Species richness, genetic diversity and preliminary population structure of common snappers and groupers in Aruban waters, estimated through DNA barcoding

Christian Romero, Susana Caballero

Abstract

Overfishing is a great threat to marine ecosystems and the people who depend on them. Even artisanal fishing of high intensity can lead to reduction of fish biomass and loss of coral cover. Aruba is a Caribbean island highly dependent on its marine ecosystem for its tourism industry, yet there is little to no published data on the current state of its fish stocks. Here, we use DNA sequencing of the cytochrome oxidase I gene to identify the species of snappers and groupers commonly caught by local artisanal fishermen, and research preliminary genetic information for these species in Aruban waters, including genetic identification, genetic diversity and population structure. Most fish caught were above maturity length (L_m), but all fish identified as *Lutjanus vivanus* were not. One of the groupers sampled was identified as *Mycteroperca xenarcha,* a species that has only been reported in the Pacific Ocean; its possible presence in the Atlantic must be researched further. We found significant population structure for *Lutjanus synagris* between Brazil and Aruba, Belize, Colombia, Florida and Mexico. This contradicts the results of previous studies, highlighting the need for more research and sampling around the Caribbean and Atlantic. Finally, the most commonly caught grouper in Aruba is *Hyporthodus flavolimbatus*, a vulnerable species of which little is known, but that shows a high level of local genetic diversity.

Keywords

Population structure, genetic diversity, fisheries, cytochrome oxidase, barcode gene, haplotype

Introduction

Fish are an essential resource for millions of people worldwide, providing food and economic security for many communities, especially in developing nations (Hauge *et al*, 2009). Even so, there is clear evidence that many commercially important fish species are not harvested sustainably. The rate of collapsed fisheries has accelerated since the 1950s, with 29% of global fish stocks considered collapsed in 2003, and a total collapse of all commercially fished species by 2048 has been predicted should this tendency continue (Worm *et al*, 2006). Although these collapses (defined as the loss of 90% or more of the original biomass) rarely result in total extinction (Le Pape *et al*, 2017), they can still have disastrous effects on marine ecosystems and the people that depend on them (Worm *et al*, 2006; Hauge *et al*, 2009). Overfishing is currently the greatest threat to marine recourses from coral reefs, along with habitat degradation caused by climate change (Wilson *et al*, 2010).

Aruba is a small Caribbean island with an economy nearly completely dependent on tourism, mostly concentrated on the white-sand beaches and clear waters of the island (Derix, 2016).

Thus, protecting the island's coastal ecosystems is of outmost importance. There is a strong link between overfishing and loss of coral reef health (Kroeker, 2016). Loss of biodiversity in marine ecosystems leads to decreased water quality, resource stability and recovery potential (Worm *et al,* 2006) and overfishing can affect whole ecosystems (Zhou *et al,* 2015). Although there is no real commercial fishery in Aruba, it has been shown that intense artisanal fishing can decrease the biomass of important fish taxa for humans, decrease coral cover and complexity, and lead to algae overgrowth (Hawkins & Roberts, 2004). On the other hand, local fishermen argue that recent legislation prohibiting certain species was done without proper studies and for species that are not under any threat. Currently, there are little to no published records available on the intensity of fishing in Aruba. For all the aforementioned reasons it is important to start researching Aruban fish stocks.

Genetic information is of great importance in making the best conservation decisions possible (DeSalle & Amato, 2004). The key role of genetic diversity in species and ecosystem conservation is widely recognized (Hoban *et al,* 2013), while knowledge of population genetics can help minimize loss of this diversity and understand interconnectivity among populations (DeSalle & Amato, 2004). Here, we use DNA sequencing of the cytochrome oxidase subunit 1 gene (COI) to genetically identify snappers and groupers species caught by artisan fishermen in Aruban waters, and to estimate genetic diversity, and initial population structure with other locations in the Caribbean. Although the COI gene can be used to infer population dynamics, the results obtained must be considered preliminary (Escobar *et al*, 2019).

Methods

Sampling and DNA extraction

With the permission of Aruban fishermen, samples were collected on June 30th 2019 in Aruba, after the weight-in of an artisanal fishing tournament. Additional samples were collected through the following days from recently-caught local snappers and groupers. Small, white muscle pieces were obtained from fish that have been preliminary identified as snappers (Lutjanidae) or groupers and stored in 1,5 mL Eppendorf tubes with absolute ethanol. We measured the fork length and total length of the sampled fish and identified their species tentatively by the morphological characteristics mentioned in Fishbase, mainly color, body shape and size. All samples were stored at 4 °C until DNA extraction.

DNA was extracted with the Chelex protocol of Cardeñosa *et al* (2018). Small pieces of the samples (~2mm) were cut and placed in 200µL of 10% Chelex solution. The samples in solution were heated in a thermocycler to 60 °C for 20 minutes, then 99 °C for 25 minutes, before being cooled to 4 °C for storage. The successful extraction of DNA was verified through a Nanodrop 2000C instrument.

PCR amplification and sequencing

We used two PCR protocols to amplify the COI gene: a modified protocol using ThermoFisher MyTaq DNA polymerase (Reaction mix per tube: 16,8µL H_2O; 5,0 µL MyTaq Reaction Buffer; 1,0 µL of forward and reverse primers at 10 mM; 0,2 µL MyTaq DNA Polymerase; 1,0 µL template DNA) and a standard PCR protocol using generic taq polymerases (Reaction mix per tube: 16,6 µL H_2O; 2,5 µL Reaction Buffer; 1,0 µL of forward and reverse primers at 10 mM; 0,2 µL dNTPs; 0,2 µL Taq polymerase; 1,0 µL template DNA). In all cases, the cycling conditions were the following: initial denaturation for 2 minutes at 92 °C, denaturation at 94 °C, annealing at 55 °C, and extension at 72 °C for 60 s in each case for 35 cycles, and a final extension at 72 °C for ten minutes. Ee used the universal primers for the COI gene of fish by Ward *et al* (2005). Successful amplification was verified by electrophoresis in 1.5% agarose gel (run conditions: 80 V for 50 minutes). Expected amplicon size was around 700 pb. PCR products were purified using magnetic beads or enzymatic methods and sequenced at Gencore (Universidad de los Andes) using an ABI 3500 sequencer.

Data analyses

We visualized the raw chromatograms using FinchTV and eliminated low quality regions. The resulting sequences were compared using BLAST against the NCBI GenBank Database. Identifications were compared to the previous morphological identification. Additional COI sequences from the species identified were downloaded from the NCBI GenBank database. For each species with sample sizes ≥5, sequences were aligned with MUSCLE, and those alignments were visualized with Jalview. Genetic diversity and genetic structure was analyzed using Arlequin V 3.5 to calculate nucleotide (π) and haplotype (h) genetic diversity, as well as F_{st} and Φ_{st} values between Aruban fish and other fish sampled throughout the Caribbean whose COI sequences are available on NCBI GenBank. The data files for analysis with Arlequin were created with DNAsp6 and edited as necessary. Haplotype networks between locations were constructed using TCS V. 1.21 with standard parameters (Clement *et al,* 2000). Finally, we used the measurements taken during sample collection to see how many of the caught fish were juveniles, following the information available in Fishbase.

Results

Species richness and juvenile capture

A total of 105 samples were collected: 81 belonging to snappers (Lutjanidae) and 23 to groupers. One sample was discarded since it belonged to a jack (Caranxidae). We obtained 35 COI sequences from the remaining 104 samples. Table 1 shows the identified species, and the

percentage of individuals above maturity length (L_m) according to Fishbase. Unfortunetly, this has not been reported for *Hyporthodus flavolimbatus* nor *Mycteroperca xenarcha*. In the case of *Rhomboplites aurorubens* an L_m of 20 cm SL (Standard Length) is reported in Fishbase. Since we did not measure SL, we cannot make an effective comparison.

Species	***Sequences (n=35)***	***L_m (cm)***	***L avg (cm)***	***% < L_m***
Lutjanus synagris	8	23.6 FL	31.32 FL	0
Hyporthodus flavolimbatus	8	?	53.38 FL	-
Rhomboplites aurorubens	7	?	27.64 FL	-
Lutjanus vivanus	5	50 FL	34.7 FL	100
Lutjanus apodus	2	25 FL	34.27 FL	0
Lutjanus purpureus	1	43 TL	44 TL	0
Lutjanus griseus	1	28 FL	29 FL	0
Hyporthodus niveatus	1	54 TL	49 TL	100
Seriola drumerili	1	99.5	-	-
Mycteroperca xenarcha	1	?	53 FL	-

Table 1: Species found in our samples, identified through COI sequences, as well as their corresponding maturation length (L_m), average length in our samples (L avg), and percentage of individuals smaller than L_m (% < L_m). The specimen in red was actually a grouper and this result was dismissed as cross contamination. All measurements are in cm and are of either fork length (FL) or total length (TL)

One sample sequence belonging to a grouper (tentatively identified as a Warsaw grouper) was identified as *Seriola drumerili* (Caranxidae), according to BLAST. We discarded this sequence, as this result was probably due to contamination during sampling. Additionally, one sample tentatively identified as a black grouper (*Mycteroperca bonaci*) had a COI sequence 539 pb long matching to *Mycteroperca xenarcha* (ID: 99.81%, E value: 0.0). As of August 2020, there are no records of *M. xenarcha* anywhere in the Atlantic Ocean.

Genetic diversity and population structure

We analyzed population structure for *L. synagris* and *L. vivanus,* comparing the sequences generated for Aruba to sequences obtained from previous studies in other regions of the Caribbean. We also calculated genetic diversity at the nucleotide (π) and haplotype (h) level for those species, along with *R. aurorubens* and *H. flavolimbatus* caught in Aruban waters.

Population structure and genetic diversity values for *L. synagris* and *L. vivanus* are shown in table 2A and 2B, respectively. For *L. vivanus*, the only place with more than five COI sequences was St. Thomas. We found no statistically significant population differentiation compared to Aruba. Aruba appeared to have a higher genetic diversity than St. Thomas, but there was an overlap between the standard deviation range; this was true for both π and h. For *L. synagris*, there was significant population differentiation ($p<0,05$) at both F_{st} and Φ_{st} levels between Brazil and all other locations, except St. Thomas. Brazil had the lowest genetic diversity of all

locations, and Florida the highest. However, once again the standard deviations of all genetic diversity values overlap with each other.

L. vivanus	Aruba	St Thomas
Aruba	π = 0.1181% ± 0.1294% *h* =0.6000 ± 0.1753	F_{st} = 0.06613
St Thomas	Φ_{st} = 0.22372	π = 0.0937% ±0.1046% *h* = 0.4762± 0.1713

Table 2A: Genetic structure and diversity for *Lutjanus vivanus* in Aruba and St. Thomas. Nucleotide (π) and haplotype (h) genetic diversity for each site is shown as well as standard deviation (± SD). Population structure values are given as Φ_{st} and F_{st}. Statistically significant values of population structure (AMOVA, $p < 0.05$) are indicated with an asterisk.

L. synagris (n= 101)	Aruba (n= 8)	Belize (n= 6)	Brazil (n= 46)	Colombia (n= 18)	Florida (n= 5)	Mexico (n= 10)	St Thomas (n= 8)
Aruba (n= 8)	*π = 0.1360% ± 0.1311%* *h = 0.6071 ± 0.1640*	F_{st} = -0.10599	F_{st} = 0.43778*	F_{st}= 0.01269	F_{st} = -0.13208	F_{st} = -0.06400	F_{st} = -0.07143
Belize (n= 6)	Φ_{st}= -0.08615	*π = 0.1737% ± 0.1622%* *h* = 0.7333 ± 0.1552	F_{st} = 0.60666*	F_{st} = 0.11166	F_{st} = -0.13230	F_{st} = 0.00483	F_{st} = -0.01659
Brazil (n= 46)	Φ_{st} = 0.34898*	Φ_{st} = 0.50338*	π = 0.0087% ± 0.0250% *h* = 0.0435 ± 0.0412	F_{st} = 0.10773*	F_{st} = 0.50995*	F_{st} = 0.24095*	F_{st} = 0.25663
Colombia (n= 18)	Φ_{st} = -0.00081	Φ_{st} = 0.06494	Φ_{st} = 0.08642*	π = 0.0642% ± 0.0763% *h* = 0.3072 ± 0.1316	F_{st} = 0.01012	F_{st} = -0.06282	F_{st} = -0.05671
Florida (n= 5)	Φ_{st} = -0.05205	Φ_{st} = -0.06339	Φ_{st} = 0.43273*	Φ_{st} = 0.05548	π = 0.2405% ± 0.2128% *h* = 0.7000 ± 0.2184	F_{st} = -0.06762	F_{st} = -0.09453
Mexico (n= 10)	Φ_{st} = -0.08228	Φ_{st} = -0.05116	Φ_{st} = 0.26631*	Φ_{st} = -0.04682	Φ_{st} = -0.02041	π = 0.0713% ± 0.0851% *h* = 0.3556 ± 0.1591	F_{st} = -0.09056
St Thomas (n= 8)	Φ_{st} = -0.04762	Φ_{st} = -0.01302	Φ_{st} = 0.18629	Φ_{st} = -0.04118	Φ_{st} = -0.02914	Φ_{st} = -0.06582	π = 0.1002% ± 0.1077% *h* = 0.4643 ± 0.2000

Table 2A: Genetic structure and diversity for *Lutjanus synagris* across the Caribbean. Nucleotide (π) and haplotype (h) genetic diversity for each site is shown as well as standard deviation (± SD). Population structure values are given as Φ_{st} and F_{st}. Statistically significant values of population structure (AMOVA, $p < 0.05$) are indicated with an asterisk.

The haplotype networks constructed for *L. synagris* and *L. vivanus* are shown in figure 1A and 1B, respectively. Each place had one unique haplotype per species, except for *L. synagris* in Mexico. There was only one mutational step between haplotypes, except for haplotype G of *L. synagris* unique to Florida.

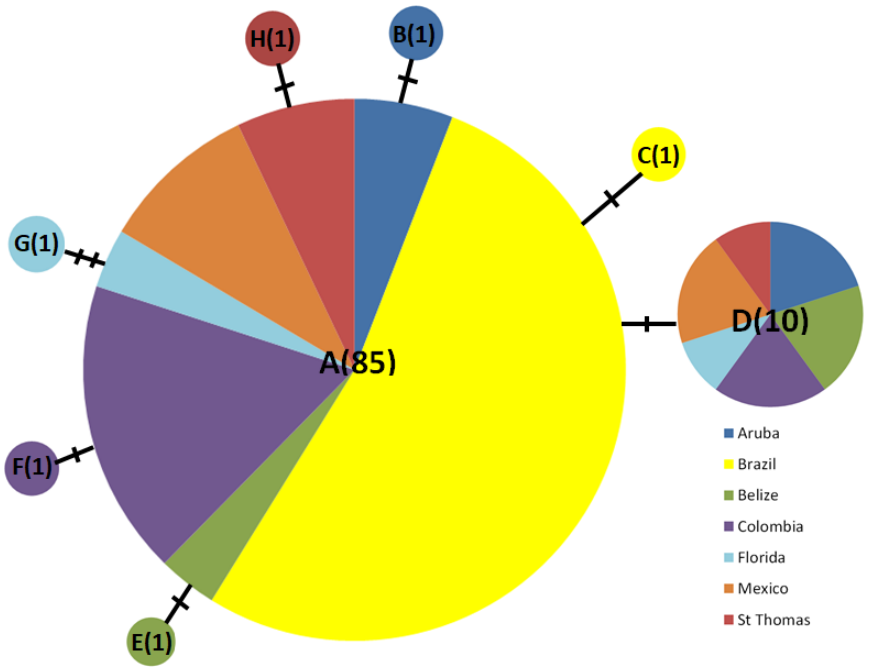

Fig. 1A: Haplotype network of *Lutjanus synagris* across the Caribbean. The number of mutational steps between haplotypes are represented by the horizontal bars and the number of samples with that haplotype is between brackets (n).

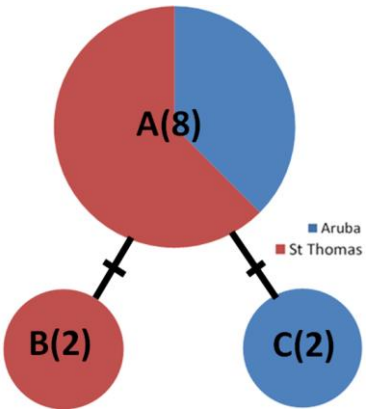

Fig. 1B: Haplotype network of *Lutjanus vivanus* for Aruba and St Thomas. The number of mutational steps between haplotypes are represented by the horizontal bars and the number of samples with that haplotype is between brackets (n).

The genetic diversity of *R. aurorubens* and *H. flavolimbatus* in Aruba was considerably higher than the one estimated for *Lutjanus* species. For *R. aurorubens*, π = 0.261% ± 0.048% and h = 0.810 ± 0.130. *H. flavolimbatus* presented alignment gaps. Without taking the loci with gaps into account, π = 0.285% ± 0.078% and h = 0.750 ± 0.139. The haplotype networks constructed for these species (Aruban samples only) are presented in figure 2A and 2B.

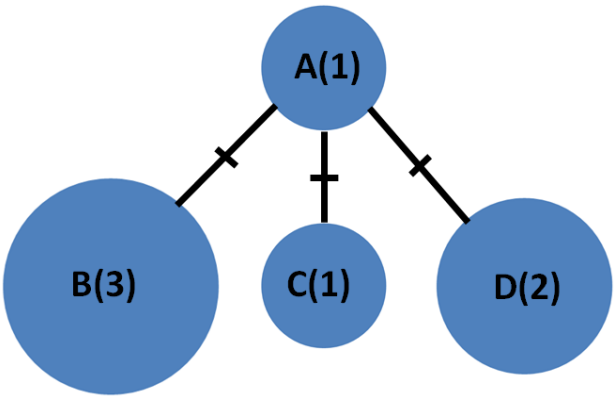

Fig. 2A: Haplotype network of *Rhomboplites aurorubens* (Aruban samples only). The number mutational steps between haplotypes are represented by the horizontal bars.

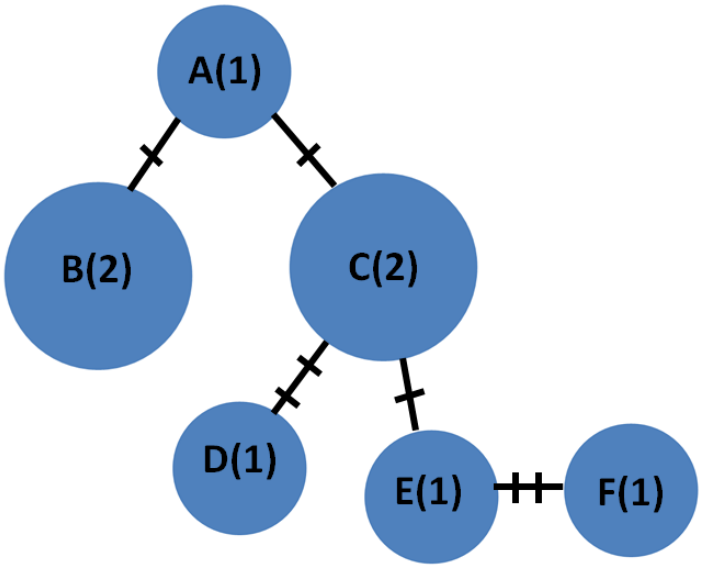

Fig. 2B: Haplotype network of *Hyporthodus flavolimbatus* (Aruban samples only). The number mutational steps between haplotypes, including indels, considered here as a fifth state, are represented by the horizontal bars.

Discussion

Species richness and juvenile capture

Molecular identification through COI sequencing revealed 6 species of snapper and 3 species of grouper in 34 samples. Most fish caught locally are above L_m. The clear exception is *L. vivanus*, in which case all individuals were below L_m. It is difficult to know how many individuals of *R. aurorubens*, the most common fish in our sampling, are mature, since the only reported L_m for this species is given in SL (standard length), a measurement we did not take. Moreover, there are

no reported L_m values for *H. flavolimbatus*, the most commonly-caught grouper in Aruban waters. Seeing that all *L. synagris* individuals were above L_m is encouraging, but a larger sample size must be taken in all cases to estimate the amount of caught juveniles more accurately. Considering that L_m values can vary considerably among locations, it is important to research the life cycle of various species more carefully, especially those that have no reported L_m values.

Surprisingly, one of the groupers sampled appears to belong to the species *Mycteroperca xenarcha,* which so far has only been reported in the Pacific Ocean (Heemstra & Randall, 1993; Froese & Paully, 2000). This specimen was initially identified as a black grouper, *Mycteroperca bonaci,* a species that is almost identically morphologically to *M. xenarcha,* both having extremely similar coloration, body shape, and the same number of fin rays and spines (Craig *et al,* 2011; Smith, 1997; Froese & Paully, 2000). Although Song *et al* (2008) showed that universal primers for barcode genes like COI can amplify similar pseudogenes, causing errors in estimating species' diversity, we do not believe this is the case here, considering the high sequence quality and match scores. However, we are not confirming the presence of *M. xenarcha* in the Atlantic Ocean. The specimen sampled was not preserved, since we did not want to take away fishermen's right to sell their catch, and we only have a single sequence. We do propose that *M. xenarcha* might inhabit the Atlantic Ocean and suggest future studies to confirm this possibility.

Genetic diversity and population structure

We found significant population structure between *L. synagris* individuals from Brazil and the other six locations, except St Thomas. This is different to the results of Escobar *et al* (2019), who found significant population structure for this species between Mexico, Florida, Brazil and Colombia, also using COI sequences. The differences between our studies might be due to the larger number of samples used here, and the fact that we used at least five sequences per location.

Furthermore, Landinez-García *et al* (2009) found very slight but significant population structure for this species when comparing 3 locations on the Colombian Atlantic coast using microsatellite data. It is possible that the level of genetic differentiation on this scale is too small to be detected without a large number of samples, which might explain why we did not find much genetic structure across much of the Caribbean. These different results highlight the importance of using a larger number of samples, even if population genetics statistics are often optimized for relatively small sample sizes and there is little gain in statistical power when n>50 (Hahn, 2018). Ultimately, using multiple genetic markers results in more robust results, and genomic techniques should be considered more often when researching population genetics.

As far as we know, this is the first study to look into the population structure of *L. vivanus.* Despite the differences in haplotypes in Aruba and St. Thomas, and a relatively high value for Φ_{st} (0.22373), there was no significant population structure. This might be due to the small

sample size. Also, there is very little data for this species compared to *L. synagris*. It is clear that more intense sampling across multiple locations is necessary to obtain truly robust results.

We did not find any locations with n>5 for *R. aurorubens* and *H. flavolimbatus,* so we looked at the local haplotype networks and genetic diversities. Bagley *et al* (1999) analyzed the population structure of *R. aurorubens* on the Atlantic coast of the USA, and found very little genetic differences along the region, concluding that there was only one genetic stock. If this pattern persists in the Caribbean, we would expect to find little to no population structure along the region; however, only research aimed specifically at this species in this region can confirm or disprove this possibility.

H. flavolimbatus is the most genetically diverse species in our survey and shows the most complex haplotype network. Curiously, it is also the only species to possess indels in a COI sequence. Although this high level of genetic diversity is encouraging, the lack of information regarding maturation size, the vulnerable conservation status of this species, and its long generational length of 46.5 years (Padovani-Ferreira *et al,* 2018) make further research imperative. This is the first study of its kind in Aruba and it is meant as a first stepping stone for more detailed research in this field. Hopefully, this will help the Aruban people in managing their marine resources in the near future.

References

Cardeñosa, D., Quinlan, J., Shea, K. H., & Chapman, D. D. (2018). Multiplex real-time PCR assay to detect illegal trade of CITES-listed shark species. *Scientific Reports, 8*(1). doi:10.1038/s41598-018-34663-6

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology, 9*(10), 1657-1659. doi:10.1046/j.1365-294x.2000.01020.x

Craig, M.T., YJ. Sadovy de Mitcheson and P.C. Heemstra, 2011. Groupers of the world: a field and market guide. North America: CRC Press/Taylor and Francis Group. 356 p.

Derix, R. R. (2016). *The history of resource exploitation in Aruba*(Aruba, Central Bureau of Statistics).

Desalle, R., & Amato, G. (2004). The expansion of conservation genetics. *Nature Reviews Genetics, 5*(9), 702-712. doi:10.1038/nrg1425

Escobar, R., Luna-Acosta, A., & Caballero, S. (2019). DNA barcoding, fisheries and communities: What do we have? Science and local knowledge to improve resource management in partnership with communities in the Colombian Caribbean. *Marine Policy, 99*, 407-413. doi:10.1016/j.marpol.2018.08.025

Froese, R. and D. Pauly. Editors. 2019. FishBase. World Wide Web electronic publication. www.fishbase.org

Hahn, M. W. (2019). Chapter 2: Experimental Design. In M. W. Hahn (Author), *Molecular population genetics* (pp. 25-42). New York: Sinauer Associates.

Hauge, K. H., Cleeland, B., & Wilson, D. C. (2009). *Fisheries Depletion and Collapse*(Rep.). Geneva, Switzerland: International Risk Governance Council.

Hawkins, J. P., & Roberts, C. M. (2004). Effects of Artisanal Fishing on Caribbean Coral Reefs. *Conservation Biology,18*(1), 215-226. doi:10.1111/j.1523-1739.2004.00328.x

Hoban, S. M., Hauffe, H. C., Pérez-Espona, S., Arntzen, J. W., Bertorelle, G., Bryja, J., . . . Bruford, M. W. (2013). Bringing genetic diversity to the forefront of conservation policy and management. *Conservation Genetics Resources,5*(2), 593-598. doi:10.1007/s12686-013-9859-y

Kroeker, K. J. (2016). Fishing for lessons on coral reefs. *Nature,535*(7612), 361-362. doi:10.1038/535361a

Landínez-García, R. (2009). Genetic analysis of *Lutjanus synagris* populations in the Colombian Caribbean. *Ciencias Marinas, 35*(4), 321-331. doi:10.7773/cm.v35i4.1240

Le Pape, O., Bonhommeau, S., Nieblas, A., & Fromentin, J. (2017). Overfishing causes frequent fish population collapses but rare extinctions. *Proceedings of the National Academy of Sciences,114*(31). doi:10.1073/pnas.1706893114

Padovani-Ferreira, B., Bertoncini, A.A. & Aguilar-Perera, A. 2018. *Hyporthodus flavolimbatus*. *The IUCN Red List of Threatened Species* 2018: e.T64400A46915591. https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T64400A46915591.en. Downloaded on 07 August 2020.

Smith, C.L., 1997. National Audubon Society field guide to tropical marine fishes of the Caribbean, the Gulf of Mexico, Florida, the Bahamas, and Bermuda. Alfred A. Knopf, Inc., New York. 720 p

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B, 360*(1462), 1847-1857.

Wilson, S. K., Fisher, R., Pratchett, M. S., Graham, N. A., Dulvy, N. K., Turner, R. A., . . . Polunin, N. V. (2010). Habitat degradation and fishing effects on the size structure of coral reef fish communities. *Ecological Applications,20*(2), 442-451.

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., . . . Watson, R. (2006). Impacts of Biodiversity Loss on Ocean Ecosystem Services. *Science,314*(5800), 787-790. doi:10.1126/science.1132294

Zhou, S., Smith, A. D., & Knudsen, E. E. (2014). Ending overfishing while catching more fish. *Fish and Fisheries,16*(4), 716-722. doi:10.1111/faf.12077