

11th MEETING OF THE SCIENTIFIC COMMITTEE

11 to 16 September 2023, Panama City, Panama

SC11 – JM14

**Population genomics and environmental associations in *Trachurus murphyi* in
the South Pacific Ocean**

Republic of Chile

Population genomics and environmental associations in Chilean jack mackerel (*Trachurus murphyi*) in the South Pacific Ocean

Sandra Ferrada Fuentes, Laboratorio de Genética y Acuicultura, Departamento de Oceanografía, Universidad de Concepción, Chile. Programa de Doctorado en Sistemática y Biodiversidad, Universidad de Concepción, Chile, sferrada@udec.cl

Cristian B. Canales-Aguirre, Centro i~mar, Universidad de Los Lagos, Camino a Chiquihue 6 km, Puerto Montt, Chile, cristian.canales@ulagos.cl

Victoria Herrera Yáñez, Laboratorio de Genética y Acuicultura, Departamento de Oceanografía, Universidad de Concepción, Chile.

Aquiles Sepúlveda, Instituto de Investigación Pesquera (INPESCA), Chile.

Sebastián I. Vásquez, Instituto de Investigación Pesquera (INPESCA), Chile.

Abstract

The assessment of the genetic structuring of biodiversity is crucial for management and conservation. For species with large population sizes and migratory behaviour, a low number of molecular markers may fail to identify population structure. A solution of this shortcoming can be high-throughput sequencing that allows genotyping thousands of markers on a genome approach while facilitating the detection of genetic structuring shaped by selection. This is the case of Chilean Jack mackerel, *Trachurus murphyi*, which is a pelagic fish widely distributed in the South Pacific Ocean and one of the commercially most relevant resources for Chile. In this study, we used high-quality biallelic SNPs and mitochondrial haplotypes from 376 samples collected from 35 localities, to investigate its genetic population structure across the South Pacific Ocean. We found low population structure at neutral loci, but high differentiation at adaptive loci distinguishing a location at New Zealand from the other locations of the Pacific Ocean. Associations between adaptive and neutral genetic distance with environmental distance were evaluated. Our results reveal a pattern of spatial genetic divergence between adaptive loci, probably reflecting adaptations to local environments such as turbulence and mesoscale activity and secondarily with aspects of biological productivity such as chlorophyll-a, and no evidence for differences related to sea surface temperature. Overall, the results obtained suggest that Chilean Jack mackerel shows population structure and adaptation despite considerable gene flow in the South Pacific Ocean. The connectivity (i.e. gene flow) and environmental variables plays a key role for the contrasting patterns of spatial structure found for neutral and adaptive loci.

Keywords

Population genomics, environmental association, management fisheries, South Pacific Ocean, *Trachurus murphyi*

Introduction

The Chilean Jack mackerel (*Trachurus murphyi* Nichols, 1920) is a widely distributed species in the South Pacific Ocean and one of Chile's most commercially relevant pelagic resources. Its extraction is focused on three main fishing areas: A northern zone (18°21' - 24°00' SE), the Caldera-Coquimbo zone (24°00'-32 00'S), and the south-central zone (32°00' - 43°30' S). Furthermore, next to the south-central zone, but outside the Exclusive Economic Zone (EEZ), it is exploited by foreign fleets. In South America, this species is fished at the coastal areas off Ecuador, Perú and Chile. The Chilean Jack mackerel fisheries management in the South Pacific Ocean is carried out in the context of the South Pacific Regional Fisheries Management Organisation (OROP-PS/SPRFMO). The Scientific Committee of the OROP-PS/SPRFMO carries out an annual assessment and determination of the state of exploitation of this resource for the whole of the South-East Pacific. Chile's management of jack mackerel in its EEZ follows the precautionary approach and the principle of compatibility, also following the management recommendations of the OROP-PS/SPRFMO. In this context, in the "9TH SCIENTIFIC COMMITTEE MEETING" of the SPRFMO held in September 2021, Chile agreed to lead a study that aims at genetic monitoring of the population structure of Chilean Jack mackerel in the South Pacific Ocean (SPO).

The knowledge of the population genetics of *T. murphyi* in the SPO have been based, so far, using several methodologies i) Protein electrophoresis, ii) PCR (nuclear and mitochondrial genes, microsatellite data), and iii) Sanger sequencing. Chronologically, Galleguillos and Torres (1988) identified polymorphic enzyme loci in samples from Chile (Chiloé, Talcahuano, Juan Fernández, Iquique), Perú, and one oceanic area (39°24'S; 76°45'W), not detecting significant differences between them. Later, Arancibia et al. (1996), using 23 enzymatic loci for samples from Chile, also failed to observe differences. Sepúlveda et al. (1996), using PCR-RFLP in ITS2 (Internal Transcribed Spacer) with the *MspI* enzyme and analyzing samples from Chile (Isla Mocha, Iquique, Juan Fernández), New Zealand, and Australia, showed no significant genetic differences among the samples, and patterns of genetic homogeneity were observed in the study area for these genetic markers. Cardenas et al. (2009) by analyzing mitochondrial DNA (control region 772 bp) and nuclear DNA (four heterologous microsatellite loci) for samples from the eastern South Pacific Ocean and western South Pacific Ocean (Chile, New Zealand, open Pacific Ocean) indicated low genetic variability for both types of molecular markers, with no significant genetic differences between localities. Subsequently, the FIP 2007-27 project, led by Serra et al. 2010, adopted a multidisciplinary approach, considering a set of methods: genetics, parasites, morphometry (body and otoliths), otolith microchemistry, and natural history patterns. More than 1,020 samples from the geographical distribution of *T. murphyi* were analyzed, with samples from Chile (7 localities) and the collaboration of the Instituto del Mar del Perú (IMARPE; three localities) and the Ministry of Fisheries of New Zealand (1 locality). The genetic analysis used six heterologous microsatellite loci (Cárdenas et al. 2009) and three species-specific loci (Canales-Aguirre et al. 2010). These microsatellites showed no differences among the localities analyzed. Then, Serra et al. (2014), through the FIPA N°2010-18 project, extended temporal and spatially the previous study of Serra et al. 2010, again using a multidisciplinary approach. For the genetic analysis, ten microsatellite loci (three heterologous and seven species-specific) were used for 852 individuals from eight localities of the South Pacific Ocean distribution of *T. murphyi* (3 localities from Perú, four localities from Chile, and one locality from New Zealand). In addition, comparisons were made by season (spring-summer). The FST index for localities and sampling seasons showed no evidence of genetic structuring (low and non-significant FST). For the temporal analysis, between samples

from the FIPA N°2007-27 and FIPA N° 2010-18 projects, the F_{ST} index showed no significant genetic differences. Ferrada et al., (2023) re-analyzed the estimating genetic diversity and population structure with polymorphic markers of Chilean jack mackerel, *Trachurus murphyi*, in the South Pacific Ocean. A total of 522 samples were collected from 11 locations in 2008, to be analyzed using six loci microsatellites. The results showed a high genetic diversity across locations ($H_o = 0.551$ to 0.980 ; $H_e = 0.676$ to 0.959). Estimates for the population structure showed a low and non-significant pairwise F_{ST} in all comparisons. We supported the non-genetic differentiation previously reported with the used microsatellite panel. Further comparative temporal studies should be conducted to identify the stability of this pattern. Overall, this study reinforces the hypothesis that, in the South Pacific Ocean, Chilean Jack mackerel correspond to a large single population.

All these previous studies indicates that *T. murphyi*, using several molecular markers, does not shown population structure in their entire geographic distribution. Nonetheless, single nucleotides polymorphisms are still not used, and there are several studies where they identified structure where previous markers were not able to recognize some structure. In this study, we investigated the population genomics of *T. murphyi* across their entire geographic distribution in the South Pacific Ocean using Single Nucleotide Polymorphism (SNPs) obtained from DArTseq. For SNPs we assessed both neutral and adaptive variation. We specifically aim to (i) evaluate genetic diversity, population differentiation between areas, and estimated the relative migration rates, (ii) correlate putative adaptive loci to oceanographic environmental variables, and (iii) discuss the implications of our results for conservation and management of the species.

Materials and methods

Sample collection

Fresh samples of Chilean Jack mackerel were collected across the South Pacific Ocean during 2022. Localities sampled along South Pacific Ocean included samples from Chile (FIPA N°2021-28). For those samples, a small piece of tissue was excised of each specimen and stored at 4°C in absolute ethanol for further molecular analyses. To complement the samples and extend temporal and geographic coverage of their distribution, we gathered samples from two previous projects. These samples came from two projects funded by Chilean government (i.e., FIPA N°2007-27 and FIPA N°2010-18). Finally, we got a total of 368 individual from 35 sampling locations (Figure 1). Given that sampling size was uneven by sampling location, we created new areas gathered locations (see caption Figure 1) for further analyses that sample size ranged between 14 and 45 (Figure 1).

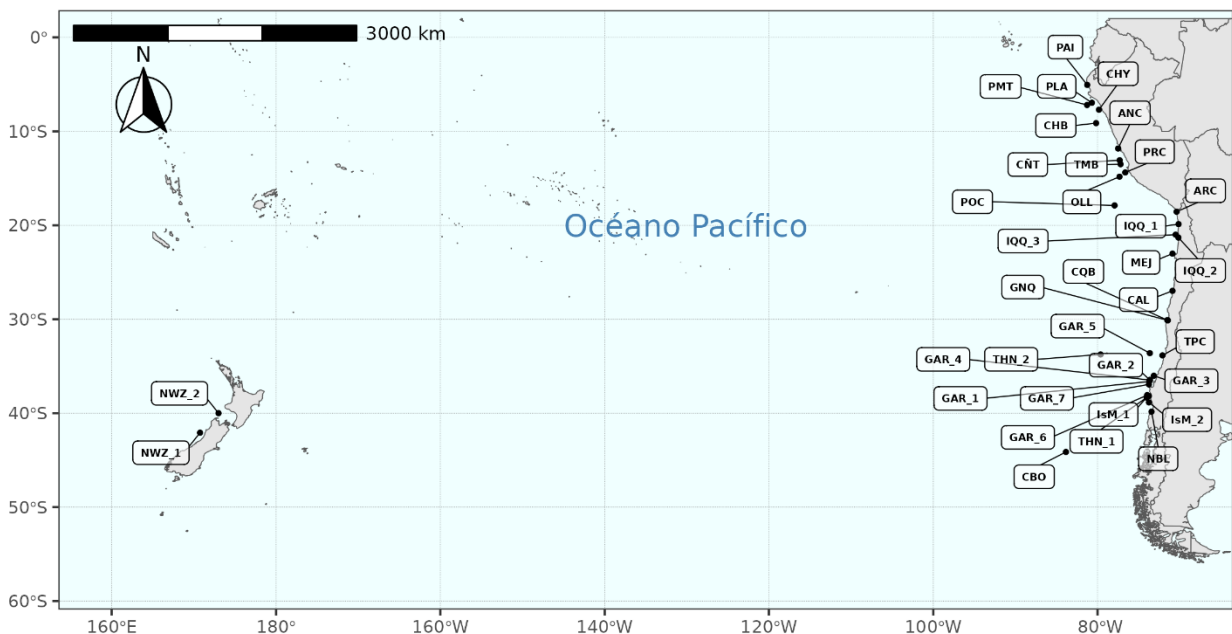


Figure 1. Sampling locations of the Chilean Jack mackerel, *Trachurus murphyi*, in the South Pacific Ocean. The areas were classified as follows: **PerA**: PAI, PLA, PMT (n=25); **PerB**: CHY, CHB, ANC, CÑT, TMB (n=25); **PerC**: OLL, PRC (n=20); **PerD**: POC (n=18); **ChiA**: ARC, IQQ_1 (n=28); **ChiB**: IQQ_2, IQQ_1 (n=30); **ChiC**: MEJ (n=15); **ChiD**: CAL (n=31); **ChiE**: CQB, GNQ (n=14); **ChiF**: TOP, GAR_5 (n=26); **ChiG**: GAR_1, GAR_2, GAR_3, GAR_4, GAR_7 (n=31); **ChiH**: GAR_6, IsM_1, IsM_2, THN_1 (n=31); **ChiI**: NBL (n=14); **ChiJ**: THN2 (n=15); **ChiK**: CBO (n=15); **NWZ**: NWZ_1, NWZ_2 (n=30). **Per**=Perú, **Chi**= Chile, and **NWZ**=New Zealand.

DNA purification for SNPs and Reduced representation SNP genotyping: DArTseq

For all samples, the total genomic DNA was extracted by the Diversity Arrays Technologies Pty Ltd, Australia. The isolation of genomic DNA was conducted using proteinase K for digest tissue and NucleoMag B-Beads for DNA binding, all include in NucleoMag Tissue kit (Macherey-Nagel).

We used DArTseq™ as approach to reduce the complexity of the genome for *Trachurus murphyi*. In brief, DArTseq™ represents a combination of a DArT complexity reduction methods and next generation sequencing

platforms (Sansaloni et al, 2011; Kilian et al, 2012; Courtois et al, 2013; Raman et al. 2014; Cruz et al. 2013). The Diversity Arrays Technology Pty Ltd selected the PstI-SphI method for *T. murphyi* and followed DNA samples were processed in digestion/ligation reactions principally as per Kilian et al (2012) but replacing a single PstI-compatible adaptor with two different adaptors corresponding to two different Restriction Enzyme (RE) overhangs. Sequences generated from each lane were processed using proprietary DArT analytical pipelines and first filtering provided by DArTseq™ was performed on the raw sequences using the following parameters: Min Phred pass score 30, Min pass percentage 75, for Barcode region and Min Phred pass score 10, Min pass percentage 50 for the whole read. A total of 48,065 SNPs were obtained from DArTseq quality controls.

Re-filtering and SNPs calling and Identification of potentially neutral and adaptive loci

Following the DArTseq quality controls, we conducted more stringent filters to remove sex linked loci, poor quality loci and samples. After remove sex-linked loci, we removed loci that i) did not reach 99% reproducibility, ii) had a calling rate per locus lower than 0.9, iii) individuals that showed a calling rate lower than 0.8, and iv) all monomorphic loci generated subsequent to the filtering of individuals. All these filters were performed with the dartR v2.7.2 package (Mijangos et al. 2022). To evaluate whether neutral or selective process better explains the observed patterns of genome differentiation, we separate on two SNPs subsets or database. We identified neutral and adaptive loci using three R packages: i) pcadapt, ii) OutFLANK, and a iii) redundancy analysis (RDA). These programs were chosen because they have different assumptions, which allowed us to increase confidence in identifying adaptive loci. The pcadapt and RDA was conducted similarly as described by Canales-Aguirre et al. 2022, with exception for RDA. In our case, we did not use environmental variables, but used spatial variation, i.e. the geographic coordinates of the centroids of each cluster of localities. Latitude and longitude were the predictor variables used. For OutFLANK, we used the function gl.outflank included in the dartR package that identifies loci under selection per population using the method of Whitlock and Lotterhos (2015). We used the default described parameters. Finally, all unique putative adaptive loci (hereafter called adaptive) obtained from pcadapt, OutFLANK and RDA were filtered from the 12,072 loci database and included in a single dataset resting loci correspond to the potentially neutral loci (hereafter called neutral). The neutral database have 11,588 and the adaptive 484 loci for 332 individuals. Both neutral and adaptive databases were used for subsequent analyses of genetic diversity and population structure.

Genetic diversity and population structure for SNPs

We estimated the following summary statistics of genetic diversity by locality, observed heterozygosity (H_o), expected heterozygosity (H_e), private number of alleles (P_A) and inbreeding coefficient (F_{IS}). All these parameters were estimated in the hierfstat v0.04-10 package (Goudet 2005).

Three approaches were used to estimate the population structure of the neutral and adaptive data obtained: i) a pairwise F_{ST} , ii) a discriminant analysis of principal components (DAPC), and iii) a Bayesian genetic clustering analysis. i) The pairwise F_{ST} values were calculated for the pairs of populations, and we performed significance tests using 10,000 permutations in the STAMPP package (Pembleton et al. 2013). ii) The DAPC was performed in the R package adegenet v.2.1.10 (Jombart 2008; Jombart and Ahmed 2011). It summarizes the genetic differentiation between areas, minimizing inter-areas variation, achieving the best discrimination of individuals into predefined groups (i.e., areas). iii) The Bayesian analysis of genetic clustering was conducted using the program STRUCTURE v.2.3.3 (Pritchard et al. 2000; Falush et al. 2003). To identify the most

probable number of populations (K), we tested a range of K between 1 and 6 ($K = 1 - 6$; maximum number observable in DAPC), where $K=1$ means that the genetic data account for a single population (no structuring) and $K=n$ means that the data account for " n " populations in the analyzed dataset. The parameters used were: a model with admixture, of no admixture, and with correlated allele frequencies, each K was tested for five replicates independently, 50000 initial iterations not used in the analysis (burning), and 100000 iterations of analysis sampling every 1000. All analyses were performed using the areas as a priori information to improve structure detection when these could be weak (Hubisz et al. 2009). To identify which tested K is the most likely in the dataset, Evanno's ΔK index was estimated (Evanno et al. 2005).

Relative migration estimations

We calculated the asymmetrical bidirectional migration rates among areas using the `divMigrate` function (Sundqvist et al., 2016) included in the `diveRcity` package (Keenan et al., 2013). We estimated the migration rates using the effective number of migrants index (N_m ; Alcalá et al., 2014). We used 1000 bootstraps to test the statistical significance of directional migration. We conducted this procedure following Canales-Aguirre et al. (2022). This approach was conducted for the whole data set, as well as the three sampling periods each (i.e. 2007, 2010, and 2021). This was conducted for the neutral and adaptive datasets.

Environmental variables

To characterize environmental conditions that approximate the ecological niche of Chilean jack mackerel throughout the South Pacific Ocean, five satellite-derived bio-oceanographic variables were analyzed in $2^\circ \times 2^\circ$ quadrats around each genetic sampling point. Monthly time series were obtained for each quadrant. Sea surface temperature was retrieved from the Ostia-Copernicus product (Good et al., 2020) at 4 km horizontal resolution for the period 1985-2022. Sea surface chlorophyll, as a proxy for phytoplankton biomass, was retrieved from the 4 km horizontal resolution Copernicus-GlobColour product for the period 1997-2022 from when ocean color information is available. To cover dynamic aspects of the ocean, the mean sea level anomaly provided by the Data Unification and Altimeter Combination System (DUACS) for the period 1993-2022 with $0.25^\circ \times 0.25^\circ$ horizontal resolution was analyzed. From the geostrophic currents derived from sea level anomalies, the Eddy Kinetic Energy (EKE) was calculated as a proxy for the mesoscale activity of the studied regions. Finally, to investigate aspects of ocean-atmosphere interaction, the wind-induced-turbulence calculated as the cube of the wind speed was analyzed from the ERA5 hourly reanalysis product at 31 km of horizontal resolution for the period 1985-2022. (Hersbach et al., 2020).

Environmental and genetic distances

Gower distance (dissimilarity) matrices (Gower, 1971) were calculated for each environmental variable, using the `StatMatch` package for R. The environmental dissimilarity matrix was related to the matrix of genetic, adaptive and neutral distances, following Cavalli-Sforza and Edwards (1967) using the Mantel test (Mantel, 1967) to address the statistical relationship between distances from different locations using the `vegan` package for R. Finally, these relationships were used to explore the association between environmental distance and distance genetics of the sampled areas along the South Pacific Ocean.

Results

Re-filtering and SNPs calling and Identification of potentially neutral and adaptive loci

The initial quality filters performed by DArTseq PL retained a total of 354 individuals and a total of 48,065 loci. Subsequent filters excluded: i) 1,075 sex chromosome loci, ii) 14,298 loci with reproducibility values less than or equal to 0.99, iii) 17,163 loci had a per locus call rate less than 0.9, iv) 22 individuals that had a per individual call rate less than 0.8, and vi) 3,457 monomorphic (i.e., non-informative) loci. The database for subsequent identification of neutral and adaptive loci was 12,072 loci in 332 individuals.

The three packages used to identify adaptive loci (i.e., pcadapt, OutFLANK, and RDA) identified different numbers of loci. The pcadapt package identified 28 loci, whereas OutFLANK 7 loci and RDA 451 loci. No loci were shared by the three R packages, however, one locus was shared between OutFLANK-RDA and one locus between pcadapt-RDA. Putative adaptive loci (hereafter called adaptive) were 484 loci, while the putative neutral loci (hereafter called neutral) were 11,588 loci.

Genetic diversity and population structure for SNPs

The genetic diversity was similar among areas, with some small differences (Table S1). For the neutral data set, the observed heterozygosity (H_o) ranged from 0.0284 (ChiA and ChiC) to 0.0339 (Chil). Expected heterozygosity (H_e) ranged from 0.0348 (NWZ) to 0.0384 (ChiK). The inbreeding coefficient (F_{IS}) ranged from 0.0648 (Chil) to 0.1470 (PerD). The number of private alleles (P_A) ranged from 0 (NWZ) to 659 (ChiB). For the adaptive dataset, the observed heterozygosity (H_o) ranged from 0.0180 (Chil) to 0.0510 (NWZ). Expected heterozygosity (H_e) ranged from 0.0138 (ChiC) to 0.0693 (NWZ). The inbreeding coefficient (F_{IS}) ranged from 0.0605 (Chil) to 0.2621 (PerD). The number of private alleles (P_A) ranged from 0 (Perú and Chile, except PerA with 3, Chil and ChiK with 4 each) to 450 (NWZ). Overall, the number of private alleles for the New Zealand samples in both datasets showed considerable qualitative differences, with 0 for the neutral data and 450 for the adaptive data.

The neutral dataset showed low pairwise F_{ST} values (Table S2), where significant comparisons were identified between areas Chil, ChiJ, and ChiK with the rest of the areas tested. NWZ showed differences only with these same three areas but not with the rest of Chile and Perú. The adaptive dataset also showed low pairwise F_{ST} values, although higher than the neutral data (one-fold increase) (Table S2). Significant differences were identified between the NWZ area and the rest of the areas of Chile and Perú, except for the Chil, ChiJ, and ChiK areas.

The discriminant analysis of principal component (DAPC) showed differences between neutral and adaptive dataset (Figure 2). The neutral dataset showed a pattern where at least four clusters are identified, i) one that included areas from Perú, Chile (except three areas) and New Zealand (i.e. central cloud of data points Figure 3A), ii) another that included individuals from Chil, iii) another that included individuals from ChiJ, and iv) one that included individuals from area ChiK. The variance explained in the neutral dataset was 14.1% for the first, while 12.3% for the second axis. In contrast, the adaptive dataset showed a clear pattern of two areas (Figure 2C), i) one included individuals from Perú and Chile and ii) one included individuals from New Zealand. The variance explained by the main axes in the adaptive dataset was 82.6% for the axis 1, while 3.9% for the axis 2.

Bayesian analysis for the neutral data set showed a $K=2$ but not differentiate any area (Figure 2B), while for the adaptive data set that showed also a $K=2$ but differentiate NWZ from the rest (Figure 2D).

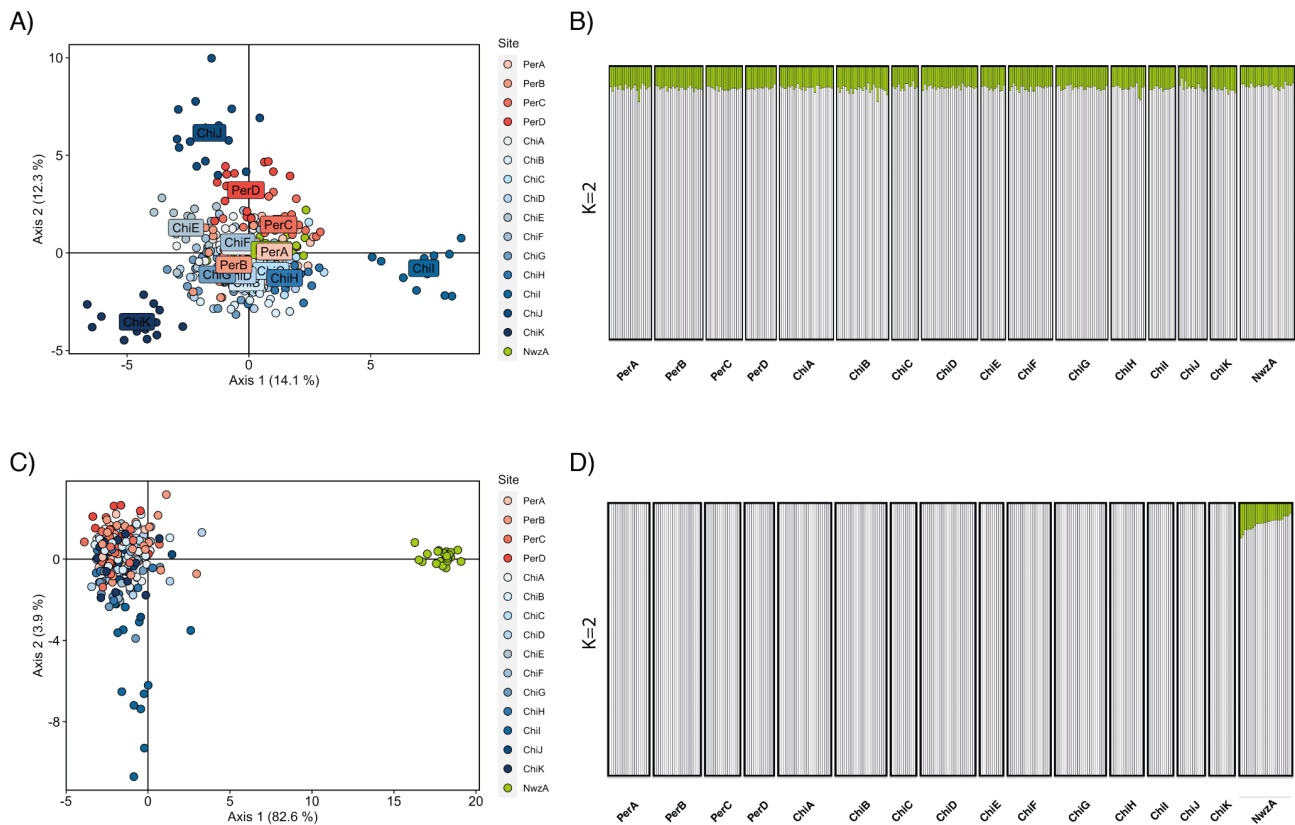
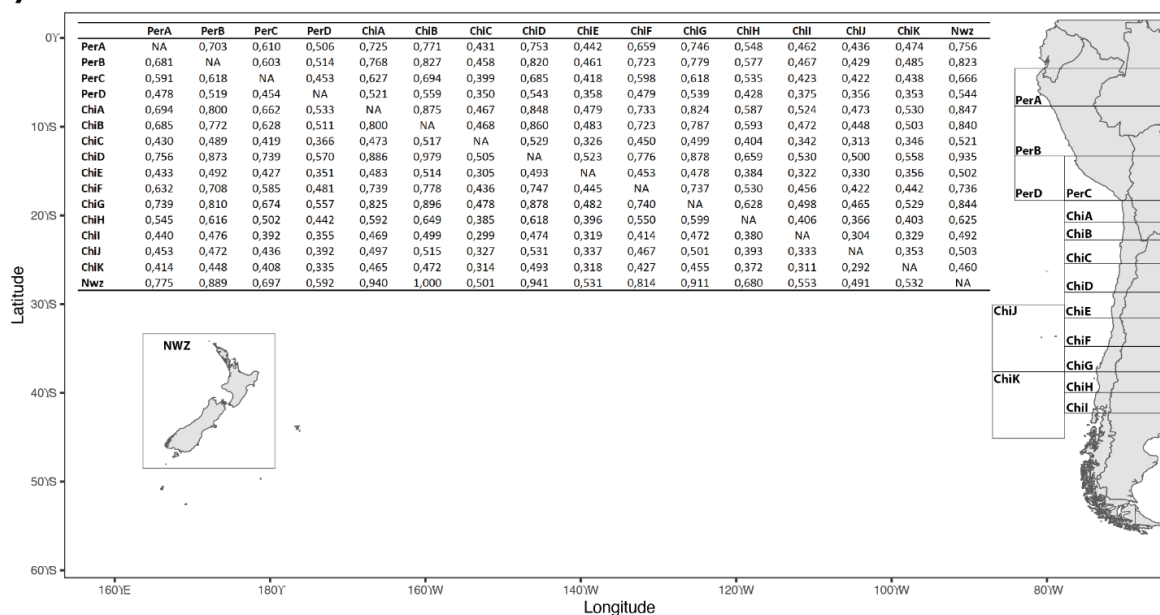


Figure 2. Results of discriminant analysis of principal component and Bayesian analysis. For DAPC, the scatterplot showed individual genotypes using information from 11588 neutral loci (A) and 484 adaptive loci (C). Each circle corresponds to an individual while the different colors correspond to the different areas used in the analysis. Both diagrams show the percentages of variance explained by axes 1 and 2 of the analysis. For the Bayesian analysis the vertical bars represent a vector of membership coefficients (Q values) of individual genotypes to the most probable number of genetic clusters using an admixture model (K=2) for (B) neutral and (D) adaptive dataset. Gray bars include all areas of Chile (ChiA - K) and Perú (PerA - D), while green bars include the New Zealand area (NWZ). The white vertical lines separate the different areas analyzed. The acronym for each area is below the bar plot.

Relative migration estimations

The bidirectional relative migration rates for neutral loci showed minimum values of 0.292 and maximum of 1. The average was 0.34 and the median was 0.296 (Figure 3A). No significant differences were observed in any of the comparisons, indicating the absence of significant asymmetrical migration rates. Conversely, the adaptive data set showed minimum values of 0.103 and maximum of 1. The average was 0.552 and the median was 0.503 (Table S3). Significantly different directional migration estimates occurred between locations from Perú and Chile (i.e., PerA, PerB, PerC, ChiA, ChiB, ChiD, ChiF, ChiG, and ChiH) to New Zealand (Table S3). Get together the samples as regions used to compare Chilean Jack mackerel biomass trends (Dragon et al., 2018), the neutral loci showed minimum values of 0.15 and maximum of 1. The average was 0.418 and the median was 0.36 (Table S4) (Figure 3B). Also, significant directional migration estimation was found from regions 1 - 6b to 5 (New Zealand). Estimations for bidirectional relative migration rates for year 2007 (Table S5), 2010 (Table S6), and 2021 (Table S7) for neutral and adaptive loci showed similar values.

A)



B)

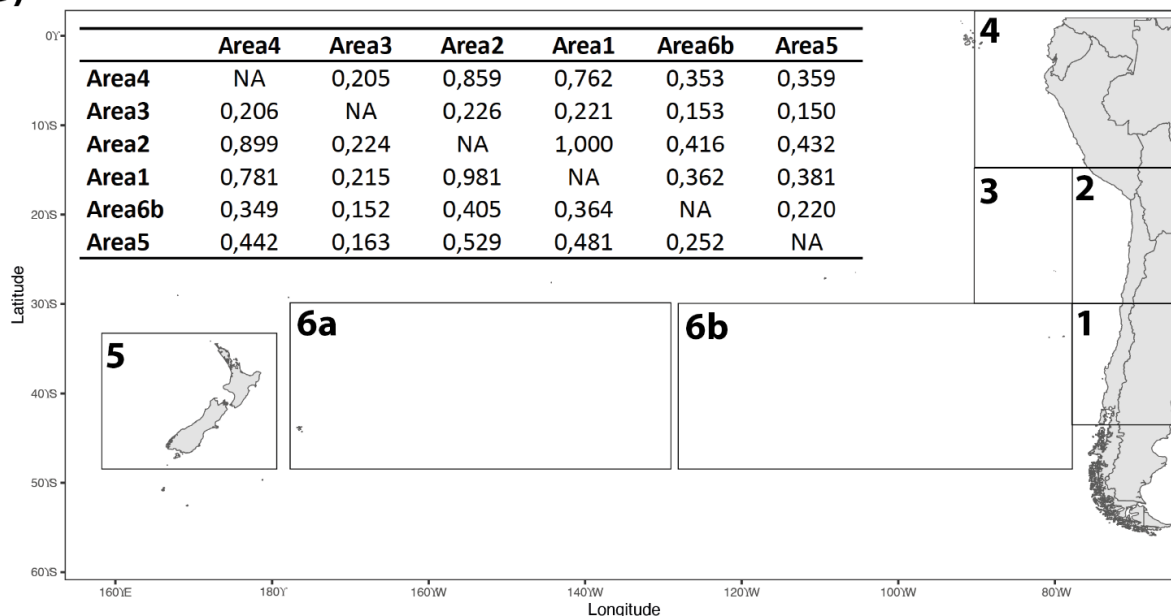


Figure 3. Bidirectional relative migration rates for neutral loci at a geographical context. A) overall data for FIPA N°2021-28 and B) grouped in the six geographical regions used to compare Chilean Jack mackerel biomass trends (Dragon et al., 2018).

Environmental and genetic distances

Figure 4A shows the relationships between the neutral genetic distances and the environmental distances statistically analyzed using the Mantel test. The sea level anomaly did not show significant relationships with neutral genetic distance (mantelr= -0.21; p=0.894), as did sea surface temperature (mantelr= 0.18; p=0.140), chlorophyll-a (mantelr = 0.18; p=0.057), wind-induced turbulence (mantelr= 0.11; p=0.099) and EKE (mantelr= -0.02; p=0.590).

The results for the adaptive genetic distance are shown in Figure 4B revealing positive relationships with the environmental distances: sea level anomaly (mantelr= 0.84; p=0.01), the EKE (mantelr= 0.73; p=0.008), the wind induced turbulence (mantelr= 0.60; p=0.01) and chlorophyll-a (mantelr= 0.29; p=0.009). The relationship between the sea surface temperature distance and the adaptive genetic distance was not positive (mantelr= 0.19; p=0.135) suggesting thermal homogeneity throughout the sampled locations.

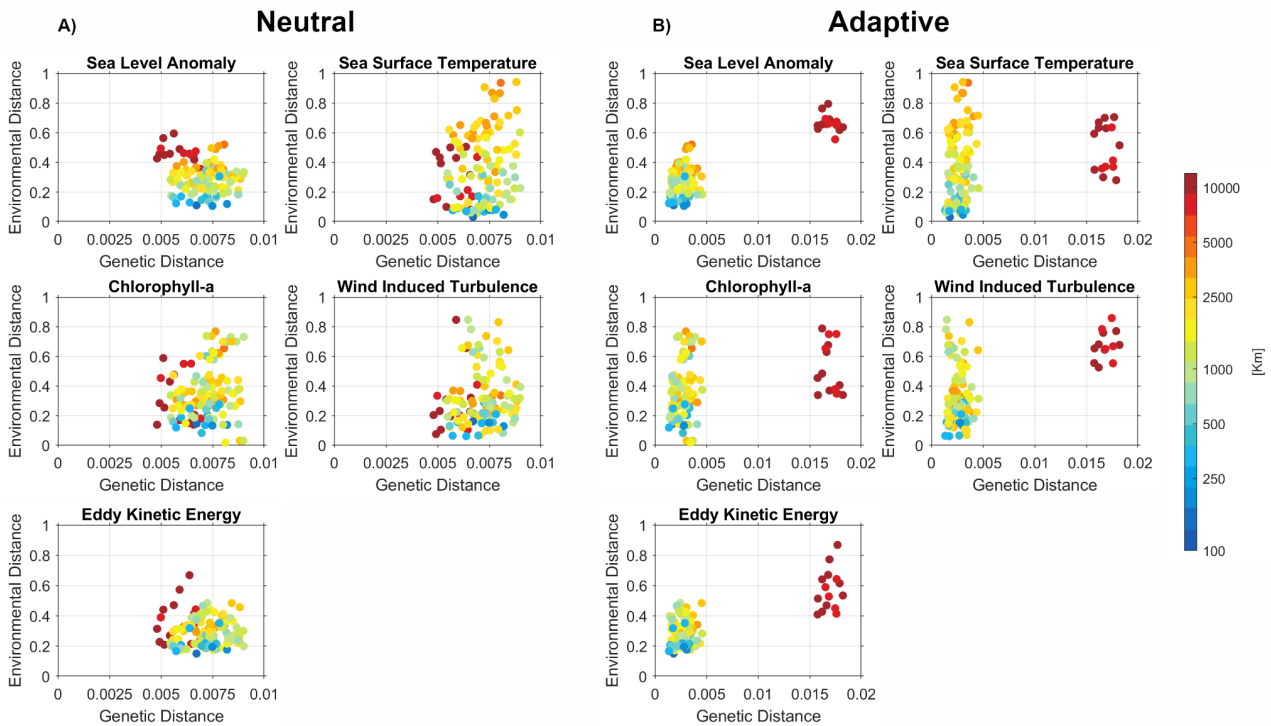


Figure 4. Relationships between the genetic distances and the environmental distances statistically analyzed using the Mantel test for Chilean jack mackerel in South Pacific Ocean. A) SNPs neutral dataset y B) SNPs adaptive data set. FIPA N°2021-28.

Concluding remarks

In this study, we conducted the first study of genomic variation in Chilean Jack mackerel, *Trachurus murphyi*, sequencing 368 individuals distributed throughout the South Pacific Ocean. Our results revealed low genome-wide differences among areas, but high differentiation at a relatively small number of putative adaptive loci. Altogether, our results revealed a pattern of spatial selection in a marine fish with high dispersal that will help support the management of this species.

1. Summary statistics of genetic diversity (i.e., H_0 and H_E) of Chilean jack mackerel differs at least two-fold between neutral and adaptive loci, being neutral larger than adaptive. Nonetheless, the values were very low, for instance the H_E for the neutral loci ranged from 0.0348 to 0.0384, while adaptive loci between 0.0138 to 0.0639.

2. Different number of private alleles were found among areas. Neutral private loci showed an overall tendency to decrease southward in Southeastern Pacific Ocean, and with none private alleles in New Zealand. Conversely, New Zealand showed almost all adaptive private loci found and nearly none of them in other areas in Southeastern Pacific Ocean.

3. Both data set give us information about the population genetic structure for Chilean Jack mackerel in the South Pacific Ocean. Neutral loci support differences for oceanic areas and south of Chile based on F_{ST} and DAPC, but not by the Bayesian genetic clustering analysis. Conversely, adaptive loci, New Zealand showed significant differences for all analyses (F_{ST} , DAPC, and the Bayesian genetic clustering).

4. The estimations of bidirectional and asymmetrical relative migration rates suggest a high connectivity among areas. Overall areas and samples used, the estimations for relative migration rates in neutral loci showed a median of 0.503 (range 0.292 - 1) with no significant asymmetrical comparison. Adaptive loci showed lower relative migration rates a median of 0.296 (range 0.103 - 1), but with 9 of 15 comparisons showing significant differences in the relative migration rates from Southeastern Pacific Ocean to New Zealand. Similar results were found to regroup the samples in the regions used to compare Chilean Jack mackerel biomass trends, where the median of relative migration rates for neutral was 0.36 (range 0.15 - 1) while for adaptive was 0.291 (range 0.081 - 1). Also, significant differences were found in the relative migration rates from the Southeastern Pacific Ocean to New Zealand.

5. Associations between oceanographic variables along whole South Pacific Ocean and genetic distances showed no differentiation for neutral loci, but high differences for adaptive loci. The sea level anomaly, the EKE, the wind induced turbulence, and chlorophyll-a, showed positive and significant association, mainly for samples of New Zealand.

6. Overall, the results obtained suggest that Chilean jack mackerel shows population structure and adaptation despite considerable gene flow in the South Pacific Ocean. The connectivity (i.e. gene flow) and environmental variables plays a key role for the contrasting patterns of spatial structure found for neutral an adaptive loci.

Acknowledgements

This study was financed by the Fondo de Investigación Pesquera and Aquaculture of Chile. FIPA N°2021-28. We thank to the professionals of Instituto de Investigación Pesquera (INPESCA) and Instituto de Fomento Pesquero of Chile (IFOP), Ministry for Primary Industries, New Zealand, Instituto del Mar del Perú and Museo de Historia Natural, Universidad de San Marcos, Perú, for your support in the sampling, and the fishers for facility in access to the samples.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE transactions on automatic control*. 19(6): 716-723.
- Anderson, G., Lal, M., Hampton, J., Smith, N., & 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. *Frontiers in Marine Science*, 6. Retrieved from <https://www.frontiersin.org/articles/10.3389/fmars.2019.00341>
- Anderson, G., Lal, M., Stockwell, B., Hampton, J., Smith, N., Nicol, S., & Rico, C. (2020). No Population Genetic Structure of Skipjack Tuna (*Katsuwonus pelamis*) in the Tropical Western and Central Pacific Assessed Using Single Nucleotide Polymorphisms. *Frontiers in Marine Science*, 7. Retrieved from <https://www.frontiersin.org/articles/10.3389/fmars.2020.570760>
- Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, et al. (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407: 513-516.
- Arancibia H, Cubillos L, Alay F, González F, Gavilán J, Hernández A, Alarcón R, Núñez S, Miranda L, George-Nascimento M (1996) Unidades de stock en el recurso jurel de la I a XI Regiones. Informe Final Proyecto FIP 1994-19: 165 pp.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: e3376.
- Canales-Aguirre CB, Ferrada S, Galleguillos R (2010). Isolation and characterization of microsatellite loci for the jack mackerel (*Trachurus murphyi* Nichols, 1920). *Conservation Genetics*. 11(3): 1235-1237.
- Cárdenas L, Silva A, Magoulas A, Cabezas J, Poulin E, Ojeda P (2009) Genetic population structure in the Chilean jack mackerel, *Trachurus murphyi* (Nichols) across the South-eastern Pacific Ocean. *Fisheries research*. 100(2): 109-115.
- Courtois B, Audebert A, Dardou A, Roques S, Ghneim- Herrera T, et al. (2013) Genome-Wide Association Mapping of Root Traits in a Japonica Rice Panel. *PLoS ONE* 8(11): e78037
- Cruz VM, Kilian A, Dierig DA (2013) Development of DArT marker platforms and genetic diversity assessment of the U.S. collection of the new oilseed crop lesquerella and related species. *PLoS One* 8(5): e64062.
- Darriba D, Taboada GL, Doallo R & Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*. 9(8): 772-772.
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5): e19379.
- Fuentes-Pardo, A. P., Farrell, E. D., Pettersson, M. E., Sprehn, C. G., & Andersson, L. (2022, May 18). The genomic basis and environmental correlates of local adaptation in the Atlantic horse mackerel (*Trachurus trachurus*) (p. 2022.04.25.489172). p. 2022.04.25.489172. bioRxiv. doi:

10.1101/2022.04.25.489172.

- Fuentes-Pardo A, Pettersson M, Sprehn C, Anderson L (2020) Population structure of the Atlantic horse mackerel (*Trachurus trachurus*) revealed by whole-genome sequencing. Report Northern Pelagic Working Group and the Pelagic Advisory. 38 pp.
- Galleguillos R, Torres A (1988) Identificación de unidades poblacionales pelágicas. Tercera etapa. Informe Final. SUBPESCA. 67 p.
- Gerlotto F, Gutierrez M, Bertrand A (2012) Review 2012. Insight on population structure of the Chilean jack mackerel (*Trachurus murphyi*). Aquatic Living Resources 25:341-355.
- Goudet J. 2005. hierfstat, a package for r to compute and test hierarchical F-statistics. Molecular Ecology Notes 5:184–186.
- Grijalva-Chon, J., Kumachi, J., Sosa-Nishizaki, K. & de la Rosa-Vélez, J. (1994). Mitochondrial DNA analysis of North Pacific swordfish *Xiphias gladius* population structure. Marine Ecology Progress Series 115, 15-19.
- Gruber B, Unmack PJ, Berry OF, Georges A. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Molecular ecology resources 18:691–699.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W. & Gascuel O (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology. 59(3): 307-321.
- Hollenbeck, C. M., Portnoy, D. S., & Gold, J. R. (2019). Evolution of population structure in an estuarine-dependent marine fish. Ecol. Evol., 9(6), 3141–3152. doi: 10.1002/ece3.4936
- Kawahara, S., Uozumi, Y., & Yamada, H. (1988). First record of a carangid fish, *Trachurus murphyi* from New Zealand. Japanese Journal of Ichthyology, 35, 212–214
- Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., et al. (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. Methods in Molecular Biology, 888, 67-89. doi:10.1007/978-1-61779-870-2
- Librado P & Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25 (11): 1451-1452.
- Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. Nat. Rev. Genet., 4(12), 981–994. doi: 10.1038/nrg1226
- Nei M. 1987. Molecular evolutionary genetics. New York, NY, USA.: Columbia University Press.
- Raman H, Raman R, Kilian A, Detering F, Carling J, et al. (2014) Genome-Wide Delineation of Natural

Variation for Pod Shatter Resistance in *Brassica napus*. PLoS ONE 9(7). e101673.

- Salzburger W, Ewing GB & Von Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular ecology*. 20(9): 1952-1963.
- Sansaloni C, Petrolì C, Jaccoud D, Carling J, Detering F, Grattapaglia D, Kilian A (2011) Diversity Arrays Technology (DArT) and next generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. *BMC Proceedings* 5(Suppl 7): P54.
- Seehausen O, Butlin R, Keller I. et al. (2014) Genomics and the origin of species. *Nat Rev Genet* 15, 176–192. <https://doi.org/10.1038/nrg3644>
- Sepúlveda A, Cubillos L, Grechina A, et al (1996) Migración de jurel desde y hacia la ZEE de Chile central. *Informes Técnicos*. FIP/IT N°96-15, 273.
- Serra, R Galleguillos R, Canales C, Ferrada S et al (2010) Estudio Migración de jurel en Chile (Fase I), 2007. Informe Final FIP N°2007-27. Fondo de Investigación Pesquera- IFOP, Valparaíso, Chile. 339p+Anexos. (www.fip.cl).
- Serra, R Galleguillos R, Canales C, Ferrada S et al (2014) Estructura poblacional de jurel. Informe Final FIP N°2010-18. Fondo de Investigación Pesquera- IFOP, Valparaíso, Chile. 277p+Anexos. (www.fip.cl).
- Takezaki N, Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144:389–399.
- Thompson JD, Gibson J, Plewniak F, Jeanmougin F. & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*. 25: 4876-4882.

Figure Supplementary

Table S1. Genetic diversity summary statistics for neutral and adaptive SNPs loci for Chilean jack mackerel in South Pacific Ocean.

Area	N	neutral				adaptivos			
		HO	HE	PA	FIS	Ho	He	PA	FIS
PerA	22	0.0305	0.0369	594	0.1170	0.0158	0.0186	3	0.1344
PerB	25	0.0307	0.0368	429	0.1108	0.0158	0.0196	0	0.1586
PerC	19	0.0307	0.0372	328	0.1159	0.0147	0.0179	0	0.1085
PerD	16	0.0287	0.0366	323	0.1470	0.0128	0.0193	0	0.2621
ChiA	28	0.0284	0.0359	481	0.1512	0.0106	0.0160	0	0.2143
ChiB	27	0.0316	0.0381	659	0.1252	0.0154	0.0190	0	0.1222
ChiC	14	0.0284	0.0351	239	0.1238	0.0111	0.0138	0	0.1299
ChiD	29	0.0305	0.0361	509	0.1070	0.0133	0.0177	0	0.1424
ChiE	13	0.0308	0.0364	201	0.1007	0.0136	0.0202	0	0.1987
ChiF	23	0.0308	0.0370	411	0.1218	0.0120	0.0162	0	0.1839
ChiG	27	0.0301	0.0360	425	0.1133	0.0155	0.0180	0	0.0959
ChiH	18	0.0313	0.0373	358	0.1165	0.0146	0.0168	0	0.1128
ChiI	14	0.0339	0.0372	191	0.0648	0.0183	0.0204	4	0.0605
ChiJ	15	0.0291	0.0354	280	0.1199	0.0124	0.0164	0	0.1957
ChiK	14	0.0320	0.0384	272	0.1154	0.0131	0.0189	4	0.2060
Nzwa	28	0.0292	0.0348	0	0.1132	0.0510	0.0693	450	0.1863

Table S2. Fst for neutral loci (below diagonal) and adaptive loci (above diagonal) for Chilean jack mackerel in South Pacific Ocean. Significant comparison at 0,01 was highlighted in red.

	PerA	PerB	PerC	PerD	ChiA	ChiB	ChiC	ChiD	ChiE	ChiF	ChiG	ChiH	ChiI	ChiJ	ChiK	NwzA
PerA	-	-0,005	-0,012	0,011	0,004	-0,007	0,009	0,008	0,012	0,003	0,005	-0,002	0,019	0,010	0,014	0,013
PerB	0,000	-	-0,005	-0,001	-0,004	-0,005	0,011	0,002	0,007	-0,002	-0,001	-0,006	0,008	0,010	0,003	0,015
PerC	-0,002	0,001	-	0,004	-0,002	-0,011	0,007	0,009	0,004	-0,006	-0,003	-0,005	0,012	0,008	0,010	0,013
PerD	0,000	0,001	0,000	-	-0,004	0,007	0,007	0,005	-0,004	0,004	0,006	0,006	0,019	0,008	0,000	0,011
ChiA	0,000	0,000	0,000	0,001	-	0,001	-0,006	-0,002	-0,004	-0,001	-0,001	-0,007	0,009	-0,006	-0,005	0,019
ChiB	-0,001	0,000	-0,001	0,001	0,000	-	0,006	0,005	-0,002	0,000	0,002	-0,001	0,013	0,005	0,004	0,015
ChiC	0,002	0,002	0,001	0,001	0,002	0,001	-	-0,004	0,015	0,005	0,019	0,013	0,020	0,006	0,007	0,004
ChiD	0,000	0,000	-0,001	0,001	0,000	-0,001	0,001	-	0,009	0,008	0,013	-0,006	0,008	0,008	0,002	0,017
ChiE	0,000	0,000	-0,001	0,001	0,000	-0,001	0,003	0,000	-	0,019	0,011	0,006	0,014	0,002	-0,011	0,004
ChiF	0,000	0,000	0,000	0,002	0,000	0,000	0,002	0,000	0,000	-	-0,008	-0,002	0,008	0,010	0,009	0,016
ChiG	-0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000	0,001	0,000	-	-0,004	0,004	0,014	0,009	0,020
ChiH	-0,001	0,000	-0,002	-0,001	0,001	0,000	0,001	0,000	0,000	0,001	0,000	-	0,001	0,009	-0,006	0,008
ChiI	0,002	0,004	0,004	0,003	0,003	0,003	0,006	0,004	0,005	0,004	0,004	0,003	-	0,010	0,002	0,005
ChiJ	0,002	0,004	0,000	0,001	0,002	0,003	0,005	0,003	0,003	0,003	0,003	0,004	0,008	-	0,004	0,006
ChiK	0,001	0,003	0,001	0,003	0,001	0,002	0,003	0,001	0,001	0,003	0,002	0,002	0,006	0,005	-	0,004
NwzA	0,000	0,000	0,000	0,001	-0,001	-0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,003	0,003	0,004	-

Table S3. Bidirectional relative migration rates for adaptive loci for all database for Chilean jack mackerel in South Pacific Ocean.

	PerA	PerB	PerC	PerD	ChiA	ChiB	ChiC	ChiD	ChiE	ChiF	ChiG	ChiH	ChiI	ChiJ	ChiK	Nwz
PerA	NA	0.707	0.697	0.311	0.375	0.589	0.166	0.383	0.220	0.312	0.384	0.374	0.226	0.230	0.228	0.440
PerB	0.538	NA	0.518	0.392	0.475	0.518	0.156	0.482	0.186	0.356	0.389	0.430	0.296	0.189	0.268	0.436
PerC	0.713	0.633	NA	0.279	0.452	0.839	0.169	0.365	0.298	0.523	0.516	0.476	0.274	0.241	0.265	0.411
PerD	0.183	0.355	0.253	NA	0.342	0.255	0.140	0.263	0.202	0.226	0.251	0.235	0.154	0.146	0.228	0.261
ChiA	0.480	1,000	0.595	0.566	NA	0.728	0.325	0.639	0.411	0.479	0.631	0.598	0.326	0.359	0.713	0.421
ChiB	0.687	0.860	0.773	0.386	0.496	NA	0.225	0.496	0.382	0.403	0.455	0.405	0.280	0.230	0.368	0.444
ChiC	0.267	0.327	0.288	0.361	0.392	0.330	NA	0.429	0.184	0.295	0.255	0.229	0.259	0.233	0.297	0.316
ChiD	0.415	0.627	0.353	0.448	0.600	0.537	0.229	NA	0.249	0.377	0.419	0.559	0.393	0.225	0.396	0.506
ChiE	0.154	0.229	0.199	0.226	0.244	0.244	0.108	0.260	NA	0.156	0.187	0.178	0.152	0.165	0.237	0.242
ChiF	0.484	0.705	0.466	0.433	0.484	0.752	0.214	0.470	0.200	NA	0.704	0.444	0.323	0.316	0.323	0.449
ChiG	0.328	0.741	0.489	0.342	0.516	0.603	0.123	0.320	0.247	0.821	NA	0.501	0.346	0.195	0.327	0.396
ChiH	0.467	0.576	0.482	0.280	0.437	0.527	0.159	0.618	0.266	0.469	0.508	NA	0.393	0.250	0.376	0.561
ChiI	0.142	0.199	0.162	0.142	0.189	0.205	0.103	0.217	0.151	0.188	0.218	0.198	NA	0.167	0.176	0.238
ChiJ	0.253	0.273	0.240	0.266	0.392	0.372	0.204	0.307	0.249	0.189	0.244	0.170	0.202	NA	0.275	0.295
ChiK	0.160	0.249	0.179	0.219	0.229	0.236	0.123	0.295	0.211	0.204	0.212	0.310	0.231	0.132	NA	0.320
Nwz	0.145	0.156	0.134	0.131	0.156	0.175	0.122	0.164	0.123	0.141	0.146	0.135	0.143	0.130	0.132	NA

Table S4. Bidirectional relative migration rates for adaptive loci for the groups used to compare Chilean Jack mackerel biomass trends (Dragon et al., 2018).

	Area4	Area3	Area2	Area1	Area6b	Area5
Area4	NA	0,257	0,741	0,600	0,306	0,297
Area3	0,211	NA	0,234	0,208	0,156	0,161
Area2	1,000	0,419	NA	0,806	0,689	0,321
Area1	0,726	0,245	0,706	NA	0,459	0,285
Area6b	0,308	0,253	0,558	0,352	NA	0,243
Area5	0,110	0,081	0,120	0,111	0,099	NA

Table S5. Bidirectional relative migration rates for neutral and adaptive loci from sampling in 2007. Chilean jack mackerel in South Pacific Ocean.

	Neutral						Adaptive					
	PerA	PerC	ChiB	ChiJ	ChiK	NWZ	PerA	PerC	ChiB	ChiJ	ChiK	NWZ
PerA	NA	0,975	0,808	0,893	0,939	0,913	NA	0,708	1,000	0,587	0,650	0,507
PerC	0,970	NA	0,741	0,827	0,856	0,831	0,957	NA	0,819	0,571	0,687	0,615
ChiB	0,819	0,752	NA	0,666	0,783	0,743	0,912	0,722	NA	0,508	0,826	0,498
ChiJ	0,902	0,834	0,733	NA	0,826	0,780	0,748	0,567	0,626	NA	0,832	0,424
ChiK	0,832	0,794	0,696	0,683	NA	0,707	0,425	0,408	0,427	0,399	NA	0,401
NWZ	1,000	0,855	0,801	0,810	0,855	NA	0,215	0,216	0,219	0,218	0,215	NA

Table S6. Bidirectional relative migration rates for neutral and adaptive loci from sampling in 2010. Chilean jack mackerel in South Pacific Ocean.

	Neutral						Adaptive					
	PerA	PerB	PerC	ChiB	ChiH	NWZ	PerA	PerB	PerC	ChiB	ChiH	NWZ
PerA	NA	0,343	0,156	0,273	0,080	0,279	NA	0,305	0,111	0,332	0,057	0,349
PerB	0,371	NA	0,331	0,916	0,112	0,928	0,375	NA	0,410	0,772	0,102	1,000
PerC	0,156	0,327	NA	0,286	0,086	0,272	0,149	0,542	NA	0,448	0,051	0,399
ChiB	0,294	0,857	0,291	NA	0,115	0,642	0,294	0,946	0,274	NA	0,089	0,786
ChiH	0,046	0,079	0,048	0,065	NA	0,064	0,055	0,077	0,052	0,070	NA	0,191
NWZ	0,309	1,000	0,305	0,760	0,118	NA	0,209	0,334	0,194	0,319	0,150	NA

Table S7. Bidirectional relative migration rates for neutral and adaptive loci from sampling in 2021. Chilean jack mackerel in South Pacific Ocean.

	Neutral									Adaptive								
	PerD	ChiA	ChiC	ChiD	ChiE	ChiF	ChiG	ChiH	ChiI	PerD	ChiA	ChiC	ChiD	ChiE	ChiF	ChiG	ChiH	ChiI
PerD	NA	0,587	0,395	0,613	0,404	0,541	0,609	0,446	0,423	NA	0,416	0,171	0,320	0,247	0,275	0,306	0,239	0,187
ChiA	0,601	NA	0,527	0,957	0,541	0,827	0,930	0,600	0,591	0,689	NA	0,396	0,778	0,500	0,584	0,768	0,594	0,397
ChiC	0,413	0,533	NA	0,597	0,368	0,508	0,564	0,422	0,386	0,440	0,478	NA	0,523	0,224	0,360	0,310	0,250	0,315
ChiD	0,643	1,000	0,570	NA	0,590	0,876	0,990	0,681	0,598	0,546	0,731	0,279	NA	0,304	0,460	0,511	0,514	0,479
ChiE	0,396	0,545	0,344	0,556	NA	0,512	0,539	0,402	0,364	0,275	0,297	0,132	0,317	NA	0,191	0,228	0,180	0,185
ChiF	0,543	0,834	0,492	0,843	0,502	NA	0,831	0,544	0,514	0,527	0,590	0,260	0,572	0,244	NA	0,858	0,549	0,393
ChiG	0,628	0,931	0,540	0,990	0,543	0,835	NA	0,637	0,562	0,417	0,629	0,150	0,390	0,301	1,000	NA	0,648	0,422
ChiH	0,480	0,655	0,420	0,685	0,433	0,604	0,657	NA	0,446	0,323	0,482	0,180	0,662	0,302	0,620	0,601	NA	0,546
ChiI	0,400	0,530	0,337	0,535	0,360	0,468	0,533	0,398	NA	0,173	0,230	0,125	0,264	0,184	0,229	0,265	0,206	NA