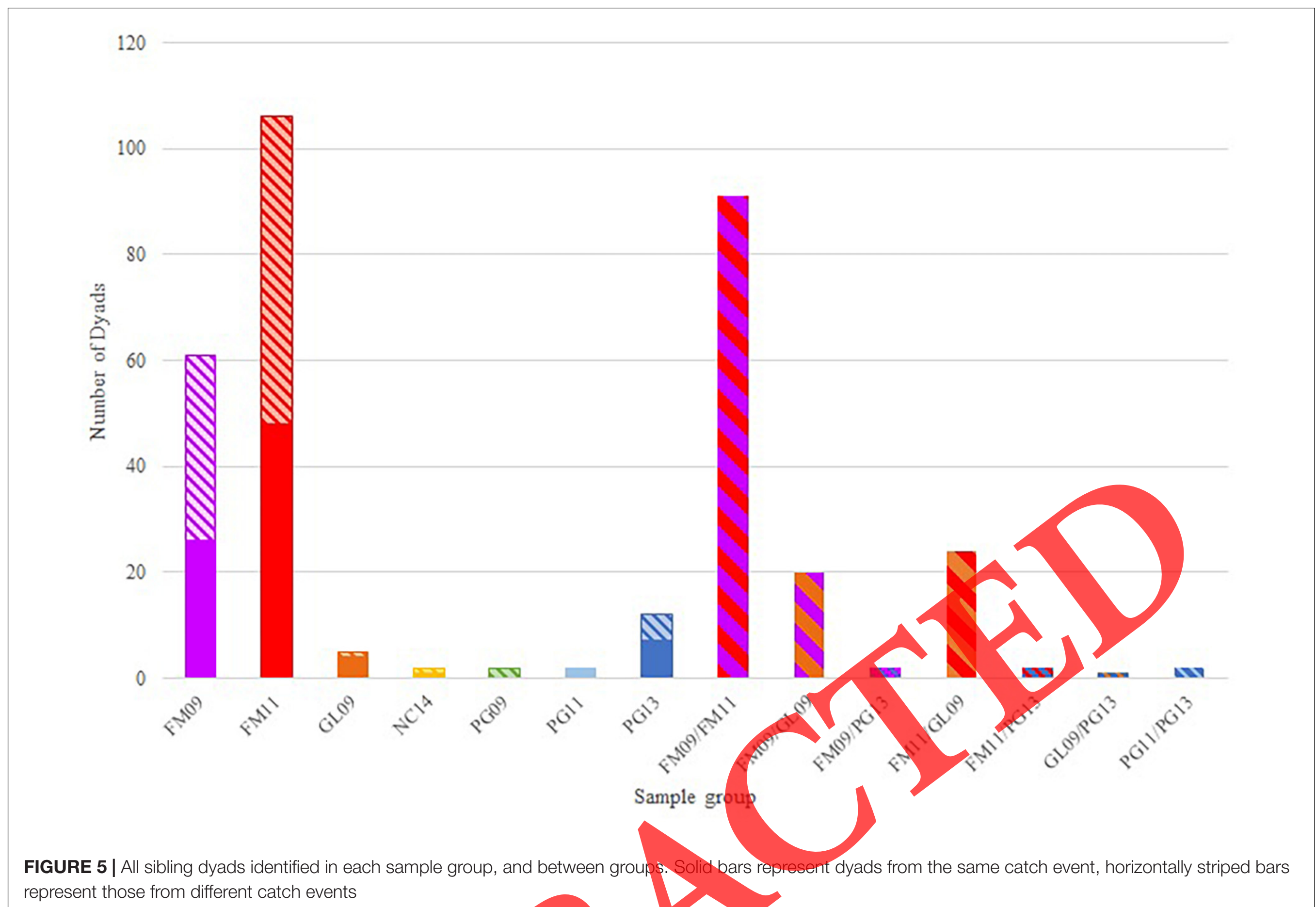

FIGURE 5 | All sibling dyads identified in each sample group, and between groups. Solid bars represent dyads from the same catch event, horizontally striped bars represent those from different catch events

RETRACTED

TABLE 6 | The distribution of kin, organized by type of associated objects.

Grouping	Full sibs	% of possible dyads	Half and full sibs	% of possible dyads
aFAD	7	0.4	24	1.3
Currentline	0	0	0	0
dFAD	3	1.2	20	9.0
Drifting object	1	0.2	2	0.5
No info	4	0.2	61	2.5
Seamount	0	0	0	0
Unassoc	2	0.3	5	0.8
Whale Shark	0	0	0	0
Between	12	0.04	220	0.7

calculations reflected the distribution of full and half-sib dyads retained in the consensus dataset: large numbers of sibling dyads lack association data, a moderate number of siblings were found around dFADs, a disproportionate number of full sibs were collected near aFADs, 65% of all sibling dyads occurred between association groups, and less than 1% originated from all other categories (**Table 6**). In short, there was a great deal of noise in the data when organized this way, which potentially confounded more subtle statistical trends.

DISCUSSION

In the process of exploring population genetic structure of yellowfin tuna in the WCPO, we identified 332 dyads that are genetically similar enough to be half sibs, 29 of which are similar enough to be full sibs. Individual dyads can also often be amalgamated into complex family groups that contain up to 24 full or half sibs. The presence of sibling groups in multiple yellowfin tuna population samples is cross-validated by heterozygosities and inbreeding coefficients. Population samples with few related individuals have slight but significant and positive inbreeding coefficients and little to no heterozygote deficit, which is consistent with other yellowfin studies that specifically sampled across multiple age groups per location, used genome-wide neutral SNP loci, and found effectively no heterozygosity deficit (Grewe et al., 2015). In summary, our identification of a limited number of kin coincides with the expected genetic diversity of explicitly mixed sample groups. Conversely, those sample groups with numerous sibling groups show the opposite trend: a heterozygous excess, and negative and significant F_{IS}. While a negative F_{IS} value is normally associated with outbreeding, Pudovkin et al. (2010) suggests that the smaller the number of breeders, the larger the possibility that genetic differences between parents result in a large heterozygote

excess, and consequently a negative F_{IS}, in the progeny. Waples and Allendorf (2015) concur that the F_{IS} of a group will be increasingly negative as it is produced by fewer parents, and the authors highlight that this may particularly apply to species with a highly fecund, broadcast spawning life history (which includes yellowfin tuna). This supports current observations that samples from purportedly outbred populations are also composed of large numbers of related individuals. The alignment of basic population structure analyses and the presence of families confirm that the observations made during kinship assessments are not mere artifacts produced by low genetic diversity in the larger populations, but represent actual relatedness.

The presence of kin is also clearly an important factor creating population structure among sampled individuals. When individuals were separated as recommended by ADMIXTURE and the *adegenet* package, one group was made up exclusively of related individuals from a complex family group. Given that population structure is fundamentally a measure of ancestral relatedness, the clustering algorithms' interpretation of a large number of highly similar individuals as a distinct population group rather than a family is an artifact of the unexpected size of the group (Astle and Balding, 2009). Without the extracted family group, and even more so when one individual from each sibling pair across all samples was deleted from the dataset, pairwise F_{ST} values between remaining sample groups dropped markedly and the recommended number of clusters reduced to $k = 1$. It is very likely that the study area should be treated as a single population with moderate substructure, and the instances of major structure recognized as a predictable result of cryptic relatedness within sample groups.

That said, using comparisons of average and median relatedness per group, several revealing and consistent trends arise that could explain the presence and distribution of kin. These are most easily explained in light of a life history that divides juveniles and adults based on behavior, physical distribution, and fishing techniques.

The Importance of FADs to Juveniles

Yellowfin tuna in the WCPO reach maturity at roughly 100 cm fork length (Lehodey and Leroy, 1999). Below sizes of 60 cm, individuals are most commonly caught aggregating in association with objects at the surface, and are therefore susceptible to pole and line and purse seine fishing. As they grow, fish depend less on floating objects and spend more time in free swimming schools or at depth (Robert et al., 2012). This relationship holds true for the specimens observed in the current study.

When dyads were grouped based on catch event within each population sample, COANCESTRY and RelateAdmix results concurred that only FM09 consistently showed higher relatedness within a school than between schools in the same population, with RelateAdmix also finding the trend significant in PG13. These groups consist almost exclusively of small fish caught around either dFADs or aFADs, whereas all other population samples include more diverse catch event descriptions. The precise placement of individuals of interest at single FADs, more than between FADs as close as 20 km from each other, has precedent in the literature. Numerous studies have discussed the dynamics of yellowfin movement around FADs. Most fundamentally, modeling of tuna aggregations under floating objects acknowledges the importance of social dynamics and the presence of other fish, more than that of prey, in attracting tuna to a FAD (Robert et al., 2014). Synchronized movement to and from anchored FADs by multiple tagged individuals, interpreted as school membership, has been observed via acoustic tagging in the Philippines (Mitsunaga et al., 2013), the Maldives (Govinden and Jauhary, 2013), and Okinawa (Ohta and Kakuma, 2005). Other studies have focused more on the impacts of FADs, themselves; when fish were tracked within a network of FADs, more than 80% of fish showed strong FAD fidelity, never visiting other FADs only 20 km away over more than a 100 days of observation (Dagorn et al., 2007). Collectively, the literature suggests a moderately structured use of FADs, largely determined by behavioral mechanisms. The current study supports the prevailing view with clear evidence of kin aggregations around FADs.

Previous tagging studies have not been able to discern if the observed cohesion among caught individuals results more from site or school fidelity. Observations of fish returning exclusively to the site of tagging, even if it was hundreds of meters away from the associated FAD, suggest that yellowfin are spatially aware and capable of precise homing (Klimley and Holloway, 1999). Thus, site fidelity is a viable mechanism to explain synchronized movement in this species. However, our observations of kin around specific FADs in FM09 and PG13 provide evidence for persistence of schools, which confounds the importance of location. If individuals showed site fidelity without any form of social cohesion, larval cohorts would disperse across a range that incorporated adjacent FADs, such that highly related individuals would develop site fidelity to different FADs within an area, and show no significant difference in relatedness at the observed scale. Furthermore, the equal performance of drifting and anchored FADs in demonstrations of school-specific relatedness negates the exact importance of location, since drifting FADs, by definition, move over time.

Therefore, there are mechanisms at work that influence small yellowfin tuna to maintain strong social, as well as geographic, bonds over long periods. It is beyond the power of the current study to determine if the disproportionate association with closely related individuals is a result of kin recognition mechanisms and preference, or of general school fidelity that develops before larval cohorts disperse, and incidentally incorporates a majority of brood-mates. The latter has precedence, including observations by Govinden and Jauhary (2013) of instances in which small tagged yellowfin left FADs in synchrony with adult skipjack tuna, indicating that individual size and proximity override species differentiation to prompt schooling behavior. Regardless, these results suggest that yellowfin tuna may remain in close kin aggregations around FADs at least through their first year of life in the WCPO. This includes the 40–70 cm life stage when yellowfin tuna experience peak fishing pressure (Tremblay-Boyer et al., 2017).

The use of purse seining on young, closely related fish groups may disproportionately impact population genetic diversity. Persistent kin association may concentrate the presence of rare

RETRACTED

RETRACTED

alleles in single schools. Purse seining extracts large portions of a school in one catch event (Majkowski, 2003), and thus potentially removes entire families, and their rare alleles, from the population. This study therefore suggests careful management of young yellowfin tuna schools, during a particularly vulnerable life stage that both demonstrates family (and, consequently, population) structuring and is targeted by both intensive and non-selective fishing techniques.

It is interesting to note that, while drifting FADs attract schools composed of fish with the same degree of relatedness as fish around aFADs, the results are not comparable to those of other drifting objects. This could reflect the lower number of catch events that occur around natural drifting objects, which produces fewer opportunities to catch related fish. However, the linear regression and Spearman correlation suggested incomplete correlations between the two factors, leaving room for other explanations. Among these, dFADs differ from natural drifting objects in multiple ways that may carry significance. Foremost, drifting FADs are specifically constructed to attract fish by including sub-surface structure like tethered palm leaves (Scott and Lopez, 2014), which may not exist on or may decay and break off of natural drifting objects. Likewise, FADs can be repositioned into high-priority fishing areas if they drift into areas of low productivity (Davies et al., 2014). Therefore, based on location, there are more likely fish to catch around dFADs than natural drifting objects, in general, and large groupings of fish that are more likely to include siblings. Drifting FADs are also often deployed in locations with few natural objects, in efforts to prompt unassociated fish to begin schooling at the surface (Davies et al., 2014; Scott and Lopez, 2014). It is regrettable that we cannot confirm the importance of FADs compared to other surface objects, given the current distribution of samples, however we encourage deeper exploration of this question in the future.

The comparison of relatedness of schools found around FADs and that of free-swimming schools is also worth noting due to historical context. The exponential increase in the number of FADs present in the Pacific in the last three decades has greatly affected the distribution and habits of young tuna. Prior to 1990, most young yellowfin were found in unassociated schools, rather than in tight, FAD-centric units (Fonteneau et al., 2000; Davies et al., 2014). The relatedness of individuals found in free-swimming schools in this study is significantly lower than that found around FADs, suggesting that kin aggregations among young fish may not have always been as significant as we currently observe. This could have implications for our perception of the impacts of FADs on yellowfin tuna, and reinforces the need to carefully regulate FAD-based purse seining.

Site Fidelity in Adults

Adult yellowfin tuna live and are caught differently from juveniles. Like most tuna species, yellowfin tend to live deeper and school more loosely as they grow. Consequently, rather than being caught via pole and line or purse seining of surface schools, adults are susceptible to longline fishing. Longline fishing involves setting up to 3,000 baited hooks at depths of 50–300 m for 5–12 h over a hundred or more kilometers (FAO, 2003). Because of the extensive area covered and prolonged soak time, the exact time and location, and conditions under which each animal was caught cannot be precisely marked. Consequently, there is a tight agreement in our data between the presence of fish greater than one meter in fork length, the use of longlining, and a lack of data concerning the presence of any surface or other aggregating object like a seamount.

Unfortunately, the technicalities of longlining limit the observations we can make about adult trends in relatedness because, even though highly related fish were observed in the same hauls, it cannot be confirmed that the fish were schooling together before being caught. However, there is already evidence in the literature that large yellowfin demonstrate population structure at the school level. Klimley and Holloway (1999) observed fish up to 90 cm in length that displayed both very strong synchronized movements and site fidelity, with 73% of fish returning to their exact catch location, in the same company, even at the same time of day, numerous times over a year of observation.

Although our observations only circumstantially support Klimley and Holloway's observation of long term school fidelity, our data effectively demonstrate site fidelity that both carries implications for population genetic structure and could still inadvertently create close proximity among kin. Longlined (adult) fish have the same or higher mean and median relatedness than pole and lined (juvenile) specimens, overall, but no significant difference in mean or median relatedness of dyads from the same catch event or between catch events from the same EEZ and year. Of 225 half or full sib dyads that incorporate at least one longlined specimen, 49 were from the same catch event, 58 between catch events in the same year, and 92 from different years in the same EEZ. In the latter case, specimens displayed appropriate size disparity to be members of the same cohort, caught 2 years apart. Still another 24 dyads followed the same explanation but were from neighboring EEZs, representing an appropriately drastic but not absolute drop in relatedness with expanding geographic area. Of 225 relevant dyads, only 2 cannot be explained by continual residence in or repetitive return to an area of 500 km radius for more than 2 years. This level of site fidelity alone may explain the species' population structure through reduced migration and population connectivity, provided that spawning aggregations are composed of individuals that show similar patterns of regional fidelity. F_{ST} values calculated in this study indicate much stronger spatial than temporal population structure, which further supports the persistence of local populations with a limited influx of migrants.

Given strong and precise site fidelity as adults, long-term kin association in yellowfin tuna only requires young sibling groups to actively persist until an age when individuals develop loyalty to the same location. Diverse research has described mechanisms that could produce the hypothesized effect. Modeling studies have suggested that we currently underestimate the passive co-dispersal of marine planktonic larvae. For example, standard oceanographic features like eddies and fronts concentrate larvae into patches, rather than allowing them to diffuse evenly and

randomly across a region (Siegel et al., 2008). One tuna-specific model using Mediterranean dynamics observed a correlation between patchiness of larvae and the strength of mesoscale features, with the highest concentrations of larvae consistently found in proximity to a semi-stable salinity front (Mariani et al., 2010). Since eddies can be as small as 10 km across and can persist for weeks or months, and larvae from a single spawning event are likely to encounter the same eddy, related individuals can be retained within a few kilometers of each other throughout their larval phase merely through environmental forcing (Siegel et al., 2008). The process is admittedly random and does not guarantee co-dispersal, but does negate the impossibility of it in broadcast spawning species with long larval stages.

Meanwhile, yellowfin tuna life history includes numerous mechanisms that could build upon environmental factors to promote cohesion among related individuals beyond larval stages. Mature captive fish display mating behavior in which one female and a small number of males swim in tight circles while releasing gametes, creating a vortex that fertilizes most eggs with sperm from a limited paternal gene pool (Margulies et al., 2007). Compared to less structured forms of synchronized broadcast spawning, this creates a disproportionately high number of half- and full-siblings. The resulting fertilized eggs are buoyant, but sink and hatch in 18–28 h (Margulies et al., 2007), well within the 3 day particle de-correlation time estimated for off-shore oceanic conditions (Siegel et al., 2003). The original mating behavior also can produce sweepstakes reproductive success, which again promotes the disproportionate presence of kin in a cohort (Hedgecock and Pudovkin, 2011).

Larvae begin feeding (and thus displaying effective movement control) in less than 4 days (Kaji et al., 1999). In general, scombrids have remarkable nektonic development rates (Hunter, 1980). For example, skipjack tuna juveniles as small as 1 cm, corresponding to 10 days old, demonstrate competent swimming and depth control (Tanabe et al., 2017). Boehlert and Mundy (1994) observed a vertical distribution of *Thunnus* spp. larvae in which larger, more effective swimmers were found in higher-energy shallow waters, and smaller individuals were more common at depths greater than 20 m, where the environment tends to be more constant. There is also a long-running observation of increased larval concentrations around islands and landmasses, especially along leeward coasts where dispersive wind and wave energies are limited (Miller, 1979; Leis et al., 1991; Boehlert and Mundy, 1994). Furthermore, tuna demonstrate schooling behavior as both juveniles and adults, with size of fish inversely correlated to the strength of schooling behaviors (Goujon and Majkowski, 2000). If social behavior arises as soon as individuals are capable of nektonic movement, the strength of which can be extrapolated from juvenile and adult observations, there is a high probability of successful cohesion among brood-mates throughout larval development.

Therefore, a combination of passive environmental mechanisms and active biological impulses could plausibly result in larval and juvenile co-dispersal that persists into adulthood due to related individuals' overlapping site fidelity.

Micronesia as a Unique System

Alternatively, the most consistent and strong trends we describe all concern fish from the Federated States of Micronesia. It is possible that the FM location presents an exceptional circumstance that promotes both juvenile and adult aggregation, and that the observed genetic relatedness is not entirely relevant to other locations. Even with sibling groups removed, population differentiation analyses performed using putatively neutral loci suggest that FM09 is indeed genetically distinct (F_{ST} = 0.007 to 0.017). FM11 cannot be confidently evaluated due to the inappropriately small number of samples that remain after kin groups have been disassembled. When compared against dyads from single, FAD-associated catch events in Papua New Guinea, the median *r* from within FM09 catch events was still 130% larger and significant (p < 2.2e-16). A similar difference was measured between catch events from FM11 and NC14, as the only groups collected with longlines and without association data (p = 0.0002). Although the trend of larger median relatedness exists in FAD- and longline-associated dyads outside of the FM sample group, it is clearly magnified within the Federated States of Micronesia's EEZ.

The large size and topographic diversity of Micronesia makes it difficult to identify oceanographic characteristics that are both unique to and ubiquitous within the EEZ of this country. In general, FM is touted as "the most productive area for tuna production in the world" (Vali et al., 2014). The West Pacific Warm Pool and Pacific Equatorial Divergence converge in this area, making it highly attractive to tuna during El Niño Southern Oscillation (ENSO) neutral years, and even more so during La Niña events (Bell et al., 2011). Habitat selection modeling highlights the importance across numerous taxa of inconsistently attractive conditions in retaining individuals' site fidelity within unpredictable habitats (Switzer, 1993), which includes the oceanic conditions associated with ENSO inter-annual variability. The presence of many small islands also provides good inshore conditions for larval and young tuna. However, other island groups in the Pacific, including the Gilbert Islands, have similar topography and experience similar primary productivity levels, SSTs, and ENSO impacts. If the Micronesia tuna population is unique, the underlying mechanisms that drive the cohesiveness of it and the limited connectivity with other genetically differentiated demes are likely complex and interconnected, and incorporate behavioral, biological, and oceanographic influences.

Confirmation of increased tuna residency in Micronesia could have extensive implications. First, comparing a location's environmental conditions against those found in FM could help predict residency and migration rates. Second, the conservation measures recommended from the current analyses are largely based on trends epitomized by the Micronesian sample groups. If over-generalized, the recommended conservation measures could unnecessarily limit the industry in other parts of the Pacific.

Regardless of underlying mechanisms, the observations in this study indicate that instances occur in which yellowfin tuna school in close kin groups well after they would have dispersed

RETRACTED

if not actively maintaining proximity. Although modeling larval dispersal is very species- and region-specific (Largier, 2003), and vanishingly few studies exist for the pelagic environment, generalized algorithms suggest that cohorts can disperse across hundreds of kilometers just during yellowfin's month-long larval life stage (Shanks et al., 2003). Instead, we repeatedly found 50 cm fish [estimated age 10–12 months (Lehodey and Leroy, 1999)] still schooling with full siblings, and 100 cm fish in close proximity to both full- and half-sibs. While numerous tagging studies have observed yellowfin moving in tandem for prolonged periods, this is the first study to establish genetic relatedness within such groups.

Relying on externally designed software and algorithms potentially exposes this study to issues of automated, inappropriate model assumptions and the false positive identification of sibling pairs. However, studies that use custom algorithms have recently tested for genetic relatedness in other marine charismatic megafauna, including southern bluefin tuna (Bravington et al., 2016a) and great white sharks (Hillary et al., 2018). Unfortunately, the results cannot be compared directly, given the significant differences in life histories between species. Regardless, these studies demonstrate that it is not impossible to sample kin among highly migratory marine species. Hillary et al. (2018) found four full sib and 20 half sib dyads among 4950 possible combinations of sampled juvenile white sharks in Eastern Australia using 2186 SNPs and algorithms of relatedness likelihoods developed by the authors. Using similar methods, Bravington et al. (2016a) managed to find 45 parent-offspring pairs among 38,000,000 possible combinations of southern bluefin tuna. Downstream analyses that hinge on the identification of bluefin parent-offspring dyads were ultimately used in management decisions (Bravington et al., 2016b), indirectly supporting the quality of the results. Bravington et al. (2016a) also acknowledges the limitation of having used microsatellites for the analysis, and expresses confidence in the improved performance of kin recognition and automatic avoidance of false positive dyads with the use of 1000+ SNPs. It can be noted that the number of loci used in the current study exceeds this critical threshold.

However, the current study methodology is not without flaws. Superficially, the direct cross-validation of COANCESTRY and RelateAdmix software programs, which rely on different base algorithms, would seem an effective means of reducing Type I error. However, these efforts could still be undermined by our selection of kin group cut-off values, which relied on a single, pre-made algorithm in COANCESTRY. A more independent and user-controlled algorithm may produce alternative conclusions about the level of relatedness that defines a full- or half-sibling from more distant relationships, which would significantly impact our observed trends. Likewise, the selection criteria used as quality filters is also debatable. The current call rate and MAF, 99 and 5%, respectively, produce a final, all-sample dataset with 0.001% of genotypes missing and enough loci to produce confident analyses. While quality loss is miniscule, filtering at 100% call rate during preliminary data exploration resulted in half as many loci, which may produce noticeably different trends.

Finally, the current study was conducted in response to unusual results during an assessment of genetic population structure in yellowfin tuna in the Western and Central Pacific. Future analyses should use sample groups specifically selected to control variables that currently overlap, like the disproportionate representation of FAD-associated fish in FM. Likewise, they should regulate the number of individuals analyzed per catch event. And, given the potential implications of the current observations for fisheries management of a major market species, follow-up studies should take much larger samples that incorporate fish of more sizes, especially very small juveniles, and from more locations, years, and seasons within years. The proposed explanations are likely part of a much larger and highly complex biological and behavioral system, which deserves further exploration using specifically designed sampling methods.

CONCLUSION

We identified a total of 29 full sib dyads and 303 half sib dyads across seven sample groups of yellowfin tuna by cross-referencing results from two software programs, COANCESTRY and RelateAdmix. High levels of relatedness were observed within schools of small (<50 cm) fish associated with FADs, and among large fish caught by longline in open ocean conditions, although the exact proximity of kin cannot be assessed due to the collection method. The majority of all kin were collected within the waters of the Federated States of Micronesia.

Collectively, the data indicate that yellowfin tuna display strong school and FAD fidelity through their first year, which may generalize to regional fidelity as they mature. Close kin proximity in young animals could be an inadvertent outcome of stochastic collective dispersal, strong schooling instincts from the larval stage, overlapping site fidelity, massive variance in reproductive success resulting in a disproportionate number of closely related recruits, or a mix of all four. Direct recognition of and preference for associating with genetically similar fish is also possible, but requires more observations to confirm. Mature yellowfin tuna continue to demonstrate site fidelity, and possibly other forms of synchronized movement among kin. The observed close proximity of related fish at both life stages results in increased population structure and elevated susceptibility of population genetic diversity to purse seine fishing.

Alternatively, the observed trends may result from processes unique to the Federated States of Micronesia and should not be generalized to the wider WCPO. More relatedness studies, specifically targeting larvae and mature adults and increased sampling of all life stages outside FM, are needed to validate the observed trends.

Whatever the relevant geographic scale, the confirmation of persistent family groups in yellowfin tuna also challenges management assumptions of panmixia in this species [see Grewe et al. (2015)], and highlights the under-appreciated social dynamics available to cosmopolitan fishes.

RETRACTED

DATA AVAILABILITY

The datasets generated for this study can be retrieved from: https://doi.org/10.17605/OSF.IO/WTZD8.

AUTHOR CONTRIBUTIONS

CR wrote the research proposal and obtained funding. CR and GA designed the study. GA carried out the research, analyzed the data, and wrote the manuscript. ML confirmed data integrity and analyses. JH and NS, as agents of the Secretariat of the Pacific Community (SPC), critically assessed the results from a fisheries perspective. SPC was also responsible for sample collection, facilitated sample access, and prepared **Figure 1**. All authors contributed to the preparation of 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 number 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 bank samples; Simon Nicol and Valerie Allain for establishing the tissue bank; Sylvain Caillot 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. We are also grateful to Jinliang Wang and Anders Albrechtsen for their in-depth guidance with software programs, as well as Karuna Reddy for aiding with statistical analyses. Finally, we thank Carlo Pecoraro for reviewing earlier versions of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00341/full#supplementary-material

RETRACTED

REFERENCES

Aguila, R., Perez, S., Catacutan, B., Lopez, G., Barut, N., and Santos, M. (2015). Distinct yellowfin tuna (Thunnus albacares) stocks detected in western and central pacific ocean (WCPO) using DNA microsatellites. *PLoS One* 10:e138292. doi: 10.1371/journal.pone.0138292

Alexander, D., Shringarpure, S., Novembre, J., and Lange, K. (2015). *Admixture 1.3 Software Manual*. Available at: http://software.genetics.ucla.edu/admixture/admixture-manual.pdf (accessed August 05, 2018).

Astle, W., and Balding, D. (2009). Population structure and cryptic relatedness in genetic association studies. *Stat. Sci.* 24, 451–471. doi: 10.1214/09-STS307

Belkir, L., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. (2004). *1996-2004 GENETIX 4.05 Logiciel Sous Windows TM Pour la Genetique des Populations. Laboratoire Genome, Populations, Interactions*. Montpellier: Universite de Montpellier II.

Bell, J., Johnson, J., Ganachaud, A., Gehrke, P., and Hobday, A. (2011). *Vulnerability of Tropical Pacific Fisheries and Aquaculture to Climate Change. Summary for Pacific Island Countries and Territories*. Noumea: Secretariat of the Pacific Community.

Boehlert, G., and Mundy, B. (1994). Vertical and onshore-offshore distributional patterns of tuna larvae in relation to physical habitat features. *Mar. Ecol. Prog. Ser.* 107, 1–14. doi: 10.3354/meps107001

Bravington, M., Grewe, P., and Davies, C. (2016a). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nat. Commun.* 7:13162. doi: 10.1038/ncomms13162

Bravington, M., Skaug, H., and Anderson, E. (2016b). Close-kin mark-recapture. *Stat. Sci.* 31, 259–274. doi: 10.1214/16-STS552

Dagorn, L., Holland, K., and Itano, D. (2007). Behavior of yellowfin (Thunnus albacares) and bigeye (T. obesus) tuna in a network of fish aggregating devices (FADs). *Mar. Biol.* 151, 595–606. doi: 10.1007/s00227-006-0511-1

Davies, T. K., Mees, C. C., and Milner-Gulland, E. J. (2014). The past, present and future use of drifting fish aggregating devices (FADs) in the Indian Ocean. *Mar. Policy* 45, 163–170. doi: 10.1016/j.marpol.2013.12.014

Eldon, B., Riquet, F., Yearsley, J., Jollivet, D., and Broquet, T. (2016). Current hypotheses to explain genetic chaos under the sea. *Curr. Zool.* 62, 551–566. doi: 10.1093/cz/zow094

Ely, B., Viñas, J., Alvarado Bremer, J., Black, D., Lucas, L., Covello, K., et al. (2005). Consequences of the historical demography on the global population structure of two highly migratory cosmopolitan marine fishes: the yellowfin tuna (Thunnus albacares) and the skipjack tuna (Katsuwonus pelamis). *BMC Evol. Biol.* 5:19. doi: 10.1186/1471-2148-5-19

Excoffier, L., and Lischer, H. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under linux and windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

FAO (2003). *Fishing Technologies. Industrial Tuna Longlining. Technology Fact Sheets*. Available: http://www.fao.org/fishery/ (accessed August 02, 2018).

Fonteneau, A., Ariz, J., Gaertner, D., Nordstrom, V., and Pallares, P. (2000). Observed changes in the species composition of tuna schools in the Gulf of Guinea between 1981 and 1999, in relation with the fish aggregating device fishery. *Aquat. Living Resour.* 13, 253–257. doi: 10.1016/S0990-7440(00)01054-8

Goujon, M., and Majkowski, J. (2000). *Biological Characteristics of Tuna*. Rome: Fisheries and Aquaculture Department.

Govinden, R., and Jauhary, R. (2013). Movement behaviour of skipjack (Katsuwonus pelamis) and yellowfin (Thunnus albacares) tuna at anchored fish aggregating devices (FADs) in the Maldives. *Aquat. Living Resour.* 26, 69–77. doi: 10.1051/alr/2012022

Grewe, P., Feutry, P., Hill, P., Gunasekera, R., Schaefer, K., Itano, D., et al. (2015). Evidence of discrete yellowfin tuna (Thunnus albacares) populations demands rethink of management for this globally important resource. *Sci. Rep.* 5:16916. doi: 10.1038/srep16916

Hedgecock, D., and Pudovkin, A. (2011). Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bull. Mar. Sci.* 87, 971–1002. doi: 10.5343/bms.2010.1051

Hillary, R., Bravington, M., Patterson, T., Grewe, P., Bradford, R., Feutry, P., et al. (2018). Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Sci. Rep.* 8, 1–9. doi: 10.1038/s41598-018-20593-w

Hunter, J. (1980). *Feeding Ecology and Predation of Marine Fish Larvae*. Available at: http://swfsc.noaa.gov/publications/CR/1981/8122.PDF (accessed November 09, 2017).

ISSF (2017). *Status of the World Fisheries for Tuna: November 2017*. Washington, DC: International Seafood Sustainability Foundation.

Kaji, T., Tanaka, M., Oka, M., Takeuchi, H., Ohsumi, S., Teruya, K., et al. (1999). Growth and morphological development of laboratory-reared yellowfin tuna thunnus albacares and early juveniles, with special emphasis on the digestive system. *Fish. Sci.* 65, 700–707. doi: 10.2331/fishsci.65.700

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., et al. (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. *Data Prod. Anal. Popul. Genomics* 888, 67–89. doi: 10.1007/978-1-61779-870-2_5

Klimley, A., and Holloway, C. (1999). School fidelity and homing synchronicity of yellowfin tuna, Thunnus albacares. *Mar. Biol.* 133, 307–317. doi: 10.1007/s002270050469

Largier, J. (2003). Considerations in estimating larval dispersal distances from oceanographic data. *Ecol. Appl.* 13, S71–S89.

Lehodey, P., and Leroy, B. (1999). *Age and Growth of Yellowfin Tuna (Thunnus albacares) from the Western and Central Pacific Ocean as Indicated by Daily Growth Increments and Tagging Data*. Noumea: Secretariat of the Pacific Community.

Leis, J., Trnski, T., Harmelin-Vivien, M., Renon, J., Dufour, V., Moudni, K., et al. (1991). High concentration of tuna larvae (Pisces: Scombridae) in near-reef waters of French Polynesia (Society and Tuamotu Islands). *Bull. Mar. Sci.* 48, 150–158.

Macfadyen, G., and Defaux, V. (2016). *Estimate of the Global Sales Values from Tuna Fisheries - Phase 2*. Hampshire: Poseidon Aquatic Resource Management Ltd.

Majkowski, J. (2003). *Fishing Techniques - Tuna Purse Seining: Technology Fact Sheets*. Rome: FAO.

Margulies, D., Suter, J., Hunt, S., Olson, R., Scholey, V., Wexler, J., et al. (2007). Spawning and early development of captive Yellowfin Tuna (Thunnus albacares). *Fish. Bull.* 105, 249–265.

Mariani, P., MacKenzie, B., Iudicone, D., and Bozec, A. (2010). Modelling retention and dispersion mechanisms of bluefin tuna eggs and larvae in the northwest Mediterranean Sea. *Prog. Oceanogr.* 86, 45–58. doi: 10.1016/j.pocean.2010.04.027

Miller, J. (1979). Nearshore abundance of tuna (Pisces: Scombridae) larvae in the hawaiian islands. *Bull. Mar. Sci.* 29, 19–26.

Mitsunaga, Y., Endo, C., and Babaran, R. (2013). Schooling behavior of juvenile yellowfin tuna thunnus albacares around a fish aggregating device (FAD) in the Philippines. *Aquat. Living Resour.* 84, 79–84. doi: 10.1051/alr/2012031

Moltke, I., and Albrechtsen, A. (2014). RelateAdmix: a software tool for estimating relatedness between admixed individuals. *Bioinformatics* 30, 1027–1028. doi: 10.1093/bioinformatics/btt652

Ohta, I., and Kakuma, S. (2005). Periodic behavior and residence time of yellowfin and bigeye tuna associated with fish aggregating devices around Okinawa Islands, as identified with automated listening stations. *Mar. Biol.* 146, 581–594. doi: 10.1007/s00227-004-1456-x

Pecoraro, C., Zudaire, I., Bodin, N., Murua, H., Taconet, P., Díaz-Jaimes, P., et al. (2016). Putting all the pieces together: integrating current knowledge of the biology, ecology, fisheries status, stock structure and management of yellowfin tuna (Thunnus albacares). *Rev. Fish Biol. Fish.* 27, 811–841. doi: 10.1007/s11160-016-9460-z

Pudovkin, A., Zhdanova, O., and Hedgecock, D. (2010). Sampling properties of the heterozygote-excess estimator of the effective number of breeders. *Conserv. Genet.* 11, 759–771. doi: 10.1007/s10592-009-9865-5

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ren, R., Ray, R., Li, P., Xu, J., Zhang, M., Liu, G., et al. (2015). Construction of a high-density DArTseq SNP-based genetic map and identification of genomic regions with segregation distortion in a genetic population derived from a cross between feral and cultivated-type watermelon. *Mol. Genet. Genomics* 290, 1457–1470. doi: 10.1007/s00438-015-0997-7

Robert, M., Dagorn, L., and Deneubourg, J. (2014). The aggregation of tuna around floating objects: what could be the underlying social mechanisms? *J. Theor. Biol.* 359, 161–170. doi: 10.1016/j.jtbi.2014.06.010

Robert, M., Dagorn, L., Deneubourg, J., Itano, D., and Holland, K. (2012). Size-dependent behavior of tuna in an array of fish aggregating devices (FADs). *Mar. Biol.* 159, 907–914. doi: 10.1007/s00227-011-1868-3

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., and Grattapaglia, D. (2011). Diversity arrays technology (DArT) and next- generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proc.* 5(Suppl. 7):54. doi: 10.1186/1753-6561-5-S7-P54

Scott, G., and Lopez, J. (2014). *The Use of FADs in Tuna Fisheries*. Available at: https://www.wcpfc.int/system/files/IPOL-PECH_NT%282014%29514002_EN.pdf (accessed December 06, 2018).

Selwyn, J., Hogan, J., Downey-Wall, A., Gurski, L., Portnoy, D., and Heath, D. (2016). Kin-aggregations explain chaotic genetic patchiness, a commonly observed genetic pattern, in a marine fish. *PLoS One* 11:e153381. doi: 10.1371/journal.pone.0153381

Shanks, A., Grantham, B., and Carr, M. (2003). Propagule dispersal distance and the size and spacing of marine reserves. *Ecol. Appl.* 13, S159–S169.

Sibert, J., and Hampton, J. (2003). Mobility of tropical tunas and the implications for fisheries management. *Mar. Policy* 27, 87–95. doi: 10.1016/S0308-597X(02)00057-X

Siegel, D., Kinlan, B., Gaylord, B., and Gaines, S. (2003). Lagrangian descriptions of marine larval dispersion. *Mar. Ecol. Prog. Ser.* 260, 83–96. doi: 10.3354/meps260083

Siegel, D. A., Mitarai, S., Costello, C. J., Gaines, S. D., Kendall, B. E., Warner, R. R., et al. (2008). The stochastic nature of larval connectivity among nearshore marine populations. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8974–8979. doi: 10.1073/pnas.0802544105

Suzuki, Z., Tomlinson, P., and Honma, M. (1978). Population structure of pacific yellowfin tuna. *Inter Am. Trop. Tuna Commission Bull.* 17, 274–441.

Switzer, P. (1993). Site fidelity in predictable and unpredictable habitats. *Evol. Ecol.* 7, 533–555. doi: 10.1007/bf01237820

Tanabe, T., Kiyofuji, H., Shimizu, Y., and Ogura, M. (2017). Vertical distribution of juvenile skipjack tuna katsuwonus pelamis in the tropical western pacific ocean. *Jpn. Agric. Res. Q.* 51, 181–189. doi: 10.6090/jarq.51.181

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017). *Stock Assessment of Yellowfin Tuna in the Western and Central Pacific Ocean*. Available at: https://www.wcpfc.int/system/files/SC13-SA-WP-06%20YFT-stock-assessment_2017_REV1.pdf

Vali, S., Rhodes, K., Au, A., Zylich, K., Harper, S., and Zeller, D. (2014). *Reconstruction of Total FIsheries Catches for the Federated States of Micronesia (1950-2010). Fisheries Centre Working Paper #2014-06*. Vancouver: University of British Columbia.

Walker, M., Kirschvink, J., Chang, S., and Dizon, A. (1984). A candidate magnetic sense organ in the yellowfin tuna, thunnus albacares. *Science* 224, 751–753. doi: 10.1126/science.224.4650.751

Wang, J. (2011). coancestry: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol. Ecol. Resour.* 11, 141–145. doi: 10.1111/j.1755-0998.2010.02885.x

Waples, R., and Allendorf, F. (2015). "Testing for hardy-weinberg proportions: have we lost the plot?" *J. Hered.* 106, 1–19. doi: 10.1093/jhered/esu062

Ward, R., Elliott, N., Innes, B., Smolenski, A., and Grewe, P. (1997). Global population structure of yellowfin tuna, Thunnus albacares, inferred from allozyme and mitochondrial DNA variation. *Fish. Bull.* 95, 566–575.

Weir, B., and Cockerham, C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x

Wild, A., and Hampton, J. (1993). "A review of the biology and fisheries for skipjack tuna, Katsuwonus pelamis, in The Pacific Ocean," in *Interactions of Pacific tuna Fisheries*, Vol. 2, eds R. Shomura and J. Majkowski (Rome: FAO).

Conflict of Interest Statement: 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.

Copyright © 2019 Anderson, Lal, Hampton, Smith 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.

RETRACTED