Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Biodiversità ed Evoluzione (Università di Bologna)

Ciclo XXVIII

Settore Concorsuale di afferenza: 05/B1 Zoologia e Antropologia

Settore Scientifico disciplinare: BIO/05 Zoologia

TITOLO TESI

Global population genomic structure and life history trait analysis of yellowfin tuna (*Thunnus albacares*)

Presentata da: **Carlo Pecoraro**

Coordinatore Dottorato:

Prof. Barbara Mantovani

Relatore:

Prof. Fausto Tinti

Co-Relatore:
Dr. Alessia Cariani
Dr. Emmanuel Chassot
Dr. Nathalie Bodin

Esame finale anno: 2016

"Global population genomic structure and life history trait analysis of yellowfin tuna (*Thunnus albacares*)"

Doctoral Thesis

Submitted to

Alma Mater Studiorum - Università di Bologna

Academic Advisors

Fausto Tinti, PhD
Alessia Cariani, PhD
Emmanuel Chassot, PhD
Nathalie Bodin, PhD

Submitted by

Carlo Pecoraro, MSc
UNIBO Matricola: 0000677562

Dipartimento di Scienze Biologiche, Geologiche ed Ambientali, Alma Mater Studiorum – Università di Bologna, Italia

” Our greatest glory is not in never falling,
but in rising every time we fall.”
Confucius

Preface

The global biomass of tuna species has been halved over the past half century by industrial fisheries but yielding considerable human benefits. Annually, fisheries catch around 10-15% of tunas and the global demand is still increasing as well as the trajectories of fishing mortality of most populations, even though they vary widely across species, oceans, biogeography, life history strategies and level of exploitation (Juan-Jordá et al 2011; 2015). Although the effort of the tuna Regional Fishery Management Organizations (tRFMOs) for developing more realistic stock dynamics models, the stock assessment of these species is mostly based on fisheries-dependent catch data, which have the potential for bias due to lack of reporting of catch and/or effort and variations in the distribution of tuna species that may cause changes in the interaction rates with individual fisheries (Collette et al., 2011). Moreover, data on basic biological parameters necessary for accurate stock assessments are often poorly known and widely neglected, thus also affecting the accuracy of the assessments and the possibility to develop more realist stock management and conservation strategies.

For instance, despite the relevance in resolving the population dynamics and structure of tuna and tuna-like species, their genetic population structure is not well resolved yet, with several studies leading to discordant lines of evidence (i.e. Laconcha et al., 2015; Riccioni et al., 2010). This is particularly true for yellowfin tuna, *Thunnus albacares* (YFT), whose population structure is still surrounded by uncertainties both at the global and local scale (Ely et al., 2005; Dammanagoda et al., 2008; Pecoraro et al., 2016; Grewe et al., 2015). Fishery-independent data are essential for a better understanding fish stock productivity and resistance to environmental changes and fisheries. However, the productivity is currently misestimated by the use of the spawning stock biomass (SSB) as a proxy for stock reproductive potential (Kell et al., 2015). In fact, this parameter assumes that fecundity is only related to the mass-at-age of the sexually mature portion of the population irrespective of the demographic composition of adults, without taking in account a variety of fundamental attributes, such as the relationship between the fish size and the reproductive potential (De Lara et al., 2007).

Within such context, there is a clear need to improve the realism about YFT stock assessment through the collection and incorporation of more fishery-indipendent data for implementing effective management and conservation strategies.

Research objectives

The rationale of this Ph.D. project was identified by prioritizing key issues as objectives for contributing to the conservation of YFT and helping to develop a more realistic stock assessment and sustainable management of this species. Specifically the project's objectives were to: 1) comprehensively understand and ascertain the current state of knowledge on YFT at the global scale, identifying future research priorities needed for better assisting its management; 2) assess the YFT genetic population structure using a traditional genetic approach (e.g. microsatellite markers); 3) examine the applicability of 2b-RAD genotyping technique for future investigations in this highly migratory species; 4) investigate the YFT global population structure using this Genotyping-By-Sequencing (GBS) technique; and 5) evaluate the existence of any maternal effect that can affect the reproductive patterns and dynamics of YFT females, linking such phenotypic traits to the genomic variation assessed.

Funding for this joint Doctoral Research Program was provided by ANCIT and by the EMOTION project (ANR JSV7 007 01).

In order to delineate future research priorities for establishing a more realistic YFT stock assessment and management, key-existing information about YFT biology, ecology, stock status and stock structure were reviewed in **Chapter 1**. In this Chapter current mismatches between species' biology/ecology and the management strategies were also widely discussed, pointing out those factors that are still delaying and negatively affecting the development of an effective conservation and management strategy. For instance, the mismanagement is the major risk of ignoring the structure of YFT populations, which must be properly assessed with all the scientific and technological tools available, in order to avoid any potential extinction risk of over-harvested populations, especially at the local scale. According to the importance of characterizing YFT populations, in **Chapter 2** the global genetic population structure was investigated over the entire species distribution, employing a panel of microsatellite markers. The results obtained confirmed the discordant patterns of differentiation and the high degree of uncertainty that still characterizes the global population structure, which might be a direct consequence of the limited genetic resolution of classical molecular markers. On the contrary, the access to more powerful and cost effective genetic tools (e.g. Next Generation Sequencing technologies) would represent the first step for resolving YFT population structure at both local and global scale. In this context, the applicability and the potential of 2b-RAD technique for investigating population genetic

structure in this non-model, large pelagic and highly migratory fish species were tested and assessed in **Chapter 3**. After having evaluated the efficiency and usefulness of 2b-RAD technique, in **Chapter 4** a large dataset of high quality 2b-RAD markers was generated. Based on the SNP markers discovered, the global YFT population structure was assessed, revealing a higher level of population structure, especially at the intra-oceanic scale, than assumed by the current tRFMOs' management.

Finally, in **Chapter 5** the reproductive patterns of the Atlantic YFT spawning females were examined in relation to their size, energy allocation strategy and genotypic profiles for evaluating the maternal effect hypothesis in this marine fish species. In this Chapter a deep reflection about the inappropriateness of the spawning stock biomass (SSB) to measure the stock productivity was also undertaken.

Reference

Collette B. B., Carpenter K. E., Polidoro B. A., Juan-Jordá M. J., Boustany A., Die D. J., Elfes C. et al., "High value and long life—double jeopardy for tunas and billfishes." Science 333, (2011): 291–92.

De Lara M., Doyen, L., Guilbaud T. and Rochet M.-J. "Is a management framework based on spawning-stock biomass indicators sustainable? A viability approach." ICES J. Mar. Sci. 64 (2007): 761–767.

Dammannagoda S. T., Hurwood D. A. and Mather P. B. "Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin tuna (*Thunnus albacares*) from the North Indian Ocean around Sri Lanka." Fisheries Research 90, (2008): 147–57.

Ely B., Viñas J., Alvarado Bremer J. R., Black D., Lucas L., Covello K., Labrie A. V. and Thelen E. "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 Evolutionary Biology 5, (2005): 19.

Grewe P. M., Feutry P., Hill P. L., Gunasekera R. M., Schaefer K. M., Itano D. G., Fuller D. W., Foster S. D. and Davies C. R. "Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource." Scientific Reports 5, (2015).

Juan-Jordá M. J., Mosqueira I., Freire J. and Dulvy N. K. "Life in 3-D: life history strategies in tunas, mackerels and bonitos." Reviews in Fish Biology and Fisheries 23, (2013): 135–55.

Juan-Jordá M. J., Mosqueira I., Cooper A. B., Freire J. and Dulvy N. K. "Global population trajectories of tunas and their relatives." Proceedings of the National Academy of Sciences, (2011).

Kell L.T., Nash R.D., Dickey-Collas M., Mosqueira I. and Szuwalski C. "Is spawning stock biomass a robust proxy for reproductive potential?". Fish and Fisheries, (2015).

Laconcha U., Iriondo M., Arrizabalaga H., Manzano C., Markaide P., Montes I., Zarraonaindia I., Velado I., Bilbao E., Goñi N. et al. "New nuclear SNP markers unravel the genetic structure and effective population size of albacore tuna (*Thunnus alalunga*)". PloS One 10, (2015): e0128247.

Pecoraro C., Babbucci M., Villamor A., Franch R., Papetti C., Leroy B., Ortega-Garcia S. et al., "Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (*Thunnus albacares*)." Marine Genomics 25, (2016)., 43–48. DOI:10.1016/j.margen.2015.12.002

Riccioni G., Landi M., Ferrara G., Milano I., Cariani A., Zane L., Sella M., Barbujani G., Tinti F. "Spatio-temporal population structuring and genetic diversity retention in depleted Atlantic Bluefin tuna of the Mediterranean Sea". Proc. Natl. Acad. Sci. 107, (2010): 2102–2107.

Table of Contents

1 Chapter 1

General introduction

Publication note:

An improved version of this Chapter has been submitted to the journal "Reviews in Fish Biology and Fisheries".

Title: Putting all pieces together: integrating information on ecology, biology, fisheries, stock structure and management of yellowfin tuna (*Thunnus albacares*)

Authors: C. Pecoraro[1,2*], E. Chassot[2], N. Bodin[2], H. Murua[3], I. Zudaire[4], A. Cariani[1], P. Díaz-Jaimes[5], F. Tinti[1].

Affiliations:

[1]: University of Bologna, Dept. Biological, Geological and Environmental Sciences (BIGEA), Via Selmi 3, 40126 Bologna, Italy;

[2]: Institut de Recherche pour le Développement (IRD), UMR MARBEC (IRD/Ifremer/UM2/CNRS) SFA, Fishing Port, BP570 Victoria, Seychelles;

[3]: AZTI, Marine Research Division, Herrera Kaia-Portualdea z/g, Pasaia, 20110 Gipuzkoa, Spain;

[4]: Ikerbasque Fundazioa, Maria Diaz de Haro, 3-6°, Bilbao, 48013 Bizkaia, Spain;

[5]: Laboratorio de Genética de Organismos Acuáticos, Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Apdo. Postal 70-305, México D.F. 04510, Mexico;

1.1 Introduction

Tunas and their relatives (family Scombridae) count 51 species of tunas, Spanish mackerels, bonitos and mackerels. Tunas are highly mobile pelagic fishes characterized by a widespread distribution and a wide range of life-history attributes (Juan-Jordá et al., 2013). For centuries fisheries have targeted tunas and tuna-like species, maintaining a continuously growing trend since the early 1950s over the subsequent decades, resulting in a total annual catch exceeding 6 million metric tons in the last decade (FAO 2014). The global status assessment of tunas suggests that stocks of tropical tunas are close to maximum sustainable yield levels while long-living high value species such as Atlantic bluefin tuna (ABFT; *Thunnus thynnus*) have been subjected to overfishing (Collette et al., 2011; Juan-Jordá et al., 2011; 2013).

The monitoring and management of populations of tuna and their relatives are under the jurisdiction of five tuna Regional Fisheries Management Organizations (tRFMOs), namely the Inter-American Tropical Tuna Commission for the Eastern Pacific (IATTC, http://www.iattc.org), the Western-Central Pacific Fisheries Commission (WCPFC, http://www.wcpfc.int), the International Commission for the Conservation of Atlantic Tunas (ICCAT, http://www.iccat.int), the Indian Ocean Tuna Commission (IOTC, http://www.iotc.org), and the Commission for the Conservation of the Southern Bluefin Tuna (CCSBT, http://www.ccsbt.org). The primary goal of tRFMOs is to maintain the stocks of each tuna and tuna-like species at levels allowing the maximum sustainable yields. Tuna management and conservation strategies rely on a large range of regulations such as catch quotas, time-area closures and fishing capacity limit (Aranda et al., 2012; Hillary et al., 2015), which are mostly derived from annual or multi-annual stock assessments (for details see Juan-Jordá et al., 2011).

Currently, the stock management boundaries employed by tRFMOs are essentially defined as large areas that generally cover the whole RFMOs' competence range, occasionally split into a few sub-discrete stocks separate by general geographic boundaries, such as North vs. South or East vs. West (http://iccat.int/Data/ICCATMaps2011.pdf). The rationale of such stock delimitation is generally rooted in the statistical area boundaries historically defined for statistical data collection, combined with expert judgment (Fromentin and Powers, 2005). It is indeed traditional practice for fisheries managers to define as stock units those fishes that occur in a specified area at a specified time, reflecting practical management necessities with little regard to information on biological

stock structure and genetic integrity (Carvalho and Hauser, 1995; Ward, 2000). For this reason, mismatches between species' biology and/or ecology and the realized management strategies frequently occur, since managers need to include a variety of partly conflicting factors in their management strategy, such as biological, economic, social or even political factors (Reiss et al., 2009). Underestimating the appropriate spatial management scale for tunas and their relatives, in relation to their spatial heterogeneity and biology, can lead to a reduction of productivity, stability of tuna populations as well as of their ability to adapt and respond to environmental variation and fishing pressures (Reusch et al., 2005; Worm et al., 2006, Mace and Purvis, 2008).

Among the principal commercially valuable tuna and tuna-like species, yellowfin tuna (*Thunnus albacares*, YFT) constitutes the second largest tuna fishery worldwide, representing about one fourth of all tuna fished commercially (around 1.6 million metric tons; ISSF 2015). From the 1950s YFT catches increased constantly until the early 2000s, since then they have started declining or maintained at the same level. Despite substantial global interest around YFT stocks and the significant decrease of its catches in various regions during recent years, there are few and fragmentary information about the biology and ecology of this species and, especially, about its stock structure in different oceans with different conclusions published in the literature (Ely et al., 2005; Pecoraro et al., 2016).

This review aims to bring together key-existing information about YFT biology, ecology, stock status and structure in order to comprehensively understand and ascertain the current state of knowledge on yellowfin tuna at the global scale, providing valuable information for assisting stock assessment and sustainable management of this globally important tuna species.

1.2 Systematics and taxonomy

Yellowfin tuna was for the first time taxonomically described by Bonnaterre in 1788, under the name of *Scomber albacares.* Many other synonyms have been used for naming this species, until 1953 when Ginsburg named it as *Thunnus albacares*. YFT is one of the eight nominal species of the genus *Thunnus* (Collette et al., 2001) and specifically it belongs to the tropical *Neothunnus* subgenus which includes *T. albacares, T. tongol and T. atlanticus* and it is separated morphologically and genetically from the more cold-tolerant *Thunnus* subgenus constituted by *T. thynnus, T. maccoyii, T. alalunga, T. obesus,* although there is some discussion on which group

locate *T. obesus* (Collette and Nauen 1983; Collette 1978). Such systematic subdivision reflects differences in morphological and physiological traits and for example, species of *Neothunnus* subgenus have both central- and lateral-body heat exchangers, which instead are absent in the cold-water species. Moreover, the three tropical species of the yellowfin group do not present striations on the ventral surface of the liver and no vascular retia mirabilia are present on the dorsal surface (Collette et al., 2001).

As outlined by Collette and Nauen (1983), YFT has a fusiform and elongate body that becomes deepest under its first dorsal fin, tapering towards the caudal peduncle, which is, in turn, slim and includes three sets of bony keels. In some large individuals the second dorsal and anal fins are very long, becoming over 20% of fork length, and seven to ten dorsal and ventral finlets are present behind them. The pectoral fins are moderately long and they usually reach the base of the second dorsal fin. The swim bladder is present only in anterior half of body cavity and it can be deflated or slightly inflated. The body has a metallic dark blue changing through silvery white on belly and lower side, crossed by many faint vertical interrupted lines. The lateral parts of the body are crossed by longitudinal lines, alternated with rows of dots. A bright yellow band and a lateral blue streak bordering the dark dorsal area stretch from the eye to the caudal peduncle. Yellow highlights also typically characterize dorsal and anal fins in adults (Fig.1.1).

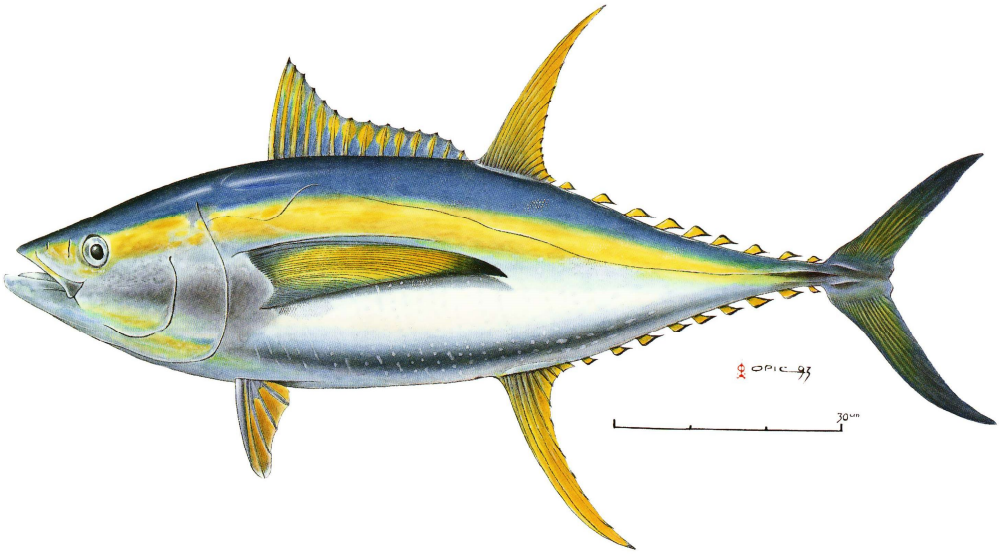


Fig 1.1_Yellowfin tuna (*Thunnus albacares*). Pierre Opic©

1.3 Ecology and biology

1.3.1 Distribution, habitat and behaviour

Yellowfin tuna have a cosmopolitan and warm-temperate distribution in tropical and sub-tropical waters of the Pacific, Atlantic and Indian Oceans, spending most of their time either in the surface mixed layer or in the top of the thermocline with a temperature ranging from 18 to 31 °C (Reygondeau et al., 2012).

Sea water temperature is one of the main factor affecting YFT habitat utilization (Schaefer et al*.,* 2007). YFT inhabit the top of the water column, using their reduced endothermic and heat-conservation strategies to raise tissue temperatures a few degrees up the water sea temperature (Block et al., 1997). YFT spend only a small fraction of total time at temperature in waters colder than 8°C because of limitations of their heart functions (Brill and Lutcavage, 2001). Indeed, their heart is not supplied with a circulatory counter-current heat exchanger and it receives coronary blood from the gills at ambient water temperature (Brill and Bushnell, 2001). Therefore, the thermal sensitivity of the heart at ambient water temperature might limit YFT ecological niche (Galli et al., 2009), with low temperature that may cause bradycardia and an associated reduction in cardiac output.

YFT are indicated to be a stenothermal species, and their temperature tolerance is determined by temperature differences relative to the SST rather than by absolute temperature itself. Hence, they do not have the thermoregulatory physiological capacity to stay for prolonged periods below the mixed layer as other tuna species do (Brill and Bushnell, 2001). Rapid deep dives have been also recorded occasionally in YFT (Block et al., 1997). The coldest temperature reached with a single dive (around 1 minute to 300 m) by a YFT sonically tracked was 7°C (Block et al., 1997), but with pop-up satellite archival tags were recorded deeper dives (984 m) at colder temperatures (5.4 °C) than with other tagging technologies (Dagorn et al., 2006; Hoolihan et al., 2014). Besides, Schaefer et al., (2009, 2011) recorded in the Eastern Pacific Ocean some dives at 1,600 m depth as well as some repetitive bounce-diving behaviour during the day to depths of 200–400 m depth.

Although temperature is the main factor that limits YFT movements, some other factors can affect its vertical distribution. For instance, their fast growth rate and high fecundity require oxygen delivery rates higher than those rates required for routine metabolic functions (Hoolihan et al., 2014). Therefore YFT, as well as other large pelagics, remains in water layers to find adequate levels of dissolved oxygen (the oxygen minimum layer), for maintaining these functions (Hoolihan

et al., 2014). The YFT habitat utilization is also strongly influenced by the presence of high prey density (Schaefer et al., 2007). In the North-Eastern, Central and Western Pacific it was demonstrated that the high levels of prey density were responsible of the YFT high degree of residency in proximity of those areas (Itano and Holland, 2000; Schaefer et al., 2007; Sibert and Hampton, 2003). Within this context, YFT preference in spending most of the time in warm waters is also related to the high mean concentration of phytoplankton biomass that concentrates epipelagic prey in euphotic surface waters (Reygondeau et al., 2011). The high concentrations of forage organisms, detected by YFT eyesight (Reygondeau et al., 2011), in the vicinity of islands and seamounts is also the main reason of the higher abundance in those areas as compared with in the surrounding oceanic waters (Schaefer et al., 2007). These seasonal oceanographic features strongly influence YFT populations, therefore understanding their effects is a fundamental step toward ecosystem-based management of fisheries, which has become a standard approach in management policies (Link et al., 2014).

1.3.2 Age, Growth and Mortality

Yellowfin tuna can grow to a maximum size of 205 cm fork length (F_L) and 194 kg of weight (ISSF 2013). The estimation of YFT age and growth is a challenge comparing to other temperate pelagic species, for three main reasons: (i) otoliths are less marked because of tropical habitat chemical and physical conditions, (ii) reproduction occurs all year round, (iii) otoliths' preparation and analysis is a time-consuming process and many biases and uncertainties (e.g., miscounting increments can lead to errors in age estimations) can occur during the analysis (Sardenne et al., 2015).

Different ageing procedures have been used for determining YFT tuna growth curve, such as: 1) modal progressions in length-frequency distributions from commercial catches or scientific monitoring (Fonteneau, 1980; Gascuel et al., 1992; Le Guen and Sakagawa, 1973; Moore, 1951); 2) direct aging of a fish from periodic deposits in calcified structures, such as scales (Yabuta et al., 1960), dorsal spines (Lessa and Duarte Neto, 2004) and otoliths (Stéquert et al., 1996); 3) estimation of growth over a specific range of time, from mark-recapture experiments (Bard, 1984; Bard et al., 1991); 4) analyses of tag-recapture in combination with otoliths data (Dortel et al., 2013; Eveson et al., 2015; Hallier et al., 2005).

YFT growth was firstly modeled with the Von Bertalanffy model (1938) which assumes a linear decrease in growth rate over the lifespan of a fish, based on modal progressions and juveniles

tagging studies (Yang et al., 1969; Le Guen and Sakagawa, 1973). However, growth studies conducted over the last decades support a two-stanza growth model with a significant change in the growth rate between juveniles and adults (Sardenne et al., 2015). According to this model, there are two growing phases: 1) a first phase in which young YFT show a slow growth rate, around 2.1 cm per $month^{-1}$, until when individuals reach 56-70 cm FL (Dortel et al., 2013; Lumineau, 2002; Marsac, 1991; Marsac and Lablache, 1985); 2) a second phase, in which YFT grow faster (4.1 $cm.month^{-1}$), until reaching a FL of around 145 cm with a decreasing of the growing rate, after that size, around 0.01 $cm.month^{-1}$ (Dortel et al., 2013). The two-stanzas model could be physiologically explained by using the different metabolic requirements in relation to fish size, the development of the swim bladder and their reproduction. Lumineau (2002) indicated that larger individuals have a relatively lower metabolic rate than small individuals and specifically the development of the gas bladder likely reduces drastically the energy required to swim. In fact, the swim bladder grows allometrically in YFT and it does not contain gas until the fish reaches 50-60 cm F_L. This implies a higher energetic availability for growing faster up. Moreover, the development of the gas bladder associated with the increase in size allows YFT to move deeper, increasing their capability to catch preys (Lehodey and Leroy, 1999). Even if, the two-stanza growth pattern has been confirmed for YFT in different oceans (Fonteneau and Hallier, 2015), no conclusive evidences on their growth model have been obtained yet, due to the lack of age validation studies and the difficulties associated with tracking YFT cohorts over time because of their extended spawning period (Dortel et al., 2013).

Sexual dimorphism has been also highlighted in YFT in terms of growth rate in length and weight. For instance, Pacific YFT young females are larger than males of the same age, but after reaching about 120 cm F_L (around 4 years old) males become larger than females (Shih et al., 2014). These results also corroborated those reported by Stéquert et al., (1996) for the YFT inhabiting the Indian Ocean and they indicated a size discrepancy among sexes, with males larger than females at three (male: 120.4cm F_L; female: 118cm F_L) and four years (male: 146.5 cm FL; female: 142.2 cm FL). Besides, it was observed a rapid decline in the percentage of females with a predominance of males at around 145-155 cm F_L (Zhu et al., 2008; Zudaire et al., 2013a).

This higher natural mortality rate (M) in females with F_L > 130 cm than in midsized individuals (Hampton and Fournier, 2001) might be due to the high energetic demand for spawning in YFT females (Schaefer, 1996), which represents the most metabolically demanding activity in their lives, occurring over multiple-spawning seasons.

Moreover, M or total mortality (Z) during YFT larval stage is due primarily to starvation and predation, and estimate of total mortality rate (Z d^{-1}) during the larval stage has been assessed from declines in abundance over time (Hampton, 2000). Lang et al., (1994) estimated a Z of 0.33 d^{-1} in YFT larvae, which may be also related to feeding and nutrition during the first development phases. However, the current available information indicate that mortality of tunas during the juvenile phase is largely a function of size or age rather than species or habitat, highlighting therefore, that YFT natural mortality during the first year of life is probably within the range reported for other tunas (Hampton 2000; Fonteneau and Pallares, 2005).

1.3.3 Reproduction

Yellowfin tuna is an iteroparous, gonochoristic and oviparous fish species, without displaying sexual dimorphism in the external morphology. Fertilization of eggs occurs externally and there is not parental care during hatching. Oocytes development pattern in YFT ovaries has been described as asynchronous (Schaefer, 1998; 2001; Zudaire et al., 2013a), and the fecundity regulation strategy is indeterminate (Zudaire et al., 2013a). The oocyte development and recruitment at the ovary is continuous during the whole spawning season, with an overlapping between oocyte recruitment and spawning events (Schaefer, 1998; Zudaire et al., 2013a). Females have a protracted spawning period (i.e., multiple batch spawner) in which mature eggs are released into the sea in multiple batches throughout the reproductive season (Itano, 2000; Schaefer, 1998; 2001; Zudaire et al., 2013a). Spawning activity has been described to occur mainly at night, between 21:00 and 03:30 h (McPherson, 1991; Schaefer, 1996) or in the early hours of the morning (Itano, 2000), with SST > 24°C (Schaefer, 1998). However, spawning in captivity took place earlier, from 13:30 to 21:30h, showing a direct relationship between the spawning diel activity and the water temperature; i.e. when water temperature increased, spawning occurred later in the day (Margulies et al., 2007). According to these authors, the variation on the spawning diel activity is related to hatching and egg-stage duration, which are inversely related to water temperature. The observed pattern might be an adaptive trait for increasing the chance of survival of newly hatched yellowfin larvae.

Fecundity in YFT was estimated to be between 1.6 and 3.1 million oocytes per batch (Batch Fecundity (BF)) in different Oceans (Table 1.1), and the relative batch fecundity (BFrel) between 55 and 74 oocytes per gram of body weight depending on the study. The inter- and intra-population fecundity variability detected in these studies seem to be affected by the geographic location and

environmental conditions (Schaefer, 2001). However, other factors, such as energy acquisition from feeding during the spawning season, play important roles in the oocyte development and, thus, in fecundity (Itano, 2000; Margulies et al., 2007; Zudaire et al., 2015). YFT females can spawn with a frequency of around 1.52 days (McPherson, 1991; Schaefer, 2001) over large areas of the tropical zone throughout the year (Itano, 2000; Stéquert et al., 2001).

Table 1.1_Batch fecundity (BF) and the relative batch fecundity (BFrel) estimates for YFT. Values are expressed in millions for BF and in oocytes per gram of gonad-free weight for BFrel. For each estimate the minimum and maximum values are also reported.

Studies	Areas	BF	BFrel
Schaefer, 1996	Eastern Pacific Ocean	1.57 (0.5-3.5)	68.0 (36.0-99.7)
Schaefer, 1998	Eastern Pacific Ocean	2.5 (0.1-8.0)	67.3 (4.9-174.0)
Itano, 2000	Western-Central Pacific Ocean	2.16 (0.4-10.6)	54.7 (31.9-147.1)
Sun et al., 2005	Western-Central Pacific Ocean	2.71 (0.9-4.7)	62.1 (31.0-98.0)
Zudaire et al., 2013	Western Indian Ocean	3.1 (0.3-6.9)	74.4 (9.2-180.8)
Diaha et al., 2015	Eastern Atlantic Ocean	2.91 (0.6-7.5)	54.39 (12.7-125.6)

Sex-ratio analysis showed differences in the proportion of male and female YFT related to size in all oceans. Similar to other tuna species (*Thunnus thynnus*: Clay, 1991), males YFT predominate among large-size individuals (Capisano, 1991; Schaefer, 1998). For instance, in the Indian Ocean males become dominant at 145-154 cm F_L (Nootmorn et al., 2005; Zhu et al., 2008; Zudaire et al., 2013). In contrast, a dominance of females was observed at sizes from 115 to 130 cm F_L (Nootmorn, 2005; Zudaire et al., 2013).

The size-related variation of the sex-ratio could be the consequence of differences in the growth rate, and/or sex-dependent natural and fishing mortalities (Fonteneau, 2002; Timochina and Romanov, 1991). Moreover, the size at which females YFT reach maturity, becoming capable of reproducing, varies among oceans. This size was identified at 102 cm F_L in the Indian Ocean (Zudaire, 2013b), at 108 cm in the Western Pacific Ocean (McPherson, 1991), at 92 cm in the Eastern Pacific (Schaefer, 1998), and at 104 cm for the equatorial Western Pacific (Itano, 2000). The size and/or age at which the 50% of the population (L_{50}) gets mature is an essential life history trait (Schaefer, 2001) and females and males YFT exhibited differences at L_{50} (92 and 69 cm F_L, respectively; Schaefer, 1998).

Nevertheless, the lack of a standardization of the criteria and methods makes any comparison between and within oceans difficult (Itano, 2000; Schaefer, 2001). YFT populations reached

maturity later in the Central-Western than in the Eastern Pacific Ocean, and this may be related to the minor primary productivity of western waters (Schafer 1998; Sun et al., 2005). It also seems that there may be a delay in YFT maturity at higher latitudes related to the lower SSTs (Itano, 2000). Although YFT females with advanced vitellogenic oocytes were considered mature in the L_{50} studies presented here, this observation could overestimate this parameter, because maturing individuals (i.e. individuals with cortical alveolar stage oocyte, CA) are classified as immature (Zudaire et al., 2013b). This stage represents the earliest evidence of oocyte maturation (Brown-Peterson et al., 2011), and females with ovaries in CA stage should be considered for maturity estimates (Lowerre-Barbieri et al., 2011). Following this recommendation Zudaire et al., (2013b), estimated L_{50} at 75 cm F_L in the Indian YFT sampled in western part.

Conversely to females YFT, few and fragmented information are available about the male reproductive biology, representing a limit for the assessment and management of YFT worldwide. Moreover, there is an evident gap of knowledge on chromosome and karyotype features for a better understanding of the YFT reproductive biology. The current state of knowledge indicates that there are not morphologically differentiated sex chromosomes in this species as in general for the genus *Thunnus* (Soares et al., 2013; Ratty et al.,1986).

1.3.4 Feeding and energetic investment

Thunnus albacares is a non-selective, opportunistic predator, feeding upon a great variety of prey species, such as crustaceans, fish, cephalopod, and gelatinous organism (Potier et al., 2004). Prey preference refers to the nomeid *Cubiceps pauciradiatus*, the swimming crab *Charibdis smithii*, and the Indian endemic species of stomatopod *Natosquilla investigatoris* (Potier et al., 2004; Potier et al., 2007; Romanov et al., 2009; Zudaire et al., 2015). The diet composition of YFT seems to be related to the fish size (Ménard et al., 2006; Graham et al., 2007; Zudaire et al., 2015), with an ontogenic feeding shift from euphasiids and planktonic prey in small individuals (<40 cm) to fish species in larger specimens (>50cm) (Maldeniya, 1996). This shift might be related to the size-increasing endothermic capability (i.e. system for conserving the heat in muscle, viscera and brain), which is functional to the independence of the YFT from thermal constraints, allowing a wide feeding vertical behaviour (Graham et al., 2007). Prey availability also affects both YFT distribution (Potier et al., 2007; Bertrand et al., 2002) and feeding success in the epipelagic environment, where prey are patchily distributed (Krosmeyer and Dewar, 2001).

Even though YFT is a day-feeder (Maldeniya, 1996) with higher activity during the dawn and sunset (Júnior et al., 2003), feeding in the night-time was also reported (Olson and Boggs, 1986). The mean daily ratio for YFT was estimated at about 2.8-4.5% of the body weight and YFT are able to acquire this ratio only in 30 minutes of feeding (Olson and Bogg, 1986). This mean daily ratio seems to increase with size until 60-70 cm F_L, after that it starts declining (Maldeniya, 1996). As described above, the decrease of the mean daily ratio is mainly linked to the higher metabolic rate in small fish than in the larger ones (Lumineau, 2002).

Yellowfin tuna, similar to other tuna species, is considered an "energy speculator" due to its high rates of energy turnover in a nutrient poor environment as the open-ocean (Korsmeyer et al., 1996). Little is known about the energy allocation strategy developed by YFT to cope with high demand of its rapid growth, early maturation and high reproductive output (Schaefer, 1998; 2001; Juan-Jorda et al.,2013). The reproductive processes require a high amount of energy, which is mainly provided by the catabolism of lipids and their constituent fatty acids, which represent the main energetic resources in fishes (Tocher, 2003). Lipids are also transferred via serum as vitellogenin proteins and very low-density lipoproteins to be deposited as yolk reserves in the oocytes (Sargent, 1995; Zudaire et al., 2014).

In YFT ovaries, it has been described a prevalence of neutral lipids over the polar lipids during ovary maturation (Zudaire et al., 2014). Instead both neutral and polar lipids are involved as endogenous energetic resources in embryogenesis and larvae development (Ortega and Mourente, 2010). YFT has shown a low capacity of energy storage in tissues during reproduction, having very low lipid contents in gonads in comparison with other tunas (Zudaire et al., 2014). This low accumulation of energy during reproduction speaks in favour of considering YFT an income-capital breeder. Since he energy stored is not enough to balance the energetic cost of reproduction, YFT also finance the high spawning activity by feeding prey during the spawning period (Zudaire et al., 2015). Species that feed during ovarian maturation acquired a great amount of dietary lipids for oocyte development (Johnson, 2009), and the energy incorporated can influence both the reproductive investment and the way in which the stored energy is employed to sustain the energetic demand (Aristzabal, 2007). This strategy of storing energy before reproduction was confirmed by the lipid composition in the liver where triacyglycerols and phospholipids are mainly accumulated (Zudaire et al., 2014).

Although the relationship between relative fecundity and body weight in YFT females is still unclear and specifically if larger females, having more energetic resources available, invest

consequently more energy for reproduction than smaller females. Obviously, the energetic cost of spawning differs between females and males, with a higher energetic investment in eggs (around 1.06% of body wet weight for a single spawning) than in sperms (Schaefer, 1996; 1998).

1.3.5 Early-life stages

The featuring of YFT early life stages come from the different scientific programs developed in the last decades (i.e. the Achotines Laboratory in Panama) to induce spawning and to rear individuals in captivity (Margulies et al., 2007). It has been demonstrated that YFT larvae during the yolk-sac stage are totally dependent on the favourable biological and physical conditions of the water column for survival, such as water temperature and dissolved oxygen, until their development for feeding is completed (Margulies et al., 2001). Sea water temperature mostly controls the duration of this developing stage, the distribution and occurrence of YFT larvae and the duration of the ingestion, metabolic and subsequent growth and mortality rates (Boehlert and Mundy, 1993; Sabatés et al., 2007; Margulies et al., 2007; Wexler et al., 2001). The YFT larvae are mostly found, in all oceans, at SST ≥ 24°C and at depths less than 50 m within the mixed layer (Wexler et al., 2007, 2011). The physiological thermal window for yolk-sac and first-feeding YFT larvae ranges between 21°C and 33°C (Wexler et al., 2011).

Limiting physical conditions (e.g. low dissolved oxygen levels, low water temperatures) negatively influence the vertical and horizontal distribution of YFT larvae and, thereby, their spatial and temporal overlap with food and predators, which might negatively affect their survival (Blaxter, 1991). In addition, the sea water density controls the development of larval swim bladder and the ontogenetic changes in its density, which in turn controls their capability to migrate vertically within the water column (Wexler et al., 2011). The fully functional development of their swim bladder in larvae may occur at a similar period as in Pacific bluefin tuna (*Thunnus orientalis*) larvae (around 7-10 days of hatching; Takashi et al., 2006).

Morphologically, YFT larvae can be identified by the presence of a single black pigment under the chin and a lack of pigment on the tail. In profile, the eye is under the line of the body axis. Instead, YTF post-larvae and small juveniles are very similar to those of their related species, making their morphological discrimination very difficult. This has promoted a number of studies aiming at the molecular identification of tuna larvae species based on molecular protocols (Hyde et al., 2005; Richardson et al., 2006; Paine et al., 2008; Viñas et al., 2009; Puncher et al., 2015a; 2015b). In fact, the proper identification of fish eggs and larvae can provide crucial information (i.e.

spawning seasons and locations, spawning stock biomass (SSB) and recruitment trends) for improving and validating stock assessment models, which are often distorted by fishery-based data (Kuparinen et al., 2014).

1.4 Fishery

1.4.1 Catches by fisheries

The global annual YFT catches amount to more than 1.6 million metric tons (using the 2009-2013 average; http://iss-foundation.org/about-tuna/status-of-the-stocks/), placing this species within the top ten positions in the ranking of marine species harvested worldwide. YFT support diverse commercial fisheries throughout their distributions, ranging from large-scale industrial to small-scale artisanal fisheries in tropical and subtropical areas (Davies et al., 2014). The YFT catches, as well as its size-selectivity, change greatly in accordance to the different fishing methods employed. Globally, purse seining (PS) is the main fishery targeting YFT and it is leading in the Eastern Pacific Ocean (EPO), whose amount (95.9%) is much higher than in the other fishing areas (61% in 2012 in the Western Central Pacific Ocean (WCPO); 68% in 2010 in the Atlantic Ocean (AO), and 35% in 2013 in the Indian Ocean (IO); Fig. 1.2).

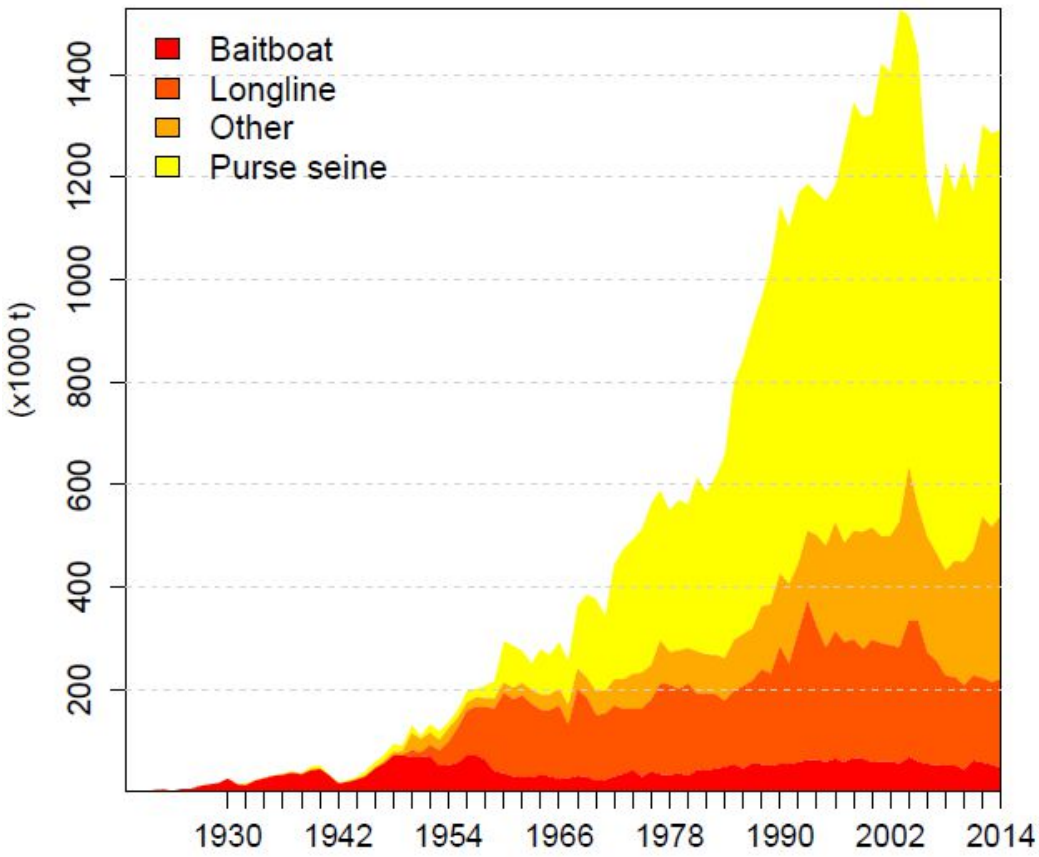


Fig.1.2_ Spatial distribution of global catch of yellowfin tuna by fishing gear over the last decades.

Tuna PS has undergone significant evolution since its inception, representing now one of the most modern and powerful fishing fleets in the world (Gaertner and Pallarés, 2002; Miyake, 2005). As it begun in EPO, PS fisheries used groups of dolphins pods predominantly of the spotted dolphin, *Stenella attenuata*, spinner dolphin, *S. longirostris*, and common dolphin, *Delphinus delphis*, to detect YFT schools, exploiting their propensity to be associated together. Scott et al., (2012) demonstrated that this association is not based on feeding advantages but mainly in reducing the risk of predation by forming large and mixed-species groups. The result of those so called "dolphin sets" was the by-catch of hundreds of thousands dolphins between 1960 and 1970. After that period, the growing public concerns led to the creation of the U.S. Marine Mammal Protection Act in 1972. In the next decades, alternative PS fishing strategies were developed worldwide. Nowadays, PS fishery, which typically targets fish ranging from 40 to 140 cm F_L, is performed with two different fishing modes (except in the eastern Pacific where dolphin sets are also common), harvesting YFT of different size: i) fishing on free-swimming schools (FSC; unassociated sets) and ii) fishing around floating objects (associated sets). Fishing on FSC, YFT of large (over 140 cm) or intermediate size (120 to 135 cm) are mainly caught on mixed or pure sets. On the contrary,

fishing on floating objects, which globally account around 40% of the catches of tropical tunas, mainly takes skipjack *Katsuwonus pelamis* (75%) and juveniles of yellowfin (16%) and bigeye *Thunnus obesus* (9%) tunas, with regional and seasonal variation in the species composition (Dagorn et al., 2012).

The term floating objects contains two types of fishing items: 1) Fish Aggregating Devices (FADs); and 2) the natural floating objects (Logs). Logs are often in areas with high load of nutrients in the sea (e.g. abundant river discharges), where they are moved by currents toward offshore waters. Instead FADs are built and deployed by fishermen. Since the early 1990s, the use of FADs for fishing has grown exponentially in the PS tuna fisheries (Scott and Lopez, 2014) because it considerably increased the catchability of tunas. On average, fishing sets around floating objects displayed higher success rates (90%) than those made on free swimming schools (50%) (Fonteneau et al., 2000; Miyake et al., 2010; Fonteneau et al., 2013).

The second largest YFT fishery is represented by longlining (LL), which mainly targets adult YFT. The fleets that mainly contribute to the global longline commercial catches are those from Japan, Republic of Korea and Taiwan and the largest amounts of catches are made in the equatorial Indonesian and Mexican waters of the Pacific Ocean (FAO 2014).

Although several mitigation measures have been proposed to effectively reduce the bycatch rate during fishing procedures (e.g. technical improvements of the fishing gear), during fishing procedures PS-FAD and LL fisheries still brought high by-catch levels (Amandè et al., 2010; Gilman et al., 2006) and many concerns have been raised on their impacts on the conservation of marine biodiversity (Gilman et al., 2011). For these reasons, it is recently growing the request for the replacement of these two fishing methods with the low-impact pole-and-line fishery, which however requires a high amount of live baits for catching mainly tuna juveniles. In general, the artisanal fishing gears have a special importance in catching YFT in Indian Ocean differently from the other oceans, where they take about 30-40% of the total catches (ca. 140,000–160,000 mt). Among the different artisanal fishing gears, gillnets are the most used, with a total catch of ca. 50,000 mt and a maximum of 170,000 mt reached in 2012.

1.4.2 Catches by Oceans

Since the 1960s the YFT catches increased until the early 1990s, when they have seemed to have levelled off at 1 to 1.6 million metric tons (mt) worldwide, representing now around the 27% of

total tuna catches worldwide (http://iss-foundation.org/about-tuna/status-of-the-stocks/). However, the trends and amounts of their catches vary among oceans (Fig. 1.3).

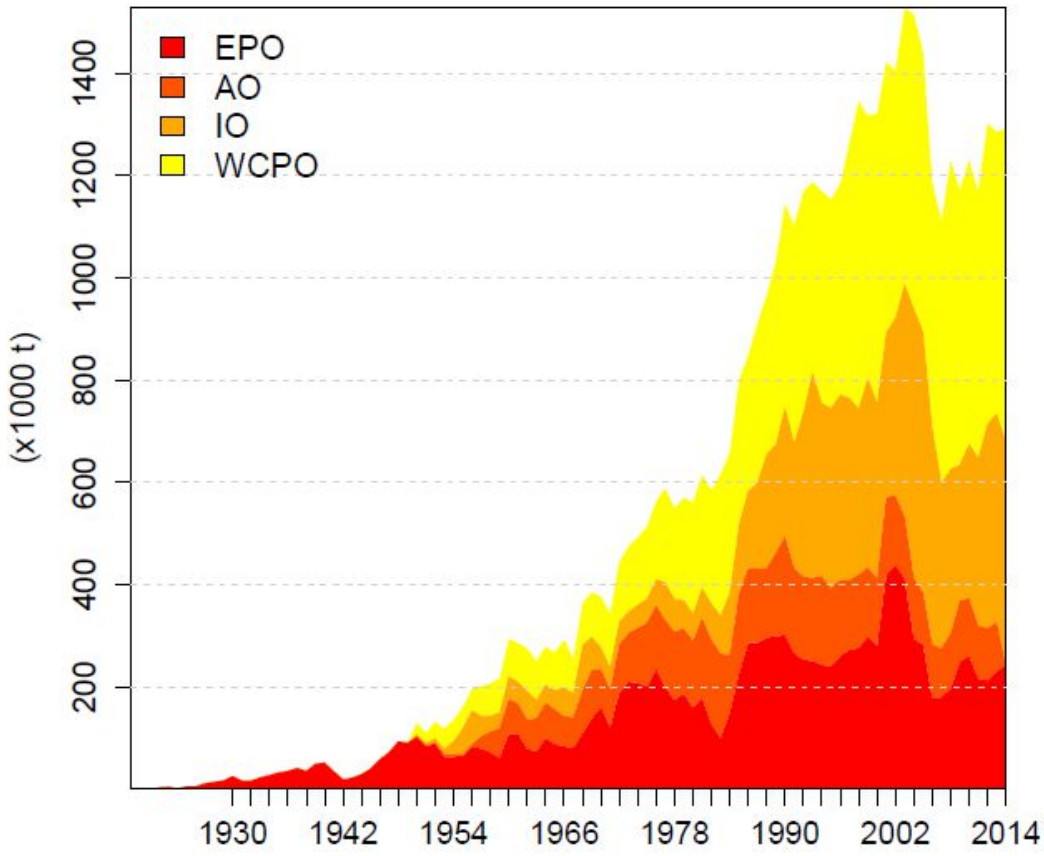


Fig. 1.3_ Catches of yellowfin tuna in the Eastern Pacific Ocean (EPO), Atlantic Ocean (AO), Indian Ocean (IO) and Western Pacific Ocean over the last decades.

In the AO, YFT are caught in tropical and sub-tropical waters, between 45°N and 40°S. Catches started increasing very quickly since the 1950s due to the rapid growth of the large-scale Asian LL fishery. YFT catches were the largest amount in this ocean until the 1990s, reaching a peak of 192,114 mt. Since then YFT catches started declining, probably due to decrease of the fishing pressure (Majkosky, 2007), and YFT has been replaced at the leading position by skipjack tuna. However, since 2011 catches has remained at similar levels of 100,000 mt.

The total annual YFT catches in IO increased significantly from 30,000 to 70,000 mt from mid-1950s to the early- 1980s with a predominance of the activities of longline and gillnet vessels. Thereafter, with the arrival of the European purse seiners YFT catches further increased, reaching over 400,000 mt in 1993 and fluctuating around that level until 2002. The YFT catches increased from 2003 to 2006 with an average annual amount of catches around 480,000 mt with a peak over 525,000 mt in 2004. After that period catches decreased noticeably, dropping down in the recent

years in areas off Somalia, Kenya and Tanzania as a consequence of the effect of piracy in the Western IO (IOTC-2013-SC16-ES04). The piracy activities have modified the normal fishing operations of the purse seiners and lead to a decrease in catch rates and effort in the Northwest of Indian Ocean (Herrera et al., 2013). However, catches increased again in 2013 and 2014 (430,000 mt; IOTC 2015).

YFT catches have increased in the whole PO from late 1950s, with two important increases in the late 1960s and the mid- to late 1980s. Until the late 1960s the catches were roughly at the same level in the eastern and western part of the Pacific Ocean. This situation changed at the end of 1970s. In 2013 the eastern catches were around 222,500 mt, with an increase of 6% from 2012, and with a higher proportion of YFT catches during El Niño years and lower during La Niña years. The bulk of YFT catches in the Eastern Pacific Ocean is taken by purse-seine sets on associated sets with dolphins and in unassociated schools (IATTC Stock Assessment Report 14). The Central-Western Pacific Ocean represents the main worldwide fishing area with more than one-quarter of YFT catches harvested at the global scale (ISSF 2013). The annual YFT catches in the Western-Central Pacific Ocean increased exponentially from the 1970 (100,000 mt) until now (550,000 mt), with a catch peak of 650,000 mt in 2008 (Davies et al., 2014).

1.5 Movement and stock structure

1.5.1 Tagging studies

Mark-recapture studies with identification tags ("conventional tagging") and the most recent electronic tagging studies have enhanced our current understanding of *T. albacares* movements and stock structure. However, the comprehension of YFT spatial stock structure is extremely hard to be fathomed, being influenced by many biotic and abiotic factors, such as geographic features, oceanographic processes and prey concentration.

YFT is able to undertake long-distance horizontal migrations with the highest mobility observed in the Atlantic Ocean, covering an average of 1574 nautical miles per month. The only case of transatlantic movements observed for this species was described for 45 individuals tagged (18,210 in total) by anglers off the east coast of the USA (Bard and Hervé, 1994), which were recovered in the Gulf of Guinea. The size of these individuals ranged from 40 to 135 cm F_L and their time at

liberty was between 18 months and almost 10 years, traveling an average of 4,100 nautical miles from their tagging sites (Fontaneau and Hallier, 2015). This migration pattern was probably linked to their homing migration from the western to the eastern side of the Atlantic Ocean. In fact, Zagaglia et al., (2004) observed that YFT juveniles were swimming along African coastal waters in the Gulf of Guinea until they reach the pre-adult stage (between 60 and 80 cm and ages between 1.5 and 2 years). Generally, after that, they started moving from the eastern to western part of the Atlantic Ocean for feeding and then they returned to the eastern part for spawning, when they reached a size of approximately 110 cm FL and about 3 years of age (Zagaglia et al., 2004). These migration patterns, having biological meanings, are indicated as advective. Comparing to the Atlantic specimens, higher degree of fidelity to their release positions and more restricted movements have been indicated in the WCPO (Itano and Holland, 2000; Sibert and Hampton, 2003) as well as in the EPO (Schaefer et al., 2007; 2011). For instance, in the EPO a persistent seasonal cycle was suggested in the movements between central and northern Baja California, with northward movements occurring during summer and southward movements occurring during winter (Schaefer et al., 2007; 2011). Whereas, in the Indian Ocean the results obtained about YFT mixing patterns were less clear than in the other oceans, as reconstructed by the Indian Ocean Tuna Tagging Programme (IOTTP; 2005–2009), and specifically its main phase, the Regional Tuna Tagging Project-Indian Ocean (RTTP-IO)). Most of the data of this tagging program came from the western equatorial region and a high proportion (>20%) of the tag releases occurred in the second and third quarters of 2006 (IOTC 2008; Langley and Million, 2012). Tagged individuals were assumed to be homogeneously distributed throughout the yellowfin population at the regional levels, showing evidence of incomplete mixing a year after their release within the Western Indian Ocean (Langley and Million, 2012). However, these results are consistent with the notion that the Indian Ocean basin probably supports a single well-mixed YFT population (IOTC-2013-WPTT15-13). Although tagging results provided evidences about the potential of YFT to undertake large movements, with an average distance travelled estimate of 710 nautical miles, they exclude any assumption about complete mixing over large regional areas (Schaefer et al., 2011), thus indicating that this species exhibits high level of structuring and site fidelity within each Ocean (Sibert and Hampton, 2003; Schaefer et al., 2009; 2011; Langley and Million, 2012; WCPFC-SC9-2013/ SA-IP-11).

1.5.2 Genetic population structure studies

Yellowfin tuna display life-history traits (LHT), such as their high dispersal capability, spawning patterns and large population sizes, which counteract evolutionary forces driving to population differentiation, favouring those conditions under which panmixia could exist. According to their life-history characteristics, which do not have a direct impact on gene flow but certainly play a role in neutralizing the effects of forces driving to divergence among populations, it is therefore quite logical to imagine that these species should exhibit a little intra-specific genetic population structuring (Ward, 2000). Nevertheless, YFT population structure is intensely conditioned by environmental constraints (i.e. sea surface temperature, salinity, ocean fronts, currents, bathymetry etc.) which affect their connectivity and dispersal capability. Moreover, as for other tuna and tuna-like species, climatic changes and paleoceanographic conditions could have promoted vicariance, population size fluctuations, and secondary contact of previously allopatric populations, influencing the patterns of genetic differentiation and geographic distribution of lineages within the species (Alvarado Bremer et al., 2005).

The rapid growth of the number of research studies that investigated the genetic population structure of yellowfin tuna reflects the increasing interest and applicability of genetics focused on the fishery management of tuna and tuna-like species in the past six decades. In fact, since 1960's the investigation of the complex tuna and tuna-like population structure started considering the use of genetic markers, with a sharp growth experienced by the number of related peer-reviewed articles over the past 50 years (Fig. 1.4). Specifically, since the mid-nineties the number of studies on the genetic population structure of tuna and tuna-like species has almost quadrupled, mainly due to the augmented availability of genetic tools and ever increasing awareness of the worth of genetic data into the stock assessment for management purposes.

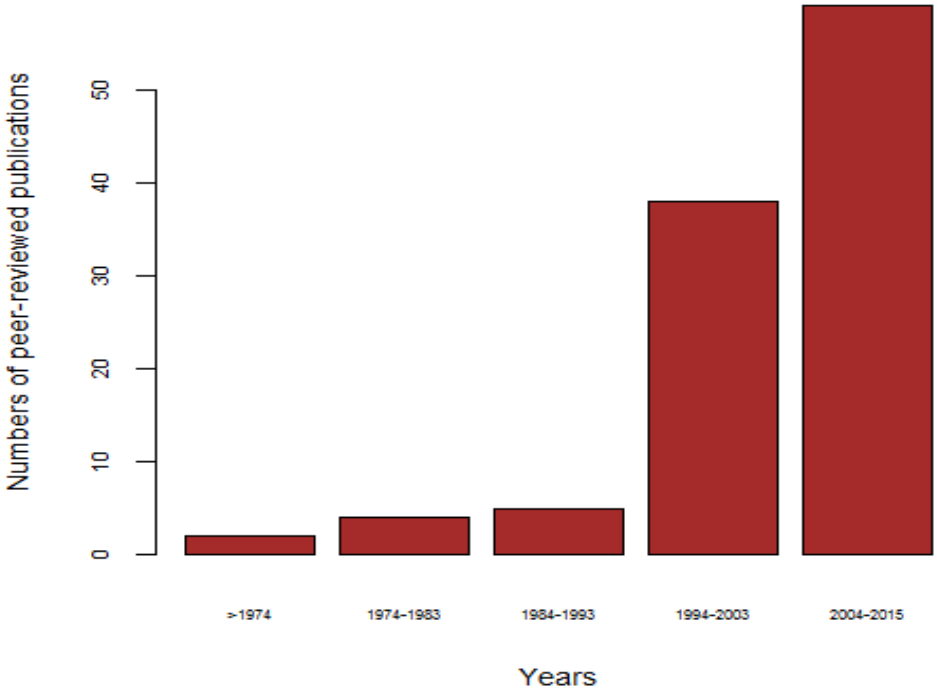


Fig. 1.4_Number of peer-reviewed articles by ten years group published since the 1960 about tuna and tuna-like species genetic population structure. Source: *Web of Science* and *Science Direct* ((TITLE: (tuna*) and TITLE: (tuna-like species*) and TOPIC: (population structure*) TIME SPAN (all years*).

Since the mid-80s the mitochondrial DNA (mtDNA) has been the first type of molecular marker largely used to identify the population structure of tuna and tuna-like species, according to its facility to be isolated and the availability of universal primers to amplify several mtDNA genes but also due to some of its features, such as its simple mode of transmission avoiding recombination, high mutation rate, maternally inherited and capacity to reveal evolutionary processes (Graves et al., 1984; Hauser and Ward, 1998). In the first period, the mitochondrial genome was analyzed by using Restricted Fragment Length Polymorphisms (RFLPs) but since the discovery of the Polymerase chain reaction (PCR), researchers have directly amplified and sequenced specific gene regions (Graves and McDowell, 2003). Consequently, over the past 30 years, there has been observed an ever-increasing of the number of studies on tuna and tuna-like population structure employing mtDNA as molecular marker, reaching their peak in the middle of the twenty-first century (Fig. 1.5).

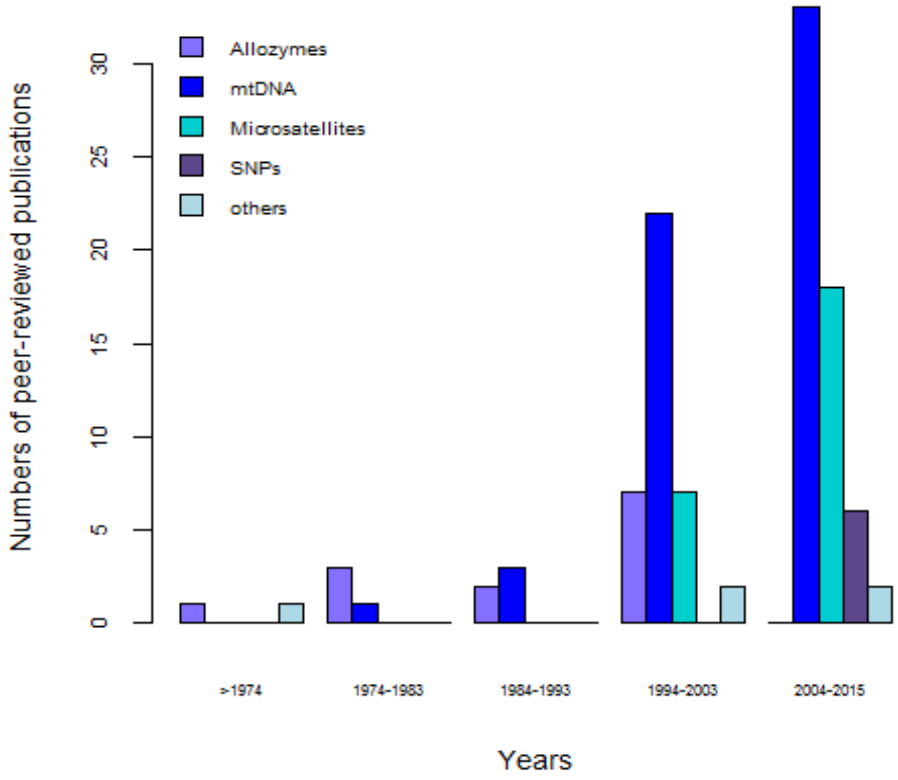


Fig. 1.5_ Number of peer-reviewed articles, employing the different molecular markers, by 10 years group published since the 1960 about tuna and tuna-like species genetic population structure. Source: *Web of Science* and *Science Direct* ((TITLE: (tuna*) and TITLE: (tuna-like species*) and TOPIC: (population structure*) and TIME SPAN (all years*).

In spite of the biological and economic relevance of this species, its genetic population structure is currently still poorly understood and not yet well resolved, with several studies leading to discordant evidences at the intra or inter-oceanic level (Table 1.2).

Table 1.2_ YFT population genetic studies, using traditional molecular markers, between and within RFMOs using different methods: (1: allozymes; 2: mtDNA; 3: microsatellites; 4: other). Symbols (x: significant differentiation; - : absence of differentiation; ^: not analyzed).

Source	Method_1	Method_2	Between RFMOs	Within RFMOs
Suzuky, (1962)	4		-	-
Sharp, (1978)	1		x	-
Scoles and Graves, (1993)	2		-	-
Ward et al., (1994)	1	2	x	-
Ward et al., (1997)	1	2	x	-
Chow et al., (IOTC 2000)	2		-	-
Nishida et al., (IOTC 2001)	4		^	-
Applayerd et al., (2001)	3		-	-
Ely et al., (2005)	2		-	-
Díaz-Jaimes and Uribe-Alcocer, (2006)	3		^	x
Dammanagoda et al., (2008)	2		^	x
Wu et al., (2010)	2		-	-
Kunal et al., (2013)	2		^	x
Li et al., (2015)	2		^	x
Aguila et al., (2015)	3		^	x
Pecoraro et al., (see Chapter 2)	3		-	-

The first genetic attempt to delineate YFT stock structure date back to Suzuki (1962), employing TG2 blood group antigens in samples of the Pacific and the Indian Oceans, without showing any sign of differentiation between both locations. This first work demonstrated the serology's unsuitableness as a tool for detecting tuna and tuna-like stocks. The analysis of the protein electrophoretic variation (i.e. allozymes) was the next molecular markers employed for revealing YFT stock structure, thanks to their relative ease of application and speed of screening programs. Using this molecular tool, for instance Sharp (1978) indicated the first signals of genetic differentiation between one Western Pacific and two Eastern Pacific YFT populations using the Glucose Phosphate Isomerase (GPI) locus. This result was corroborated by Ward et al., (1994), who detected significant heterogeneity at GPI-F* allozyme locus. Nevertheless, this pattern of differentiation was not confirmed by the examination of the whole mitochondria DNA (mtDNA) of the Pacific yellowfin tuna using a RFLP approach (Scoles and Graves, 1993; Ward et al., 1994). The lack of spatial differentiation in these two studies could be related to the small size of the analyzed samples (20 samples from each of five locations of Pacific Ocean and one sample from Atlantic Ocean) and to the only two restriction enzymes employed (BclI and EcoRI), or to the inability of this protocol to detect enough variation at the nucleotide level. Ward et al., (1997) underlined the possible presence of four YFT stocks (Atlantic Ocean, Indian Ocean, W-C Pacific and E Pacific

Ocean) employing allozyme and mtDNA RFLP markers. However later studies, employing different markers (microsatellites, PCR-sequencing of the mtDNA control region and PCR-RFLP of the ATCO segment) have confirmed the presence of one panmictic population across the different oceans and RFMOs (Appleyard et al., 2001; Nishida et al., 2001; Ely et al., 2005). These studies suggested that the Cape of Good Hope might not represent a geographical barrier to gene flow for YFT populations, as it does for bigeye tuna (Alvarado-Bremer et al*.,* 1998; Chow et al., 2000a; Martinez et al., 2006; Chiang et al., 2008) as well as for other tuna and tuna-like species (i.e. blue marlin: Finnerty and Block, 1992; Graves and McDowell, 1995; swordfish: Alvarado Bremer et al., 1996; Chow et al., 1997; Chow and Takeyama, 2000; sailfish: Graves and McDowell, 1995).

Other studies also failed to find signals of genetic differentiation between Indian and Pacific samples and between westernmost and easternmost Indian samples, employing mtDNA and nuclear markers (Wu et al., 2010; Chow et al., 2000b; Nishida et al., 2001). These studies validated the possibility to have a single large panmictic population worldwide. The lower levels of genetic differentiation for YFT stocks among oceans than other tuna and tuna-like species, although in contrast with its distribution pattern, are probably linked to its large population size, which might be associated with high levels of standing genetic variation (Ward et al., 1994; Ely et al., 2005).

This hypothesis has also been supported by our findings. We analysed 5 geographical samples employing 6 microsatellite loci recently isolated for YFT (Antoni et al., 2014) and 6 microsatellite loci isolated for *Thunnus maccoyi* (Dr. P. Grewe, personal communication). The obtained results showed very low and not significant levels of genetic differentiation at both intra- and inter-oceanic level among Atlantic, Indian and Pacific Oceans (for details see Chapter 2).

On the contrary, some recent studies have suggested the possibility to face with YFT population structure at a much smaller scale in basins of the Indian and Pacific Oceans than what assumed in its assessment and management process so far (Díaz-Jaimes and Uribe-Alcocer, 2006; Dammannagoda et al., 2008, Kunal et al., 2013; Li et al., 2015; Aguila et al., 2015). For instance, the analysis of 7 SSR loci showed a significant differentiation between the northern and the southern region in the Eastern Pacific YFT (FCT = 0.016; P = 0.011) but no significant differences among samples from the same area (FSC = 0.0017; P = 0.315) (Díaz-Jaimes and Uribe-Alcocer, 2006). These results coincide with those obtained with tagging studies for individuals from the North-Eastern Pacific and speaking in favor of the site fidelity behavior (Schaefer et al., 2007; 2011). Moreover, possible existence of different stocks has been detected both within the Central

Pacific Ocean, using the mtDNA COI gene sequences (Li et al., 2015) and between Central and Western Pacific Ocean employing 10 microsatellite loci (Aguila et al., 2015).

In the north-western Indian Ocean, Dammannagoda et al*.,* (2008), revealed the possible presence of three genetically discrete YFT sub-populations (N=285), detecting significant genetic differentiation among sites using both the mitochondrial ATPase 6 and 8 regions (498 bp) (Φst = 0.1285, P <0.001) and at two microsatellite loci (Fst = 0.0164, P <0.001 and Fst = 0.0064, P <0.001). The hypothesis of presence of different stocks in the northern part of the Indian Ocean was corroborated by Kunal et al*.,* (2013), detecting the presence of at least three YFT genetic stocks in Indian waters by analyzing the sequence of mtDNA D-loop. The presence of different stocks revealed in these studies might be also explained here by a site fidelity to different spawning areas of YFT within the Indian Ocean, with the main spawning grounds west of 75°E and secondary spawning grounds exist off Sri Lanka and the Mozambique Channel and in the eastern Indian Ocean off Australia (IOTC 2015).

The high degree of differentiation gathered at these small geographical scales has raised many doubts about the actual YFT population structure, which may have wide-ranging implications for local management, underlining the necessity to investigate deeper the population structure of this species within and between each ocean. In this context, the access to more powerful and cost effective genetic tools (e.g. next generation sequencing technologies) would represent the first step for resolving the YFT stock structure at both local and global scale.

1.5.3 Sampling strategies and the misidentification issue between juveniles

The discrepancy in the genetic results obtained for yellowfin tuna might not be linked to the type and number of genetic markers employed only but also to the sampling design considered in the different studies. In fact, in order to identify proper management units for such highly migratory species, an intensive sampling effort should be performed to enable comprehensive assessment of the population structure. The first main factor to take into account is the sample size (Viñas et al., 2011). It is typically considered of around 50 individuals for population genetic studies so as to limit the sampling noise when estimating the genetic distance among samples (Nei, 1987). For instance, the sampling of few individuals, which are close relatives, have been resulted in the overestimation of population divergence and lead to misinterpretation of meaningful population structure (Allendorf and Phelps, 1981; Waples, 1998).

Since tuna and tuna-like species are able to perform long-distance migrations (Block and Stevens, 2001), a sampling focused only on adults or sub-adults in a specific area might lead to gather a group of individuals originated from different spawning areas, nullifying any possibilities to find out signals of genetic population structure. By contrast, the collection of larvae and of young-of-the year (i.e. post-larval and early-stage juvenile) characterized by limited ability to swim and as such to be caught close to their spawning areas is expected to increase the likelihood to analyse specimens that better reflect the genetic composition of the spawning populations (Carlsson et al., 2007).

However, post-larval and early-stage juveniles are morphologically similar among many tuna and tuna-like species, especially within the same genus (Robertson et al., 2007) and their identification requires elevated taxonomic skills. Species misidentification at post-larval and early-stage juveniles could be also caused by geographical variation in their morphological features (Chow and Inoue, 1993). The use of morphological characters (i.e. body shape, pigments, characteristics of the fins, etc.) are deceptive when for instance individuals are frozen or exposed to other processes (i.e. canning, filleting), which make the species identification almost impossible (Itano, 2000). This is particularly true for small-sized bigeye and yellowfin tuna, which are caught together in the FAD-based purse-seine fishery. Although there are some external and internal characteristics that allow discrimination between them, these morphological differences are almost imperceptible when the specimens are below a certain size (<40 cm F_L), vanishing the utility of these characters. Discriminating between the juveniles of the two species is essential to provide more accurate data about their recruitment trends as well as population structuring, spawning season and location.

Many different protocols based on molecular markers have been used to help in assessing with the very challenging discrimination between YFT and BET due to their low evolutionary distances, which might confound the results if a marker with a low genetic variability is employed. For discriminating properly between them, the mtDNA Control Region has been indicated as a robust marker, in association with the liver morphology test (Pedrosa-Gerasmio et al., 2012), as well as the nuclear genetic marker ITS1 (first internal transcribed spacer of the nuclear rDNA genes; Viñas and Tudela, 2009). Besides, the use of the ATCO gene region (the flanking region between ATPase and cytochrome oxidase subunit III genes; Chow and Inoue, 1993), using restriction endonucleases or nucleotide sequence analysis, was verified as one of the most performing marker in discriminating between these two species. Moreover, Michelini et al., (2007) demonstrated the

utility of the multiplex-based PCR as a rapid molecular instrument for discriminating between different tuna species.

1.5.4 Future perspectives for genetic studies

The previous section on genetic structure points out one central question: " Why are there conflicting evidences about genetic patterns of this species both at the local and global scale?". The more logical answer may be linked to the fact that YFT display combinations of life history traits (e.g. high fecundity, large population sizes, etc.) that make detecting patterns of genetic differentiation among population samples very difficult (Ely et al., 2005; Juan-Jordá et al., 2013; Pecoraro et al., 2016). These dispersal patterns together with the fact that YFT population genetic studies, focusing on a relatively small number molecular markers distributed in a limited portion of the genome (Appleyard et al., 2001; Díaz-Jaimes and Uribe-Alcocer, 2006), have led to inadequate results for testing and validating the appropriateness of the current stock structure for assessment and management adopted by each tRFMO (Pecoraro et al., 2016). The limits of these approaches are underlined by the fact that genes are not islands but they are part of a genomic community, combined both by physical closeness on chromosomes and by various evolutionary processes (Bonin, 2008). In this context, the advent of NGS technologies and the consequent decreasing of high-throughput sequencing cost have opened new possibility of identifying thousands of co-dominant single nucleotide polymorphisms (SNPs) at a genome-wide scale (Davey et al., 2011), increasing the power of genetics to discriminate weakly differentiated YFT populations (Grewe et al., 2015; Pecoraro et al., 2016). The NGS approaches have thus provided a better representation of the species´ genomes, allowing to distinguish between neutral evolutionary processes and those influenced by selection. Therefore, there is an urgent need to switch from genetics to genomics for defining proper management units in highly migratory species (Grewe et al., 2015). The need for adopting a genome-wide perspective for disentangling YFT population structure is corroborated by the evidence about the presence of genetically distinct populations at ocean basin-scale using SNP markers (Grewe et al., 2015; Pecoraro et al., 2016). Pecoraro et al., (2016) preliminarily pointed out YFT population structure among Atlantic, Indian and Pacific oceans (F_{ST} =0.0273; P-value < 0.01), demonstrating the efficiency of 2b-RAD genotyping technique (Wang et al., 2012) in assessing genetic divergence in a marine fish with high dispersal potential. Whereas, Grewe et al., (2015) firstly confirmed the presence of two populations between the Eastern and Western Pacific basins but, after having discovered a panel of 18 SNPs putatively under positive

selection, they also underlined the level of structure within the western samples, i.e. Coral Sea and Tokelau.

These results demonstrated the potential of genomic approaches but also the need to have a representative sampling design, especially at local scale, for detecting population structuring in highly pelagic species. On the other hand, they suggested the need of redefining YFT conservation units in the Pacific Ocean that should be assessed and, potentially, managed independently (Grewe et al., 2015), even if interannual stability of the genetic structure should be verified. This study also demonstrates the need to carry out finer scale population structure studies, employing outlier loci putatively under selection. In fact, for species with high migration rate and large population size, just multiplying the amount of neutral markers detected could be not enough, with the main consequence of failing to reveal the current level of demographic connectivity, especially at the local scale (Gagnaire et al., 2015; Hauser and Carvalho, 2008).

Another future research priority is represented by the estimation of the effective population size (Ne), using both temporal or single-sample estimators (Palstra and Fraser, 2012), which determines the genetic properties of a population (Hauser et al., 2002). The Ne plays a central role in defining how demographically, environmentally or genetically stochastic events can affect the species' persistence and fitness (Palstra and Fraser, 2012). In marine fishes, this is particularly true because Ne is generally much lower than the adult census size (N) in natural populations (Palstra and Ruzzante, 2008).

Populations of millions of individual fish may be dependent to an effective population size of only hundreds or thousands individuals (Hauser et al., 2002). Therefore, there is an evident need to determine the *Ne* in YFT populations in order to understand how fishery and environmental changes are influencing the genetic diversity and thus *Ne*, foreseeing possible short- and long-term evolutionary scenarios.

Despite the importance of this parameter for populations' conservation, it has never been estimated for YFT populations. Based on the lack of these parameters, the genomic approach represents life-blood for the conservation of this species, increasing the precision and accuracy of Ne estimation by genotyping hundreds to thousands of neutral loci in numerous individuals. In fact, the high number of loci examined will increase the accuracy of Ne estimates, facilitating the identification and exclusion of outlier loci which cause biased estimates of this parameter (Allendorf et al., 2010). Consequently, a casualness of these information would represent a loss of an important perspective in our attempt to match assessment with true biological boundaries.

1.6 Current status and management

Yellowfin tuna is listed as Near Threatened (IUCN 2011), and it has been considered to be divided in 4 stocks, which are separately and independently managed and assessed by the 4 major RFMOs: ICCAT in the Atlantic, IOTC in the Indian Ocean, IATTC in the Eastern Pacific Ocean and the WCPFC in the Central and Western Pacific Ocean. The current status of each stock is based on the last stock assessment carried out by the scientific bodies of each commission, with the main objective to estimate population parameters, such as time series of recruitment, biomass and fishing mortality, in relation to biological reference points, which indicate the status of the stock in function of the fishing impacts. Thus, the status of each stock is summarized in Table 1.3, using reference points, such as the ratios of stock biomass to the biomass at maximum sustainable yield (*SBlatest/SBMSY* but also *SBcurrent/SBMSY*) and fishing mortality to the fishing mortality at MSY (*Fcurrent/FMSY*).

Table 1.3_ For each of the four RFMOs, there are reported the total amounts of catches in metric tons (Recent catch), the average of the total amount of catches in the last five years before the last stock assessment (5-yr catch (mt)), the Maximum Sustainable Year (MSY), the reference points for the Fishing Mortality (F/FMSY).

	Estimate	Years	Note
WCPFC			
Recent catch (mt)	599	2014	7%>2013
5-yr catch (mt)	565	2010-14	
MSY	586.00	2008-11	
F/FMSY	0.72 (0.58-0.90)	2008-11	F<FMSY
B/BMSY	1.24 (1.05-1.51)	2012	B>BMSY
IATTC			
Recent catch (mt)	243	2014	5%>2013
5-yr catch (mt)	232	2012-2014	
MSY	275.00	2012-2014	
F/FMSY	0.86	2012-2014	F<FMSY
B/BMSY	0.99	Start of 2015	B≤BMSY
ICCAT			
Recent catch (mt)	104	2014	3%>2013
5-yr catch (mt)	106	2010-2014	
MSY	145 (114-155)	2010	
F/FMSY	0.86(0.68-1.40)	2010	F<FMSY
B/BMSY	0.96 (0.61-1.12)	2010	B<BMSY
IOTC			
Recent catch (mt)	430	2014	6%>2013
5-yr catch (mt)	373	2010-2014	
MSY	421 (114-155)	2014	
F/FMSY	1.34 (1.02-1.67)	2014	F>FMSY
B/BMSY	0.66 (0.58-0.74)	2014	B<BMSY

YFT stock assessment in the Eastern Pacific Ocean (EPO) performed by IATTC has the longest history than any other tuna species. In the most recent YFT stock assessment carried out in the EPO (IATTC Stock Assessment Report 15), as in the previous one, an integrated statistical age-structured stock assessment model was used, indicating the YFT population has experienced two, or possibly three, different recruitment productivity regimes (1975-1982, 1983-2002, and 2003-2012). This model is based on the statement of a single YFT stock in the EPO. However, the high regional fidelity for this species assessed by tagging studies together with the geographic variation in its phenotypic and genotypic characteristics revealed in some studies, advise the possibility to have multiple stocks in the EPO and throughout the Pacific Ocean. Moreover, this hypothesis is in agreement with the fact that longline catch-per-unit-of-effort (CPUE) trends change within this area. The model indicates the existence of a stock-recruit relationship with a rather variable recruitment, which was reflected in a change of the biomass levels. In the last assessment conducted in 2015, the recent YFT spawning biomass ratio (SBR; 2011- 2013) was estimated to be 0.26 at the beginning of 2015 slightly below the MSY level (0.27), which indicated a sharp decline in spawning biomass after 2009, probably due to the higher fishing mortality of middle-aged YFT,

followed by an increase in 2012 to above the level corresponding to the MSY (IATTC Stock Assessment Report 15). Recent biomass is larger than B_{MSY} and fishing mortality is below F_{MSY}. Lastly, the MSY was stable for the entire assessment period (1975-2014), while the overall level of fishing effort changed in respect to the MSY level (IATTC 2015).

A more recent history has the YFT stock assessment in the Western-Central Pacific Ocean (WCPO, west of 150°W), in which the first assessment was conducted in 1999 and they were repeated annually until 2007 (Hampton et al., 2004; 2005 and 2006; Langley et al., 2009; 2011). YFT is considered to constitute a single stock within the WCPO (Hampton and Fournier, 2001). The methodology used for the stock assessment is based on a length-based, age-structured model (MULTIFAN-CL modelling software), which provides an integrated method for estimating catch age composition, growth parameters, mortality rate, recruitment and other parameters from time series of fishery catches, effort and length frequency data (Hampton and Fournier, 2001; Davies et al., 2014). The results of the last stock assessment in 2014 (Davies et al., 2014) are roughly similar to those in 2011, with the same magnitude order of spawning biomass (999,000 mt and 845,000 mt, respectively), and estimates of *Fcurrent/FMSY* (around 0.7), but with moderately lower spawning biomass (1.24 and 1.30, respectively).

In the Atlantic Ocean the last assessment of YFT was conducted in 2011, employing a both age-structured (VPA) model and a non-equilibrium production model (ASPIC), based on the available catch data from 2010. The results of the age structured analysis indicated that the MSY (~144,600 mt) may be lower than that estimated in the past due to the shift of the fishing selectivity through smaller fish. Stock biomass was estimated to be below B_{MSY} and, thus, the stock has been considered to be overfished. However, fishing mortality was estimated to be below F_{MSY} and, thus, overfishing was not occurring. The Standing Committee on Research and Statistics (SCRS) recommended to reduce F of small YFT to increase the long-term sustainable yield.

The most alarming situation concerns the status of the YFT stock in Indian Ocean, which was determined to be overfished and subject to overfishing during the last assessment (IOTC 2015). The low level of stock biomass in 2014 is in agreement with the long-term decline in the primary stock abundance indices (longline CPUE indices) due to the increase in longline, gillnet, handline and purse seine efforts and associated catches in recent years. Catches passed from 266,848 mt landed in 2009 to 430,327 mt in 2014. Subsequently, F exceeds the F_{MSY}, and consequently the SSB in 2014 was around the 23% (21–36%) of the unfished levels and 66% (58–74%) of the level which can support MSY.

1.7 Conclusions and directions for future research priorities

Yellowfin tuna are multi-billion dollar common resources, and with the increasing fishing pressure over the past 20 years, there is a growing need to make some progresses in the definition of its stock structure for a proper and sustainable management. In fact, our overview indicates that YFT fishery management is not based on the biological structure outlined by tagging and genetic studies, which have indicated a more structured population than typically assumed in the assessment and management process. Therefore, there is a need for establishing more accurate and realistic population boundaries for this species, incorporating multi-approach fishery independent data as those from population genomics, tag-recapture or chemical tags in otoliths. Otherwise, the main risk of ignoring the appropriate population structure would be the local overexploitation and the failure in achieving objectives related to YFT conservation and optimal economic use in those areas where a mismatch between management units and population structure is undeniable. Thus, the spatial population structure and its temporal stability, within the separated spawning grounds and the areas where populations are mixed, provided by those multidisciplinary studies has to be integrated into the management strategy of each RFMO in order to ensure a sustainable fishery of this valuable natural and economic resource. This review also underlines the great importance of inserting other important biological information into the stock assessment of this species, for developing more realistic population dynamics models and for developing effective fishery management strategies. For instance, the productivity of YFT stocks has to take in account a variety of fundamental attributes, such as the onset of maturity, fecundity, atresia, duration of reproductive season, daily spawning behaviour and spawning fraction (Murua et al., 2003). Otherwise, if YFT productivity is only measured by the estimation of the SSB, which is based on the assumption that fecundity is only related to the mass-at-age of the sexually mature portion of the population irrespective of the demographic composition of adults (Murawski et al., 2001; Kell et al., 2015), there is an evident risk to overestimate populations' productivity and consequently their resistance to fisheries and environmental changes.

Moreover, crucial progresses are needed in the improvement of data collection and the adaptation of reference points (limit and target) and the Harvest Control Rules (HCRs), which have to take into account each aspect of the biology of this species. In particular, the adaptation of HCRs, which are a set of well-defined management actions in response to changes in the stock status with respect to target and limit reference points, is a crucial aspect of modern fisheries

management. They have to be identified and adopted by the four RFMOs in order to manage YFT stocks in a way that does not depend on long negotiations that end up almost always in a weak consensus. All tuna RFMOs need to adopt stock-specific limit and target reference points and HCRs to avoid over-fishing and to rebuild depleted stocks. Although relevant frameworks for the development of reference points and harvest control rules have been already developed within Commission documents, such as CMM 2014-06 (WCPFC), Resolution 13/10 (IOTC) and Recommendation 11-13 (ICCAT), the adoption of measures is slow and most of the work is conducted at the scientific level without a formal mandate from the Commissions. In fact, at present there are no explicit statements on how to achieve targets or to avoid limits, which might reflect Commission HCRs. RFMOs have also to adopt the Management Strategy Evaluation (MSE), to support stock assessments and management advice and to obtain some understanding of relative benefits and risks in adopting HCRs under a series of simulated plausible scenarios. But also in this case the level of integration of MSE activities in management differs among RFMOs and generally these management procedures are mainly developed and integrated at the scientific level without gathering any feedback from stakeholders and commissioners.

In the near future, more effort has to be put in improving the interaction and dialogue among researchers, stakeholders and policy-makers specifically in regards to the setting of objectives given at any particular harvest strategy. This is particularly important because improvements in scientific research alone are not enough for improvements in long-term management of YFT.

References

Aguila R. D., Sweedy Perez K. L. , Catacutan B. J. N., Lopez G. V., Barut N. C. and Santos M. D. "Distinct yellowfin Tuna (*Thunnus albacares*) stocks detected in Western and Central Pacific Ocean (WCPO) using DNA microsatellites." PloS One 10, (2015): e0138292.

Allendorf F. W., England P. R., Luikart G., Ritchie P. A. and Ryman N. "Genetic effects of harvest on wild animal populations." Trends in Ecology & Evolution 23, (2008): 327–37.

Allendorf F. W. and Phelps S. R. "Use of allelic frequencies to describe population structure." Canadian Journal of Fisheries and Aquatic Sciences 38, (1981): 1507–14.

Alvarado Bremer, J. R., Viñas J., Mejuto J., Ely B., and Pla C. "Comparative phylogeography of atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes." Molecular Phylogenetics and Evolution 36, (2005): 169–87.

Alvarado Bremer JR, Stéquert B., Robertson N. W. and Ely B. "Genetic evidence for inter-oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations." Marine Biology 132, (1998): 547–57.

Alvarado Bremer J. R., Mejuto J., Greig T. W. and Ely B. "Global population structure of the swordfish (*Xiphias gladius* L.) as revealed by analysis of the mitochondrial DNA Control Region." Journal of Experimental Marine Biology and Ecology 197, (1996): 295–310.

Amandè M. J., Ariz J., Chassot E., Delgado de Molina A., Gaertner D., Murua H., Pianet R., Ruiz J.and Chavance P. "Bycatch of the European purse seine tuna fishery in the Atlantic Ocean for the 2003–2007 period." Aquatic Living Resources 23, (2010): 353–62.

Antoni L., Luque P. L., Naghshpour K. and Saillant E. A. "Polymorphic microsatellite markers for the yellowfin tuna (*Thunnus albacares*)." Conservation Genetics Resources 6, (2014): 609–11.

Appleyard S., Grewe P., Innes B. and Ward R. "Population structure of yellowfin tuna (*Thunnus albacares*) in the Western Pacific Ocean, inferred from microsatellite loci." Marine Biology 139, (2001): 383–93.

Aranda M., Murua H. and de Bruyn P. "Managing fishing capacity in tuna Regional Fisheries Management Organisations (RFMOs): development and state of the art." Marine Policy 36, (2012): 985–92.

Aristizabal E. O. "Energy investment in the annual reproduction cycle of female red porgy, *Pagrus Pagrus* (L.)." Marine Biology 152, (2007): 713–24.

Bard F. X. and Hervé A. "Structure du stock de l'albacore (*Thunnus albacares*) atlantique d'après les marquages comparés aux lieux de ponte." Rec. Doc. Sci 42 (1994): 204–8.

Bard F., Chabanet N. and Caouder N. "Croissance du thon albacore (*Thunnus albacares*) en Ocean Atlantique estim´ees par marquages". ICCAT Collection Volume of Scientific Papers 36, (1991): 182–204.

Bard F. “Croissance de l’albacore (*Thunnus albacares*) Atlantique d’apr´es 25 les donnees des marquages”. ICCAT Collective Volume of Scientific Papers 20, (1984): 104–116.

von Bertalanffy L. A quantitative theory of organic growth (Inquiries on growth laws. II). Human Biology 10, (1938): 181–213.

Bertrand A., Bard F. X. and Josse E. “Tuna food habits related to the micronekton distribution in French Polynesia.” Marine Biology 140, (2002): 1023–37.

Blaxter J. H. S. “The effect of temperature on larval fishes.” Netherlands Journal of Zoology 42, (1991): 336–57.

Block B. A. and Stevens E. D. Tuna: Physiology, Ecology, and Evolution. Vol. 19. Gulf Professional Publishing, 2001.

Block, B. A., Keen J. E., Castillo B., Dewar H., Freund E. V., Marcinek D. J., Brill R. W. and C. Farwell. “Environmental preferences of yellowfin tuna (*Thunnus albacares*) at the Northern Extent of Its Range.” Marine Biology 130, (1997): 119–32.

Boehlert G. W. and Mundy B. C. “Ichthyoplankton assemblages at seamounts and oceanic islands.” Bulletin of Marine Science 53, (1993): 336–61.

Bonin A. “Population genomics: a new generation of genome scans to bridge the gap with functional genomics.” Molecular Ecology 17, (2008): 3583–84.

Brill R. W. and Bushnell P. G. “The cardiovascular system of tunas.” Fish Physiology 19 (2001): 79–120.

Brill R. W. and Lutcavage M. E. “Understanding environmental influences on movements and depth distributions of tunas and billfishes can significantly improve population assessments.” In American Fisheries Society Symposium (2001): 179–98. American Fisheries Society.

Brown-Peterson N. J., Wyanski D. M., Saborido-Rey F., Macewicz B. J. and Lowerre-Barbieri S. K. “A standardized terminology for describing reproductive development in fishes.” Marine and Coastal Fisheries 3, (2011): 52–70.

Capisano C. “Analysis of length frequencies, sex ratio and reproduction zones of the yellowfin *Thunnus albacares* in the Atlantic.” ICCAT Coll. Vol. Sci. Pap 36, (1991): 214–79.

Carlsson J., McDowell J. R., Carlsson J. E. L. and Graves J. E. “Genetic identity of YOY bluefin tuna from the eastern and western atlantic spawning areas.” Journal of Heredity 98, (2007): 23–28.

Carvalho G. R. and Hauser L. “Molecular genetics and the stock concept in fisheries.” In Molecular Genetics in Fisheries, (1995): 55–79.

Chiang H.-C., Hsu C.-C., Wu G. C.-C., Chang S.-K. and Yang H.-Y. “Population structure of bigeye tuna (*Thunnus obesus*) in the Indian Ocean inferred from mitochondrial DNA.” Fisheries Research 90, (2008): 305–12.

Chow S., Okamoto H., Miyabe N., Hiramatsu K. and Barut N. “Genetic divergence between Atlantic and Indo-Pacific stocks of bigeye tuna (*Thunnus obesus*) and admixture around South Africa.” Molecular Ecology 9, (2000a): 221–27.

Chow S., Hazama K., Nishida T., Ikame S. and Kurihara S. “A preliminary genetic analysis on yellowfin tuna stock structure in the Indian Ocean using mitochondrial DNA variation”. WPTT00–11 IOTC Proc., 3, (2000b): 312–316.

Chow S. and Takeyama H. “Nuclear and mitochondrial DNA analyses reveal four genetically separated breeding units of the swordfish.” Journal of Fish Biology 56, (2000): 1087–98.

Chow S., Okamoto H., Uozumi Y., Takeuchi Y. and Takeyama H. “Genetic stock structure of the swordfish (*Xiphias gladius*) inferred by PCR-RFLP analysis of the mitochondrial DNA Control Region.” Marine Biology 127, (1997): 359–67.

Chow S. and Inoue S. “Intra-and interspecific restriction fragment length polymorphism in mitochondrial genes of *Thunnus* tuna species.” Bull Natl Res Inst Far Seas Fish 30 (1993): 207–25.

Clay D. “Atlantic bluefin tuna (*Thunnus thynnus* L.): a review.” IATTC Spec Rep 7 (1991): 89–180.

Collette B. B., Carpenter K. E., Polidoro B. A., Juan-Jordá M. J., Boustany A., Die D. J., Elfes C. et al. “High value and long life—double jeopardy for tunas and billfishes.” Science 333 (2011): 291–92.

Collette B. B., Reeb C. and Block B. A. “Systematics of the tunas and mackerels (*Scombridae*).” Fish Physiology 19 (2001).

Collette, B. B. and Nauen C. E. "Scombrids of the words." FAO, Rome (Italy), 1983.

Collette B. B. “Adaptations and systematics of the mackerels and tunas.” The Physiological ecology of tunas, (1978): 7–39.

Dagorn L., Filmalter J. D., Forget F., Amandè M. J., Hall M. A., Williams P., Murua H. et al. “Targeting bigger schools can reduce ecosystem impacts of fisheries.” Canadian Journal of Fisheries and Aquatic Sciences 69, (2012): 1463–67.

Dagorn L., Holland K. N., Hallier J.-P., Taquet M., Moreno G., Sancho G., Itano D. G., et al. “Deep diving behavior observed in yellowfin tuna (*Thunnus albacares*).” Aquatic Living Resources 19, (2006): 85–88.

Dammannagoda S. T., Hurwood D. A. and Mather P. B. “Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin tuna (*Thunnus albacares*) from the North Indian Ocean around Sri Lanka.” Fisheries Research 90, (2008): 147–57.

Davey J. W., Hohenlohe P. A., Etter P. D., Boone J. Q., Catchen J. M. and Blaxter M. L. “Genome-wide genetic marker discovery and genotyping using next-generation sequencing.” Nature Reviews Genetics 12, (2011): 499–510.

Davies Y. K., Mees C. C. and Milner-Gulland E. J. “The past, present and future use of drifting Fish Aggregating Devices (FADs) in the Indian Ocean.” Marine Policy 45, (2014): 163–70.

Davies N., Harley S., Hampton J. and McKechnie S. “Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean”. WCPFC-SC10-2014/SA-WP-04, (2014).

Díaz-Jaimes P. and Uribe-Alcocer M. “Spatial differentiation in the Eastern Pacific yellowfin tuna revealed by microsatellite variation.” Fisheries Science 72, (2006): 590–96.

Dortel E., Massiot-Granier F., Rivot E., Million J., Hallier J., Morize E., Munaron J., Bousquet N. and Chassot E. “Accounting for age uncertainty in growth modeling, the case study of yellowfin tuna *Thunnus albacares* of the Indian Ocean”. PLoS ONE 8, (2013): 1–12.

Dragovich A. “The Food of skipjack and Yellowfin Tunas in the Atlantic Ocean.” Fish. Bull 68, (1970): 445–60.

Ely B., Viñas J., Alvarado Bremer J. R., Black D., Lucas L., Covello K., Labrie A. V. and Thelen E. “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 Evolutionary Biology 5, (2005): 19.

Eveson, J. P., Million J., Sardenne F. and Le Croizier G. “Estimating growth of tropical tunas in the Indian Ocean using tag-recapture data and otolith-based age estimates.” Fisheries Research 163, (2015): 58–68.

FAO (Food and Agriculture Organization). 2014. The State of World Fisheries and Aquaculture 2014, http://www.fao.org/fishery/sofia/en.

Finnerty, J. R., and B. A. Block. “Direct Sequencing of Mitochondrial DNA Detects Highly Divergent Haplotypes in Blue Marlin (*Makaira Nigricans*).” Molecular Marine Biology and Biotechnology 1, no. 3 (1992): 206–14.

Fonteneau A. and Hallier J.-P. “Fifty years of dart tag recoveries for tropical tuna: a global comparison of results for the Western Pacific, Eastern Pacific, Atlantic, and Indian Oceans.” Fisheries Research 163 (2015): 7–22.

Fonteneau A. and Chassot E. “An overview of yellowfin tuna growth in the Atlantic Ocean: von Bertalanffy or multistanza growth?” Collect. Vol. Sci. Pap. ICCAT 69, (2013): 2059–75.

Fonteneau A. Chassot E. and Bodin N. “Global spatio-temporal patterns in tropical tuna purse seine fisheries on drifting Fish Aggregating Devices (DFADs): taking a historical perspective to inform current challenges.” Aquatic Living Resources 26, (2013): 37–48.

Fonteneau A. and Pallares P. “Tuna natural mortality as a function of their age: the bigeye tuna (*Thunnus obesus*) case.” Collect. Vol. Sci. Pap, ICCAT 57, (2005): 127–41.

Fonteneau, A. “Estimated sex ratio of large yellowfin taken by purse seiners in the Indian Ocean; comparison with other Oceans IOTC Proceedings,” 2002.

Fonteneau A., Pallares P. and Pianet R. “A worldwide review of purse seine fisheries on FADs.” In pêche thonière et dispositifs de concentration de poissons, caribbean-martinique, 2000.

Fonteneau A. 1980. Croissance de l'albacore (*Thunnus albacares*) de l'Atlantique Est. Col. Vol. Sci. Pap., ICCAT 9(1): 152-168.

Fromentin J. M, and Powers J. E. “Atlantic bluefin tuna: population dynamics, ecology, fisheries and management.” Fish and Fisheries 6, (2005): 281–306.

Gaertner D. and Pallares P. “The European union research project, Efficiency of Tuna Purse-Seiners and Effective Effort (ESTHER): scientific report of project. Doc.” SCTB15-FTWG-3, 2002.

Gagnaire P.-A., Broquet T., Aurelle D., Viard F., Souissi A., Bonhomme F., Arnaud-Haond S. and Bierne N. “Using Neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era.” Evolutionary Applications 8, (2015): 769–86.

Galli G. LJ., Shiels H. A. and Brill R. W. “Temperature Sensitivity of cardiac function in pelagic fishes with different vertical mobilities: yellowfin tuna (*Thunnus albacares*), bigeye tuna (*Thunnus obesus*), mahimahi (*Coryphaena hippurus*), and swordfish (*Xiphias gladius*).” Physiological and Biochemical Zoology 82, (2009): 280–90.

Gascuel D., Fonteneau A. and Capisano C.” Modélisation d’une croissance en deux stances chez l’albacore (*Thunnus albacares*) de l’Atlantique Est”. Aquat. Living Resour. 5, (1992): 155–172.

Gillett R. “Replacing purse seining with pole-and-line fishing in the central and western pacific: some aspects of the baitfish requirements.” Marine Policy 35, (2011): 148–54.

Gilman E. L. “Bycatch governance and best practice mitigation technology in global tuna fisheries.” Marine Policy 35, (2011): 590–609.

Gilman E. L., Dalzell P. and Martin S. “Fleet communication to abate fisheries bycatch.” Marine Policy 30, (2006): 360–66.

Graham B. S., Grubbs D., Holland K. and Popp B. N. “A rapid ontogenetic shift in the diet of juvenile yellowfin tuna from hawaii.” Marine Biology 150, (2007): 647–58.

Graves J. E. and McDowell J. R. “Stock Structure of the World’s Istiophorid Billfishes: A Genetic Perspective.” Marine and Freshwater Research 54, (2003): 287–98.

Graves J. E. and McDowell J. R. “Inter-ocean genetic divergence of istiophorid billfishes.” Marine Biology 122, (1995): 193–203.

Graves J. E., Ferris S. D., Dizon A. E. “Close genetic similarity of Atlantic and Pacific skipjack tuna (*Katsuwonus pelamis*) demonstrated with restriction endonuclease analysis of mitochondrial DNA”. Marine Biology, 79, (1984): 315–319.

Grewe P. M., Feutry P., Hill P. L., Gunasekera R. M., Schaefer K. M., Itano D. G., Fuller D. W., Foster S. D. and Davies C. R. “Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource.” Scientific Reports 5, (2015).

Hallier J.-P. and Gaertner D. “Drifting Fish Aggregation Devices could act as an ecological trap for tropical tuna species.” Marine Ecology Progress Series 353, (2008): 255–64.

Hampton J., Langley A., Kleiber P. “Stock assessment of yellowfin tuna in the western and central Pacific Ocean, including an analysis of management options”. WCPFC SC2 SA WP-1, Manila, Philippines, (2006).

Hampton J., Kleiber P., Langley A., Takeuchi Y. and Ichinokawa M. “Stock assessment of yellowfin tuna in the western and central Pacific Ocean”. WCPFC SC1 SA WP-1, Noumea, New Caledonia, (2005).

Hampton J., Kleiber P., Langley A. and Hiramatsu M. “Stock assessment of yellowfin tuna in the western and central Pacific Ocean”. WP SA-1, SCTB 17, Majuro, Republic of Marshall (2004).

Hampton J. and Fournier D. A. “A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (Thunnus albacares) in the Western and Central Pacific Ocean.” Marine and Freshwater Research 52, (2001): 937–63.

Hampton J. “Natural mortality rates in tropical tunas: size really does matter”. Canadian Journal of Fisheries and Aquatic Sciences 57, (2000): 1002–1010.

Hauser L. and Carvalho G. R. “Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts.” Fish and Fisheries 9, (2008): 333–62.

Hauser L., Adcock G. J., Smith P. J., Ramírez B. J. H. and Carvalho G. R. “Loss of microsatellite diversity and low effective population size in an overexploited population of new zealand snapper (*Pagrus auratus*).” Proceedings of the National Academy of Sciences 99, (2002): 11742–47.

Hauser L. and Ward R. D. “Population identification in pelagic fish: the limits of molecular markers.” NATO ASI Series a Life Sciences 306 (1998): 191–224.

Herrera M.; Pierre L. and Geehan J. “Review of the statistical data and fishery trends for tropical tunas”. IOTC-2013-WPTT15-07 Rev_1, (2013).

Hillary R. M., Preece A. L, Davies C. R, Kurota H., Sakai O., Itoh T., Parma A. M., Butterworth D. S., Ianelli J. and Branch T. A. “A scientific alternative to moratoria for rebuilding depleted international tuna stocks.” Fish and Fisheries, 2015.

Hoolihan J. P., Wells R. J. D., Luo J., Falterman B., Prince E. D. and Rooker J. R. “Vertical and horizontal movements of yellowfin tuna in the gulf of mexico.” Marine and Coastal Fisheries 6, (2014): 211–22.

Hyde J. R., Lynn E., Humphreys R. Jr, Musyl M., West A. P. and Vetter R. “Shipboard identification of fish eggs and larvae by multiplex pcr, and description of fertilized eggs of blue marlin, shortbill spearfish, and wahoo.” Marine Ecology Progress Series 286 (2005): 269–77.

ISSF (International Seafood Sustainability Foundation) (2013) ISSF tuna stock status update 2013(2): status of the world fisheries for tuna. ISSF Tech Rep 2013-04A, International Seafood Sustainability Foundation, Washington, DC.

Itano D. G. The reproductive biology of yellowfin tuna (*Thunnus albacares*) in hawaiian waters and the western tropical Pacific Ocean: Project Summary. University of Hawaii, Joint Institute for Marine and Atmospheric Research Hawaii, 2000.

Itano D. G. and Holland K. N. "Movement and vulnerability of bigeye (*Thunnus obesus*) and yellowfin tuna (*Thunnus albacares*) in relation to FADs and natural aggregation points." Aquatic Living Resources 13, (2000): 213–23.

IUCN. (2011). *Thunnus albacares:* Collette B., Acero A., Amorim A.F., Boustany A., Canales Ramirez C., Cardenas G., Carpenter K.E., Chang S.-K., de Oliveira Leite Jr. N., Di Natale A., Die D., Fox W., Fredou F.L., Graves J., Guzman-Mora A., Viera Hazin F.H., Hinton M., Juan Jorda M., Minte Vera C., Miyabe N., Montano Cruz R., Masuti E., Nelson R., Oxenford H., Restrepo V., Salas E., Schaefer K., Schratwieser J., Serra R., Sun C., Teixeira Lessa R.P., Pires Ferreira Travassos P.E., Uozumi Y. & Yanez E.: The IUCN Red List of Threatened Species 2011: e.T21857A9327139. Retrieved from http://www.iucnredlist.org/details/21857/0.

Jaquemet S., Potier M. and Ménard F. "Do drifting and anchored Fish Aggregating Devices (FADs) similarly influence tuna feeding habits? A case study from the Western Indian Ocean." Fisheries Research 107, (2011): 283–90.

Johnson R. B. "Lipid deposition in oocytes of teleost fish during secondary oocyte growth." Reviews in Fisheries Science 17, (2009): 78–89.

Juan-Jordá M. J., Mosqueira I., Freire J. and Dulvy N. K. "Life in 3-D: life history strategies in tunas, mackerels and bonitos." Reviews in Fish Biology and Fisheries 23, (2013): 135–55.

Juan-Jordá M. J., Mosqueira I., Cooper A. B., Freire J. and Dulvy N. K. "Global population trajectories of tunas and their relatives." Proceedings of the National Academy of Sciences, (2011).

Júnior T. V., Vooren C. M., and Lessa R. P. "Feeding strategy of yellowfin tuna (*Thunnus albacares*), and wahoo (*Acanthocybium solandri*) in the Saint Peter and Saint Paul Archipelago, Brazil." Boletim Do Instituto de Pesca, São Paulo 29 (2003): 173–81.

Kell, L. T., Nash R. DM, Dickey-Collas M., Mosqueira I. and Szuwalski C. 'Is spawning stock biomass a robust proxy for reproductive potential?', Fish and Fisheries, (2015).

Korsmeyer K. E., Dewar H., Lai N. C. and Graham J. B. "Tuna aerobic swimming performance: physiological and environmental limits based on oxygen supply and demand." Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 113, (1996): 45–56.

Korsmeyer K. E. and Dewar H. "Tuna metabolism and energetics." Fish Physiology 19 (2001): 35–78.

Kunal S. P., Kumar G., Menezes M. R. and Meena R. M. "Mitochondrial DNA analysis reveals three stocks of yellowfin tuna *Thunnus albacares* (Bonnaterre, 1788) in Indian Waters." Conservation Genetics 14, (2013): 205–13.

Kuparinen A. Keith D. M. and Hutchings J. A. “Allee effect and the uncertainty of population recovery.” Conservation Biology 28, (2014): 790–98.

Lang K. L., Grimes C. B. and Shaw R. F. “Variations in the age and growth of yellowfin tuna larvae, *Thunnus albacares*, collected about the Mississippi river plume.” Environmental Biology of Fishes 39, (1994): 259–70.

Langley A. and Million J. “Determining an appropriate tag mixing period for the Indian ocean yellowfin tuna stock assessment.” Indian Ocean Tuna Commission 14th Working Party on Tropical Tuna Working Paper IOTC-2012-WPTT14-31, 2012.

Langley A., Hoyle S. and Hampton J. “Stock assessment of yellowfin tuna in the western and central Pacific Ocean”. WCPFC SC7 SA WP-3, Pohnpei, Federated States of Micronesia, (2011).

Langley A., Harley S., Hoyle S., Davies N., Hampton J. and Kleiber P. “Stock assessment of yellowfin tuna in the western and central Pacific Ocean”. WCPFC SC5 SA WP-3, Port Vila, Vanuatu, (2009).

Le Guen, J. and G. Sakagawa. “Apparent growth of yellowfin tuna from the Eastern Atlantic Ocean”. Fishery Bulletin 71; (1973): 175–187.

Lehodey P.and Leroy B. “Age and growth of yellowfin tuna (*Thunnus albacares*) from the western and central Pacific Ocean as indicated by daily growth increments and tagging data.” *WP YFT-2, SCTB* 12 (1999): 16–23.

Lessa R. and Duarte-Neto P. “Age and growth of yellowfin tuna (*Thunnus albacares*) in the western equatorial Atlantic, using dorsal fin spines”. Fisheries Research 69, (2004) 157–170.

Li W., Xinjun C., Qianghua X., Jiangfeng Z., Xiaojie D., and Liuxiong X. “Genetic population structure of *Thunnus albacares* in the Central Pacific Ocean based on mtDNA COI Gene Sequences.” Biochemical Genetics 53, (2015): 8–22.

Link J. S. and Browman H. I. “Integrating what? Levels of marine ecosystem-based assessment and management.” ICES Journal of Marine Science: Journal Du Conseil 71, (2014): 1170–73.

Lowerre-Barbieri S. K., Ganias K., Saborido-Rey F., Murua H. and Hunter J. R. “Reproductive timing in marine fishes: variability, temporal scales, and methods.” Marine and Coastal Fisheries 3, (2011): 71–91.

Lumineau O. “Study of the growth of yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean based on length frequency data”. IOTC Proceedings 5, (2002): 316–327.

Mace G. M. and Purvis A. “Evolutionary biology and practical conservation: bridging a widening gap.” Molecular Ecology 17, (2008): 9–19.

Majkowski J. Global Fishery Resources of Tuna and Tuna-like Species. 483. FAO, (2007).

Maldeniya R. “Food consumption of yellowfin tuna, *Thunnus albacares*, in Sri Lankan waters.” Environmental Biology of Fishes 47, (1996): 101–7.

Margulies D., Sutter J. M., Hunt S. L., Olson R. J., Scholey V. P., Wexler J. B. and Nakazawa A. “spawning and early development of captive yellowfin tuna (*Thunnus albacares*).” Fishery Bulletin 105, (2007): 249–65.

Margulies D., Wexler J. B., Bentler K. T., Suter J. M., Masuma S., Tezuka N., Teruya K., Oka M., Kanematsu M. and Nikaido H. “Food selection of yellowfin tuna, *Thunnus albacares*, larvae reared in the laboratory.” Collect. Vol. Sci. Pap, ICCAT 22, (2001): 9–33.

Marsac F., Fonteneau A. and Ménard F. “Drifting FADs used in tuna fisheries: an ecological trap?” In pêche thonière et dispositifs de concentration de poissons, Caribbean-Martinique, (2000).

Marsac F. “Preliminary study of the growth of yellowfin estimated from purse seine data in the Western Indian Ocean”. IPTP Collective volume of working documents 6, (1991): 35–39.

Marsac F. and Lablache G. “Preliminary study of the growth of yellowfin estimated from purse seine data in the Western Indian Ocean”. IPTP Collective volume of working documents TWS/85/31:91–110, (1985).

Martínez P., González E. G., Castilho R. and Zardoya R. “Genetic diversity and historical demography of atlantic bigeye tuna (*Thunnus obesus*).” Molecular Phylogenetics and Evolution 39, (2006): 404–16.

McPherson G. R. “Reproductive biology of yellowfin tuna in the eastern australian fishing zone, with special reference to the north-western coral sea.” Marine and Freshwater Research 42, (1991): 465–77.

Ménard F., Labrune C., Shin Y.-J., Asine A.-S., Bard F.-X., et al. “Opportunistic predation in tuna: a size-based approach.” Marine Ecology Progress Series 323 (2006): 223–31.

Michelini E., Cevenini L., Mezzanotte L., Simoni P., Baraldini M., De Laude L. and Roda A. “One-step triplex-polymerase chain reaction assay for the authentication of yellowfin (*Thunnus albacares*), bigeye (*Thunnus obesus*), and skipjack (*Katsuwonus pelamis*) tuna DNA from fresh, frozen, and canned tuna samples.” J. of Agricultural and Food Chemistry 55, (2007): 7638–47.

Miyake P. M., Guillotreau P., Sun C.-H. and Ishimura G. "Recent Developments in the Tuna Industry: Stocks, Fisheries, Management, Processing, Trade and Markets". FAO, (2010).

Miyake P. M. “A brief history of the tuna fisheries of the world.” In second meeting of the technical advisory committee of the FAO project management of tuna fishing capacity: conservation and socio-economics. FAO Fish. Proc, (2005): 23–50.

Moore H. “Estimation of age and growth of yellowfin tuna (*Neothunnus macropterus*) in Hawaiian waters by size frequencies”. Fishery Bulletin 52, (1951): 132– 802 149.

Murawski S. A., Rago P. J. and Trippel E. A."Impacts of demographic variation in spawning characteristics on reference points for fishery management", ICES Journal of Marine Science: Journal Du Conseil, 58 (2001), 1002–14.

Murua H., Kraus G., Saborido-Rey F., Witthames P. R., Thorsen A. and Junquera S. "Procedures to estimate fecundity of marine fish species in relation to their reproductive strategy". J. Northw. Atl. Fish. Sci., 33, (2003): 33-54.

Nei M. Molecular Evolutionary Genetics. Columbia university press, 1987.

Nishida T., Chow S., Ikame S., Kurihara S. "RFLP analysis on single copy nuclear gene loci in yellowfin tuna (*Thunnus albacares*) to examine the genetic differentiation between the western and eastern samples from the Indian Ocean". Proc. IOTC. 4, (2001): 437-441.

Nootmorn P., Yakoh A. and Kawises K. "Reproductive biology of yellowfin tuna in the eastern indian ocean." IOTC-WPTT 14 (2005): 379–85.

Olson R. J., and Boggs C. H. "Apex predation by yellowfin tuna (*Thunnus albacares*): independent estimates from gastric evacuation and stomach contents, bioenergetics, and cesium concentrations." Canadian Journal of Fisheries and Aquatic Sciences 43, (1986): 1760–75.

Ortega A. and Mourente G. "Comparison of the lipid profiles from wild caught eggs and unfed larvae of two scombroid fish: northern bluefin tuna (*Thunnus thynnus* L., 1758) and Atlantic Bonito (*Sarda sarda* Bloch, 1793)." Fish Physiology and Biochemistry 36, (2010): 461–71.

Paine M. A. McDowell J. R. and Graves J. E. "Specific identification using COI sequence analysis of Scombrid larvae collected off the Kona coast of Hawaii island." Ichthyological Research 55, (2008): 7–16.

Palstra F. P., and Fraser D. J. "Effective/census population size ratio estimation: a compendium and appraisal." Ecology and Evolution 2, (2012): 2357–65.

Palstra F. P. and Ruzzante D. E. "Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence?" Molecular Ecology 17, (2008): 3428–47.

Pecoraro C., Babbucci M., Villamor A., Franch R., Papetti C., Leroy B., Ortega-Garcia S. et al., "Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (*Thunnus albacares*)." Marine Genomics 25, (2016), 43–48. DOI:10.1016/j.margen.2015.12.002

Pedrosa-Gerasmio I. R., Babaran R. P. and Santos M. D. "Discrimination of juvenile yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) tunas using Mitochondrial DNA Control Region and liver morphology". PLoS ONE (2012).

Potier M., Marsac F., Cherel Y., Lucas V., Sabatié R., Maury O. and Ménard F. "Forage fauna in the diet of three large pelagic fishes (lancetfish, swordfish and yellowfin tuna) in the Western equatorial Indian Ocean." Fisheries Research 83, (2007): 60–72.

Potier M., Marsac F., Lucas V., Sabatie R., Hallier J.-P., Menard F. ´ Feeding partitioning among Tuna taken in surface and mid-water layers: the case of yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) in the western tropical Indian Ocean. Western Ind. Ocean J. Mar. Sci. 3, (2004): 51–62.

Puncher G.N., Arrizabalaga H., Alemany F., Cariani A., Oray I.K., Karakulak F.S., Basilone G., Cuttitta A., Mazzola A., Tinti F. “Molecular identification of atlantic bluefin tuna (*Thunnus thynnus, Scombridae*) Larvae and development of a DNA character-based identification key for Mediterranean ccombrids”. PLoS ONE 10, (2015a): 7.

Puncher G.N., Alemany F., Arrizabalaga H., Cariani A., Tinti F**.** “Misidentification of bluefin tuna larvae: a call for caution and taxonomic reform” Reviews in Fish Biology and Fisheries. (2015).

Ratty F. J., Song Y. C. and Laurs R. M. "Chromosomal analysis of albacore*, Thunnus-alalunga,* and yellowfin*, Thunnus-alabacares,* and skipjack*, Katsuwonus-pelamis, tuna*". Fish B-NOAA, 84, (1986): 469.

Reiss H., Hoarau G., Dickey-Collas M. and Wolff W. J. “Genetic population structure of marine fish: mismatch between biological and fisheries management units.” Fish and Fisheries 10, (2009): 361–95.

Reusch T. BH., Ehlers A., Hämmerli A. and Worm B. “Ecosystem recovery after climatic extremes enhanced by genotypic diversity.” Proceedings of the National Academy of Sciences of the United States of America 102, (2005): 2826–31.

Reygondeau G., Maury O., Beaugrand G., Fromentin J. M., Fonteneau A. and Cury P. “Biogeography of tuna and billfish communities.” Journal of Biogeography 39, (2012): 114–29.

Robertson M. D., Ovenden J. R. and Barker S. C. “Identification of small juvenile scombrids from northwest tropical australia using mitochondrial dna cytochrome b sequences.” Ichthyological Research 54, (2007): 246–52.

Sabatés A., Olivar M. P., Salat J., Palomera I. and Alemany F. “Physical and biological processes controlling the distribution of fish larvae in the NW Mediterranean.” Progress in Oceanography 74, (2007): 355–76.

Sardenne F., Dortel E., Le Croizier G., Million J., Labonne M., Leroy B., Bodin N. and Chassot E. “Determining the age of tropical tunas in the Indian ocean from otolith microstructures.” Fisheries Research 163 (2015): 44–57.

Sargent J.R. “Origin and functions of egg lipids: nutritional implications”. In: Bromage, N.R., Roberts, R.J. (Eds.), Broodstock Management and Egg and Larval Quality: Blackwell Science, London, (1995): pp. 353-372.

Schaefer K. M., Fuller D. W. and Block B. A. “Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the Pacific Ocean off Baja California, Mexico, determined from archival tag data analyses, including unscented kalman filtering.” Fisheries Research 112, (2011): 22–37.

Schaefer K. M., Fuller D. W. and Block B. A. “Vertical movements and habitat utilization of skipjack (*Katsuwonus pelamis*), yellowfin (*Thunnus albacares*), and bigeye (*Thunnus obesus*) tunas in the equatorial Eastern Pacific Ocean, ascertained through archival tag data.” In Tagging and Tracking of Marine Animals with Electronic Devices, (2009): 121–44.

Schaefer K. M., Fuller D. W. and Block B. A. “Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the northeastern pacific ocean, ascertained through archival tag data.” Marine Biology 152, (2007): 503–25.

Schaefer K. M. Reproductive biology of yellowfin tuna (*Thunnus albacares*) in the Eastern Pacific Ocean. Inter-American Tropical Tuna Commission, (1998).

Schaefer K. “Spawning time, frequency, and batch fecundity of yellowfin tuna, *Thunnus albacares*, near Clipperton.” Fishery Bulletin 94, (1996).

Scoles, D. R., and Graves J E. “Genetic-analysis of the population-structure of yellowfin tuna, *Thunnus-albacares*, from the pacific-ocean.” Fishery Bulletin 91, (1993): 690–98.

Scott G. P. and Lopez J. The Use of FADs in Tuna Fisheries. (2014).

Scott M. D., Chivers S. J., Olson R. J., Fiedler P. C. and Holland K. “Pelagic predator associations: tuna and dolphins in the Eastern Tropical Pacific Ocean.” Marine Ecology Progress Series 458, (2012): 283–302.

Sharp G. D. “Behavioral and physiological properties of tunas and their effects on vulnerability to fishing gear”. In: Sharp GD, Dizon AE (eds) The physiological ecology of tunas. Academic Press, New York, (1978): pp 397±449.

Shih C.-L., Hsu C.-C. and Chen C.-Y. “First attempt to age yellowfin tuna, *Thunnus albacares*, in the Indian Ocean, based on sectioned otoliths.” Fisheries Research 149, (2014): 19–23.

Sibert J. and Hampton J. “Mobility of tropical tunas and the implications for fisheries management.” Marine Policy 27, (2003): 87–95.

Soares R. X., Bertollo L. A. C., Félix da Costa G. W. W. and Franco Molina W. “Karyotype stasis in four atlantic Scombridae fishes: mapping of classic and dual-color fish markers on chromosomes.” *Fisheries Science* 79, (2013): 177–83.

Stéquert B., Rodriguez J. N., Cuisset B. and Le Menn F. “Gonadosomatic Index and seasonal variations of plasma sex steroids in skipjack tuna (*Katsuwonus pelamis*) and yellowfin tuna (*Thunnus albacares*) from the Western Indian Ocean.” Aquatic Living Resources 14, (2001): 313–18.

Stéquert B., Panfili J. and Dean J. “Age and growth of yellowfin tuna, *Thunnus albacares*, from the western Indian Ocean, based on otolith microstructure”. Fishery Bulletin 94, (1996): 124–134.

Sun C. L., Wang W.-R and Yeh S. “Reproductive biology of yellowfin tuna in the central and western Pacific Ocean.” WCPFC-SC1. BI WP-1, 2005.

Suzuki A. “On the blood types of yellowfin and bigeye tuna.” The American Naturalist 96, no. 889 (1962): 239–46.

Takashi T., Kohno H., Sakamoto W., Miyashita S., Murata O. and Sawada Y. “Diel and ontogenetic body density change in Pacific bluefin tuna, *Thunnus orientalis* (Temminck and Schlegel), larvae.” Aquaculture Research 37, (2006): 1172–79.

Timochina O. I. and Romanov E. V. “Notes on reproduction biology of yellowfin tuna in the Western Indian Ocean,” Vol. Work. Doc., IPTP, TWS (1991).

Tocher D. R. “Metabolism and functions of lipids and fatty acids in teleost fish.” Reviews in Fisheries Science 11, (2003): 107–84.

Viñas J., Gordoa A., Fernández-Cebrián R., Pla C., Vahdet Ü. and Araguas R. M. “Facts and uncertainties about the genetic population structure of atlantic bluefin tuna (*Thunnus thynnus*) in the Mediterranean. Implications for fishery management.” Reviews in Fish Biology and Fisheries 21, (2011): 527–41.

Viñas J. and Tudela S. “A validated methodology for genetic identification of tuna species (genus *Thunnus*).” PLoS ONE 4, (2009): e7606.

Wang S., Meyer E., McKay J. K. and Matz M. V. “2b-RAD: a simple and flexible method for genome-wide genotyping.” Nature Methods 9, (2012): 808–10.

Waples R. S. “Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species.” Journal of Heredity 89, (1998): 438–50.

Ward R. D. “Genetics in Fisheries Management.” Hydrobiologia 420, (February 1, 2000): 191–201.

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

Ward R. D. “Population genetics of tunas.” Journal of Fish Biology 47, (1995): 259–80.

Ward R. D., Elliott N. G., Grewe P. M. and Smolenski A. J. “Allozyme and mitochondrial DNA variation in yellowfin tuna (*Thunnus albacares*) from the Pacific Ocean.” Marine Biology 118, (1994): 531–39.

Wexler J. B., Chow S., Wakabayashi T., Nohara K. and Margulies D. “Temporal variation in growth of yellowfin tuna (*Thunnus albacares*) larvae in the Panama Bight, 1990-1997.” Fishery Bulletin 105, (2007): 1–18.

Wexler J. B., Margulies D. and Scholey V. P. “Temperature and dissolved oxygen requirements for survival of yellowfin tuna, *Thunnus albacares*, larvae.” Journal of Experimental Marine Biology and Ecology 404, (2011): 63–72.

Worm B., Barbier E. B., Beaumont N., Duffy J. E., Folke C., Halpern B. S., Jackson J. BC, et al. “Impacts of biodiversity loss on ocean ecosystem services.” Science 314, (2006): 787–90.

Wu G. C.-C., Chiang H.-C., Chou Y.-W., Wong Z.-R., Hsu C.-C., Chen C. Y. and Yang H.-Y.” Phylogeography of yellowfin tuna (*Thunnus albacares*) in the Western Pacific and the Western Indian Oceans inferred from mitochondrial DNA”. Fish. Res., 105, (2010): 248–253).

Yabuta Y., Yukinawa M. and Warashina Y. “Growth and age of yellowfin tuna: age determination (scale method)”. Report of Nankai Regional Fisheries Research Laboratory 12 (1960): 63–74.

Yang R., Nose Y. and Hiyama Y. “A comparative study on the age and growth of yellowfin tunas from the Pacific and Atlantic Oceans”. Far Seas Fisheries Research Laboratory Bulletin 2, (1969): 1–21.

Zagaglia C. R., Lorenzzetti J. A. and Stech J. L. “Remote sensing data and longline catches of yellowfin tuna (*Thunnus albacares*) in the equatorial Atlantic.” Remote Sensing of Environment 93, (2004): 267–81.

Zhu G., Xu L., Zhou Y. and Song L. “Reproductive biology of yellowfin tuna *T. albacares* in the West-Central Indian Ocean.” Journal of Ocean University of China 7, (2008): 327–32.

Zudaire I., Murua H., Grande M., Goñi N., Potier M., Ménard F., Chassot E. and Bodin N. “Variations in the diet and stable isotope ratios during the ovarian development of female yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean.” Marine Biology 162, (2015): 2363–77.

Zudaire I., Murua H., Grande M., Pernet F. and Bodin N. “Accumulation and mobilization of lipids in relation to reproduction of yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean.” Fisheries Research, (2014).

Zudaire I., Murua H., Grande M., Korta M., Arrizabalaga H., Areso J. J. and Delgado-Molina A. “Fecundity regulation strategy of the yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean.” Fisheries Research 138 (2013a): 80–88.

Zudaire I., Murua H., Grande M. and Bodin N. “Reproductive potential of yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean.” Fishery Bulletin 111, (2013b): 252–64.

Zhu G., Xu L., Zhou Y. and Song L. “Reproductive biology of yellowfin tuna *T. albacares* in the West-Central Indian Ocean”. Journal of Ocean University of China 7, (2008): 327–332.

2 SECTION.1

Section.1 preludes the next three chapters, which aim to:

Chapter 2) investigate the global YFT population structure using a panel of 14 microsatellite loci recently isolated for *Thunnus albacares* and for *Thunnus maccoyi*;

Chapter 3) assess the 2b-RAD genotyping technique (Wang et al., 2012) in providing reliable genomic data to analyze the genetic diversity and differentiation patterns of this highly pelagic species;

Chapter 4) reveal the YFT genomics population structure across its entire distribution range, using genome-wide single-nucleotide polymorphisms (SNPs).

The main goal of this section is to report those phases shared among the three next Chapters (i.e. sampling design, DNA extraction) in order to avoid redundancy, which of course would decrease the readability of each specific Chapter.

2.1 Sampling design

A sampling design, unheard-of for YFT population structure studies, was realized at the global scale (between- and within-ocean) in order to assess stock boundaries and stock units, covering the entire species distribution range (Fig. 2.1).

Fig. 2.1_Location of *Thunnus albacares* geographical population samples analyzed in this study. Sample codes are given as in Table 2.1.

This sampling design relied upon a variety of sampling sources: (i) boarding of collaborator members on fishing vessels, (ii) collaborations with other research projects including scientific campaigns-at-sea, and (iii) sampling at port of unloading and at tuna canneries. The complementarity of these approaches has ensured the safe running of the project. A total of 500 juveniles YFT from ten geographic population samples were collected with exact information about their fishing locations during 2013 and 2014 (Table 2.1). I am aware that a multi-annual sampling to generate temporal replicates from the same areas is necessary to test and validate the population structure detected. However, due to the global scale of sampling required a great effort across oceans and institutions and the relatively limited time allowed for the genetic and genomic analyses to be accomplished within the three-years Doctoral program, I have targeted a spatial sampling with a unique temporal replicate. This represents a caveat that should be taken into account when results will be obtained and discussed. Skeletal muscle or finclip specimens were collected from individual YFT and fixed in 96% ethanol or in DMSO/EDTA/saturated sodium chloride (DESS). Specimens were stored at 4°C for few weeks and at -20° C for several months until DNA extraction. The targets of the sampling were YFT juveniles ranged in size from 35 to 55 cm FL. Targeting juveniles YFT increases the likelihood to catch them close to their spawning areas, due to their reduced capability of long distance migrations.

Table 2.1_The table summarizes the sampling location, sample code, the coordinates, number of individuals, the tissue sample, the fishing gear used per each geographic population sample.

Location	Sample code	Latitude	Longitude	N. of individuals	Tissue	Fishing gear	Provider	*Institute*
W Atlantic Ocean	31_1	28.44	-88.22	50	fin clip	Hook and line	Jay Roocker	Texas A&M University
E Atlantic Ocean	34_1	-11.13	11.61	50	fin clip	Purse seiner	Carlo Pecoraro	UNIBO
E Atlantic Ocean	34_2	-2.14	-16.12	50	fin clip	Purse seiner	Carlo Pecoraro	Unibo
W Atlantic Ocean	41_1	11.14	-65.3	50	muscle	Bait boat	Freddy Arocha	Instituto Oceanográfico de Venezuela
W Indian Ocean	51_1	-11.47	42.38	50	fin clip	Purse seiner	Nathalie Bodin	IRD
W Indian Ocean	51_2	1.02	49.97	50	fin clip	Purse seiner	Nathalie Bodin	IRD
WC Pacific Ocean	71_1	-4.64	154.05	50	fin clip	Pole and line	Bruno Leroy	Secretariat of the Pacific Community
WC Pacific Ocean	71_2	-3.4	166.36	50	fin clip	Purse seiner	Jeff Muir	Hawaii University/ISSF
E Pacific Ocean	77_1	24.5	-113.25	50	muscle	Purse seiner	Sofía Ortega	Institute Politécnico National Maxico City
E Pacific Ocean	77_2	28.86	-116.06	50	muscle	Purse seiner	Sofía Ortega	Institute Politécnico National Maxico City

2.2 DNA extraction

From each individual the total genomic DNA (gDNA) was extracted from approximately 20 mg of tissue using the commercial Invisorb® Spin Tissue Mini Kit (Invitek, STRATEC Biomedical, Germany). Tissue samples were lysed, adding 400 µL of Lysis Buffer G and 40 µL of Proteinase K, in a thermomixer overnight at 50°C. After lysate centrifugation for 2 min at 11.000 rpm, the supernatant was transferred into a new tube, adding 40 µL of RNAse (10 mg/ml) and 200 µL of Binding Buffer A. The lysate was then transferred onto an Invisorb Spin Filter to absorb the gDNA into its membrane, discarding the filtrate containing possible contaminants. In order to efficiently remove any remaining contaminants and enzyme inhibitor, two washing steps with 500 µL of Wash Buffer were performed. In addition, for eliminating any trace of ethanol, the filtrate was firstly centrifuged for 4 min at 14.000 rpm and then left at room temperature for 1 h. Lastly, the Spin Filter was placed into a 1.5 mL Receiver Tube and added with 80 µL of pre-warmed Elution Buffer to increase the final extracted gDNA volume. The eluted DNA was incubated at room temperature for 3 min and centrifuged for 1 min at 11.000 rpm, discarding the Spin Filter at the end. For each sample, the concentration (ng/µL) and purity of the genomic DNA extracted were quantified by both a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and a Qubit 2.0 Fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, Massachusetts, USA). This procedure ensured to work with high quality samples and comparable DNA concentration.

2.3 Genetic species identification of yellowfin tuna juveniles

Given that juvenile YFT, especially those with F_L < 40 cm, are very similar morphologically with bigeye tuna (BET) juveniles (Pedrosa-Gerasmio et al., 2012), and we did not know personally the sampling provider in the Central-Western Pacific Ocean (sample code = 71-2-Y), each individual from that area was specifically identified using the mitochondrial DNA segment flanking ATPase and cytochrome oxidase subunit III genes (ATCO region; Chow and Inoue, 1993). Sixty-six specimens from that area were selected for genetic analysis and compared with five BET sampled in the Eastern Atlantic Ocean by trained researchers. Moreover, two homologous YFT sequences

(Acc. N. KM055398.1 and AF115278.1), one of BET (Acc. N. AF115274.1) and one of skipjack (*Katsuwonus pelamis;* SKJ; Acc. N. GU256527.1, used as outgroup) were retrieved from the GenBank (www.ncbi.nlm.nih.gov/genbank).

The primers used in this study were from Chow and Inoue (1993): L8562 (5'-CTTCGACCAATTTATGAGCCC-3') and H9432 (5'-GCCATATCGTAGCCCTTTTTG-3').

The PCR amplification was carried out by a 50 µL reaction mixture; 1x GoTaq Flexi Buffer (Promega) 1.5 mM $MgCl_2$, 0.4mM dNTPs, 1µM each primer, 2 U of *Taq* Dna Polymerase (Promega), 4 µL (~20 ng) of DNA template. The reaction mixture was pre-heated at 94°C for 2 minutes followed by 35 cycles of amplification (93°C for 1 min, 52-57,8°C for 1 min and 72°C for 45 sec) with a final extension at 72°C for 8 min (Fig. 2.2).

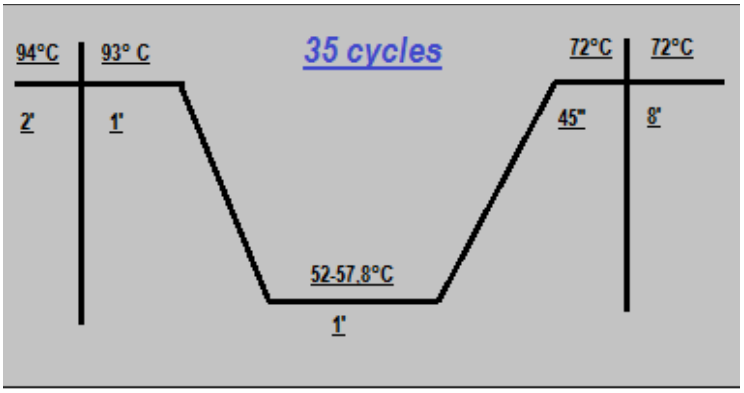


Fig. 2.2. PCR cycling conditions for ATCO amplification

The ATCO sequences obtained from Macrogen Europe (Amsterdam, Netherlands) were edited and aligned with the ClustalW algorithm implemented in the software MEGA6 (Tamura et al., 2013) and trimmed to 827 bp.

The sequences obtained in this study were homologous with those retrieved from GenBank, allowing the species identification of each individual. We reconstructed a Neighbour-Joining tree (Fig. 2.3) based on the Tamura-Nei distances. Such phylogenetic tree clearly discriminated individual sequences in the YFT and BET clusters with very high bootstrap support (95% and 99% for YFT and BET clades, respectively).

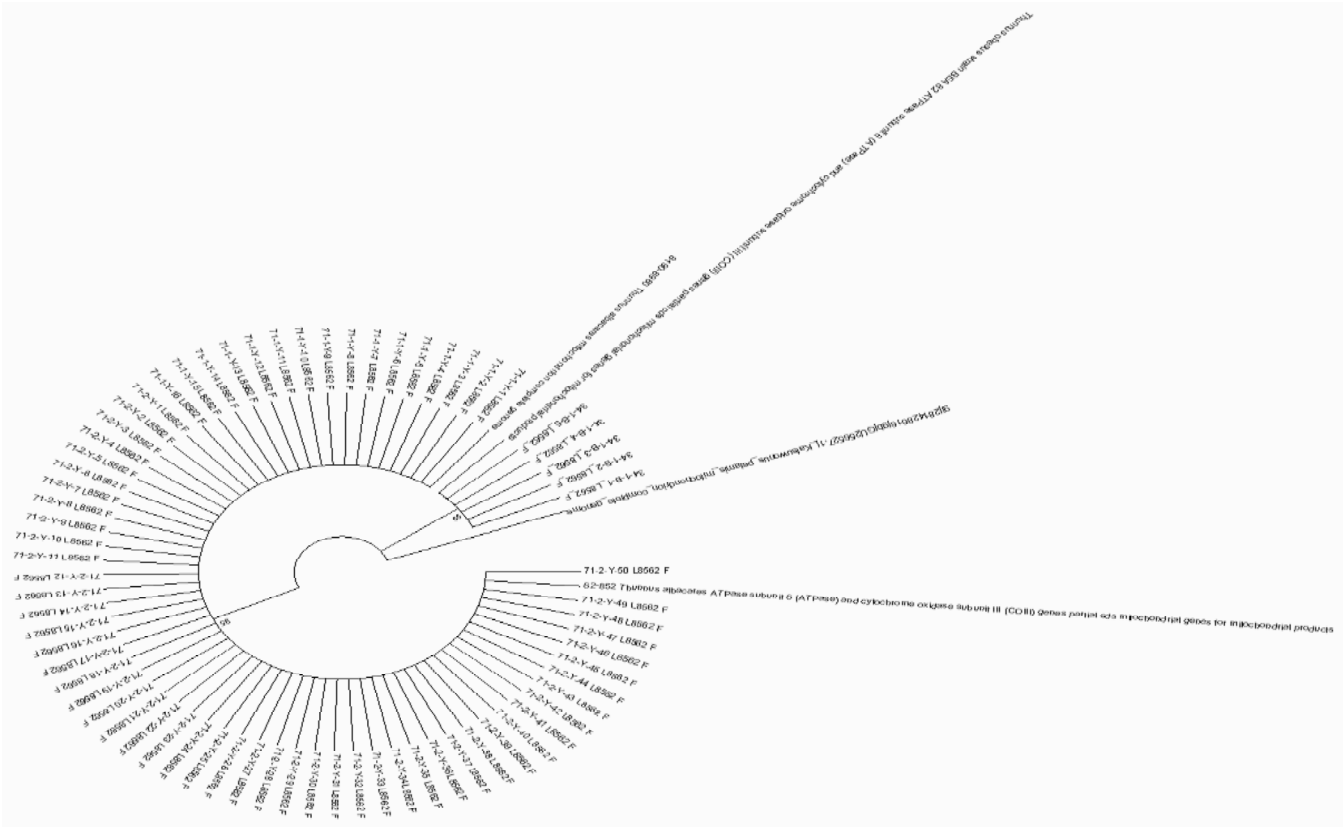


Figure 2.3. Neighbour-Joining tree based on the ATCO sequences (827 bp) obtained in this study and retrieved from GenBank. The Tamura-Nei (1993) distance method was used and bootstrap analysis, based on 1,000 re-samplings, was carried out.

Our results confirmed and supported the use of the ATCO sequence marker to specifically identify and discriminate between YFT and BET and the complete correspondence between the molecular and morphological identification of the specimens.

References

Chow S. and Inoue S. "Intra-and interspecific restriction fragment length polymorphism in mitochondrial genes of *Thunnus* tuna species." Bull Natl Res Inst Far Seas Fish 30 (1993): 207–25.

Pedrosa-Gerasmio I. R., Babaran R. P. and Santos M. D. "Discrimination of juvenile yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) tunas using Mitochondrial DNA Control Region and liver morphology". PLoS ONE (2012).

Tamura K., Stecher G., Peterson D., Filipski A., Kumar, S. "MEGA6: Molecular Evolutionary Genetics Analysis version 6.0". Mol Biol Evol 30, (2013): 2725-2729.

Wang S., Meyer E., McKay J. K. and Matz M. V. "2b-RAD: a simple and flexible method for genome-wide genotyping." Nature Methods 9, (2012): 808–10.

3 Chapter 2

YFT population structure inferred from microsatellite loci

3.1 Introduction

The definition of proper management units is crucial for a sustainable exploitation of fish stocks (Cadrin et al., 2014). However, fish management units are traditionally based on territorial jurisdictions (Reiss et al., 2009) and the same definition of "stock" reflects practical administrative necessities rather than biological realities. This management approach trivializes the complex biological and ecological processes involved in shaping the population dynamics and structure (Cuéllar-Pinzón et al., 2015). The major risk of ignoring the proper population structure is the resource mismanagement, which could have potential negative effects on the biological capacity of fish populations to adapt to changing environmental conditions and, thereby, to recover from over-exploitation (Worm et al., 2006). In such a context, it is crucial to properly assess the fish stock status with all available scientific and technological tools, in order to avoid any potential extinction risk of over-harvested populations (Antoniou and Magoulas, 2014).

The application of genetics to fisheries management has increased over the past six decades (Waples et al., 2008), together with the concerns about the potential genetic changes in harvested marine fish induced by their exploitation (Laikre and Ryman, 1996; Law, 2007; Allendorf et al., 2008; Palkovacs, 2011). Since the 1970s the continuous development of new molecular techniques has improved our capability to investigate useful DNA variations, providing new insights into the spatial dynamics, migratory movements and population structure of marine exploited species, in order to address critical conservation and fisheries management issues, such as stock structure and population demography (Hauser and Carvalho, 2008). However, the great contribution that genetic studies might lead to fish stock management is counterbalanced by some methodological limits in the capability of detecting significant heterogeneity among populations. In marine fishes, population differentiation through genetic drift is generally weak (Hauser and Carvalho, 2008), and very low degree of genetic differentiation represents a role rather than an exception (Ward et al., 1995; Ward, 2000). A common measure for describing the distribution of genetic diversity within a species are the fixation indices (Figure 3.1), first defined by Wright (1951).

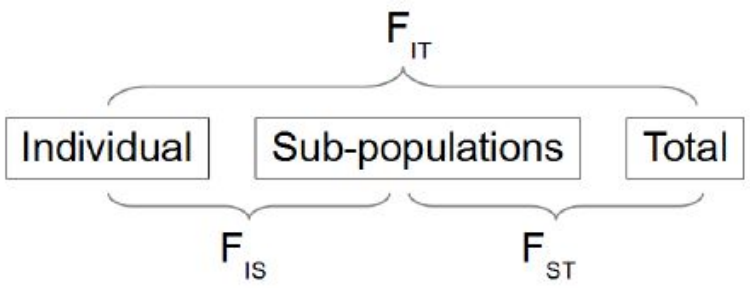


Figure 3.1. The fixation indices used to estimate the differentiation of populations (F_{ST}), inbreeding within sub-populations (F_{IS}), and genetic diversity of an entire species (F_{IT}).

The lower degree of genetic differentiation of marine species than freshwater or anadromous species (Ward, 2000) was demonstrated in 57 marine fish species, in which the mean and median F_{ST} values were of 0.062 and 0.02, respectively (Waples, 1998). These levels of genetic differentiation are certainly expected to be lower, at least within each oceanic basin, in highly migratory species, such as YFT (Ward, 2000).

Regarding the different genetic markers employed for revealing population structure in marine fish species, microsatellites have become the marker of choice in the last decade, according to the high polymorphism. These markers are highly variable and are characterized by short tandemly repeated di-, tri-, or tetra-nucleotide sequences of 1–6 bp, which have a characteristic mutational behaviour (Kelkar et al., 2010). They have a high allelic diversity and are widely distributed across the nuclear genomes of eukaryotes (Bhargava and Fuentes, 2010). The repetitive elements are typically composed by 2-4 nucleotides (DeWoody and Avise, 2000) and the dinucleotide repeats -GT- and -CA- are considered the most common microsatellites in vertebrates (Zardoya and Mayer, 1996). Moreover, these markers are characterized by a relative ease of transfer between closely related species (Guichoux et al., 2011). The high rate of variation in microsatellite is explained by the strand slippage (slipped strand mispairing) during DNA replication and unequal crossing-over during recombination, which is manifested by a change of repeat unit(s) (Levinson and Gutman, 1987). The stepwise mutation model (SMM), in which each mutational event is reflected by the addition or the loss of a single repeat unit, is the most applicable theoretical description of the microsatellite mutation process (Slatkin, 1995). The direct consequence of SMM is the higher homoplasy for microsatellites, because of their allele size constraints and high mutation rates (Estoup et al., 2002), compared to other markers, which is often cited as a significant drawback in the use of microsatellites as genetic markers (Estoup et al., 2002). Briefly, homoplasy occurs when two individuals with different ancestries at a locus mutate to the same allele and become identical only in state, and not by descent. Moreover, because microsatellites are exclusively genotyped by amplicon length variation, extra causes of homoplasy, that would otherwise be detectable by

direct sequencing, need to be considered (Putman and Carbone, 2014). Another common issue in using microsatellite loci is represented by the presence of null alleles, which are variation in the nucleotide sequences of flanking regions that may prevent the primer annealing to template DNA during amplification of the microsatellite locus by PCR. As consequence, the heterozygote will show a homozygote profile, while the homozygote, for the null allele, will not produce any amplicon.

To date, there are only few published microsatellites studies on YFT stock structure and most of them were focused on the Pacific Ocean. In the first work, Appleyard et al., (2001) used five dinucleotide microsatellite loci for investigating the YFT population structure within the Pacific Ocean, indicating the total absence of genetic differentiation between the eastern and the western samples. Later, Dìaz-Jaimes and Uribe-Alcocer (2006), focusing only on the population structure in the Eastern Pacific Ocean, employed seven microsatellite markers. Although their results suggested the existence of discrete northern and southern populations separated by the equator, the authors cautioned, however, that this spatial differentiation could be just related to the temporal variation or to the non-random sampling (Diaz-Jaimes and Uribe-Alcocer, 2006). More recently, significant genetic differentiation was detected between central and western Pacific samples (F_{ST} = 0.034; p value = 0.016) using nine microsatellite loci (Aguila et al., 2015) and in the north-western part around Sri Lanka (Dammannagoda et al., 2008), employing three microsatellite loci together with the mitochondrial ATPase.

Given the importance of the proper characterization of the genetic population structure and the paucity of published microsatellite studies on YFT, especially at wide geographical scale, in this Chapter YFT global population structure was inferred from a panel of microsatellite loci.

3.2 Materials and methods

3.2.1 Microsatellites loci and genotyping

In this study we used 14 microsatellite loci and 163 individuals from five geographic population samples (Fig. 3.2 and Table 3.1).

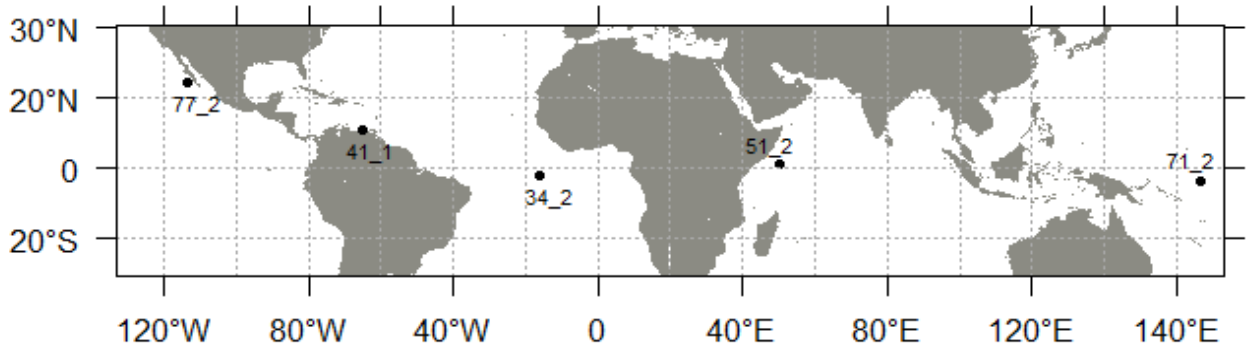


Fig. 3.2. Location of *Thunnus albacares* geographical population samples analyzed in this study.

Table 3.1_The table summarizes the corresponding sample codes and the number of individuals analysed per sample.

Sampling location	Sample code	N° of individuals
WC Pacific Ocean	**71_2**	30
E Pacific Ocean	**77_2**	37
E Atlantic Ocean	**34_2**	33
W Atlantic Ocean	**41_1**	32
W Indian Ocean	**51_2**	31

These samples were genotyped at the University of Bologna (GenoDream laboratory) using eight species-specific markers recently isolated for YFT by Antoni et al., (2014) and at CSIRO (Marine and Atmospheric Research; Hobart, Australia) using other six loci isolated for *Thunnus maccoyii* (Southern Bluefin tuna; SBT; Peter Grewe, personal communication). Among the different loci published by Antoni et al., (2014), we chose those with the same annealing temperature in order to perform multiplex amplifications with the QIAGEN Multiplex kit. The forward primer of each locus was labeled with a fluorescent probe for allowing the simultaneous electrophoretic run of more fragments in the same capillary of the automatic sequencer. These different fluorophore-labelled primers were chosen depending on their different fragments' size. Two multiplex reactions with four loci each (Multiplex 1 and Multiplex 2; Table 3.2) were optimized, after performing the preliminary optimization of PCR conditions with a few individuals.

Table 3.2. Microsatellite panel: Locus: name of the locus; Repeat: motif repeat; Primer sequence: sequence composition; AT: Annealing temperature; Label: sample labeling dyes.

Locus	Repeat	Primer sequence	AT°	Label
Multiplex 1				
YT4	(CA)	F: CGACTCCTGGCCTTGGTTA	62	PET
		R: GGAGAACAGGCCCTGAATTT		
YT84	(GATA)	F: TTACTTACTGCCGCCTACTGG	62	NED
		R: AAGTTGGAACCAGAGAACCATT		
YT87	(CATA)	F: CGAGATACGCGAAGGACAA	62	6FAM
		R: CAGTGGCATTCTTGGCAATAG		
YT111	(CA)	F: CGTCCAGTAGTGACGAGGAT	62	VIC
		R: GACTGTATCTGCTGAAACCAAA		
Multiplex 2				
YT121	(CA)	F: GAAGGCTCTTCAGCTGGTTG	62	PET
		R: AGTTGTGGCACGATTGTCTG		
YT29	(CA)	F: TGCGTCTTTGAATGGCTAA	62	NED
		R: GGTTGTCAGTCAGGACAAAGG		
YT92	(GA)	F: CCTCAGCCAAGGTGAGAAGA	62	6FAM
		R: CGCTCGCTACTACTCACTCCA		
YT12	(CA)	F: GAGATGTGGCTTCCTCCAAA	62	VIC
		R: GGCGATTTATGGCATCAGTT		

For each reaction, 2 µL of DNA were amplified in a 10 µL final volume with the following concentrations: Master mix Quiagen 1X and primers Forward and Reverse 0.20 µM each.
The Master mix contains also pre-optimized concentrations of HotStarTaq DNA Polymerase, MgCl2, dNTPs and Polymerase buffer. PCR amplifications were carried out in a T-Gradient thermocycler (Biometra) and the same temperature profile was used for all of the loci. Following Qiagen's guidelines the thermal profile consisted of: 1) an initial heat activation step at 95 °C for 15 minutes; 2) 35 cycles (denaturation: 94°C for 30 seconds, annealing: 62°C for 90 seconds, extension: 72°C for 90 seconds); 3) final extension: 72 °C for 10 minutes.

The success of the PCR reactions was verified by 1.8% agarose gel electrophoresis (0.5X TAE (20 mM Tris, 10 mM acetic acid, 0.5 mM EDTA, pH 8)), at 3 V/cm for 30 minutes.

The fragment analysis was performed by a commercial provider (Macrogen Inc, Seoul, Korea), using the GS-500LIZ size standard. The data obtained were analyzed by using the software PEAK SCANNER 1.0 (Applied Biosystems).

3.2.2 Allele calling (binning)

The next step consisted in assigning an integer, or bin, to the alleles observed according to their length in base pairs (bp). In order to optimize the binning phase and to obtain reliable binned values from the raw dataset, we used the software TANDEM (Matschiner and Salzburger, 2009). This software operates through an algorithm that fills a gap of the microsatellite workflow by rounding allele sizes to valid integers, depending on the microsatellite repeat units. The module repeat was manually established on the basis of loci motifs. For the other six microsatellite loci isolated for SBT, allele calling and the binning were carried out at CSIRO using the software GeneMapper 4. Finally, the two datasets were integrated for analyzing the genetic differentiation between and within geographic samples.

3.2.3 Data analysis

Prior to the genetic differentiation analyses, MICROCHECKER ver. 2.2.3 (van Oosterhout et al., 2004) was used to identify scoring errors as null alleles, stutter and large allele drop-out. After this quality check, the dataset was imported (in txt format) into the software GENETIX 4.05.2 (Belkhir et al*.,* 1996) to create the input file for the software FSTAT (Goudet, 2001). Using this software, it was possible to estimate the overall genetic diversity, mean number of alleles per locus (A), observed heterozygosity (Ho) and expected heterozygosity (He) per each locus and population, mean and single-population estimation of allelic richness per locus, and inbreeding coefficient (FIS) for each population. The probabilities of Hardy-Weinberg equilibrium for each locus within each population were estimated using the exact probability test, by using the software GENEPOP (on line version 4.2; Raymond and Rousset, 1995; Rousset, F., 2008). Moreover, Jackknifing over loci (GENETIX 4.05.2) was carried out for examining single-locus effects of all microsatellites loci, using the Weir & Cockerham's F-statistics estimators. Pairwise F_{ST} estimates and analysis of molecular variance (AMOVA), with 20,000 permutations and a significance level of 0,01, were obtained by using the software ARLEQUIN 3.5.1.2 (Excoffier and Lischer, 2010).

STRUCTURE software (v2.3.4; Pritchard et al., 2000; Hubisz et al., 2009), which employs a Bayesian model-based clustering algorithm, was used to further investigate the population structure, without making any *a priori* assumptions on the possible existence of different populations, and to estimate the extent of admixture by individuals and populations. The determination of the most likely number of homogenous population (K) and the estimation of the posterior probability

represent two fundamental steps for the correct setting of STRUCTURE algorithm. K value is based on the best adaptation of the genotypic data under the conditions of minimizing divergence from Hardy-Weinberg equilibrium and minimal degree of linkage disequilibrium. The posterior probability depends on the likelihood that each individual's genotype belongs to each defined cluster. On the y-axis of the resulting barplot is described the membership percentage value for each individual to a given cluster. This analysis was carried out assuming an admixture ancestry model with the geographical origin of samples as prior information (LOCPRIOR models), associated with a correlated allele frequencies model. For each simulation of K (1-12), ten independent replicates were run, setting a burn-in period of 50,000 iterations and 100,000 iterations for the MCMC. The assessment of the number of cluster K was inferred using Evanno's Δk and Pritchard's average log probability methods (Pritchard et al., 2000; Evanno et al., 2005), both implemented in the STRUCTURE HARVESTER v.0.6.93 web application (Earl and Von Holdt, 2012).

3.3 Results

The results of the MICROCHECKER analysis indicated the presence of null alleles in samples 71-2-Y (at loci YT12, YT29, YT84), 77-2-Y (at loci YT12, YT84), 34-2-Y (at loci YT12 and YT84), 41-1-Y (at loci YT12, YT29 and YT84) and 51-2-Y (at loci YT12 and YT84). However, no stuttering problems or large dropout were detected as summarized in Table 3.3.

Table 3.3_MICROCHECKER Scoring errors test: NA= Null allele presence, ST= Stuttering LD= Large Allele Dropout.

SCORING ERRORS	71-2	77-2	34-2	41-1	51-2
YT4	0	0	0	0	0
YT87	0	0	0	0	0
YT111	0	0	0	0	0
YT29	**NA**	0	0	**NA**	0
YT92	0	0	0	0	0
YT121	0	0	0	0	0
YT84	**NA**	**NA**	**NA**	**NA**	**NA**
YT12	**NA**	**NA**	**NA**	**NA**	**NA**
D235	0	0	0	0	0
3D4	0	0	0	0	0
D4D6	0	0	0	0	0
D10	0	0	0	0	0
D12	0	0	0	0	0
D211	0	0	0	0	0

Since the presence of null alleles in all samples, the loci YT12 and YT84 were removed from the dataset before going further with the analyses. The resulting dataset consisted of 163 individuals and 12 loci, which were all polymorphic. Locus YTD211 (Fig. 3.3) showed the lowest number of alleles (8) while YT111 the highest (26).

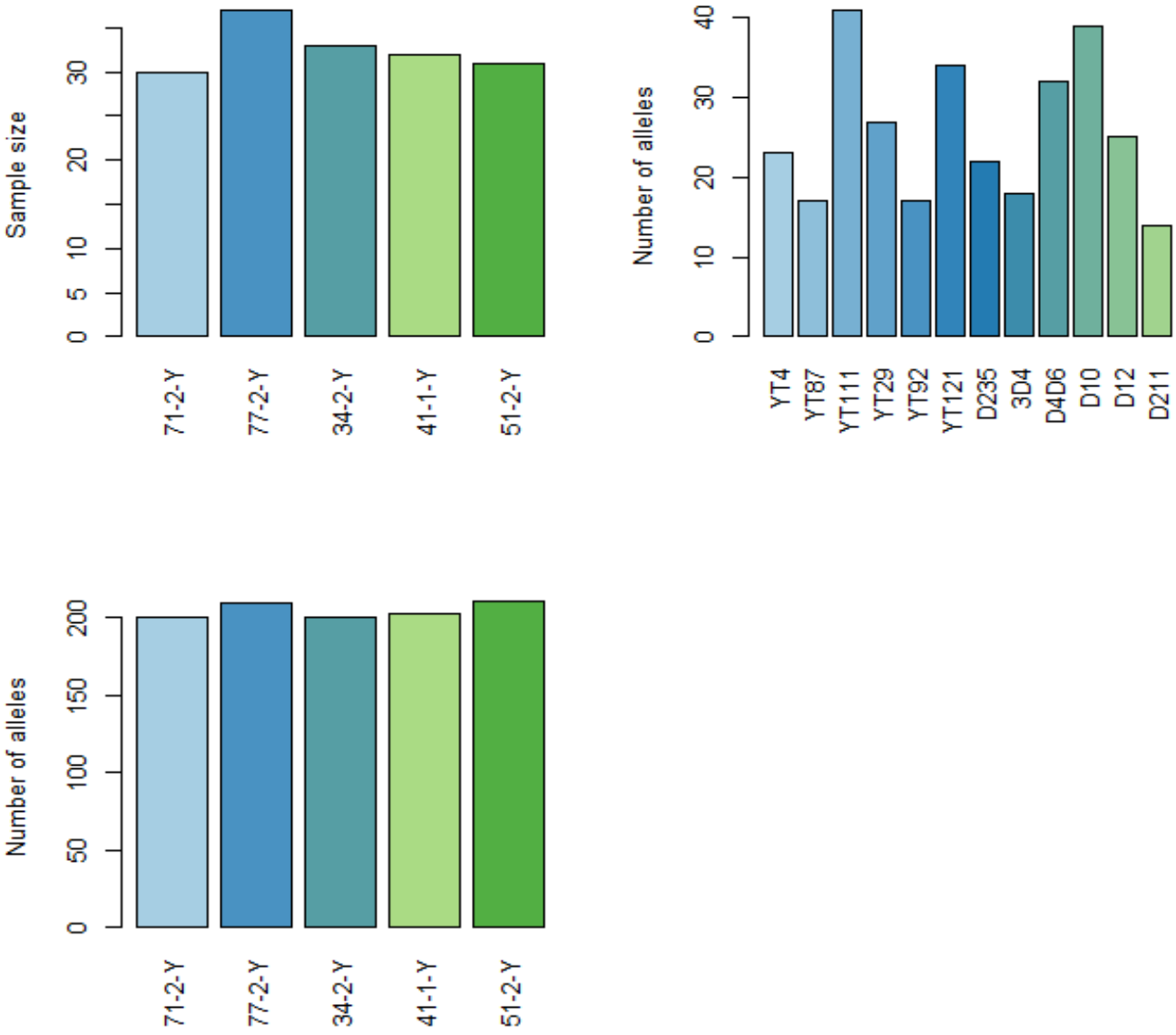


Fig. 3.3_Plots of: 1) sample size; 2) number of alleles per locus; 3) number of alleles per sample.

In order to compare different samples with different size, we calculate the allelic richness (Ar), which provides a measure of the number of alleles standardized to smaller sample size in our dataset (Table 3.4). All five geographic samples analyzed have a similar mean number of alleles per locus, ranging from 16.66 (34-2-Y) to 17.58 (51-2-Y) and a mean value of allele richness ranging from 15.78 (34-2-Y) to 16.86 (51-2-Y).

Table 3.4_Number of alleles (An) on the left compared with Allelic richness (Ar) on the right per each locus per each population sample.

An	71-2-Y	77-2-Y	34-2-Y	41-1-Y	51-2-Y	Ar	71-2-Y	77-2-Y	34-2-Y	41-1-Y	51-2-Y
YT4	15	18	16	17	15	**YT4**	14.65	16.38	15.20	16.28	14.44
YT87	10	11	12	10	12	**YT87**	9.77	10.10	11.20	9.90	11.20
YT111	19	24	24	21	25	**YT111**	18.48	21.09	22.34	19.76	23.40
YT29	17	18	15	16	18	**YT29**	16.76	16.19	14.49	15.75	17.18
YT92	11	11	7	10	10	**YT92**	10.30	9.74	6.64	9.22	9.71
YT121	23	22	23	24	23	**YT121**	22.14	19.51	21.78	22.26	22.13
D235	18	15	16	18	16	**D235**	17.48	14.16	15.32	17.35	15.44
3D4	11	10	10	12	12	**3D4**	10.58	9.17	9.42	12.00	11.81
D4D6	23	24	24	21	21	**D4D6**	22.43	21.88	22.55	21.00	20.36
D10	26	28	25	25	27	**D10**	26.00	24.86	23.35	23.77	25.66
D12	18	17	16	18	20	**D12**	17.51	15.63	15.61	17.82	19.17
D211	10	12	12	11	12	**D211**	9.89	11.32	11.39	10.71	11.85

The mean observed heterozygosity ranged from 0.82 (71-2-Y) to 0.88 (41-1-Y) instead the lower Ho value (0.60) was detected at loci YT87 and 3D4 in sample 71-2-Y, while the highest value (1.00) was identified at locus D4D6 in sample 41-1-Y. After Bonferroni correction, deviations from HWE were detected only at locus YT4 (in sample 41-1-Y) and at locus D12 (in sample 77-2-Y) (Appendix 3.2). The Jackknife test showed that all of the loci have similar F statistics values, confirming their reliability for this study (Table 3.5).

Table 3.5_Jackknife statistics analysis used for variance and bias estimation: FIS; FIT; FST.

Jackknife	F_{IS}	F_{IT}	F_{ST}
YT4	0.03036	0.03093	0.00059
YT87	0.03388	0.03478	0.00093
YT111	0.02949	0.02997	0.0005
YT29	0.02997	0.03006	0.00009
YT92	0.03273	0.03348	0.00078
YT121	0.03288	0.03418	0.00135
D235	0.03314	0.03458	0.00149
3D4	0.02653	0.02809	0.0016
D4D6	0.03035	0.03145	0.00113
D10	0.03244	0.03345	0.00104
D12	0.02471	0.02592	0.00125
D211	0.03186	0.03287	0.00104
Mean (±SD)	0.03076 (±0.008)	0.03171 (±0.008)	0.00098 (±0.001)

Pairwise Fst analysis did not reveal any significant divergence among population samples (F_{ST} = 0.00076, P value > 0.05) (Table 3.6). In fact, as pointed out by the analysis of molecular variance (AMOVA; Table 3.7) more than 99% of the variation was explained within samples.

Table 3.6_Pairwise F_{ST} values calculated among population samples of yellowfin tuna are reported (below diagonal) with their associated P-values (below diagonal). NS: not significant.

	71-2-Y	77-2-Y	34-2-Y	41-1-Y	51-2-Y
71-2-Y	*	NS	NS	NS	NS
77-2-Y	0.00066	*	NS	NS	NS
34-2-Y	0.00134	-0.00038	*	NS	NS
41-1-Y	0.00107	-0.00024	-0.00302	*	NS
51-2-Y	0.00032	0.00012	0.00111	0.00372	*

Table 3.7_AMOVA statistical analysis.

Source of variation	Variation (%)	Fixation index	Pvalue
Between Oceans	0.20	0.00076	0.43657
Within Oceans	-0.12	-0.00121	0.80546
Within samples	99.92	0.00198	0.43657

YFT population structure was further investigated running the software STRUCTURE (Table 3.8).

Table 3.8_STRUCTURE HARVESTER summary.

K	Reps	MeanLnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	ΔK
1	10	-10282.9400	1.8338	-	-	-
2	10	-10291.8600	14.5011	-8.920000	295.520000	20.379090
3	10	-10596.3000	181.1186	-304.440000	318.320000	1.757523
4	10	-10582.4200	269.5946	13.880000	357.860000	1.327400
5	10	-10926.4000	558.5932	-343.980000	123.600000	0.221270
6	10	-11146.7800	1078.4448	-220.380000	30.940000	0.028689
7	10	-11398.1000	440.7671	-251.320000	853.680000	1.936805
8	10	-12503.1000	929.5943	-1105.000000	1099.560000	1.182839
9	10	-12508.5400	1086.1228	-5.440000	357.760000	0.329392
10	10	-12871.7400	1830.7899	-363.200000	1161.140000	0.634229
11	10	-12073.8000	730.3310	797.940000	1405.200000	1.924059
12	10	-12681.0600	1215.6994	-607.260000	-	-

The number of clusters (K) that best fit our data was inferred with the Pritchard method (Pritchard et al., 2000). This approach analyzes LnP(K) trends (the logarithm of the probability of the data given K), identifying the possible K value as the nearest value to the “plateau” (Fig. 3.4). In fact, if several values of K give similar estimates of LnP(K), the smallest seems to be the most real. Any “plateau” was detected in our analysis, suggesting the lack of samples differentiation (Fig. 3.4).

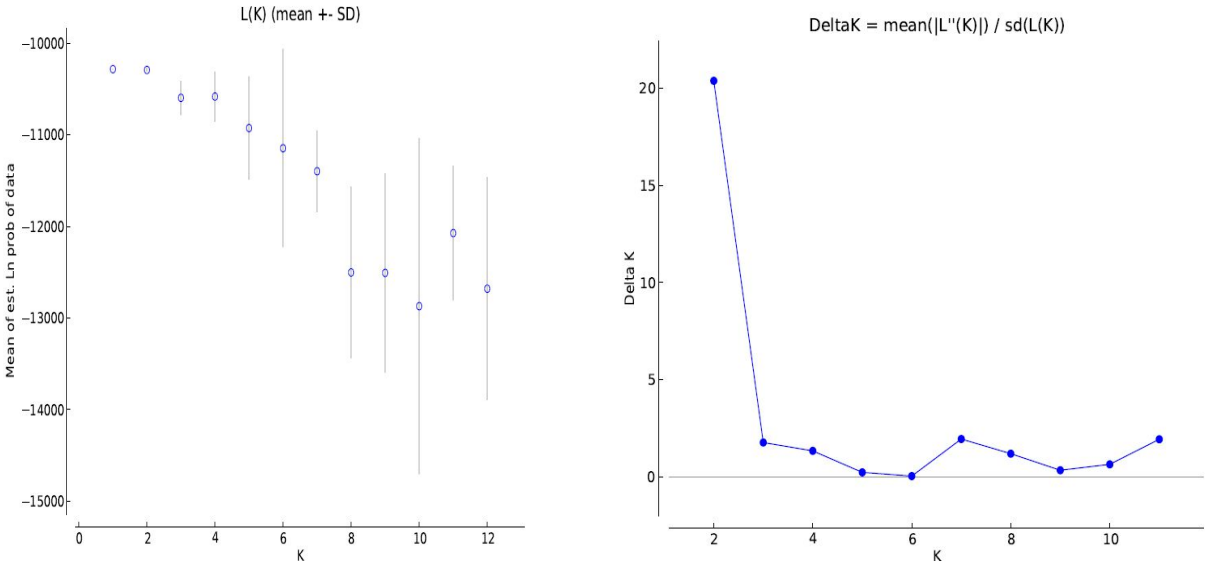


Figure 3.4. LnP(K) and Evanno's absolut ΔK graph graphs.

According to Evanno et al., (2005) the likely number of K populations can also be inferred by correlating the second rates of change of LnP" (K) with K (ΔK). This method seems to show a clear peak at the true value of K. In Fig. 3.4 ΔK values are compared to the K value obtained in this study, which indicated a maximum ΔK value at K=2. However, this result is not consistent with the previous LnP(K) assessment, representing a not realistic picture of YFT population structure revealed in the present study. In fact, as clearly shown in the resulting barplot, at K=2 it does not correspond any structuring between YFT samples (Fig.3.5).

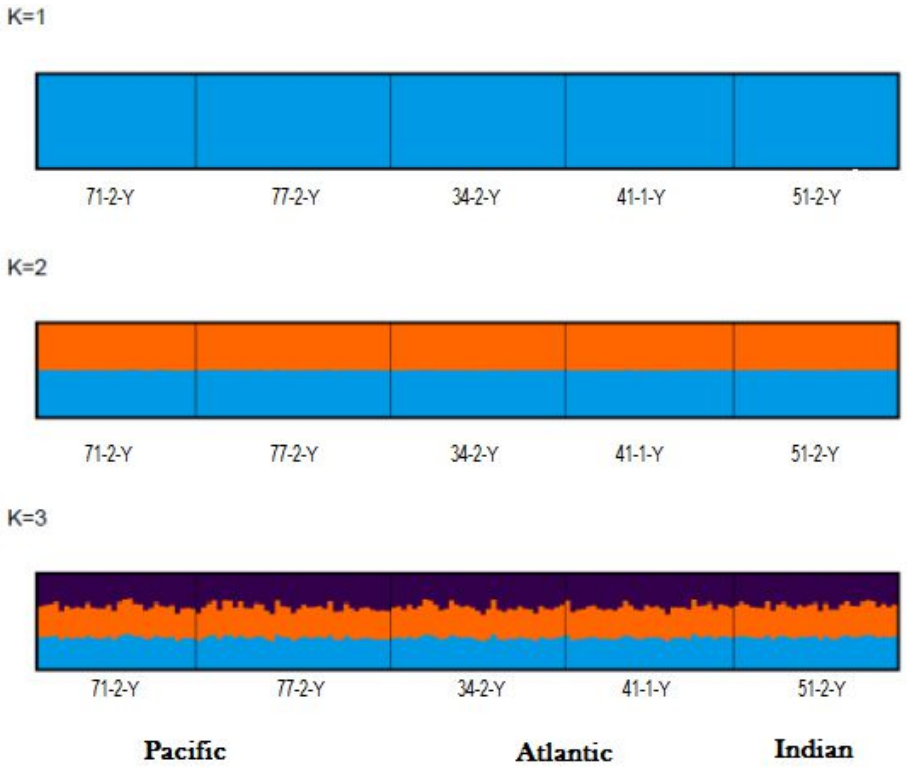


Fig. 3.5_ Estimated membership fraction of individuals from the five sampling areas.

3.4 Discussion and conclusions

In this study the global genetic population structure of YFT was inferred from microsatellite loci. All the population samples considered in this study showed a high genetic diversity, with mean observed and expected heterozygosity per locus and per population equivalent to 0.854 and 0.882, respectively. All markers were polymorphic with the mean number of alleles per locus and per population of 17.08. Significant deviations from Hardy-Weinberg equilibrium expectations (after Bonferroni correction) were detected at locus YFT4 in the sample 41-1-Y and at locus D12 in the sample 77-2-Y.

Our results showed almost null and not significant levels of genetic differentiation at both intra- and inter-oceanic levels, indicating a not significant rate of differentiation between YFT samples collected from the three oceanic basins and supporting the presence of one panmictic population worldwide. The absence of genetic differentiation, especially between Atlantic and Indo-Pacific Oceans, is in contrast with the distribution patterns identified in other tuna and tuna-like species, for which the Cape of Good Hope represents a potential genetic point break, causing significant

genetic drift. For instance, this pattern has been underlined in BET *Thunnus obesus* (Alvarado-Bremer et al*.,* 1998; Chow et al., 2000; Martinez et al*.,* 2006; Chiang et al., 2008) as well as in *Makaira nigricans* (Finnerty and Block, 1992; Graves and McDowell, 1995), *Xiphias gladius* (Alvarado Bremer et al., 1996; Chow et al., 1997; Chow and Takeyama, 2000), *Istiophorus albicans* (Graves and McDowell, 1995) and *Thunnus alalunga* (Chow and Ushiyama, 1995). However, the genetic differentiation between Atlantic and Indo-Pacific Oceans could reflect inter-oceanic migrations occurred in the past as underlined by the presence of "Pacific" mitochondrial DNA clades in Atlantic sub-populations of many tuna species (Alvarado-Bremer et al., 1998).

Our results have confirmed the patterns of homogeneity previously revealed in other studies (Ely et al., 2005; Appleyard et al., 2001), suggesting that the lack of genetic differentiation at both intra- and inter-oceanic levels could be related to a higher gene flow in YFT comparing to other pelagic species. Such issue also reflects a large effective population size (*Ne*). However, this hypothesis has not been confirmed by any studies so far.

A second possible explanation of the absence of genetic differentiation observed here, could be associated to the more recent time, compared to other pelagic species, since YFT population expansion began. Doing so, the separation time among populations was not enough to make them reproductively isolated. This alternative explanation was rejected by Ely et al., (2005), according to the calculation of YFT expansion for the hypervariable control region I (CR-I). The authors, considering a conservative mutation rate of 4.9% per million years, a generation time of 3.5 years and the tau value of 8.52, assumed that the YFT expansion started about 522,000 years ago. Using the same mutation rate, the authors suggested that a similar or even shorter time for other species, such as *Xiphias gladius* or *Thunnus thynnus*, was enough to explain their genetic differentiation between Atlantic and Mediterranean populations. So even this second hypothesis can be rejected for explaining the homogeneity revealed among oceans.

A third and more consistent explanation of lack of genetic differentiation at the global scale, in contrast with the results of some recent studies (Dammanagoda et al., 2008; Kunal et al., 2013; Li et al., 2015; Pecoraro et al., 2016; Grewe et al., 2015), may be linked to the inadequate resolution power of the markers used in the present study. In fact, although Antoni et al., (2014) stated that their six markers employed here are reliable for forthcoming YFT conservation studies, our results pointed out their inefficiency in detecting possible genetic differentiation among geographic population samples at the global scale. Even though, in this study we confirmed the relative ease of transfer of microsatellite loci between closely related species, especially within the genus

Thunnus, however the same discussion about the markers' inefficiency can be extended to the other six markers, recently isolated and still unpublished for Southern Bluefin tuna by P. Grewe at CSIRO.

In conclusion, our results cannot reject the null hypothesis of the existence of only one panmictic population at the global scale. According to the contrasting information gathered from other studies so far, there is a need to develop appropriate and non-neutral (adaptive) DNA markers (i.e. SNPs) to investigate YFT population structure both at the local and also at the global scale. In fact, a major hazard of assuming global and local panmixia, due to the markers' inability in pointing out possible genetic differentiation, is the over-exploitation of some populations with profound implications for their management and conservation.

References

Allendorf F., England P. R., Luikart G., Ritchie P. A. and Ryman N. “Genetic effects of harvest on wild animal populations.” *Trends in Ecology & Evolution* 23, (2008): 327–37.

Alvarado-Bremer J. R., Mejuto J., Greig T. W. and Ely B. “Global population structure of the swordfish (*Xiphias gladius* L.) as revealed by analysis of the mitochondrial dna control region.” *Journal of experimental marine biology and ecology* 197, (1996): 295–310.

Alvarado-Bremer J. R., Stéquert B., Robertson N. W. and Ely B. “Genetic evidence for inter-oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations.” *Marine Biology* 132, (1998): 547–57.

Aguila R. D., Sweedy Perez K. L. , Catacutan B. J. N., Lopez G. V., Barut N. C. and Santos M. D. “Distinct yellowfin Tuna (*Thunnus albacares*) stocks detected in Western and Central Pacific Ocean (WCPO) using DNA microsatellites.” PloS One 10, (2015): e0138292.

Antoni L., Luque P. L., Naghshpour K. and Saillant E. A. “Polymorphic microsatellite markers for the yellowfin tuna (*Thunnus albacares*).” Conservation Genetics Resources 6, (2014): 609–11.

Antoniou A. and Magoulas A. “Application of mitochondrial dna in stock identification.” *Stock Identification Methods, 2nd Edition. Academic Press, San Diego, California*, 2014, 257–95.

Appleyard S., Grewe P., Innes B. and Ward R. “Population structure of yellowfin tuna (*Thunnus albacares*) in the Western Pacific Ocean, inferred from microsatellite loci.” *Marine Biology* 139, (2001): 383–93.

Belkhir K., Borsa P., Chikhi L., Raufaste N. and Bonhomme F. “GENETIX 4.05, logiciel sous windows tm pour la génétique des populations.” *Laboratoire génome, populations, interactions, CNRS UMR* 5000 (1996): 1996–2004.

Bhargava A. and F. Fuentes F. “Mutational dynamics of microsatellites.” *Molecular biotechnology* 44, (2010): 250–66.

Cadrin S. X., Karr L. A. and Mariani S. “Stock Identification Methods.” In *Stock Identification Methods*, 1–5. Elsevier, 2014.

Chow S., Okamoto H., Miyabe N., Hiramatsu K. and Barut N. “Genetic Divergence between atlantic and indo-pacific stocks of bigeye tuna (*Thunnus obesus*) and admixture around South Africa.” *Molecular Ecology* 9, (2000): 221–27.

Chow S., Okamoto H., Uozumi Y., Takeuchi Y. and Takeyama H. “Genetic stock structure of the swordfish (*Xiphias gladius*) inferred by pcr-rflp analysis of the mitochondrial dna control region.” *Marine Biology* 127, (1997): 359–67.

Chow, S. and Takeyama H. “Nuclear and mitochondrial dna analyses reveal four genetically separated breeding units of the swordfish.” *Journal of Fish Biology* 56, (2000): 1087–98.

Chow, S. and Ushiama H. “Global population structure of albacore (Thunnus alalunga) inferred by RFLP analysis of the mitochondrial ATPase gene.” *Marine Biology* 123, (1995): 39–45.

Cuéllar-Pinzón J., Presa P., Hawkins S. J. and Pita A. “Genetic Markers in Marine Fisheries: Types, Tasks and Trends.” *Fisheries Research* 173 (2016): 194–205.

Dammannagoda S. T., Hurwood D. A. and Mather P. B. “Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin tuna (*Thunnus albacares*) from the North Indian Ocean around Sri Lanka.” *Fisheries Research* 90, (2008): 147–57.

DeWoody J. A. and Avise J. C. “Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals.” *Journal of Fish Biology* 56, (2000): 461–73.

Díaz-Jaimes P. and Uribe-Alcocer M. “Spatial differentiation in the Eastern Pacific yellowfin tuna revealed by microsatellite variation.” Fisheries Science 72, no. 3 (2006): 590–96.

Ely B., Viñas J., Alvarado-Bremer J. R., Black D., Lucas L., Covello K., Labrie A. V. and Thelen E. “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 Evolutionary Biology* 5, (2005): 19.

Estoup A., Jarne P., Cornuet J.M. “Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis”. *Molecular Ecology*. 11, (2002): 1591-1604.

Evanno G., Regnaut S. and Goudet J. “Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.” *Molecular Ecology* 14, (2005): 2611–20.

Excoffier L. and Lischer H. E. “Arlequin suite Ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows.” *Molecular Ecology Resources* 10, (2010): 564–67.

Finnerty, J. R. and Block B. A. “Direct sequencing of mitochondrial DNA detects highly divergent haplotypes in blue marlin (*Makaira nigricans*).” *Molecular Marine Biology and Biotechnology* 1, (1992): 206–14.

Goudet, J. *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3)*, 2001.

Graves J. E. and McDowell J. R. “Inter-ocean genetic divergence of istiophorid billfishes.” *Marine Biology* 122, (1995): 193–203.

Grewe P. M., Feutry P., Hill P. L., Gunasekera R. M., Schaefer K. M., Itano D. G., Fuller D. W., Foster S. D and Davies C. R. “Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource.” *Scientific Reports* 5 (2015).

Guichoux E., Lagache L., Wagner S., Chaumeil P., Léger P., Lepais O., Lepoittevin C. et al. “Current trends in microsatellite genotyping.” *Molecular Ecology Resources* 11, (2011): 591–611.

Hauser L. and Carvalho G. R. “Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts.” Fish and Fisheries 9, (2008): 333–62.

Hubisz M. J., Falush D., Stephens M. and Pritchard J. K. “Inferring weak population structure with the assistance of sample group information.” *Molecular Ecology Resources* 9, (2009): 1322–32.

Kelkar Y. D., Strubczewski N., Hile S. E., Chiaromonte F., Eckert K. A. and Makova K. D. “What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats.” *Genome Biology and Evolution*, (2010): evq046.

Kunal S. P., Girish Kumar, Maria Rosalia Menezes, and Ram Murti Meena. “Mitochondrial DNA Analysis Reveals Three Stocks of Yellowfin Tuna Thunnus Albacares (Bonnaterre, 1788) in Indian Waters.” *Conservation Genetics* 14, (2013): 205–13.

Laikre, L. and Ryman N. “Effects on intraspecific biodiversity from harvesting and enhancing natural populations.” *Ambio*, (1996): 504–9.

Law R. “Fisheries-induced evolution: present status and future directions.” *Marine Ecology Progress Series* 335, (2007): 271–77.

Levinson G. and Gutman G. A. “Slipped-strand mispairing: a major mechanism for DNA sequence evolution.” *Molecular Biology and Evolution* 4, (1987): 203–21.

Li W., Chen X., Xu Q., Zhu J., Dai X. and Xu L. “Genetic population structure of *Thunnus albacares* in the Central Pacific Ocean Based on mtDNA COI Gene Sequences.” *Biochemical Genetics* 53, (2015): 8–22.

Martínez P., González E. G., Castilho R. and Zardoya R. “Genetic diversity and historical demography of atlantic bigeye tuna (*Thunnus obesus*).” *Molecular Phylogenetics and Evolution* 39, (2006): 404–16.

Matschiner M. and Salzburger W. “TANDEM: Integrating automated allele binning into genetics and genomics workflows.” *Bioinformatics* 25, (2009): 1982–83.

Palkovacs E. P. “The overfishing debate: an eco-evolutionary perspective.” *Trends in Ecology & Evolution* 26, (2011): 616–17.

Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly. “Inference of Population Structure Using Multilocus Genotype Data.” *Genetics* 155, no. 2 (2000): 945–59.

Putnam A. I. and Carbone I. “Challenges in analysis and interpretation of microsatellite data for population genetic studies”. Ecology and Evolution, 4, (2014): 4399-4428.

Raymond M. and Rousset F. “GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism.” *Journal of Heredity* 86, (1995): 248–49.

Reiss H., Hoarau G., Dickey-Collas M. and Wolff W. J. “Genetic population structure of marine fish: mismatch between biological and fisheries management units.” Fish and Fisheries 10, (2009): 361–95.

Rousset F. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol. Ecol. Resources 8, (2008): 103-106.

Slatkin M.” A measure of population subdivision based on microsatellite allele frequency”. Genetics, 139, (1995): 457-462.

Van Oosterhout C., Hutchinson W. F., Wills D. PM and Shipley P. “MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data.” *Molecular Ecology Notes* 4, (2004): 535–38.

Waples R. S. “Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species.” *Journal of Heredity* 89, (1998): 438–50.

Waples R. S, Punt A. E. and Cope J. M. “Integrating Genetic data into management of marine resources: how can we do it better?” *Fish and Fisheries* 9, (2008): 423–49.

Ward R. D. “Genetics in Fisheries Management.” Hydrobiologia 420, (2000): 191–201.

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

Ward R. D. “Population genetics of tunas.” Journal of Fish Biology 47, (1995): 259–80.

Worm B., Barbier E. B., Beaumont N., Duffy J. E., Folke C., Halpern B. S., Jackson J. BC, et al. “Impacts of biodiversity loss on ocean ecosystem services.” Science 314, (2006): 787–90.

Wright, S.” The genetical structure of populations”. Ann. Eugen. 15, (1951): 323-354.

Zardoya R. and Meyer A. “Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates.” Molecular Biology and Evolution 13, (1996): 933–42.

Appendix 2.1

Allelic frequencies of the 12 microsatellites analyzed

LOCUS	*Geographic samples*				
	71-2-Y	77-2-Y	34-2-Y	41-1-Y	51-2-Y
YT4					
(N)	30	37	33	32	31
Allele					
200				0.0313	0.0161
202		0.0405	0.0152	0.0156	
206		0.0135		0.0156	
208		0.0135			0.0161
210	0.1667	0.1757	0.1061	0.2188	0.2419
212	0.0833	0.0946	0.0606	0.0313	0.1129
214	0.0333	0.0135			0.0484
216	0.1	0.0541	0.0758	0.0625	0.0484
218	0.0333	0.0405	0.1364	0.0781	0.0323
220	0.2	0.1892	0.0909	0.125	0.1452
222	0.1167	0.027	0.0303	0.0625	0.0645
224	0.0167	0.0541	0.0152	0.0313	0.1129
226	0.0833	0.0405	0.1212	0.0156	0.0645
228		0.0676	0.0152	0.0469	0.0161
230	0.0167	0.0135	0.0606	0.0313	0.0323
232	0.0333	0.027	0.0606	0.0469	
234	0.0333	0.0946	0.1212	0.125	0.0161
236		0.0135		0.0469	0.0323
238	0.0333	0.027	0.0455		
242				0.0156	
244	0.0333		0.0152		
246			0.0303		
250	0.0167				
YT87					
(N)	30	37	33	32	31
Allele					
269	0.65	0.5135	0.5606	0.5156	0.5484
273		0.0676	0.0455	0.0313	0.0484
277	0.05	0.0676	0.0758	0.0469	0.129
281	0.05	0.0405	0.0152		0.0323
285	0.0833	0.0541	0.0455	0.0625	0.0323
289	0.0167	0.0811	0.0909	0.1406	0.0161

293	0.0333	0.0135		0.0313	0.0161
295					0.0161
297	0.0333	0.1081	0.0606	0.0313	0.1129
301	0.0333		0.0303	0.0625	0.0161
305	0.0333	0.027	0.0152	0.0469	
309			0.0152		
313		0.0135	0.0152		0.0161
317		0.0135		0.0313	
321	0.0167				
325			0.0303		
329					0.0161
YT111					
(N)	30	37	33	32	31
Allele					
144	0.0333				
156				0.0156	
158		0.027	0.0303		
160		0.027			
162	0.05		0.0152	0.0469	
164			0.0303	0.0156	0.0645
166	0.0167	0.0405	0.0152	0.0469	0.0161
168	0.05	0.027	0.0303	0.0156	
170	0.0667	0.1081	0.0606	0.0469	0.1129
172		0.0405	0.0758	0.0313	
174	0.0667	0.0811	0.0152	0.0625	0.0806
176	0.0833	0.1351	0.1818	0.1719	0.0645
178	0.1167	0.0811	0.0303	0.1719	0.0645
180	0.0667	0.0135	0.0606	0.0156	0.0161
182	0.0167	0.027	0.0152		
184		0.0135	0.0455		0.0161
186	0.0667	0.0135	0.0303	0.0313	0.0323
188	0.05	0.1081	0.0455	0.0313	0.0484
190	0.0667	0.0135		0.0156	0.0806
192	0.0167	0.027	0.0606	0.0938	0.0161
194	0.1167	0.0135	0.0455	0.0625	0.0323
196		0.0811	0.0303	0.0313	0.0161
198			0.0606	0.0156	0.0645
200	0.05	0.027	0.0606	0.0313	0.0806
202	0.0333	0.0405	0.0152	0.0313	0.0484
204				0.0156	
206	0.0167	0.0135			
208					0.0161
212			0.0152		0.0161
214		0.0135			
216		0.0135			

218					0.0161
222					0.0161
224					0.0161
226		0.0135			
228					0.0161
238			0.0152		
260	0.0167				
262			0.0152		
264					0.0161
268					0.0323
YT29					
(N)	30	37	33	28	31
Allele					
161				0.0179	
163				0.0714	
165		0.0135			
167		0.0135			
171	0.0333				
173				0.0179	
177	0.0667	0.0405	0.0303	0.0179	0.0484
179	0.05	0.1081	0.2273	0.3214	0.0806
185	0.1167	0.1081	0.1212	0.0357	0.2097
187					0.0161
189	0.1	0.0541	0.0909	0.0714	0.0806
191	0.0333	0.0541			0.0161
193	0.0667	0.027	0.0303	0.0179	0.0161
195	0.0333			0.0179	
197			0.0152		0.0161
199	0.0667	0.0405	0.0455	0.0179	0.0323
201	0.05	0.1081	0.0455	0.0714	0.0161
203	0.0667	0.027	0.0303	0.0536	0.0323
205	0.15	0.1622	0.1515	0.1071	0.129
207	0.0667	0.1081	0.0455	0.1071	0.0806
209	0.0333	0.0811	0.0455		0.0806
211	0.0333	0.0135	0.0758	0.0357	0.0484
213	0.0167		0.0303	0.0179	0.0323
215	0.0167	0.0135			0.0484
217		0.0135			
219			0.0152		0.0161
223		0.0135			
YT92					
(N)	30	37	33	32	31
Allele					
204	0.0167				

206				0.0156	
208	0.0167			0.0156	
210	0.0167				0.0161
212	0.0167				0.0484
214	0.0167	0.0676	0.0909	0.1094	0.0161
216	0.0167	0.0135	0.0152		0.0323
218	0.4833	0.4595	0.4545	0.2656	0.4194
220	0.2167	0.2838	0.2576	0.3125	0.2419
222	0.0833	0.0405	0.0758	0.1094	0.0806
224	0.1	0.0405	0.0909	0.125	0.0645
226	0.0167	0.027			
228		0.027		0.0156	0.0323
230		0.0135			
232		0.0135		0.0156	
234			0.0152		0.0484
238		0.0135		0.0156	
YT121					
(N)	30	37	33	32	31
Allele					
148	0.0167				
150					0.0323
154	0.0167	0.027	0.0455	0.0313	0.0484
156	0.0667	0.1351	0.0758	0.0313	0.0645
158	0.0167	0.0405	0.0455	0.0313	0.0484
160	0.0833	0.1216	0.0606	0.1406	0.0645
162	0.1333	0.1351	0.0758	0.1406	0.1129
164	0.0667	0.0541	0.0758	0.0781	0.0645
166	0.0333	0.0405		0.0313	0.0806
168	0.1333	0.0811	0.0758	0.125	0.0806
170	0.0167	0.0541	0.0455	0.0313	0.0323
172	0.0333		0.0152	0.0313	
174		0.027		0.0156	
176	0.0333		0.0303	0.0156	0.0161
178	0.0167		0.0455	0.0156	0.0484
180	0.0333	0.0541	0.0455	0.0156	0.0323
182		0.027	0.0455	0.0156	0.0323
184	0.05	0.0135	0.0758	0.0156	0.0161
186	0.0333	0.0676	0.0455	0.0625	0.0161
188	0.05	0.0135	0.0152	0.0313	0.0484
190			0.0303		0.0484
192			0.0606	0.0156	0.0161
194		0.0135		0.0156	
198	0.05				0.0323
200	0.0333	0.027		0.0156	
202	0.0333		0.0152	0.0313	0.0323

204			0.0152		0.0161
206	0.0167	0.0135	0.0152	0.0469	
208	0.0167		0.0303		
210			0.0152		
212		0.0135			
214	0.0167	0.0135			0.0161
216		0.0135		0.0156	
222		0.0135			
D235					
(N)	30	37	32	31	31
Allele					
174				0.0161	
178			0.0469	0.0161	0.0323
182	0.0167		0.0156	0.0161	0.0323
186	0.0167			0.0161	
190	0.0333	0.0676	0.0781	0.0645	0.0806
194	0.0667	0.0405	0.0313	0.0645	0.0645
198	0.05	0.0541	0.1563	0.0645	0.0161
202	0.05	0.0541	0.0313	0.0645	0.0806
206	0.0667	0.0676	0.1094	0.0645	0.0968
210	0.0667	0.0946	0.0469	0.0484	0.0645
214	0.1333	0.0946	0.0781	0.1129	0.0645
218	0.0667	0.0811	0.1094	0.0484	0.129
222	0.1	0.1622	0.1094	0.129	0.1613
226	0.1	0.0811	0.0781	0.0484	0.0968
230	0.1	0.0946	0.0625	0.0806	0.0323
234	0.0333	0.0676	0.0156	0.0645	0.0161
238		0.0135	0.0156	0.0161	
242	0.05	0.0135	0.0156	0.0645	
246	0.0167				
250					0.0161
254	0.0167	0.0135			0.0161
262	0.0167				
3D4					
(N)	30	36	33	27	31
Allele					
212			0.0152		
228	0.0167	0.0556	0.0606	0.0556	0.0484
232	0.45	0.4167	0.4091	0.3704	0.371
236	0.0667	0.0139	0.0909	0.0741	0.0645
240	0.0333	0.0278	0.1061	0.1111	0.0806
244	0.0167				0.0323
248	0.1667	0.125	0.0758	0.0926	0.0645
252	0.15	0.2361	0.1818	0.1667	0.1774

256	0.0333	0.0556	0.0303	0.0185	0.0484
260		0.0417	0.0152	0.037	0.0323
264	0.0167				
268					0.0323
272	0.0167	0.0139	0.0152	0.0185	0.0323
276					0.0161
280				0.0185	
284				0.0185	
288	0.0333	0.0139			
292				0.0185	
D4D6					
(N)	29	37	33	27	30
Allele					
310	0.0345		0.0152	0.0556	0.0167
314	0.0172	0.027	0.0455	0.0185	
318	0.1034	0.0811	0.0909	0.0556	
322	0.069	0.0541	0.0455	0.037	0.0833
326		0.027	0.0152		0.0167
334	0.0172	0.0811	0.0758	0.037	0.0833
338	0.069	0.0405	0.0152	0.037	0.0167
342	0.0345	0.0541	0.0303	0.0741	0.05
346	0.0345	0.0541	0.0606	0.037	0.05
350	0.0517	0.027	0.0152	0.0556	0.0333
354	0.0345		0.0152	0.0185	0.05
358		0.0135	0.0455		0.05
362	0.069		0.0303	0.0556	
366		0.0405	0.0303	0.0185	0.0167
370	0.0345	0.0676	0.0758	0.0556	0.0333
374	0.069		0.0303	0.037	0.05
378	0.0172	0.027	0.0152		0.0167
382	0.1034	0.0541	0.0303	0.0556	
386	0.0345	0.0541	0.0606	0.037	0.0833
390	0.0172	0.1351	0.0909	0.0741	0.0333
394	0.069	0.0405	0.0455	0.1111	0.1167
398	0.0517	0.027	0.0606	0.0741	0.05
402	0.0172	0.027	0.0455	0.037	0.1
406	0.0172		0.0152	0.0185	0.0333
410		0.0135			
418		0.0135			
422					0.0167
434		0.0135			
438	0.0172				
450		0.0135			
490		0.0135			
498	0.0172				

D10					
(N)	27	37	33	31	30
Allele					
244			0.0152		
260					0.0167
276	0.0185				
280	0.0185	0.0135			
292	0.0185				
296	0.0185	0.0135	0.0152	0.0161	
300		0.0135	0.0606	0.0484	
304	0.0556	0.027	0.0303		0.05
308	0.0185	0.027	0.0303	0.0484	0.0333
312		0.0405	0.0303	0.0645	0.0167
316	0.037	0.0405	0.0455	0.0161	0.0333
320	0.037	0.027	0.0606	0.0806	0.05
324	0.0556	0.0676	0.0303	0.0645	0.0167
328			0.0606	0.0161	0.05
332	0.0185	0.0811	0.1061	0.0806	0.05
336	0.0556	0.0676	0.0606	0.0161	0.1
340	0.0185	0.0946	0.0303	0.0161	0.0167
344	0.037	0.0405	0.0758	0.0323	0.0167
348	0.037	0.027	0.0152	0.0161	0.0167
352	0.0556	0.1081	0.0606	0.0645	0.0167
356	0.037	0.0405	0.0606	0.0484	0.0833
360	0.1111	0.0135	0.0303	0.0161	0.05
364	0.0741	0.027	0.0455	0.0645	0.0833
368	0.0741	0.027	0.0152	0.0645	0.0667
372		0.0135	0.0152	0.0484	0.0667
376	0.0185	0.0135	0.0606	0.0323	0.0333
380		0.0135	0.0152	0.0484	0.0167
384	0.0556	0.027		0.0323	0.0167
388		0.027			0.0167
392	0.0185		0.0152		0.0167
396	0.037	0.0541			
400	0.037			0.0323	
404	0.0185	0.0135			0.0333
408		0.027			0.0167
412			0.0152		
416				0.0161	
420					0.0167
424	0.0185			0.0161	
448		0.0135			
D12					
(N)	29	37	32	28	30

Allele					
248	0.0172				0
292					0.0167
316		0.0135			
320	0.0172	0.027			0.0333
324		0.0541	0.0469	0.0714	0.05
328				0.0179	
332	0.0345	0.027	0.0938	0.0179	0.0667
336	0.0862	0.0676	0.0938	0.0893	0.0167
340	0.069	0.0676	0.0625	0.0714	0.0167
344	0.0517	0.0946	0.0781	0.1429	0.0167
348	0.1034	0.1486	0.0625	0.1071	0.1
352	0.1207	0.0676	0.1094	0.0536	0.1167
356	0.1379	0.1351	0.1406	0.0714	0.0833
360	0.1034	0.1486	0.0625	0.0714	0.15
364	0.0862	0.027	0.0313	0.0714	0.0333
368	0.0172	0.027	0.0313	0.0179	0.0833
372	0.0345		0.0781	0.0357	0.0667
376	0.0172	0.0541	0.0156	0.0179	0.05
380	0.0172	0.0135	0.0469	0.0536	0.0333
384	0.0517		0.0156	0.0357	0.0167
388	0.0172			0.0179	0.0167
392	0.0172			0.0357	
396					0.0167
400		0.0135	0.0313		
408		0.0135			0.0167
D211					
(N)	30	37	33	31	31
Allele					
171			0.0152		
183	0.0167		0.0303	0.0161	0.0484
187	0.0667	0.0135	0.0152	0.0323	0.0484
191	0.1	0.0946	0.0606	0.0806	0.0806
195	0.1167	0.2027	0.1515	0.129	0.0968
199	0.1333	0.1351	0.1515	0.1935	0.1129
203	0.1167	0.1351	0.2273	0.0806	0.129
207	0.2333	0.0676	0.1364	0.1935	0.0968
211	0.1167	0.1081	0.1212	0.1613	0.1935
215	0.0667	0.0946	0.0455	0.0645	0.0806
219	0.0333	0.0811	0.0303	0.0161	0.0645
223		0.027	0.0152	0.0323	0.0161
227		0.027			0.0323
231		0.0135			

Appendix 2.2

Genetic diversity estimates: Single-locus and mean values are given. N= number of samples analysed; Na= number of alleles; Ar= allelic richness; He (n.b) = expected non biased heterozygosity; Ho= observed heterozygosity; Fis= inbreeding coefficient. Bold and underlined values indicated a significant HW disequilibrium (P<0.05); Significance after sequential Bonferroni's correction is described as follows: *P < 0.05; **P < 0.01.

YT111					YT87					YT4				
Fis	0.05	-0.01	0.01	0.05	Fis	-0.05	0.05	0.01	-0.15	Fis	0.15	-0.01	-0.05	0.07
Pvalue	0.34	0.73	0.84	0.06	Pvalue	0.81	0.73	0.44	1	Pvalue	0.08	0.05	0.71	**0.000** **
AR	18.48	21.09	22.34	19.76	AR	9.77	10.1	11.2	9.9	AR	14.65	16.38	15.2	16.28
Na	19	24	24	21	Na	10	11	12	10	Na	15	18	16	17
H obs.	0.9	0.95	0.94	0.88	H obs.	0.6	0.68	0.67	0.81	H obs.	0.77	0.92	0.97	0.84
H n.b.	0.95	0.94	0.94	0.92	H n.b.	0.57	0.71	0.67	0.71	H n.b.	0.9	0.91	0.92	0.91
N	30	37	33	32	N	30	37	33	32	N	30	37	33	32
Sample	71_2	77_2	34_2	41_1	Sample	51_2								

YT121					YT92					YT29				
Fis	0.12	0.05	0.01	-0.01	Fis	0.05	-0.15	0.01	0.05	Fis	0.02	0.07	-0.05	0.01
Pvalue	0.21	0.06	0.84	0.73	Pvalue	0.06	1	0.44	0.73	Pvalue	0.64	**0.000** **	0.71	0.05
AR	23.4	19.76	22.34	21.09	AR	11.2	9.9	11.2	10.1	AR	14.44	16.28	15.2	16.38
Na	25	21	24	24	Na	12	10	12	11	Na	15	17	16	18
H obs.	0.84	0.88	0.94	0.95	H obs.	0.65	0.81	0.67	0.68	H obs.	0.87	0.84	0.97	0.92
H n.b.	0.96	0.92	0.94	0.94	H n.b.	0.67	0.71	0.67	0.71	H n.b.	0.89	0.91	0.92	0.91
N	31	32	33	37	N	31	32	33	37	N	31	32	33	37
Sample					Sample									

-0.02	-0.04	0.05	0.07	-0.01		Fis	0.06	0.04	0.09	-0.04	0.02
0.55	0.45	0.03	0.29	0.95		Pvalue	0.08	0.92	0.13	0.58	0.6
22.14	19.51	21.78	22.26	22.13		AR	22.43	21.88	22.55	21	20.36
23	22	23	24	23	D4D6	Na	23	24	24	21	21
0.97	0.97	0.91	0.88	0.97		H obs.	0.9	0.92	0.88	1	0.93
0.95	0.93	0.96	0.94	0.96		H n.b.	0.96	0.95	0.96	0.96	0.95
30	37	33	32	31		N	29	37	33	27	30
-0.08	0.09	0.03	-0.13	0.11		Fis	0.2	0.16	-0.01	-0.04	0.1
0.68	0.24	0.86	0.29	0.19		Pvalue	0.02	0.04	0.18	0.41	0.29
10.3	9.74	6.64	9.22	9.71		AR	10.58	9.17	9.42	12	11.81
11	11	7	10	10	3D4	Na	11	10	10	12	12
0.77	0.65	0.7	0.91	0.68		H obs.	0.6	0.64	0.79	0.85	0.74
0.71	0.71	0.72	0.8	0.76		H n.b.	0.75	0.76	0.78	0.82	0.82
30	37	33	32	31		N	30	36	33	27	31
0.15	-0.03	-0.01	0.14	-0.02		Fis	0.04	0.01	-0.01	0.01	-0.01
0.02	0.96	0.34	0.01	0.04		Pvalue	0.3	0.5	0.88	0.3	0.6
16.76	16.19	14.49	15.75	17.18		AR	17.48	14.16	15.32	17.35	15.44
17	18	15	16	18	D235	Na	18	15	16	18	16
0.8	0.95	0.91	0.75	0.94		H obs.	0.9	0.92	0.94	0.94	0.94
0.94	0.92	0.9	0.87	0.92		H n.b.	0.94	0.92	0.92	0.94	0.92
30	37	33	28	31		N	30	37	32	31	31
71_2	**77_2**	**34_2**	**41_1**	**51_2**		**Sample**	**71_2**	**77_2**	**34_2**	**41_1**	**51_2**

Mean	FIS	0.05	0.04	0.01	0	0.04
	Pvalue	0.33	0.46	0.57	0.43	0.4
	H obs.	0.82	0.84	0.87	0.88	0.86
	H n.b.	0.87	0.88	0.88	0.89	0.89
D211	Fis	-0.02	0.06	-0.01	0.04	0
	Pvalue	0.73	0.6	1	0.01	0.83
	AR	9.89	11.32	11.39	10.71	11.85
	Na	10	12	12	11	12
	H obs.	0.9	0.84	0.88	0.84	0.9
	H n.b.	0.88	0.89	0.87	0.88	0.91
	N	30	37	33	31	31
D12	Fis	0.08	0.18	0.1	0.01	0.08
	Pvalue	0.17	**0.000****	0.43	0.86	0.15
	AR	17.51	15.63	15.61	17.82	19.17
	Na	18	17	16	18	20
	H obs.	0.86	0.76	0.84	0.93	0.87
	H n.b.	0.93	0.92	0.93	0.94	0.94
	N	29	37	32	28	30
D10	Fis	0.08	0.01	-0.01	0	0
	Pvalue	0.21	0.32	0.97	0.93	0.24
	AR	26	24.86	23.35	23.77	25.66
	Na	26	28	25	25	27
	H obs.	0.89	0.95	0.97	0.97	0.97
	H n.b.	0.97	0.96	0.96	0.96	0.96
	N	27	37	33	31	30
	Sample	**71_2**	**77_2**	**34_2**	**41_1**	**51_2**

4 Chapter 3

Switching from genetics to genomics: 2b-RAD genotyping technique

Publication Note:

This Chapter has been published in the journal "Marine Genomics". doi: 10.1016/j.margen.2015.12.002

Title: Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (*Thunnus albacares*)

Authors: Carlo Pecoraro[a,b*], Massimiliano Babbucci[c], Adriana Villamor[a], Rafaella Franch[c], Chiara Papetti[k], Bruno Leroy[d], Sofia Ortega-Garcia[e], Jeff Muir[f], Jay Rooker[g], Freddy Arocha[h], Hilario Murua[i], Iker Zudaire[j,b], Emmanuel Chassot[b], Nathalie Bodin[b], Fausto Tinti[a], Luca Bargelloni[c], Alessia Cariani[a]

Affiliations:

[a]: University of Bologna, Dept. Biological, Geological and Environmental Sciences (BIGEA), Via Selmi 3, 40126 Bologna, Italy;

[b]: Institut de Recherche pour le Développement (IRD), UMR MARBEC (IRD/Ifremer/UM2/CNRS) SFA, Fishing Port, BP570 Victoria, Seychelles;

[c]: University of Padova, Comparative Biomedicine and Food Science, viale dell'Università 16, 35020, Legnaro (PD), Italy;

[d]: Secretariat of the Pacific Community, Oceanic Fishery Programme, BP D5, 98848 Noumea, New Caledonia;

[e]: Instituto Politécnico Nacional-CICIMAR, Departamento de Pesquerias, Avenida IPN s/n, La Paz, BCS, Mexico;

[f]: Pelagic Fisheries Research Program, University of Hawaii, Marine Science Building 312, Honolulu, Hawaii 96822, USA;

[g]: Department of Marine Biology, Texas A&M University, 1001 Texas Clipper Road, Galveston, TX 77553, USA;

[h]: Instituto Oceanográfico de Venezuela, Universidad de Oriente, Avda. Universidad Cerro Colorado, Cumana 6101, Venezuela;

[i]: AZTI, Marine Research Division, Herrera Kaia-Portualdea z/g, Pasaia, 20110 Gipuzkoa, Spain;

[j]: Ikerbasque Fundazioa, Maria Diaz de Haro, 3-6°, Bilbao, 48013 Bizkaia, Spain;

[k]: Section of Integrative Ecophysiology, Alfred-Wegener-Institute for Polar and Marine Research, Am Handelshafen 12, Bremerhaven 27570, Germany.

4.1 Introduction

Yellowfin tuna (*Thunnus albacares*, YFT) has relevant biological and economic importance at the global scale, being an apex predator in oceanic ecosystem and representing the second largest tuna fishery worldwide (FIGIS, 2010-2015). Currently, YFT is managed in four distinct stocks under the jurisdiction of four independent Regional Fisheries Management Organizations (RFMOs). Although a proper fish stock management needs accurate knowledge on the stock structure and its genetic variation with respect to environmental and ecological conditions (Papetti et al., 2013), YFT genetic population structure has not been resolved yet. Different studies provided discordant patterns of YFT global-scale genetic differentiation (Ward et al., 1997; Ely et al., 2005; Appleyard et al., 2001), together with a genetic structuring detected at the regional level (Dammanagoda et al., 2008; Kunal et al., 2013; Li et al., 2015). This discordance was likely due to the YFT life history traits (e.g. high fecundity, large population sizes), which make detecting patterns of genetic differentiation among population samples very difficult (Ely et al., 2005; Juan-Jordá et al., 2013). Moreover, population genetic studies reporting significant differences relied upon a relatively small number of molecular markers, hence, covering only a very limited portion of the genome (Appleyard et al., 2001; Díaz-Jaimes and Uribe-Alcocer, 2006). Failing to detect population structure, due to limited genetic resolution of classical markers, can potentially be misleading for management purposes, driving to local overfishing and severe stock decline (Ying et al., 2011).

According to the uncertainty about both population structure and size of YFT stocks, there is an evident need for developing alternative approaches based on genomics, that allow screening a larger number of markers across the entire genome, including neutral and non-neutral loci. This might enable detecting YFT population structure, quantifying the extent of spatial demographic changes and discover imprints of local adaptation, which represent priority focus for implementing any effective management plan.

The rapid advent of next-generation sequencing (NGS)-based genotyping methods has significantly improved our ability to analyse thousands of Single Nucleotide Polymorphism (SNP) markers across the entire genome, increasing the precision in detecting small genetic differentiation among geographical populations (Waples, 2008; Allendorf et al., 2010; Davey et al., 2011; Narum et al., 2013; Andrews and Luikart, 2014). Although SNPs are characterized by a low diversity due to the only four possible allelic states, this limitation is largely outweighed by their abundance, being as frequent as one SNP every few hundred base pairs (Morin et al., 2004; 2009). Moreover, SNPs

are becoming the marker of choice for many applications in population ecology, evolution and conservation genetics, having a high potential for genotyping efficiency, data quality and low-scoring error rates, genome-wide coverage and analytical simplicity (Milano et al., 2014).

Here, for the first time, we applied the 2b-RAD Genotyping-By-Sequencing (GBS) technique (Wang et al., 2012) for testing its potential for investigating population genetic structure in a non-model, large pelagic and highly migratory fish species. This novel genomic tool is based on sequencing reduced representation libraries produced by type IIB restriction endonucleases, which cleave genomic DNA upstream and downstream of their target site, generating tags of uniform length that are ideally suited for sequencing on existing NGS platforms (Wang et al., 2012). This method permits parallel and multiplexed sample sequencing of tag libraries for the rapid discovery of thousands of SNPs across the entire individuals' genome, with a very cost-effective procedure resulting in high genome coverage. The 2b-RAD method allows to screen in parallel almost every restriction site in the genome, whereas other GBS methods can only target a subset of total restriction sites to counterbalance loss of PCR amplification and sequencing efficiency due to large size of restriction fragments. This technique also allows fine-tuning the marker density by means of selective adapters in order to sequence fewer loci with higher coverage, for applications such as population genetics (Puritz et al., 2014; Andrews and Luikart, 2014). Given these attributes, the 2b-RAD method has the potential to discriminate the existence of genetic differentiation with a high statistical power, generating genome-wide data for genetic structure analysis at different spatial scales for YFT populations.

In this study, we: i) first examine the utility of Technical Replicates (TRs) for optimizing genotyping procedure, comparing the results obtained running the *denovo_map.pl* and the *ref_map.pl* programs in *Stacks* (Catchen et al., 2011; 2013); and ii) finally assess the applicability of 2b-RAD for future investigations in this highly migratory species.

4.2 Materials and methods

4.2.1 Sampling design, libraries preparation and sequencing

A total of 100 juvenile YFT (35-55 cm of fork length, F_L) from Atlantic, Indian and Pacific geographic population samples (Table 4.1) were analysed, covering the entire species distribution (Fig. 4.1).

Table 4.1_The table summarizes the sampling location, sample code and number of individual per each geographic population sample.

Sampling location	Sample code	Number of individuals
W Atlantic Ocean	**31_1**	10
E Atlantic Ocean	**34_1**	10
E Atlantic Ocean	**34_2**	10
W Atlantic Ocean	**41_1**	10
W Indian Ocean	**51_1**	10
W Indian Ocean	**51_2**	10
WC Pacific Ocean	**71_1**	10
WC Pacific Ocean	**71_2**	10
E Pacific Ocean	**77_1**	10
E Pacific Ocean	**77_2**	10

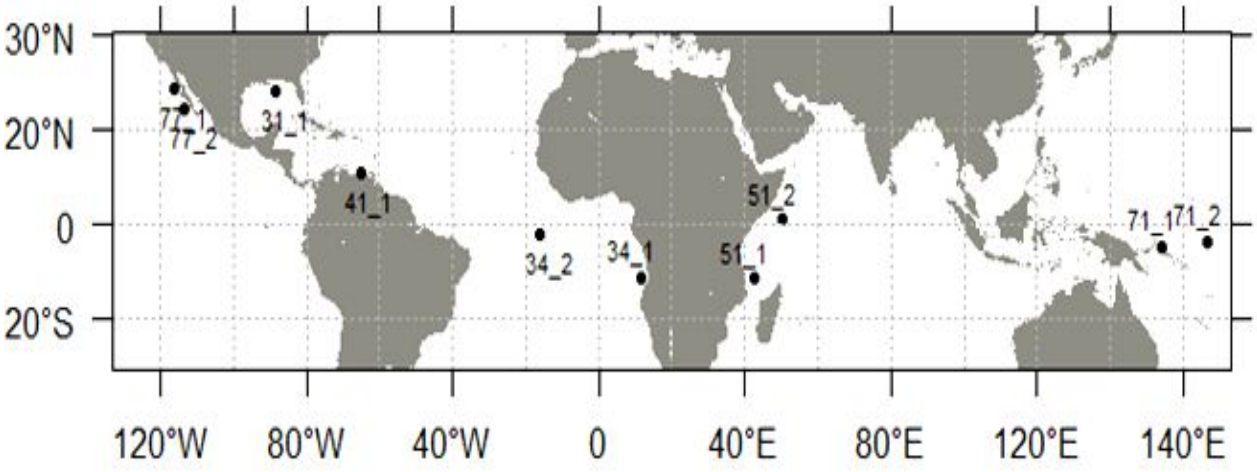


Fig. 4.1. Location of *Thunnus albacares* geographical population samples analyzed in this study. Sample codes are given as in Table 4.1.

Genomic DNA (gDNA) was extracted from approximately 20 mg of tissue (skeletal muscle or finclip) using the commercial kit Invisorb® Spin Tissue Mini Kit (Invitek, STRATEC Biomedical,

Germany) as described in Section 1. Since high-quality gDNA is required in the 2b-RAD genotyping technique, its concentration and purity, in terms of ratios of absorbance at 260/230 nm and at 260/280 nm, were quantified by both a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and a Qubit 2.0 Fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, Massachusetts, USA). This procedure ensured to work with high quality samples and comparable DNA concentration.

The 2b-RAD libraries were constructed for each individual following the protocol from Wang et al., (2012) with minor modifications (see below). To assess the robustness of the method and subsequent data analyses, three libraries were replicated (Technical Replicates, TRs) for two individuals (34_2_Y_2 and 77_2_Y_15) and two for an additional third specimen (51_1_Y_7). gDNA (300 ng) was digested with 2 U of the enzyme *CspC*I (New England Biolabs, NEB, Ipswich, Massachusetts, USA) for 1 h at 37°C. The digested DNA was ligated in a 25µL total volume reaction consisting of 0.4 µM for each of the two library-specific adaptors, 0.2 mM ATP (New England Biolabs, NEB, Ipswich, Massachusetts, USA) and 1 U T4 DNA ligase (SibEnzyme Ltd., Academ town, Siberia). To reduce marker density, one adaptor with fully degenerate 3' overhangs NN and one with reduced 3'degeneracy NG were chosen. Sample-specific barcodes were designed with Barcode Generator (http://comailab.genomecenter.ucdavis.edu/index.php/Barcode_generator) and introduced by PCR with platform-specific barcode-bearing primers. 2b-RAD tags were amplified by PCR in two separate 25 µL-reactions, in order to minimize PCR amplification bias (Mastretta-Yanes et al., 2014). Each amplification consisted of 6.25 µL of ligated DNA, 0.5 µM each primer (P4 and P6-BC, Eurofins Genomics S.r.l, Italy), 0.2 µM each primer (P5 and P7, Eurofins Genomics), 0.3 mMdNTP (New England Biolabs, NEB, Ipswich, Massachusetts, USA), 1X Phusion HF buffer and 1 U TaqPhusion high-fidelity DNA polymerase (NEB). Cycling conditions were: 98°C for 4min; 98 °C for 5 s, 60° C for 20 s, 72° C for 5 s for 14 cycles, 72°C for 5 min. The reduced number of amplification cycles (n=14) is crucial to produce a negligible amount of PCR amplification errors, comparing to those needed to reach the plateau phase.

PCR products were purified with the SPRIselect purification kit (Beckman Coulter, Pasadena, California, USA), to exclude any high-molecular weight DNA remaining after the enzyme digestion and any incorrect constructs that may emerge during PCR amplification. The concentration of purified individual libraries was quantified using Qubit®ds DNA BR Assay Kit (Invitrogen–ThermoFisher Scientific, MA, USA) and Mx3000P qPCR instrument, and the quality checked on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA). Individual libraries

were pooled into equimolar amounts and resulting pools' quality was re-verified on Agilent 2100 Bioanalyzer. Pooled libraries were sequenced on an Illumina HiSeq2500 platform with a 50 bp single-read module at the Genomix4Life S.r.l. facilities (Baronissi, Salerno, Italy), which also performed data demultiplexing.

4.2.2 Technical Replicates analysis and optimization of genotyping procedure

Demultiplexed reads were returned by the sequencing facility in Fastq format and their quality was checked by FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). After this, a custom-made Perl script was run for quality filtering and adaptors trimming of the reads, obtaining sequences of 34bp (Fasta files available at SRA Bioproject: PRJNA294940). Filtered reads were analyzed with the software *Stacks* v. 1.32 (Catchen et al., 2011; 2013), which allows genotype inference through the identification of SNP loci without a reference genome (*denovo_map.pl* program) or aligning reads against a reference genome (*ref_map.pl* program). Different settings were tested on the TRs dataset to fine-tune the *de novo Stacks* pipeline' parameters and to assess the consistency of results, in terms of total number of identified SNPs; the error rate calculated counting discordant genotypes between TRs and, among the concordant data, the percentage of heterozygous SNPs. Following the *Stacks* author guidelines, multiple combinations were considered for: a) the minimum number of identical reads necessary to call an allele (*-m* value set to: 5, 8, 10, 15); b) mismatches between reads within a locus (*-M* value set to: 2, 3, 4, 6); and c) mismatches among loci when comparing across individuals (*-n* value set to: 0, 2, 3, 4, 6). Only one parameter was varied at a time while keeping the others fixed. The default values for *min_het_seqs* and *max_het_seqs* were used. In addition, we compared the default and the bounded SNP calling models (*--bound_high* value set to: 0.05, 0.1, 0.15, 0.2, 0.5) to evaluate the percentage of heterozygous genotypes correctly assessed, in order to make the genotype calling between them as much concordant as possible. In the bounded SNP calling model, *Stacks* employs a multinomial-based likelihood model to identify SNPs and to estimate the maximum-likelihood value of the sequencing error rate ε at each nucleotide position, in order to proper call each possible genotype (for details see Catchen et al., 2011; 2013).

The reads were also mapped against the genome of *Thunnus orientalis* (GenBank accession numbers BADN01000001-BADN01133062; Nakamura et al., 2013) using CLC Genomics Workbench v. 5.1 (CLC Bio) program. The following parameters settings were applied *length fraction*= 1.0 and *similarity fraction*= 0.9 (all remaining parameters as default), retaining only uniquely mapped

reads. Mapping results were exported in SAM format and were used as input files for *refmap_map.pl* in *Stacks*. To further evaluate the robustness of the approach a similar testing was performed on mapped data, using the same settings as for the *denovo_map.pl* for *-m*, *-n* and *--bound_high* of the bounded SNP calling model.

4.2.3 Preliminary analysis of YFT population structure

Once identified the *Stacks*' parameter set which minimized differences among TRs, the *denovo_map.pl* program was run on the entire YFT dataset (Table 1). Using the program *populations* available from *Stacks* software, different combinations of *-p* (4, 6, 9) and *-r* (0.4, 0.7, 0.9) parameters were tested, in order to investigate changes in the number of SNPs obtained, and in the percentage of missing values among samples. Following these tests, we selected from the resulting catalogue of loci only those containing one bi-allelic SNP (*-F snps_l*=1 *snps_u*=2), and those values of *-p* and *-r* rendering the highest number of SNPs with the lowest percentage of missing data.

In order to increase the sample size and to improve robustness of the genetic analyses, several grouping of the geographic samples were tested, especially due to their ocean basin distance, performing an analysis of molecular variance (AMOVA) with the software Arlequin 3.5.1.2 (Excoffier and Lischer, 2010) with 10,000 permutations and $p \leq 0.01$ significance level.

Based on the SNPs dataset and AMOVA results obtained, F_{ST} estimates for pairwise comparison among pooled samples, were calculated with the software Arlequin 3.5.1.2 using the same settings as above.

A preliminary assessment of YFT genetic structure was performed using the Discriminant Analysis of Principal Components (DAPC, Jombart et al., 2010) implemented in the R package Adegenet (Jombart, 2008, R version 3.1.2, R Development Core Team, 2014; http://www.r-project.org). The function *find.clusters* was used to identify the optimal number of clusters (k) that maximizes the variation between groups (Jombart et al., 2010). The cross-validation test was also carried out in order to validate the number of Principal Components (PCs) retained in the first transformation step of DAPC analysis, because a wrong choice of the number of PCs might negatively impact the DAPC results and produce unstable output due to over-parameterization.

4.3 Results and discussion

A similar number of reads was obtained among TRs, before and after quality filtering (Table 4.2), which underlines the reliability of this technique in genotyping individuals.

Table 4.2. Details on the technical replicates: acronym (Sample ID), Oceanic origin, genomic DNA concentration in ng/µL, library concentration in nm/µL, number of raw reads obtained, retained reads after quality filtering, and their corresponding percentage.

Sample ID	Oceanic origin	gDNA ng/µL	Library nm/µL	N° raw reads	N° filtered reads	% retained reads
34_2_Y_2R1	Atlantic Ocean	333.80	185.38	2,276,239	1,772,927	78 %
34_2_Y_2R2	Atlantic Ocean	333.80	207.83	2,672,917	1,914,181	72 %
34_2_Y_2R3	Atlantic Ocean	333.80	197.03	2,309,039	1,805,181	78 %
77_2_Y_15R1	Pacific Ocean	218.02	238.81	2,212,559	1,788,988	81 %
77_2_Y_15R2	Pacific Ocean	218.02	240.53	2,292,834	1,850,871	81 %
77_2_Y_15R3	Pacific Ocean	218.02	208.91	2,085,658	1,802,820	86 %
51_1_Y_7R1	Indian Ocean	177.54	166.05	2,330,292	1,767,345	76 %
51_1_Y_7R2	Indian Ocean	177.54	170.81	2,300,342	1,824,635	79%

Among the different *Stacks* settings considered, the *-m* value was the parameter that most affected the genotyping results, in particular the number of detected SNPs. Sensitivity tests performed on the TRs showed a decrease in the number of SNPs, from 5,753 to 4,490, when increasing *-m* from 5 to 15 (Fig. 4.2 and Supplementary Material 1 for values with associated Standard Error). The percentage of error rate varied approximately from 1% to 5%, with a decreasing trend when increasing the *–m* values (Fig. 4.2 and Supplementary Material 1). The percentage of heterozygous SNPs remained constant with increasing *-m* values (Fig. 4.2 and Supplementary Material 1).

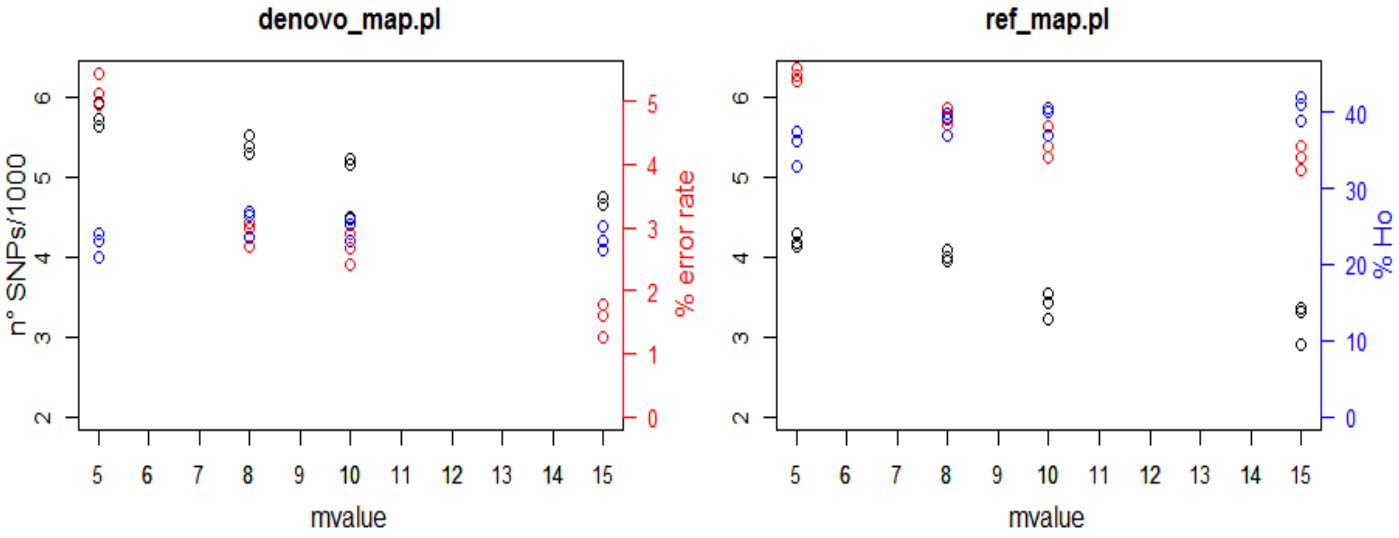


Fig. 4.2. Comparison between d*enovo_map.pl* (Left panel) and ref_map.pl (Right panel) performance in terms of (a) number of SNPs (black dots) (b) error rate (red dots) and (c) percentage of heterozygous loci (blue dots), using different -m values. Each dot represents the average value among the three individual's TRs. The three y axes (n° SNPs/1000, % error rate, % Ho) are shared between the two plots.

An increase in true heterozygous SNPs calls was observed using the *bounded SNP calling model* compared to the *default SNP model* and reducing the upper bound values, in agreement with the results obtained by Mastretta-Yanes et al., (2014). In fact, reducing the upper bound on the maximum-likelihood of ε decreases the possibility of calling a homozygote instead of a true heterozygous genotype (Catchen et al., 2013). The proper genotype calling was further checked for a sub-sample of the total reads obtained, in the *Stacks* web interface, verifying the sequences alignment and monitoring the genotyping inference when the results were exported. This procedure was repeated each time when changing the different model's *upper bound* values.

By relaxing the number of mismatches within each locus (*-n*) and among loci (*-M*), an increase in the number of SNPs and error rate was observed (Table 4.3).

Table 4.3.- The effects of different *-M* values (2, 3, 4, 6; Left Table); and *-n* values (0, 2, 3, 4, 6; Right Table) on the: 1) number of SNPs obtained (SNPs/1000); (2) percentage of error rate; and (3) percentage of heterozygous loci).

ID Sample	- M value	SNPs/1000	%error rate	% Ho
34_2_Y_2	2	5.02 (± 0.34)	3.02 (± 0.43)	21.32 (± 0.25)
	3	5.27 (± 0.26)	3.08 (± 0.26)	23.70 (± 0.18)
	4	5.35 (± 0.33)	3.43 (± 0.47)	23.72 (± 0.36)
	6	5.37 (± 0.16)	3.57 (± 0.21)	22.98 (± 0.13)
77_2_Y_15	2	5.24(± 0.21)	2.14(± 0.12)	24.45(± 0.43)
	3	5.37(± 0.14)	2.71(± 0.32)	27.00(± 0.34)
	4	5.43(± 0.37)	2.91(± 0.42)	27.02(± 0.21)
	6	5.50(± 0.26)	3.21(± 0.31)	27.12(± 0.38)
51_1_Y_7	2	5.23(± 0.43)	2.97(± 0.39)	26.31(± 0.45)
	3	5.52(± 0.12)	2.99(± 0.23)	26.42(± 0.07)
	4	5.56(± 0.25)	3.15(± 0.42)	27.02(± 0.13)
	6	5.86(± 0.48)	3.65(± 0.37)	27.31(± 0.57)

ID Sample	- n value	SNPs/1000	%error rate	% Ho
34_2_Y_2	0	5.19 (± 0.13)	3.04 (± 0.27)	22.63 (± 0.31)
	2	5.27 (± 0.26)	3.08 (± 0.26)	23.70 (± 0.18)
	3	5.31(± 0.46)	3.25(± 0.18)	23.82 (± 0.34)
	4	5.33 (± 0.28)	3.29(± 0.32)	23.84 (± 0.21)
	6	5.49 (± 0.13)	4.01 (± 0.34)	23.96 (± 0.27)
77_2_Y_15	0	5.24(± 0.21)	2.58(± 0.27)	26.98(± 0.15)
	2	5.37(± 0.14)	2.71(± 0.32)	27.00(± 0.34)
	3	5.42(± 0.17)	2.92(± 0.28)	27.04(± 0.16)
	4	5.51(± 0.37)	3.18(± 0.19)	27.15(± 0.07)
	6	5.59(± 0.45)	3.31(± 0.09)	27.22(± 0.22)
51_1_Y_7	0	5.41(± 0.04)	2.93(± 0.18)	25.97(± 0.39)
	2	5.52(± 0.12)	2.99(± 0.23)	26.42(± 0.07)
	3	5.54(± 0.45)	3.11(± 0.03)	26.45(± 0.13)
	4	5.63(± 0.45)	3.23(± 0.15)	26.48(± 0.26)
	6	5.67(± 0.31)	3.24(± 0.27)	26.68(± 0.19)

Mapping 2b-RAD reads against the genome of *T. orientalis* allowed a high percentage of successfully mapped sequences (86.59%). The outputs obtained on the mapped data from TRs with the *ref_map.pl* program, confirmed the trends observed with the *denovo_map.pl* program (Fig. 4.2). However, the absolute number of SNPs was lower than that obtained with the *denovo_map.pl* program, likely due to the incompleteness of the reference genome used (the only *Thunnus sp.* genome available to date, Nakamura et al., 2013) and the phylogenetic distance between YFT and *Thunnus orientalis.*

Aligning reads to the reference genome, before calling a locus, can filter out erroneous stacks generated by contaminants (e.g. bacteria) possibly present in very small amount in the starting

gDNA sample. Moreover, the error rate also showed a less evident decreasing pattern when increasing *-m*, confirming however a low error rate in the genotyping call (<5%). On the contrary, the percentage of heterozygous SNPs identified using *T. orientalis* genome as reference, showed a slight increase from 35.6% to 40.7%, when higher values of *m* where used (Fig. 4.2).

Based on the number of SNPs identified with the two different approaches (with or without using the reference genome), the low error rates and the consistent percentage of heterozygous SNPs obtained among TRs, the *denovo_map.pl* program was run applying the following parameter settings *-m*= 8, *-M* = 3, *-n* = 2 and a bounded *SNP calling model* with an upper bound of 0.1 (all remaining *Stacks* settings as default), to obtain the final dataset (see Supplementary Material 2 for details about each individual). The AMOVA results (Supplementary Material 3) revealed that pooling samples into the three major oceanic regions produced the highest percentage of variation explained by groups subdivision (2.73% P-value <0.01) and at the same time very low and not significant differences were observed among populations within groups (0.97% P-value >0.01). Pooling YFT individuals in these three groups corresponding to the three oceans (Table 4.3), allowed to increase the sample size and to obtain more robust and reliable inferences of population structure.

Table 4.3. Summary statistics of the three *Thunnus albacares* oceanic groups. The table reports: the sampling origin (location) the sample size (N° individuals), the mean number (millions) of raw reads with the associated standard error (SE), the corresponding mean number (millions) of filtered reads (with SE), the percentage of reads retained, the mean value of unique tags, polymorphic SNPs and number of SNPs found.

Location	N° individuals	Raw reads (mln)	Filtered reads (mln)	% of reads retained	Unique tags	Polymorphic SNPs	SNPs found
Atlantic Ocean	40	3.80 (± 0.29)	3.09 (± 0.37)	80%	30,776	3,264	5,693
Indian Ocean	20	3.98 (± 0.46)	2.91 (± 0.15)	79%	31,430	3,695	6,516
Pacific Ocean	40	3.30 (± 0.26)	2.78 (± 0.24)	84%	31,573	3,363	5,906

The optimization process combining different *-r* and *-p* parameters' values of the *Stacks* population module, led to *-r = 0.7* and *p =* 6 as best middle ways between the number of SNPs and the percentage of missing values obtained (Table 4.4).

Table 4.4. Number of SNPs and percentage of missing value (NA %) obtained for the entire dataset according to the *-r* and *-p* parameters' values of the *Stacks* population program.

		-p					
		4		6		9	
		SNPs	NA %	SNPs	NA %	SNPs	NA %
-r	0.4	8,158	14.3	7,871	11.49	7,560	9.33
	0.7	7,049	9.33	6,772	6.3	6,430	4.69
	0.9	5,981	8.62	5,673	5.44	5,187	2.58

This set of parameters produced a panel of 6,772 SNPs. Pairwise F_{st} distances, calculated with this dataset, were highly significant, suggesting genetic differences occurring among oceanic groups (Table 4.5).

Table 4.5_Pairwise F_{st} values calculated among geographic pools (from Atlantic, Indian and Pacific Ocean) of yellowfin tuna are reported (below diagonal) with their associated P-values (below diagonal). Significant values after Bonferroni standard correction are in bold (nominal significant threshold α = 0.01).

	Atlantic	**Indian**	**Pacific**
Atlantic	*	**<0.01**	**<0.01**
Indian	0.04736	*	**<0.01**
Pacific	0.02932	0.01714	*

The DAPC confirmed the genetic differentiation among oceanic basins. The graph of the Bayesian Information Criterion (BIC) values for increasing values of the number of clusters (k) showed that k=3 corresponded to the lowest associated BIC value. In the data transformation step for PCA analysis, 30 principal components (PCs) were retained, accounting for approximately the 92% of the total genetic variability. The eigenvalues of the DAPC indicated that the first two components explained most of the variation. The resulting scatterplot (Fig. 4.3) showed three genetic clusters corresponding to Atlantic, Indian and Pacific YFT groups. Moreover, the cross-validation of DAPC performed on our dataset, indicated that the number of PCs associated to the highest mean success and lowest mean squared error corresponded to 35 PCs. This result supported our choice to retain 30 PCs during the dimension-reduction step of the DAPC.

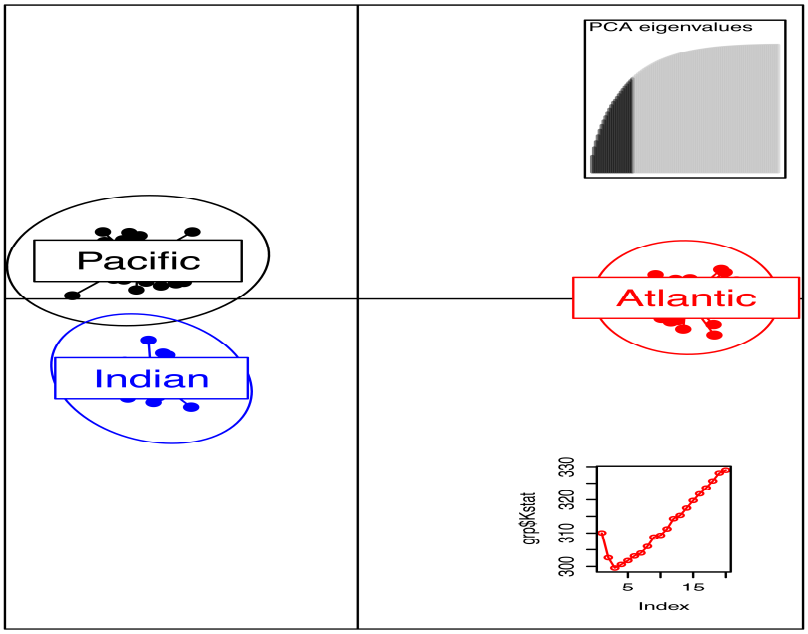


Fig. 4.3. Scatterplot of the DAPC results identifying three genetic clusters of *Thunnus albacares.*

These results are in agreement with previously observed signatures of genetic heterogeneity among oceans found by Ward et al., (1997) by means of significant allele frequency differences at the locus GPI-A (PGI-F). Although, this scenario necessarily needs to be confirmed by increasing the sample size, it validated 2b-RAD genotyping technique as a powerful tool to assess YFT genetic structure and diversity at the global scale.

4.4 Conclusions

This methodological study confirmed that TRs are useful for optimizing genotyping procedure and that they are crucial to reduce the amount of statistical error introduced in allele frequency estimation due to PCR artefacts. We unambiguously mapped the TRs' tags against the reference genome of *T. orientalis* with a high percentage of success (86,59%), in spite of the small size of fragments (Puritz et al., 2014), and the evolutionary distance between these two species. The methodological approach showed that the lack of a reference genome, although undesirable, does not evidently compromise the reproducibility and accuracy of the data obtained, underlying the consistence of the technique in genotyping individuals. We preliminarily demonstrated that 2b-RAD is a promising tool to screen a large set of genomic loci in a marine high gene-flow species, underlying the inter-oceanic population genetic differentiation. Certainly, an increased sample size

is needed to address estimates of genetic differentiation among YFT population samples also at a smaller local geographic scale.

Acknowledgements

This study is part of the PhD research program of Carlo Pecoraro performed at the University of Bologna and the Institut de Recherche pour le Développement (IRD) with the financial support of Bolton Group and Associazione Nazionale Conservieri Ittici e delle Tonnare. We are also grateful for the great contribution in the sampling activities received from the EMOTION project (ANR JSV7 007 01), the International Seafood Sustainability Foundation, the Seychelles Fishing Authority, the Centre de Recherches Océanologiques in Abidjan and the IRD Observatoire Thonier.

References

Appleyard S., Grewe P., Innes B., and Ward R. “Population Structure of Yellowfin Tuna (*Thunnus albacares*) in the Western Pacific Ocean, inferred from microsatellite loci.” Marine Biology 139, (2001): 383-93. DOI: 10.1007/s002270100578.

Allendorf F. W., Hohenlohe P. A., and Gordon L. “Genomics and the future of conservation genetics.” Nature Reviews Genetics 11, (2010): 697–709. DOI: 10.1038/nrg2844.

Andrews K. R. and Luikart G. “Recent novel approaches for population genomics data analysis.” Molecular Ecology 23, (2014): 1661–1667. DOI: 10.1111/mec.12686.

Catchen J., Hohenlohe P. A, Bassham S., Amores A., and Cresko W. A. “Stacks: an analysis tool set for population genomics.” Molecular Ecology 22, (2013): 3124–40. DOI: 10.1111/mec.

Catchen J. M., Amores A., Hohenlohe P., Cresko W., and Postlethwait J. H. “Stacks: building and genotyping loci de novo from short-read sequences.” G3: Genes, Genomes, Genetics 1, (2011): 171–82.DOI: 10.1534/g3.111.000240.

Chow S. and Inoue S. “Intra-and interspecific restriction fragment length polymorphism in mitochondrial genes of *Thunnus tuna* species.” Bulletin National Research Institute of Far Seas Fisheries, 30, (1993): 207-225.

Dammannagoda S. T., Hurwood D. A., and Mather P. B. “Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin tuna (*Thunnus albacares*) from the north Indian Ocean around Sri Lanka.” Fisheries Research 90, (2008): 147–57. DOI: 10.1016/j.fishres.

Davey J. W., Hohenlohe P. A., Etter P. D., Boone J. Q., Catchen J. M., and Blaxter. M. L. “Genome-wide genetic marker discovery and genotyping using next-generation sequencing.” Nature Reviews Genetics 12, (2011): 499–510. DOI: 10.1038/nrg3012.

Díaz-Jaimes P. and Uribe-Alcocer M. “Spatial differentiation in the eastern Pacific yellowfin tuna revealed by microsatellite variation.” Fisheries Science 72, (2006): 590–96. DOI: 10.1111/j.1444-2906.2006.01188. x.

Ely B., Viñas J., Alvarado Bremer J. R., Black D., Lucas L., Covello K., Labrie A. V., and Thelen E. “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 Evolutionary Biology 5, (2005): 19. DOI: 10.1186/1471-2148-5-19.

Excoffier L., and Lischer H. E. “Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows.” Molecular ecology resources, 10, (2010): 564-567.

FAO 2010-2015. Fisheries Global Information System (FAO-FIGIS) - Web site. Fisheries Global Information System (FIGIS). FI Institutional Websites. In: FAO Fisheries and Aquaculture Department [online]. Rome. Updated. http://www.fao.org/fishery/figis/en.

Jombart T., Devillard S., Dufour A. B., and Pontier D. “Revealing cryptic spatial patterns in genetic variability by a new multivariate method.” Heredity 101, (2008): 92–103. DOI: 10.1038/hdy.2008.34.

Juan-Jordá M. J., Mosqueira I., Freire J., and Dulvy N. K. “Life in 3-d: life history strategies in tunas, mackerels and bonitos.” Reviews in Fish Biology and Fisheries 23, (2013): 135–55. DOI: 10.1007/s11160-012-9284-4.

Kunal S. P., Kumar G., Menezes M. R., and Meena R. M. “Mitochondrial DNA analysis reveals three stocks of yellowfin tuna *Thunnus Albacares* (Bonnaterre, 1788) in Indian Waters.” Conservation Genetics 14, (2013): 205–13. DOI: 10.1007/s10592-013-0445-3.

Li W., Xinjun C., Qianghua X., Jiangfeng Z., Xiaojie D., and Liuxiong X. “Genetic population structure of *Thunnus Albacares* in the Central Pacific Ocean based on mtDNA COI Gene Sequences.” Biochemical Genetics 53, (2015): 8–22. DOI: 10.1007/s10528-015-9666-0.

Mastretta-Yanes A., Arrigo N., Alvarez N., Jorgensen T. H., Piñero D. and. Emerson B. C. “Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference.” Molecular Ecology Resources 15, (2015): 28–41. DOI: 10.1111/1755-0998.12291.

Milano I., Babbucci M., Cariani A., Atanassova M., Bekkevold D., Carvalho G. R., Espiñeira M., Fiorentino F., Garofalo G., Geffen A. J., Hansen Jakob. H., Helyar S. J., Nielsen E. E., Ogden, R., Patarnello T., Stagioni M., FishPopTrace Consortium, Tinti F. and Bargelloni L. “Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*).” Molecular Ecology 23, (2014): 118–135. DOI: 10.1111/mec.12568.

Morin P. A., Luikart G., Wayne R. K. and the SNP workshop Group. “SNPs in ecology, evolution and conservation.” *Trends in Ecology & Evolution* 19, (2004): 208–16. DOI: 10.1016/j.tree.2004.01.009.

Morin P. A., Martien K. K., and Taylor L. B. “Assessing statistical power of SNPs for population structure and conservation studies.” Molecular Ecology Resources 9, (2009): 66–73. DOI: 10.1111/j.1755-0998.2008. 02392.x.

Nakamura Y., Mori K., Saitoh K., et al. ″Evolutionary changes of multiple visual pigment genes in the complete genome of Pacific bluefin tuna.″ Proc Natl Acad Sci USA 110, (2013): 11061–6. DOI: 10.1073/pnas.1302051110.

Narum S. R., Buerkle C. A., Davey J. W., Miller M. R., and Hohenlohe P. A. “Genotyping-by-Sequencing in ecological and conservation genomics.” Molecular Ecology 22, (2013): 2841–47. DOI: 10.1111/mec.12350.

Pedrosa-Gerasmio I.R., Babaran R.P., and Santos M.D. ″Discrimination of juvenile yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) tunas using mitochondrial DNA control region and liver morphology.″PLoS ONE 7, (2012): e35604. DOI: 10.1371/journal.pone.0035604.

Papetti C., Di Franco A., Zane L., Guidetti P., De Simone V., Spizzotin M., Zorica B., Kec V. C., Mazzoldi C. “Single population and common natal origin for Adriatic *Scomber scombrus* stocks: evidence from an integrated approach.” ICES Journal of Marine Science: Journal Du Conseil, 70, (2013): 387 –398. DOI: 10.1093/icesjms/fss201.

Puritz J. B., Matz M. V., Toonen R. J., Weber J. N., Bolnick D. I., and Bird. C. E. “Demystifying the RAD Fad.” Molecular Ecology 23, (2014): 5937–42. DOI: 10.1111/mec.12965.

Ward R. D., Elliot N. G., Innes B. H., Smolenski A. J., and Grewe P. M. ″Global population structure of yellowfin tuna, *Thunnus albacares,* inferred from allozyme and mitochondrial DNA variation”. ″Fishery Bulletin 95, (1997): 566-575. DOI: 10.1007/bf00347499.

Wang S., Meyer E., McKay J. K., and Matz M. V. “2b-RAD: a simple and flexible method for genome-wide genotyping.” Nature Methods 9, (2012): 808–10. DOI: 10.1038/nmeth.2023.

Waples R. S., Punt A. E., and Cope J. M. “Integrating genetic data into management of marine resources: how can we do it better?”. Fish and Fisheries 9, (2008): 423–49. DOI: 10.1111/j.1467-2979.2008. 00303.x.

Ying Y., Chen Y., Lin L., Gao T., and Quinn T. “Risks of ignoring fish population spatial structure in fisheries management.” Canadian Journal of Fisheries and Aquatic Sciences 68, (2011): 2101–20. DOI: 10.1139/F2011-116.

Supplementary Materials

Supplementary Material 1- Comparison between *denovo_map.pl* (Left table) and *ref_map.pl* (Right table) performance in terms of: (1) number of SNPs obtained (SNPs/1000) with the associated standard error (SE); (2) error rate (SE); and (3) percentage of heterozygous loci (SE), using different *-m* values (5, 8, 10, 15).

denovo_map.pl				
ID Sample	mvalue	SNPs/1000	error rate	% Ho
34_2_Y_2	15	4.08 (± 0.19)	1.51(± 0.39)	22.13(± 0.37)
	10	4.50(± 0.14)	2.90(± 0.15)	23.12(± 0.15)
	8	5.27(± 0.26)	3.08(± 0.26)	23.70(± 0.18)
	5	5.92(± 0.13)	5.43(± 0.29)	21.12(± 0.27)
77_2_Y_15	15	4.65(± 0.32)	1.27(± 0.21)	25.23(± 0.24)
	10	5.13(± 0.15)	2.42(± 0.13)	26.10(± 0.11)
	8	5.37(± 0.14)	2.71(± 0.32)	27.00(± 0.34)
	5	5.63(± 0.29)	4.95(± 0.24)	24.61(± 0.31)
51_1_Y_7	15	4.74(± 0.27)	1.77(± 0.15)	23.23(± 0.23)
	10	5.22(± 0.29)	2.68(± 0.27)	25.31(± 0.14)
	8	5.52(± 0.12)	2.99(± 0.23)	26.42(± 0.07)
	5	5.71(± 0.09)	5.13(± 0.17)	23.21(± 0.12)

ref_map.pl				
ID Sample	mvalue	SNPs/1000	error rate	% Ho
34_2_Y_2	15	2.89(± 0.26)	3.91(± 0.19)	39.00(± 0.33)
	10	3.21(± 0.23)	4.11(± 0.37)	37.11(± 0.41)
	8	4.08(± 0.14)	4.63(± 0.25)	36.98(± 0.25)
	5	4.18(± 0.16)	5.53(± 0.21)	33.10(± 0.24)
77_2_Y_15	15	3.35(± 0.08)	4.11(± 0.15)	41.00(± 0.15)
	10	3.54(± 0.13)	4.29(± 0.08)	40.00(± 0.18)
	8	3.99(± 0.25)	4.89(± 0.16)	39.80(± 0.23)
	5	4.27(± 0.28)	5.32(± 0.27)	37.40(± 0.16)
51_1_Y_7	15	3.30(± 0.15)	4.28(± 0.14)	42.11(± 0.09)
	10	3.42(± 0.27)	4.61(± 0.18)	40.50(± 0.22)
	8	3.94(± 0.33)	4.76(± 0.23)	39.21(± 0.12)
	5	4.13(± 0.28)	5.41(± 0.24)	36.32(± 0.11)

Supplementary Material 2- Summary statistics of each *Thunnus albacares* individual. The table reports: the sampling origin (Ocean), the mean number (millions) of raw reads, the corresponding mean number (millions) of filtered reads, the percentage of reads retained, the mean value of unique tags, polymorphic SNPs and the number of SNPs found.

ID Sample	Ocean	Raw reads	Filtered reads	Unique tags	Polymorphic loci	SNPs found
T_31_1_Y_12	Atlantic	3105528	2563455	30838	3235	5699
T_31_1_Y_13	Atlantic	5637152	4557441	33788	3402	5824
T_31_1_Y_14	Atlantic	2644769	2285761	29602	3007	5260
T_31_1_Y_15	Atlantic	3288097	2451980	30016	3107	5365
T_31_1_Y_16	Atlantic	6336440	5450801	34926	3611	6249
T_31_1_Y_23	Atlantic	3654650	3263003	31607	3314	5712
T_31_1_Y_35	Atlantic	2721757	2521783	29673	3205	5565

T_31_1_Y_39	Atlantic	2740004	2261642	29808	3215	5694
T_31_1_Y_43	Atlantic	2603891	1676058	27211	2905	5223
T_31_1_Y_34	Atlantic	2122040	985219	22810	3024	5248
T_34_1_Y_13	Atlantic	3131345	2517429	31023	3344	5891
T_34_1_Y_15	Atlantic	3347614	2590822	31639	3333	5867
T_34_1_Y_16	Atlantic	3019207	2308602	30701	3338	5940
T_34_1_Y_17	Atlantic	4386713	3073122	33114	3537	6198
T_34_1_Y_28	Atlantic	6107652	5081470	37117	3964	6819
T_34_1_Y_29	Atlantic	4130588	3147206	33245	3602	6162
T_34_1_Y_3	Atlantic	2846057	2056458	29254	3092	5644
T_34_1_Y_31	Atlantic	4025468	3185735	34137	3617	6281
T_34_1_Y_35	Atlantic	2717490	2111695	30702	3255	5802
T_34_1_Y_37	Atlantic	3145853	2203273	29257	2927	5185
T_34_2_Y_10	Atlantic	4650572	3494452	15529	1347	2493
T_34_2_Y_11	Atlantic	4032665	3227973	32985	3605	6243
T_34_2_Y_14	Atlantic	3029896	2359699	20612	1991	3342
T_34_2_Y_16	Atlantic	4998909	4301682	35583	3732	6550
T_34_2_Y_17	Atlantic	4023014	3294257	32566	3449	5896
T_34_2_Y_20	Atlantic	2376762	1948145	29748	3208	5647
T_34_2_Y_21	Atlantic	4679065	4050647	35695	3752	6534
T_34_2_Y_23	Atlantic	4300413	3598795	34985	3663	6322
T_34_2_Y_27	Atlantic	3494322	2747041	31605	3358	5864
T_34_2_Y_38	Atlantic	2469862	1839548	29513	3118	5454
T_41_1_Y_13	Atlantic	2885225	2229266	29375	3148	5557
T_41_1_Y_14	Atlantic	4112269	3887296	34064	3574	6042
T_41_1_Y_18	Atlantic	3845104	3104426	32644	3423	5868
T_41_1_Y_19	Atlantic	3750752	3150911	32874	3411	5912

T_41_1_Y_20	Atlantic	7135055	6366517	18207	2407	4535
T_41_1_Y_27	Atlantic	5182575	4736811	35547	3742	6319
T_41_1_Y_30	Atlantic	2463935	2048843	29966	3186	5569
T_41_1_Y_31	Atlantic	3984362	3443014	33264	3422	5852
T_41_1_Y_33	Atlantic	3405123	2981862	31292	3412	5969
T_41_1_Y_36	Atlantic	4903957	4112642	34534	3599	6123
T_51_1_Y_11	Indian	4197832	3271746	30808	3259	5746
T_51_1_Y_12	Indian	5106469	3943972	32425	3204	5508
T_51_1_Y_13	Indian	3807089	2659014	30458	3347	5782
T_51_1_Y_15	Indian	5553686	4150067	33413	3463	5822
T_51_1_Y_16	Indian	4224280	3640016	31420	3275	5564
T_51_1_Y_19	Indian	3013892	2312388	30010	3178	5547
T_51_1_Y_28	Indian	3210746	2295315	28312	2962	5234
T_51_1_Y_3	Indian	2884006	2052075	29288	3368	6223
T_51_1_Y_4	Indian	5210472	3911316	26135	2810	4962
T_51_1_Y_5	Indian	2632718	1931146	28444	2763	4917
T_51_2_Y_1	Indian	3522515	3040289	32751	3432	6043
T_51_2_Y_10	Indian	6441211	6003918	31732	3619	6442
T_51_2_Y_11	Indian	2641836	2141771	30187	3242	5766
T_51_2_Y_27	Indian	2888780	2238918	29561	3295	5923
T_51_2_Y_34	Indian	2543649	2073189	29620	3215	5718
T_51_2_Y_44	Indian	3485004	2841056	35543	5501	9745
T_51_2_Y_45	Indian	3135178	2553943	40053	8862	15727
T_51_2_Y_46	Indian	2404384	1941743	29545	3300	5904
T_51_2_Y_47	Indian	2260180	1850287	33939	3679	6280
T_51_2_Y_50	Indian	3937595	3282788	34957	4138	7470
T_71_1_Y_11	Pacific	3897551	3283367	33373	3534	5988

T_71_1_Y_12	Pacific	3394137	2893282	32166	3287	5842
T_71_1_Y_13	Pacific	3604370	3047045	32116	3432	5895
T_71_1_Y_14	Pacific	4000961	3396357	33478	3500	6093
T_71_1_Y_16	Pacific	4727965	3962636	34144	3593	6179
T_71_1_Y_18	Pacific	3818853	3285112	33031	3446	5970
T_71_1_Y_2	Pacific	3440711	2874443	32188	3408	5963
T_71_1_Y_21	Pacific	2508039	2155027	34264	4617	8628
T_71_1_Y_22	Pacific	4044866	3428059	33271	3502	6084
T_71_1_Y_23	Pacific	5676853	4942237	35273	3714	6391
T_71_2_Y_15	Pacific	2656551	2123387	30040	3279	5745
T_71_2_Y_17	Pacific	3377858	2825637	31733	3351	5814
T_71_2_Y_2	Pacific	3159899	2710391	32016	3435	5923
T_71_2_Y_24	Pacific	4837695	4034112	34695	3706	6474
T_71_2_Y_25	Pacific	2863301	2313035	30574	3221	5797
T_71_2_Y_26	Pacific	3121822	2545630	31697	3416	5977
T_71_2_Y_27	Pacific	3391479	2748050	32031	3498	6099
T_71_2_Y_29	Pacific	3345330	2801175	32150	3377	5895
T_71_2_Y_30	Pacific	3719775	2980475	30729	3235	5603
T_71_2_Y_31	Pacific	3951818	3299857	42885	4454	8100
T_77_2_Y_11	Pacific	3618576	3216232	33632	3618	6333
T_77_2_Y_17	Pacific	2403163	1914776	29096	3118	5607
T_77_2_Y_18	Pacific	3866009	3386384	33704	3516	6125
T_77_2_Y_19	Pacific	2806651	2363857	21917	2321	4269
T_77_2_Y_2	Pacific	2780059	2340856	29617	3101	5546
T_77_2_Y_20	Pacific	2549065	2108849	29540	3193	5573
T_77_2_Y_27	Pacific	3398439	3004995	31710	3308	5713
T_77_2_Y_28	Pacific	2919624	2420989	30259	3243	5710

T_77_2_Y_32	Pacific	4317852	3725254	33004	3452	6035
T_77_2_Y_33	Pacific	4371397	3896432	33500	3513	5920
T_77_1_Y_30	Pacific	5765636	5148952	35233	3706	6160
T_77_1_Y_17	Pacific	2794243	2470957	31793	3342	5927
T_77_1_Y_19	Pacific	2723470	2254595	29008	2990	5105
T_77_1_Y_45	Pacific	2400483	2096522	30429	3239	5747
T_77_1_Y_1	Pacific	2155938	1855371	29363	3100	5584
T_77_1_Y_15	Pacific	2085535	1770279	29350	3182	5679
T_77_1_Y_5	Pacific	1857567	1612989	27970	2936	5210
T_77_1_Y_16	Pacific	2035715	1556718	28019	2930	5242
T_77_1_Y_22	Pacific	1796311	1478481	27735	2953	5322
T_77_1_Y_50	Pacific	1990282	1464790	26208	2779	4986

Supplementary Material 3- Results of the *analysis* of molecular variance (*AMOVA*), testing different grouping of the geographic samples: *1) #group1 (34_1, 34_2, 41_1, 31_1); #group2 (51_1 and 51_2); #group3 (71_1, 71_2, 77_1, 77_2); *2) #group1 (34_2, 51_1, 31_1, 77_2); #group2 (71_1, 41_1); #group3 (71_2, 77_1, 34_1, 51_2); *3) #group1 (31_1, 77_2, 51_2, 41_1); #group2 (34_2; 77_1); #group3 (71_2, 34_1, 71_1, 51_2); *4) #group1(34_1, 34_2, 41_1, 31_1); #group2 (51_1, 51_2, 71_1, 71_2, 77_1, 77_2).

*1 Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
---				
Among groups	2	1815.099	2.50242	2.73
P value				<0.01
Among populations within groups	7	304.905	0.88479	0.97
P value				>0.01
Within populations	704	30660.433	88.22814	96.3
P value				<0.01

Total	199	18036.935	91.61535	

*2 Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
---				
Among groups	2	236.373	-0.47002	-0.52
P value				>0.01
Among populations within groups	7	1037.216	2.99888	3.3
P value				<0.01
Within populations	190	16763.346	88.22814	97.21
P value				<0.01

Total	199	18036.935	90.757	

*3 Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
----------				
Among groups	2	226.016	-0.57412 Va	-0.63
P value				<0.01
Among populations within groups	7	1047.573	3.07290 Vb	3.39
P value				<0.01
Within populations	190	16763.346	88.22814 Vc	97.25
P value				>0.01

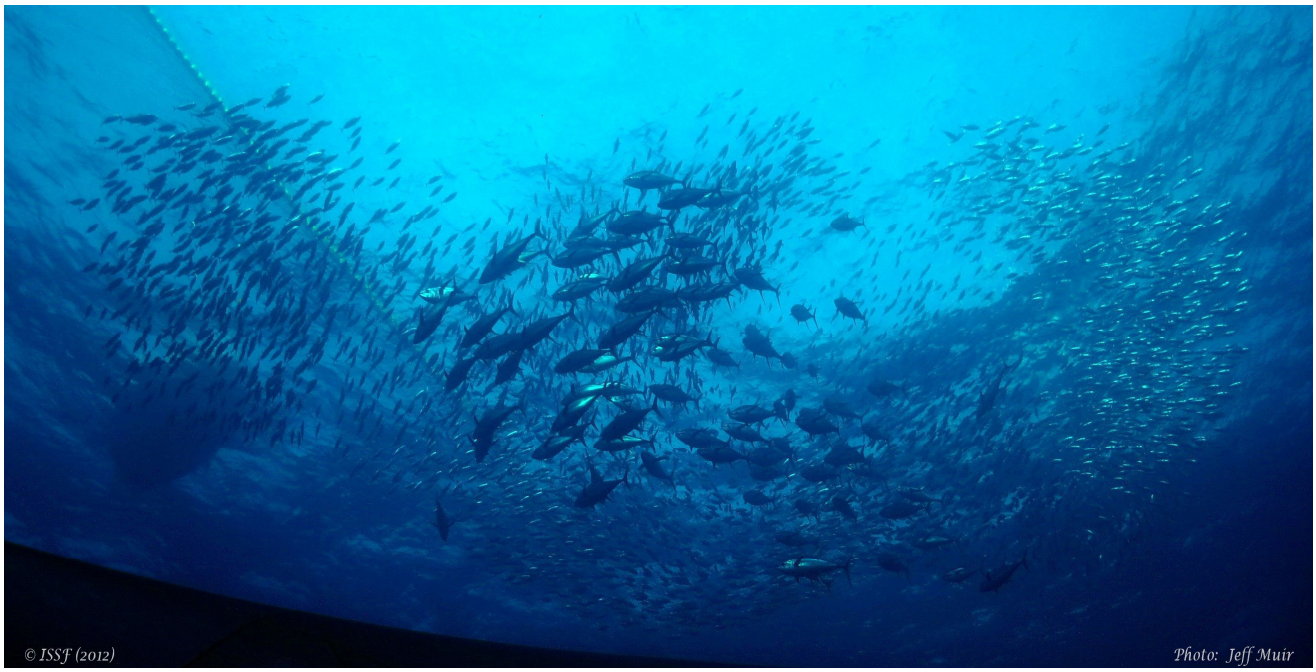
Total	199	18036.935	90.72692	

*4 Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
----------				
Among groups	1	344.227	2.37453 Va	2.58
P value				<0.01
Among populations within groups	8	929.362	1.39798 Vb	1.52
P value				<0.01
Within populations	190	16763.346	88.22814 Vc	95.9
P value				>0.01

Total	199	18036.935	90.72692	

5 Chapter 4

YFT population genomics



5.1 Introduction

The fish stock assessment needs to be based on valid biological assumptions, otherwise management may fail to achieve goals of stock conservation and optimal economic use of fishery resources. The definition of proper management units is a crucial objective for developing proper conservation and management plans (Palsbøll et al., 2007), which need to take in consideration that even low levels of genetic structure can have significant conservation implications (Graves and McDowell, 2003). However, a very central query to apply genetic tools into fishery management is the evaluation of the proper level of divergence among populations in order to explain their separation in terms of management (Waples et al., 2008). In such a context, elucidating the genetic structure of marine fish populations with a high dispersal capacity and large effective population sizes, such as tuna and tuna-like species, remains a big challenge.

In one of the last tropical tuna stock assessment in the Indian Ocean (IOTC-2013-WPTT15-13) it was stated: "if there are different sub-populations managed as a unique panmictic population, some of them could be locally over-exploited and management measures might be directed toward the wrong populations". This sentence becomes particularly significant for yellowfin tuna (*Thunnus albacares*; YFT), whose population genetic structure is a riddle still unresolved with more doubts than certainties. In fact, the results obtained so far, using traditional molecular markers, have been inadequate for testing and validating the appropriateness of the current assessment and management strategy adopted by the four tuna Regional Fishery Management Organizations (tRFMOs; Pecoraro et al., 2016). Precious help to overcome this issue are represented by the advent of Next Generation Sequencing (NGS) technologies and by the consequent decreasing of high-throughput sequencing costs, which have opened the possibility to screen large panels of co-dominant single nucleotide polymorphisms (SNPs) at a genome-wide scale (Davey et al., 2011), even in studies of non-model organisms, such as YFT (Grewe et al., 2015; Pecoraro et al., 2016). Being single base change in the DNA sequence (and often bi-allelic), SNPs are less polymorphic than microsatellite loci. However, their low resolution power, related to the low number of independent alleles, is outweighed by their abundance across the genome that makes them perfect for investigating the inheritance of genomic regions (Baird et al., 2008). Moreover, SNPs are commonly found within or adjacent to coding and regulatory regions of a genome, and so their allele frequencies may be influenced by selection.

NGS have made the Whole Genome Sequencing (WGS) of some marine fish species widely accessible but still prohibitively expensive with extreme intensive computational burden for most laboratories. In the near future, according to the continue development of sequencing technology and bioinformatic methods, WGS might become the main approach for ecological and conservation genomic studies (Bernardi et al., 2012; Wit and Palumbi, 2013). Although the availability of genomic resources for tuna and tuna-like species is expected to increase in the next years, nowadays the draft genome of Pacific bluefin tuna (*Thunnus orientalis*) (Genome size: 800 Mb; Nakamura *et al.*, 2013) is the only genomic resource available for these pelagic marine species. Besides, WGS are unnecessary for population structure studies, especially due to genome size and complexity in eukaryotes (Davey and Blaxter, 2010). Alternatively, genotyping-by-sequencing (GBS) approaches have made population-scale sequencing possible at a fraction of the cost of WGS (Davey et al., 2011). These approaches that selectively target only a fraction of the genome (reduced representation library-RRL) followed by NGS analysis, represent an optimal trade-off for time- and cost-effective high-resolution molecular typing of a sizable set of strains (Narum et al., 2013), making possible to identify and score thousands of polymorphisms across the entire genome and to obtain genotypes for a large number of individuals directly from the sequences (Davey et al., 2011; Hemmer-Hansen et al., 2014). One of the most exploited GBS methods, in a broad range of eukaryotic species, is the Restriction site-Associated DNA sequencing (RAD-seq) genotyping method (Miller et al., 2007; Baird et al., 2008), which links enzymatic fragmentation of the genome with high throughput sequencing for scoring high coverage of homologous SNP loci. Briefly, this method reduces the genome complexity by focusing the sequencing on only the same subset of DNA regions adjacent to recognition sites of a single restriction endonuclease (a modification was implemented using also two restriction endonucleases; Peterson et al., 2012), across multiple individuals (widely discussed in Davey et al., 2011), allowing efficient high-throughput identification of thousands of SNPs spread evenly throughout the genome, in model and non-model species. Therefore, RAD-seq represents a flexible, inexpensive and powerful tool for mapping and analyzing quantitative trait loci (Baird et al., 2008; Chutimanitsakun et al., 2011; Pujolar et al., 2013; Wang et al., 2013), adaptation (Hohenlohe et al., 2010) and phylogeography (Emerson et al., 2010). From the original RAD-seq technique, other methods have been developed, such as double digest RAD (ddRAD, Peterson et al., 2012), type IIB endonucleases RAD (2b-RAD, Wang et al., 2012) and ezRAD (Toonen et al., 2013). Each method is certainly characterized by specific strengths and weaknesses (Puritz et al.,

2014). Among the different RAD-seq approaches, in this work we employed 2b-RAD genotyping technique, as already extensively described in Chapter 3, according to its extreme protocol simplicity and cost-efficiency (Wang et al., 2012). In fact, 2b-RAD library preparation procedure involves consecutive addition of reagents to the same 96-well plate and there are not intermediate purification steps or need for special instrumentation beyond a PCR instrument and a standard agarose gel (Puritz et al., 2014). This protocol requires only 50-bp single-end Illumina sequencing and 2b-RAD tags are sequenced at both strands allowing also the use of strand bias as a quality filtering criteria. Moreover, according to goals of population genetic structure studies, it is possible to select a less amount of markers by means selective adaptors (Puritz et al., 2014).

However, sometimes just multiplying the number of neutral loci might be not enough to assess correctly the population structure of these highly migratory marine species especially at the local scale (Gragnaire et al., 2015), representing a failure in assisting their management and stock assessment. On the contrary, the identification of candidate loci, which exhibit patterns of variation extremely divergent from the rest of the genome (hereafter referred to as 'outlier' loci; Luikart et al., 2003) might provide new opportunities for identifying reproductively isolated populations, upon which the current fisheries management practices should be based.

The main objective of the present study was to assess how using many SNP markers derived from 2b-RAD loci can contribute to disentangle YFT genetic structure and increase the assignment success of individuals to sites of origin within and between RFMOs. Subsequently, we concentrated on outlier loci and explored subsets of these markers loci to improve the resolution of population structure patterns of this pelagic fish species.

5.2 Materials and Methods

5.2.1 2b-RAD libraries preparation and sequencing

A total of 500 juveniles YFT from ten geographic population samples (Fig. 5.1; Section.1) were analyzed, covering the entire species distribution.

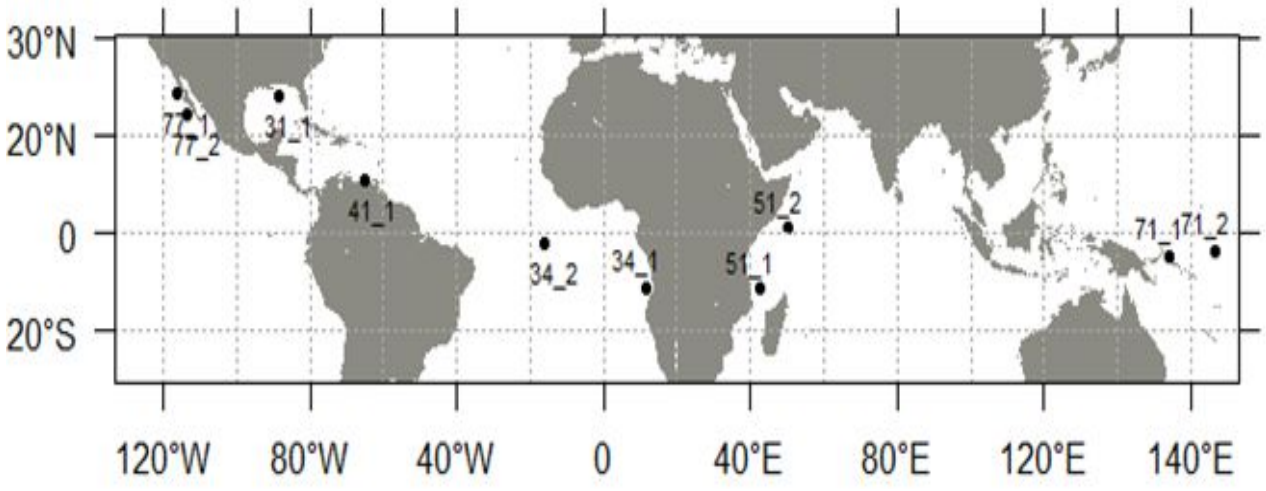


Fig. 5.1. Location of *Thunnus albacares* geographical population samples analyzed in this study. Sample codes are given as in Table 1.

The 2b-RAD libraries were constructed for each individual following the protocol from Wang et al., (2012) that is schematically illustrated in Figure 5.2 (see Chapter 3 for details).

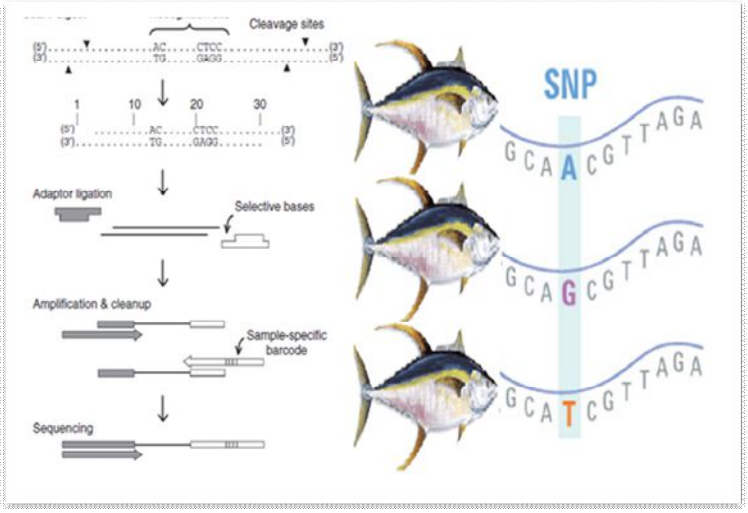
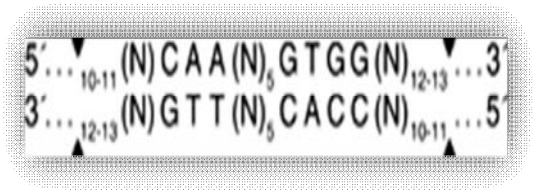


Figure 5.2. Different steps of 2b-RAD library preparation.

Minor modifications were introduced during library preparation in order to exclude samples with low concentration and suboptimal quality profile and minimize possible PCR artefacts.

An important modification is represented by the greater digestion volume used comparing to that proposed in the original protocol (10 µL vs. 6 µL), keeping the same final concentration of buffer R (10X) and S-adenosylmethionine (SAM; 150 µM). This modification was linked to the low quality of our DNA samples. In fact, in this way we were able to increase the efficiency of the reaction in two ways: 1) diluting potential inhibitors; 2) reducing potential pipetting error.

The restriction endonuclease used for digesting the genomic DNA (gDNA) was CspCI, which has a 7 bp recognition site (5'-CAA-GTGG-3' and its reverse complement 3'-GTT-CACC-5'; Fig. 5.3). CspCI cleaves DNA fragments twice to excise its recognition site, producing 34-36 base-pair fragment



with 2-base 3′ overhangs. The cleavage position may shift of one base, depending on the DNA sequence context between the recognition site and the point of cleavage.

Figure 5.3. CspCI's recognition site.

In order to avoid possible amplification errors, the 2b-RAD tags amplification was split in two separate 25 µL-reactions, which were pooled together only at the end. In fact, being the PCR a random process, it is very unlikely that the same error will be repeated in the same position during the reaction. PCR products were loaded on a 2% agarose gel to confirm that their molecular weight was ~165 bp (Fig. 5.4).

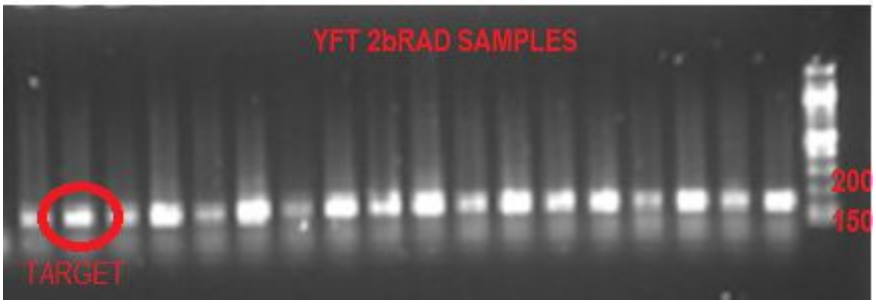


Figure 5.4. Gel electrophoresis of PCR products .

PCR products were purified with the SPRIselect purification kit (Beckman Coulter, Pasadena, California, USA) to exclude any high-molecular weight DNA remained after the enzyme digestion, and any incorrect constructs that may rise during PCR. The purification begins adding 1.5X of SPRIselect beads (85 µL) to the amplified solution. After removing the supernatant and washing with ethanol, the elution buffer was added. The eluted material was then transferred into a clean plate and the concentration of purified individual libraries was quantified both using Qubit®ds DNA BR Assay Kit (Invitrogen–ThermoFisher Scientific, MA, USA) and Mx3000P qPCR instrument. Finally, their quality was also checked on the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA). Individual libraries were pooled into equimolar amounts, according to their unique sample specific barcode and the remaining steps were carried out on the pools, reducing the laboratory's labor and costs. In fact, sample multiplexing allows to pool large number of libraries, which can be sequenced simultaneously during a single sequencing run. The unique index added to each sample during library preparation allows identifying and sorting them before data analysis. The resulting four

pools' quality were re-verified on Agilent 2100 Bioanalyzer, before sending them for sequencing (Fig. 5.5).

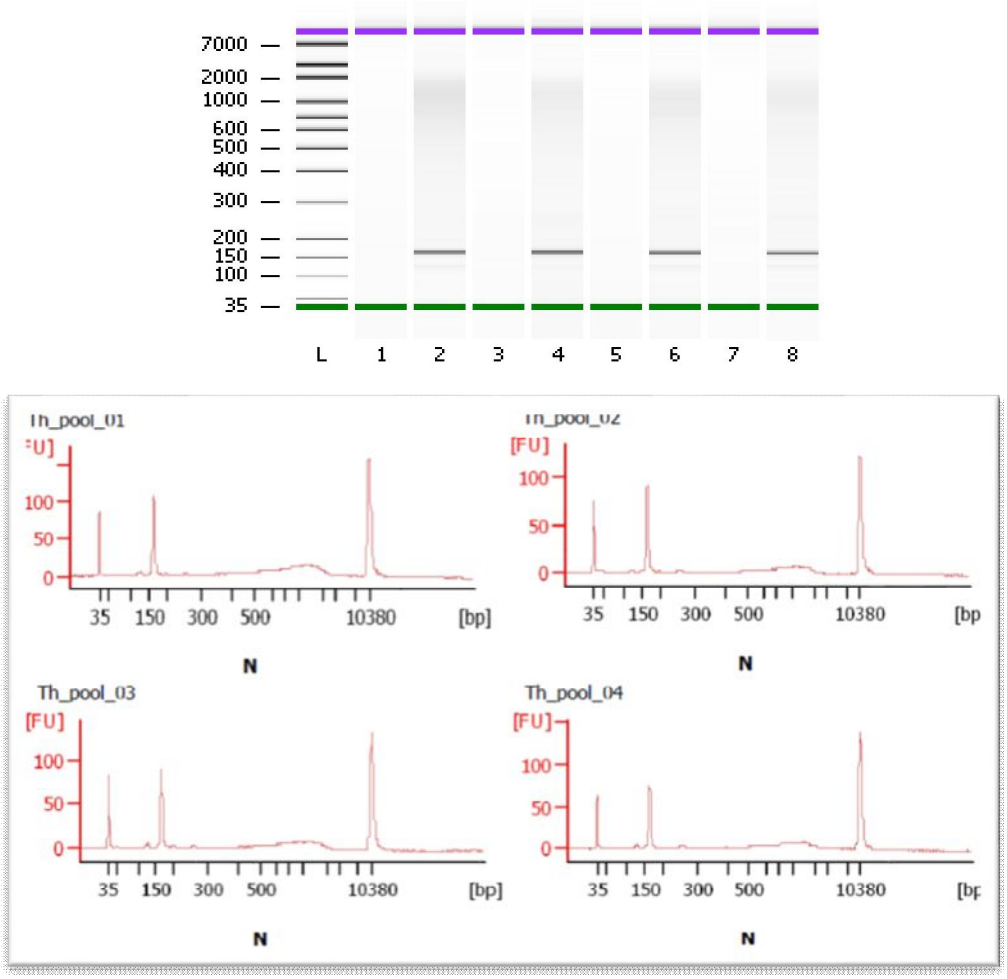


Fig.5.5_Bioanalyzer's quality control results for the four pools.

5.2.2 Sequencing and quality control

The sequencing process was performed through an Illumina platform, which uses the Sequencing By Synthesis (SBS) technology that is one of the most successful and widely Next-Generation Sequencing (NGS) technology used so far. Illumina sequencing reagents support massively parallel sequencing using a method that detects single bases as they are incorporated into growing DNA strands. Illumina reads are normally 25-250 nucleotide long sequences produced by a reversible-terminator cyclic reaction associated to base-specific colorimetric signals within the sequencing machine (Del Fabbro et al., 2013). Besides, SBS technology supports separated reads "single read" or paired-end reads runs. The main difference between the two runs is that with single read the

sequencing instrument reads the fragment from one end to the other one, instead in paired-end runs the instrument reads the fragment twice from both sides. A fluorescently labeled reversible terminator is imaged as each deoxynucleotide (dNTP) is added and then split to allow incorporation of the next base. Since all four reversible terminator-bound dNTPs are present during each sequencing cycle, natural competition minimizes incorporation bias.

Pooled libraries were sequenced on an Illumina HiSeq2500 platform with a 50 bp single-read module at the Genomix4Life S.r.l. facilities (Baronissi, Salerno, Italy). The expected number of reads generated during the DNA sequencing reaction using an Illumina HiSeq2500 instrument was 150,000,000 reads per lane. Two lanes were dedicated per each pool (filling an entire Illumina Genome Analyzer flowcell) in order to increase the number of reads obtained and the expected coverage per each library. The theoretical or expected coverage is the average number of times that each nucleotide is expected to be sequenced, considering a certain number of reads of a given length and assuming that reads are randomly distributed across the genome (Sims et al., 2014). Instead the depth of coverage is referred to its redundancy (Sims et al., 2014).

Data demultiplexing, which is the computational process by which reads are assigned to the proper sample according to their specific barcode, was also performed by the sequencing provider. It occurred during Bcl (files generated by Real Time analysis-RTA software, that performs primary analysis for Illumina's sequencing instruments) to Fastq processing. This step was performed with the script *ConfigureBclToFastq.pl* included in CASAVA package version 1.8.2, which is an internal Illumina software, allowing 0 mismatches in the index sequence.

Demultiplexed data were returned in Fastq format, which includes quality scores in addition to the sequence itself. Fastq files contain four lines: 1) "@" and the sequence ID; 2) the sequence; 3) "+"; and 4) an ASCII-encoded quality number corresponding to a PHRED score (Q). These data were easily checked by using the FastQC software (www.bioinformatics.babraham.ac.uk/projects/fastqc/). This software provides a Quality Check (QC) report, which reveals any possible bias or problems in the raw data that may affect upstream the analysis. In fact, ignoring the existence of low quality base calls may in fact be disadvantageous for any NGS analysis (Del Fabbro et al., 2013). A custom-made Perl script was run to merge the sequences obtained from the two lanes and to perform the quality filtering of the reads and adaptors trimming. This script turned those reads with the restriction site in reverse complement and filtered out all of those ones that do not 100% match the restriction site. In the remaining reads the script removes adaptors and cuts every sequence before and after the restriction sites.

This procedure, without changing the original read dataset, allowed to eliminate only low quality regions of the sequence.

5.2.3 STACKs analysis and optimization of genotyping procedure

2b-RAD data were analyzed with the software *Stacks v. 1.32* (http://creskolab.uoregon.edu/stacks/), which is a computational pipeline implemented by component programs written in C++ and Perl and designed to work with any restriction enzyme–based GBS data (Catchen et al., 2011; 2013). *Stacks* allows genotype inference through the identification of SNP loci without a reference genome (*denovo_map.pl* program) or aligning reads against a reference genome (*ref_map.pl* program). This software allows to work with thousands of individuals and incorporates a MySQL database and web front end to efficiently visualize, manage and modify the data. Data analysis consisted of different steps such as acquiring raw sequence data, filtering out low-quality reads, assembling or aligning reads, and inferring SNPs and genotypes (Catchen et al., 2013).

The Stacks pipeline was run to perform the following processes:

1) grouping data from each individual into loci, identifying polymorphic nucleotide sites (*ustacks* and *pstacks*);

2) creating a catalogue of loci across individuals (*cstacks*);

3) matching loci from each individual against the catalogue to determine the allelic state at each locus in each individual (*sstacks*);

4) subjecting allelic states to population genetic statistics, exporting the results in formats usable for other software (i.e. Genepop, Structure).

Each process is characterized by different challenges and uncertainties associated with the genomic attributes (i.e. number of loci identified, the degree of repetitive sequences throughout the genome, the level of polymorphism and divergence among populations) as well as with biological factors (i.e. quality of DNA and degree of sample multiplexing) that may interact with sequencing characteristics (Catchen et al., 2013). However, the Stacks' module "rxstacks" (associated with both *denovo_map.pl* and *ref_map.pl* programs) corrects the genotype data based on population-level inferences. For instance, a log likelihood is assigned to each locus in the population according to their coverage in terms of primary and secondary reads. This has the effect of removing loci with sequencing errors or error produced during PCR for certain types of loci, such as those with lots of short repeats. The expected result for non-repetitive genomic

regions is that a monomorphic locus will generate a single stack because the two sequences on the two homologous chromosomes are identical and thus indistinguishable. On the contrary, a polymorphic locus will produce two different stacks. Possible confounding cases can be limited by the use of Stacks parameters (See Chapter 3 for details about preliminary tests conducted on TRs).

5.2.4 SNPs selection and genetic diversity analysis

Since each *de novo* sample may bring a small amount of error into Stacks' catalog of loci, working with hundreds of samples could bring lots of error in the catalog. For this reason and according to the promising results gathered with our technical replicates (Chapter 3), library's reads were mapped against the *Thunnus orientalis* genome, filtering out possible erroneous SNPs from the catalog. The final analysis was carried out running the *ref_map.pl* program on the entire dataset, using the most reliable Stacks' parameter identified on the mapped TRs (see Chapter 3 for more details). Using the Stacks module *populations,* which has a number of filtering parameters, from the resulting catalog of loci we selected only those containing (i) one bi-allelic SNP; (ii) scored in at least two geographic samples (*-p*); and (iii) represented in more than 60% of the individuals within each sample *(-r*). The resulting dataset was exported in a Genepop formatted file.

Some quality filtering criteria were used to the resulting dataset in order to remove possible sequencing, genotyping errors and uninformative polymorphisms using the last version of the Rpackage Adegenet 2.0.1 (Jombart, 2008, R version 3.1.2, R Development Core Team, 2014; http://www.r-project.org). First, any putative SNP that failed to be genotyped in >80% of individuals was removed as well as any individuals that was not genotyped in >80% of loci. Then those loci with a minor allele frequency (MAF) lower than 1% in all samples were filtered out, using the new function minorAllele implemented in the last version of Adegenet 2.0.1 (https://github.com/thibautjombart/adegenet). These polymorphisms are likely to be uninformative and they can distort signals of selection and drift in natural populations, biasing tests for selection (Roesti et al., 2012). We also conducted exact tests of Hardy-Weinberg equilibrium, using the function hw.test of the Rpackage Pegas (Paradis et al., 2015), removing loci that were out of equilibrium in 9 or 10 samples ($P < 0.05$). The dataset was then converted to the specific file formats required by the different software using PGD Spider (Lischer and Excoffier, 2012) in order to perform the population structure analyses.

5.2.5 Population structure analyses

The genetic diversity within samples in terms of observed (Ho) and expected heterozigosity (He) was calculated using the software Arlequin 3.5.1.2 (Excoffier and Lischer, 2010). Linkage disequilibrium (LD) between each pair of loci in each population was calculated using Genepop (1,000 iterations per batch for 100 batches; Raymond and Rousset, 1995).

To get more insight into the genetic relationship and differentiation among all the geographical samples, we first estimated the F_{ST} values for pairwise comparison of all samples and then we performed the analysis of molecular variance (AMOVA) with Arlequin 3.5.1.2, using 20,000 permutations and 0.01 significance as settings.

Based on the SNPs dataset obtained, different multivariate analyses were performed using the Rpackage Adegenet 2.0.1 to assess the genetic structure. Firstly, we performed a Principal Component Analysis (PCA), which is the most common multivariate analysis used in genetics. On the genind object the allelic data (as frequencies) were extracted with the accessor *tab* and then using the function *dudi.pca* () we replaced missing values with mean values. After this step, we chose a number of principal components (PC) from the screeplot displayed (barplot of eigenvalues) that mostly explained the variance of our data. In fact, each eingenvalue represents the variance explained by each PC. Being more interested in the diversity among groups of individuals than in the total diversity of our data, we also performed a Discriminant Analysis of Principal Components (DAPC; Jombart et al., 2010). This method identifies an optimal number of genetic clusters that maximizes the variation among groups by running a k-means clustering algorithm. The first step is the transformation of the data through a PCA analysis, which has the major advantages of reducing the number of variables and speeding up the clustering algorithm. In addition, during this step all the PCs can be retained without losing information. Since the number of populations represented by our data was not a priori known, the *find. clusters* () function was used for running successive number of clusters (K)-means clustering of the individuals for K= 1:20, and identified the best supported number of clusters through comparison of the Bayesian Information Criterion (BIC) for the different values of K. The best value of K that maximizes the variation between groups is usually indicated by an elbow in the curve of BICs (Jombart et al., 2010). I also used *table.value* () function to build a contingency table and to check how the inferred groups matched with our geographical samples. After having found the best number of clusters that described the genetic variance of the data, we run a Discriminant Analysis

(DA) on the retained principal components using the *dapc* () function. The function displayed a similar graph of cumulated variance as in *find.cluster* (), but in this case we can choose the minimum number of components that maximizes the cumulative variance without losing too much information. Otherwise, the retention of too many components may make unstable the membership probabilities returned by the method.

In order to avoid over-fitting, in this analysis 40 principal components (PCs) were retained during the data transformation step for PCA analysis (indicated to be the optimal number based on the *optim.a.score* () function), which accounted approximately for the 81% of the total variation in the data set. The cross-validation test is crucial because a wrong choice of the number of PCs might negatively impact the DAPC results and produce unstable output due to over-parameterization.

Finally, after having chosen the best number of eigenvalues for the DA analysis, the DAPC results (DAPC scatterplots) were visualized graphically by using the *scatter.dapc* () function, adding some graphical possibilities to improve the display of the analysis. The function *compoplot* () function was also run to summarize the overall picture of the clusters' composition obtained, categorizing samples with mean membership probability of >0.6 to one of the clusters as 'pure' sample.

YFT population genetic structure was further assessed by using the software STRUCTURE v.2.3.4 (Pritchard et al., 2000). STRUCTURE was run ten times at K=1-12, setting a burn-in period of 50,000 iterations and 100,000 iterations for the MCMC. Correlated allele frequencies (Falush et al., 2003), admixture ancestry and sampling locations were assumed using the LOCPRIOR model. The most likely number of K was assessed using the DK statistics (Evanno et al., 2005) implemented in the online program STRUCTURE HARVESTER (Earl and von Holdt, 2012). This analysis was first performed at the global scale using the entire data set. Then, in order to investigate the population structure within each oceanic basin, the analysis was separately re-computed on the Atlantic, Pacific and Indian samples. Finally, the software CLUMPAK (Cluster Markov Packager Across K.; Kopelman et al., 2015) was used to determine the optimal assignment of clusters for the individuals analyzed.

The software GENECLASS 2.0 (Piry et al., 2004) was employed to assign individuals to each population sample, according to the highest likelihood of their multilocus genotype. We employed the Bayesian approach described by Rannala and Mountain (1997) to evaluate whether a certain multilocus genotype could occur in one or several populations using the Monte-Carlo re-sampling algorithm described by Paetkau et al., (2004). The likelihood of individual genotypes from each sample locality was calculated by comparing individual genotypes to 10,000 simulated individuals

per locality. A random selection with the 75% of individuals was selected from each sample and pooled by the three geographic regions to serve as assignment references, while the other 25% was used as queries.

5.2.6 Outlier loci detection analysis

Finally, we searched for loci that showed divergent patterns of differentiation compared to neutral expectations, and therefore potentially affected by selection (outlier loci), using two different approaches.

First, we used the software BAYESCAN 2.1 (Foll, 2012), a Bayesian approach that allows the estimation of the posterior probability of a given locus being under the effect of selection using differences in allele frequencies between populations (Foll, 2012). It is based on the multinomial-Dirichlet model and assumes that allele frequencies among sub-populations are correlated through a common migrant gene pool, therefore allowing complex ecological scenarios to be modeled adequately (Foll, 2012). We ran Bayescan with 20 pilot runs of each 5000 iterations followed by an additional burn-in of 50 000 iterations and then 5000 samplings with a thinning interval of 10. Correcting for multiple testing, the program computes q-values based on the posterior probability for each locus. We defined as "outliers", those loci with an alpha-value significantly >0 (i.e. with q-values smaller than 0.05). Instead those with an alpha-values significantly smaller than 0 were considered under balancing selection. The remaining loci were classified as neutral (Moore et al., 2014).

Then, we used a hierarchical Fdist model (Excoffier et al., 2009) implemented in Arlequin 3.5 (Excoffier and Lischer, 2010), to identify loci that maximize assignment to the sampling location of origin. Loci with significantly (at the 0.05 significance level) higher FCT or FST values were classified as outliers potentially under directional selection among groups. Instead loci with significantly lower FCT or FST values were classified as loci putatively under balancing selection, while all other ones were classified as putatively neutral loci (Moore et al., 2014).

To assess how much the YFT population structure is driven by loci under selection, a DAPC analysis was performed with the subset of those loci detected as outliers using the two different approaches.

5.3 Results

5.3.1 Sequencing and quality control

The overall FastQC result pointed out from each QC report indicated a good quality of the libraries without spotting any problems originated either in the sequencing process or in the starting library material (i.e. contaminations, overrepresented sequences, adaptor sequence presence). In fact, all of our libraries had a quality value comprised between 28 and 40, locating them in the green portion of the graph (Fig. 5.6).

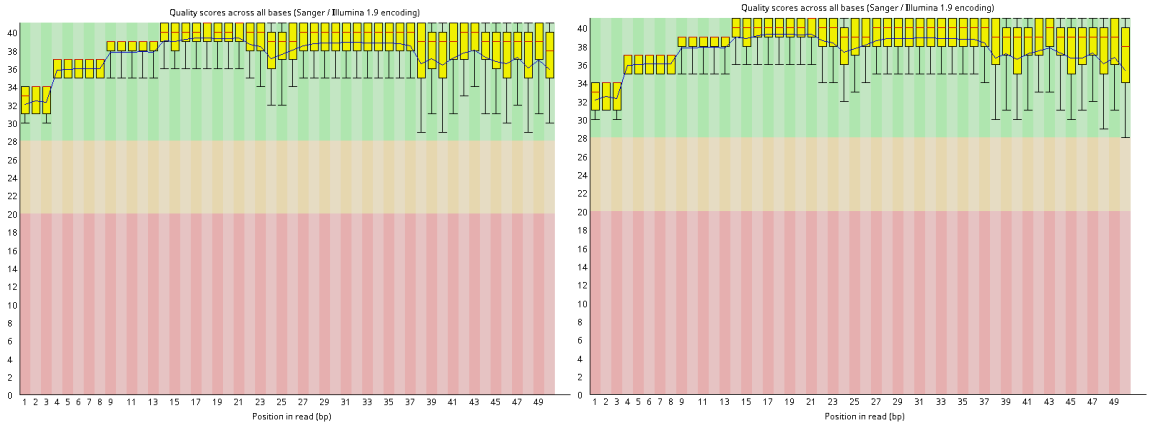


Fig. 5.6_Quality scores of all nucleotides in the two samples sequenced.

The fig. 5.6 represents the general trend of all nucleotides' quality scores for all of the libraries sequenced; however negligible differences among them were detected. In general, a higher quality value was identified in the central portion of the graph, with a decrease moving towards the extremities as pointed out by the mean quality (blue line).

This general qualitative pattern in the reads was also confirmed by the nucleotide composition graph (Fig. 5.7), which showed a clear decrease of their variability in correspondence of two specific positions: 1) at the CspCI restriction sites; and 2) at the extremities of the graph (i.e. adaptors' position).

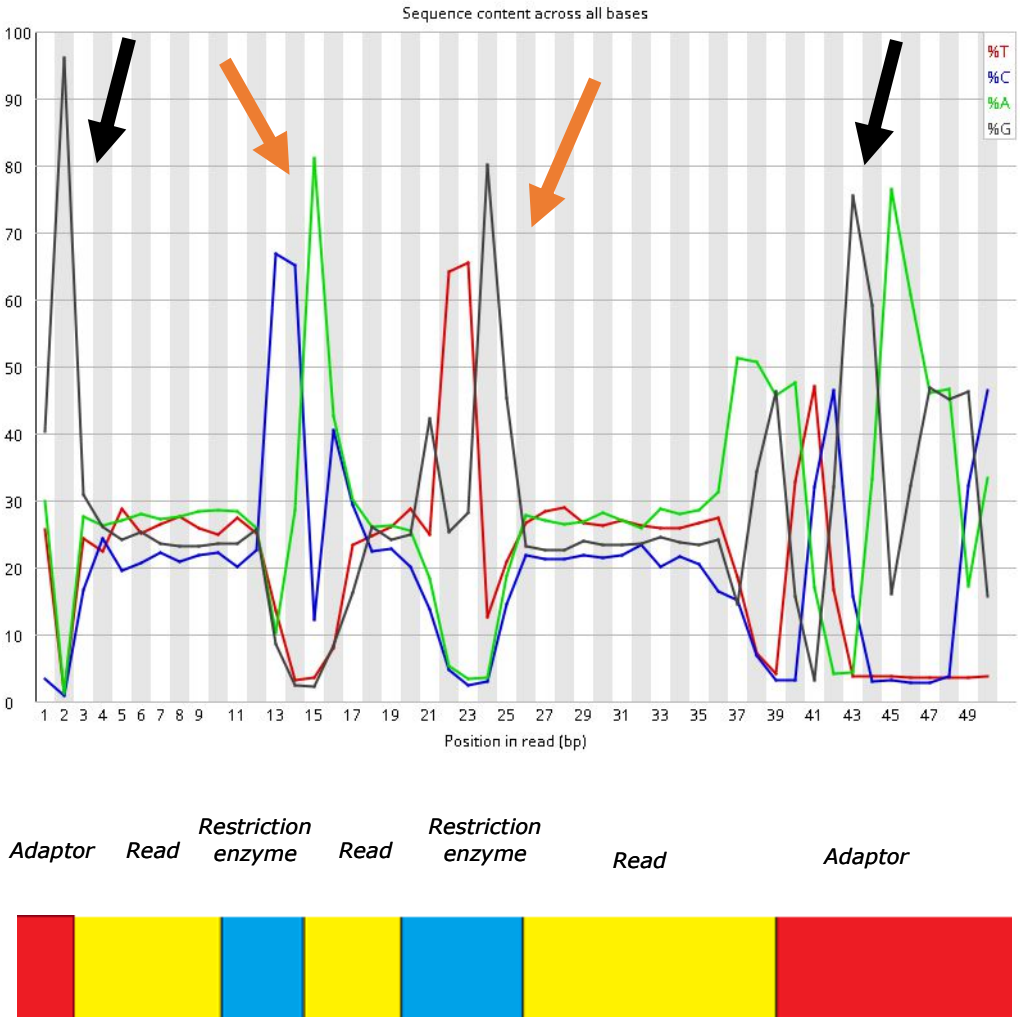


Fig. 5.7_ Graph showing variability of nucleotides in the read.

After quality filtering and adaptors trimming, we obtained sequences of 34 bp and in average a total of 2,636,921 reads per each individual from each pool (Fig. 5.8). In average about the 20% of raw reads per pool were lost after the filtering and trimming step. After this process, eight individuals with less than 500,000 reads were removed from the dataset.

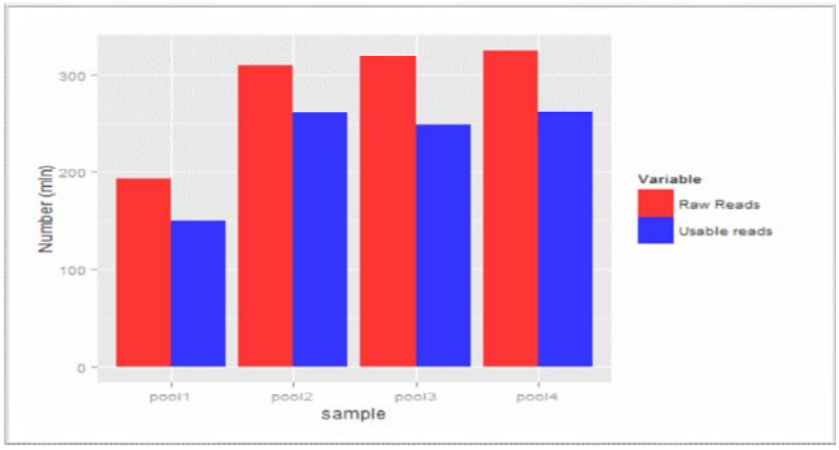


Fig. 5.8_ Number of raw reads obtained per each pool after sequencing and the corresponding number obtained after quality filtering.

5.3.2 SNPs selection and genetic diversity analysis

The *ref_map.pl* program was run on the resulting dataset obtained after the filtering process (Table 5.1).

Table 5.1. Summary statistics of the 10 *Thunnus albacares* geographical population samples. The table reports: The samples code (ID), the sampling origin (location), the mean number (millions) of raw reads, the corresponding mean number (millions) of filtered reads, the percentage of reads retained, the mean value of unique tags (with SE), polymorphic SNPs and number of SNPs found with the associated standard error. Moreover, observed and expected heterozygosity per each geographical sample are reported.

ID code	Sampling location	Raw reads	Filtered reads	% of reads retained	Unique Tags	Polymorphic Loci	SNPs Found	Obs Het	Exp Het
31_1	W Atlantic	3312337	2526644	24.9	26885(± 0.42)	1824(± 0.03)	2080(± 0.03)	0.12	0.14
34_1	E Atlantic	3433692	2670318	34.1	23770(± 0.78)	1617(± 0.05)	1851(± 0.06)	0.10	0.13
34_2	E Atlantic	3159125	2481523	23.5	26624(± 0.79)	1806(± 0.06)	2066(± 0.07)	0.13	0.14
41_1	W Atlantic	3244529	2713851	17.8	26663(± 0.76)	1782(± 0.04)	2031(± 0.05)	0.14	0.16
51_1	W Indian	3674435	2683577	27.8	27876(± 0.46)	2213(± 0.09)	2614(± 0.12)	0.14	0.19
51_2	W Indian	4081030	3527438	14.5	30110(± 0.37)	2009(± 0.02)	2294(± 0.03)	0.12	0.13
71_1	WC Pacific	3623456	3160638	13.2	27588(± 0.50)	1900(± 0.04)	2167(± 0.05)	0.10	0.12
71_2	WC Pacific	2878098	2363162	19.1	23015(± 0.50)	1564(± 0.03)	1786(± 0.04)	0.12	0.13
77_1	E Pacific	3328447	2950684	17.3	28110(± 0.31)	1909(± 0.06)	2094(± 0.04)	0.13	0.15
77_2	E Pacific	2768277	2464614	18.2	24494(± 0.42)	1677(± 0.42)	1903(± 0.42)	0.12	0.13

Using the Stacks' module *population*, we exported a dataset of 370 individuals genotyped at 6,896 loci. In the Stacks' output file together with each individual's genotype, the position of each SNP in the read was also reported, which made possible to verify that no SNPs were present at the

restriction sites as well as at the extremities, highlighting the good success of the filtering process (Fig. 5.9).

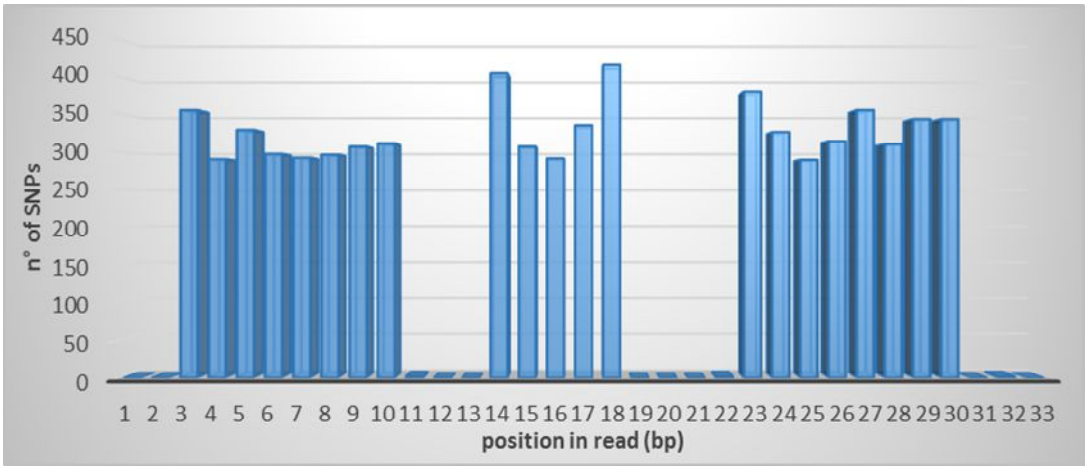


Fig.5.9_SNPs' number at every bp in the read.

Removing putative SNPs that were genotyped in <80% of individuals and individuals that were genotyped in <80% of SNPs, we eliminated 542 loci and 13 individuals: seven individuals belonged to the population sample 41-1-Y, four to the population sample 77-1-Y, and two from the population sample 31-1-Y. After removing SNPs with minor allele frequency (MAF) <0.01, 1,443 SNPs were retained. Significant deviations from Hardy–Weinberg equilibrium were observed in 471 SNPs, which were also removed. The final filtered data set consisted of 357 individuals genotyped at 972 SNPs with a similar total number of alleles per locus among samples, independently from their sample size (Fig. 5.10).

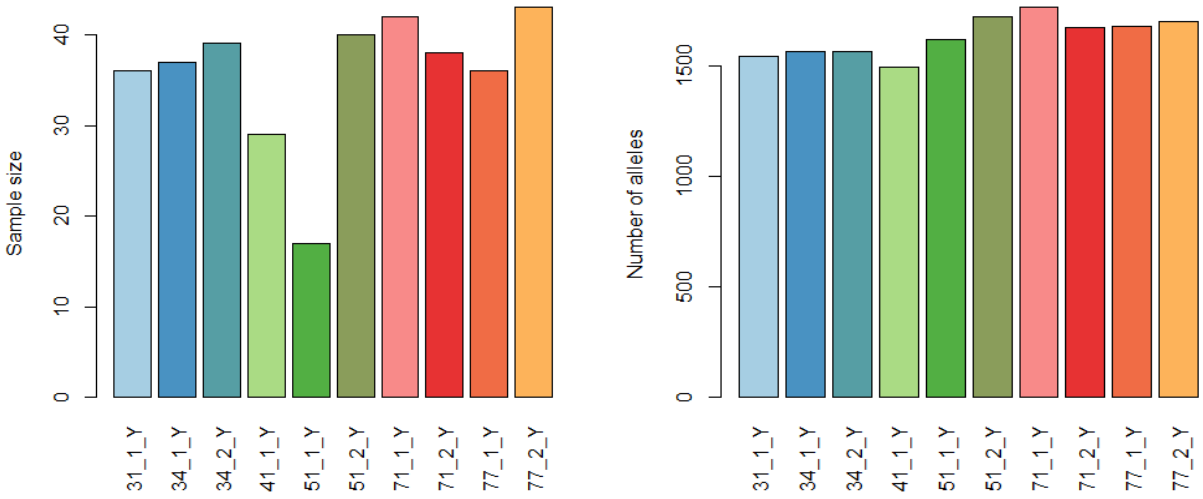


Fig. 5.10_Number of individuals and respective number of alleles per population sample.

The mean observed heterozygosity per variable SNP site was lower (0.10-0.14) than that expected (0.12-0.19) within each population sample (Summary statistics in Table 5.1). This pattern was clearly observed plotting the average observed and expected heterozygosities per each locus (Fig. 5.11).

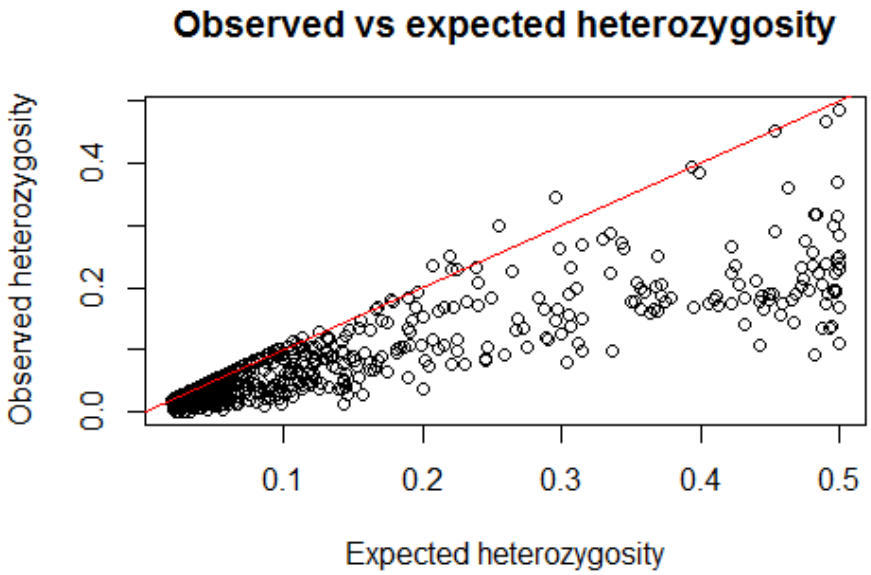


Fig. 5.11_The average observed and expected heterozygosity per each locus.

5.3.3 Assessment of the population structure

Pairwise F_{ST} distances, being highly significant ($P<0.01$), indicated that genetic differentiation occurs among the three oceans. Instead panmixia was revealed within each oceanic basin (Table 5.2).

Table 5.2_Pairwise F_{ST} values calculated among population samples of yellowfin tuna are reported (below diagonal) with their associated P-values (below diagonal). Significant values after Bonferroni standard correction are underlined (nominal significant threshold α = 0.01). NS: not significant.

Pairwise FST values										
	1	2	3	4	5	6	7	8	9	10
1	*	NS	NS	NS	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
2	0.01053	*	NS	NS	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
3	0.01217	0.00780	*	NS	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
4	0.00983	0.00903	0.00876	*	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
5	0.13639	0.11697	0.11648	0.12461	*	NS	<0.01	<0.01	<0.01	<0.01
6	0.13077	0.11208	0.11166	0.11967	0.00078	*	<0.01	<0.01	<0.01	<0.01
7	0.08871	0.07493	0.07400	0.08012	0.03633	0.03985	*	NS	NS	NS
8	0.07722	0.06633	0.06629	0.06897	0.04159	0.04463	0.00221	*	NS	NS
9	0.10268	0.08597	0.08493	0.09521	0.04080	0.04412	0.00571	0.00745	*	NS
10	0.10569	0.08820	0.08697	0.09749	0.04221	0.04535	0.00300	0.01165	0.00888	*

These patterns of differentiation were also confirmed by the hierarchical AMOVA analysis, grouping the population samples per oceanic basin. The AMOVA results (Table 5.3) indicated that the 8.39 % of the total variation was explained by the oceanic groups subdivision (P-value <0.01) and at the same time not significant differences were observed among populations within groups (0.08% P-value >0.01).

Table 5.3_AMOVA table of the spatial genetic variation of yellowfin tuna samples grouped per Oceanic basins.

Source of variation	Sum of squares	Fixation indices	Percentage of variation	Pvalue
Among groups	1912.937	FST: 0.08478	8.39699	<0.01
Among populations within groups	304.905	FSC: 0.00089	0.08112	>0.01
Within populations	30660.433	FCT: 0.08397	91.52189	<0.01
Total	33609.561	50.01571		

Principal Component Analysis (PCA) results confirmed the genetic heterogeneity among the geographic samples considered in the study (Fig. 5.12). PCA grouped individuals in clusters according to their populations of origin.

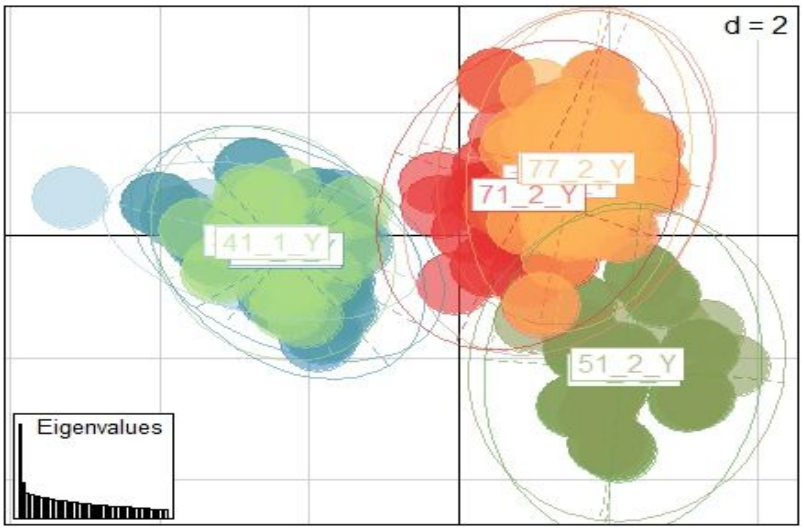


Fig.5.12_The genetic clusters, corresponding to the three oceanic groups, detected by PCA results.

The presence of these three clusters was also confirmed by plotting the correspondence between inferred groups and the geographic population samples (Fig.5.13).

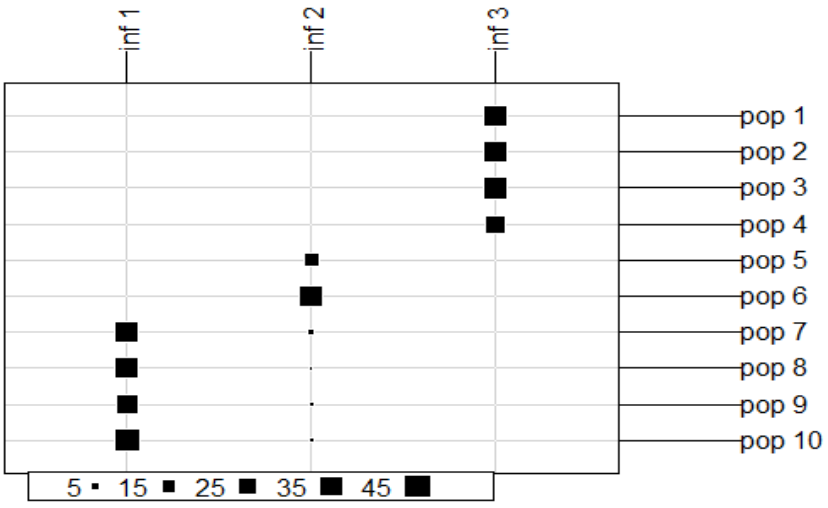


Figure 5.13. Correspondence between inferred groups and population samples.

The resulting scatterplot (Fig. 5.14) of the DAPC analysis confirmed the inter-oceanic genetic differentiation among Atlantic, Indian and Pacific YFT populations. As shown in Fig. 5.14, the elbow in the curve matches the smallest Bayesian Information Criterion (BIC), and clearly indicates the presence of three genetic clusters.

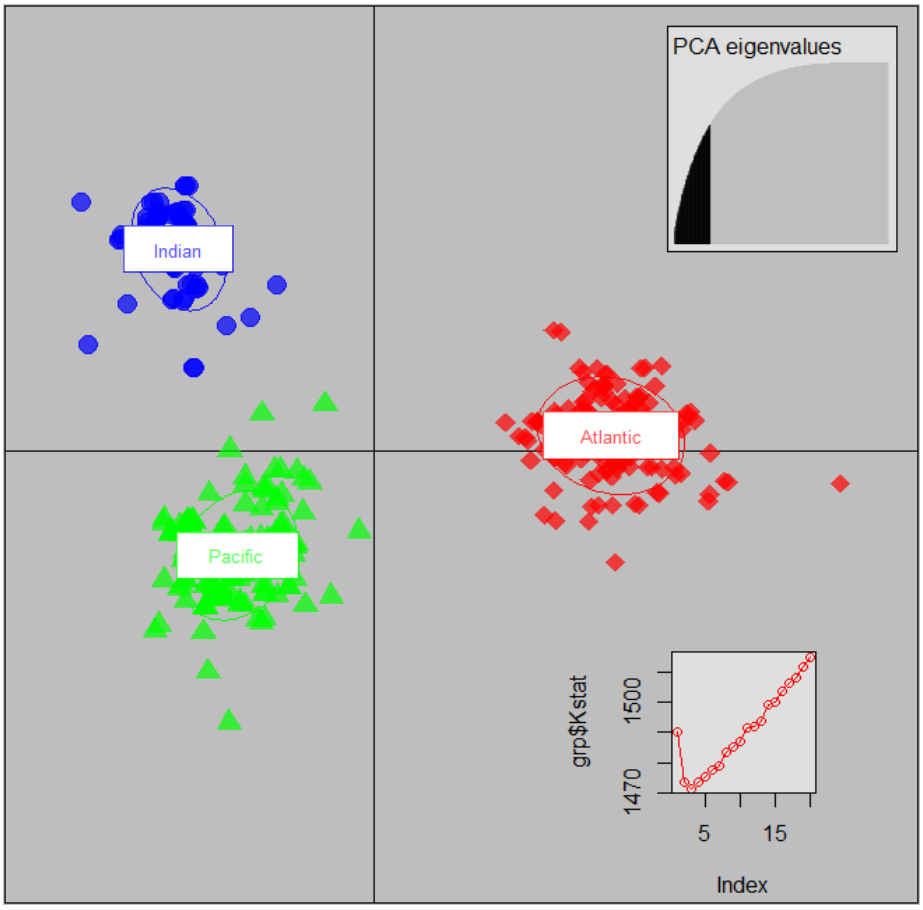


Fig. 5.14. Scatterplot of the DAPC results identifying three genetic clusters of *Thunnus albacares*. There are also graphically represented the number of PC retained during the data trasformation step and corresponding elbow of the BIC vs. number of clusters.

Group membership of the DAPC analysis, which provides a membership probability of each individual for the different populations based on the retained discriminant functions, were used as indicators for underlining how consistent are the genetic clusters detected. The resulting compoplot (in a STRUCTURE-like way) based on the full data set (Fig. 5.15),

indicated that >90% of individuals showed posterior membership probability of >0.95 to one of the three clusters, confirming the reliability of the membership of individuals and groups. However, it indicated that there are three individuals of the Cluster 1 (Pacific) assigned to the Cluster 3 (Indian), and 5 individuals of the Cluster 3 assigned to the Cluster 1. Instead all of the individuals were properly and without margin of error assigned to the Atlantic (Cluster 2).

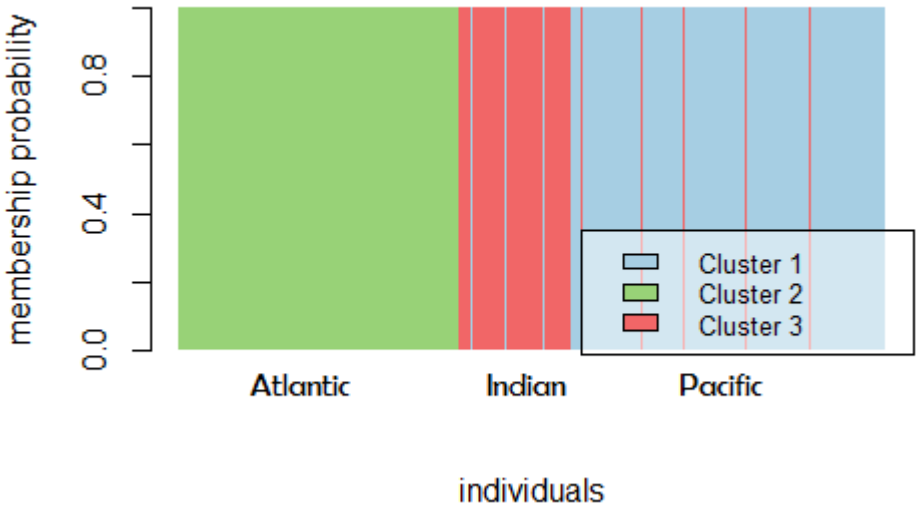


Fig. 5.15_Compoplot of the individuals' membership probability analysis. Cluster 1:3: Pacific, Atlantic, Indian.

According to the ΔK statistics, the presence of three clusters (K=3) was also confirmed by the Cluster analysis based on the Bayesian algorithm performed by using the software STRUCTURE (Fig. 5.16). Instead any sign of differentiation was detected within each of the three clusters.

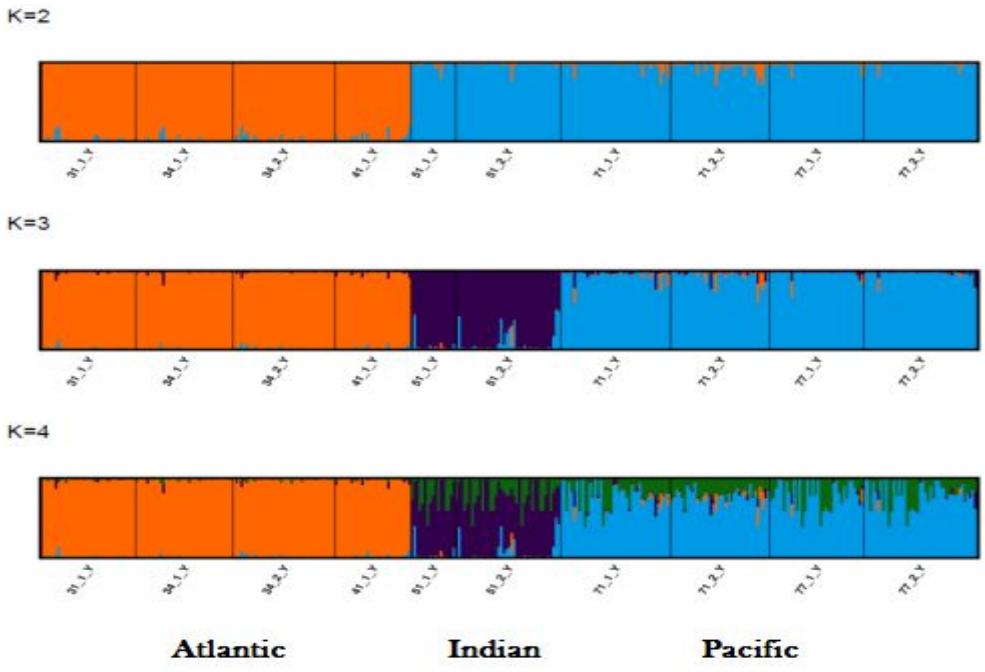


Fig.5.16_ The three oceanic clusters (Atlantic, Indian and Pacific) detected by Structure and visualized by CLUMPAK.

Using the maximum likelihood method implemented in GeneClass2, 79 simulated genotypes out of 95 were correctly assigned (85.3%). Any genetic cluster had a percentage of assignment with a certainty of 99%. The rate of correct assignment was of 87.5% for the Atlantic individuals with 2 and 3 individuals incorrectly assigned to the Indian and Pacific groups, respectively. For the individuals sampled in the Indian Ocean, the percentage of individuals correctly assigned was of the 85.72% with four individuals wrongly assigned to the other two Oceans, and finally the percentage of correct assignment was of 82.93% in the Pacific Ocean, in which the 17.07% of individuals was wrongly assigned to the Indian Ocean population (Fig. 5.17).

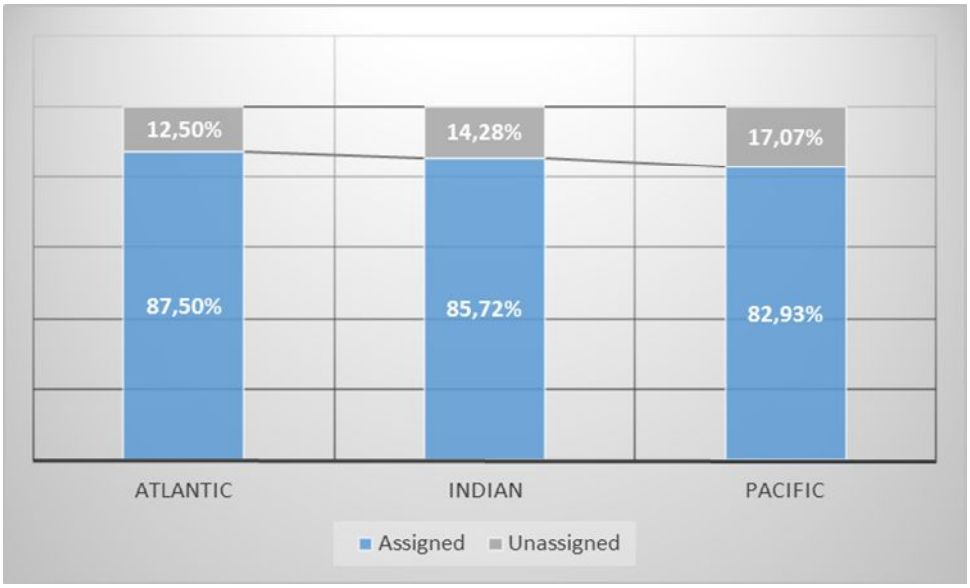


Fig. 5.17_ Percentage of correctly assigned and unassigned individuals to each oceanic populations

5.3.4 Outlier loci detection analysis

In all analyses, BAYESCAN detected considerably more outliers than ARLEQUIN (often more than twice as many), but ARLEQUIN outliers were almost exclusively a subset of BAYESCAN outliers. Here, we describe only results on outliers identified by both methods.

A total of 33 outlier loci were identified using the two different approaches (Table 5.4). All those loci have an alpha-value significantly higher than 0 and Q-values approximately of 0 in Bayescan and FST with a significant P value in Arlequin.

Table 5.4_Results of outlier detection with the software Arlequin and Bayescan.

Locus	Obs. Het. BP	Obs FST	FST P-value	1-FST quantile	Locus	prob	log10(PO)	qval	alpha	FST
734	0.61	0.54	0.00	0.00	**734**	1.0000	1000.0	0.00	3.11	0.35
486	0.56	0.50	0.00	0.00	**486**	1.0000	1000.0	0.00	3.23	0.37
800	0.55	0.49	0.00	0.00	**800**	1.0000	1000.0	0.00	3.45	0.40
584	0.12	0.41	0.00	0.00	**584**	1.0000	1000.0	0.00	2.32	0.23
826	0.50	0.37	0.00	0.00	**826**	1.0000	1000.0	0.00	1.86	0.17
373	0.57	0.35	0.00	0.00	**373**	1.0000	1000.0	0.00	2.14	0.21
352	0.56	0.32	0.01	0.01	**352**	1.0000	1000.0	0.00	2.02	0.19
582	0.09	0.31	0.00	0.00	**582**	1.0000	1000.0	0.00	2.03	0.20
68	0.56	0.28	0.02	0.02	**68**	1.0000	1000.0	0.00	1.89	0.18
263	0.49	0.28	0.00	0.00	**263**	1.0000	1000	0.00	3.29	0.17
755	0.43	0.28	0.01	0.01	**755**	1.0000	1000.0	0.00	2.04	0.20
306	0.52	0.26	0.02	0.02	**306**	1.0000	1000.0	0.00	2.00	0.19
972	0.41	0.26	0.01	0.01	**972**	1.0000	1000.0	0.00	1.78	0.16
55	0.52	0.26	0.02	0.02	**55**	1.0000	1000.0	0.00	1.79	0.16
428	0.54	0.25	0.03	0.03	**428**	1.0000	1000.0	0.00	1.89	0.18
246	0.32	0.24	0.02	0.02	**246**	1.0000	1000.0	0.00	2.11	0.20
268	0.29	0.23	0.03	0.03	**268**	1.0000	1000.0	0.00	1.96	0.18
854	0.40	0.23	0.03	0.03	**854**	1.0000	1000.0	0.00	1.67	0.15
329	0.49	0.22	0.03	0.03	**329**	1.0000	1000.0	0.00	1.85	0.17
440	0.48	0.22	0.03	0.03	**440**	1.0000	1000.0	0.00	1.96	0.19
240	0.44	0.22	0.03	0.03	**240**	1.0000	1000.0	0.00	1.90	0.18
146	0.37	0.21	0.04	0.05	**146**	1.0000	1000.0	0.00	1.89	0.18
219	0.21	0.17	0.04	0.05	**219**	1.0000	1000.0	0.00	1.83	0.17
601	0.15	0.17	0.02	0.02	**601**	1.0000	1000.0	0.00	1.97	0.19
141	0.09	0.15	0.00	0.00	**141**	1.0000	1000.0	0.00	2.00	0.19
202	0.08	0.14	0.00	0.00	**202**	0.9916	2.072	0.00	2.05	0.10
837	0.07	0.12	0.00	0.00	**837**	1.0000	1000.0	0.00	1.86	0.17
650	0.47	0.09	0.00	0.00	**650**	0.98876	1.9445	0.00	1.18	0.10
635	0.11	0.08	0.00	0.00	**635**	1.0000	1000.0	0.00	1.50	0.16
313	0.07	0.08	0.00	0.00	**313**	1.0000	1000.0	0.00	1.97	0.19
251	0.04	0.06	0.00	0.00	**251**	1.0000	1000.0	0.00	1.87	0.08
864	0.04	0.05	0.00	0.00	**864**	1.0000	1000.0	0.00	2.78	0.09
346	0.12	0.04	0.00	0.00	**346**	1.0000	1000.0	0.00	1.75	0.11

The pairwise F_{ST} coefficients using those loci potentially under selection for the three oceanic clusters were considerably higher than using the entire dataset. Besides using those outlier loci, significant F_{ST} were detected within the Atlantic and Pacific Oceans (ranging from 0,021 to 0,056) and specifically between the estern and astern part of each oceanic basin. The resulting scatterplot (Fig. 5.18) of the DAPC analysis confirmed the inter- and intra-oceanic genetic differentiation among Atlantic, Indian and Pacific YFT sub-populations.

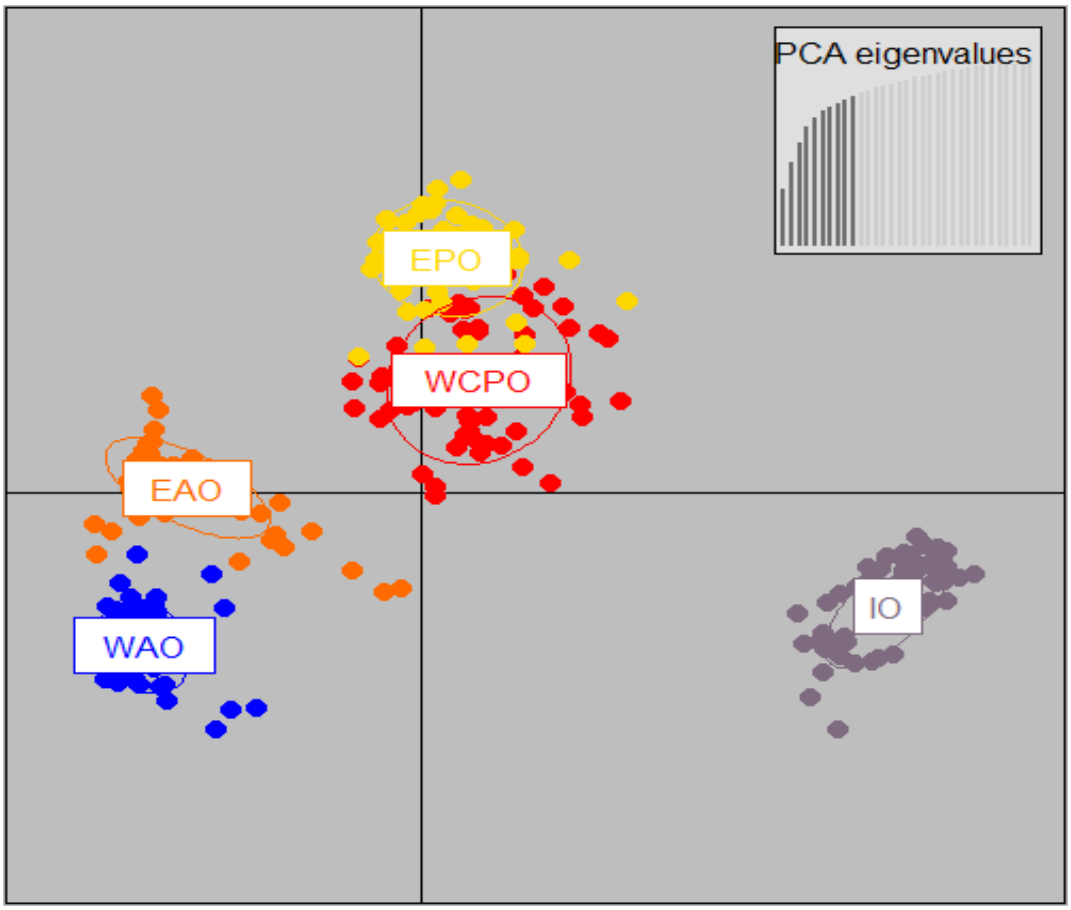


Fig. 5.18. Scatterplot of the DAPC results identifying five genetic clusters of *Thunnus albacares*: WAO (Western Atlantic Ocean), EAO (Eastern Atlantic Ocean), WCPO (Western Central Pacific Ocean), EPO (Eastern Pacific Ocean) and IO (Indian Ocean).

5.4 Discussion and Conclusions

Here we provide the most exhaustive and complete work on the global population structure of yellowfin tuna employing a new genomic approach (2b-RAD genotyping technique) for discovering and at the same time genotyping thousands of SNP markers across hundreds of individuals. Although this genotyping technique was proven as a simple and flexible method for genome-wide genotyping (Wang et al., 2012; Guo et al., 2014; Pauletto et al., 2015; Dou et al., 2016), however it has never been used in published marine fish population structure studies so far. The trustworthiness of this approach lies in the fact that many quality checks are required during the library preparation and sequence data processing, in order to exclude samples with low concentration and suboptimal quality profile and to minimize possible PCR artefacts. In our study,

the reliability of this approach was further demonstrated by the comparable number of reads obtained after sequencing per each pool. In fact, the mean number of reads obtained from each individual of each pool was roughly higher than 2 million, providing a high depth of coverage per individual over all SNPs. The lower number of reads obtained from the pool 1 comparing to those obtained from the other three pools could be linked to different issues, such as the starting tissue quality/quantity, the barcodes used and the lower sequencing efficiency for the first two lanes. However, the removal of all those individuals with less than 500,000 reads have guaranteed to perform the subsequent analyses only on those individuals with a similar and sufficient mean coverage. The efficiency of this genomics approach was also demonstrated by the FastQC results, which underlined a similar quality pattern for all of our samples. Checking the quality of the samples with FastQC is a crucial step for identifying possible problems (i.e. contaminations, overrepresented sequences and presence of adaptors), before going further with the analyses. Moreover, by using barcodes adaptors, it was possible to pool together 98 libraries and to make 2b-RAD library preparation simpler and faster than other GBS approaches, reducing also sequencing costs.

An important finding of this study was the possibility to map with a high percentage of success (86,59%) our short 2b-RAD reads against the genome of *Thunnus orientalis*, which represents to date the only available genome for the genus *Thunnus*. In fact, using a reference genome it is always advisable for unmistakably mapping 2b-RAD SNPs, enhancing considerably their chance of being mapped unambiguously in the genome (Puritz et al., 2014). In doing so, we also demystified one of the biggest disadvantages of this technique, which is the expected impossibility to be used for building genome contigs and to be cross-mappable across large genetic distances when a reference genome is missing (Puritz et al., 2014). Running Stacks with a reference genome increases the accuracy in data processing (Catchen et al., 2013), reducing for instance possible false-positive SNP calls. Genotype biases were also avoided identifying and removing those loci and individuals which went beyond some conservative filtering criteria (see Material and Methods for more details).

On one hand, in this study we have confirmed that the analysis of a large amount of molecular markers has the great advantage of increasing the resolution in detecting genetic differentiation among populations with large effective sizes and high migration rates (Waples, 1998). Our results, based on the entire dataset, revealed the existence of significant structuring among the three main oceanic basins, with the pairwise F_{ST} values calculated from genome-aligned markers ranging

from 0.036 to 0.136. The presence of these three populations was also confirmed and visualized by different multivariate and individual-based analyses (i.e. PCA, DAPC, STRUCTURE etc.), indicating limited gene flow and dispersal among the three oceans. These findings confirmed the suggestions on the YFT global genetic structure proposed by Ward et al., (1997), based on gene frequencies at a single polymorphic allozyme Locus *GPI-A**, which advised to reject the null hypothesis of a single panmictic population worldwide (Ely et al., 2005). In fact, the significant genetic differentiation and the high rate of self-assignment assessed here (>80%), have shed light on the existence of at least three putatively isolated populations among the Atlantic, the Indian and the Pacific Oceans.

On the other hand, we pointed out that within the Pacific and Atlantic Oceans just multiplying the number of neutral markers was not enough to reveal the proper level of structuring at the local scale (Gagnaire et al., 2015). In doing so, it was necessary to discovery a subset of outlier loci putatively under selection to delineate and separate locally adapted stocks within these two oceans. In fact, the effects of selection might influence much more the population structure than the genetic drift in species as YFT characterized by a large effective size and high dispersal capacity. This underlines how much useful are outlier SNPs for resolving fine-scale population connectivity, admixture, mixed stock analysis, and for identifying local adaptation for conservation and management purposes (Milano et al., 2014; Allendorf et al., 2010; Nielsen et al., 2009). Therefore, detecting outlier loci is essential to understand the YFT local adaptation's mechanisms as they may be directly linked to gene regions associated with ecotype divergence over very recent (ecological) timescales (Russello et al., 2011). In such a context, biogeography and barriers to dispersal can play a crucial role in isolating reproductively YFT populations between and within the three oceans. On a large scale, the complete closure of the Isthmus of Panamà (IOP) dated to ca. 3.5 Myr (Coates and Obando 1996; Coates et al., 1992) and the closure of the Tethys seaway (Terminal Tethyan Event, TTE) (12–18 Myr) represent the two main physical barriers between the Indo-Pacific and the Atlantic realms. In opposition, the Atlantic and Indo-Pacific populations are still in contact through the Cape of Good Hope and their migrations are facilitated by the unidirectional flow of warm waters to the Southeast Atlantic carried along the west coast of Africa by the Agulhas current (Shannon et al., 1990). However, this gene flow was maybe interrupted during the peak of the major glacial periods, in which waters were significantly lower than today enhancing their genetic divergence. Other soft barriers might also take on relevant importance in creating and maintaining the genetic differentiation between and within oceans (Cowman and

Bellwood, 2013). Specifically, the East Pacific Barrier (EPB) seems to affect the dispersal capacity of several Indian and the Western Pacific marine species from the Eastern Pacific population. In addition, two incomplete barriers in both sides of the Indo-Australian Archipelago (IAA), might intervene in differentiating genetically Indian and Pacific populations as well as Western and Eastern Pacific populations. In addition, an important soft barrier in the Atlantic Ocean is represented by the Mid-Atlantic Barrier (MAB; the stretch of tropical ocean between equatorial America and Africa), which is a deep-ocean barrier and the main geographic barrier between the two sides of the tropical Atlantic Ocean (Fig. 5.19). MAB was generated by the formation of the Atlantic Ocean basin as Africa and South America separated over the past 85 Myr (Luiz et al., 2012).

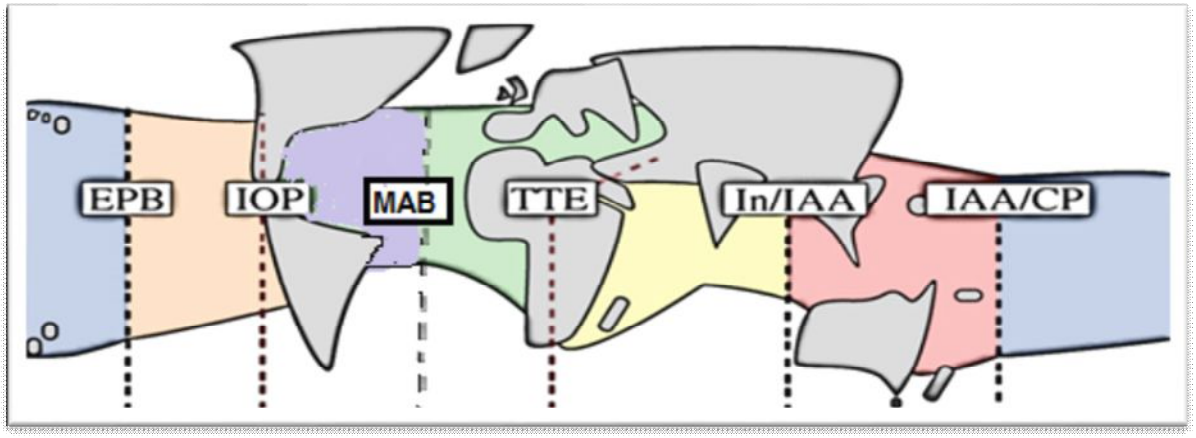


Figure 5.19. Schematic diagram of world map identifying boundaries between regions (modified from Cowman and Bellwood 2013).

However, the genetic separation detected in the present study can be linked to a wide variety of evolutionary and biological mechanisms apart from the local adaptation to specific environmental constraints, which of course interact among them to module and define the population structure of this biological and economic valuable species. The combination of those evolutionary and biological mechanisms seem to get the better of YFT high potential gene flow, making their populations reproductively isolated among the three oceans as well as between the western and eastern sides of the Atlantic and the Pacific basins. Instead the failure of finding differences within the Indian Ocean might be due to the limits of our sampling design, lacking of eastern samples. However, our results are in agreement with those obtained by the largest tagging program never realized in the Indian Ocean (IOTTP; 2005–2009), which indicated the possible presence of a single well-mixed YFT population in the entire basin. As underlined by the commission (IOTC–2015–

SC18–ES04) future studies should focus efforts to areas of admixture at a much finer-scale here reported, by employing multidisciplinary approaches, including population genomic structure assessment together with other population identification methodologies.

Our results are partially in contrast with those obtained by Grewe et al., (2015), which also detected two distinct populations within the Western Pacific Ocean (Coral Sea and Tokelau). This different result is mainly due to the geographic proximity of our population samples, which came from areas identified here as admixture zones within the eastern and western parts of each ocean. The genetic differentiation observed in this study, as already indicated in other pelagic species (i.e. swordfish) might also be linked to a strong natal fidelity, with admixture primarily confined to feeding grounds (Alvarado-Bremer et al., 2005). This homing behavior associated with many biotic and abiotic factors, such as geographic features, oceanographic processes, prey concentration, might also play an important role in shaping YFT population structure. Homing migration patterns in YFT have been also confirmed by some tagging studies. For instance, restricted movements and a high degree of fidelity were also indicated within the Central and Western Pacific (Itano and Holland, 2000; Sibert and Hampton, 2003), and within the Eastern Pacific (Schaefer et al., 2007; 2011) excluding any assumption about complete mixing over large regional areas (Schaefer et al., 2011). Furthermore, in the Atlantic it was observed that YFT juveniles move along African coastal waters in the Gulf of Guinea, which is their main spawning ground, until they reach the pre-adult stage. After that, they start swimming from the eastern to the western part of the Atlantic Ocean for feeding to return later to the eastern part (i.e. Gulf of Guinea) for spawning, when they reach 110 cm F_L and about 3 years of age (Zagaglia et al., 2004). However, there are still inadequate information about the migration patterns from the East to the West Atlantic for spawning reasons, that can speak in favor of two reproductively isolated populations in the Gulf of Guinea and Gulf of Mexico/southeastern Caribbean, respectively (Arocha, 2001). This east-west population division in the Atlantic Ocean, as pointed out by our results, supports a possible higher level of structuring than generally assumed. Those possible sub-populations should be taken into account by ICCAT, re-examining previous biological assumptions, in order to better understand YFT populations' response to fishing, which can have different productivities and levels of fishing mortality experienced in the Gulf of Guinea and in the Gulf of Mexico/Caribbean Sea.

.

References

Allendorf F.W., Hohenlohe P.A. and Luikart G. Genomics and the future of conservation genetics. Nat. Rev. Genet. 11, (2010): 697–709.

Alvarado Bremer J.R., Naseri I. and Ely, B. Orthodox and unorthodox phylogenetic relationships among tunas revealed by the nucleotide sequence analysis of the mitochondrial DNA control region. J. Fish Biol. 50, (1997): 540–554.

Alvarado Bremer J.R., Viñas J., Mejuto J., Ely B. and Pla C. Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. Mol. Phylogenet. Evol. 36, (2005): 169–187.

Arocha F., Lee D.W., Marcano L.A. and Marcano J.S. Update information on the spawning of yellowfin tuna, Thunnus albacares, in the western central Atlantic. Col Vol Sci Pap ICCAT 52, (2001): 167–176.

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A. and Johnson E.A. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS One 3, (2008).

Bernardi G., Wiley E.O., Mansour H., Miller M.R., Orti G., Haussler D., O'Brien S.J., Ryder O.A. and Venkatesh B. The fishes of Genome 10K. Mar. Genomics 7, (2012): 3–6.

Chutimanitsakun Y., Nipper R.W., Cuesta-Marcos A., Cistué L., Corey A., Filichkina T., Johnson E.A. and Hayes P.M. Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. BMC Genomics 12, (2011): 4.

Coates A.G. and Obando J.A. The geologic evolution of the Central American Isthmus. Evol. Environ. Trop. Am. (1996): 21–56.

Coates AG, et al. Closure of the Isthmus of Panama: The near-shore marine record of Costa Rica and western Panama.Geol Soc Am Bull 104(7), (1992): 814–828.

Cowman P.F. and Bellwood D.R. Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. Proc. R. Soc. Lond. B Biol. Sci. 280, (2013): 20131541.

Davey J.W. and Blaxter M.L. RADSeq: next-generation population genetics. Brief. Funct. Genomics 9, (2010): 416–423.

Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M. and Blaxter M.L. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. 12, (2011): 499–510.

Dou J., Li X., Fu Q., Jiao W., Li Y., Li T., Wang Y., Hu X., Wang S. and Bao Z. Evaluation of the 2b-RAD method for genomic selection in scallop breeding. Sci. Rep. 6, (2016).

Ely B., Viñas J., Bremer J.R.A., Black D., Lucas L., Covello K., Labrie A.V. and Thelen, E. 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, (2005): 19.

Emerson K.J., Merz C.R., Catchen J.M., Hohenlohe P.A., Cresko W.A., Bradshaw W.E. and Holzapfel C.M. Resolving postglacial phylogeography using high-throughput sequencing. Proc. Natl. Acad. Sci. 107, (2010): 16196–16200.

Excoffier L., Hofer T. and Foll M. Detecting loci under selection in a hierarchically structured population. Heredity 103, (2009): 285–298.

Excoffier L. and Lischer H.E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, (2010): 564–567.

Foll, M. Bayescan v2. 1 user manual. Ecology 20, (2012): 1450–1462.

Gagnaire P.-A., Broquet T., Aurelle D., Viard F., Souissi A., Bonhomme F., Arnaud-Haond S. and Bierne N. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. Evol. Appl. 8, (2015): 769–786.

Graves J.E. and McDowell J.R. Stock structure of the world's istiophorid billfishes: a genetic perspective. Mar. Freshw. Res. 54, (2013): 287–298.

Grewe P.M., Feutry P., Hill P.L., Gunasekera R.M., Schaefer K.M., Itano D.G., Fuller D.W., Foster S.D., Davies C.R. Evidence of discrete yellowfin tuna (Thunnus albacares) populations demands rethink of management for this globally important resource. Sci. Rep. 5 (2015).

Guo Y., Yuan H., Fang D., Song L., Liu Y., Liu Y., Wu L., Yu J., Li Z., Xu X. and others. An improved 2b-RAD approach (I2b-RAD) offering genotyping tested by a rice (Oryza sativa L.) F2 population. BMC Genomics 15, (2014): 1.

Hemmer-Hansen J., Therkildsen N.O., Meldrup D. and Nielsen E.E. Conserving marine biodiversity: insights from life-history trait candidate genes in Atlantic cod (Gadus morhua). Conserv. Genet. 15, (2014): 213–228.

Hohenlohe P.A., Bassham S., Etter P.D., Stiffler N., Johnson E.A., Cresko W.A. et al. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet 6, (2010): e1000862.

Itano D.G. and Holland K.N. Movement and vulnerability of bigeye (Thunnus obesus) and yellowfin tuna (Thunnus albacares) in relation to FADs and natural aggregation points. Aquat. Living Resour. 13, (2000): 213–223.

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

Jombart T., Devillard S., Dufour A.-B. and Pontier D. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101, (2008): 92–103.

Kopelman N.M., Mayzel J., Jakobsson M., Rosenberg N.A. and Mayrose I. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Resour, (2015).

Luikart G., England P.R., Tallmon D., Jordan S. and Taberlet P. The power and promise of population genomics: from genotyping to genome typing. Nat. Rev. Genet. 4, (2003): 981–994.

Luiz O.J., Madin J.S., Robertson D.R., Rocha L.A., Wirtz P. and Floeter S.R. Ecological traits influencing range expansion across large oceanic dispersal barriers: insights from tropical Atlantic reef fishes. Proc. R. Soc. Lond. B Biol. Sci. 279, (2012): 1033–1040.

Milano I., Babbucci M., Cariani A., Atanassova M., Bekkevold D., Carvalho G.R., Espiñeira M., Fiorentino F., Garofalo G., Geffen A.J. et al. Outlier SNP markers reveal fine-scale genetic

structuring across European hake populations (Merluccius merluccius). Mol. Ecol. 23, (2014): 118–135.

Miller M.R., Dunham J.P., Amores A., Cresko W.A. and Johnson E.A. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res. 17, (2007): 240–248.

Moore J.-S., Bourret V., Dionne M., Bradbury I., O'Reilly P., Kent M., Chaput G. and Bernatchez L. Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. Mol. Ecol. 23, (2014): 5680–5697.

Nakamura Y., Mori K., Saitoh K., Oshima K., Mekuchi M., Sugaya T., Shigenobu Y., Ojima N., Muta S., Fujiwara A. et al. Evolutionary changes of multiple visual pigment genes in the complete genome of Pacific bluefin tuna. Proc. Natl. Acad. Sci. 110, (2013): 11061–11066.

Narum S.R., Buerkle C.A., Davey J.W., Miller M.R. and Hohenlohe P.A. Genotyping-by-sequencing in ecological and conservation genomics. Mol. Ecol. 22, (2013): 2841–2847.

Nielsen E.E., Hemmer-Hansen J., Poulsen N.A., Loeschcke V., Moen T., Johansen T., Mittelholzer C., Taranger G.-L., Ogden R. and Carvalho G.R. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (Gadus morhua). BMC Evol. Biol. 9, (2009): 276.

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

Palsbøll P.J., Berube M. and Allendorf F.W. Identification of management units using population genetic data. Trends Ecol. Evol. 22, (2007): 11–16.

Pauletto M., Carraro L., Babbucci M., Lucchini R., Bargelloni L. and Cardazzo B. Extending RAD tag analysis to microbial ecology: a comparison between MultiLocus Sequence Typing and 2b-RAD to investigate Listeria monocytogenes genetic structure. Mol. Ecol. Resour.(2015).

Pecoraro C., Babbucci M., Villamor A., Franch R., Papetti C., Leroy B., Ortega-Garcia S. et al., "Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (*Thunnus albacares*)." Marine Genomics 25, (2016)., 43–48. DOI:10.1016/j.margen.2015.12.002

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS One 7, (2012): e37135.

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

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

Pujolar J.M., Jacobsen M.W., Frydenberg J., Als T.D., Larsen P.F., Maes G.E., Zane L., Jian J.B., Cheng L. and Hansen M.M. A resource of genome-wide single-nucleotide polymorphisms generated by RAD tag sequencing in the critically endangered European eel. Mol. Ecol. Resour. 13, (2013): 706–714.

Puritz J.B., Matz M.V., Toonen R.J., Weber J.N., Bolnick D.I. and Bird C.E. Demystifying the RAD fad. Mol. Ecol. 23, (2014): 5937–5942.

Rannala B. and Mountain J.L. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. 94, (1997): 9197–9201.

Raymond M. and Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Hered. 86, (1995): 248–249.

Roesti M., Hendry A.P., Salzburger W. and Berner D. Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. Mol. Ecol. 21, (2012): 2852–2862.

Russello M.A., Kirk S.L., Frazer K.K. and Askey P.J. Detection of outlier loci and their utility for fisheries management. Evol. Appl. 5, (2012): 39–52.

Schaefer K.M., Fuller D.W., Block B.A. Movements, behavior, and habitat utilization of yellowfin tuna (Thunnus albacares) in the Pacific Ocean off Baja California, Mexico, determined from archival tag data analyses, including unscented Kalman filtering. Fish. Res. 112, (2011): 22–37.

Schaefer K.M., Fuller D.W. and Block, B.A. Movements, behavior, and habitat utilization of yellowfin tuna (Thunnus albacares) in the northeastern Pacific Ocean, ascertained through archival tag data. Mar. Biol. 152, (2007): 503–525.

Shannon L.V., Agenbag J.J., Walker N.D. and Lutjeharms J.R.E. A major perturbation in the Agulhas retroflection area in 1986. Deep Sea Res. Part Oceanogr. Res. Pap. 37, (1990): 493–512.

Sibert, J. and Hampton, J. Mobility of tropical tunas and the implications for fisheries management. Mar. Policy 27, (2003): 87–95.

Toonen R.J., Puritz J.B., Forsman Z.H., Whitney J.L., Fernandez-Silva I., Andrews K.R. and Bird, C.E. ezRAD: a simplified method for genomic genotyping in non-model organisms. PeerJ 1, (2013): e203.

Wang S., Meyer E., McKay J.K., Matz M.V. 2b-RAD: a simple and flexible method for genome-wide genotyping. Nat. Methods 9, (2012): 808–810.

Wang X.Q., Zhao L., Eaton D.A.R., Li D.Z. and Guo Z.H. Identification of SNP markers for inferring phylogeny in temperate bamboos (Poaceae: Bambusoideae) using RAD sequencing. Mol. Ecol. Resour. 13, (2013): 938–945.

Waples R.S. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J. Hered. 89, (1998): 438–450.

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

Wit P.and Palumbi S.R. Transcriptome-wide polymorphisms of red abalone (Haliotis rufescens) reveal patterns of gene flow and local adaptation. Mol. Ecol. 22, (2013): 2884–2897.

Zagaglia C.R., Lorenzzetti J.A. and Stech J.L. Remote sensing data and longline catches of yellowfin tuna (Thunnus albacares) in the equatorial Atlantic. Remote Sens. Environ. 93, (2004): 267–281.

6 Chapter 5

Maternal effect

6.1 Introduction

Understanding the productivity and resilience for fish stocks, which define the level of fishing mortality (F) that they can tolerate and their ability to recover from a depleted state, is a crucial point to provide sound scientific advice for fisheries management (Morgan et al., 2009). However, the estimation of the productivity, which is based on the stock–recruit (S–R) relationship, is still a difficult challenge in the study of marine fish stock's dynamics and management. This relationship is traditionally measured by the estimation of the spawning stock biomass (SSB) that is used as a proxy for the stock reproductive potential (SRP, Trippel, 1999; Tomkiewicz et al., 2003; Lowerre-Barbieri et al., 2011). SBB implies that survival rates of offspring are independent from parental age, body size or condition (Cardinale and Arrhenius, 2000), and that total relative egg production per unit weight of adult stock is invariant over time (Morgan et al., 2009). Thus, SSB does not take into account a variety of fundamental attributes, such as the onset of maturity, fecundity, atresia, duration of reproductive season, daily spawning behaviour and spawning fraction (Murua et al., 2003). Accurate knowledge of those reproductive characteristics that have a direct influence on the productivity and resilience of commercial, threatened and endangered fish species, such as tropical tunas, are fundamental components for developing effective and realistic fishery management and conservation strategies (Trippel, 1999; Morgan et al., 2009; Brown-Peterson et al., 2011). In such a context, many doubts have been raised about the appropriateness of SSB (Marshall et al., 1998; 2009), which therefore assumes that fecundity is only related to the mass-at-age of the sexually mature portion of the stock irrespective of the demographic composition of adults (Murawski et al., 2001; Kell et al., 2015). The resulting supposition is that many small individuals will produce as many offspring as a few large individuals. On the contrary, there is an increasing consensus in fishery science that spawning stocks are composed by individuals with a different range of sizes and ages that may contribute differently to spawning and recruitment (Marshall et al., 1998; Scott et al., 1999¸ Kell et al., 2015). In addition, there is a general acknowledgement on the relevant impacts that maternal effects, which correspond to an increase in allocating reproductive resources for postnatal use with the females' size (Berkeley et al., 2004), might have on fecundity and viability of eggs and larvae (Kjesbu et al., 1998; Scott et al., 1999; Trippel, 1999; Berkeley et al., 2004). The positive relationship between the mothers' size/age and both the potential productivity and the survival rates of the recruits have been already

demonstrated for other fish species (Marshall et al., 1998; 1999; Cardinale and Arrhenius, 2000; Shelton et al., 2015; Berkeley et al., 2004; Bobko and Berkeley, 2004; Riveiro et al., 2004). Moreover, older and larger females, having a wider spatial and temporal window for spawning than smaller individuals, enhance the perspectives for their larvae to encounter advantageous conditions to survive (Birkeland and Dayton, 2005). In doing so, larger females should invest a higher amount of energy for reproduction than smaller fish and this energy is mainly provided by the metabolization of lipids and their constituent fatty acids, which represent the main energetic resources in fishes (Tocher, 2003). Fatty acids in fishes are the favorite source of metabolic energy in the form of ATP via mitochondrial *β*-oxidation (Sargent et al*.,* 1995). Lipids can be divided into two main groups according to their chemical properties and functions: 1) neutral lipids (NLs); and 2) polar lipids (PLs). PLs mainly correspond to the lipid class of phospholipids and to a lesser extend to the ketones and wax-esters. PLs are important constituents of membranes and they have an important role as precursors in eicosanoid metabolism (structural fat). Instead NLs serve primarily as a depot of lipids (depot fat) mainly used as an energy source (Henderson and Tocher, 1987); they regrouped the lipid classes of triacylglycerols and sterols.

Fatty acids, and particularly polyunsaturated fatty acids (PUFAs), are functionally essential for the reproduction in fish, influencing egg quality, spawning, hatching and survival of larvae (Sargent et al., 1989; 2002). PUFAs also intervene in regulating the production of eicosanoids (prostaglandins), steroid hormones and gonad development (Izquierdo et al., 2001). Hence, fatty acid composition in both NLs and PLs must be investigated for understanding fishes' energetic investment for reproduction. The acquisition of lipids and their fatty acids by fishes is driven by the two main strategies of resource use for reproduction described so far (Alonso-Fernández and Sabórido-Rey, 2012; Aristizabal, 2007): 1) capital breeders, who store the energy required before the onset of reproductive period; 2) income breeders, who acquire the energy required by feeding during reproduction. In the latter group, the fatty acid composition of the female gonad is greatly affected by the dietary fatty acid content, which, in turn, directly influences the egg quality in a short period of time (Izquierdo et al., 2001). However, the separation between these two strategies is not clear and there are shades of grey between them. An illustrative case is represented by yellowfin tuna (*Thunnus albacares;* YFT) that is described as an income-capital breeder (Zudaire et al*.*, 2013b). YFT takes on the cost of reproduction both by feeding during the spawning period and using some energetic resources acquired before that. According to the operational difficulties (e.g. the sampling) in investigating these complex effects in wild and large

pelagic species, to date very few studies have been conducted to understand how the size of YFT females affects the fecundity (i.e. Zudaire et al., 2014), which is linked to many biotic and abiotic cues. It is also very important to understand if the development of a possible phenotype is just driven by the environment that the organism faces ("phenotypic plasticity") or if there are genetic components that can play an important role in regulating these reproductive patterns as well as in the way in which YFT acquire energy for meeting the high energetic demands (McBride et al., 2013). Understanding the genetic basis of those physiological and metabolic activities would provide new insights into the YFT reproductive success in relation to female's size/age. Within this scenario, Genome-Wide Association Study (GWAS) permits to screen a very large number of genes simultaneously for understanding the genetic contribution to possible variations of the reproductive outputs and of the energy allocation strategy related to the females 'size. In fact, in GWAS a group of individuals are phenotyped and then genotyped using a large number of single nucleotide polymorphisms (SNPs) in order to detect statistical association between genetic markers and the phenotypic traits of interest (Bush and Moore, 2012). GWAS explores the association of each SNP marker with the phenotypic trait of interest, making possible the identification of specific alleles associated with the trait. However, due to the large number of genetic markers, certain multiple testing corrections are required to reduce the increased probability of false positive findings.

The investigation of the potential existence of maternal effects in YFT can help in better understanding the consequences of age truncation induced by removing individuals fish via fishing (i.e. industrial tuna longlining, sport fishing). Such studies can lead to taking in consideration possible length-based fishery regulations by tRFMOs in order to protect YFT stocks from overfishing and to meet ecological and social objectives. In fact, fisheries removing larger individuals can have a direct impact on the YFT productivity by: (i) reducing the time and changing spawning locations (Bobko and Berkeley 2004, Berkeley et al., 2004, Poisson et al., 2009); and (ii) decreasing the production and quality of eggs (Hutchings and Myers, 1993; Marteinsdottir and Steinarsson, 1998; Vallin and Nissling 2000; Berkeley et al., 2004; Birkeland and Dayton, 2005; Raventos and Planes, 2008).

In this chapter, I investigated how the reproductive patterns and the spawning quality of the Atlantic YFT females are influenced by the fish size and thus by the possible variation in their energy allocation strategy. Additionally, I examined whether there is any genetic variation related to those "phenotypic" traits, using a GWAS approach. Given the fundamental importance to

evaluate the ubiquity of the relationship between fish size and offspring quality, this study can have important implications for verifying the potential for increased females' productivity in the Atlantic Ocean, where YFT is still considered to consist of only one panmictic population for management purposes (Arocha, 2001; see Chapter 4 for details).

6.2 Materials and Methods

6.2.1 Fish sampling

YFT were caught by purse-seine vessels in the Gulf of Guinea (Eastern Atlantic Ocean) from April 2013 to January 2014 and sampled during their process at the cannery "Pêche et Froid" of Abidjan, Ivory Coast (Fig. 6.1).

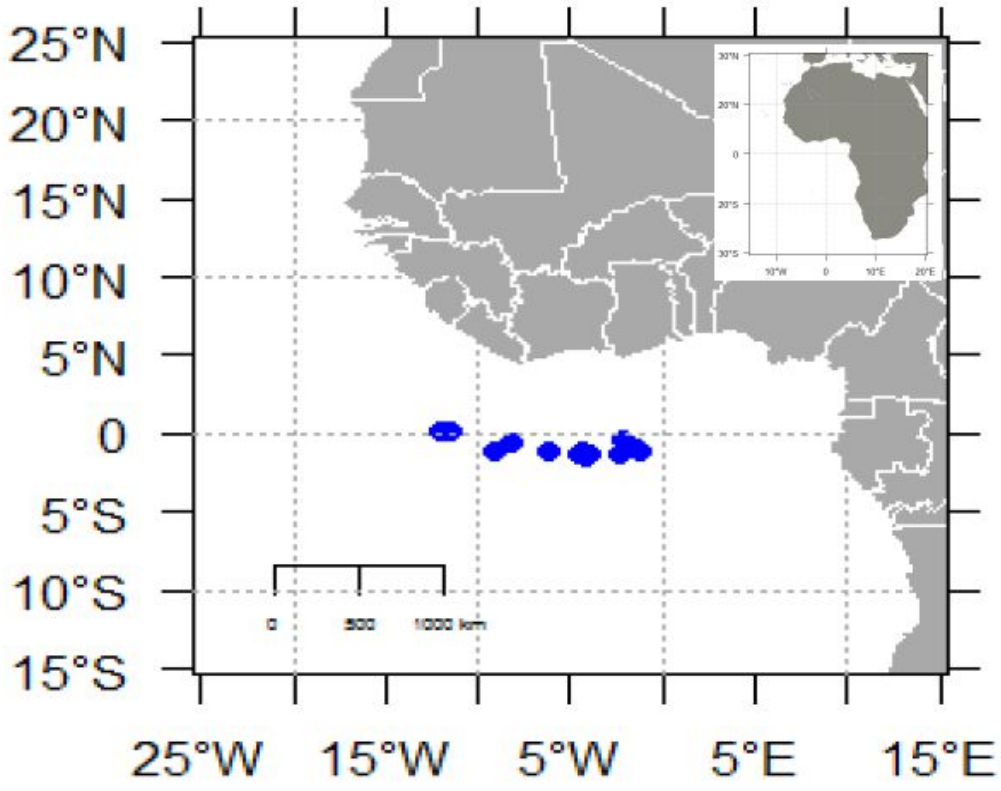


Fig. 6.1_Geographic origin of the yellowfin females sampled. Each dot indicates a fishing set.

From each fish the fork length (F_L; cm), the total fish weight (W; kg) and the gonad weight (GW ;g) were recorded, and a macroscopic maturity stage was assigned following the maturity reference

scale for yellowfin tuna (see Appendix 6.1). For the purpose of the present study, only spawning capable females (i.e., with late maturing or ripe ovaries) off the total number of females YFT, ranging in size from 125.8 to 154.5 cm FL (Fig. 6.2).

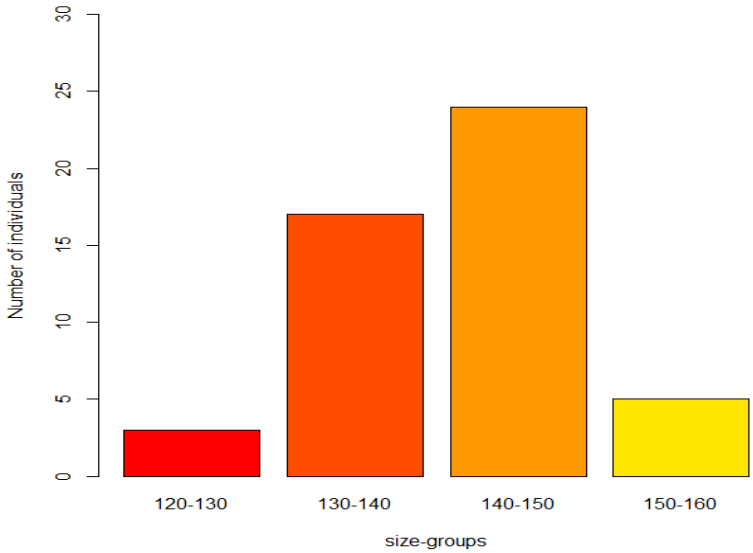


Fig. 6.2_Number of YFT females collected at the cannery and caught by purse-seiner in the Gulf of Guinea (Eastern Atlantic Ocean) from April 2013 to January 2014. Females are gathered by 10 cm fork length (F_L) size-group.

For the 50 selected females, a cross section of the gonad of 4-5 cm was cut between the middle and end part of the right or left lobe and preserved in 4% buffered formaldehyde for reproductive analysis. In addition, a 2g sample of gonads was collected and stored frozen in a labeled Eppendorf for fatty acid analysis, and a 1g sample of white muscle was cut and placed in a labeled Eppendorf tube filled with 96% ethanol (with a tissue/ethanol volume ratio of 1:10) for genomics analysis.

6.2.2 Reproductive analysis

6.2.2.1 Histological analysis

A cross section of ca. 1 cm of thickness was collected from the formaldehyde stored tissue and dehydrated in graded series of ethanol of increasing concentrations as described in Diaha et al., (2015). These dehydrated portions of the ovaries were embedded in paraffin and cut into sections of 5-6 µm in thickness with a microtome (MICROM, HM 350 S). The sections were stretched out in a bath with warm water (40°C) and subsequently gathered with glass slides properly labeled. After

24h the slides were stained with hematoxylin and eosin, using an automated Slide Stainer (Thermo Scientific™ Varistain™ 24-4). Once stained, the slides were mounted and then examined under a light microscope to be classified histologically following the criteria of Wallace and Selman (1981) and modified by Schaefer (1996; 1998). The histological analysis was performed at Azti tecnalia (Sukarrieta; Spain). Being characterized by an asynchronous ovarian development, each ovary was classified according to the most advanced oocyte stage present in the ovary (Murua and Motos, 2006) applying the terminology proposed by Brown-Peterson et al., (2011), and established for YFT in Zudaire et al., (2013a): (i) immature phase (primary growth stage [PG]); (ii) developing phase (cortical alveolar [CA], primary vitellogenesis [Vtg1], and secondary vitellogenesis [Vtg2] stages); (iii) spawning-capable phase (tertiary vitellogenesis [Vtg3], germinal vesicle migration [GVM], and hydration stages[HYD]), and (iv) regenerating phase. For the purposes of this work, only the ovary in spawning-capable phase, containing oocytes in the stages Vtg3, GVM and HYD (i.e. the most advanced oocyte development stages) were selected (Fig. 6.3).

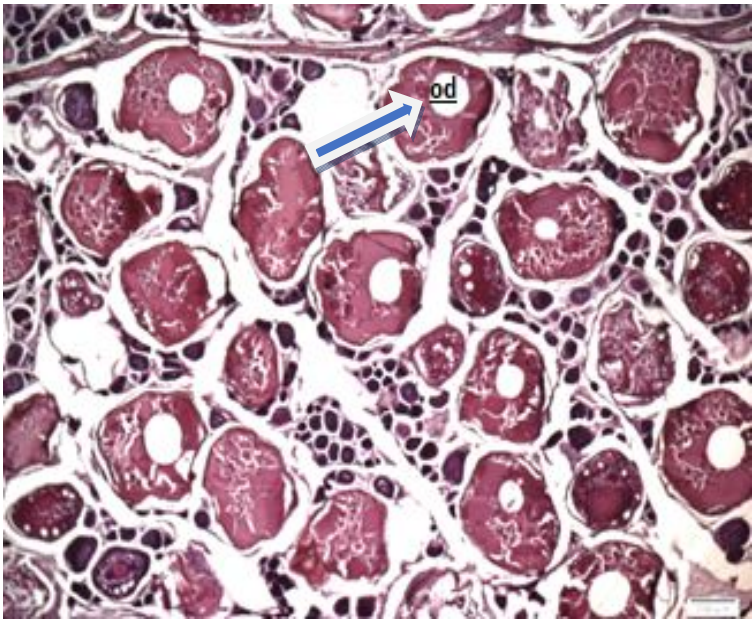


Fig. 6.3_YFT ovary in actively spawning phase with the oocytes in the hydration stage. Oil droplets (od) are clearly evident and marked. Bar represents 50 µm

Besides, atresia was not considered due to the brine conservation process used in the purse seine fleet that resulted in the break of the follicle wall and chorion, making not possible the accurate quantification of alpha-atresia (Zudaire et al., 2013a).

6.2.2.2 Oocyte size-frequency distribution

A portion of the preserved ovary of 0.04 g (±0.0001 g) was collected and analysed for oocyte size-frequency distribution at AZTI Technalia (San Sebastian, Spain). The sub-sample was placed into a filter with a mesh size of 125 µm and sprayed with high pressure water in order to separate the oocytes from the connective tissue. The separated oocytes were located in a gridded plate, photographed with a digital camera attached to the stereoscope, and analyzed with ImageJ free software (Rasband, W.S., ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA, http://rbs.info.nih.gov/ij/,1997-2012) to count and measure automatically all the oocytes of each sub-sample. The number of developing oocytes (NDO) was calculated between the minimum threshold of CA oocyte size (estimated at 120 µm; Zudaire et al., 2013b) and the Vtg3 stage. The separation between the oocytes in a developing stage from those belonging to the current batch (i.e. GVM or hydrated oocytes) was performed for each individual. The diameter of the developing oocytes and that of the oocytes forming the current batch were also calculated.

6.2.2.3 Batch fecundity estimation

The gravimetric method (Hunter et al., 1989) was used to estimate the batch fecundity (BF), which is the number of oocytes spawned per batch, for the selected ovaries at actively spawning phase without presence of new post-ovulatory follicles. This method consists of counting the total number of oocytes in the most advanced maturation stage (i.e. GVM or hydrated oocytes). For this purpose, three subsamples of 0.1 g (±0.01) were collected from each sample of formalin-preserved ovary. Each sub-sample was placed on a slide and covered with 3-4 drops of glycerin (Schaefer, 1987) to make them translucent, simplifying their selection from the non-hydrated oocytes, that are relatively opaque, in order to be counted under the stereomicroscope. After having counted the total number of oocytes in the most advanced maturation stage, the BF was estimated as the weighted mean density of the three subsamples multiplied by the total weight of the ovary, applying a threshold of 10% as coefficient of variance (CV) among them. However, when the CV went beyond this threshold, more subsamples were taken in order to reach this value of CV. To estimate the relative batch fecundity (BFrel), the value of BF calculated was divided by the gonad-free weight of the fish. Analyses were performed at AZTI Technalia (San Sebastian, Spain).

6.2.3 Genomics analysis

After having extracted the genomic DNA from each female (see Section.1), 2b-RAD libraries were prepared following the protocol from Wang et al., (2012) (see Chapter 3 and 4 for details). Two important modifications were carried out, with regard to the protocol used for the population structure study, in order to increase the number of reads obtained per each library: 1) the use of the restriction endonuclease AlfI (Thermo Fisher Scientific, Waltham, Massachusetts, USA); and 2) the use of both library-specific adaptors with fully degenerate cohesive ends (5'-NN-3'). In fact, AlfI, which recognizes palindromic sequences, has more restriction recognition sites than CspCI, thus increasing the number of available sites for the two non-selective adaptors. Purified libraries were pooled together and sequenced on an Illumina HiSeq2500 platform, with a 50 bp single-read module at the Genomix4Life S.r.l. facilities (Baronissi, Salerno, Italy), dedicating two lanes to the pool. The sequencing provider performed also the data demultiplexing. Quality and adapters trimming of the sequence reads was performed by running a custom-made Perl script, obtaining 34-bp fragments ready to be evaluated for SNPs presence in Stacks (Catchen et al., 2013). After having mapped the reads against the *Thunnus orientalis'* genome (see details in Chapter 3), individual genotypes were constructed using the *ref_map.pl* Stacks pipeline by setting the following parameters: *-m*= 8, *-n* = 2 and a bounded *SNP calling model* with an upper bound of 0.1 (all remaining *Stacks* settings as default). From the resulting catalog of loci, using the Stacks module *populations*, we only selected those that have just one bi-allelic SNP. The resulting SNP dataset (in a Genepop formatted file) was imported into the Rpackage Adegenet 2.0.1 (Jombart, 2008, R version 3.1.2, R Development Core Team, 2014; http://www.r-project.org). Prior to association analyses, SNP markers were filtered for minor allele frequencies (MAF) of > 0.5, and call rate >0.9 (both at the individual and SNP level). Markers were also tested for deviations from Hardy Weinberg Equilibrium (HWE) expectations. Significance for HWE deviations was determined at the $\alpha = 0.05$ level adjusted for multiple testing using a Bonferroni correction (Rice, 1989). The library preparation was performed at the University of Padova (Italy) while the bioinformatics analysis was performed at the University of Bologna.

6.2.4 Fatty acids analysis

Fatty acids analysis was performed at INRA, UMR TOXALIM (Research Centre in Food Toxicology), in Toulouse (France) according to Bodin et al., (2014) and Sardenne et al., (2016). Firstly, each 2-g sample of gonads was subjected to cryogenic grinding by using a mixer mill, MM400 Retsch® (Verder, France), obtaining a homogenized powder. From this, a subsample (0.1±0.001g) was weighed under a nitrogen atmosphere and extracted following a Folch method (Folch et al., 1957). An aliquot of the sample extracted was additional separated by adsorption chromatography on a silica gel micro-column (Kieselgel 70 to 230 mesh, heated at 450°C and deactivated with 6% water). Neutral and Polar lipids were eluted with 10 ml chloroform: methanol mixture (98:2 v/v) and with 20 ml of methanol, respectively. After adding a known amount of C23:0 fatty acid as internal standard, each fraction was transmethylated at 100°C with 10 wt% boron trifluoride-methanol (Metcalfe and Schmitz, 1961). The fatty acid methyl esters were analyzed on a TRACE 1310 gas chromatograph equipped with an on-column injector and a flame-ionization detector (GC-FID, Thermo Scientific). Compounds were separated on a FAMEWAX^{TM} column (30 m, 0.32 mm internal diameter, Restek) using helium as carrier gas at a constant flow of 15 ml/min. The injector temperature was set at 225°C and the oven temperature was raised from 130°C to 245°C at 2°C/min after a stationary phase at 130°C for 1 min. Helium was used as the carrier gas at a constant flow of 15 ml/min. Peaks were identified by comparing sample retention times to those of commercial standard mixtures (Menhaden oil and Food Industry FAME Mix, Restek) with Xcalibur 2.2 software. Results were expressed in % as a relative abundance of total identified compounds in each lipid fraction. According to their degree of unsaturation (number of ethylenic or “double” bonds), fatty acids were grouped and estimated in saturated fatty acids (SFAs), mono-unsaturated fatty acids (MUFAs) and poly-unsaturated fatty acids (PUFAs). In addition, PUFAs, having more than one double bond present within the molecule, were further classified into two groups: omega-3 (n-3) and omega-6 (n-6). This classification is based on their chemical structure and specifically if their first double bond is located in 3 or 6 carbons from the methyl end of the molecule.

6.2.5 Multiple regression analysis

For the 50 selected spawning capable females, the reproductive variables were correlated with the fish size through a linear regression analysis in R statistical software (R Development Core Team 2013).

The reproductive variables, that resulted significantly correlated with the fish size, were selected as response variables for building a multivariate linear regression model in the R statistical software. In this model, the females' fork length (F_L) and the fatty acid profiles were used as explanatory variables:

$$G \sim F_L + FA1 + FA2 + \varepsilon$$

Where: G is the reproductive output to explain, F_L is the fish fork length, FA1 and FA2 are the fatty acids profiles from the neutral and polar fractions that explain a major part of the variability of G. ε are the residuals.

A problem with multiple regression model is the multi-collinearity effects that occur when the predictor variables are too strongly correlated to each other, making the parameter estimates unstable and difficult to interpret. For this reason, the analysis was performed separately for each group of fatty acids (SFA, MUFA and PUFA). Moreover, the significant effects of PUFAs from PL and NL were firstly tested separately. Then, fatty acids from the two lipid fractions with a significant effect on the response variable were aggregated together in the model. However, the possible presence of multi-collinearity among explanatory variables was investigated through the use of variance inflation factors (vif) calculations. VIF was calculated using the function vif () in the R package car.

6.2.6 Genome Wide Association Study (GWAS)

The association between phenotype and genotype was examined by comparing allelic patterns between individuals and different phenotypic traits (Fig 6.4).

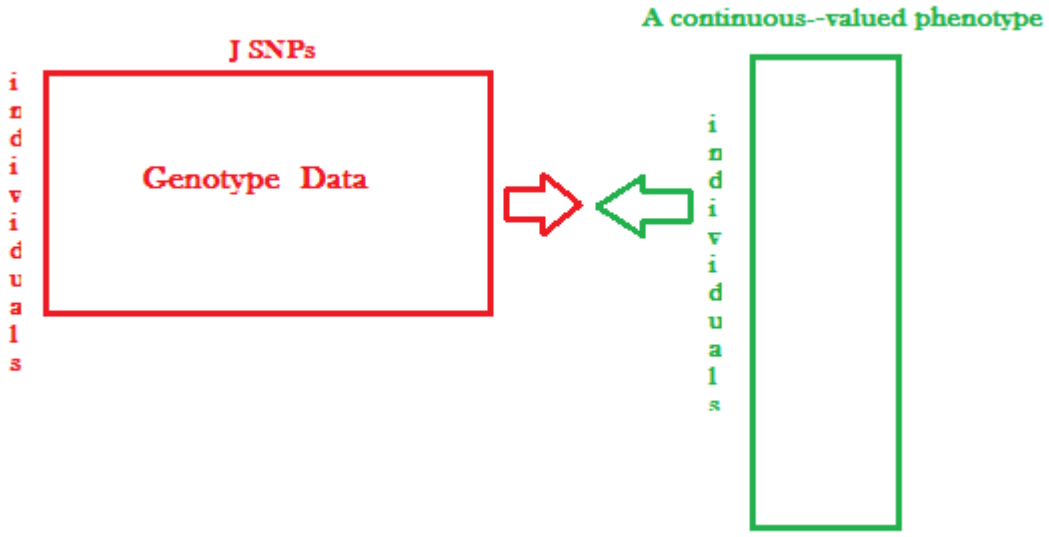


Fig. 6.4_Schematic representation of the association between the j-th SNPs and continuous phenotype traits.

Statistical inference was performed using a linear regression model in R, analyzing all SNPs separately to assess the influence of each SNP on the given phenotype. In particular, denoting with “a” the minor allele and with “AA”, “Aa” and “aa” the three corresponding SNP genotypes, the following model was considered:

$$y= \beta0+\beta1x1+ \beta2x2+\varepsilon$$

-Data:

- y: a continuous-valued phenotype
- x1: Dummy variable for coding SNP genotype Aa at a given locus
- x2: Dummy variable for coding SNP genotype aa at a given locus

–Parameters

- β0: intercept term (expected value for phenotype y for SNP genotype AA)
- β1: differential effect for SNP genotype Aa
- β2: differential effect for SNP genotype aa
- ε: noise or the part of y that is not explained by the SNP x

Association between SNP and phenotype was assessed by testing the following hyphotesis: H0: β1= β2=0, using an ANOVA F test. Due to the large number of genetic markers, corrections for multiple testing were performed using the positive false discovery rate (FDR) method (Benjamini

andYekutieli, 2001), in order to reduce the increased probability of false positive findings. We considered tests significant when the FDR was <5% (q< 0.05).

From those SNPs significantly associated with the phenotypic traits, the consensus DNA sequences (contigs) were retrived from the genome of *Thunnus orientalis*. In order to annotate the contigs, BLASTx was run against a protein sequence database. Briefly, BLASTX, using a nucleotide query against a protein database, is able to reliably identify protein-coding regions within a DNA sequence if sufficient similarity exists between the translated query and an entry in the data-base.

6.3 Results

6.3.1 Reproductive analysis

The Table 6.1 presents the main reproductive variables measured for each individual of the 50 spawning capable YFT females.

Table 6.1. Details on the reproductive output in *Thunnus albacares* females: acronym (Sample ID), fork length (F_L) in cm, fish weight (W) in Kg, gonad weight (GW) in Kg, batch fecundity (BF) in millions of oocytes, relative batch fecundity (BFrel) in oocytes per gram of gonad-free weight, total number of developing oocytes (NDOtot) in millions and their mean diameter (NDOmn) in μm. In addition, mean value and standard deviation (SD) are provided per each variable.

Sample ID	F_L	W	GW	BF	BFrel	NDOtot	NDOmn
34-Y-ME-1	132	44	1.0	1.35	31.35	6.86	336.42
34-Y-ME-3	136.6	48.8	1.1	3.23	67.66	6.88	284.25
34-Y-ME-6	142.5	55.3	1.0	1.62	29.93	7.04	323.75
34-Y-ME-7	139	51.4	0.7	2.40	47.87	9.61	280.65
34-Y-ME-8	136.5	48.7	1.0	3.03	63.73	1.23	323.66
34-Y-ME-9	139.8	52.2	1.1	1.34	26.22	7.50	321.73
34-Y-ME-17	133	45	1.5	3.48	79.99	8.10	356.61
34-Y-ME-18	149.5	63.8	1.5	5.34	86.66	2.02	294.20
34-Y-ME-20	139	51.4	1.1	1.85	36.89	5.74	352.25
34-Y-ME-21	135.3	47.1	2.2	1.88	41.28	3.45	334.54
34-Y-ME-22	141.2	53.8	1.0	7.89	14.91	9.66	260.67

34-Y-ME-23	127.7	39.9	0.9	2.91	75.40	7.30	288.68
34-Y-ME-24	145.5	58.8	1.2	2.96	51.46	1.02	289.41
34-Y-ME-25	144	57.1	1.4	2.86	51.20	7.49	320.40
34-Y-ME-26	146.2	59.7	0.9	1.45	24.78	7.94	320.07
34-Y-ME-27	141	53.6	1.1	1.25	23.84	8.75	311.88
34-Y-ME-28	139.8	52.2	1.3	3.00	59.20	7.11	346.92
34-Y-ME-29	139.3	51.7	1.5	2.08	41.41	6.40	349.56
34-Y-ME-33	154	69.7	1.4	3.27	47.92	1.28	294.47
34-Y-ME-35	139.5	51.91	1.9	3.09	61.88	3.27	345.71
34-Y-ME-36	139.9	52.35	2.0	2.84	56.45	1.31	315.33
34-Y-ME-37	142.9	55.76	1.7	2.79	51.75	1.62	301.96
34-Y-ME-39	138.1	50.37	1.3	2.20	44.69	3.32	327.14
34-Y-ME-40	143.6	56.58	1.0	2.14	38.82	8.17	326.33
34-Y-ME-45	147.1	60.79	1.0	4.52	76.70	1.42	283.95
34-Y-ME-46	140.2	52.69	1.5	1.46	28.42	8.73	312.97
34-Y-ME-47	141.5	54.15	1.0	4.67	88.92	1.98	317.70
34-Y-ME-48	146.6	60.17	0.8	1.96	33.35	7.97	336.87
34-Y-ME-50	146.7	60.29	1.0	1.71	28.77	5.00	290.83
34-Y-ME-54	146	59.44	1.2	3.35	57.98	1.45	299.23
34-Y-ME-55	144.2	57.29	1.6	4.48	79.65	6.18	350.62
34-Y-ME-56	150.5	65.06	0.8	2.33	36.48	8.44	288.76
34-Y-ME-58	130.5	42.56	1.8	1.88	45.04	6.94	261.67
34-Y-ME-60	135.6	47.71	1.5	2.83	61.06	1.35	316.03
34-Y-ME-61	137.3	49.51	0.7	2.54	52.50	8.04	316.15
34-Y-ME-62	144	57.05	1.4	2.79	50.10	1.34	305.66
34-Y-ME-63	146.5	60.05	1.2	3.60	61.70	1.39	337.75
34-Y-ME-64	143.9	56.93	0.8	2.47	44.59	1.39	293.97
34-Y-ME-65	141.1	53.7	1.6	3.89	74.62	1.41	306.52
34-Y-ME-67	148.5	62.52	1.3	2.50	40.96	1.23	299.30

34-Y-ME-69	141.9	54.61	1.3	3.13	58.77	9.61	337.48
34-Y-ME-70	150.5	65.06	1.6	2.36	37.10	1.00	290.64
34-Y-ME-72	139.7	52.13	1.4	2.49	48.88	1.35	320.05
34-Y-ME-77	128.8	40.93	1.1	1.96	48.96	6.12	316.10
34-Y-ME-81	154.5	70.35	1.3	2.78	40.51	1.08	326.27
34-Y-ME-82	125.8	38.16	1.3	1.37	36.62	5.46	301.68
34-Y-ME-83	143.6	56.58	1.4	3.11	56.47	1.57	321.23
34-Y-ME-84	138.4	50.7	1.7	2.36	48.23	1.77	327.44
34-Y-ME-85	149.7	64.04	1.1	5.03	81.62	1.45	306.26
34-Y-ME-86	151.8	66.75	2.3	2.54	38.78	5.84	250.44
Mean (± SD)	141.6 (± 6.4)	54.6 (± 7.3)	1.29 (± 0.3)	2.66(±0.99)	50.2 (± 17.6)	1.00 (±0.4)	312.4 (± 24.3)

The BF was estimated at 2.66±0.9 million oocytes and the BFrel at 50.2±17.6 oocytes per gram of gonad-free weight. None of those variables were significantly correlated (PNS) with the fish size. On the contrary, the total number of developing oocytes and the gonad weight were significantly correlated with the size ($P<0.01$). The size had a significant effect on the increasing of both the total number of developing oocytes and gonad weight, explaining the 12% and 14.5 % of their variability, respectively. The mean gonad weight was 1275.9 grams, ranging from 655.5 to 2305.3 grams.

6.3.2 Genomics analysis

The 2b-RAD reads were mapped against the genome of *Thunnus orientalis* with a percentage of success higher than 86%. The *ref_map.pl* program in Stacks was run on the resulting dataset obtained after the quality and adaptors trimming process (Table 6.2). After this process, in average the 86.3% of reads was retained with in total 21328 SNPs shared among the 50 females. However, only 7550 SNPs were retained for subsequent analyses, having a MAF > 0.5 and a percentage of missing value < 10%.

Table 6.2. Summary statistics of the 50 females of *Thunnus albacares*. The table reports: the identification code (ID), the number of raw reads and the corresponding number of filtered reads, the percentage of reads retained, the value of unique tags, polymorphic SNPs and number of loci found.

ID code	Raw reads	Filtered reads	% of reads retained	Unique Stacks	Polymorphic Loci	SNPs Found
34_Y_ME_1	5703894	4985821	87.4	138480	11924	14319
34_Y_ME_17	4327318	3827215	88.4	130918	11272	13456
34_Y_ME_18	8975762	7522896	83.8	146618	12541	15171
34_Y_ME_20	7807073	7040062	90.2	143898	12265	14714
34_Y_ME_21	4768136	4231495	88.7	139226	12101	14415
34_Y_ME_22	4845506	4043267	83.4	132089	11464	13734
34_Y_ME_23	12011988	10719107	89.2	149746	12816	15525
34_Y_ME_24	11549711	9969565	86.3	149543	12791	15486
34_Y_ME_25	12633549	11152477	88.3	151253	12973	15640
34_Y_ME_26	11023917	10006567	90.8	148577	12725	15432
34_Y_ME_27	8390702	7562676	90.1	144615	12524	14984
34_Y_ME_28	6021472	4492771	74.6	136011	11773	14181
34_Y_ME_29	10408741	9421842	90.5	148081	12387	15007
34_Y_ME_3	12003058	10690603	89.1	149539	12710	15320
34_Y_ME_33	7286778	5886898	80.8	129550	10925	13131
34_Y_ME_35	10709811	9160480	85.5	147626	12602	15194
34_Y_ME_36	6726475	5545731	82.4	139630	12063	14450
34_Y_ME_37	7195260	6108528	84.9	141004	11982	14389
34_Y_ME_39	7353003	5885137	80.0	142102	12309	14753
34_Y_ME_40	3528834	3242053	91.9	124384	10788	12918
34_Y_ME_45	4716812	4032780	85.5	147161	12748	15283
34_Y_ME_46	6057862	5354723	88.4	138829	11918	14332
34_Y_ME_47	6490810	5759821	88.7	138727	12081	14532
34_Y_ME_48	5227023	4651060	89.0	133761	11673	13933
34_Y_ME_50	6235471	5597685	89.8	138596	12046	14435
34_Y_ME_54	7370027	6289860	85.3	144455	16971	20323
34_Y_ME_55	8662128	7620591	88.0	143913	12514	15090
34_Y_ME_56	5482506	4852604	88.5	133688	11489	13657
34_Y_ME_58	8305781	7500946	90.3	142436	12241	14782
34_Y_ME_6	8171016	6691975	81.9	144611	12533	15078
34_Y_ME_60	4293580	3673621	85.6	128189	10988	13077
34_Y_ME_61	5705199	4969654	87.1	135805	11776	14125
34_Y_ME_62	5707467	4741143	83.1	136532	11798	14066
34_Y_ME_63	3493863	2923497	83.7	121074	10658	12732
34_Y_ME_64	11834977	9895634	83.6	151551	12990	15610
34_Y_ME_65	7022455	6304615	89.8	139763	11964	14353
34_Y_ME_67	4196849	3522322	83.9	126701	10979	13170
34_Y_ME_69	5809394	5025519	86.5	137076	11994	14372
34_Y_ME_7	6773079	6121034	90.4	139089	11736	14043
34_Y_ME_70	5910715	5316666	89.9	136782	11718	14020
34_Y_ME_72	10823702	9557720	88.3	147645	12732	15348
34_Y_ME_77	10202748	9222457	90.4	147077	12860	15540
34_Y_ME_8	6269692	5533403	88.3	140660	12125	14609

34_Y_ME_81	7519081	6427225	85.5	142405	12113	14562
34_Y_ME_82	4667536	4118182	88.2	130378	11434	13717
34_Y_ME_83	7605051	6061030	79.7	141737	12116	14533
34_Y_ME_84	7430267	5885223	79.2	129606	11032	13325
34_Y_ME_85	6773277	5523800	81.6	141399	12114	14507
34_Y_ME_86	10211284	8737976	85.6	147751	12737	15315
34_Y_ME_9	6724045	5703243	84.8	139295	12118	14459

6.3.3 Fatty acids analysis

Table 6.3 shows the overall fatty acid composition in the 50 YFT females' ovary measured in Polar (PL) and Neutral lipid (NL) fractions.

Table 6.3. Fatty acid composition in the neutral and polar lipid fractions of the YFT gonads.

Fatty acid	% of total FA in neutral lipid fraction (Mean ± SD.)	% of total FA in polar lipid fraction (Mean ± SD.)
C14:0	2.11±1.13	0.25±0.18
C15:0	1.26±0.44	0.34±0.27
C16:0	28.8±6.66	29.59±7.47
C17:0	0.98±0.21	0.97±0.28
C18:0	9.54±3.55	9.27±1.46
C20:0	2.67±1.22	0.49±0.3
C22:0	0.03±0.03	0.04±0.03
C16:1n-7	5.85±1.36	1.14±0.43
C17:1n-7	0.74±0.15	1.05±0.39
C18:1n-9	10.54±2.04	8.52±1.22
C18:1n-7	2.52±0.29	1.69±0.28
C20:1n-9	0.66±0.2	0.33±0.18
C22:1n-9	0.59±0.31	0.52±0.3
C24:1n-9	0.09±0.1	0.08±0.06
C18:3n-3	0.43±0.1	0.19±0.06

C18:4n-3	0.36±0.08	0.12±0.04
C20:3n-3	0.04±0.04	0.03±0.02
C20:4n-3	0.48±0.15	0.29±0.08
C20:5n-3 (EPA)	4.12±0.66	5.81±1.45
C22:5n-3	1.05±0.26	0.82±0.26
C22:6n-3 (DHA)	20.91±5.4	25.87±6.25
C18:2n-6	1.29±0.39	0.65±0.14
C18:3n-6	0.4±0.18	0.17±0.06
C20:2n-6	0.07±0.04	0.06±0.04
C20:3n-6	0.21±0.11	0.16±0.13
C20:4n-6 (AA)	2.1±0.43	5.24±0.94
C22:5n-6	1.03±0.22	1.63±0.34
Sum SFA	44.5±5.75	40.19±7.9
Sum MUFA	20.77±2.87	12.8±1.66
Sum PUFA	32.26±7.27	40.32±8.93
Sum PUFA n-3	27.2±6.71	32.64±7.87
Sum PUFA n-6	5.06±1.05	7.68±1.29

In general, the fatty acid profile in NL showed a higher variation than in PL. For NL, SFA was the most abundant group (44.5%), while in PL the most abundant groups were SFA and PUFA (40.19% and 40.32%, respectively). Overall, NL contained a higher percentage of SFAs and MUFAs but a lower percentage of PUFAs than PL. However, for both lipid fractions, the three most abundant fatty acids in the gonads were C16:0 (palmitic acid), C22:6n-3 (docosahexaenoic acid, DHA) and C18:1n-9 for NL, and C16:0, DHA, C18:0 for PL.

The fatty acid C16:0 was the most abundant SFA at both fractions (28.8% in NL vs 29.59% in PL), followed by C18:0 (stearic acid; 9.54% NL vs 9.27% PL). Regarding MUFAs, in both lipid fractions the most abundant fatty acid found was the C18:1n-9 with however a higher concentration in NL (ranged from 5.16 to 24.10%) than in PL (ranged from 7.72 to 17.58%). The primary source of total PUFAs found in gonads tissues were the namely n-3 fatty acids C20:5n-3 (Eicosapentaenoic acid;

EPA) and DHA. The arachidonic acid (AA; C20:4n–6) was the most abundant n-6 PUFA in our samples.

6.3.4 Multiple regression analysis

The gonad weight (Fig. 6.5) was chosen as explanatory variable for the multiple linear regression model, according to the fact that no fatty acids were significantly correlated (P value > 0.05) with the total number of developing oocytes.

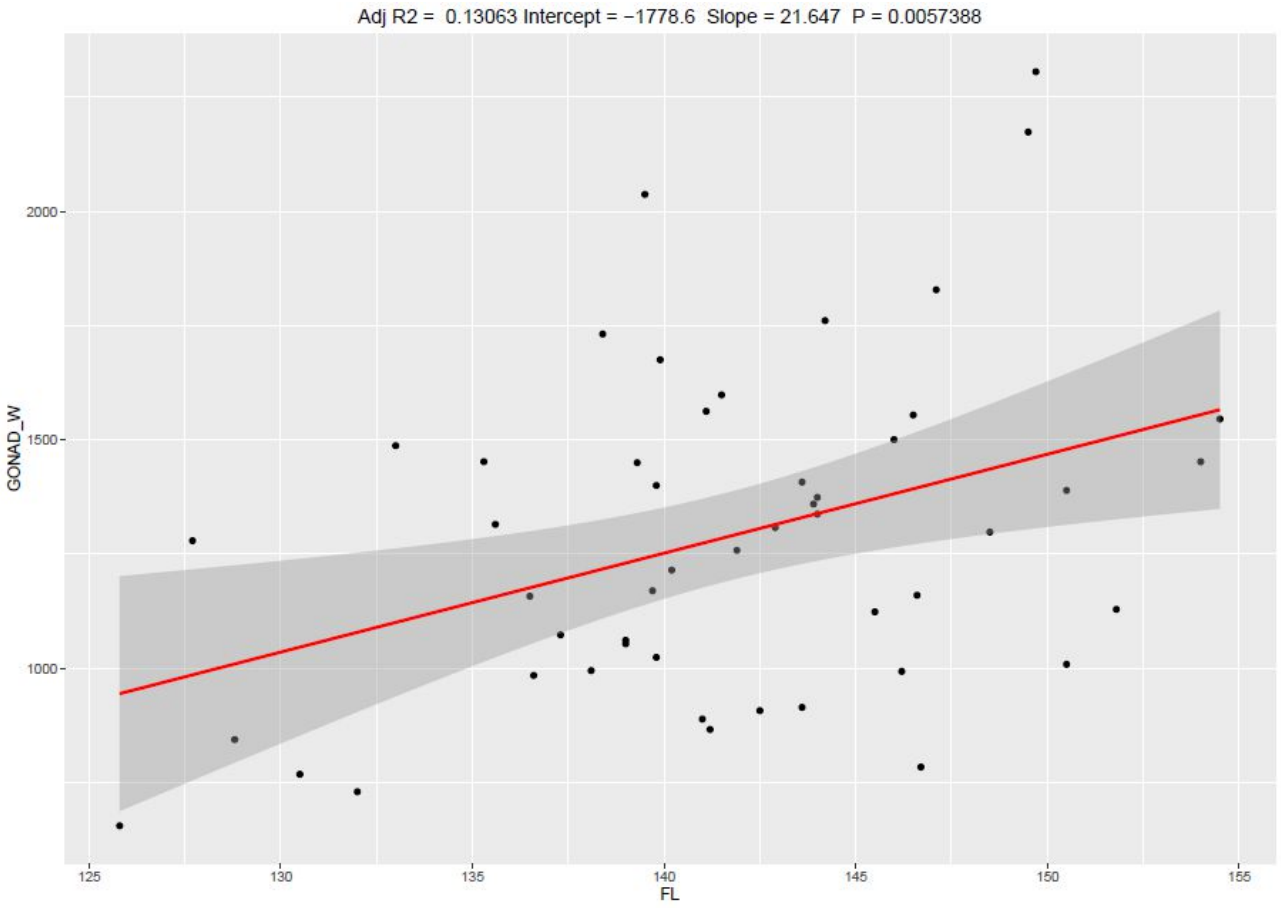


Fig. 6.6_Scatter plot of the gonad weight (GW, g) and the fork length (F_L, cm) for the 50 females of yellowfin tuna. In red is displayed the regression line.

In the polar lipids, any significant correlation was found among both SFA and MUFA groups, the gonad weight and the fish size. Instead, in the neutral lipids it was detected a significant effect of specific SFAs and MUFAs on the variability of the gonad weight. Specifically for SFAs, C17:0 and

C16:0 explained together around 33% of the variation of the gonad weight in relation to the fish size. Significant correlations (p-value<0.001) were also found with C18:1n-9 and C18:1n-7 (MUFAs), which explained the 26% of the variance in the gonad weight.

Contrarily, significant correlations were found with some combinations of PUFAs as well as with some omega-3: omega-6 ratios. For the polar fraction, a significant correlation (p-value<0.001) was detected considering as explanatory variables the sums of omega-3 (n3-PUFAs) and omega-6 (n6-PUFAs), with an Adjusted r^2 of 0.415. The regression coefficient was positive (Estimate: 3.92) for n3-PUFAs, while it was negative for n6-PUFAs (Estimate: -2.52).

Exploring the effects of each PUFA of the polar fraction, it was discovered that among all the omega-3 and omega-6 fatty acids, only DHA (RC: 51.864), AA (RC: -389.214) and C18:2n-6 (RC: 798.591) had a significant effect on the model (p-value<0.001), explaining together with the fork length around the 50% of the gonad weight variability (Adjusted R-squared: 0.5051).

Adding in the model the specific interaction of those fatty acids with the fish size, the r^2 was further increased (Adjusted R-squared: 0.5814; p-value<0.001). This result demonstrated that the female size enhances the opposite effect of these three fatty acids: underlining that the increase of the gonad weight in correspondence to the proportion of specific omega-3 and omega-6 will be higher in larger females. This pattern was also confirmed by the significant effect (pvaue<0.001) of the ratio n3-PUFAs over n6-PUFAs and the ratio EPA over AA, which explained respectively the 29% and 24% of the increase of the gonad weight in relation to the mother size.

Instead in the neutral fraction the only two PUFAs with a significant (p-value<0.001) and positive effect on the model were C20:4n-3 and C18:3n-6 (Adjusted R-squared: 0.37). These two PUFAs were added in the model (Table 6.4), together with the three significant PUFAs detected for the polar fraction (DHA, AA and C18:2n-3).

Table 6.4_Summarizing the regression coefficients of each predictor variable included in the model. The total Adjausted R-squared and the p-value are also reported. PL: polar lipid fraction; NL: neutral lipid fraction

	Estimate	Std. Error	t value	Pr(>|t|)	vif
(Intercept)	5.539	33.971	0.163	0.871271	
F_L	20.779	5.618	3.699	0.000636	1.11
C22:6n-3 PL	32.088	10.971	2.925	0.005591	4.22
C20:4n-6 PL	-321.181	65.992	-4.867	1.72E-05	3.49
C18:2n-6 PL	298.436	336.535	0.887	0.380364	1.74
C20:4n-3 NL	1016.626	336.109	3.025	0.004283	2.15
C18:3n-6 NL	479.57	243.453	1.97	0.055634	1.24
Adjusted R-squared:	0.5972				
p-value:	5.21E-08				

However, the effect of C18:2 n-6 was not significant (p-value>0.05) and hence it was removed from the model. Running the model with the other four fatty acids made possible to explain almost the 60% (Adjusted R-squared: 0.5993) of the variability of the gonad weight and all the explanatory variables considered had a significant effect on the model. In addition, adding in this model the interactions between PUFAs and the fish size allowed to further increase the percentage of variability explained (Adjusted R-squared: 0.6765).

For all the considered models, VIF values were always lower than 10, thus suggesting the absence of severe multi-collinearity.

6.3.5 Genome Wide Association Studies

Any marker was significantly associated with the gonad weight, the fork length and fatty acids included in the multiple linear regression models. However different unique SNPs were significantly associated with different fatty acids of both polar and neutral lipid fractions (see details in the Appendix 6.2). The BLASTx results reported the predicted protein sequences for each of the contig's sequence used as query with the corresponding level of similarity. However, specific functional genomics studies are needed to investigate the biological role that those protein-coding transcripts play in YFT.

6.4 Discussion and Conclusions

Our results pointed out the shortcoming of BF and relBF as response variables for investigating the variation of the fecundity and energetic allocation in YFT females in relation to their size. The inefficiency of those variables for the purposes of the present study is linked to the way in which they are obtained that is strictly dependent on the fish size. In doing so, their use as response variable in the multiple regression model would nullify any possibility to hypothesize an increase of the reproductive success related to the mother size.

On the contrary, here using the fish gonad weight, which is a good indicator of the individual reproductive effort, we demonstrated for the first time that within a YFT spawning stock, females allocate differently their available nutrient reserves and dietary resources for reproduction. This different energetic pattern is mainly related to their size. In fact, our results pointed out that larger YFT females with larger gonads not only have a greater body volume for holding eggs, but also they allocate a greater fraction of surplus energy to egg production than smaller ones. This higher energetic allocation will enhance the spawning quality, increasing the likelihood to produce higher-quality offspring that will augment, in turn, larvae survival chances through a decrease of development duration from embryogenesis to the first oral feeding (Fernández-Palacios et al., 2011). The higher quality of ova lipid reserves in larger-size YFT females is highlighted by the higher concentration of certain polyunsaturated fatty acids (PUFAs) especially in Polar Lipids (PL). The high percentage of dietary PUFAs in PL underlines their importance for YFT reproductive processes (Tocher, 2003). In fact tunas, like other animals, cannot synthesize *de novo* omega-3 and omega-6 PUFAs, lacking of the appropriate fatty acid desaturase enzymes. Thus, the proportion of the different PUFAs in the gonads reflects the amount assumed by feeding. Balance in the diet of both PUFAs n-3 and n-6 is an essential point for optimizing fish reproductive success (Acharia et al., 2000). PUFAs in general, and omega-3 in particular, actively participate in gonad maturation, egg quality (Izquierdo et al., 2001) and larval growth of fish (Tulli and Tibaldi, 1997), regulating also the production of eicosanoid (prostaglandins), steroid hormones and gonad development (ovulation; Izquierdo et al., 2001). Our results indicated that larger females had a higher concentration of docosahexaenoic acid (C22:6n-3; DHA) but a lower concentration of arachidonic acid (C20:4n-6; AA), and linoleic acid (C18:2n-6). DHA is an essential fatty acid that fishes are not able to synthesize from the essential precursors alpha-linolenic acid (Riediger et al., 2009). It has a specific structural role in nervous tissue (Sargent et al., 1993) and higher supply of DHA available

after the start of feeding would support the rapid development of membrane systems (Tocher, 2003). In NL it was observed a much higher level of DHA than EPA and their ratio increased with the female size. This result confirmed that usually tuna lipids, including triacyglycerols and phospholipids, have higher levels of DHA than EPA (Murase and Saito, 1996).

Overall, it was detected a higher PUFAs n-3/n-6 ratio in PL than in NL of larger females. Our results showed that an increase of gonad weight in larger females corresponds also to a decrease of AA and a consequent increase of EPA/AA ratio. Higher levels of EPA/AA ratio, which is crucial for determining eicosanoid actions, have been associated with a superior resistance to infection in several marine and freshwater species (Sargent et al., 1995). Therefore, dietary intake of these fatty acids can assume a relevant importance in YFT reproduction, even if there are not information about the optimal intake of omega-3 to guarantee the spawning quality and reproductive success of this species.

High levels of omega-3 in lipids and in particular of DHA and EPA are a prerogative of tuna species (Murase and Saito, 1996). For instance, in the Pacific Ocean YFT showed a total amount of n-3 PUFAs approximately around 35% of total fatty acids with DHA alone accounting for between 25% and 30% (Sunarya et al., 1995). Although, the relatively high level of omega-3 measured in this study for YFT females seems to be an inherent characteristic of tuna species, however the higher amount of omega-3 in the gonads of larger females indicates a change in the energetic strategy of retaining/accumulating those fatty acids for reproduction in relation to the fish size (Tocher, 2003).

In NL, higher levels of the palmitic acid (16:0) and oleic acid (18:1n-9) were also found in gonads of larger mothers. These fatty acids have significant quantitative and qualitative roles in structural phospholipids (Bell and Dick, 1991) and they can be biosynthesized *de novo* by fish as well as by all known organisms (Sargent et al., 1989).

Our results indicated a possible onthogenic shift in the energy allocation strategy of YFT females, highlighting that larger individuals allocate more energy for reproduction than smaller ones, increasing noticeably the spawning quality. On the contrary, smaller females invest more energetic resources for somatic growth than larger ones (Zudaire et al., 2014). These authors showed a negative relationship between the amount of total lipids in the muscle and the size of YFT females. Therefore, larger mothers will switch the energy allocation from somatic to gonadic growth for ensuring future reproductive opportunities (Wiegand et al., 2007). This size-related energetic allocation strategy might be linked to a much higher natural mortality rate (M) in females with a F_L

> 130 cm than in midsized individuals (Hampton and Fournier, 2001). In doing so, larger females could allocate all their energetic resources for reproduction before dying in order to increase the offspring quality and fitness, thus promoting the survival of future generations.

This study provided only a qualitative perspective of the maternal effect in YFT females, since it was not possible to show any increase of the production of oocytes per gram of gonad weight (i.e. relative batch fecundity) related with the mother size. However, the maternal effect could be interpreted, in terms of quantity, through the possibility to have a wider spatial and temporal window for spawning that certainly will increase the number of oocytes released during the all spawning season (Zudaire et al., 2014; Schaefer, 1998).

Many doubts were raised here about the use of the spawning stock biomass (SSB) as a proxy of YFT stock reproductive potential by the tRFMOs. In fact, if a sustainable fishery is the desired outcome of the current management of YFT in the Atlantic Ocean, then it must take into account the reproductive importance of large females that may result essential for the stock rebuilding. It is also evident that additional demographic criterions, that account for the reproductive importance of larger and most experienced spawners, are needed to properly estimate the reproductive potential of YFT stocks.

In such a context, we argue in favor of preserving such larger-size females to increase the stock productivity in the Atlantic Ocean (Berkeley et al., 2004; Birkeland and Dayton, 2005). Protecting those spawners (i.e. establishing no-take areas), which produce offspring of higher quality, would undoubtedly increase per capita reproductive output (Kaiser et al., 2007) and would provide an added bonus to YFT stocks, because the recruitment rates of the offspring should be congruently higher. Oppositely, increasing the mortality of larger and most experienced spawners through different fishing methods (i.e. tuna purse seiner and longline fisheries) might have potential negative consequences for YFT stocks, altering the time and the location of the spawning events and thus decreasing the production and quality of eggs released. Consequently, large YFT females have a crucial relative reproductive value (Grey and Law, 1987), which may intensely contribute to year class strength and surplus production under exploited conditions (Arlinghaus et al., 2010). Maintaining and protecting highly fecund large individuals can thus represent an efficient and reliable strategy for a sustainable management of YFT in the Atlantic Ocean.

Instead, our GWAS failed to find genomic regions significantly associated with the gonad weight, the fish size and the different percentage of specific fatty acids that intervene in increasing the YFT spawning quality. These results pointed out that the energetic patterns related to the female's size

in this pelagic species are not under the control of specific genomic regions. Although, the corrections for multiple testing were performed with the False Discovery Rate (FDR; pvalue<0.05), that is less conservative than a Bonferroni adjustment (Narum, 2006), however the small sample size hold the analysis so stringent that only the largest effects could be detected. Besides, the use of the draft genome of another tuna species could have left off parts of the genome in which are present SNPs involved in those energetic patterns. In summary, due to the complexity of these traits and the intrinsic limitation factors of our study, these results are not conclusive and future investigations with more samples are required in order to determine if there are candidate genes involved in regulating those reproductive and energetic processes.

Future effort is also needed for analysing the fatty acids composition in somatic tissues, such as white muscle and liver, in order to understand how the energy is transferred from those tissues to the gonads during spawning events in relation to females' size. These information are crucial to point out if the size-related fatty acids composition pattern observed in the gonads in YFT females is also confirmed by the dynamics of the somatic energy reserves during reproduction.

References

Ackman R.G. and Takeuchi T. "Comparison of fatty acids and lipids of smolting hatchery-fed and wild Atlantic salmonSalmo salar". Lipids 21, (1986): 117–120.

Alonso-Fernández A. and Saborido-Rey F. "Relationship between energy allocation and reproductive strategy in Trisopterus luscus". J. Exp. Mar. Biol. Ecol. 416, (2012): 8–16.

Aristizabal E.O. "Energy investment in the annual reproduction cycle of female red porgy, Pagrus pagrus (L.)". Mar. Biol. 152, (2007): 713–724.

Arlinghaus R., Matsumura S. and Dieckmann, U. "The conservation and fishery benefits of protecting large pike (Esox lucius L.) by harvest regulations in recreational fishing". Biol. Conserv. 143, (2010): 1444–1459.

Arocha F., Lee D.W., Marcano L.A. and Marcano, J.S. "Update information on the spawning of yellowfin tuna, *Thunnus albacares*, in the western central Atlantic". Col Vol Sci Pap ICCAT 52, (2001): 167–176.

Bell, J.G. and Sargent, J.R. "Arachidonic acid in aquaculture feeds: current status and future opportunities". Aquaculture 218, (2003): 491–499.

Bell M.V. and Dick J.R. "Molecular species composition of the major diacyl glycerophospholipids from muscle, liver, retina and brain of cod (Gadus morhua)". Lipids 26, (1991): 565–573.

Berkeley S.A., Chapman C. and Sogard, S.M. "Maternal age as a determinant of larval growth and survival in a marine fish, Sebastes melanops". Ecology 85, (2004): 1258–1264.

Birkeland C. and Dayton P.K. "The importance in fishery management of leaving the big ones". Trends Ecol. Evol. 20, (2005): 356–358.

Bobko S.J. and Berkeley S.A. "Maturity, ovarian cycle, fecundity, and age-specific parturition of black rockfish (Sebastes melanops)". Fish. Bull. 102, (2004): 418–429.

Bodin, N., Lucas, V., Dewals, P., Adeline, M., Esparon, J., Chassot, E. "Effect of brine immersion freezing on the determination of ecological tracers in fish". European Food Research and Technology 238, (2014): 1057-1062.

Brown-Peterson N.J., Wyanski D.M., Saborido-Rey F., Macewicz B.J. and Lowerre-Barbieri S.K. "A standardized terminology for describing reproductive development in fishes". Mar. Coast. Fish. 3, (2011): 52–70.

Bush W.S. and Moore J.H. Genome-wide association studies. PLoS Comput Biol 8, (2012): e1002822.

Cardinale M. and Arrhenius F. "The relationship between stock and recruitment: are the assumptions valid?". Mar. Ecol. Prog. Ser. 196, (2000): 305–309.

Catchen J., Hohenlohe P.A., Bassham S., Amores A. and Cresko W.A. "Stacks: an analysis tool set for population genomics". Mol. Ecol. 22, (2013): 3124–3140.

Folch J., Lees M., Sloane-Stanley G.H. and others. A simple method for the isolation and purification of total lipids from animal tissues. J Biol Chem 226, (1957): 497–509.

Grey D.R. and Law, R. "Reproductive values and maximum yields". Funct. Ecol. (1987): 327–330.

Hampton J. and Fournier D.A. "A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (Thunnus albacares) in the western and central Pacific Ocean". Mar. Freshw. Res. 52, (2001): 937–963.

Henderson R.J. and Tocher D.R. "The lipid composition and biochemistry of freshwater fish". Prog. Lipid Res. 26, (1987): 281–347.

Hunter J.R., Macewicz B.J. and Kimbrell C.A. "Fecundity and other aspects of the reproduction of sablefish, Anoplopoma fimbria, in central California waters". CalCOFI Rep 30, (1989): 61–72.

Hutchings J.A. and Myers R.A. "Effect of age on the seasonality of maturation and spawning of Atlantic cod, Gadus morhua, in the Northwest Atlantic". Can. J. Fish. Aquat. Sci. 50, (1993): 2468–2474.

Izquierdo M.S., Fernandez-Palacios H. and Tacon A.G.J. "Effect of broodstock nutrition on reproductive performance of fish." Aquaculture 197, (2001): 25–42.

Jombart T., Devillard S., Dufour A.-B. and Pontier D. "Revealing cryptic spatial patterns in genetic variability by a new multivariate method". Heredity 101, (2008): 92–103.

Kaiser M.J., Blyth-Skyrme R.E., Hart P.J., Edwards-Jones G. and Palmer D. "Evidence for greater reproductive output per unit area in areas protected from fishing". Can. J. Fish. Aquat. Sci. 64, (2007): 1284–1289.

Kell L.T., Nash R.D.M., Dickey-Collas M., Mosqueira I. and Szuwalski, C. "Is spawning stock biomass a robust proxy for reproductive potential?". Fish Fish. (2015).

Kjesbu O.S., Witthames P.R., Solemdal P. and Walker, M.G. "Temporal variations in the fecundity of Arcto-Norwegian cod (Gadus morhua) in response to natural changes in food and temperature". J. Sea Res. 40, (1998): 303–321.

Lowerre-Barbieri S.K., Ganias K., Saborido-Rey F., Murua H. and Hunter J.R. "Reproductive timing in marine fishes: variability, temporal scales, and methods". Mar. Coast. Fish. 3, (2011): 71–91.

Marshall C.T., Kjesbu O.S., Yaragina N.A., Solemdal P. and Ulltang Ø. "Is spawner biomass a sensitive measure of the reproductive and recruitment potential of Northeast Arctic cod?" Can. J. Fish. Aquat. Sci. 55, (1998): 1766–1783.

Marshall C.T., Yaragina N.A., Lambert Y. and Kjesbu O.S. "Total lipid energy as a proxy for total egg production by fish stocks". Nature 402, (1998): 288–290.

Marteinsdóttir G. and Steinarsson A. "Maternal influence on the size and viability of Iceland cod Gadus morhua eggs and larvae". J. Fish Biol. 52, (1998): 1241–1258.

McBride R.S., Wuenschel M.J., Nitschke P., Thornton G. and King J.R. "Latitudinal and stock-specific variation in size-and age-at-maturity of female winter flounder, Pseudopleuronectes americanus, as determined with gonad histology". J. Sea Res. 75, (2013): 41–51.

Metcalfe L.D. and Schmitz A.A. "The rapid preparation of fatty acid esters for gas chromatographic analysis". Anal. Chem. 33, (1961): 363–364.

Morgan M.J., Murua H., Kraus G., Lambert Y., Marteinsdottir G., Marshall C.T., O'Brien L. and Tomkiewicz, J. "The evaluation of reference points and stock productivity in the context of alternative indices of stock reproductive potential". Can. J. Fish. Aquat. Sci. 66, (2009): 404–414.

Murase T. and Saito H. "The Docosahexaenoic Acid Content in the Lipid of Albacore Thunnus alalunga caught in Two Separate Localities". Fish. Sci. 62, (1996): 634–638.

Murawski S.A., Rago P.J., Trippel E.A. "Impacts of demographic variation in spawning characteristics on reference points for fishery management". ICES J. Mar. Sci. J. Cons. 58, (2001): 1002–1014.

Murua H., Kraus G., Saborido-Rey F., Witthames P.R., Thorsen A., Junquera S. "Procedures to estimate fecundity of marine fish species in relation to their reproductive strategy". Northwest Atl. Fish. Sci. 33, (2003): 33–54.

Murua H. and Motos L. "Reproductive strategy and spawning activity of the European hake Merluccius merluccius (L.) in the Bay of Biscay". J. Fish Biol. 69, (2006): 1288–1303.

Poisson F. and Fauvel C. "Reproductive dynamics of swordfish (Xiphias gladius) in the southwestern Indian Ocean (Reunion Island). Part 1: oocyte development, sexual maturity and spawning". Aquat. Living Resour. 22, (2009): 45–58.

Rice, W.R. "Analyzing tables of statistical tests". Evolution 43, (1989): 223–225.

Riediger N.D., Othman R.A., Suh M., Moghadasian M.H. "A systemic review of the roles of n-3 fatty acids in health and disease". J. Am. Diet. Assoc. 109, (2009): 668–679.

Riveiro I., Guisande C., Maneiro I. and Vergara A.R. "Parental effects in the European sardine Sardina pilchardus". Mar. Ecol. Prog. Ser. 274, (2004): 225–234.

Sardenne, F., Bodin, N., Chassot, E., Amiel, A., Fouché, E., Degroote, M., Hollanda, S.J., Pethybridge, H., Lebreton, B., Guillou, G., Ménard, F. Trophic niches of sympatric tropical tuna in the Western Indian Ocean inferred by stable isotopes and neutral fatty acids. Progress in Oceanography, (2016).

Sargent J.R., Bell J.G., Bell M.V., Henderson R.J. and Tocher D.R. "Requirement criteria for essential fatty acids". J. Appl. Ichthyol. 11, (1995): 183–198.

Sargent J.R., Bell J.G., Bell M.V., Henderson R.J. and Tocher D.R. "The metabolism of phospholipids and polyunsaturated fatty acids in fish". Aquac. Fundam. Appl. Res. (1993): 103–124.

Sargent J.R., Tocher D.R., Bell J.G. and others. "The lipids". Fish Nutr. 3, (2002): 181–257.

Schaefer K. M. "Spawning time, frequency, and batch fecundity of yellowfin tuna, Thunnus albacares, near Clipperton". Fish. Bull. 94, (1996): 98Á113.

Schaefer K.M. "Reproductive biology of black skipjack, Euthynnus lineatus, an eastern Pacific tuna". Inter-Am. Trop. Tuna Comm. Bull. 19, (1987): 166–260.

Schaefer K.M. "Reproductive biology of yellowfin tuna (Thunnus albacares) in the eastern Pacific Ocean". Inter-American Tropical Tuna Commission. (1998).

Scott B., Marteinsdottir G. and Wright P. "Potential effects of maternal factors on spawning stock–recruitment relationships under varying fishing pressure". Can. J. Fish. Aquat. Sci. 56, (1999): 1882–1890.

Shelton A.O., Hutchings J.A., Waples R.S., Keith D.M., Akçakaya H.R. and Dulvy N.K. "Maternal age effects on Atlantic cod recruitment and implications for future population trajectories". ICES J. Mar. Sci. J. Cons. (1995).

Sunarya M.F. and Mulyani H. "The efect of season on fat content and fatty acid profile especially n-3 of yellowfin tuna. res. contrib. present". Ninth Sess. Work. Party Fish Technol. Mark. Cochin India 7-9 March 1994.

Tocher D.R. "Metabolism and functions of lipids and fatty acids in teleost fish". Rev. Fish. Sci. 11, (2003): 107–184.

Tocher D.R., Dick J.R., MacGlaughlin P. and Bell J.G. "Effect of diets enriched in Δ6 desaturated fatty acids (18: 3n- 6 and 18: 4n- 3), on growth, fatty acid composition and highly unsaturated fatty acid synthesis in two populations of Arctic charr (Salvelinus alpinus L.)". Comp. Biochem. Physiol. B Biochem. Mol. Biol. 144, (2006): 245–253.

Tomkiewicz J., Tybjerg L. and Jespersen AA. "Micro-and macroscopic characteristics to stage gonadal maturation of female Baltic cod". J. Fish Biol. 62, (2003): 253–275.

Trippel E.A. "Estimation of stock reproductive potential: history and challenges for Canadian Atlantic gadoid stock assessments". J. Northwest Atl. Fish. Sci. 25, (1999): 61–82.

Tulli F. and Tibaldi E. "Changes in amino acids and essential fatty acids during early larval rearing of dentex". Aquac. Int. 5, (1997): 229–236.

Vallin L. and Nissling A. "Maternal effects on egg size and egg buoyancy of Baltic cod, Gadusmorhua: implications for stock structure effects on recruitment". Fish. Res. 49, (2000): 21–37.

Wallace R.A. and Selman K. "Cellular and dynamic aspects of oocyte growth in teleosts". Am. Zool. 21, (1981): 325–343.

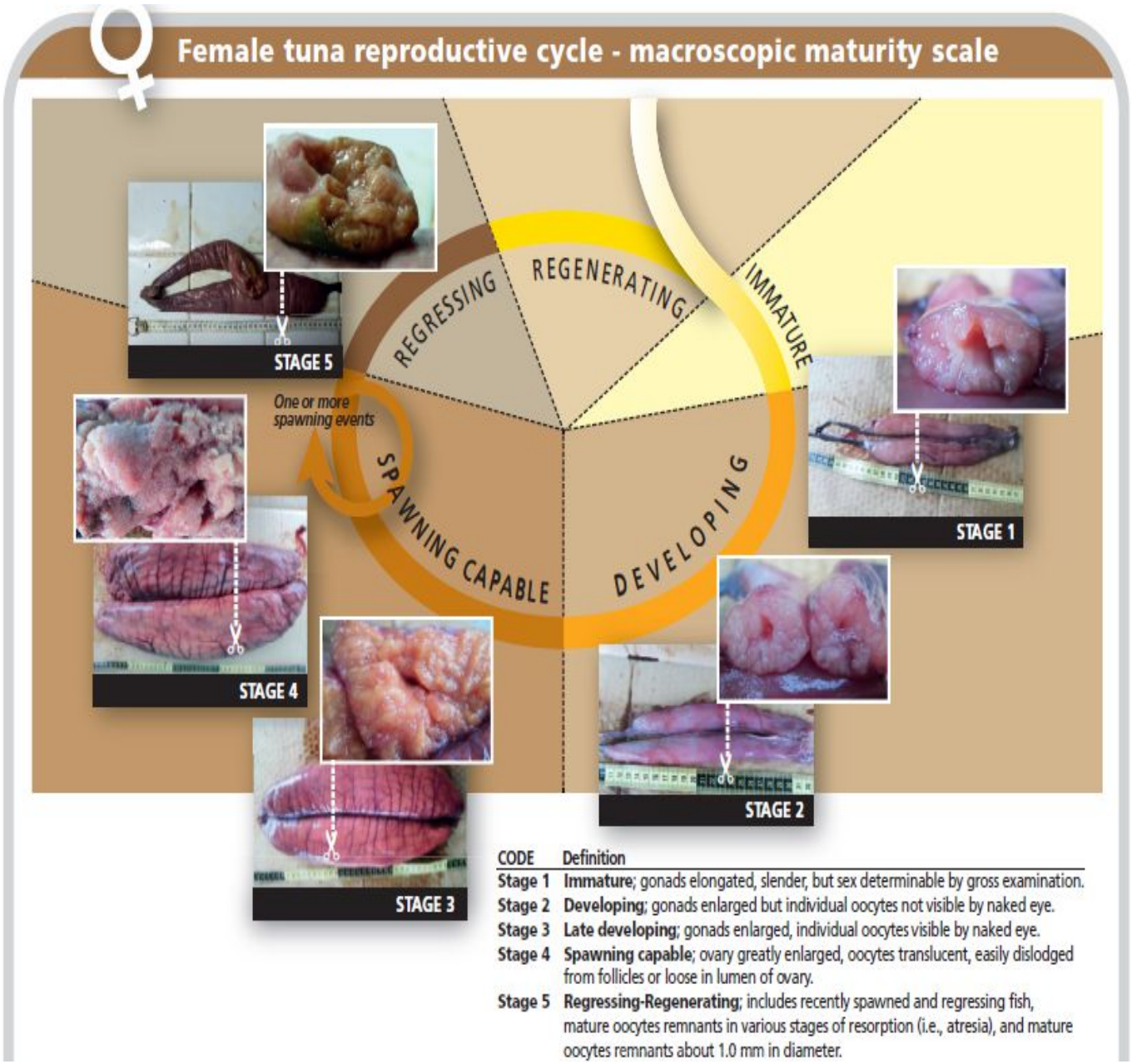
Wang S., Meyer E., McKay J.K. and Matz M.V. "2b-RAD: a simple and flexible method for genome-wide genotyping". Nat. Methods 9, (2012): 808–810.

Zudaire I., Murua H., Grande M. and Bodin, N. "Reproductive potential of yellowfin tuna (Thunnus albacares) in the western Indian Ocean". Fish. Bull. 111, (2013): 252–264.

Zudaire I., Murua H., Grande M., Pernet F. and Bodin, N. "Accumulation and mobilization of lipids in relation to reproduction of yellowfin tuna (Thunnus albacares) in the Western Indian Ocean". Fish. Res. (2014).

Appendix 6.1

Maturity stage scale for visual examination of large pelagic gonads (source: IOTC-2015-WPTT17-INO03).



Appendix 6.2

The table summerizes the contig's ID whose sequence was used as query in BLAST, the fatty acid associated with the SNP, the expected value of the alignment (e-value), the overall quality of the alignment (score), the number of identical matches (nident), the name of the protein that is similar with our query and the scientific name of the species in which it was found.

Contig	Fatty acid	e-value	Score	nident	Species	Protein
BADN01037938.1	C15:0 NL	2.16e-80	498	111	Maylandia zebra	PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like
BADN01077159.1	C15:0 NL	9.36e-123	489	101	Paralichthys olivaceus	reverse transcriptase-like protein
BADN01058733.1	C17:0 NL	4.89e-16	218	60	Neolamprologus brichardi	PREDICTED: paired mesoderm homeobox protein 1-like
BADN01010614.1	C18:0 NL	5.25e-44	323	66	Larimichthys crocea	Zinc finger BED domain-containing protein 1
BADN01020684.1	C18:0 NL	2.31e-26	312	57	Larimichthys crocea	hypothetical protein EH28_14086
BADN01088413.1	C18:0 NL	2.28e-08	169	32	Stegastes partitus	PREDICTED: G-protein coupled receptor family C group 5 member C isoform X2
BADN01091971.1	C18:0 NL	4.61e-35	262	45	Stegastes partitus	PREDICTED: presequence protease, mitochondrial
BADN01067369.1	C18:0 NL	1.30e-26	302	55	Tetraodon nigroviridis	unnamed protein product
BADN01020063.1	C18:0 PL	5.42e-32	362	114	Alpine marmot	PREDICTED: ankyrin-1-like
BADN01067004.1	C20:1w9 PL	9.21e-153	1287	322	Larimichthys crocea	hypothetical protein EH28_08179
BADN01023737.1	C20:1 w9 PL	0.0	1714	326	Larimichthys crocea	Neuropeptide Y receptor type 2
BADN01066581.1	C24:1w9 PL	4.10e-14	209	38	Larimichthys crocea	Hypothetical protein EH28_23722
BADN01045269.1	C24:1w9 PL	8.19e-58	325	76	Larimichthys crocea	PREDICTED: zinc finger BED domain-containing protein 1-like
BADN01022523.1	C18:3w3 PL	5.94e-41	296	65	Strongylocentrotus purpuratus	PREDICTED: uncharacterized protein LOC592034
BADN01068792.1	C15:0 PL	3.75e-54	478	84	Lycodichthys dearborni	CR1-3
BADN01077504.1	C15:0 PL	3.72e-12	193	37	Larimichthys crocea	Potassium voltage-gated channel subfamily KQT member 1
BADN01047352.1	C15:0 PL	2.46e-22	274	64	Larimichthys crocea	Protein FAM150B
BADN01071779.1	C15:0 PL	4.37e-120	941	223	Larimichthys crocea	putative DNA repair and recombination protein RAD26-like protein
BADN01038820.1	C15:0 PL	8.42e-51	180	32	Salmo salar	ORF2 protein
BADN01057259.1	C16:0 PL	5.42e-32	352	85	Oreochromis niloticus	PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like
BADN01057259.1	SUM.SFA PL	5.42e-32	352	85	Oreochromis niloticus	PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like

BADN01047352.1	C16:1w7 PL	2.46e-22	274	64	Larimichthys crocea	Protein FAM150B
BADN01045901.1	C16:1w7 PL	4.08e-38	426	76	Larimichthys crocea	PREDICTED: cadherin-2
BADN01036392.1	C24:1w9 PL	3.21e-47	484	97	Xenopus (Silurana) tropicalis	PREDICTED: general transcription factor II-I repeat domain-containing protein 2-like
BADN01047439.1	C24:1w9 PL	4.33e-16	219	45	Austrofundulus limnaeus	PREDICTED: receptor activity-modifying protein 3-like isoform X2 [Austrofundulus limnaeus]
BADN01043809.1	C24:1w9 PL	2.32e-13	200	42	Oncorhynchus mykiss	unnamed protein product
BADN01096355.1	C24:1w9 PL	8.04e-23	289	53	Larimichthys crocea	PREDICTED: vasoactive intestinal polypeptide receptor 2
BADN01003838.1	C24:1w9 PL	2.16e-80	498	111	Maylandia zebra	PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like
BADN01044590.1	C24:1w9 PL	1.17e-34	377	69	Austrofundulus limnaeus	PREDICTED: uncharacterized protein LOC106522257
BADN01092270.1	C24:1w9 PL	1.72e-19	248	46	Salmo salar	PREDICTED: arf-GAP with GTPase, ANK repeat and PH domain-containing protein 1-like, partial
BADN01005310.1	C24:1w9 PL	1.71e-20	176	43	Oreochromis niloticus	PREDICTED: uncharacterized protein LOC102077105
BADN01034974.1	C24:1w9 PL	1.20e-37	392	90	Oryzias latipes	PREDICTED: LOW QUALITY PROTEIN: uncharacterized protein LOC105355317, partial
BADN01086105.1	C24:1w9 PL	2.11e-07	161	41	Anguilla japonica	reverse transcriptase
BADN01063047.1	C24:1w9 PL	9.38e-35	276	61	Notothenia coriiceps	PREDICTED: zinc finger protein 346-like
BADN01017698.1	C20:3w3 PL	3.52e-47	505	117	Oreochromis niloticus	PREDICTED: integrin alpha-L-like isoform X2
BADN01095731.1	C20:3w3 PL	2.67e-25	186	36	Larimichthys crocea	Thiamine transporter 1
BADN01022467.1	C20:3w6 PL	6.68e-55	351	77	Larimichthys crocea	Retrovirus-related Pol polyprotein LINE-1
BADN01073637.1	C20:3w6 PL	5.34e-47	296	63	Tetraodon nigroviridis	unnamed protein product, partial
BADN01054561.1	C20:3w6 PL	1.21e-33	359	76	Xenopus (Silurana) tropicalis	PREDICTED: 41 kDa spicule matrix protein-like
BADN01073809.1	C20:3w6 PL	5.51e-50	480	91	Xenopus (Silurana) tropicalis	TPA: putative transposase
BADN01014200.1	C20:3w6 PL	8.68e-11	194	41	Cyprinodon variegatus	PREDICTED: uncharacterized protein LOC107091067
BADN01068319.1	C20:3w6 PL	2.37e-12	207	40	Cricetulus griseus	putative ubiquitin carboxyl-terminal hydrolase FAF-X isoform 2
BADN01071796.1	C20:3w6 PL	2.39e-41	232	69	Lycodichthys dearborni	CR1-3
BADN01006921.1	C20:3w6 PL	8.43e-55	536	133	Oncorhynchus mykiss	unnamed protein product
BADN01073732.1	C20:3w6 PL	6.91e-53	510	165	Larimichthys crocea	PREDICTED: protein naked cuticle homolog 2 isoform X1
BADN01017707.1	C20:3w6 PL	1.17e-61	569	115	Larimichthys crocea	PREDICTED: protein phosphatase 1 regulatory subunit 27-like
BADN01062649.1	C20:3w6 PL	8.99e-13	201	36	Dicentrarchus labrax	Noggin-2
BADN01090510.1	C20:3w6 PL	6.58e-07	156	30	Clupea harengus	PREDICTED: uncharacterized protein LOC105899288

BADN01079351.1	C20:3w6 PL	1.46e-37	388	104	Tetraodon nigroviridis	unnamed protein product, partial
BADN01113927.1	C20:3w6 PL	2.22e-98	731	157	Pundamilia nyererei	PREDICTED: LOW QUALITY PROTEIN: probable imidazolonepropionase
BADN01057629.1	C20:3w6	2.07e-22	284	55	Larimichthys crocea	PREDICTED: U1 small nuclear ribonucleoprotein 70 kDa
BADN01023286.1	C20:3w6	1.05e-11	182	32	Tetraodon nigroviridis	unnamed protein product
BADN01036971.1	C20:3w6	1.57e-22	271	52	Neolamprologus brichardi	PREDICTED: glutamate receptor ionotropic, kainate 4-like, partial
BADN01042171.1	C20:3w6	1.73e-23	301	59	Stegastes partitus	PREDICTED: anoctamin-4
BADN01027793.1	C20:3w6	6.25e-122	1105	281	Larimichthys crocea	PREDICTED: LOW QUALITY PROTEIN: potassium voltage-gated channel subfamily H member 5-like
BADN01046024.1	C20:3w6	1.21e-47	287	56	Danio rerio	PREDICTED: zinc finger BED domain-containing protein 4-like

7 Conclusions

Yellowfin tuna (*Thunnus albacares*; YFT) represents one of the most important seafood commodities in the world. Despite the substantial social and economic importance of YFT fisheries and the significant decrease of catches in various regions around the world (i.e. recently in Indian Ocean), there are still fragmented and controversial information regarding the ecology and biology of this important biological, natural and economic resource. In **Chapter 1**, all those aspects have been widely examined in depth for delineating the current state of knowledge and future research directions and improvements required for the governance of YFT in the tRFMOs. Even though a proper fish stock management needs accurate knowledge on the stock structure and its genetic variation, with respect to environmental and ecological conditions, it was emphasised that there are still more doubts than certainties about YFT genetic population structure both at the intra- and inter-oceanic level. Failing to detect population structure, due to limited genetic resolution of classical markers, can potentially be misleading for management purposes, driving to local overfishing and severe stock decline with profound implications for YFT management and conservation.

This general pattern was confirmed in **Chapter 2**, in which our results, mainly due to the inadequate resolution power of the neutral markers employed, cannot reject the null hypothesis of the existence of only one panmictic population at the global scale. In such a context, there is an evident need for developing alternative approaches based on genomics, that allow screening a larger number of markers across the entire genome, including neutral and non-neutral loci. This might enable resolving YFT population structure, quantifying the extent of spatial demographic changes and discover imprints of local adaptation, which represent priority focus for implementing effective management plans.

In **Chapter 3**, after having examined the utility of Technical Replicates (TRs) for optimizing genotyping procedure, I tested the potential of 2b-RAD genotyping approach in investigating YFT population genetic structure. Our results confirmed the great utility of TRs for optimizing genotyping procedure and even more importantly it was demonstrated the high potential of 2b-RAD in screening a large set of genomic loci in a high gene-flow marine species. Besides, we discovered the possibility to unambiguously map the TRs' tags against the reference genome of *T. orientalis* with a high percentage of success (86,59%), in spite of the small size of fragments and the evolutionary distance between these two species. Although our results highlighted the great advantage of using a large amount of molecular markers to detect genetic differentiation among

YFT populations however those estimates needed to be addressed and confirmed increasing the spatial distribution and size of the geographical samples employed.

This was the main objective of **Chapter 4**, in which the sample size of the samples employed was increased (N=500) as well as their spatial distribution, covering the entire species distribution. On the one hand, our results confirmed the existence of genetic structuring among the three main oceanic basins, with the pairwise F_{ST} values calculated from genome-aligned markers ranging from 0.036 to 0.136. On the other hand, we pointed out that sometimes just multiplying the number of neutral markers is not enough to reveal the proper level of structuring that occurs in these highly migratory marine species. In fact, it was necessary to discover a subset of outlier loci putatively under selection to delineate and separate locally adapted sub-populations within the Atlantic and the Pacific Oceans (east-west division). Our results showed a higher level of structuring in the Atlantic Ocean than generally assumed by ICCAT. These sub-populations should be assessed and, potentially, managed independently, re-examining previous biological assumptions about the YFT productivity and resilience to the fishing pressures in this ocean.

However, the stock productivity is traditionally measured by estimating the spawning stock biomass (SSB), which implies that the quality and thus the survival rates of offspring are independent from parental age, body size, or condition and that the total relative egg production per unit weight of adult stock is invariant over time. In **Chapter 5**, we raised many doubts about the use of SSB as a proxy of YFT reproductive potential. In fact, we emphasized for the first time that larger YFT females allocate a greater fraction of surplus energy to egg production than smaller ones, improving noticeably the spawning quality. This result shed light on the important contribution that larger and most experienced spawners have for the YFT productivity in the Atlantic Ocean.

Within this scenario, it results evident that re-evaluating YFT population structure by means the use of genomic approaches and protecting highly fecund large females could thus represent effective strategies for managing YFT fishery in a more sustainable way.

ACKNOWLEDGEMENTS

First of all, I want to express my deepest gratitude to Fausto Tinti for having given me the freedom to conduct this project on my own and at the same time the guidance to recover when my steps faltered. I would like to thank Nathalie Bodin and Manu Chassot for their supervision and especially because without them this work would have been impossible. Their passion for science has been a souce of inspiration and support for me. Thanks also to Alessia Cariani for her contribution.

Many thanks to Hilario Murua and Luca Bargelloni for welcoming me into their laboratory at AZTI Tecnalia and at the University of Padova, respectively. Many thanks to Ancit and Bolton for their financial and logistic support over the past three years.

I am also deeply indebted to Max Babbucci, Iker Zudaire and Manuelino Romeo for their huge help and great friendship. I will never forget those great moments we had together. My most sincere thanks to Rafaella Franch for having taught me how to prepare 2bRAD libraries and for having provided valuable advice and guidance over the past two years.

Many thanks to Giuliano Galimberti for his big help with the statistical analysis and to Jeff Muir for his awesome pictures.

I would like to express my warm and sincere thanks to all the friends met in Seychelles, Abidjan, San Sebastian, Padova and Ravenna. Thanks to the Unibo crew and especially to Agostino Leone and Adriana Villamor for the nice moments spent together.

Grazie mille ad Angi per la sua amorevole pazienza e per avermi supportato/sopportato durante questi ultimi tre anni…ce l'abbiamo fatta.

Un abbraccio ai miei genitori, a mia sorella e a tutta la mia famiglia dalla Sicilia a Berlino per il vostro supporto nella mia vita.

Questo lavoro è dedicato ai miei nipotini Ali e Lori con l'augurio di un brillante futuro insieme.