Supplementary Material on the Methodology Part VIII

Final report on blackspot seabream genetics

Coordinators of genetics: Alessia Cariani (UniBo), Kenza Mokhtar-Jamaï (INRH), Carolina Johnstone (IEO)

Pagellus bogaraveo genetics coordinated by UNIBO: Alessia Cariani, Alice Ferrari, Federica Piattoni, Martina Spiga

1. Introduction

Wild fishery resources currently represent a precious reservoir requiring a sustainable management strategy based on the knowledge of the target species at different scales (Ramírez et al., 2021; Schickele et al., 2021). In this perspective, the FAO-COPEMED II TRANSBORAN Project (Transboundary population structure of Sardine, European hake and Blackspot Seabream in the Alboran Sea and adjacent waters: a multidisciplinary approach) was created to evaluate the population structure and dynamics of the Blackspot Seabream, *Pagellus bogaraveo* Brünnich, 1768, across the Mediterranean Sea and with a special focus on the Alboran area. Due to its commercial importance, this valuable deep-water fish has been worryingly exploited in the past decade (Lorance 2011; Pinho et al., 2014; ICES 2018a; ICES 2019; Sanz-Fernandez et al., 2019). Thus, a stock delimitation plan based on scientific evidence is needed.

With TRANSBORAN, the fruitful cooperation established between Research Centres, Universities and local fishermen across the target area, allowed the creation of a baseline for the investigation of a) a correct spatial scale in which a sustainable management and assessment of fisheries can be implemented, b) the profitability of the fishery taking into account the resilience of marine communities and c) measures for the conservation of biodiversity.

The present report describes the application of molecular techniques to the Blackspot Seabream to investigate its population structure in terms of genetic differentiation across 14 sampling locations in five main areas, from the Atlantic Ocean, the Strait of Gibraltar, to the Western, Central and part of Eastern Mediterranean Sea. Here, 20 species-specific microsatellite loci were tested and applied across a geographical range covering almost all the species' dispersal area.

Moreover, the comparison of genetic results with the hydrodynamic modelling of the Strait of Gibraltar area (Sammartino et al., 2018) allowed to gather information on stock structure for resource management capturing different ecological and structuring processes likely acting at contrasting spatial and temporal scales (i.e., fine scale differences around the Gibraltar Strait, but also large-scale differences influencing the biogeography of the species).

2. Materials and methods

2.1. Sampling design

In this research Project, a total of 320 individuals were collected from commercial fishery, scientific surveys from the EU-funded MEDITS Project (Bertrand et al., 2002) carried out in 2018 and 2019 and from fish markets along the European and African coasts in order to obtain samples representative of the different fishing areas in the Atlantic Ocean and Mediterranean Sea and to cover most of *P. bogaraveo* dispersal area (complete list of all the individuals in Annex 1).

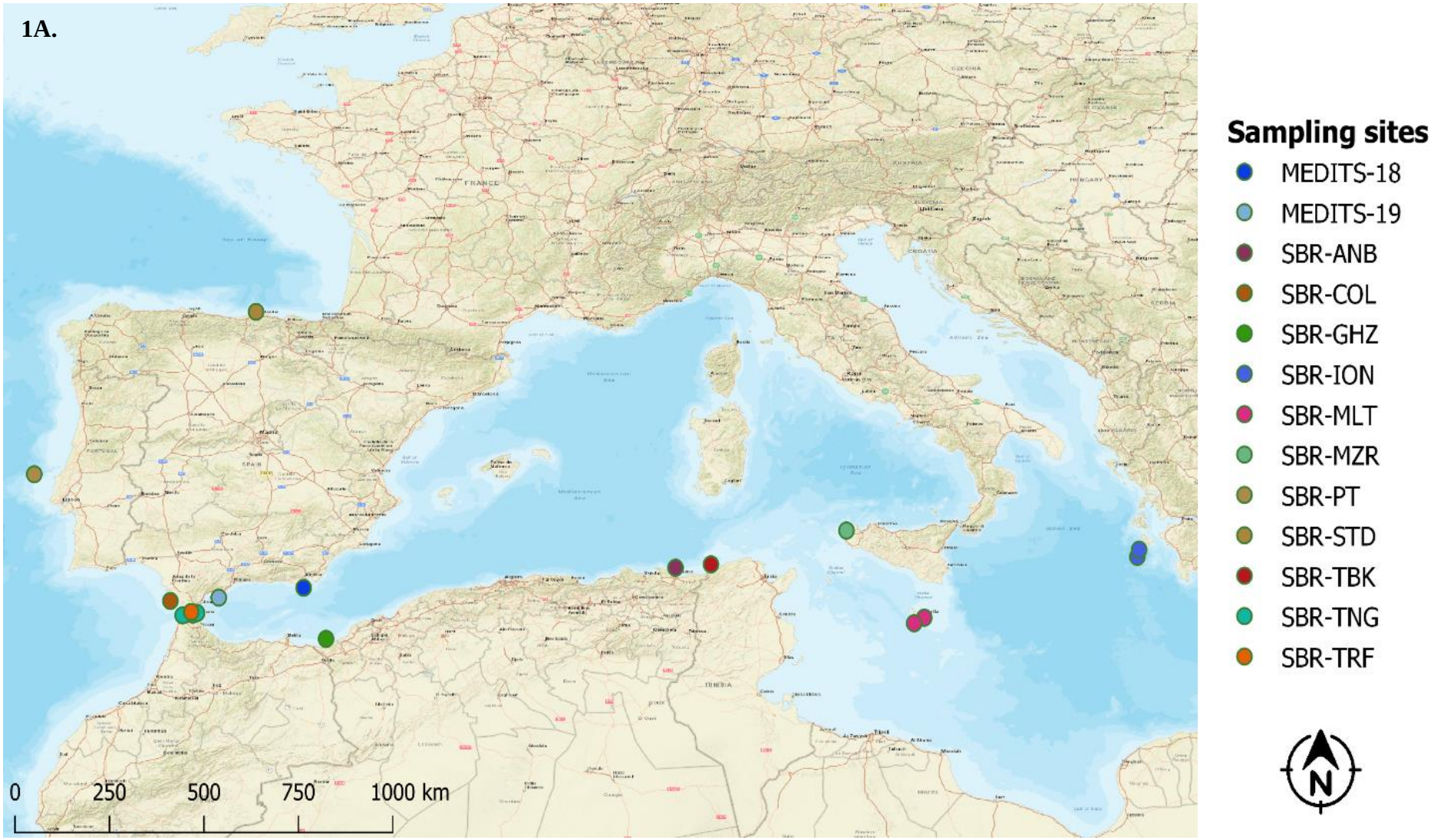
The initial TRANSBORAN project sampling design included nine different sampling ports. The sampling of Blackspot Seabream was challenging mainly for the lack of catches from the Spanish longline fishery around the Gibraltar Strait. Part of this lack was filled by other samples provided by establishing collaborative projects with Research Institutes from Portugal, Greece, Santander (Spain), Annaba (Algeria), Italy and from MEDITS survey (Spain), with the common interest in revealing a large-scale genetic structure of this target species.

The subdivision of the samples is described in Tab. 1 and the final localities are shown in Fig. 1.

Table 1. FAO major fishing areas, sampling location, number of samples and period of sampling.

FAO fishing areas	Country	Location Name	Location code	N. of Expected Individuals	N. of Received Individuals	Sampling Period
27.8.c	Spain	Santander	STD		26	November 2018
27.9.a	Portugal		PT		36	January-February 2019
27.9.a	Spain	Conil	COL	30	15	July 2018
27.9.a	Spain	Tarifa	TRF	30	3	July 2018
27.9.a	Morocco	Tanger	TNG	30	18	June 2018
27.9.a	Morocco	Ksar Sghir	TNG		15	June 2018
27.9.a	Morocco	Eddalya	TNG		35	October 2018
37.1.1	Spain	Algeciras	ALG	30	0	
37.1.1	Spain	MEDITS	SPA18		10	April-May 2018
37.1.1	Spain	MEDITS	SPA19		30	May-June 2019
37.1.1	Spain	Roquetas	RQT	30	0	
37.1.1	Algeria	Ghazaouet	GHZ	30	9	June 2018
37.1.1	Morocco	Nador	NDR	30	0	
37.1.1	Algeria	Annaba	ANB		30	July 2018
37.1.1	Tunisia	Tabarka	TBK	30	30	December 2018
37.1.3	Italy	Trapani	MZR		16	March 2019
37.2.2	Malta	Malta	MLT	30	4	September 2018
37.2.2	Malta	Malta	MLT		13	March 2019
37.2.2	Greece		ION		30	August 2018
			Total number of individuals:		320	

1A.


Sampling sites
MEDITS-18
MEDITS-19
SBR-ANB
SBR-COL
SBR-GHZ
SBR-ION
SBR-MLT
SBR-MZR
SBR-PT
SBR-STD
SBR-TBK
SBR-TNG
SBR-TRF
0
250
500
750
1000 km
N

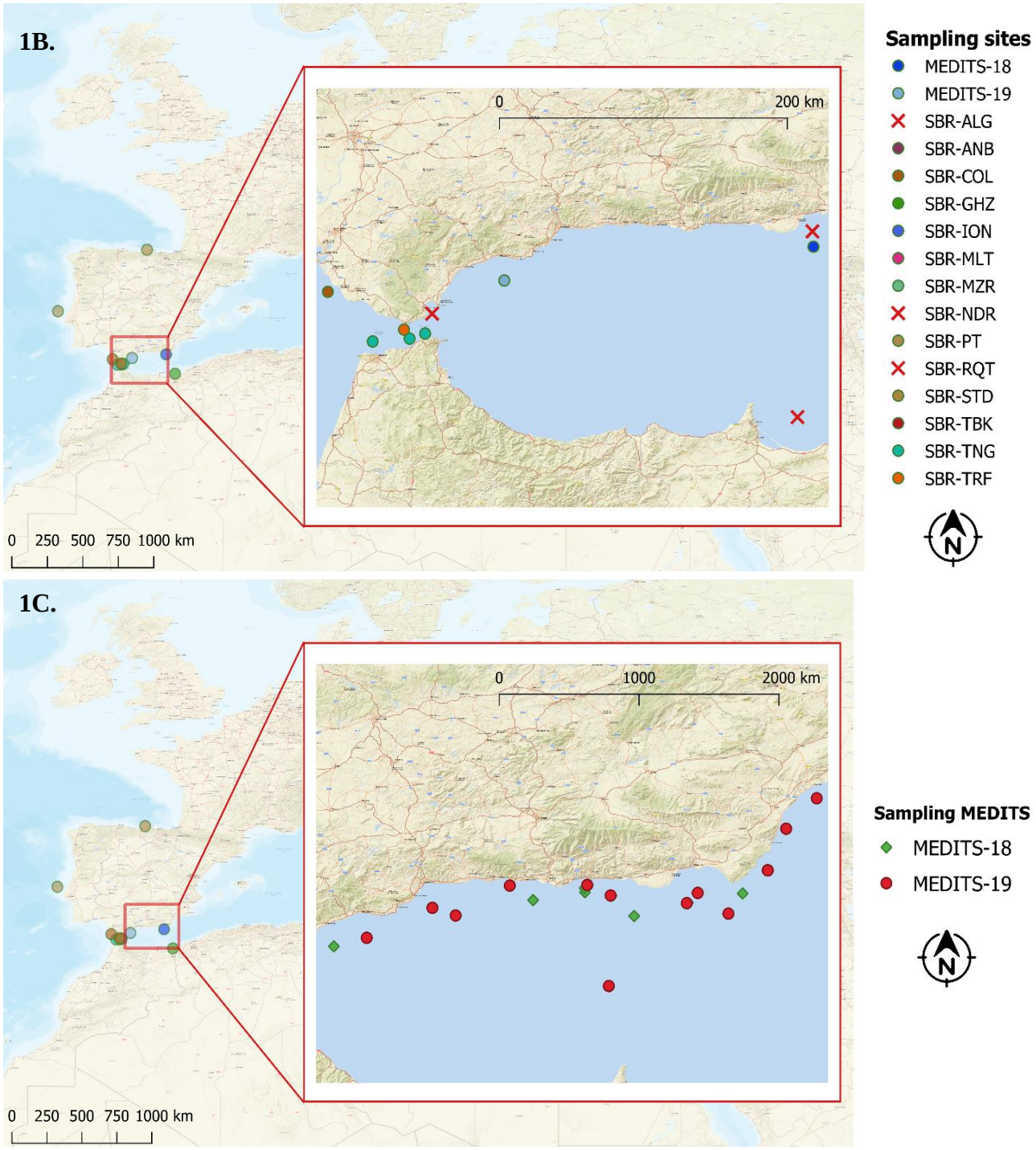


Figure 1. Maps of geographical distribution of analysed samples. A. Complete sampling ground; B. Details of Gibraltar strait and Alboran Sea samples. C. Details of MEDITS 2018 and 2019 samples. Qgis v. 3.4. Precise coordinates were available only for the MEDITS 2018 and MEDITS 2019 samples. Red crosses indicate expected sampling location where no samples were collected.

A detailed sampling protocol was implemented within the TRANSBORAN activities and distributed to Institutes and Research Centres. When needed, sampling materials were distributed as well. Due to the great commercial importance of the species, the collection of necessary individuals implied a great effort and collaborative work between all the TRANSBORAN project participants. The realized sampling, which allowed performing this study, was possible also thanks to the collaboration of additional Research Partners involved in this Project.

Details of all the individuals collected, biological data (length, weight and sex) complete with indication of sampling site (date, coordinates, zone and depth) were used to implement a punctual database of the Project and to set up the experimental design used in the following analyses.

2.2. DNA extraction

Muscle tissue (about 1 cm^3) was collected from fresh or frozen specimens and immediately preserved in 96% ethanol after collection. All the samples were shipped to the GenoDREAM laboratory in Ravenna, where either transfer of tissue samples into 2ml vial and replacement of ethanol were performed, sub-sampling of tissue was carried out and the label with the correct location code checked. The ID-labelled vials containing the tissue were stored at -20°C until shipping to a service provider (BMR Genomics) for the DNA extraction and the amplification of microsatellite loci.

Before DNA extraction procedure, tissues were pre-treated adding 600 µl of ASL buffer (Qiagen) and about 50-100 µl of zirconium beads (Zirconia/Silica Beads, 0.1mm) to a 2 ml vial with screw cap; samples were then vortexed using the Tissue Lyzer for 1 min at 25 Hz followed by 1 min at 25 Hz and finally, centrifuged at 14000 xg for 2 min. A total of 200 µl of the supernatant was then used as starting material to be aliquoted in the S-block well.

The DNA extraction was then carried out with the Qiacube cador pathogen kit for QIAcube HT System (cat N° 9001793).

Multiple DNA extractions were attempted for samples from Tangeri, using an improved protocol, in which approximately 0.02 g of tissue was homogenized in 100 µl lysis buffer and the DNA IQ™ System-Database Protocol from Promega, was applied. As a matter of fact, the extraction procedure used for the other samples, did not show enough genomic quality for the Tangeri ones, probably due to the low quality of the tissue itself.

2.3. Microsatellite loci

A total of 22 microsatellite loci previously isolated and characterized from *P. bogaraveo* by Stockley et al. (2000), using individuals from Azores area, and Pinera et al. (2006), using individuals from the European area, and seven loci isolated and characterized from *P. erythrinus* by Ramšak et al. (2003) were tested (refer to Annex 2 for the complete list of primers used). To test each locus, amplification trials were performed using 20 *P. bogaraveo* samples (from Tarifa, Conil and MEDITS-2018 survey) and four *P. erythrinus* samples, to test cross-amplification.

In total, six different multilocus reactions were designed based on loci annealing temperature and expected amplicon size, and the multiplex PCR tested (refer to Tab. 2): the first four contained only *P. bogaraveo*'s microsatellite loci and the last two contained the microsatellite loci developed for *P.*

erythrinus. The loci mixes were optimized in order to minimize the differences in annealing temperature and maximize the spacing between marker amplicons. The selection of different fluorescent labels allowed us to distinguish between fragments of overlapping microsatellite loci.

Table 2. Mix setup containing the 22 P. bogaraveo *and the 7* P. erythrinus *microsatellite loci.*

MIX A	**MIX B**	**MIX C**	**MIX D**	**MIX E**	**MIX F**
Pb-OVI-B2	Pb-OVI-A5	Pb-OVI-A3	PbMS 17	AY188339	AY188343
PbMS 2	Pb-OVI-D114	PbMS 20	PbMS 4	AY188340	AY188344
PbMS 6	Pb-OVI-D102	PbMS 18	PbMS 15	AY188341	AY188346
PbMS 16	Pb-OVI-D20	PbMS 1	Pb-OVI-C103	AY188342	
PbMS 19	Pb-OVI-D21	Pb-OVI-D22	Pb-OVI-D101		
Pb-OVI-D108	Pb-OVI-D106				

The microsatellite loci have been amplified by PCR and the separation of the products from different samples have been carried out by agarose gel electrophoresis. The PCR primers were labelled using NED, FAM, PET, VIC and LIZ fluorochromes.

The fluorescently labelled DNA fragments were mixed with the GeneScanner 500 LIZ internal size standard (used for fragments of 500 pb length).

2.4. Amplification conditions

PCRs were performed with the Platinum® Multiplex PCR Master Mix (Thermofisher Scientific cat Nr. 4464269) and oligonucleotides sequences provided by reference bibliography and labelling associated with the set-up of multiplex amplification. All multiplex PCR amplifications were performed in a total volume of 15 µL, with fixed amounts of Multiplex master mix (7.5 µL), enhancer (1.5 µL) and template DNA (1.1 µL). Volumes of water and primers have been optimized for each locus (see Annex 3 for complete amplification conditions).

The amplification conditions included an initial denaturation at 95°C for 2 min, a denaturation at 95°C for 30 s, an elongation at 60-62°C (depending on the mix) for 30 s, a second denaturation at 95°C for 30 s, an elongation at 68°C for 45 s for a total of 32 cycles and a final extension step at 68°C for 30 s.

The fluorescently labelled PCR fragments were separated using the Applied Biosystems® 3130xl Genetic Analyzer. The results of the electrophoretic runs were provided in fsa format.

2.5. Markers' evaluation

The preliminary data analysis was performed at University of Bologna and in collaboration with IEO of Malaga:

- All results were first screened with PeakScanner v.1.0 (Thermo Fisher Scientific), which provided an overview of individual samples and relative allele size (in base pairs) of nucleic acid fragments using an internal size standard and detecting the presence of a peak. Under the suggestion of BMR Genomics, a new manage size standard (GS500 (-35-50) LIZ; Fig.2) and a new analysis method (minimum peak heights ORANGE: 10) were created and applied.
- Then a more detailed analysis was performed with GeneMarker v.2.7.4 (SoftGenetics). This software allowed the allele binning and the allele calling through the creation of a "Panel" for each assessed marker. The purpose of the "Panel" is to outline and verify the position of expected alleles. The same size standard was used for both approaches.

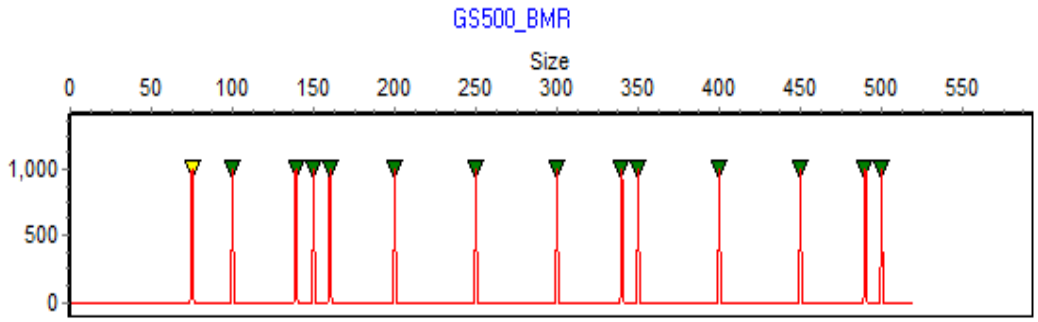


Figure 2. Pattern of peaks in modified GS500 LIZ used in this research work. Peaks are present at: 75, 100, 139, 159, 160, 200, 250, 300, 340, 350, 400, 450, 490 and 500 pb.

2.5.1. Allele calling or binning

After importing the samples' raw data in GeneMarker, a preliminary run was performed without a Panel. This allowed the software to compare peaks in samples with those of the internal size standard (GS500 (-35-50) LIZ; Fig.2).

In the Data Process Analysis, some important parameters were defined:

- The range of "Raw Data Analysis" was set from 1500 to 8000. This allowed the software to correctly read the size standard without translation or shift. This parameter was defined on the range of the raw data
- The range of "Allele Call" was set depending on the range of the locus of interest
- The minimum intensity of "Peak Detection Threshold" was set at 100 in order to allow allele calling even with low intensity peaks.

After the first run, using the software's Panel editor function, a Panel was initialised including all possible alleles for each locus and their bins (expected size limit for each allele; Fig. 3). Once entered into the Panel, the software was able to call peaks observed within each sample into bins referring to the individual designated alleles (Flores-Renteria & Krohn 2013).

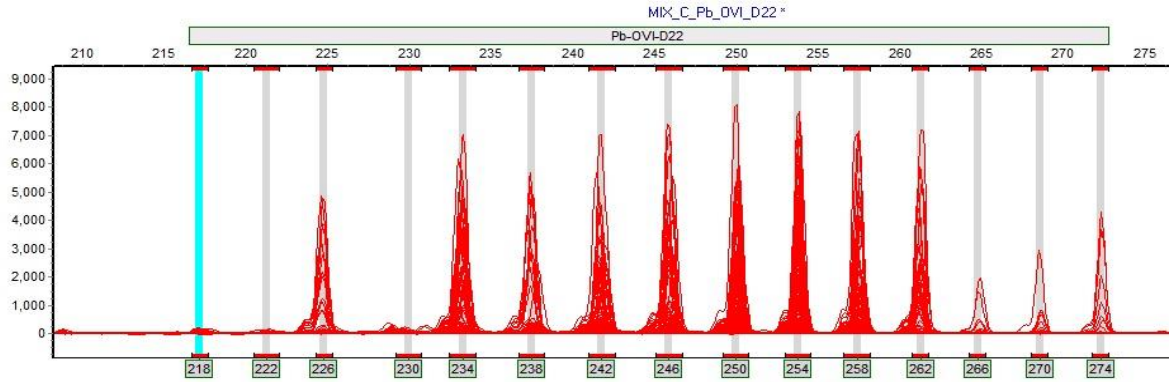


Figure 3. Example of a Panel for a tetra-nucleotide microsatellite locus. Overlay Trace off all samples; the Markers bars appear above the electropherogram and the Bins appear within the electropherogram as brackets at the top and bottom.

Once the definition of the Panel was optimized, data were re-processed. Allele calling was performed manually. When some alleles were not captured in the sample, the Panel was adjusted and the analysis re-run before finalizing the genotyping data.

Besides, one or more internal size standard peaks needed to be adjusted using a manual calibration, to avoid miscalled alleles when applying the Panel.

Post-processing editing of automated allele calls can be necessary because of the presence of confounding peaks caused by PCR artefacts (e.g., dinucleotide stutter or pull-ups due to spectral overlap of two fluorophores, null alleles and imperfect repeats).

Univocal allele calling has been performed following Guichoux et al. (2011), Flores-Renteria & Krohn (2013), the advice of BMR Genomics experts and TRANSBORAN Genetic groups.

After the completion of individual genotyping, results were exported as an Excel file table.

The individuals were considered adequate for the statistical analyses, if a) they successfully amplified across the individuals and b) they amplified in more than 90% of total microsatellite loci.

In order to validate raw data obtained using multilocus and single locus amplifications, a total of eight individuals were randomly chosen for additional amplification and run in parallel for fragment analysis.

2.6. Data set assessment

MICROCHECKER (van Oosterhout et al., 2004) was used to infer the presence of null alleles (using the Brookfield method also implemented in the software), PCR stuttering and large allele drop-out. The default settings (350 base pairs as the maximum expected allele size, 95% confidence interval and 1000 as the number of randomizations) were applied. Moreover, the software FreeNA (Chapuis & Estoup, 2007). was used to determine the frequencies of null allele across population and loci and the F_{ST} value using or not the ENA (Excluding Null Alleles) correction using 1000 runs.

Departures from Hardy-Weinberg Equilibrium (HWE) for each locus and population were performed using Fisher's exact test implemented in GENEPOP v. 4.2 (Raymond & Rousset 1995; Rousset 2008). The Markov Chain Monte Carlo (MCMC) approximation involved 10,000 dememorization steps, 1000 batches and 10,000 iterations per batch. Probability tests were also performed, and the probability values were corrected for multiple testing at alpha: 0.05 using the Bonferroni correction. The same software was used also to evaluate the Linkage Disequilibrium among couples of loci across all populations and the percentage of private alleles [Mean Frequency of Private Alleles p (1)].

Allele frequencies, mean number of alleles per locus, expected and observed heterozygosity values and polymorphism index were calculated using GENETIX software package v. 4.05 (Belkhir et al., 2004).

Allelic richness was estimated using FSTAT v. 2.9.3 (Goudet 2001). The same software was used to calculate the coefficients of inbreeding (F_{IS}).

2.7. Genetic analyses

The pairwise F_{ST}, following Weir and Cockeram model, were calculated with ARLEQUIN v. 3.5 (Excoffier & Lischer 2010) with 20000 permutation and alpha thresholds: 0.05. The same software was used to test hierarchical genetic differentiation by AMOVA analysis (Excoffier et al., 1992) performed in different classes based on the results from pairwise F_{ST} analysis.

The adegenet and ade4 packages (Jombart & Ahmed 2011) for R software were used to perform a Discriminant Analysis of Principal Components (DAPC) to identify and describe genetic clusters.

GenAlEx v. 6.5 (Peakall & Smouse 2012) were used to perform a Principal Coordinates Analysis (PCoA) to explore and to visualize similarities or dissimilarities of data.

STRUCTURE (Pritchard et al., 2000) allowed to create a Bayesian cluster of the individual genotypes. A burning period of 50,000 iterations followed by 500,000 Markov Chain Monte Carlo (MCMC) reps, Admixture Model using Sample Location Information as a prior (LOCPRIOR),

alpha= 1, and Allele frequencies Correlated option were used to perform the analysis. In order to verify consistency of response across runs, 10 runs for each given K (1-14) were tested.

The evaluation of the best K after STRUCTURE analysis, was performed according to the Pritchard criterion (Pritchard et al., 2000) through the web-based STRUCTURE HARVESTER (Earl & vonHoldt 2012). It takes into consideration the trend of LnP(K); the best value of K is represented by the one next to the reaching of the plateau. In situations where numerous values of K give similar estimates of LnP(K), the smallest one was considered correct. The web based CLUMPAK (Kopelman et al., 2015) was used to obtain the bar plots resulting from STRUCTURE analyses.

The Wilcoxon sign-rank test (Luikart et al., 1998) for heterozygosity excess was applied to detect recent bottlenecks, under the two-phase model (TPM) with 10,000 iterations and 95% Single Step Mutations (SMMs), using the program BOTTLENECK (Piry et al., 1999). The qualitative descriptor of the allele frequency distribution, "mode-shift" indicator was also used for assessing distortion in allele frequency, indicative of possible bottleneck.

The statistical power for detecting genetic differentiation at various levels of F_{ST} was evaluated using the software POWSIM (Ryman & Palm, 2006; Ryman et al., 2006). We used sample sizes, number of loci and allele frequencies obtained from the samples in this study.

3. Results

During the first amplification trials, locus Pb-OVI-D101 (mix D) did not amplify in any sample and was then excluded from subsequent analyses. Moreover, the locus PbMS17 was excluded due to an incorrect amplification of the alleles that did not allow a truthful peak calling of the genotype.

The mix E and F, containing microsatellite loci developed on *P. erythrinus*, were excluded from the analysis. In fact, according to previous results, these loci showed low level of polymorphism (AY188339, AY188343), and in the current study, the primer did not pair the same fragment in both the species (AY188342, AY188344, AY188346) or amplification simply failed in *P. bogaraveo* individuals (AY188340, AY188341; complete report of the *P. bogaraveo* loci in Annex 4).

The test for Linkage Disequilibrium did not show any significant results after using Bonferroni correction, so no further loci were excluded at this stage.

Due to cross-contamination between samples, three individuals (SBR-TBK12, SBR-PT42 and SBR-MLT9) were removed from the final data set. Moreover, eight samples were removed from the final dataset due to lack of amplification for more than 10% across all the loci (SBR-TNG8, SBR-TNG12, SBR-TNG21, SBR-TNG22, SBR-TNG29, SBR-TBK3, SBR-MZR8, SBR-MZR13). The dataset used to proceed with further quality check and genetic diversity contained 309 individuals and 20 microsatellite loci.

The results described were obtained from data analyses performed testing:

- 14 different geographic locations STD, PT, COL (including the three samples from Tarifa as the closest sampling port), TNG-SK, TNG-T, TNG-E, MEDITS2018, MEDITS2019, GHZ, ANB, TBK, MZR, MLT (including samples from 2018 and 2019) and ION;
- Six different macroareas: ATL (containing samples from STD, PT and COL+TRF sampling location), GS (containing TNG-SK, TNG-T and TNG-E samples), N-ALB (containing SPAS18 and SPA19 samples), S-ALB (containing GHZ samples), C-MED (containing ANB, TBK, MZR and MLT samples) and E-MED (containing ION samples);
- Nine areas based on GSAs limits: N-ATL (containing samples from STD and PT sampling port); GSA1_NWGS (containing COL+TRF samples), GSA3_SWGS (containing TNG-SK, TNG-T and TNG-E samples); GSA 1 (containing SPAS18 and SPA19 samples), GSA 4 (containing GHZ and ANB samples), GSA 12 (with TBK samples), GSA 15 (with MLT samples); GSA16 (with MZR samples) and GSA20 (with Ionian samples; Tab. 4).

The three scenarios were selected to study the genetic structure considering different level of geographic origin. The second analysis performed using only six geographical area allowed us to study the genetic difference between more balanced populations which contain an equal or major

number of 30 samples (excluding GHZ); while using 9 areas based on GSAs limits, we can better compare the results obtained from different analyses and method performed for the TRANSBORAN project.

Table 4. The population scenarios used in this project to test genetic structure.

Scenario 1	N. of individuals	Scenario 2	N. of individuals	Scenario 3	N. of individuals
STD	26	**ATL**	79	**N-ATL**	61
PT	35				
COL	18			**GSA1_NWGS**	18
TNG-SK	13	**W-GS**	63	**GSA3_SWGS**	63
TNG-T	15				
TNG-E	35				
SPA18	10	**N-ALB**	40	**GSA 1**	40
SPA19	30				
GHZ	9	**S-ALB**	9	**GSA4**	39
ANB	30	**CEN-MED**	88		
TBK	28			**GSA12**	28
MZR	16			**GSA15**	16
MLT	14			**GSA16**	14
ION	30	**E-MED**	30	**GSA20**	30

3.1. Marker quality

The percentage of private alleles [Mean Frequency of Private Alleles p (1)] calculated on 309 individuals and 20 loci, resulted in a value of 3.1%.

Significant deviations from Hardy-Weinberg expectations, after Bonferroni correction, were found in all locations and in 13 microsatellite loci (PbMS2, PbMS6, PbMS16, PbMS19, Pb-OVI-D108, Pb-OVI-A5, Pb-OVI-D102 Pb-OVI-A3, PbMS20, PbMS18, PbMS1, PbMS4, and PbMS15; Tab. 6).

Analyses of these data by means of the MICROCHECKER software indicated the presence of null alleles as the most probable reason for departures from HWE. The errors for stuttering (ST), defined as the highly significant shortage of heterozygous genotypes with alleles of one repeat unit difference, have been detected only in few loci and localities: PbMS20 for STD, PT, ANB, MLT and ION locations, PbMS18 for STD, PT and TNG-T locations Pb-OVI-D114 for TNG-SK location (Tab. 7).

No evidence for large allele dropout was detected.

The results obtained with FreeNA about the frequencies of presence of Null Allele, are comparable with the output of MICROCHECKER, although the former software appeared more stringent.

Evidence of scoring errors were compared with the results of the HWE test. This highlighted the overlapping of the results obtained with the different approaches. In Tab. 7 cells in yellow represent

the loci, within a location, which are in HW disequilibrium (alpha: 0.05), also without scoring error (both null allele and stuttering).

Table 5. Genetic diversity at 14 geographic locations for all 20 loci. N ind: number of analysed samples within a population; Hexp: expected heterozygosity; Hobs: observed heterozygosity. Mean number of alleles for each population, mean AR: mean allelic richness; FIS: inbreeding coefficient; HWE: Hardy-Weinberg Equilibrium: ***significant result after Bonferroni correction (alpha: 0.0001786).***

Pop	N. of individuals	Hexp	Hobs	Mean N. allele	Mean AR	Fis	HWE
STD	26	0.833	0.701	14.25	8.327	0.178	**< 7.92e-23**
PT	34	0.850	0.730	14.90	8.367	0.156	**< 7.82e-18**
COL+TRF	18	0.827	0.720	11.55	8.082	0.157	**< 3.73e-15**
TNG-SK	13	0.802	0.623	10.00	7.984	0.263	**< 2.27e-14**
TNG-T	15	0.816	0.686	10.80	7.953	0.193	**< 3.07e-18**
TNG-E	35	0.852	0.710	15.10	8.355	0.181	**< 3.76e-22**
SPA18	10	0.800	0.733	9.15	8.205	0.137	**< 3.01e-06**
SPA19	29	0.863	0.703	15.80	8.872	0.202	**< 4.20e-26**
GHZ	9	0.817	0.731	8.70	8.216	0.164	**< 2.78e-12**
ANB	30	0.846	0.731	14.70	8.370	0.152	**< 2.91e-16**
TBK	27	0.851	0.726	14.20	8.385	0.165	**< 1.93e-21**
MZR	14	0.812	0.702	10.20	8.079	0.173	**< 2.43e-12**
MLT	14	0.825	0.751	11.35	8.189	0.122	**< 2.00e-10**
ION	29	0.848	0.750	14.30	8.411	0.133	**< 2.08e-13**

Table 6. Genetic diversity at 20 microsatellite loci over 14 Pagellus bogaraveo *geographic locations. Number of alleles, Hexp: expected heterozygosity; Hobs: observed heterozygosity, F_{is}: inbreeding coefficient, HWE: Hardy-Weinberg Equilibrium:* ***significant result after Bonferroni correction (alpha: 0.0001786).***

Locus	N. alleles	Hexp	Hobs	Mean AR	Fis	HWE	Locus	N. alleles	Hexp	Hobs	Mean AR	Fis	HWE
Pb-OVI-B2	23	0.862	0.823	8.653	0.074	0.004119	**Pb-OVI-D21**	14	0.842	0.871	7.378	-0.004	0.978329
PbMS2	20	0.848	0.556	8.381	0.368	**< 0.00e+00**	**Pb-OVI-D106**	11	0.788	0.810	5.753	0.001	0.647982
PbMS6	37	0.937	0.866	12.386	0.104	**< 4.91e-12**	**Pb-OVI-A3**	22	0.884	0.712	9.161	0.222	**< 2.32e-15**
PbMS16	42	0.937	0.861	12.556	0.110	**< 6.52e-15**	**PbMS20**	21	0.830	0.446	7.324	0.485	**< 0.00e+00**
PbMS19	10	0.665	0.587	4.240	0.147	**< 9.43e-07**	**PbMS18**	23	0.832	0.502	8.373	0.422	**< 0.00e+00**
Pb-OVI-D108	7	0.456	0.462	2.893	0.004	**7.27E-06**	**PbMS1**	38	0.917	0.463	11.036	0.517	**< 0.00e+00**
Pb-OVI-A5	41	0.914	0.826	10.952	0.125	**< 5.30e-11**	**Pb-OVI-D22**	12	0.821	0.825	6.889	0.023	0.568561
Pb-OVI-D114	11	0.812	0.736	6.355	0.121	0.00568	**PbMS4**	29	0.905	0.899	10.295	0.035	**< 1.18e-05**
Pb-OVI-D102	20	0.879	0.718	8.855	0.212	**< 6.98e-13**	**PbMS15**	40	0.924	0.740	11.702	0.226	**< 3.24e-31**
Pb-OVI-D20	12	0.820	0.861	6.703	-0.025	0.790455	**Pb-OVI-C103**	11	0.758	0.719	5.534	0.080	0.02496

Table 7. MICROCHECKER results compared with HW test results. The yellow cell represented the loci, within a population, in HW disequilibrium (alpha: 0.05). Text within the cell indicates scoring errors. NA: presence of Null Allele, ST: presence of stuttering.

	STD	PT	COL + TRF	TNG-SK	TNG-T	TNG-E	SPA18	SPA19	GHZ	ANB	TBK	MZR	MLT	ION
Pb-OVI-B2			NA					NA			NA			
PbMS2	NA	NA	NA		NA	NA		NA	NA	NA	NA	NA	NA	NA
PbMS6	NA	NA			NA			NA		NA	NA		NA	
PbMS16			NA	NA				NA					NA	NA
PbMS19				NA		NA								
Pb-OVI-D108								NA				NA		
Pb-OVI-A5		NA			NA	NA		NA						
Pb-OVI-D114		NA		NA/ST			NA			NA	NA			
Pb-OVI-D102	NA		NA	NA	NA	NA	NA			NA	NA			
Pb-OVI-D20														
Pb-OVI-D21				NA										
Pb-OVI-D106				NA										
Pb-OVI-A3	NA			NA	NA	NA		NA				NA		NA
PbMS20	NA/ST	NA/ST	NA	NA	NA	NA	NA	NA	NA	NA/ST	NA	NA	NA/ST	NA/ST
PbMS18	NA/ST	NA/ST	NA	NA	NA/ST	NA	NA	NA	NA	NA	NA	NA	NA	NA
PbMS1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Pb-OVI-D22														
PbMS4	NA							NA		NA				
PbMS15	NA	NA	NA	NA		NA	NA	NA		NA	NA		NA	NA
Pb-OVI-C103	NA	NA												

3.2. Genetic diversity

The number of individuals for each location, expected and observed heterozygosity, mean values of number of alleles, mean value of allelic richness, F_{IS} values, and HWE results for location are presented in Tab. 5.

The number of alleles, the expected and observed heterozygosity, F_{IS} values, and HWE results for microsatellite loci are presented in Tab. 6.

The observed heterozygosity (Hobs) did not vary significantly across the sampling location and ranged between 0.751 (MLT) and 0.623 (TNG-SK). This index, together with the positive F_{IS} value at all the sampling locations, pinpointed a heterozygote deficiency.

The Hobs did vary across microsatellite loci, ranging from 0.899 for PbMS4 to 0.462 for Pb-OVI-D108. The comparison with the expected heterozygosity (Hexp) and the positive F_{IS} value at almost all the loci, except Pb-OVI-D20 and Pb-OVI-D21, indicated heterozygote deficiency.

The mean value of allelic richness has been normalized using the minimum number of individuals within a location (nine individuals for GHZ). Sample SPA19 showed the highest value of mean allelic richness (8.872), while TNG-T showed the lowest one (7.953).

However, the results indicated high levels of genetic diversity, with an average of 22.20 alleles per locus. The mean number of alleles per locus estimated in each location fluctuated between 15.80 and 8.70 in SPA19 and GHZ locations, respectively. The results appear to be related to the sample size, since allelic richness and expected heterozygosity are similar for all locations (Fig.4).

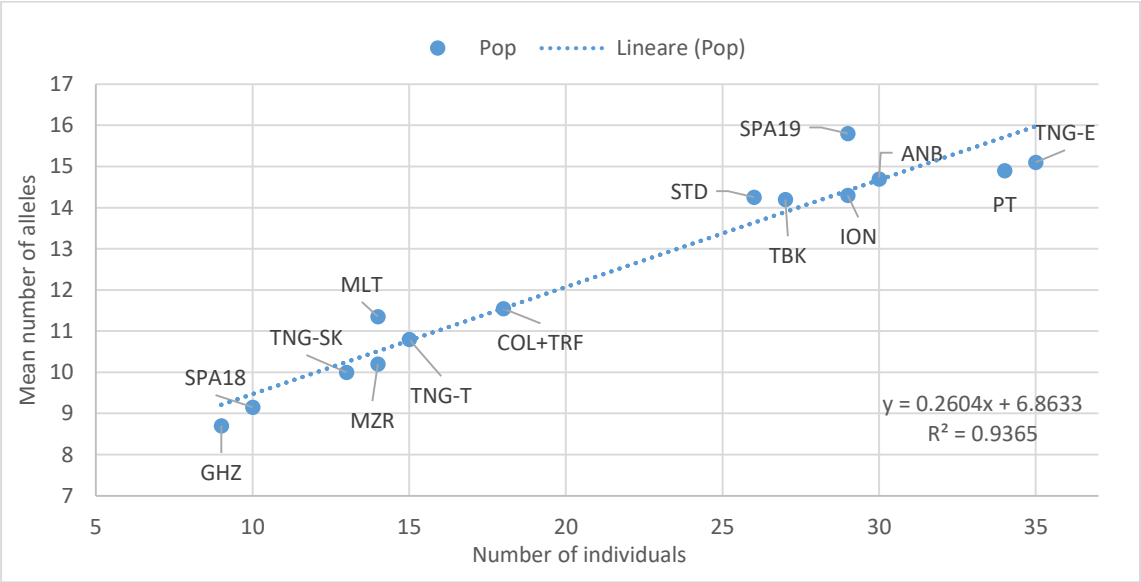


Figure 4. Rarefaction curve representative of the mean number of alleles and sampling units for 14 geographic locations.

From the results of descriptive analyses, tests of genetic population structure were performed using two different microsatellite loci dataset:

- 14 microsatellite loci: Pb-OVI-B2, PbMS6, PbMS16, PbMS19, Pb-OVI-D108, Pb-OVI-A5, Pb-OVI-D114, Pb-OVI-D20, Pb-OVI-D21, Pb-OVI-D106, Pb-OVI-A3, Pb-OVI-D22, PbMS4 and Pb-OVI-C103. These markers were chosen because they presented null alleles and HW disequilibrium, for less than 50% of the 14 locations.
- seven microsatellite loci: Pb-OVI-B2, Pb-OVI-D114, Pb-OVI-D20, Pb-OVI-D21, Pb-OVI-D106, Pb-OVI-D22 and Pb-OVI-C103. These markers were chosen because they presented very low null alleles frequencies and were at HW equilibrium for all locations.

The expected and observed heterozygosity, F_{IS} values, and HWE results for these two sets of microsatellite loci are presented in Tab. 8.

Table 8. Genetic diversity at 14 geographic locations for 14 and 7 microsatellite loci. Hexp: expected heterozygosity; Hobs: observed heterozygosity; FIS: inbreeding coefficient; HWE: Hardy-Weinberg Equilibrium: ***significant result after Bonferroni correction.***

	14 loci			**7 loci**		
Pop	**Hexp.**	**Hobs.**	**Fis /HWE**	**Hexp.**	**Hobs.**	**Fis /HWE**
STD	0.822	0.781	**0.069**	0.832	0.852	-0.004
PT	0.839	0.795	**0.067**	0.835	0.792	0.067
COL+TRF	0.807	0.816	0.017	0.825	0.849	-0.001
TNG-SK	0.801	0.667	**0.208**	0.806	0.663	0.219
TNG-T	0.801	0.771	0.072	0.807	0.837	-0.003
TNG-E	0.838	0.783	**0.08**	0.843	0.868	-0.015
SPA18	0.788	0.821	0.011	0.807	0.856	-0.006
SPA19	0.845	0.772	**0.104**	0.843	0.848	0.012
GHZ	0.803	0.818	0.041	0.802	0.857	-0.011
ANB	0.823	0.804	**0.04**	0.824	0.852	-0.016
TBK	0.840	0.809	**0.056**	0.832	0.795	**0.062**
MZR	0.810	0.776	0.079	0.814	0.845	0.006
MLT	0.817	0.840	0.006	0.822	0.875	-0.032
ION	0.833	0.816	0.038	0.832	0.843	0.004

3.3. Population differentiation and demography

3.3.1. Analysis using 14 microsatellite loci

A pattern of genetic differentiation was observed using pairwise F_{ST} values between ANB versus STD, TNG-SK SPA19, MZR and ION (F_{ST}: 0.007, 0.011, 0.011, 0.014 and 0.006 respectively). Significant values of F_{ST} involved the comparison of TBK with COL+TRF, TNG-SK, TNG-T, SPA19, ANB and MZR (F_{ST}: 0.008, 0.014, 0.008, 0.006, 0.010 and 0.009 respectively). Moreover, GHZ resulted significantly different from COL-TRF (F_{ST}: 0.012; Tab.9).

Three AMOVA analyses were conducted dividing samples based on results of pairwise F_{ST} (Tab. 10):

AMOVA test 1: One Group, no geographical definition.

AMOVA test 2: Group 1: STD; Group 2: PT; Group 3: COL, TNG-SK, TNG-T, TNG-E; Group 4: SPA18, SPA19, GHZ; Group 5: ANB, TBK; Group 6: MZR, MLT; Group 7: ION

AMOVA test 3: Group 1: STD, PT, TNK-SK, TNG-T, TNG-E, COL, SPA18, SPA19, MLT, MZR, ION; Group 2: GHZ, TBK, ANB.

Low percentage of variation was observed among groups (0.00 and 0.26 for AMOVA test 2 and test 3, respectively) and was significant (alpha: 0.05) only for AMOVA test 3. The percentage of variation among locations within groups ranged from 5.17 to 0.10 and was significant in AMOVA test 1 and 2. The largest variation was observed “within individuals” and ranged from 94.64 to 94.48 and was significant in all tests.

*Table 9. Pairwise F_{ST} values, conducted for 14 geographic locations and 14 loci in **bold significant result at alpha: 0.05**.*

	STD	PT	COL+TRF	TNG-SK	TNG-T	TNG-E	SPA18	SPA19	GHZ	ANB	TBK	MZR	MLT	ION
STD	0.000													
PT	-0.003	0.000												
COL+TRF	-0.004	-0.001	0.000											
TNG-SK	0.006	0.006	0.009	0.000										
TNG-T	0.007	0.000	0.006	0.002	0.000									
TNG-E	0.000	-0.001	0.002	0.004	-0.002	0.000								
SPA18	0.003	-0.002	0.002	0.015	-0.002	-0.001	0.000							
SPA19	0.000	0.001	0.004	0.010	0.008	0.002	0.002	0.000						
GHZ	0.007	0.000	**0.012**	0.013	0.008	0.007	0.015	0.001	0.000					
ANB	**0.007**	0.003	0.004	**0.011**	0.001	0.003	0.008	**0.011**	-0.001	0.000				
TBK	0.002	0.001	**0.008**	**0.014**	**0.008**	0.004	0.009	**0.006**	0.002	**0.010**	0.000			
MZR	-0.003	0.002	-0.004	0.000	0.008	0.002	0.004	0.001	0.015	**0.014**	**0.009**	0.000		
MLT	0.002	-0.002	0.001	0.004	0.004	0.001	0.003	0.003	0.001	0.006	0.001	0.003	0.000	
ION	0.000	-0.002	0.004	0.004	0.002	0.001	0.005	0.003	0.005	**0.006**	0.004	0.000	0.004	0.000

*Table 10. Proportion of molecular variance among groups, among populations and within individuals resulting from AMOVA Analysis performed on 14 geographic locations and 14 loci with three different groupings. In **bold significant result at alpha: 0.05**.*

AMOVA Test1					**AMOVA Test2**					AMOVA Test3				
Group1: all samples					**Group1: STD; Group2: PT; Group3: COL, TNG-SK, TNG-T, TNG-E; Group4: SPA18, SPA19, GHZ; Group5: ANB, TBK; Group6: MZR, MLT; Group7: ION**					**Group1: STD, PT, TNK-SK, TNG-T, TNG-E, COL, SPA18, SPA19, MLT, MZR, ION; Group2: GHZ, TBK, ANB**				
Source of Variation	**% variation**	**F statistics**		**P**	**Source of Variation**	**% variation**	**F statistics**		**P**	**Source of Variation**	**% variation**	***F* statistics**		**P**
					Among groups	-0.12	FCT	-0.001	0.815	**Among groups**	0.26	FCT	0.003	**0.014**
Among population	0.20	FST	0.002	1.000	**Among population within groups**	0.31	FSC	0.003	**0.037**	**Among population within groups**	0.10	FSC	0.001	0.162
Among individuals Within population	5.17	FIS	0.052	**0.000**	**Among individuals Within population**	5.17	FIS	0.052	**0.000**	**Among individuals Within population**	5.16	FIS	0.052	**0.000**
Within individuals	94.63	FIT	0.054	**0.000**	**Within individuals**	94.64	FIT	0.054	**0.000**	**Within individuals**	94.48	FIT	0.055	**0.000**

The results from DAPC analyses with 14 microsatellite loci were performed based on 14 locations, 6 macro-areas and 9 GSA and they are represented in Figs. 5, 6 and 7. In all the resulting clusters, the points were scattered, and it was not possible to observe a clear sign of differentiation between locations, macroareas or GSAs.

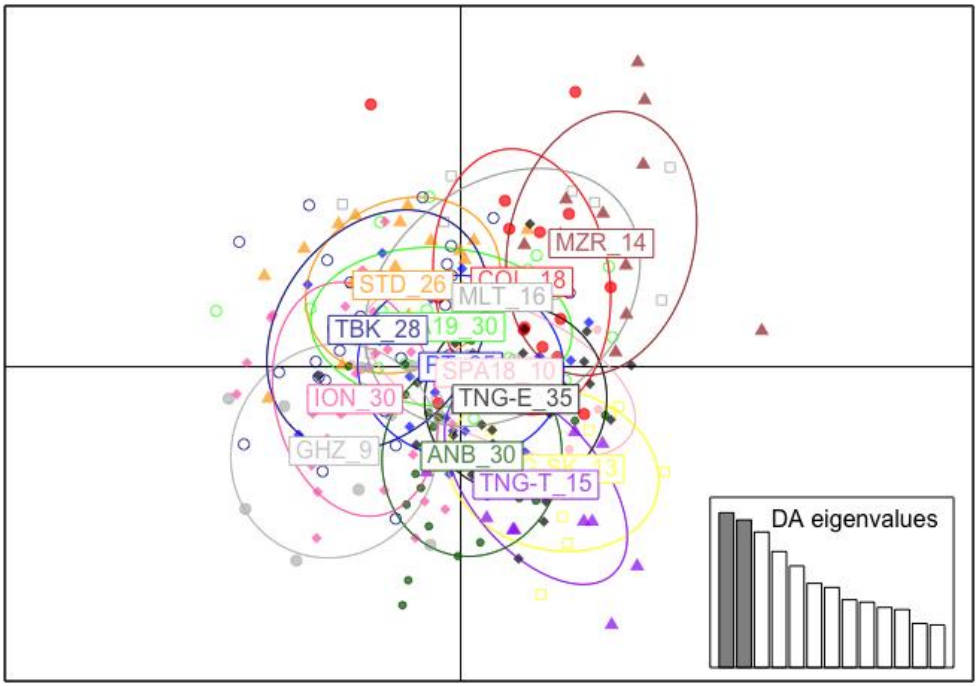


Figure 5. Result from DAPC analysis performed on 14 locations with 14 microsatellite loci using R software adegenet package.

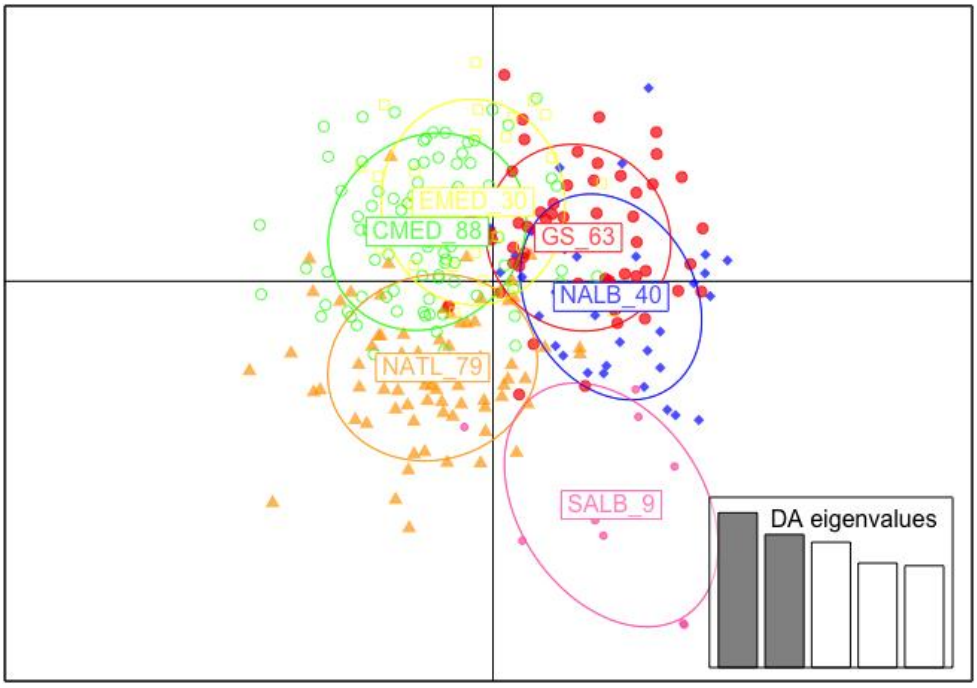


Figure 6. Result from DAPC analysis performed on 6 macroareas with 14 microsatellite loci using R software adegenet package.

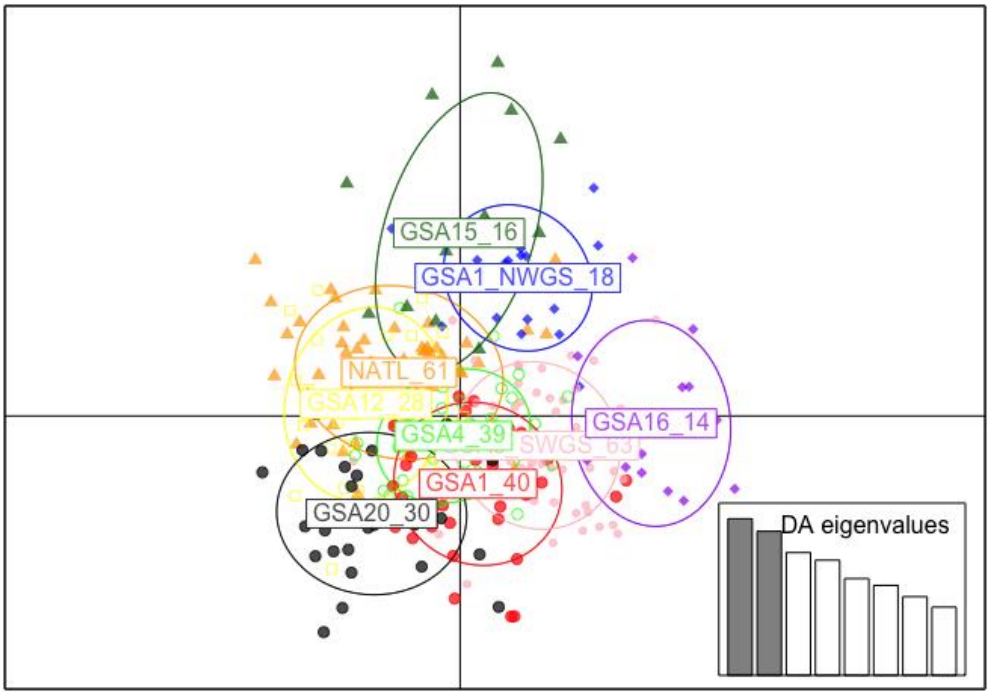


Figure 7. Result from DAPC analysis performed on 9 GSAs with 14 microsatellite loci using R software adegenet package.

The results from PCoA analyses with 14 microsatellite loci, based on 14 locations, 6 macro-areas and 9 GSA, are represented in Figs. 8, 9 and 10.

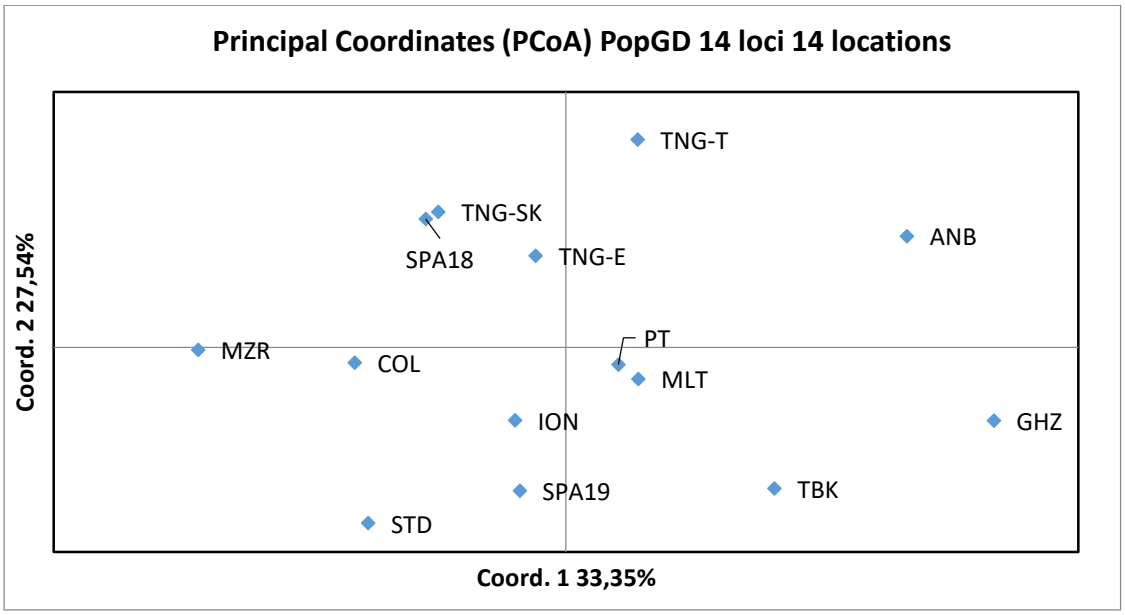


Figure 8. Result from PCoA analysis performed on 14 locations with 14 microsatellite loci using GenAlEx v. 6.5.

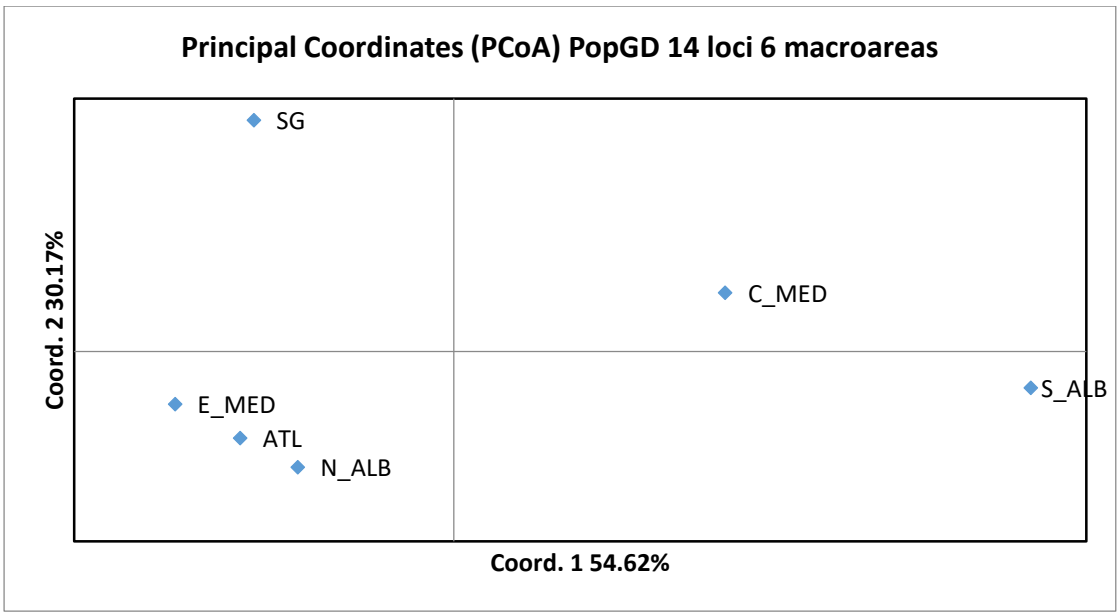


Figure 9. Result from PCoA analysis performed on 6 macroareas with 14 microsatellite loci using GenAlEx v. 6.5.

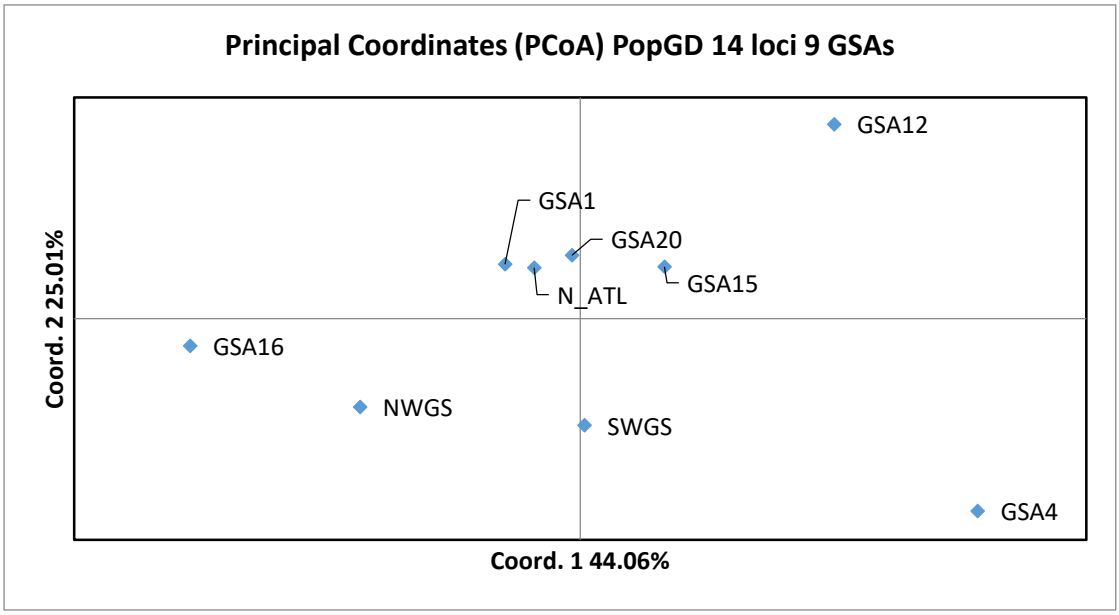


Figure 10. Result from PCoA analysis performed on 9 GSAs with 14 microsatellite loci using GenAlEx v. 6.5.

The results from STRUCTURE were analysed using the LnP(K) trend to evaluate best K by STRUCTURE HARVESTER. The plot did not show the expected plateau usually displayed when a population structure is present. For best K=1 no genetic differentiation was detected between the geographic populations. The plot showed an increase of LnP(K) variance between runs in relation to the increase of K (Fig.11).

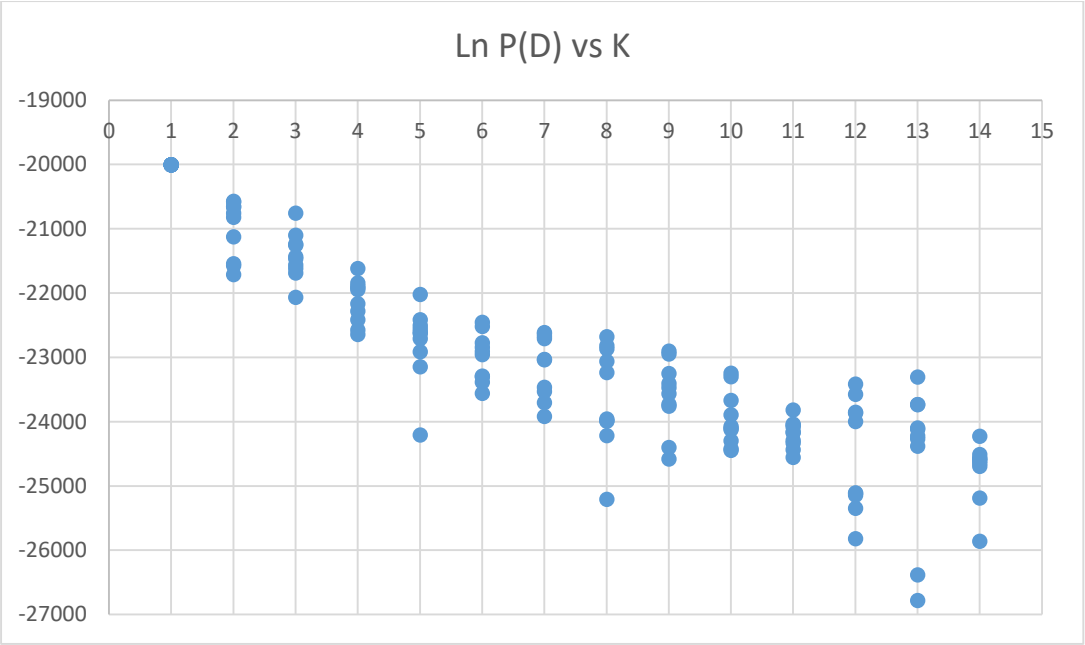


Figure 11. LnP(K) graph resulting from STRUCTURE HARVESTER.

Individual bar plots grouped from K = 2 to K = 5 population clusters, were obtained with the web based CLUMPAK (Figure 12). The results showed an admixed genetic component across all the investigated areas.

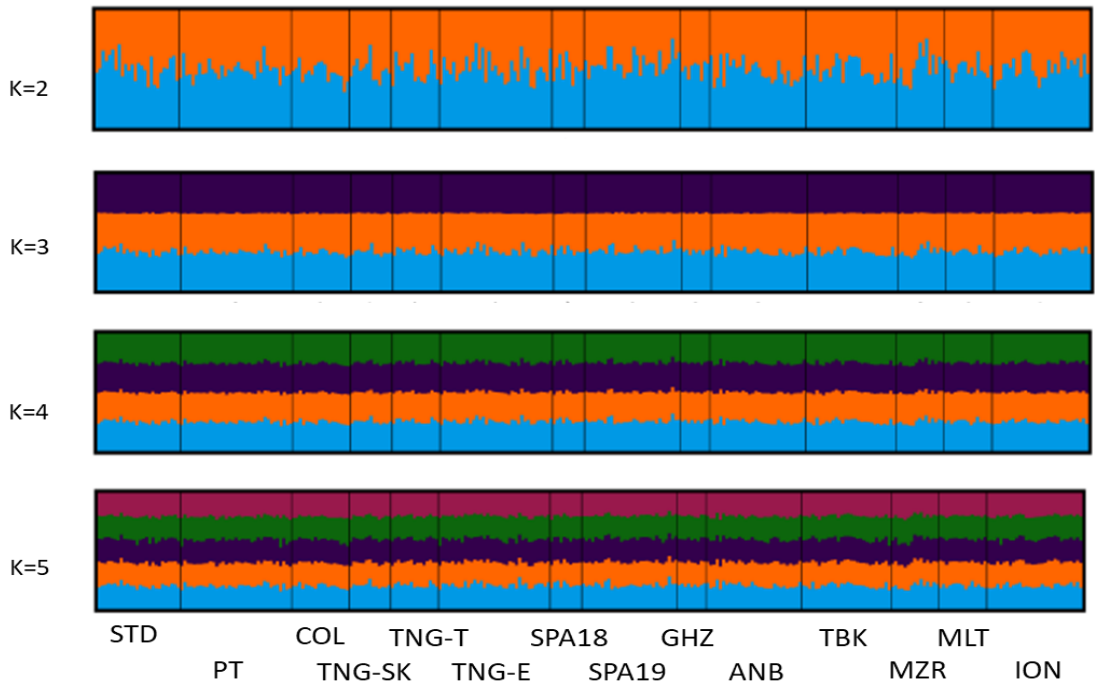


Figure 12. Summary plot of the STRUCTURE analysis from K=2 to K=5. Each individual is represented by a single vertical line on the x-axis. Coloured segments represent the membership of individuals to the population clusters. Results are shown for the 14 locations subdivision using 14 microsatellite loci

BOTTLENECK software output gives back a single significant result (with p-value:0.05) for tail 1 (deficit of heterozygotes) at TNG-E population (Tab. 11).

*Table 11 Results from BOTTLENECK software; in **bold significant result at alpha: 0.05**.*

Population	Mean N	Mean k	Mean He	p_W_1t_TPM	p_W_2t_TPM
STD	50.9	13.9	0.839	0.500	1.000
PT	69.6	14.9	0.852	0.097	0.194
COL+TRF	35.9	11.2	0.830	0.148	0.296
TNG-SK	25.1	9.7	0.834	0.068	0.135
TNG-T	29.9	10.4	0.829	0.134	0.268
TNG-E	69.6	15.2	0.850	**0.029**	0.058
SPA18	19.9	9.1	0.829	0.121	0.241
SPA19	58.6	15.6	0.860	0.749	0.542
GHZ	18.0	8.8	0.850	0.163	0.326
ANB	59.4	13.6	0.837	0.524	1.000
TBK	55.0	13.4	0.855	0.213	0.426
MZR	27.6	10.6	0.840	0.452	0.903
MLT	31.4	11.0	0.844	0.163	0.326
ION	58.6	14.5	0.848	0.548	0.952

3.3.2. Analysis using 7 microsatellite loci

A pattern of genetic differentiation was observed using pairwise F_{ST} values between ANB versus STD, TNG-SK, TNG-E, SPA19, MZR and MLT (F_{ST}: 0.007, 0.020, 0.006, 0.013, 0.011 and 0.013 respectively). Significant values of F_{ST} involved the comparison of TBK with TNG-SK and ANB (F_{ST}: 0.016 and 0.012; Tab. 12).

Three AMOVA analyses were conducted dividing samples based on results of pairwise FST (Tab. 13):

AMOVA test 1: Group 1, all samples inside a unique location.

AMOVA test 2: Group 1: STD; Group 2: PT; Group 3: COL, TNG-SK, TNG-T, TNG-E; Group 4: SPA18, SPA19, GHZ; Group 5: ANB, TBK; Group 6: MZR, MLT; Group 7: ION

AMOVA test 3: Group 1: STD, PT, TNK-SK, TNG-T, TNG-E, COL, GHZ, SPA18, SPA19, MLT, MZR, ION; Group 2: TBK, ANB.

Low percentage of variation was observed among groups (0.00 and 0.10 for AMOVA test 2 and test 3, respectively) but none of the tests were significant considering alpha: 0.05. The percentage of variation among locations within groups ranged from 0.40 to 0.13 and were not significant. The largest variation was detected “within individuals”, varying from 99.45 to 99.37 and showing not significant values in all tests.

*Table 12. Pairwise F_{ST} values, conducted for 14 geographic locations and 7 loci in **bold significant result at alpha: 0.05***

	STD	PT	COL+TRF	TNG-SK	TNG-T	TNG-E	SPA18	SPA19	GHZ	ANB	TBK	MZR	MLT	ION
STD	0.000													
PT	-0.003	0.000												
COL+TRF	-0.005	-0.004	0.000											
TNG-SK	0.006	0.008	0.009	0.000										
TNG-T	0.006	-0.003	0.003	0.009	0.000									
TNG-E	0.003	-0.002	-0.001	0.007	-0.006	0.000								
SPA18	0.003	-0.005	-0.001	0.019	-0.003	0.000	0.000							
SPA19	-0.004	0.001	0.000	0.011	0.006	0.001	-0.001	0.000						
GHZ	0.001	0.003	0.005	0.018	0.011	0.008	0.016	0.000	0.000					
ANB	**0.007**	0.003	-0.003	**0.020**	0.001	**0.006**	0.010	**0.013**	-0.001	0.000				
TBK	-0.001	-0.002	0.000	**0.016**	0.007	0.002	-0.001	0.007	0.006	**0.012**	0.000			
MZR	-0.006	-0.002	-0.008	-0.017	-0.004	-0.006	0.001	-0.002	0.007	**0.011**	-0.001	0.000		
MLT	-0.003	-0.002	-0.005	0.007	0.004	0.001	-0.001	0.000	0.012	**0.013**	-0.001	-0.008	0.000	
ION	-0.001	-0.005	0.002	0.002	-0.003	-0.001	0.000	-0.001	0.001	0.006	0.000	-0.004	0.000	0.000

*Table 13. Proportion of molecular variance among groups, among populations and within individuals resulting from AMOVA Analysis performed on 14 geographic locations and 7 loci with three different groupings. In **bold significant result at alpha: 0.05**.*

AMOVA Test1					**AMOVA Test2**					**AMOVA Test3**				
Group1: all samples					**Group1: STD; Group2: PT; Group3: COL, TNG-SK, TNG-T, TNG-E; Group4: SPA18, SPA19, GHZ; Group5: ANB, TBK; Group6: MZR, MLT; Group7: ION**					**Group1: STD, PT, TNK-SK, TNG-T, TNG-E, COL, GHZ, SPA18, SPA19, MLT, MZR, ION; Group2: TBK, ANB**				
Source of Variation	**% variation**	**F statistics**		**P**	**Source of Variation**	**% variation**	**F statistics**		**P**	**Source of Variation**	**% variation**	**F statistics**		**P**
					Among groups	-0.13	FCT	-0.001	0.742	**Among groups**	0.10	FCT	0.001	0.230
Among population	0.16	FST	0.002	1.000	**Among population within groups**	0.28	FSC	0.003	0.091	**Among population within groups**	0.13	FSC	0.001	0.204
Among individuals Within population	0.40	FIS	0.004	0.320	**Among individuals Within population**	0.40	FIS	0.004	0.315	**Among individuals Within population**	0.40	FIS	0.004	0.324
Within individuals	99.43	FIT	0.006	0.259	**Within individuals**	99.45	FIT	0.006	0.258	**Within individuals**	99.37	FIT	0.006	0.266

The DAPC performed using 7 microsatellite loci and 14 locations, 6 macroareas and 9GSAs are represented in Figs. 13, 14 and 15. All the grouping did not show a delineated differentiation.

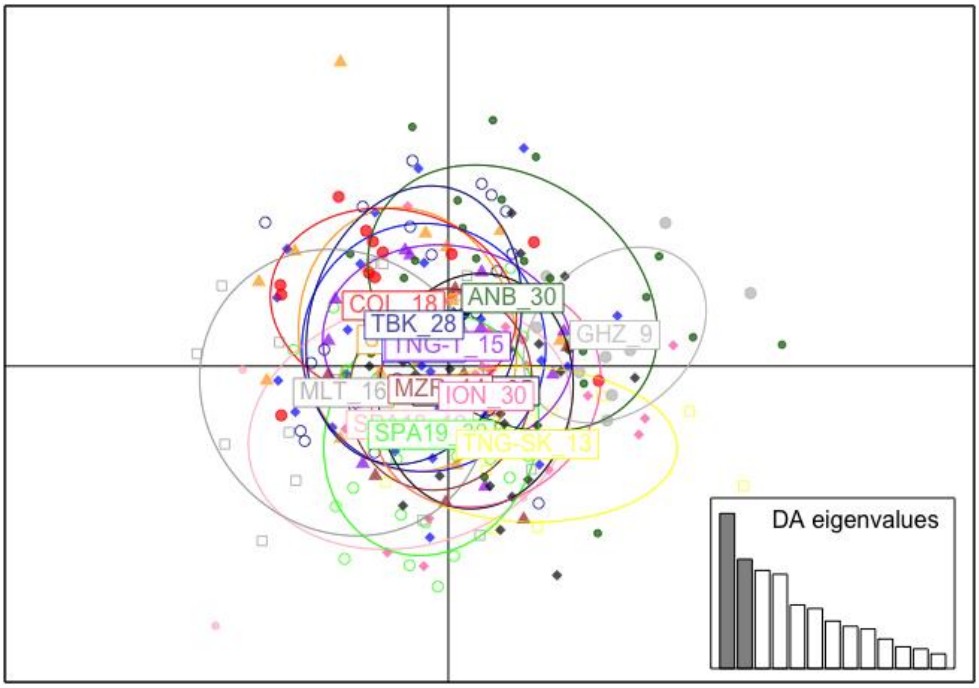


Figure 13. Result from DAPC analysis performed on 14 locations with 7 microsatellite loci using R software adegenet package.

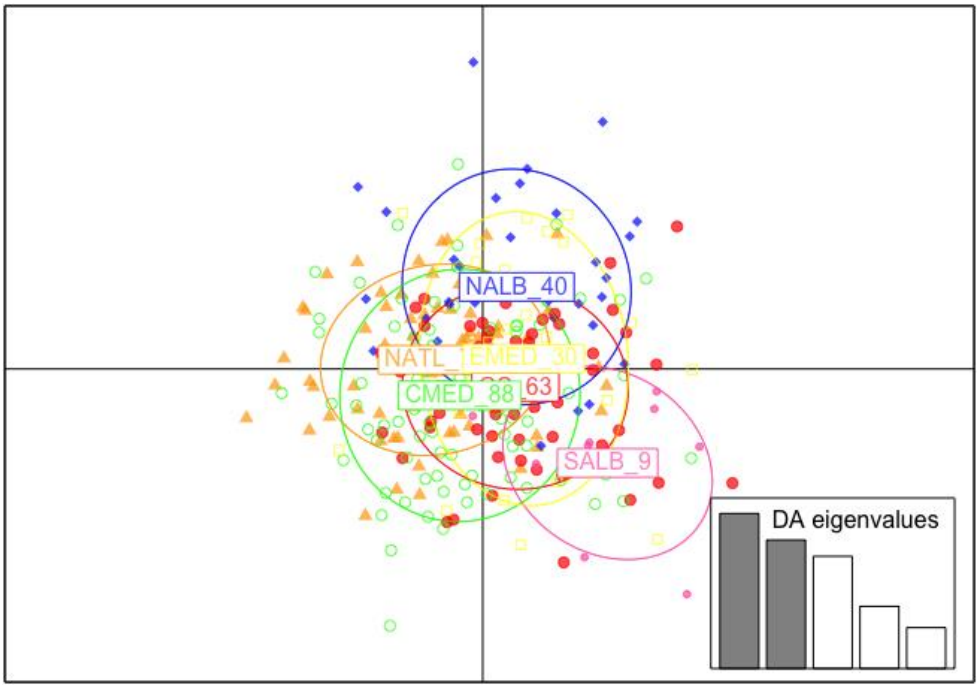


Figure 14. Result from DAPC analysis performed on 6 macroareas with 7 microsatellite loci using R software adegenet package.

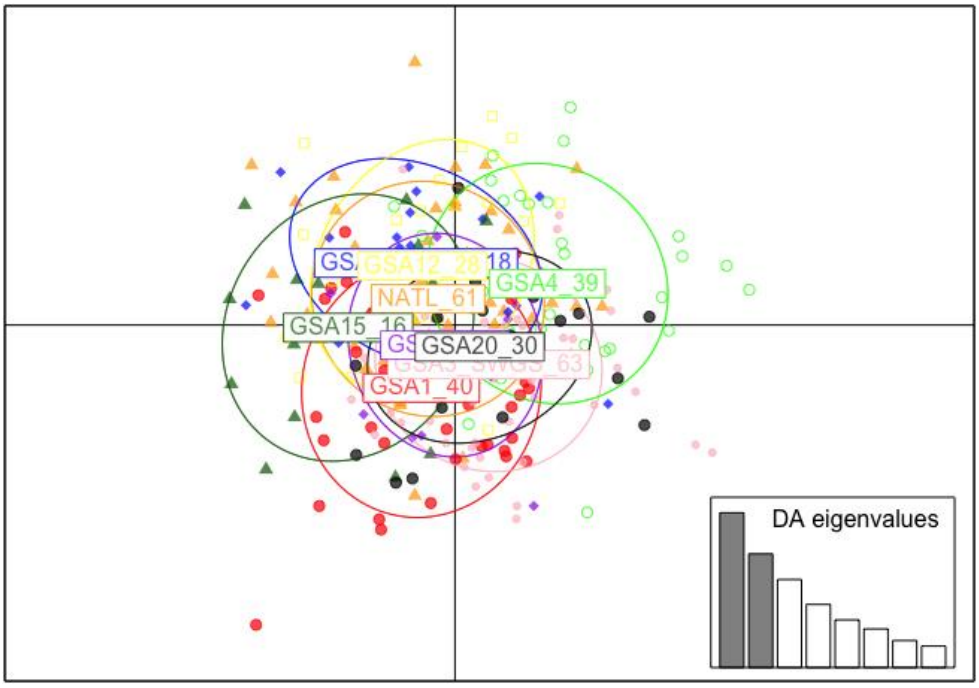


Figure 15. Result from DAPC analysis performed on 9 GSAs with 7 microsatellite loci using R software adegenet package.

The results from PCoA analyses with seven microsatellite loci, based on 14 locations, 6 macro-areas and 9 GSA, are represented in Figs. 16, 17 and 18.

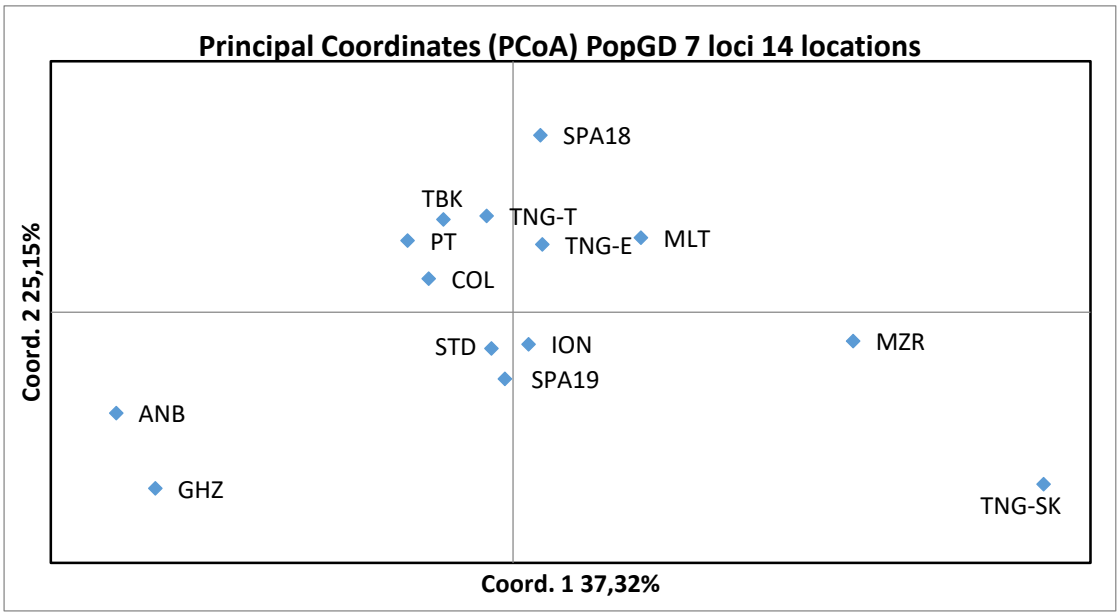


Figure 16. Result from PCoA analysis performed on 14 locations with 7 microsatellite loci using GenAlEx v. 6.5.

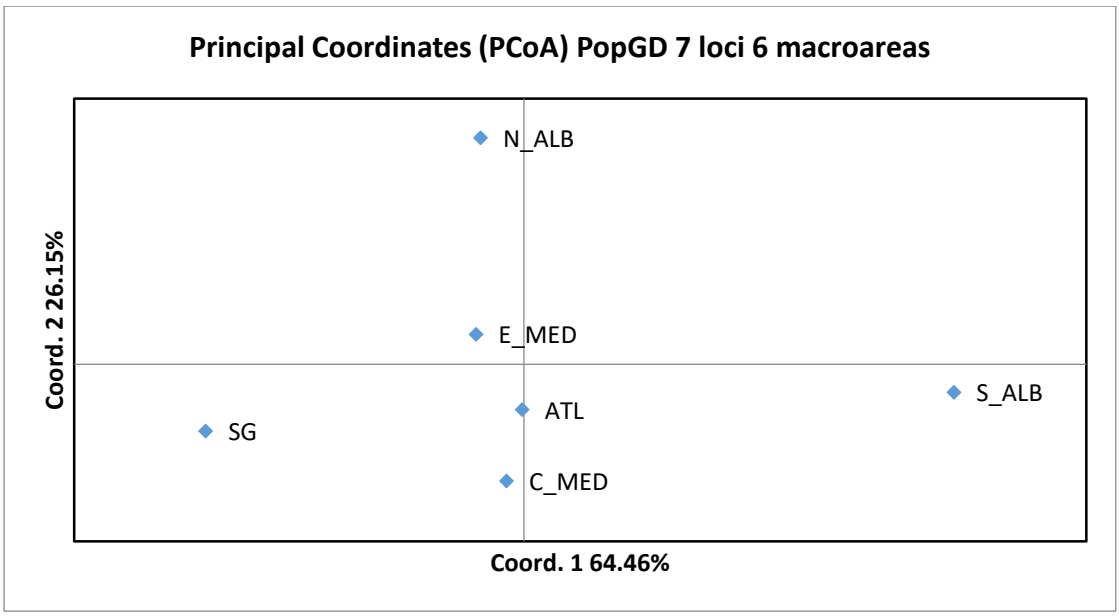


Figure 17. Result from PCoA analysis performed on 6 macroareas with 7 microsatellite loci using GenAlEx v. 6.5.

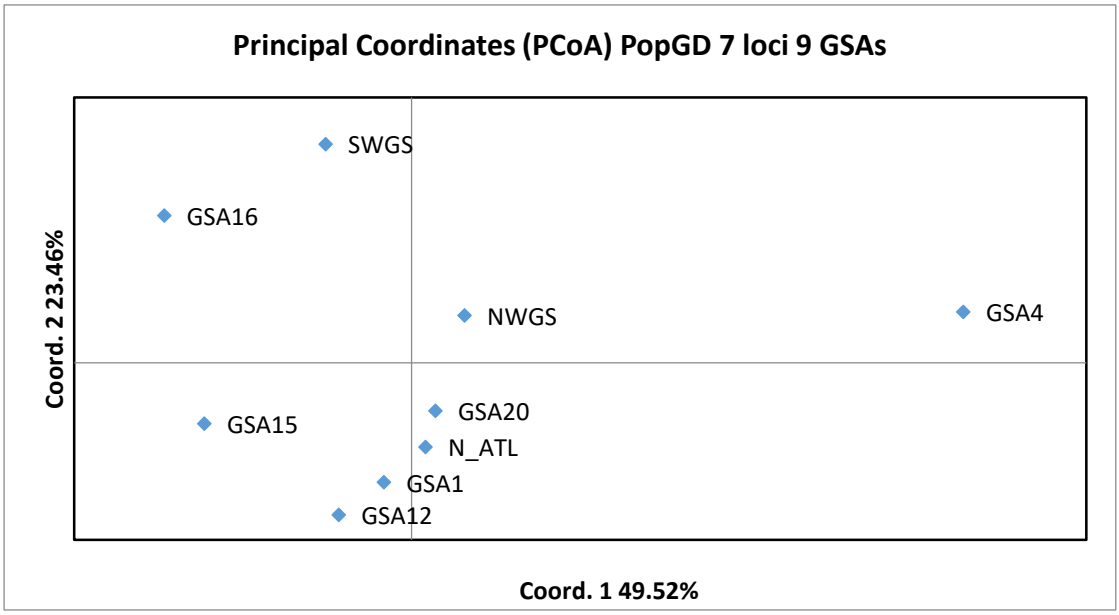


Figure 18. Result from PCoA analysis performed on 9 GSAs with 7 microsatellite loci using GenAlEx v. 6.5.

The results from STRUCTURE were analysed using LnP(K) trend to evaluate best K by STRUCTURE HARVESTER. The plot did not display a plateau, that it's expected when a population structure is detected. This would seem to indicate K=1 as a correct value and a lack of genetic differentiation between the geographic populations. The plot showed an increase of LnP(K) variance between runs in relation to the increase of K (Fig.19).

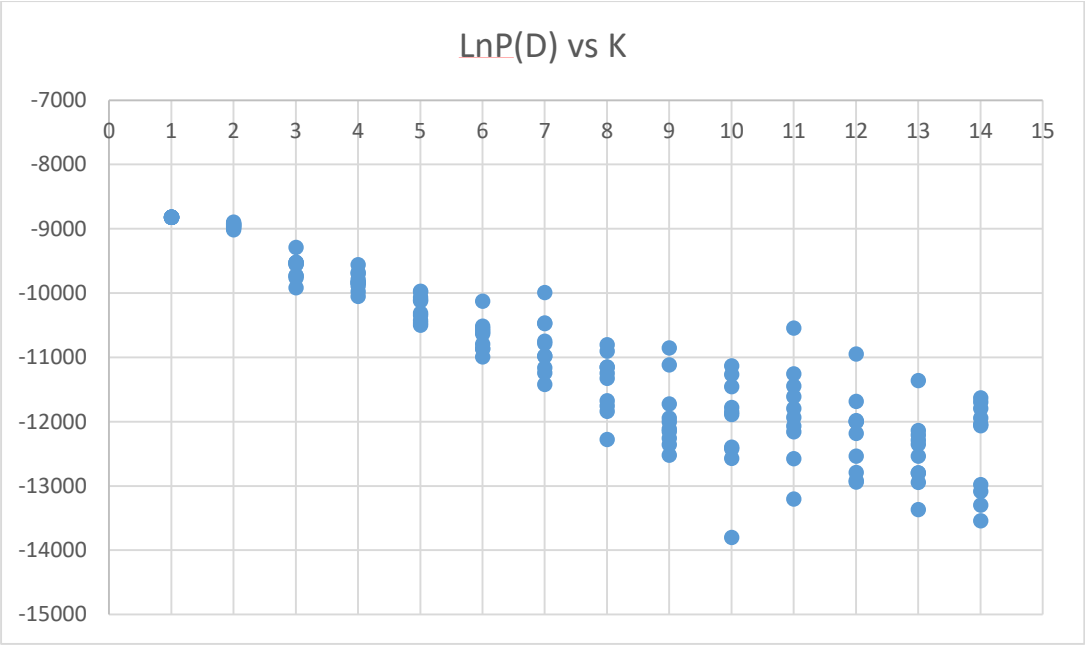


Figure 19. LnP(K) graph resulting from STRUCTURE HARVESTER.

Individual bar plots from K = 2 to K = 5, obtained with the web based CLUMPAK, are presented in Fig. 20. The results showed an admixed genetic component across all the investigating areas.

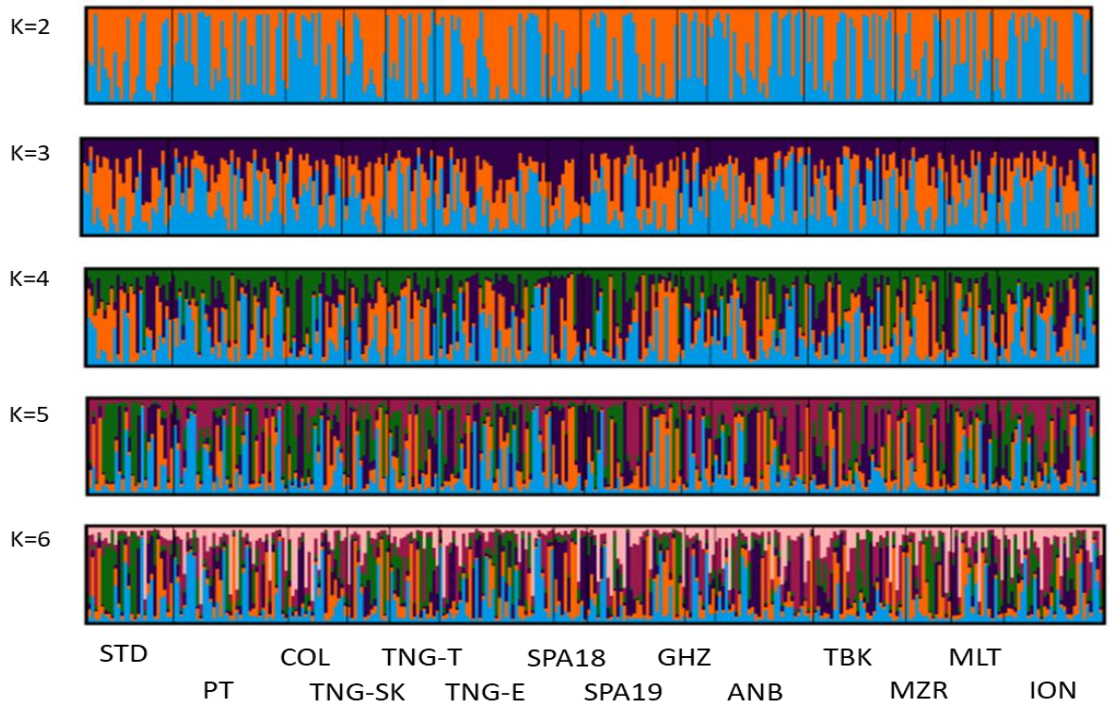


Figure 20. Summary plot of the STRUCTURE analysis from K=2 to K=5. Each individual is represented by a single vertical line on the x-axis. Coloured segments represent the membership of individual to the population clusters. Results are show for the 14 locations subdivision using 14 microsatellite loci

BOTTLENECK analysis showed two significant results (with p-value:0.05) at COL+TRF and TNG-T populations both for tail 1 (deficit of heterozygotes) and tail 2 (excess of heterozygotes; Tab. 14).

*Table 14 Results from BOTTLENECK software; in **bold significant result at alpha: 0.05**.*

Population	Mean N	Mean k	Mean He	p_W_1t_TPM	p_W_2t_TPM
STD	51.7	9.4	0.839	0.469	0.938
PT	69.7	9.7	0.838	0.234	0.469
COL+TRF	36.0	8.1	0.839	**0.020**	**0.039**
TNG-SK	24.3	7.9	0.829	0.148	0.297
TNG-T	29.7	7.4	0.828	**0.012**	**0.023**
TNG-E	69.1	10.6	0.847	0.148	0.297
SPA18	19.7	7.4	0.845	0.406	0.813
SPA19	58.6	10.3	0.852	0.188	0.375
GHZ	18.0	7.0	0.842	0.148	0.297
ANB	58.9	9.9	0.832	0.594	0.938
TBK	55.4	10.0	0.839	0.656	0.813
MZR	27.4	8.4	0.832	0.594	0.938
MLT	31.7	8.1	0.837	0.148	0.297
ION	59.1	9.4	0.838	0.289	0.578

3.3.3. Statistical power of microsatellite loci

Analyses of statistical power conducted with POWSIM indicated that all marker sets will detect F_{ST}= 0.005 or larger with a probability close to one, for present sample sizes and allele frequencies. At lower levels of divergence, 20 and 14 microsatellites are more powerful for detecting structuring than 7 microsatellite loci (Tab. 15).

The Chi-square and Fisher approach α-error is close to 5% at all marker sets even if the Fisher approach appears overly conservative with an α error of only 0.46 – 0.82%.

Table 15. Statistical power of different marker sets for detecting various levels of population differentiation (FST) by both Chi-square and Fisher's method when using samples size and allele frequencies obtained by the samples in this study.

	20 loci		14 loci		7 loci	
Fst	**Chi-square**	**Fisher**	**Chi-square**	**Fisher**	**Chi-square**	**Fisher**
0.001	0.644	0.577	0.524	0.464	0.265	0.243
0.0025	0.998	0.997	0.966	0.937	0.662	0.615
0.005	1.000	1.000	1.000	1.000	0.990	0.987
0.010	1.000	1.000	1.000	1.000	1.000	1.000
0.02	1.000	1.000	1.000	1.000	1.000	1.000

4. Discussion

This study assessed the genetic variation of microsatellite loci across different geographical samples of *P. bogaraveo* caught from different fishing areas in the Atlantic Ocean and the Mediterranean Sea, covering the majority of *P. bogaraveo* dispersal area.

The main aim was to study the genetic population structure of *P. bogaraveo* inside the Alboran Sea in order to start setting the basis for a correct management based on the scientific approach and to avoid the overexploitation and depletion of resident stocks.

4.1. Quality of microsatellite loci

The failure of cross amplification test on *P. bogaraveo* of those microsatellite loci originally isolated in *P. erythrinus* could be related to the species-specific trait of the markers. This represents a considerable disadvantage in relation to other molecular markers as SNPs (Single Nucleotide Polymorphisms; Pinera et al., 2007b). In general, this condition underlines the need for developing new microsatellites for target species since, as in this case, those developed from the most close-related species resulted ineffective.

Moreover, the use of a multiplex amplification could result in problematic scoring of the microsatellite loci because of issues occurring during the PCR. As a matter of fact, the low intensity of peaks, the presence of stutter bands and split peaks, pull-outs of different fluorochromes and primer competition may interfere with the downstream allele calling. To avoid this inconvenience, a good arrangement of the loci during the experimental design and the optimisation of amplification conditions are strongly recommended. The test on the congruence between single-locus and multiplex amplification in terms of allele calling, did not highlight differences between 8 individuals, except one single case, where the first allele did not amplify.

FInally, the analysis performed with MICROCHECKER and FreeNA for the assessment of Null Allele (NA) frequencies, showed that the highest values were mostly observed in the microsatellite loci isolated and characterized in individuals collected from the Azores area (Stockley et al., 2000). In fact, previous studies detected a marked differentiation between Azores individuals and European ones (Bargelloni et al., 2003; Stockley et al., 2005; Lemos et al., 2006 and Piñera et al., 2007a). We hypothesize that the use of microsatellite loci developed from a genetically distinct population may increase the frequencies of NA because of possible mutations at primers binding sites (Jones et al., 1998, Wen et al., 2013). Thus, including samples from the Azores area may be effective in testing this hypothesis.

As a consequence of the results on NA and Hardy-Weinberg disequilibrium, we decided to test statistical power when reduced marker sets were used to evaluate genetic differentiation. The results obtained with POWSIM confirmed the high power of all the marker sets on F_{ST} values equal or greater of 0.05 and even 0.0025 when 14 microsatellite loci are used (Tab. 15). This value is satisfactory when we considered that significant values of F_{ST} found in this study had the same magnitude (Tabs 9 and 12).

4.2. Genetic diversity of *P. bogaraveo*

The genetic diversity of *P. bogaraveo* populations in terms of the number of alleles per locus per population and expected heterozygosity was similar or slightly major to values from literature (on the genetic population structure) for the majority of loci (Tab. 15; Stockley et al., 2005; Piñera et al., 2007a).

Table 15. Genetic diversity value found in this thesis compared with reference data. N allele: number of alleles per locus, Hobs: observed heterozygosity; Hexp: expected heterozygosity; FIS: fissation index; ND: no data available. Reference data from Pinera et al. 2007a (range for Hobs and Hexp), Stockley et al. 2005.

	Reference Data				Study Data			
Locus	**N. samples**	**N. alleles**	**Hobs**	**Hexp**	**N. samples**	**N. alleles**	**Hobs**	**Hexp**
Pb-OVI-B2	123	24	0.853	0.905	309	23	0.823	0.862
PbMS2	350	ND	0.728-0.611	0.899-0.873	309	20	0.556	0.848
PbMS6	350	ND	1.000-0.873	0.969-0.938	309	37	0.866	0.937
PbMS16	350	ND	0.983-0.915	0.964-0.945	309	42	0.861	0.937
PbMS19	350	ND	0.516-0.458	0.628-0.486	309	10	0.587	0.665
Pb-OVI-D108	123	5	0.407	0.496	309	7	0.462	0.456
Pb-OVI-A5	123	37	0.713	0.946	309	41	0.826	0.914
Pb-OVI-D114	123	11	0.780	0.833	309	11	0.736	0.812
Pb-OVI-D102	123	21	0.851	0.899	309	20	0.718	0.879
Pb-OVI-D20	123	10	0.811	0.840	309	12	0.861	0.820
Pb-OVI-D21	123	13	0.882	0.863	309	14	0.871	0.842
Pb-OVI-D106	123	10	0.765	0.800	309	11	0.810	0.788
Pb-OVI-A3	123	18	0.868	0.910	309	22	0.712	0.884
PbMS20	350	ND	0.918-0.833	0.937-0.850	309	21	0.446	0.830
PbMS18	350	ND	ND	ND	309	23	0.502	0.832
PbMS1	350	ND	0.577-0.353	0.962-0.931	309	38	0.463	0.917
Pb-OVI-D22	123	ND	0.91	0.855	309	12	0.825	0.821
PbMS4	350	ND	ND	ND	309	29	0.899	0.905
PbMS15	350	ND	0.932-0.840	0.955-0.923	309	40	0.740	0.924
Pb-OVI-C103	123	9	0.762	0.784	309	11	0.719	0.758

Higher values of number of alleles in Pb-OVI-D108, Pb-OVI-A5, Pb-OVI-D20, Pb-OVI-D21, Pb-OVI-D106, Pb-OVI-A3 and Pb-OVI-C103, could be related to the higher number of samples analysed in the present study. As a matter of fact, we highlighted the positive correlation between the mean number of alleles and the number of samples for populations (Fig.4).

Significant Hardy Weinberg disequilibrium was observed in almost the loci, except for Pb-OVI-B2, Pb-OVI-D114, Pb-OVI-D20, Pb-OVI-D21, Pb-OVI-D106, Pb-OVI-D22 and Pb-OVI-C103. This deviation from panmixia can be caused by various processes like selective forces against heterozygotes, spatial and/or temporal Wahlund effects, inbreeding or presence of NA. In our study, the presence of NA across loci and populations was confirmed both by FreeNA and MICROCHECKER and it seems the most probable driver for the Hardy Weinberg disequilibrium observed.

Regarding the hypothesis of inbreeding, Fassatoui et al. (2011), explained that "the phenomenon must involve the whole genome and thus we should expect to observe a deficit in heterozygotes for all the loci analysed" which is not the case in the present study. Moreover, the results from BOTTLENECK did not identify a bottleneck event that could justify the reduction of population size. Nevertheless, this hypothesis cannot be totally excluded since *Pagellus bogaraveo* is subject to overexploitation that can increase the effects of genetic drift and inbreeding over long timescales.

4.3. Population differentiation

All the analyses performed on both datasets involving 14 and seven loci, agreed on the lack of genetic structure across all the investigating populations. The only exception was represented by the significant but low F_{ST} values measured among pairs of samples (Tabs. 9 and 12). According to AMOVA (Tabs. 10 and 13) the highest variation was observed at the individual level: the percentage of variation attributed to differences "within individuals" class are explained by the 94.63 – 99.43% (in AMOVA test 1 with 14 loci and AMOVA test 1 with 7 loci, respectively). The high variation within individuals measured in the present study likely reflects the high genetic diversity values found.

The DAPC plot resulting from analyses performed with 14 microsatellite loci and the three different groupings (14 locations, 6 macroareas and 9 GSAs), showed a scattered pattern, confirming the lack of genetic differentiation between populations and the great variability within samples (Figs. 5, 6 and 7). In the analysis performed using 6 macroareas, only S-ALB seemed to be more differentiated, but this group contains only 9 individuals. Similar, in the analysis performed using 9 GSAs, GSA15, GSA16 and GSA1_NWGS seemed to be more differentiated, but they contain the lowest number of individuals: 16, 14 and 18 respectively.

The same lack of differentiation was obtained using seven microsatellite loci and the same groupings (Figs. 13, 14 and 15). The slightly differentiation found between macroareas and GSAs groups previously described, were no more present.

The most important results obtained using the PCoA analyses is the similarity between the samples from Atlantic area (like PT, STD and COL) with the ones from Central and, especially, East Mediterranean areas (like TBK, MLT and MZR in the first case and ION in the second). This pinpointed the lack of genetic differentiation between samples collected very far away from each other (Figs. 8, 9 and 10)

The same results were obtained using seven microsatellite loci and the same groupings (Figs. 16, 17 and 18).

The structure output and the results from STRUCTURE HARVESTER and CLUMPAK (Figs. 11, 12, 19 and 20) revealed a homogeneous distribution of genetic components across sampling locations, underling the lack of distinct clusters of *P. bogaraveo,* regardless of the dataset used for the analyses.

The absence of population structure resulting when using seven markers is congruent with the findings obtained when considering 14 microsatellite loci. The use of POWSIM and the high power of our marker sets resulting, allow us to say that the choice of loci based on the HWE principle and absence of scoring errors didn't lead to the exclusion of informative loci, undermining the resolution power of this dataset.

To date, there is no consensus about the allowable percentage of null alleles across loci and geographical samples. According to Carlsson, (2008), the bias caused by null alleles may lead only to a slight reduction in the power of Bayesian tests as STRUCTURE in assigning individuals. On the other hand, Wen et al. (2013) described how null alleles have significant effects on the results of genetic analysis, potentially decreasing population genetic diversity and increasing genetic differentiation among populations. Previous studies concluded that microsatellite loci affected by null alleles would probably not alter the overall outcome of assignment testing and could therefore be included in genetic studies. In all, it seems that the number of loci and level of genetic differentiation have greater effects on the accuracy of assignment tests than the effects related to the presence/absence of null alleles (Carlsson 2008, Wen et al., 2013, Wu et al., 2019).

Stockley et al. (2005), Lemos et al. (2006) and Piñera et al. (2007) hypothesized different reasons to explain the lack of differentiation between *P. bogaraveo* samples:

- *P. bogaraveo* has been subjected to heavy fishing pressure that has impacted the larger size classes. As the species is a protandrous hermaphrodite, fishing activities led to higher mortality rates amongst females and consequently led to low diversity in mitochondrial genes.

- The dispersal biology of blackspot seabream allows gene flow among close regions within the slope, even though the Oran-Almeria Front, whereas it is not capable of crossing large hydrographical barriers such as ocean basins.
- The divergence among samples is temporally recent and differentiation events have not occurred yet.

For what concerns this last hypothesis, Stockley et al., 2005, Piñera et al., 2007 and Bargelloni et al., 2003 suggested that a recent re-colonization of the Mediterranean Sea from the Atlantic Ocean could have happened after a strong bottleneck in Mediterranean populations. This would have implied that, despite the low migration rate across the Strait of Gibraltar, an insufficient time allowed the accumulation of genetic differences among the populations of the two basins (Patarnello et al. ,2007). Furthermore, individuals from the Azores Islands could be genetically different from continental individuals of blackspot seabream, whereas partial and shallow differences have been found within the continental slope to date. We can hypothesize that including Azores' individuals may help in confirming the pattern of differentiation previously suggested.

5. Conclusions

Evidence produced within this work suggested that if any restriction to *P. bogaraveo* gene flow is present between Mediterranean and Atlantic areas, it may not be strong enough to drive significant genetic differentiation among populations.

It is the opinion of the authors that the differences previously reported between Atlantic and Mediterranean regions are the consequence of the experimental design including (or not) individuals collected in the Azores Islands.

More exhaustive analyses on the biology of blackspot seabream and its dispersal capability should be performed in the next future. Moreover, we strongly suggest the use of mitochondrial DNA on one hand (Robalo et al., 2021) and Single Nucleotide Polymorphisms (SNPs) on the other, in order to detect any genetic differentiation of this species.

Lastly, new studies using morphometric measurements could identify significant differences between Atlantic and Mediterranean samples, as reported by Palma & Andrade (2004). This could contribute to the holistic approach requested for stock assessment, particularly recommended for *P. bogaraveo* to avoid the collapse of the species in areas where fishery is not adequately regulated.

6. REFERENCES

BARGELLONI L., ALARCON J. A., ALVAREZ M. C., PENZO E., MAGOULAS A., REIS C., PATARNELLO T. (2003) DISCORD IN THE FAMILY SPARIDAE (TELEOSTEI): DIVERGENT PHYLOGEOGRAPHICAL PATTERNS ACROSS THE ATLANTIC–MEDITERRANEAN DIVIDE. J. EVOL. BIOL. 16, 1149–1158.

BELKHIR K.; BORSA P.; CHIKHI L.; RAUFASTE N.; BONHOMME F. (2004). GENETIX4. 05, LOGICIEL SOUS WINDOWS TM POUR LA GÉNÉTIQUEDES POPULATIONS. LABORATOIRE GÉNOME, POPULATIONS, INTERACTIONS, CNRS UMR. 5000, 1996-2004

BERTRAND J.A.; DE SOLA L.G.; PAPACONSTANTINOU C.; RELINI G.; SOUPLET A. (2002) THE GENERAL SPECIFICATION OF THE MEDITS SURVEYS. SCI.MAR., 66 (SUPPL. 2) :9-17.

CARLSSON J. (2008) EFFECTS OF MICROSATELLITE NULL ALLELES ON ASSIGNMENT TESTING. JOURNAL OF HEREDITY 99(6):616–62

CHAPUIS, M.P., AND A. ESTOUP. (2007). MICROSATELLITE NULL ALLELES AND ESTIMATION OF POPULATION DIFFERENTIATION. MOL. BIOL. EVOL. 24(3): 621-631.

EARL, D. A. & VONHOLDT, B. M. (2012) STRUCTURE HARVESTER: A WEBSITE AND PROGRAM FOR VISUALIZING STRUCTURE OUTPUT AND IMPLEMENTING THE EVANNO METHOD. CONSERVATION GENETICS RESOURCES VOL. 4 (2) PP. 359-361.

EXCOFFIER L.; SMOUSE P.; QUATTRO J. (1992) ANALYSIS OF MOLECULAR VARIANCE INFERRED FROM METRIC DISTANCES AMONG DNA HAPLOTYPES: APPLICATION TO HUMAN MITOCHONDRIAL DNA RESTRICTION DATA. GENETICS, 131, 479-491.

EXCOFFIER L. & LISCHER H. (2010) ARLEQUIN SUITE VER 3.5: A NEW SERIES OF PROGRAMS TO PERFORM POPULATION GENETICS ANALYSES UNDER LINUX AND WINDOWS. MOLECULAR ECOLOGY RESOURCES, 10, 564-577.

FASSATOUI C., MDELGI E., MS R.. (2011). SHORT-TERM TEMPORAL GENETIC INVESTIGATION OF TWO POPULATIONS OF *PAGELLUS ERYTHRINUS* (LINNAEUS 1758, SPARIDAE) FROM WESTERN AND EASTERN MEDITERRANEAN BASINS OFF THE TUNISIAN COAST. MARINE BIOLOGY RESEARCH. 7. 147–158.

FLORES-RENTERIA L. AND KROHN A. (2013) SCORING MICROSATELLITE LOCI. IN: KANTARTZI S. (EDS) MICROSATELLITES. METHODS IN MOLECULAR BIOLOGY (METHODS AND PROTOCOLS), VOL 1006, 319- 336. HUMANA PRESS, TOTOWA, NJ

GOUDET J. (2001) FSTAT, VERSION 2.9. 3, A PROGRAM TO ESTIMATE AND TEST GENE DIVERSITIES AND FIXATION INDICES. LAUSANNE UNIVERSITY, LAUSANNE, SWITZERLAND

GUICHOUX E., LAGACHE L., WAGNER S., CHAUMEIL P., LÉGER P., LEPAIS O., LEPOITTEVIN C., MALAUSA T., REVARDEL E., SALIN F. PETIT R.J. (2011) CURRENT TRENDS IN MICROSATELLITE GENOTYPING. MOLECULAR ECOLOGY RESOURCES 11, 591–611

ICES (2018A.) REPORT OF THE WORKING GROUP ON THE BIOLOGY AND ASSESSMENT OF DEEP-SEA FISHERIES RESOURCES (WGDEEP), 11–18 APRIL 2018, ICES HQ, COPENHAGEN, DENMARK. ICES CM 2018/ACOM:14. 771 PP.

ICES. (2019) WORKING GROUP ON THE BIOLOGY AND ASSESSMENT OF DEEP-SEA FISHERIES RESOURCES (WGDEEP). ICES SCIENTIFIC REPORTS. 1,21. 988 PP

JOMBART T., AHMED I. (2011). "ADEGENET 1.3-1: NEW TOOLS FOR THE ANALYSIS OF GENOME-WIDE SNP DATA." BIOINFORMATICS

JONES A., STOCKWELL C., WALKER D., AVISE J.C. (1998). THE MOLECULAR BASIS OF A MICROSATELLITE NULL ALLELE FROM THE WHITE SANDS PUPFISH. JOURNAL OF HEREDITY. 89.

KOPELMAN N. M.; MAYZEL J.; JAKOBSSON M.; ROSENBERG N. A; MAYROSE, I. (2015) CLUMPAK: A PROGRAM FOR IDENTIFYING CLUSTERING MODES AND PACKAGING POPULATION STRUCTURE INFERENCES ACROSS K. MOLECULAR ECOLOGY RESOURCES 15(5): 1179-1191

LEMOS A., FREITAS A.I., FERNANDES A.T., GONÇALVES R., JESUS J., ANDRADE C., BREHM A. (2006) MICROSATELLITE VARIABILITY IN NATURAL POPULATIONS OF THE BLACKSPOT SEABREAM PAGELLUS BOGARAVEO (BRÜNNICK, 1768): A DATABASE TO ACCESS PARENTAGE ASSIGNMENT IN AQUACULTURE. AQUACULTURE RESEARCH, 1-6.

LORANCE, P. (2011) HISTORY AND DYNAMICS OF THE OVEREXPLOITATION OF THE BLACKSPOT SEA BREAM (PAGELLUS BOGARAVEO) IN THE BAY OF BISCAY. ICES JOURNAL OF MARINE SCIENCE, 68, 290–301.

LUIKART, G., & CORNUET, J. M. (1998). EMPIRICAL EVALUATION OF A TEST FOR IDENTIFYING RECENTLY BOTTLENECKED POPULATIONS FROM ALLELE FREQUENCY DATA. CONSERVATION BIOLOGY, 12(1), 228-237.

PALMA J. & ANDRADE J.P. (2004). MORPHOLOGICAL STUDY OF *PAGRUS PAGRUS*, *PAGELLUS BOGARAVEO*, AND *DENTEX DENTEX* (SPARIDAE) IN THE EASTERN ATLANTIC AND THE MEDITERRANEAN SEA. JOURNAL OF THE MARINE BIOLOGICAL ASSOCIATION OF THE UK. 84. 449 - 454.

PATARNELLO T.; VOLCKAERT F. A. M. J.; CASTILHO R. (2007) PILLARS OF HERCULES: IS THE ATLANTIC–MEDITERRANEAN TRANSITION A PHYLOGEOGRAPHICAL BREAK? MOLECULAR ECOLOGY 16, 4426–4444

PEAKALL, R. & SMOUSE P.E. (2012) GENALEX 6.5: GENETIC ANALYSIS IN EXCEL. POPULATION GENETIC SOFTWARE FOR TEACHING AND RESEARCH-AN UPDATE. BIOINFORMATICS 28, 2537-2539.

PIÑERA J.A.; BERNARDO D.; BLANCO G.; VAZQUEZ E.; SANCHEZ J.A. (2006) ISOLATION AND CHARACTERIZATION OF POLYMORPHIC MICROSATELLITE MARKERS IN PAGELLUS BOGARAVEO, AND CROSS-SPECIES AMPLIFICATION IN SPARUS AURATA AND DICENTRARCHUS LABRAX. MOLECULAR ECOLOGY NOTES 6, 33-35

PIÑERA J. A., BLANCO G., VÁZQUEZ E., SÁNCHEZ J.A. (2007A) GENETIC DIVERSITY OF BLACKSPOT SEABREAM (PAGELLUS BOGARAVEO) POPULATIONS OFF SPANISH COASTS: A PRELIMINARY STUDY. MAR BIOL 151,2153–2158.

PIÑERA J. A; BERNARDO D.; BLANCO G.; VÁZQUEZ E., SÁNCHEZ J.A. (2007B) USEFULNESS OF MICROSATELLITE MARKERS DEVELOPED FROM PAGELLUS BOGARAVEO TO GENETICALLY STUDY FIVE DIFFERENT SPECIES OF SPARIDAE. MARINE ECOLOGY 2007, 28 (SUPPL. 1), 184–187.

PINHO M., DIOGO H., CARVALHO J., PEREIRA J.G. (2014) HARVESTING JUVENILES OF BLACKSPOT SEA BREAM (PAGELLUS BOGARAVEO) IN THE AZORES (NORTHEAST ATLANTIC): BIOLOGICAL IMPLICATIONS, MANAGEMENT, AND LIFE CYCLE CONSIDERATIONS. ICES JOURNAL OF MARINE SCIENCE, 71(9), 2448–2456.

PIRY S., LUIKART G., & CORNUET J.M. (1999). BOTTLENECK: A PROGRAM FOR DETECTING RECENT EFFECTIVE POPULATION SIZE REDUCTIONS FROM ALLELE DATA FREQUENCIES. JOURNAL OF HEREDITY, 90(4), 502-503.

PRITCHARD J.; STEPHENS M.J.; DONNELLY P.J. (2000) INFERENCE OF POPULATION STRUCTURE USING MULTILOCUS GENOTYPE DATA. GENETICS, 155, 945-959

RAMÍREZ F, PENNINO MG, ALBO-PUIGSERVER M, STEENBEEK J, BELLIDO JM, COLL M. 2021. SOS SMALL PELAGICS: A SAFE OPERATING SPACE FOR SMALL PELAGIC FISH IN THE WESTERN MEDITERRANEAN SEA. SCIENCE OF THE TOTAL ENVIRONMENT, 756, 144002

RAMŠAK A.; GAROIA F.; GUARNIERO I.; MANNINI P.; TINTI F. (2003) NOVEL POLYMORPHIC MICROSATELLITE MARKERS FOR THE COMMON PANDORA (PAGELLUS ERYTHRINUS). MOLECULAR ECOLOGY NOTES, 3, 553-555

RAYMOND M. & ROUSSET F. (1995) GENEPOP (VERSION 1.2): POPULATION GENETIC SOFTWARE FOR EXACT TESTS AND ECUMENICISM. J HERED, 86.

ROBALO J., FARIAS I., FRANCISCO S., AVELLANEDA K., CASTILHO R., FIGUEIREDO I. (2021). GENETIC POPULATION STRUCTURE OF THE BLACKSPOT SEABREAM (*PAGELLUS BOGARAVEO*): CONTRIBUTION OF MTDNA CONTROL REGION TO FISHERIES MANAGEMENT. MITOCHONDRIAL DNA. PART A, DNA MAPPING, SEQUENCING, AND ANALYSIS. 32. 1-5.

ROUSSET F. (2008) GENEPOP'007: A COMPLETE RE-IMPLEMENTATION OF THE GENEPOP SOFTWARE FOR WINDOWS AND LINUX. MOLECULAR ECOLOGY RESOURCES 8, 103–106

RYMAN N., PALM S. (2006). POWSIM: A COMPUTER PROGRAM FOR ASSESSING STATISTICAL POWER WHEN TESTING FOR GENETIC DIFFERENTIATION. MOL ECOL NOTES 6: 600–602.

RYMAN N., PALM S., ANDRE´ C., CARVALHO G.R., DAHLGREN T.G., JORDE P.E. (2006). POWER FOR DETECTING GENETIC DIVERGENCE: DIFFERENCES BETWEEN STATISTICAL METHODS AND MARKER LOCI. MOL ECOL 15: 2031–2045.

SAMMARTINO S.; SÁNCHEZ-GARRIDO J. C.; NARANJO C.; GARCÍA LAFUENTE J.; RODRÍGUEZ RUBIO P.; SOTILLO M. (2018) WATER RENEWAL IN SEMI-ENCLOSED BASINS: A HIGH RESOLUTION LAGRANGIAN APPROACH WITH APPLICATION TO THE BAY OF ALGECIRAS, STRAIT OF GIBRALTAR. LIMNOLOGY AND OCEANOGRAPHY: METHODS, 16, 106 -118.

SANZ-FERNÁNDEZ V., GUTIÉRREZ-ESTRADA J.C., PULIDO-CALVO I., GIL-HERRERA J., BENCHOUCHA S., EL ARRAF S. (2019) ENVIRONMENT OR CATCHES? ASSESSMENT OF THE DECLINE IN BLACKSPOT SEABREAM (PAGELLUS BOGARAVEO) ABUNDANCE IN THE STRAIT OF GIBRALTAR. JOURNAL OF MARINE SYSTEMS 190, 15– 24.

SCHICKELE A, GOBERVILLE E, LEROY B, BEAUGRAND G, HATTAB T, FRANCOUR P, & RAYBAUD V. 2021. EUROPEAN SMALL PELAGIC FISH DISTRIBUTION UNDER GLOBAL CHANGE SCENARIOS. FISH AND FISHERIES.

STOCKLEY B.M; ROGERS A.D.; IYENGAR A.; MENEZES G.; SANTOS R.; LONG A. (2000) TEN MICROSATELLITE LOCI ISOLATED AND DEVELOPED FOR THE BLACKSPOT SEABREAM, PAGELLUS BOGARAVEO (BRÜNNICH 1768). MOLECULAR ECOLOGY 9, 993-1011.

STOCKLEY B., MENEZES G., PINHO M.R., ROGERS A.D. (2005) GENETIC POPULATION STRUCTURE IN THE BLACK-SPOT SEA BREAM (PAGELLUS BOGARAVEO BRÜNICH, 1768) FROM THE NE ATLANTIC. MARINE BIOLOGY 146, 793–804.

VAN OOSTERHOUT C., HUTCHINSON W.F., WILLS D.P.M., SHIPLEY P. (2004) MICRO-CHECKER: SOFTWARE FOR IDENTIFYING AND CORRECTING GENOTYPING ERRORS IN MICROSATELLITE DATA. MOLECULAR ECOLOGY NOTES 4, 535–538

WEN Y., UCHIYAMA K., WENJUN H., UENO S., WEIDONG X., XU G., TSUMURA Y. (2013). NULL ALLELES IN MICROSATELLITE MARKERS. BIODIVERSITY SCIENCE. 21. 117-126.

WU X., WANG L., ZHANG D., WEN Y. (2019). MICROSATELLITE NULL ALLELES AFFECTED POPULATION GENETIC ANALYSES: A CASE STUDY OF MAIRE YEW (TAXUS CHINENSIS VAR. MAIREI). JOURNAL OF FOREST RESEARCH. 24. 1-5.