

**Establishment of genetic markers for  
breeding the white grouper (*Epinephelus  
aeneus*)**

**Thesis**

**Submitted to the Robert H. Smith Faculty of Agriculture, Food  
and Environment, the Hebrew University of Jerusalem, in  
partial fulfillment of the requirements for M.Sc. degree in  
agricultural sciences**

**By**

**Lior Dor**

**This thesis was conducted under the supervision of Professor Gideon  
Hulata and Professor Micha Ron**

**The Institute of Animal Science, the Volcani Center, Agricultural  
Research Organization**

# Contents

	Page
Abstract	
1. Introduction	
1.1 Groupers	1
1.2 <i>Epinephelus aeneus</i> compatibility for aquaculture	2
1.3 <i>Epinephelus aeneus</i> culture in Israel	3
1.4 DNA-based identification	5
1.5 Genetic markers	6
1.6 Microsatellite markers	6
1.7 SNP markers	8
1.8 Genetic linkage map	9
1.9 Research objectives	10
2. Materials and Methods	
2.1 Abbreviations used and details of solutions	11
2.2 Research population	11
2.3 DNA extraction	13
2.3.1 Genomic DNA extraction from blood samples	13
2.3.2 Genomic DNA extraction from wild type and larvae tissue samples	14
2.3.3 Genomic DNA extraction from fingerling tissue samples	14
2.4 Concentration determination	15
2.5 COI analysis	15
2.6 PCR amplification of COI fragments	16
2.7 Fragments separation	16
2.8 Fragment sequencing	17
2.9 Phylogeny of <i>Epinephelus</i> species	17
2.10 Development of heterologous markers	17
2.11 Two-step PCR procedure	18
2.12 PCR reaction	18

2.12.1 PCR amplification program	20
2.13 PCR products detection	20
2.14 Genotyping procedure	20
2.15 Parentage analysis	21
2.16 Statistical analysis	22
2.17 Polymorphic Information Content (PIC)	22
3. Results	
3.1 Development of microsatellite markers	23
3.2 Genotyping of microsatellites	27
3.3 Polymorphism of heterologous microsatellites	29
3.4 Polymorphism of microsatellites in the wild and captive populations	30
3.5 COI sequences comparison of <i>Epinephelus</i> species	32
3.6 Parentage analysis	35
3.7 Linkage map population	36
4. Discussion	
4.1 Microsatellite markers development	37
4.2 Cost effective amplification	38
4.3 Heterologous marker polymorphism according to species of origin	39
4.4 Phylogenetic tree of species	39
4.5 Reproduction strategy	41
4.6 Genetic variation	42
4.7 Genetic mapping of markers	44
5. References	45
6. Appendix I - Sequences of polymorphic primers	56

## Abstract

The white grouper (*Epinephelus aeneus*) is a protogynous hermaphrodite species found throughout the southern Mediterranean Sea and along the west coast of Africa to southern Angola. The species has a high market value, excellent taste, rapid growth rate, hardiness and relatively high disease resistance. Therefore, it is a promising candidate for intensive aquaculture and in 2010 an *Epinephelus aeneus* domestication project was launched by the department of fisheries and aquaculture of the Israeli ministry of agriculture. The mission of the project is to develop innovative technologies aimed at domestication, breeding, and industrial rearing of the white grouper. The major obstacles in this procedure are: Accurate identification of species by relying only on morphologic characteristic, monitoring of genetic variance reduction within the captive population, and understanding of *Epinephelus aeneus* reproduction strategy i.e. how many brooders within a group actually contribute genetic material to the offspring.

The current research deals with these issues using several genomic tools. First a DNA identification method is applied based on sequence diversity in mitochondrial Cytochrome C Oxidase C subunit 1 gene (COI) in order to identify the specimens used in the research. Phylogenetic DNA- and protein-based trees of 36 different *Epinephelus* species were constructed to determine the correlation between phylogenetic distance and effective heterologous amplification of genetic markers in *E. aeneus*. The markers were used to determine the rate of polymorphism in the captive and wild populations. Finally, polymorphic microsatellite markers were used as parentage test to assign larvae (n=59) spawned by the captive broodstock to their parents. In addition, a larger population of the captive broodstocks' fingerlings was obtained in order to locate a population, using the parentage test set, which will form a segregating population for construction of the first generation linkage map.

In this research a total of 656 published sequences of microsatellites in *Epinephelus* species were found, from which 233 were suitable for primers design. Heterologous markers were developed from eight different *Epinephelus* species (*E. awoara*, *E. septemfasciatus*, *E. fuscoguttatus*, *E. itajara*, *E. guttatus*, *E. akaara*, *E. coiodes* and *E.*

*lanceolatus*). A cost effective two-step PCR method was applied for product amplification of each set of primers in which only two fluorescent primers were used for efficient detection and genotyping of all 233 PCR products. A total of 144 loci were found polymorphic and the yield of polymorphic microsatellites by species of origin ranged from 42 to 83%. The lowest yield was found for *E. septemfasciatus* ( $P = 0.001$ ) which according to the Cytochrome c oxidase subunit I gene (COI) analysis was the most phylogenetically distant from *E. aeneus*. However, all selected *Epinephelus* species were found suitable for development of heterologous markers. An average number of 6.2 alleles ranging from 2 to 14 per marker was found in the wild type population versus 3.7 from 2 to 7 per marker in the captive population ( $P = 1.85 \times 10^{-20}$ ) suggesting a reduction in genetic variation among the captive broodstock population comparing to the wild.

For parentage analysis of the offspring population a set of eight polymorphic microsatellites was developed. Among the twelve brooders two males and two females were found to contribute genetic material in various pairing combinations, whereas one pair was shown to be the parents of most of the progeny (69.2%) thus suggesting the reproduction mode of *E. aeneus* as polygamous.

In light of the fact that only a small number of brooders contributed genetic material to the offspring and the detected decrease in allelic diversity among captive broodstock compared to the wild population of *E. aeneus*, the conclusion is that polymorphism level in the captive broodstock must be further monitored for potential loss of variance. Furthermore, future culture programs must enhance the captive stock genetic diversity by analysis of genetic markers.

The parentage test is being used to establish a segregating population for linkage mapping of all available genetic markers. In the future additional markers can be developed using either next generation sequencing or heterologous amplification of known microsatellites in related species thus creating a white grouper linkage map that will be an infrastructure for domestication and breeding.

# 1. Introduction

## 1.1 Groupers

Groupers are members of the family Serranidae, sub-family Epinephelinae. There are about 159 species of groupers in 15 genera. They are bottom-associated fishes found in the tropical and subtropical waters of all oceans and their majority inhabits depths less than 100 m whereas juveniles are often found in tide-pools (Heemstra and Randall, 1993). Groupers are popular carnivorous fish with a high market demand in many parts of the world and they are ideal candidate species for intensive aquaculture particularly in the Asia–Pacific region because of high consumer demand, desirable taste, hardiness in a crowded environment, fast growth, efficient feed conversion and rapid growth (Harikrishnana et al., 2010). Groupers are cultured in many Southeast Asian countries, including Indonesia, Malaysia, Philippines, Taiwan, Thailand, Hong Kong, Mainland China and Vietnam. They are cultured also in other parts of the tropics in the south eastern USA and Caribbean, India, Sri Lanka, Saudi Arabia, Republic of Korea and Australia (Ottolenghi et al., 2004).

The most common species for culture are orange-spotted grouper (*E. coioides*), brown-marbled grouper (*E. fuscoguttatus*), Malabar grouper (*E. malabaricus*), dusky-tail grouper (*E. bleekeri*), giant grouper (*E. lanceolatus*), hump-back grouper (*Cromileptes altivelis*) and coral trout/leopard grouper (*Plectropomus leopardus* and). Other important species include red-spotted grouper (*E. akaara*) seven-band grouper (*E. septemfasciatus*) and long-tooth grouper (*E. bruneus*) (Liao and Leaña, 2008). Due to the demand for groupers meat culture technologies for additional species are being developed, especially in Taiwan and China, including species like the three-spotted grouper (*E. trimaculatus*), banded grouper (*E. amblycephalus*), areolate grouper (*E. areolatus*), potato grouper (*E. tukula*) and speckled blue grouper (*E. cyanopodus*) (Liao and Leaña, 2008).

Global grouper production increased dramatically in recent years, with 59,274, 99,298, and 214,735 tons in 1990, 2000 and 2010, respectively (FAO 2012). Although

the main part of production is captured, a significant tenfold increase has also occurred in aquaculture production through the last decade from 7,664 tons in 1990 to 74,356 tons in 2010 (FAO 2012).

### 1.2 *Epinephelus aeneus* compatibility for aquaculture

White grouper, *Epinephelus aeneus* (Saint-Hilaire, 1817) of the family Serranidae, sub-family Epinephelinae, is a species found throughout the southern Mediterranean and along the west coast of Africa to southern Angola (Heemstra and Randall, 1993) although, in the last years it has been seen also along the Adriatic sea (Glamuzina et al., 2000). Adults are usually located on rocky or mud and sand bottoms in depths of 20 to 200 meters and juveniles have been found in coastal lagoons and estuaries. The white grouper can reach the length of 120 cm and the weight of 25kg (Heemstra and Randall, 1993). The species is protogynous hermaphrodite meaning it matures first as female and later changes sex to male (Fennessy and sadovy, 2002). This kind of reproduction behavior has been described for other members of the *Epinephelus* genus such as *E. microdon* (Debas et al., 1990), *E. andersoni* (Fennessy and sadovy, 2002), *E. tauvina* (Kailasam et al., 2007) and *E. merra* (Bhandari et al., 2003).

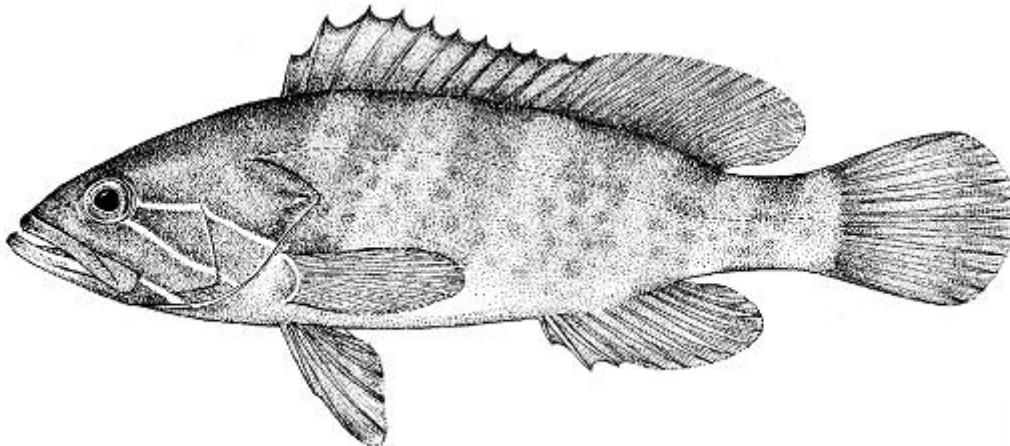


Figure 1. *Epinephelus aeneus* morphology (Heemstra and Randall, 1993)

Like many other groupers the white grouper is a promising candidate for intensive aquaculture due to its high market value, excellent taste, rapid growth rate, hardiness and disease resistance (Hassin et al., 1997; Glamuzina et al., 2000).



Although there are groupers that have been successfully bred in captivity artificial breeding often results in poor survival rate of larvae which is still a major obstacle in this industry that is widely under research (Lim, 1993; Kohno et al., 1997, Russo et al., 2009). Therefore many of the *Epinephelus* species cultured are based on wild caught fingerlings that are being stocked in floating net cages in coastal waters or in earthen ponds (Juario et al., 1993; Gorshkov, 2010).

### 1.3 *Epinephelus aeneus* culture in Israel

The first attempt to culture the *E. aeneus* in captivity has been made in Israel during the late 1990's in order to assess the species potential. The research was concentrated mainly on the potential growth rate in captivity and induced reproduction techniques (Hassin et al., 1997). However, the domestication of the white grouper is still in its infancy. Therefore, farming of the white grouper is still mostly based on wild populations captured from the natural environment and as yet there are no domesticated strains resulting from breeding activity. A white grouper domestication project was launched by the department of fisheries and aquaculture of the Israeli ministry of agriculture in 2010. The mission of the project is to develop innovative technologies aimed at domestication, breeding, and industrial rearing of the white grouper (Gorshkov, 2010).

The white grouper domestication process, similar to that of other fish species domestication procedures, is facing a few major genetic obstacles:

- Accurate identification of species - there are difficulties in relying primarily on morphology when attempting to identify fishes during various stages of their development (Ward et al., 2009). Misidentification of the species might lead to reproduction failure due to incompatibility among broodstock species or to unintentionally creation of hybrids. Therefore, in order to successfully culture and eventually commercialize the marketing of the white grouper, more accurate genetic identification method for cultured population is required. Furthermore, the possibility that sub-species occur within broodstock have to be rejected.

- Genetic variance reduction - domestication may lead to genetic deterioration of stocks due to genetic drift, inbreeding, and selection if poor broodstock management is applied (Na-Nakorn and Moeikum, 2009). Populations bred in captivity are subject to evolutionary changes occurring as direct and indirect responses to artificial selection, and as a response to natural selection in the domestic environment (Falconer, 1989). Therefore, another important objective of the domestication program is monitoring the potential reduction in genetic variation as a result of inbreeding (Trippel et al., 2009) within the white grouper cultured stock over generations.
- Understanding the reproduction strategy of *E. aeneus* – currently breeding of this species in captivity is made by culturing a group of several males and females together in a tank and collecting their fertilized eggs. However, the number of parents that contribute to a group of offspring has not been explored sufficiently both in wild and captive populations. Therefore, it is unclear whether the white grouper reproduction mode is monogamous or polygamous, i.e., a variable number of parents of both sexes contribute genetic material to the offspring population. Understanding of this process will provide important information for the establishment of optimal populating conditions for broodstock breeding in captivity.

In order to overcome these obstacles a genetic research is needed side by side with the physiological and behavioral investigations. A prerequisite for efficient genetic management of the white grouper captive broodstock is the existence of genomic tools such as DNA-based identification system, DNA markers and a linkage map.

#### 1.4 DNA-based identification

The limitations inherent in morphology-based identification systems have led to new genomic approaches that exploit diversity among DNA sequences to identify organisms. These sequences can be viewed as genetic 'barcodes' that are embedded in every cell (Hebert et al., 2003b). Among the frequently used DNA barcoding systems are methods based on sequence diversity in mitochondrial Cytochrome Oxidase C subunit 1 gene (COI) which appears to be among the most conservative protein-coding genes in the mitochondrial genome of animals (Folmer et al., 1994). Therefore the sequence varies only to a very minor degree within species, such that this variation is much less than between species (Savolainen et al., 2005). A region of about 650-bp of COI gene is the primary barcode sequence for members of the animal kingdom (Ratnasingham and Hebert, 2007) and in May 2004 the Consortium for the Barcode of Life (CBOL) was launched in an attempt to create a barcode library for all eukaryotic life using the COI sequence. As a result, the Barcode of Life Data System (BOLD) (<http://www.barcodinglife.org>) was developed. The COI sequences that are published in the BOLD system may be used for research and can be compared by phylogenetic analysis. COI gene was shown as a reliable and universal tool for the identification of species and subspecies among all animals. In aquaculture and fisheries research the sequence has various uses such as identification and revealing of mislabeled commercially marketed species like tuna (Lowenstein et al., 2009) and catfish (Carvalho et al., 2011). Other projects are taxonomic analysis of close species for instance tilapia (Shirak et al., 2009) and *Starksia atlantica* (Baldwin et al., 2011). The COI sequence was also used for analyzing the diversity of fish species inhabiting areas or environments like Antarctic regions (Dettai et al., 2011) or the North American freshwaters (April et al., 2011).

In light of this, genetic analysis of the white grouper population using COI sequence would be most valuable for preventing misidentification of broodstock fishes, controlling the possibility of sub-species within the research population and for phylogenetic studies.

### 1.5 Genetic markers

Molecular markers, revealing polymorphisms at the DNA level, are now key players in animal genetics and have revolutionized the way aquaculture genetics research is conducted (Vignal et al., 2002; Liu and Cordes, 2004). DNA markers have allowed rapid progress in investigations of genetic variability and inbreeding, parentage assignments, species and strain identification, and the construction of high-resolution genetic linkage maps for aquaculture species (Liu and Cordes, 2004). Monitoring the variation among the white grouper captive populations using genetic markers will greatly assist the species domestication process. Various genetic markers are available for this kind of research including RAPDs (random amplified polymorphic DNA), AFLPs (amplified fragment length polymorphism), RFLPs (restriction fragment length polymorphism), microsatellites and SNPs (single nucleotide polymorphisms) (Liu and Cordes, 2004). Each marker type is suitable for several kinds of studies. Today, the markers most commonly used for monitoring genetic variation among populations and for development of linkage maps are microsatellites and the latest developed system of SNP markers.

### 1.6 Microsatellite markers

Microsatellites are short DNA sequence stretches in which a motif of one to six bases is tandemly repeated; this most likely happens due to slippage, which would occur during replication or DNA repair. The sequences can differ in number of repeats among individuals. Therefore, by using the polymerase chain reaction (PCR) this property of microsatellite DNA was converted into highly versatile genetic marker (Tautz, 1989; Schlötterer, 2000) because the changes in number of repeat units can result in a large number of alleles at each microsatellite locus in a population. Thus, microsatellites are considered to have the highest Polymorphic Information Content (PIC) of any DNA marker (Liu and Cordes, 2004).

Microsatellites have been estimated to reside every 10 kb and tend to be evenly distributed in the genome on all chromosomes and all regions of the chromosome. They have been found inside gene-coding regions, introns, and in the non-gene

sequences (Liu and Cordes, 2004). Due to their high level of polymorphism, high genomic abundance and co-dominant inheritance, microsatellite markers are an excellent tool for assessing genetic information (O'Connell and Wright, 1997; Schlötterer, 2000; Liu and Cordes, 2004; Chen et al., 2007). In aquaculture, microsatellite markers have been used for population genetic studies of many species e.g., the grass carp (Zheng et al., 2007; Liu et al., 2009), Assian and European sea bass (Garcia et al., 1997; Yue et al., 2009), gilthead sea bream (Loukovitis et al., 2012) and Atlantic salmon (Norris et al., 1999). Microsatellite markers are also used for parentage and kinship analysis e.g., rainbow trout (Herbinger et al., 1995), red drum (Saillant et al., 2007), Atlantic salmon (Norris et al., 2000; Villanueva et al., 2002), common carp (Vandeputte et al., 2004) and sea bass (Wang et al., 2008; Frost et al., 2006). Additional application of microsatellite markers is genome mapping e.g., zebra fish (Shimoda et al., 1999) tilapia (Kocher et al., 1998; Lee et al., 2005), Turbot (Bouza et al., 2007), rainbow trout (Guyomard et al., 2006; Sakamoto et al., 2000), Atlantic salmon (Gilbey et al., 2004), Assian and European sea bass (Chistiakov et al., 2005; Wang et al., 2011), catfish (Kucuktas et al., 2009) and gilthead seabream (Franch et al., 2006). However, use of microsatellite markers involves a large amount of up-front investment and effort. Each microsatellite locus has to be identified and its flanking region sequenced for the design of PCR primers (Liu and Cordes, 2004).

Today, thanks to advanced new sequencing equipment, technologies and methods, the development of large number of microsatellite markers for a species of interest is available at specialized labs. However, it is still a costly procedure. An alternative method for developing microsatellite markers is by cross-species amplification that results in heterologous markers i.e., sequence information in one species is used to amplify the microsatellite in other related species. When the current study was launched, no sequence information for development of specific microsatellite primers was available for *E. aeneus*. However, this information was available for other *Epinephelus* species and was therefore used for development of heterologous markers for *E. aeneus*.

## 1.7 SNP markers

Single nucleotide polymorphism (SNP) marker is a single base change in a DNA sequence, with usually two alternative possible nucleotides (alleles) at a given position. A polymorphic SNP marker should have the least frequent allele (variant) at a frequency of 1% or greater. Although, in principle, any of the four possible nucleotide bases can be present at each position of a sequence stretch, in practice SNPs are usually bi-allelic (Vignal et al., 2002).

SNPs are the result of point mutations. There are two mechanisms for the creation of a SNP: transitions, exchanges of purine-purine (A<>G) or pyrimidine-pyrimidine (C<>T), or transversions which are exchanges of purine-pyrimidine or pyrimidine-purine (A<>C, A<>T, G<>C or G<>T) (Freese, 1959). Although the expected ratio of transition:transversion should be 1:2, observations show a favor of transitions over transversions (Collins and Jukes, 1994). SNP markers have been used for paternity testing, genetic fingerprinting, calculation of genetic distances, genome mapping and identification of candidate genes for quantitative trait loci (QTL) (Liu and Cordes, 2004; Rengmark et al., 2006). SNP markers have low polymorphism (2-4 alleles) and their relative advantage is achieved by their high abundance across the genome. Therefore, a relatively large number of SNP markers are needed for projects such as paternity tests (Rengmark et al., 2006) and genetic mapping (Bradley et al., 2007). DNA sequencing has been the most accurate and most used approach for SNP discovery. Random shotgun sequencing, amplicon sequencing using PCR, and comparative EST analysis are among the most popular sequencing methods (Vignal et al., 2002). Despite the progress in development of efficient equipment such as genetic chips and advanced sequencing methods, the discovery and efficient genotyping of SNP markers for *E. aeneus* genomic project will require large economic effort. In light of this, although it is predicted that SNPs will be the future marker of choice in biotechnology-related industries due to their abundance and adaptability to automation (Liu and Cordes, 2004), at this stage it was decided to discover microsatellite markers for the project which are more cost effective and efficient for the current research objectives.

## 1.8 Genetic linkage map

A genetic linkage map is the positioning of genes and markers onto linkage groups and chromosomes using linkage analysis. It is based on the assumption that the greater the distance between two linked genes or markers, the greater the chance that non-sister chromatids would cross over in the region between them and, hence, the greater the proportion of recombinants that would occur. Therefore, by determining the frequency of recombinants, a measure of the map distance between the genes and markers can be obtained (Griffiths et al., 2000). A linkage map does not show the physical distance between genes and markers but their relative recombination distance presented in centiMorgan (cM) units in honor of Thomas Hunt Morgan. Families are needed as mapping populations in order to make a segregation analysis of alleles across generations.

Linkage maps were constructed using multi-site genetic markers such as RFLPs (Bostein et al., 1980), AFLPs (Kocher et al., 1998) and RAPDs (Debener and Mattiesch, 1999) and more recently using single-site markers, e.g. microsatellites and SNPs (Love et al., 1990; Moen et al., 2008, 2009; Sanetra et al., 2009; Zheng et al., 2011).

The power of genetic maps increases with their density (Dietrich et al., 1994). Larger number of markers increases the map accuracy (Gyapay et al., 1994; Dib et al., 1996) and fine linkage map is an essential infrastructure for detection of QTL linked to genes affecting important traits and the location of candidate genes.

### 1.9 Research objectives:

- ❖ Development of genetic markers for *E. aeneus* and preliminary testing of their potential for construction of *E. aeneus* linkage map in the future.
- ❖ Analysis of phylogenetic distances of *Epinephelus* species and the efficiency of construction of heterologous polymorphic markers.
- ❖ Characterization of the Israeli National Center for Mariculture (NCM) captive broodstock compared to wild samples.
- ❖ Parentage analysis of *E. aeneus* in captivity.



## 2. Materials and methods

### 2.1 Abbreviations used and details of solutions

RNase – Ribonuclease

DTT – Dithiothreitol

dNTP – deoxy Nucleotide Tri Phosphate

EDTA – Ethylene Diamine Tetra Acetate

Tris – tris (hydroxymethyl) aminomethane

TAE – Tris acetate EDTA buffer

SDS - sodium dodecyl sulfate

DNA extraction Buffer (50ml) – 0.5 ml Tris pH 8.0 (1M)

1 ml EDTA pH 8.0 (0.5M)

5 ml SDS 1%

1 ml 100 mM NaCl

42.5 ml water

### 2.2 Research population

Research was conducted using several *E. aeneus* populations:

1. Broodstock population - was F<sub>1</sub> of wild-caught *E. aeneus* collected from the Mediterranean shore of Israel and cultured at the National Center for Mariculture (NCM). The broodstock population consisted of 12 mature *E. aeneus* individuals that were grown in captivity. Gender of the individuals was determined according to type of gametes (milt or eggs) obtained by gonadal biopsy (Meiri-Ashkenazi et al., 2010). Two thirds of the individuals were clearly determined, so the broodstock comprised of five males, three females and four individuals of unknown gender.

Table 1. Broodstock population according to gender

#	PIT-tag*	Gender
1	2898	Male
2	1B74	Male
3	2492	Male
4	1C16	Male
5	F67D	Male
6	2D75	Unclear gender
7	2D44	Unclear gender
8	5B01	Female
9	DE8C	Female
10	B7E3	Female
11	BFF8	Unclear gender
12	1313	Unclear gender

\* PIT tag - passive integrated transponder tag

2. Wild Type population - 12 Individuals representing the wild type population originated from various locations along the Mediterranean coast of Israel.
3. Progeny population - Included three batches (9 to 22 larvae each) collected on different days of spawning.

Table 2. Progeny population

Group	Spawning Tank	Hatchery Tank	Collecting date	Larvae number
A	D1	B10	14/8/10	21
B	D1	B7	19/8/10	22
C	D1	B9	6-8/8/10	9

4. Fingerling population – Included 256 fingerlings in the age of 46 days collected on 18/5/11 at the NCM to form a mapping population for linkage map construction.

### 2.3 DNA extraction

DNA was extracted from the following sources:

- Broodstock blood samples - received in Eppendorf tubes from the National Center for Mariculture (NCM) in Eilat.
- Wild type *E. aeneus* – received as fin-clip samples.
- Whole larvae samples – received in ethanol tubes from the NCM.
- Fingerling tissue samples- received in ethanol tubes from the NCM.

#### 2.3.1 Genomic DNA extraction from blood samples

600 µl of Red Cell Lysis Solution (Epicentre<sup>®</sup> Biotechnologies, WI, USA) was mixed with 200 µl of blood sample in a 2 ml tube by inverting the tube and flicking its bottom. Samples were incubated at room temperature for five minutes followed by a brief vortex and a second incubation followed by another brief vortexing. White blood cells were pelleted by centrifugation for 25 seconds in a microcentrifuge. Supernatant was removed leaving 25 µl of liquid. Mix was then vortexed to suspend the pellet. The white blood cells were resuspended in 300 µl of Tissue and Cell Lysis Solution (Epicentre<sup>®</sup> Biotechnologies, WI, USA) by pipetting the cells up and down several times. 1 µl of 5 µg/µl RNase A was added and mixed by vortex. Samples were incubated at 37°C for 30 minutes and later placed on ice for 3-5 minutes. 90 µl of MPC Protein Precipitation Reagent (Epicentre<sup>®</sup> Biotechnologies, WI, USA) were added to the lysed samples and the mix was vortexed vigorously for 10 seconds. Debris was pelleted by centrifugation for 10 minutes at 13,000 rpm in a microcentrifuge, supernatant was transferred to a clean tube and 500 µl of Isopropanol was added. Tubes were inverted 40 times for mixing purposes and DNA was pelleted by centrifugation for 10 minutes at 14000 rpm. Isopropanol was poured off without dislodging the DNA pellet followed by rinsing with 1 ml of 100% Ethanol.

DNA was pelleted by centrifugation for 4 minutes at 14000 rpm. Ethanol was poured off without dislodging the DNA pellet followed by second rinsing with 1 ml of 70% Ethanol that was later poured off. DNA was resuspended in 300 µl of Ultra pure H<sub>2</sub>O (Biological industries, Beit Haemek, Israel) and incubated at 37°C for 30 minutes.

### 2.3.2 Genomic DNA extraction from wild type and larvae tissue samples

Extracting procedure was identical for wild type fin clips samples and whole larvae samples except for smaller volumes of some materials that were used in the larvae procedure due to the smaller tissue samples.

Each sample was treated with 300µl (100 µl for larvae) of Tissue and Cell Lysis Solution (Epicentre<sup>®</sup> Biotechnologies, WI, USA) and 1 µl of Proteinase K (Epicentre<sup>®</sup> Biotechnologies, WI, USA) in a 2 ml tube. Samples were incubated at 65°C for 15 minutes and vortexed every 5 minutes. Samples were later cooled to 37°C and added with 1µl (0.5 µl for larvae) of 5 mg/ml RNase A (Epicentre<sup>®</sup> Biotechnologies, WI, USA). The following precipitation process of the DNA was as previously described for genomic DNA extraction from blood samples (section 2.3.1)

### 2.3.3 Genomic DNA extraction from fingerling tissue samples

Due to the large number of samples extraction of DNA was made following the protocol of Zilberman et al. (2006) using 2ml plate (USA Scientific, FL, USA) containing 96 wells (Riplate, PP). Tissue samples of about one cm<sup>2</sup> along with 500 µl Extraction Buffer, 5 µl DTT (1M) and 3 µl Proteinase K (20 mg/ml) were incubated for 12 hours at a temperature of 45°C. Debris was pellet by centrifugation for 10 minutes at 4000 rpm, supernatant was transferred to a new plate. 200 µl of NaCl (5 M) was added to each sample. The plates were shaken for 30 minutes in an incubator at 37°C and centrifuged for 10 minutes at 4000 rpm. The supernatants were transferred to a new plate containing 1 ml of 100% Ethanol in each well and cooled for 1 hour at -4°C. Samples were then centrifuged for another 10 minutes at 4000 rpm. Ethanol was poured off without dislodging the DNA pellet followed by second rinsing with 1ml of 70% Ethanol that was later poured off. DNA was

resuspended in 100 µl of Ultra pure H<sub>2</sub>O (Biological industries, Beit haemek, Israel) and saved in a freezer at -4°C.

## 2.4 Concentration determination

DNA concentration was determined using NanoDrop 1000 Spectrophotometer (NanoDrop products, DE, USA) and diluted to 10-40 ng/µl. Due to their size larval samples contained less DNA material. Therefore, less DNA was used for larval PCR reactions in order to have sufficient material to complete the project.

## 2.5 COI analysis

In order to verify samples' species of origin, exclude interspecies segmentation among *E. aeneus* population and development of phylogenetic trees, the reference fragment of COI gene was amplified and sequenced from four wild type and four broodstock samples using a cocktail of primers that is proved to be efficient for fishes from Ivanova et al. (2007) protocol. Cocktail name is COI-1 which amplified a read length of 631 bp and contains the following primers:

Table 3. COI-3 primer cocktail description

Name	Ratio	Primer sequence 5 '-3 '	primer position
VF2_t1	1	TGTA AACGACGGCCAGTCAACCAACCACAAAGACATTGGCAC	6448–6474
FishF2_t1	1	TGTA AACGACGGCCAGTCGACTAATCATAAAGATATCGGCAC	6448–6474
FishR2_t1	1	CAGGAAACAGCTATGACACTTCAGGGTGACCGAAGAATCAGAA	7152–7127
FR1d_t1	1	CAGGAAACAGCTATGACACCTCAGGGTGTCCGAARAAYCARAA	7152–7127

PCR reaction contained: 1-30 ng of DNA, mixture of 2.5 mM dNTPs of each nucleotide (JMR Holding, London, UK), 20 pmol of primer COI-3 cocktail, 1.2 units of Bio BIO-X-ACT Long DNA polymerase (Bioline, London, UK), 1 µl of 10 x OptiBuffer

(Bioline, London, UK), 0.7 µl of 50mM MgCl solution and Ultra pure H<sub>2</sub>O (Biological industries, Beit Haemek, Israel) for volume completion to 10 µl.

The same procedure was also performed on tissue samples of *E. itajara* and *E. gutattus* species for which no COI sequence was available in the BOLD system. Samples of these two species were obtained from Drs. Jose Bautista and Elena Gonzalez, (Universidad Complutense de Madrid, Spain) and from Dr. Rachel Graham (Wildlife Conservation Society, Gulf and Caribbean Sharks & Rays Program, Belize).

### 2.6 PCR amplification of COI fragments

Initial denaturation at 94°C for 3 min

30 cycles of:

- Denaturation at 92°C for 40 sec
- Annealing at 53°C for 40 sec
- Elongation at 72°C for 1 min

Final extension step at 72°C for 10 min.

### 2.7 Fragments separation

PCR products along with 2 µl Gel loading Blue buffer (New England Biolabs, MA, USA) were separated by electrophoresis through Agarose gel (Amresco LLC, OH, USA) in concentration of 15 mg/ml containing Modified TAE Buffer 1:50 (Merck Millipore, MA, USA). The gel contained Ethidium bromide (Amresco LLC, OH, USA) in concentration of 0.5 µg/ml. Electrophoresis was made in a voltage of 100 V for 40 minutes with a 100 bp ladder (New England Biolabs, MA, USA) was used for length detection.

### 2.8 Fragment sequencing

Successful PCR products were cut and extracted out of the gel using Montage Gel Extraction Kit (Merck Millipore, MA, USA). Centrifugation time was 10 minutes at 5000 rpm. The extracted products were sent for sequencing at the DNA sequencing unit, Weizmann Institute of Science, Rehovot, Israel.

## 2.9 Phylogeny of *Epinephelus* species

COI sequences of *E. aeneus* were analyzed along with other available *Epinephelus* COI sequences from GenBank including the seven species from which the heterologous markers were developed: *E. awoara*, *E. septemfasciatus*, *E. itajara*, *E. akaara*, *E. fuscoguttatus*, *E. guttatus*, and *E. lanceolatus*. A consensus COI sequence was determined for each species using GAP4 software (<http://staden.sourceforge.net>). The consensus sequences were used for construction of DNA- and protein-based phylogenetic trees by CLUSTALW software (<http://www.genome.jp/tools/clustalw/>) using the neighbor joining method and bootstrap analysis with 500 replicates.

## 2.10 Development of heterologous markers

At the beginning NCBI database was BLAST searched for DNA sequences of fishes belonging to the genus *Epinephelus*. The search was limited for sequences containing a minimum of (CA)<sub>10</sub>. In order to search sequences of nuclear origin the following types of sequences were excluded: Virus, Bacteria, Ribosomal and Mitochondrial. Sequences were scanned, in FASTA format, against Tilapia Repeat Masker software (Shirak et al., 2010) in order to remove putative sequence repeats.

Primers were designed simultaneously by muPlex software (<http://apps.diatomsoftware.com/muplex/html/MuPlex.html>) (Rachlin et al., 2005a, 2005b) according to the Tilapia Repeat Masker results. Only species with at least six sequences suitable for primers production were retained for developing a set of heterologous microsatellite markers.

In order to investigate the correlation between marker success and geographical relationship, as shown in Table 6, information about geographical distribution of relevant *Epinephelus* species was gathered from FishBase system (<http://www.fishbase.org>).

### 2.11 Two-step PCR procedure

A two-step PCR procedure (Wang and Godfrey, 1998) was used for amplification of fragments based on chosen primer pairs. In this procedure, the forward primers were synthesized with generic oligonucleotide linkers: CCTAGGTGATCAAGATCTCGC or CCTCGAGGTGGACGGTATCGG instead of a fluorescence label. Following the first PCR, one microliter of diluted amplified material is used as the template for the second step PCR. The second step PCR is set up as the first step. However, the forward primer used is no longer specific, it is fluorescence labeled and its sequence is identical to one of the linkers used in the first PCR. The same reverse primer is used in both amplifications. The procedure is presented schematically in Figure 2. The two fluorescent primers that were used for the research were CCTAGGTGATCAAGATCTCGC using the VIC dye and CCTCGAGGTGGACGGTATCGG using the FAM dye (Applied Biosystems, CA, USA).

### 2.12 PCR reaction

First and second stage PCR reactions were performed in a 96 wells plate. Each reaction contained a total volume of 10  $\mu$ l. materials used for the first PCR reaction:

- A mixture of 2.5 mM dNTPs of each nucleotide (JMR Holding, London, UK)
- 1.5 units of Super-Therm Taq DNA polymerase (JMR Holding, London, UK)
- 1  $\mu$ l of complementary reaction Buffer (10x) with 15 mM  $MgCl_2$  (JMR Holding, London, UK)
- 5 pmol of each primer (IDT, Jerusalem, Israel)
- 1-30 ng of genomic DNA
- Ultra pure  $H_2O$  (Biological industries, Beit Haemek, Israel) for volume completion to 10  $\mu$ l



First PCR reaction products were diluted in a 1:10 ratio and 1  $\mu$ l was used as template for the second PCR reaction. All other materials used for the second PCR reaction were similar to the first reaction with an exception of 5 pmol of FAM or VIC (Applied Biosystems, CA, USA) fluorescently labeled primers that were used instead of the original forward primers.

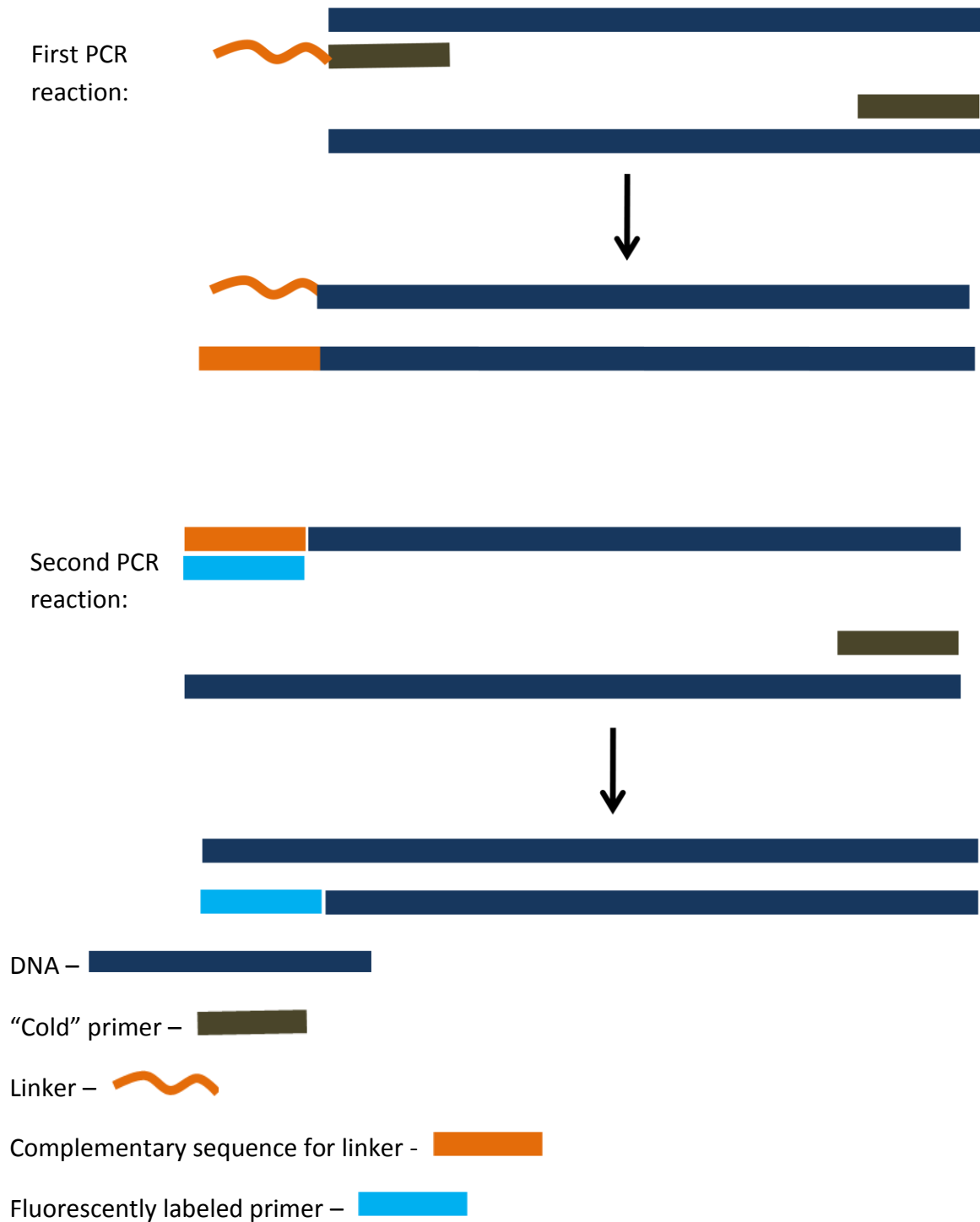


Figure 2. A scheme of two-step PCR procedure

### 2.12.1 PCR amplification program

Initial denaturation at 94°C for 3 min

30 cycles of:

- Denaturation at 93°C for 40 sec
- Annealing at 55°C for first PCR reaction and 57°C for second PCR reaction for 40 sec
- Elongation at 72°C for 1 min

Final extension step at 72°C for 10 min

### 2.13 PCR products detection

The second PCR products were diluted in a ratio of 1:100 and 1 µl of each diluted sample was transferred to a different well of a 96 well optical reaction plate (Applied Biosystems, CA, USA) to be detected by the ABI3130 Genetic Analyzer (Applied Biosystems, CA, USA). First and second PCR amplifications were processed separately for each primers pair. However, in order to make the working process more efficient, two pairs of primers of different dyes were diluted, pooled and loaded together on the ABI3130 Genetic Analyzer. Each well contained a total volume of 16.05 µl which included the PCR product, 15 µl of Hi-Di™ Formamide and 0.05 µl of GeneScan™ 500 LIZ™ Size standard (Applied Biosystems, CA, USA).

### 2.14 Genotyping procedure

Allelic sizes were evaluated using GeneMapper software v.4.0 (Applied Biosystems, CA, USA). All broodstock and wild type populations' samples were genotyped for 233 microsatellite markers. Yield of polymorphic markers was calculated separately for each species as the number of markers with >1 alleles in percent of the total number of markers analyzed.

The offspring population was genotyped for eight ARO markers shown in Table 4 for parentage analysis. These markers were selected based on their level of polymorphism, accurate allele detection and lack of null alleles.

Table 4. Primer pairs used for parental analysis

Marker	Primer sequences <sup>1</sup>	bp <sup>2</sup>
ARO1003	F: <u>CCTAGGTGATCAAGATCTCG</u> CGTGCAAGGCAAGCTGTGTTA R:AGCAGGCATCTTGTTATCTGG	206-238
ARO1045	F: <u>CCTAGGTGATCAAGATCTCG</u> CCACGAAGTATTTGGCTGAT R:GAGAAAGTGGCAATATTTGAC	218-236
ARO1083	F: <u>CCTCGAGGTGGACGGTATCGG</u> CCGGTTCTTCTCTCTCCC R:TTACTGTTGATTGAGTTGTTGT	228-258
ARO1084	F: <u>CCTCGAGGTGGACGGTATCGG</u> GGGTTTATTTCAAAGGTCAG R:CCCAATGAGGTGTTCAATAT	202-218
ARO1105	F: <u>CCTAGGTGATCAAGATCTCG</u> CTGATAGCTTTACATGCACTCA R:CTGAACCTCACCTGAAA	178-200
ARO1120	F: <u>CCTAGGTGATCAAGATCTCG</u> CCTCTGATGCTGTTTACACAAC R: TCTCCATCGAAGGTAAAGG	189-209
ARO1131	F: <u>CCTAGGTGATCAAGATCTCG</u> CTGTGTGTCAGAGGTGGGTT R: TGAATTTCACTGCATGTTTC	174-206
ARO1137	F: <u>CCTAGGTGATCAAGATCTCG</u> CATGGGTATAATTAGGACACACT R: AGGAAAGGAGGGAGGAAA	210-230

<sup>1</sup> Linker sequence is underlined

<sup>2</sup> PCR product size range

### 2.15 Parentage analysis

Attribution of parents to progeny was carried out using PAPA 2.0 software (Duchesne et al., 2002). The analysis was based on genotypes of 12 putative parents for the selected markers. Genotypes of each larva were analyzed for every marker finding putative parent pairs. Finally the pair that was positively assigned for all eight markers was considered as the larva's original pair of parents.

### 2.16 Statistical analysis

Differences in yield of polymorphic heterologous markers by species of origin were statistically analyzed using a Chi-squared test ( $\chi^2$ ). Distributions of number of failures and success to amplify heterologous polymorphic markers by origin of species were compared. Distributions not statistically different from one another were pooled and compared to that of the species that were statistically different from the others. Number of alleles per marker was compared for all polymorphic markers between the wild and captive populations using paired t-test.

### 2.17 Polymorphic Information Content (PIC)

PIC is the value of a marker for detecting polymorphism within a population and depends on the number of detectable alleles and the distribution of their frequencies (Liu and Cordes, 2004).

PIC was calculated as follows:

$$PIC = 1 - \sum_{i=1}^n p_i^2$$

where:

PIC - Polymorphic Information Content of the marker.

Pi – frequency of the allele i

### 3. Results

#### 3.1 Development of microsatellite markers

A total of 2,100 *Epinephelus* sequences were detected in NCBI database from which 1,600 were nuclear sequences. These sequences were scanned for a minimum of ten CA repeats and a total of 656 sequences originating from 20 different *Epinephelus* species were located as shown in Table 5. Seven *Epinephelus* species (*E. awoara*, *E. septemfasciatus*, *E. fuscoguttatus*, *E. itajara*, *E. guttatus*, *E. akaara* and *E. lanceolatus*) were found to have at least six nuclear sequences and were suitable for developing a set of heterologous microsatellite markers. A total of 233 primer pairs were developed.

The seven species from which microsatellite markers were developed varied in their geographical distribution as presented in Table 6. Three species: *E. akaara*, *E. awoara* and *E. septemfasciatus* are located in the Pacific Northwest area, two species: *E. fuscoguttatus* and *E. lanceolatus* are Indo-Pacific and the remaining species: *E. guttatus* and *E. itajara* are Atlantic species.

Table 5. *Epinephelus* species used for primers development

Species	Total number of sequences	Sequences suitable for primers development
<i>E. acanthistius</i>	8	
<i>E. adscensionis</i>	1	
<i>E. aeneus</i>	1	
<i>E. akaara</i>	49	11
<i>E. awoara</i>	81	52
<i>E. bleekeri</i>	1	
<i>E. bruneus</i>	2	
<i>E. coioides</i>	162	9
<i>E. fuscoguttatus</i>	145	50
<i>E. guttatus</i>	42	32
<i>E. itajara</i>	57	20
<i>E. lanceolatus</i>	15	6
<i>E. maculatus</i>	1	
<i>E. malabaricus</i>	3	
<i>E. merra</i>	1	
<i>E. morio</i>	2	
<i>E. polyphkadion</i>	1	
<i>E. quernus</i>	5	
<i>E. septemfasciatus</i>	78	53
<i>E. tauvina</i>	1	

Table 6. Geographical distribution of selected *Epinephelus* species (FishBase)

Specie	Common name	Geographic distribution
<i>E. akaara</i> (Temminck & Schlegel, 1842)	Hong Kong grouper	Northwest Pacific: southern China, Taiwan, East China Sea, Korea and southern Japan and Viet Nam.
<i>E. awoara</i> (Temminck & Schlegel, 1842)	Yellow grouper	Northwest Pacific: Korea, Japan, Taiwan, China, Viet Nam and islands in the South China Sea.
<i>E. coioides</i> (Hamilton, 1822)		Indo-West Pacific: Red Sea south to at least Durban, South Africa and eastward to Palau and Fiji, north to the Ryukyu Islands, south to the Arafura Sea and Australia. Recently reported from the Mediterranean coast of Israel.
<i>E. fuscoguttatus</i> (Forsskål, 1775)	Brown-marbled grouper	Indo-Pacific: Red Sea and East Africa to Samoa and the Phoenix Islands, north to Japan and south to Australia
<i>E. guttatus</i> (Linnaeus, 1758)	Red hind	Western Atlantic: North Carolina, USA to Paraíba, Brazil. The most common species of <i>Epinephelus</i> in the West Indies.
<i>E. itajara</i> (Lichtenstein, 1822)	Atlantic-goliath grouper	Western Atlantic: Florida, USA to southern Brazil, including the Gulf of Mexico

---

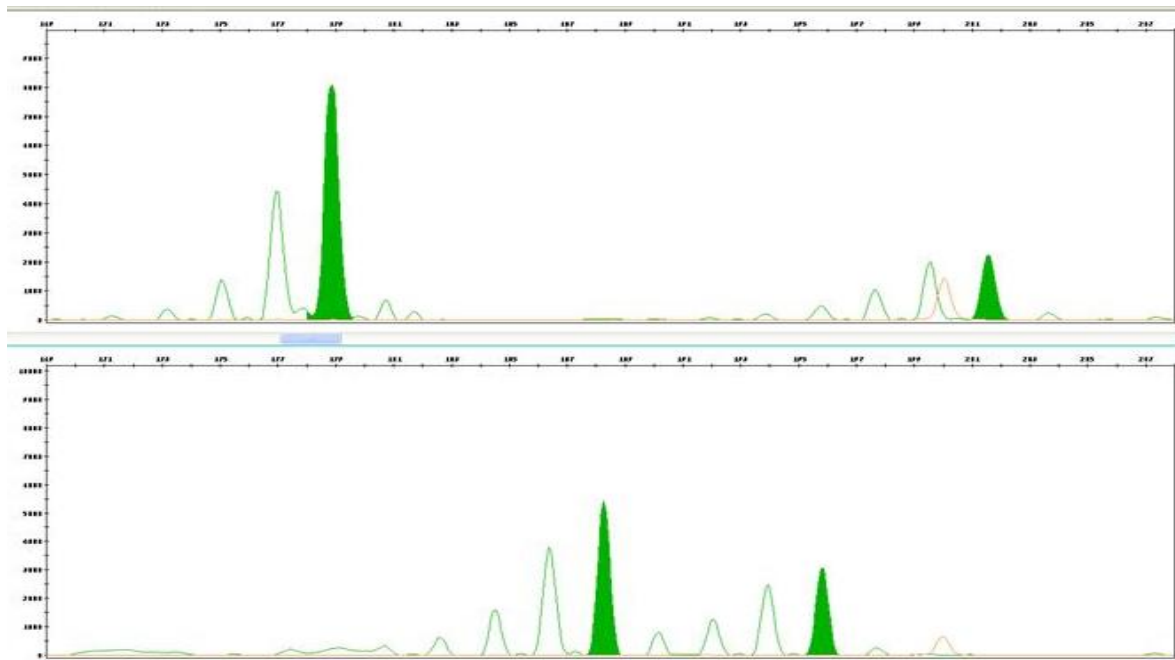
		and the Caribbean. Eastern Atlantic: Senegal to Congo, rare in Canary Islands. Discrete populations exist in the western Atlantic
<i>E. lanceolatus</i> (Bloch, 1790)	Giant grouper	Indo-Pacific: Red Sea to Algoa Bay, South Africa and eastward to the Hawaiian and Pitcairn islands, north to southern Japan, south to Australia.
<i>E. septemfasciatus</i> (Thunberg, 1793)	Convict grouper	Northwest Pacific: Japan, Korea, and China.

---

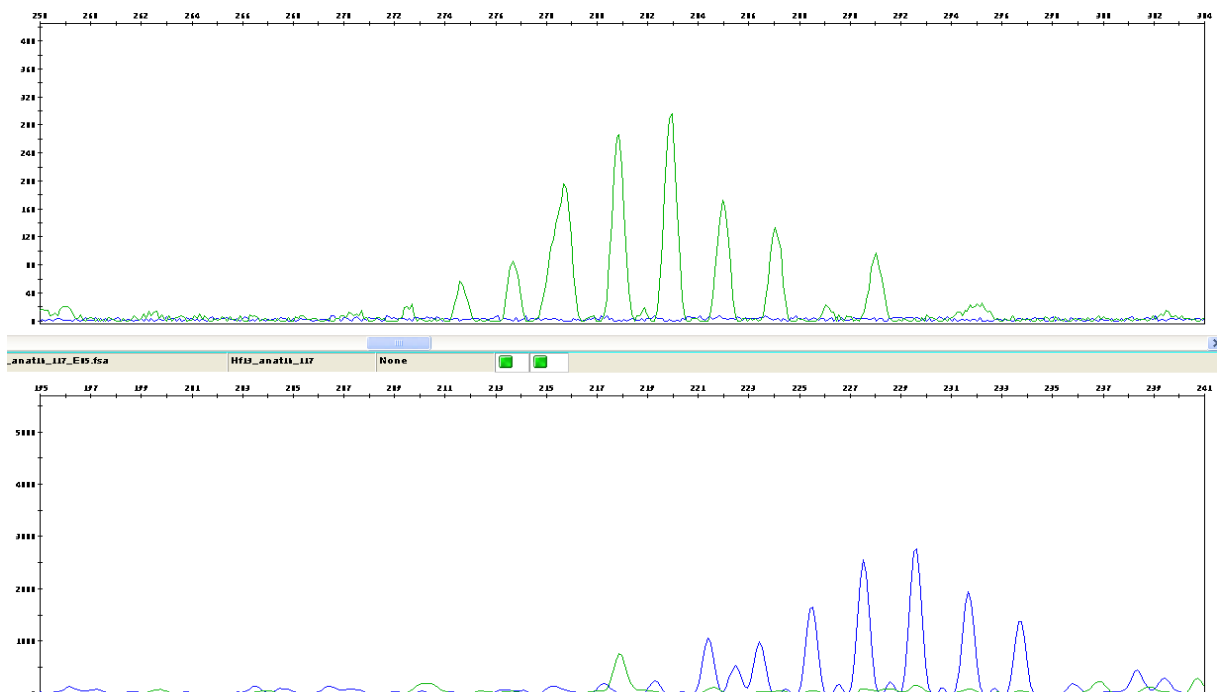


### 3.2 Genotyping of microsatellites

a.



b.



**Figure 3a.** Typical display of heterozygote genotypes of two different samples using GeneMapper software. **b.** Complex display of genotypes of two different samples using GeneMapper software.

PCR products that were analyzed showed variation in allele calling of genotypes due to: quality, allele sizes, number of alleles, peak heights and level of heterozygosity. Most of the results were clear for allele calling using the GeneMapper software and presented a unidirectional descending appearance as demonstrated in Figure 3a. However, some PCR products showed results that were less clear for allele calling as presented in Figure 3b. Nevertheless these products were also included in the analysis and every marker that included two alleles or more was considered as polymorphic.

### 3.3 Polymorphism of heterologous microsatellites

Heterologous microsatellite marker for which more than one allele was identified was considered polymorphic. The yield of polymorphic microsatellites in *E. aeneus* by species of origin ranged from 42% to 83% as shown in Table 7. Application of  $\chi^2$  test on the distributions of failures and successes of development of heterologous polymorphic markers by species of origin showed significance only between *E. septemfasciatus* and the other six species. Markers developed using sequences of *E. septemfasciatus* showed low success rate (42%) relative to that of other species.

Table 7. Polymorphism of heterologous markers in *Epinephelus aeneus* by species of origin

Species of origin	N	n	Polymorphism rate <sup>1</sup> (%)
<i>E. awoara</i>	52	32	62
<i>E. septemfasciatus</i>	53	22	42*
<i>E. fuscoguttatus</i>	50	33	66
<i>E. itajara</i>	20	16	80
<i>E. guttatus</i>	32	23	72
<i>E. akaara</i>	11	6	55
<i>E. lanceolatus</i>	6	5	83
<i>E. coioides</i>	9	7	78

n - Number of polymorphic markers; N - number of tested markers.

<sup>1</sup> Polymorphism rate =  $n/N$ ;

\* Significantly different by chi-squared test from the other six species ( $\chi^2 = 10.4$ ;  $P=0.001$ ).

### 3.4 Polymorphism of microsatellites in the wild and captive populations

COI sequences of all *E. aeneus* samples that have been tested were identical to the consensus sequence of *E. aeneus*. Out of the 233 primer pairs examined 147 markers were polymorphic in the wild type population, whereas 144 (62%) were also polymorphic in the broodstock population. The maximum number of alleles per marker found in the wild type population was 14 compared to seven in the captive population. The average number of alleles per marker was significantly different between wild and captive populations as shown in Table 8.

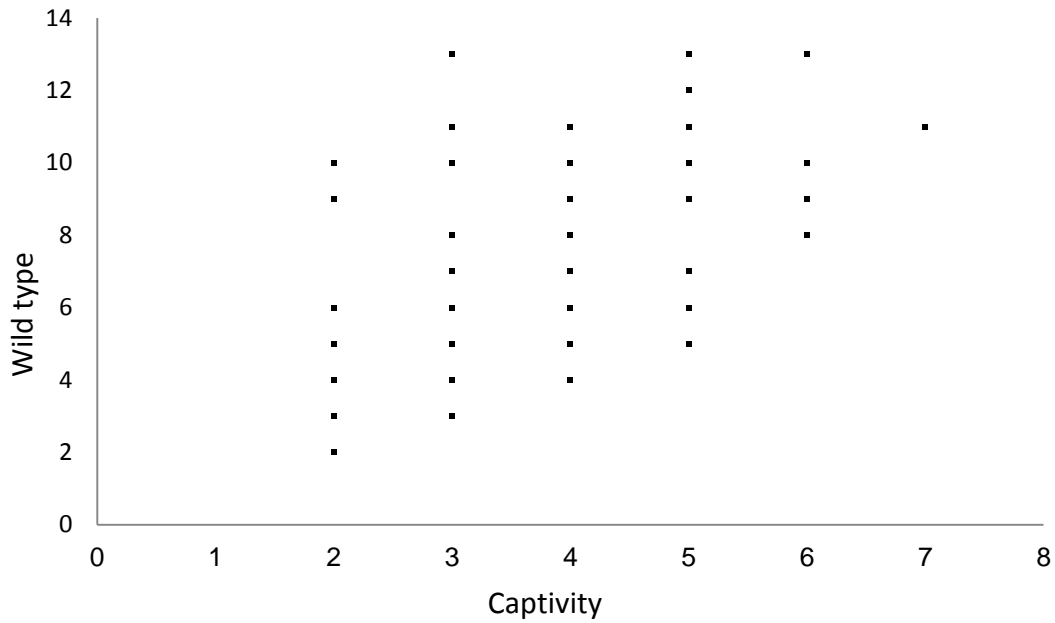
Table 8. Number of alleles per marker found in wild and captive populations for 144 polymorphic markers

Population	Range of alleles	Avg. no. of alleles <sup>1</sup>	PIC avg. <sup>2</sup>	PIC range
	No.	( $\pm$ SD) per marker		
Wild	2 – 14	6.23 $\pm$ 3.32	0.69 $\pm$ 0.2	0.08 - 0.94
Captive	2 – 7	3.7 $\pm$ 1.46	0.58 $\pm$ 0.18	0.08 - 0.83

<sup>1</sup>Significantly different (paired t test = 10.9;  $P = 1.85 \times 10^{-20}$ )

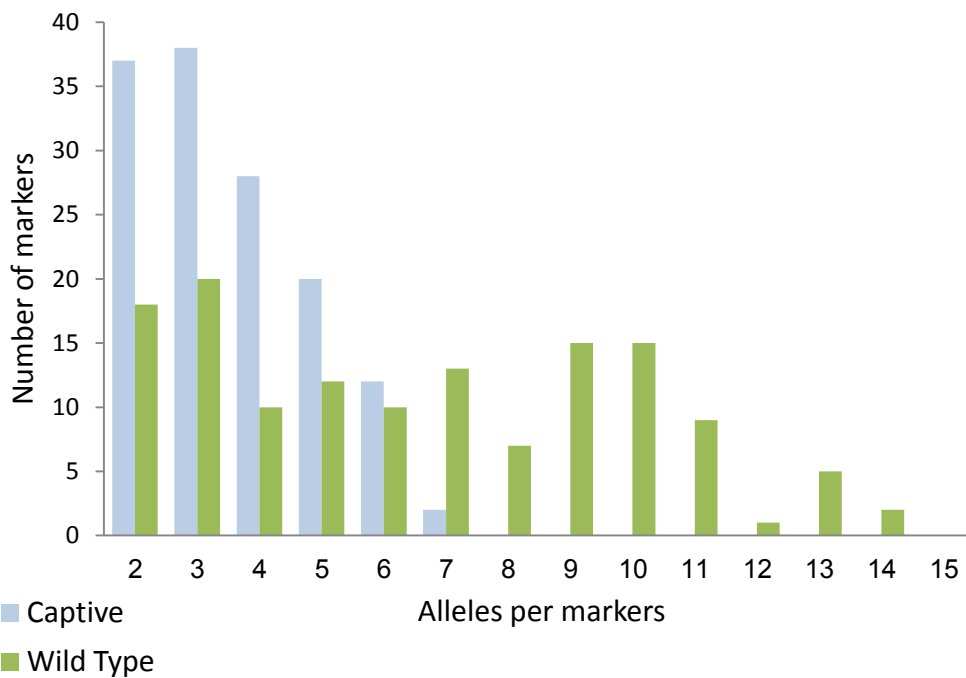
<sup>2</sup>Significantly different (paired t test = 6.8;  $P = 2.96 \times 10^{-10}$ )

Variable reduction in allele number was noted for different marker loci for the captive population. A few loci with two alleles in both wild and captive populations were detected alongside with loci with 14 alleles in the wild population and three alleles in the captive population. The number of alleles of the polymorphic markers in both populations is presented in Figure 4.



**Figure 4.** Alleles per marker: wild vs. captive population (each point may indicate more than one marker for a total of 144 markers).

For 75% of the polymorphic markers, 2 to 4 alleles were identified for the captive population whereas the number of alleles per markers for the wild type population was more varied as shown in Figure 5.



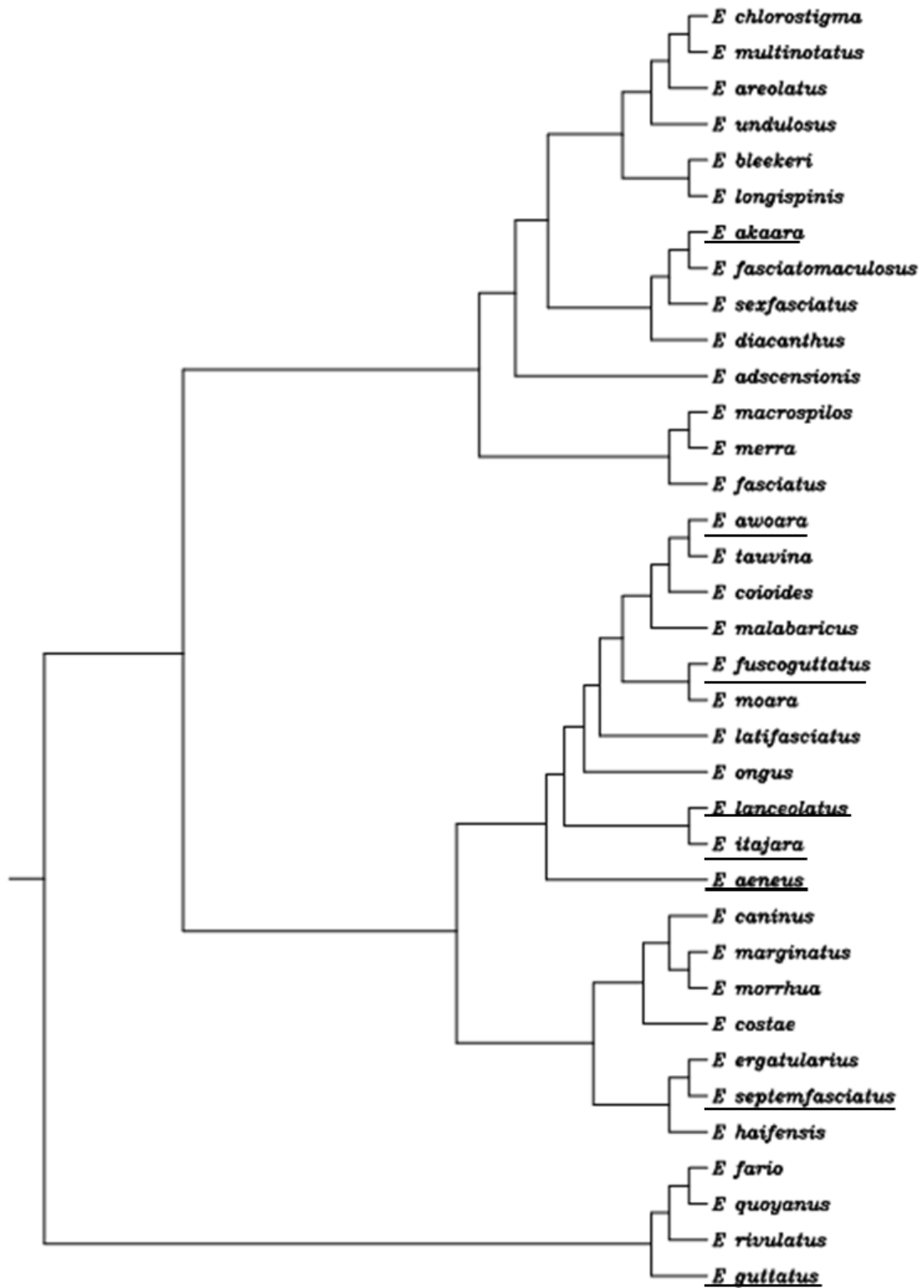
**Figure 5.** No. of alleles per marker in wild vs. captive populations

### 3.5 COI sequences comparison of *Epinephelus* species

In addition to *E. aeneus* COI sequence, sequences for 35 different *Epinephelus* species were located for phylogenetic analysis by searching the BOLD system.

In order to determine the consensus sequence for each of the *Epinephelus* species, 3-26 sequences were downloaded and aligned for every species separately. *Epinephelus* COI sequences showed small divergence among intra-specific sequences (<3 base substitutions, <0.5%) but greater differences for inter-specific sequences (>12 base substitutions; >2%) as presented in Figures 6 and 7. Comparison of the DNA- and protein-based phylogenetic trees built from the same 36 consensus sequences showed differences in clustering of species and in the relative distance of different species from *E. aeneus*.

According to the DNA-based tree, all *Epinephelus* species were clustered in three major groups in which all species used for generation of heterologous markers except *E. guttatus* and *E. akaara* belong to the same group as presented in Figure 6. However, according to the protein-based tree presented in Figure 7, all *Epinephelus* species were clustered in two major groups and three additional small groups which include one or two species. In addition, according to the protein based tree *E. septemfasciatus* is the most phylogenetically distant from *E. aeneus* among the grouper species that were analyzed.



**Figure 6.** DNA-based phylogenetic tree of *Epinephelus* species. The species used to generate heterologous markers are underlined.

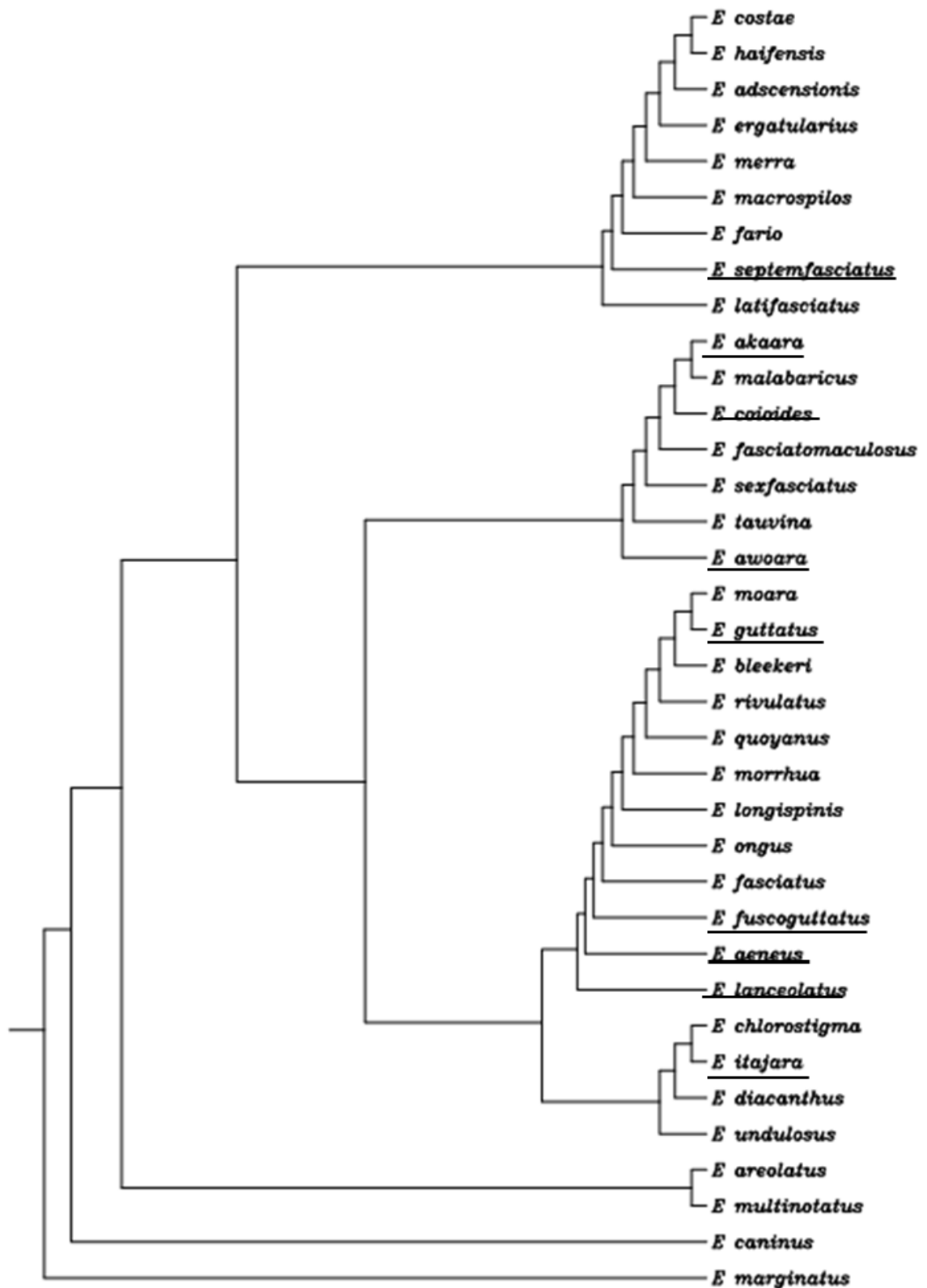


Figure 7. Protein-based phylogenetic tree of *Epinephelus* species. The species used to generate heterologous markers are underlined.



### 3.6 Parentage analysis

Genotypes of larvae for a set of eight microsatellite loci, presented in Table 9, were used for parentage assignment. Comparison of PIC values may indicate the statistical power of the various markers to detect polymorphism (Liu and Cordes, 2004).

Table 9. Polymorphism information of the parentage assignment set of markers

Marker	No. of samples (WT)	No. of alleles	PIC*
<i>ARO1003</i>	26	8	0.82
<i>ARO1045</i>	22	10	0.84
<i>ARO1083</i>	24	9	0.86
<i>ARO1084</i>	24	7	0.70
<i>ARO1105</i>	20	10	0.87
<i>ARO1120</i>	22	10	0.85
<i>ARO1131</i>	26	9	0.87
<i>ARO1137</i>	20	6	0.77

\* Polymorphic Information Content (PIC; Botstein et al., 1980)

Parental combinations were identified for a total population of 59 larvae which was comprised of three different batches: seven larvae were incompletely assigned to a single pair of parents for seven received of the eight tested markers whereas all the other 52 larvae were assigned to a single pair of parents for all eight markers. Out of 59 potential parental combinations four have been identified. Distribution of larvae assigned to each pair is presented in Table 10; only larvae with full assignment confidence (all eight markers) are presented. The identified pair combinations included two males (#2, #3) and two females (#8, #9).

Larvae of batch A were all assigned to a single pair of parents, male #2 x female #9, whereas analysis of the two other batches revealed larger variation of parental combinations. Among batch B 63.6% of larvae were assigned to the parental pair male #2 x female 9 and in batch C 77.7% of larvae were assigned to the parental pair male #3 x female #9.

Proportions of total offspring number produced by each parental combination were different. The parental pair male #2 x female #9 produced 69.2% of all offspring and the other pairs produced 0.04% to 21.2% of offspring tested. Furthermore, male #2 was responsible for 73.1% of the total male contribution and female #9 contributed 90.4% of the total female contribution.

Table 10. Frequency of larvae from each parental pair in different batches

Batch	Parental combination				
	♂2 x ♀9	♂3 x ♀9	♂2 x ♀8	♂3 x ♀8	
A	21	0	0	0	
B	14	4	1	3	
C	1	7	1	0	
Total	36	11	2	3	52

### 3.7 Linkage mapping population

DNA was extracted from 256 fingerlings and a preliminary two-step PCR reaction has been performed for the first group of 96 samples (plate No. 1) using four microsatellites of the parental test set of markers. Genotyping for the four markers was not sufficient for parentage assignment. It appears that genotyping of the whole fingerlings population for all eight markers of the parental test is needed for establishment of a population with identified parents for development of a segregating population necessary for linkage mapping.

## 4. Discussion

### 4.1 Microsatellite markers development

At the beginning of the research no genomic information about *E. aeneus* was available. Therefore, the two main options for development of novel microsatellites in a species with no sequenced genome were: A) *De novo* sequencing of DNA samples using next generation sequencing methods. This procedure has recently been demonstrated among several fish species such as sparids (Reid et al., 2012) and breams (Wang et al., 2012). The results can be a large number of sequences containing microsatellites from which only part may be suitable for primer development. However, the procedure was relatively expensive. B) Cross species amplification. Sequences containing microsatellites that have been published for closely related *Epinephelus* species may be used for designing primers for heterologous markers. This procedure had been demonstrated before for carps (Yue et al., 2004), salmonids (Williamson and May, 2002; Dehaan and Ardren, 2005), Acipenseridae (King et al., 2001) and other species (Rico et al., 1996). Cross-amplification can be very cost effective, based on available sequence data. Therefore, it was decided to use heterologous primers for development of microsatellite markers for *E. aeneus*.

The results of this study showed over 90% cross-species amplification and 62% polymorphic microsatellites were obtained. This is a relatively high rate of success compared to success rates in other projects such as 41.7% in carp species (Yue et al., 2004) or 26% in tilapia (Palti et al., 2002). At the beginning of the research 2,100 *Epinephelus* sequences were available in NCBI database out of which 656 contained a minimum of (CA)<sub>10</sub> repeats; a similar test made recently showed about 9,500 *Epinephelus* sequences and 1599 sequences containing a minimum of (CA)<sub>10</sub> repeats. Furthermore, since the beginning of the current research in 2010 the technology of next generation sequencing has been rapidly improved and today the cost of a basic library of 100 bp sequenced fragments that will eventually yield about 200 new specific microsatellite markers for *E. aeneus* is about \$3500 (Roy J. Carver

Biotechnology Center, Urbana, IL, personal communication). This is still rather expensive comparing to the current project that has yielded 144 polymorphic markers through heterologous primers. In order to advance towards development of an adequately-densed linkage map additional markers are needed. These markers can be either heterologous markers developed from new *Epinephelus* species sequences that have recently been published or specific markers that can be derived from the application of next generation sequencing.

#### 4.2 Cost effective amplification

The two-step PCR amplification method (described in section 2.12) has proven to be very efficient, and time and cost saving. The procedure has three major advantages:

1. Economic – allowing the use of only two fluorescently dyed primers instead of purchasing more than 200 different fluorescent primers.
2. Clarity – the final product after the second PCR reaction shows less background fragments which allow easier identification of microsatellites.
3. Ease of allele calling – for most of the primer pairs there was only a single dominant amplified fragment size in the final result. Using the GeneMapper software the dominant length was displayed as the fragment size with the highest peak (Figure 3a) and was relatively easy for determination.

An attempt was made to create a multiplex test kit that will eventually include several polymorphic primer pairs that will be amplified simultaneously. The results showed that the current available heterologous markers for this species cannot be amplified simultaneously because the primers' sequences are not fully compatible with the *E. aeneus* sequence, meaning that in the relatively low annealing temperature of 55°C many undesirable products are obtained and the final amplification product is not amenable for genotyping. Therefore, in the future in order to design a multiplex set *de novo* sequencing of an *E. aeneus* sample should be made for development of specific primers that will allow simultaneous amplification at higher temperatures than 55°C during the first PCR reaction, and the use of several fluorescent primers in the second reaction.

#### 4.3 Heterologous marker polymorphism according to species of origin

The yield of heterologous polymorphic microsatellites varied according to the species of origin from 42% (*E. septemfasciatus*) to 83% (*E. lanceolatus*) (Table 7). Significance was found by comparing *E. septemfasciatus* polymorphism rate to a pool of all other species polymorphism rate, meaning that *E. septemfasciatus* was the species that significantly yielded the lowest rate of polymorphic markers. The number of markers tested varied according to the species of origin because it relied on sequences availability in NCBI database and compatibility for marker development. Thus, the amount of tested markers originating from *E. septemfasciatus* sequences was the largest. A similar number of markers were developed from *E. awoara* and *E. fuscoguttatus* which yielded a polymorphism rate of over 62% for both species. Therefore, in future efforts for developing additional heterologous markers for *E. aeneus*, sequences originating from *E. septemfasciatus* should be less preferable.

#### 4.4 Phylogenetic tree of species

The COI gene is a highly effective identification system (Hebert et al., 2003a) that has revealed cryptic diversity across the animal kingdom (April et al., 2011). COI analysis of morphologically identified *E. aeneus* broodstock and wild type samples showed that all tested fishes were of the *E. aeneus* species. Success rate of cross-species amplification varied among different loci and previous knowledge suggest that polymorphism is generally high among closely related species (Rico et al., 1996). No correlation was found between DNA-based COI analysis (Figure 6) and polymorphic markers rate. However, protein-based COI analysis showed that *E. septemfasciatus* is the most phylogenetically distant from *E. aeneus*. The protein-based tree had higher discrimination power than the DNA-based tree as was found earlier for cichlids (Shirak et al., 2009). This may explain the lower yield of polymorphic markers in *E. aeneus* among markers based on sequence information from *E. septemfasciatus*. Furthermore, classification of the tribe Epinephelini of the subfamily Epinephelinae have been constantly evolving due to new research tools, thus suggesting that *E.*

*septemfasciatus* will not be included in the *Epinephelus* genus and would form a new group (*Hyporthodus*) within the Epinephelini along with several additional members of the same tribe (Craig and Hasting, 2007). The new classification had occurred in several fish databases during 2010. Moreover, during the process of phylogenetic tree development the number of sequences used can greatly affect the structure and accuracy of the tree, therefore each additional sample can change the structure of the tree. There are 86 valid *Epinephelus* species found in FishBase out of which the sequences of only 36 species were used for the current phylogenetic analysis. In the future a phylogenetic analysis including all *Epinephelus* or Epinephelinae species COI sequences may reveal higher correlation between yield of polymorphic markers and phylogenetic relationship. Such correlation can increase the efficiency of heterologous markers development process for construction of a linkage map and determination of QTL.

Information about geographical distribution of *Epinephelus* species used to generate heterologous markers (Table 6) revealed large variation. There is no clear correlation between geographic relationships among the seven *Epinephelus* species and the yield of polymorphic markers. *Epinephelus* species are strong and relatively large fishes, therefore it is possible that a group of fishes of one species will migrate across an ocean and evolve to a new sub-species or species over time. In this scenario the species will genetically resemble its ancestor more than species located in its new surroundings. All *Epinephelus* species tested were eventually suitable for heterologous markers generation regardless of their geographical distribution, therefore future procedure of development of new markers for *E. aeneus* should not strongly rely on the geographic distribution of the species.

#### 4.5 Reproduction strategy

Until now no information about *E. aeneus* reproduction strategy in the wild or in captivity was known and breeding of the broodstock was made according to morphological and behavioral characteristics. Therefore, the process of parental assignment is the first step needed toward understanding of this mechanism.

A set of eight polymorphic markers has shown sufficient statistical power to assign the majority of the offspring population to a single pair of parents (Table 10). In the future if genetic variance would decrease along further generations, additional polymorphic markers may be used to extend the number of markers in the set.

Within a relatively small number of brooders there were 59 potential parental pair combinations from which the assignment of parents to offspring needed to be made using microsatellite markers. Results showed varied paternity combinations (Table 10). While for the first batch a single parental pair was detected, parentage assignment of progeny population in the second and third batch revealed that more than a single parental pair was involved. Nevertheless, the results also indicate that when multiple pairs take part in *E. aeneus* spawning in captivity a dominant pair is present in each spawn. In view of the small number of progeny analyzed, and the low contribution of some parental pairs, it is possible that a larger sample of progeny from the first batch would have revealed that more parents were involved.

In light of these results, the conclusion is that *E. aeneus* is polygamous under the current captivity conditions with a relatively small number of brooders that participate in the actual breeding process. Multiple paternity and unequal contribution of parents have been observed in other fish species such as gilthead seabream, *Sparus auratus* (Brown et al., 2005), barramundi, *Lates calcarifer* (Frost et al., 2006), *Solea senegalensis* (Porta et al., 2006), Atlantic cod, *Gadus morhua* (Herlin et al., 2008), humpback grouper, *Melanogrammus aeglefinus* (Trippel et al., 2009), European seabass, *Dicentrarchus labrax* (Novel et al., 2010) and *Cromileptes altivelis* (Na-Nakorn et al., 2010).

These findings demonstrate a rapid decrease in allele richness in successive generations of the captive population as a result of genetic bottleneck due to a small number of parents contributing their genetic material to the offspring population. As mentioned previously, parentage test is only the first step in *E. aeneus* reproduction strategy analysis and further investigation is needed.

#### 4.6 Genetic variation

Polymorphism level, calculated as the rate of polymorphic loci, was similar in the captive and wild type. However, the captive broodstock, which is only one generation removed from the wild, had significantly ( $P = 1.85 \times 10^{-20}$ ) lower average number of alleles per marker by approximately two fold (Table 8) and lower average of PIC ( $2.96 \times 10^{-10}$ ). In addition, the maximum number of alleles per marker found in the wild type ( $n=14$ ) was double than that found in the captive population. Furthermore, the distribution of the number of alleles per marker among polymorphic markers is different (Figure 5). Among the captive population most of the markers contain two or three alleles and the number of markers with higher allelic richness is decreasing whereas among the wild type population there is no clear change in allele richness. This can occur due to errors during the genotyping process, rare alleles that have been detected among the wild population fish or due to a small sampling size which have a great effect on the observed allelic richness (Leberg, 2002; Porta et al., 2006). It is clear that there is a reduction in genetic variation among the captive broodstock population comparing to the wild. Therefore, polymorphism level in the captive broodstock must be further monitored for potential loss of allelic diversity and increased inbreeding in current and future generations (Allendorf, 1986).

Aquaculture practices, in the past and present, have consistently produced populations with significantly lower genetic diversity than their wild progenitor populations leaving substantial cause for concern (Lind et al., 2009). Decrease in genetic variability with artificial rearing was observed in other fish species, e.g. *Salmo salar* (Cross and King, 1983), *Lates calcarifer* (Frost et al., 2006), *Solea*



*senegalensis* (Porta et al., 2006), *Salmo clarkii* (Allendorf and Phelps, 1980) and *Cyprinus carpio* (Kohlmann et al., 2005).

In fishes, inbreeding depression has been manifested by reduced progeny survival, growth, food conversion efficiency and increased frequency of deformities (Trippel et al., 2009). A total of 40.6% reduction in allelic richness was observed among the captive stock compare to the wild type population, therefore it is of a great importance to increase the genetic polymorphism among *E. aeneus* broodstock population. One option for preserving as much genetic variation as possible is to have multiple broodstocks (Aho et al., 2006). These broodstocks should be managed according to genetic analysis. Currently the total *E. aeneus* capture stock available at the NCM is divided into groups according to behavioral factors, meaning the potential brooders are held in a tank, in time some of the fishes are starting to act as a school while others stand aside. The individuals that do not participate in the schooling-like behavior are removed until eventually a unified group of brooders is created. Although there is much more to investigate about *E. aeneus* reproduction process in captivity, it is necessary to combine genetic analysis in the process of broodstock management.

Usually the  $F_1$  generation is preferred as a broodstock population because it was spawned in captivity and therefore more likely adjusted. In view of the new finding of decrease in population's genetic variance throughout generations and due to unequal contribution of genetic material of parents an important conclusion is that current NCM brooders stock of the domestication project must be introgressed with new  $F_1$  individuals originating from wild type fish in order to increase population size as well as genetic variation. Using the 144 new polymorphic markers the small number of breeders may be selected for high polymorphism.

#### 4.7 Genetic mapping of markers

The first step towards mapping the 144 microsatellite markers in the white grouper is the development of a segregating population. By using the parentity test with eight polymorphic markers, several single-pair families can be identified *a posteriori*. In the future, the segregating population will be genotyped for all available markers to build the first-generation white grouper linkage map. This will provide a genetic infrastructure for breeding the white grouper following other species such as the *Danio rerio* (Shimoda et al., 2000), *Oreochromis niloticus* (Kocher et al., 1998), *Sparus aurata* (Franch et al., 2006), *Cyprinus carpio* (Sun and Liang, 2004) and others.

Microsatellite markers may be used for selection of brooders thus minimizing genetic relatedness. To enable this approach, an infrastructure of hundreds of markers that are distributed genome-wide is required.

A paper based on the results of this work was accepted for publication:

Dor, L., Shirak, A., Gorshkov, S., Ron, M., Hulata, G.,(2013). Establishment of genetic markers for breeding the white grouper (*Epinephelus aeneus*). Aquaculture, in press.  
DOI: 10.1016/j.aquaculture.2013.02.023

## 5. References

- Aho, T., Rönn, J., Piironen, J., Björklund, M., 2006. Impacts of effective population size on genetic diversity in hatchery reared Brown trout (*Salmo trutta* L.) populations. *Aquaculture*. 253, 244-248.
- Allendorf, F.W., 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*. 5, 181-190.
- Allendorf, F.W., Phelps, S.R., 1980. Loss of genetic-variation in a hatchery stock of cutthroat trout. *Transactions of the American Fisheries Society*. 109, 537-543.
- April, J., Mayden, R.L., Hanner, R.H., Bernatchez, L., 2011. Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences*. 108, 10602-10607.
- Baldwin, C., Castillo, C., Weigt, L., Victor, B., 2011. Seven new species within western Atlantic *Starksia atlantica*, *S. lepicoelia*, and *S. sluiteri* (Teleostei, Labrisomidae), with comments on congruence of DNA barcodes and species. *ZooKeys*. 79, 21-72.
- Bhandari, R.K., Komuro, H., Nakamura, S., Higa, M., Nakamura, M., 2003. Gonadal restructuring and correlative steroid hormone profiles during natural sex change in protogynous honeycomb grouper (*Epinephelus merra*). *Zoological Science*. 20, 1399-1404.
- Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*. 32, 314-331.
- Bouza, C., Hermida, M., Pardo, B.G., Fernández, C., Fortes, G.G., Castro, J., Sánchez, L., Presa, P., Pérez, M., Sanjuán, A., de Carlos, A. Álvarez-Dios, J.A., Ezcurra, S., Cal, R.M., Piferrer, F., Martínez, P., 2007. A microsatellite genetic map of the Turbot (*Scophthalmus maximus*). *Genetics*. 177, 2457-2467.
- Bradley, K., Elmore, J.B., Breyer, J., Yaspan, B., Jessen, J., Knapik, E., Smith, J., 2007. A major zebrafish polymorphism resource for genetic mapping. *Genome Biology*. 8, R55.

- Brown, C.R., Woolliams, J.A., McAndrew, B.J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. *Aquaculture*. 247, 219-225.
- Carvalho, D.C., Neto, D.A.P., Brasil, B.S.A.F., Oliveira, D.A.A., 2011. DNA barcoding unveils a high rate of mislabeling in a commercial freshwater catfish from Brazil. *Mitochondrial DNA*. 22, 97-105.
- Chen, S.L., Ma, H.Y., Jiang, Y., Liao, X.L., Meng, L., 2007. Isolation and characterization of polymorphic microsatellite loci from an EST library of turbot (*Scophthalmus maximus*) and cross-species amplification. *Molecular Ecology Notes*. 7, 848-850.
- Chistiakov, D.A., Hellemans, B., Haley, C.S., Law, A.S., Tsigenopoulos, C.S., Kotoulas, G., Bertotto, D., Libertini, A., Volckaert, F.A.M., 2005. A Microsatellite linkage map of the European Sea bass *Dicentrarchus labrax* L. *Genetics*. 170, 1821-1826.
- Collins, D.W., Jukes, T.H., 1994. Rates of transition and transversion in coding sequences since the human-rodent divergence. *Genomics*. 20, 386-396.
- Craig, M.T., Hastings, P.A., 2007. A molecular phylogeny of the groupers of the subfamily Epinephelinae (Serranidae) with a revised classification of the Epinephelini. *Ichthyological Research*. 54, 1-17.
- Cross, T.F., King, J., 1983. Genetic-effects of hatchery rearing in Atlantic salmon. *Aquaculture*. 33, 33-40.
- Debas, L., Fostier, A., Fuchs, J., Weppe, M., Nedelec, G., Benett, A., Cauty, C., Jalabert, B., 1989. The sexuality of cultured hermaphroditic fish species: Analysis of morphological and endocrinological features in a protogynous hermaphrodite, *Epinephelus microdon*, as a basis for further research to control reproduction in the grouper. *Actes de colloques. Ifremer .Brest*. 1990.
- Debener, T., Mattiesch, L., 1999. Construction of a genetic linkage map for roses using RAPD and AFLP markers. *TAG Theoretical and Applied Genetics*. 99, 891-899.
- Dehaan, P.W., Ardren, W.R., 2005. Characterization of 20 highly variable tetranucleotide microsatellite loci for bull trout (*Salvelinus confluentus*) and

cross-amplification in other *Salvelinus* species. *Molecular Ecology Notes*. 5, 582-585.

Dettai, A., Adamowicz, S.J., Allcock, L., Arango, C.P., Barnes, D.K.A., Barratt, I., Chenuil, A., Couloux, A., Cruaud, C., David, B., Denis, F., Denys, G., Díaz, A., Eléaume, M., Féral, J.-P., Froger, A., Gallut, C., Grant, R., Griffiths, H.J., Held, C., Hemery, L.G., Hosie, G., Kuklinski, P., Lecointre, G., Linse, K., Lozouet, P., Mah, C., Monniot, F., Norman, M.D., O'Hara, T., Ozouf-Costaz, C., Piedallu, C., Pierrat, B., Poulin, E., Puillandre, N., Riddle, M., Samadi, S., Saucède, T., Schubart, C., Smith, P.J., Stevens, D.W., Steinke, D., Strugnell, J.M., Tarnowska, K., Wadley, V., Ameziane, N., 2011. DNA barcoding and molecular systematics of the benthic and demersal organisms of the CEAMARC survey. *Polar Science*. 5, 298-312.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Kazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., Weissenbach, J., 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*. 380, 152-154.

Dietrich, W.F., Miller, J.C., Steen, R.G., Merchant, M., Damron, D., Nahf, R., Gross, A., Joyce, D.C., Wessel, M., Dredge, R.D., et al., 1994. A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature Genetics*. 7, 220-245.

Duchesne, P., Godbout, M.-H., Bernatchez, L., 2002. papa (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. *Molecular Ecology Notes*. 2, 191-193.

Falconer, D.S., 1989. *Introduction to Quantitative Genetics* 3ed. Longman, New York.

FAO, 2012. Fishstat plus.

<http://www.fao.org/fishery/statistics/software/fishstat/en>

Fennessy, S.T., Sadovy, Y., 2002. Reproductive biology of a diandric protogynous hermaphrodite, the serranid *Epinephelus andersoni*. *Marine and Freshwater Research*. 53, 147-158.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse

- metazoan invertebrates. *Molecular Marine Biology and Biotechnology*. 3, 294-299.
- Franch, R., Louro, B., Tsalavouta, M., Chatziplis, D., Tsigenopoulos, C.S., Sarropoulou, E., Antonello, J., Magoulas, A., Mylonas, C.C., Babbucci, M., Patarnello, T., Power, D.M., Kotoulas, G., Bargelloni, L., 2006. A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. *Genetics*. 174, 851-861.
- Freese, E., 1959. The difference between spontaneous and base-analogue induced mutations of phage T4. *Proc Natl Acad Sci USA*. 45, 622-633.
- Frost, L.A., Evans, B.S., Jerry, D.R., 2006. Loss of genetic diversity due to hatchery culture practices in barramundi (*Lates calcarifer*). *Aquaculture*. 261, 1056-1064.
- García, F.J., Chikhi, L., Bonhomme, F., 1997. Microsatellite polymorphism and population subdivision in natural populations of European sea bass *Dicentrarchus labrax* (Linnaeus, 1758). *Molecular Ecology*. 6, 51-62.
- Gilbey, J., Verspoor, E., McLay, A., Houlihan, D., 2004. A microsatellite linkage map for Atlantic salmon (*Salmo salar*). *Animal Genetics*. 35, 98-105.
- Glamuzina, B., Skaramuca, B., Glavic, N., Kožul, V., Tutman, P., 2000. Proceedings of the seminar of the CIHEAM network on technology of aquaculture in the Mediterranean (TECAM), Zaragoza (Spain), 24-28 May 1999, Cahiers Options Méditerranéennes 2000, pp. 235-239.
- Gorshkov, S., 2010. Long-term plan for domestication of the White grouper (*Epinephelus aeneus*) in Israel. *Israeli Journal of Aquaculture-Bamidgeh*. 62, 215-224.
- Griffiths, A.J., Miller, J.H., Suzuki, D.T., Lewontin, R.C., Gelbart, W.M., 2000. An introduction to genetic analysis, 7th edition. W. H. Freeman, New York.
- Guyomard, R., Mauger, S., Tabet-Canale, K., Martineau, S., Genet, C., Krieg, F., Quillet, E., 2006. A type I and type II microsatellite linkage map of Rainbow trout (*Oncorhynchus mykiss*) with presumptive coverage of all chromosome arms. *BMC Genomics*. 7, 302.
- Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., Weissenbach, J., 1994. The 1993-94 genethon human genetic linkage map. *Nature genetics*. 7, 246-339.

- Harikrishnan, R., Balasundaram, C., Heo, M.-S., 2010. Molecular studies, disease status and prophylactic measures in grouper aquaculture: Economic importance, diseases and immunology. *Aquaculture*. 309, 1-14.
- Hassin, S., deMonbrison, D., Hanin, Y., Elizur, A., Zohar, Y., Popper, D.M., 1997. Domestication of the white grouper, *Epinephelus aeneus* .1. Growth and reproduction. *Aquaculture*. 156, 305-316.
- Hebert, P.D., Ratnasingham, S., deWaard, J.R., 2003a. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci*. 270 Suppl 1, S96-99.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., Dewaard, J.R., 2003b. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 270, 313-321.
- Heemstra, P.C., Randall, J.E., 1993. Groupers of the world : (family *Serranidae*, subfamily *Epinephelinae*) : an annotated and illustrated catalogue of the grouper, rockcod, hind, coral grouper and lyretail species : known to date. FAO, Rome.
- Herbinger, C.M., Doyle, R.W., Pitman, E.R., Paquet, D., Mesa, K.A., Morris, D.B., Wright, J.M., Cook, D., 1995. DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. *Aquaculture*. 137, 245-256.
- Herlin, M., Delghandi, M., Wesmajervi, M., Taggart, J.B., McAndrew, B.J., Penman, D.J., 2008. Analysis of the parental contribution to a group of fry from a single day of spawning from a commercial Atlantic cod (*Gadus morhua*) breeding tank. *Aquaculture*. 274, 218-224.
- Ivanova, N.V., Zemplak, T.S., Hanner, R.H., Hebert, P.D.N., 2007. Universal primer cocktails for fish DNA barcoding. *Molecular ecology notes*. 7, 544-548.
- Juario, J.V., Silapan, J.R., Silapan, L.L.J., 1993. The commercial production of green grouper fingerlings, *Epinephelus suillus*, from wild caught fry - an industry experience, In: Marte CL, Quintio GF, Emata AC (Eds.). *Proceedings of the Seminar-Workshop on Breeding and Seed Production of Cultured Finfishes in the Philippines*. 4-5 May 1993, Tigbauan, Iloilo, Philippines. (pp. 132-139).

- Kailasam, M., Arasu, A.R.T., Abraham, M., Sultana, M., Shiranee, P., Balakrishnan, M., Thiagarajan, G., Karaiyan, K., Subburaj, R., 2007. Induced sex reversal and breeding of greasy grouper *Epinephelus tauvina*. Indian Journal of Fisheries. 54, 27-36.
- King, T.L., Lubinski, B.A., Spidle, A.P., 2001. Microsatellite DNA variation in Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) and cross-species amplification in the Acipenseridae. Conservation Genetics. 2, 103-119.
- Kocher, T.D., Lee, W.J., Sobolewska, H., Penman, D., McAndrew, B., 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). Genetics. 148, 1225-1232.
- Kohlmann, K., Kersten, P., Flajshans, M., 2005. Microsatellite-based genetic variability and differentiation of domesticated, wild and feral common carp (*Cyprinus carpio* L.) populations. Aquaculture. 247, 253-266.
- Kohno, H., Ordonio-Aguilar, R., Ohno, A., Taki, Y., 1997. Why is grouper larval rearing difficult?: an approach from the development of the feeding apparatus in early stage larvae of the grouper, *Epinephelus coioides*. Ichthyological Research. 44, 267-274.
- Kucuktas, H., Wang, S., Li, P., He, C., Xu, P., Sha, Z., Liu, H., Jiang, Y., Baoprasertkul, P., Somridhivej, B., Wang, Y., Abernathy, J., Guo, X., Liu, L., Muir, W., Liu, Z., 2009. Construction of genetic linkage maps and comparative genome analysis of Catfish using gene-associated markers. Genetics. 181, 1649-1660.
- Leberg, P.L., 2002. Estimating allelic richness: Effects of sample size and bottlenecks. Molecular Ecology. 11, 2445-2449.
- Lee, B.Y., Lee, W.J., Streelman, J.T., Carleton, K.L., Howe, A.E., Hulata, G., Slettan, A., Stern, J.E., Terai, Y., Kocher, T.D., 2005. A second-generation genetic linkage map of tilapia (*Oreochromis spp.*). Genetics. 170, 237-244.
- Liao, I.C., Leaño, E.M., 2008. The aquaculture of groupers. Asian Fisheries Society, Quezon City .
- Lim, L.C., 1993. Larviculture of the Greasy Grouper *Epinephelus tauvina* F. and the Brown-Marbled Grouper *E. fuscoguttatus* F. in Singapore. Journal of the World Aquaculture Society. 24, 262-274.



- Lind, C.E., Evans, B.S., Knauer, J., Taylor, J.J.U., Jerry, D.R., 2009. Decreased genetic diversity and a reduced effective population size in cultured silver-lipped pearl oysters (*Pinctada maxima*). *Aquaculture*. 286, 12-19.
- Liu, F., Xia, J.H., Bai, Z.Y., Fu, J.J., Li, J.L., Yue, G.H., 2009. High genetic diversity and substantial population differentiation in grass carp (*Ctenopharyngodon idella*) revealed by microsatellite analysis. *Aquaculture*. 297, 51-56.
- Liu, Z.J., Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*. 238, 1-37.
- Loukovitis, D., Sarropoulou, E., Vogiatzi, E., Tsigenopoulos, C.S., Kotoulas, G., Magoulas, A., Chatziplis, D., 2012. Genetic variation in farmed populations of the gilthead sea bream *Sparus aurata* in Greece using microsatellite DNA markers. *Aquaculture Research*. 43, 239-246.
- Love, J.M., Knight, A.M., McAleer, M.A., Todd, J.A., 1990. Towards construction of a high resolution map of the mouse genome using PCR-analysed microsatellites. *Nucleic Acids Res.* 18, 4123-4130.
- Lowenstein, J.H., Amato, G., Kolokotronis, S.O., 2009. The real *maccoyii*: Identifying tuna sushi with DNA barcodes – contrasting characteristic attributes and genetic distances. *PloS one*. 4, e7866.
- Meiri-Ashkenazi, I., Pagelson, G., Rosenfeld, H., 2010. Broodstock management of the White grouper (*Epinephelus aeneus*): problems and prospects, The 11th Dan Popper Symposium Eilat (Israel).
- Moen, T., Delghandi, M., Wesmajervi, M.S., Westgaard, J.I., Fjalestad, K.T., 2009. A SNP/microsatellite genetic linkage map of the Atlantic cod (*Gadus morhua*). *Animal Genetics*. 40, 993-996.
- Moen, T., Hayes, B., Baranski, M., Berg, P., Kjøglum, S., Koop, B., Davidson, W., Omholt, S., Lien, S., 2008. A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers. *BMC Genomics*. 9, 223.
- Na-Nakorn, U., Moeikum, T., 2009. Genetic diversity of domesticated stocks of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878), in Thailand: Relevance to broodstock management regimes. *Aquaculture*. 297, 70-77.

- Na-Nakorn, U., Yashiro, R., Wachirachaikarn, A., Prakoon, W., Pansaen, N., 2010. Novel microsatellites for multiplex PCRs in the Humpback grouper, *Cromileptes altivelis* (Valenciennes, 1828), and applications for broodstock management. *Aquaculture*. 306, 57-62.
- Novel, P., Porta, J.M., Porta, J., Béjar, J., Alvarez, M.C., 2010. PCR multiplex tool with 10 microsatellites for the European seabass (*Dicentrarchus labrax*) — Applications in genetic differentiation of populations and parental assignment. *Aquaculture*. 30 ,Supplement 1, S34-S38.
- Norris, A.T., Bradley, D.G., Cunningham, E.P., 1999. Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture*. 180, 247-264.
- Norris ,A.T., Bradley, D.G., Cunningham, E.P., 2000. Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture*. 182, 73-83.
- OConnell, M., Wright, J.M., 1997. Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries*. 7, 331-363.
- Ottolenghi, F., 2004. Capture-based aquaculture: the fattening of eels, groupers, tunas and yellowtails. Food and Agriculture Organization of the United Nations, Rome.
- Palti, Y., Shirak, A., Cnaani, A., Hulata, G., Avtalion, R.R., Ron, M., 2002. Detection of genes with deleterious alleles in an inbred line of tilapia (*Oreochromis aureus*). *Aquaculture* 206, 151-164
- Porta, J., Porta, J.M., Martínez-Rodríguez, G ,.Alvarez, M.C., 2006. Genetic structure and genetic relatedness of a hatchery stock of Senegal sole (*Solea senegalensis*) inferred by microsatellites. *Aquaculture*. 251, 46-55.
- Rachlin, J., Ding, C., Cantor, C., Kasif, S., 2005a. Computational tradeoffs in multiplex PCR assay design for SNP genotyping. *BMC Genomics*. 6, 102.
- Rachlin, J., Ding, C., Cantor, C., Kasif, S., 2005b. MuPlex: multi-objective multiplex PCR assay design. *Nucleic Acids Res*. 33, W544-547.
- Ratnasingham, S., Hebert, P.D., 2007. bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*. 7, 355-364.

- Reid, K., Hoareau, T.B., Bloomer, P., 2012. High-throughput microsatellite marker development in two sparid species and verification of their transferability in the family Sparidae. *Molecular Ecology Resources*. 12, 740-752.
- Rengmark, A.H., Slettan, A., Skaala, O., Lie, O., Langaas, F., 2006. Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. *Aquaculture*. 253, 229-237.
- Rico, C., Rico, I., Hewitt, G., 1996. 470 Million years of conservation of microsatellite loci among fish species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 263, 549-557.
- Russo, T., Boglione, C., De Marzi, P., Cataudella, S., 2009. Feeding preferences of the dusky grouper (*Epinephelus marginatus*, Lowe 1834) larvae reared in semi-intensive conditions: A contribution addressing the domestication of this species. *Aquaculture*. 289, 289-296.
- Saillant, E., Ma, L., Wang, X., Gatlin, D.M., Gold, J.R., 2007. Heritability of juvenile growth traits in red drum (*Sciaenops ocellatus* L.). *Aquaculture Research*. 38, 781-788.
- Sakamoto, T., Danzmann, R.G., Gharbi, K., Howard, P., Ozaki, A., Khoo, S.K., Woram, R.A., Okamoto, N., Ferguson, M.M., Holm, L.-E., Guyomard, R., Hoyheim, B., 2000. A microsatellite linkage map of Rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics*. 155, 1331-1345.
- Sanetra, M., Henning, F., Fukamachi, S., Meyer, A., 2009. A microsatellite-based genetic linkage map of the cichlid fish, *Astatotilapia burtoni* (Teleostei): a comparison of genomic architectures among rapidly speciating cichlids. *Genetics*. 182, 387-397.
- Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G.K., Lane, R., 2005. Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 360, 1805-1811.
- Schlotterer, C., 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma*. 109, 365-371.

- Shimoda, N., Knapik, E.W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H., Fishman, M.C., 1999. Zebrafish genetic map with 2000 microsatellite markers. *Genomics*. 58, 219-232.
- Shirak, A., Cohen-Zinder, M., Barroso, R.M., Seroussi, E., Ron, M., Hulata, G. 2009. DNA barcoding of Israeli indigenous and introduced cichlids. *The Israeli Journal of Aquaculture - Bamidgeh*. 61(2), 83-88.
- Shirak, A., Grabherr, M., Di Palma, F., Lindblad-Toh, K., Hulata, G., Ron, M., Kocher, T.D., Seroussi, E., 2010. Identification of repetitive elements in the genome of *Oreochromis niloticus*: Tilapia repeat masker. *Marine Biotechnology*. 12, 121-125.
- Sun, X., Liang, L., 2004. A genetic linkage map of common carp (*Cyprinus carpio* L.) And mapping of a locus associated with cold tolerance. *Aquaculture*. 238, 165-172.
- Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*. 17, 6463-6471.
- Trippel, E.A., Rideout, R.M., O'Reilly, P.T., Herbinger, C.M., Neil, S.R.E., Hamilton, L., 2009. Communal spawning leads to high potential for inbreeding in gadoid aquaculture. *Aquaculture*. 296, 27-35.
- Vandeputte, M., Kocour, M., Mager, S., Dupont-Nivet, M., De Guerry, D., Rodina, M., Gela, D., Vallod, D., Chevassus, B., Linhart, O. 2004. Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). *Aquaculture*. 235, 223-236.
- Vignal, A., Milan, D., SanCristobal, M., Eggen, A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics, Selection, Evolution : GSE*. 34, 275-305.
- Villanueva, B., Verspoor, E., Visscher, P.M., 2002. Parental assignment in fish using microsatellite genetic markers with finite numbers of parents and offspring. *Animal Genetics*. 33, 33-41.
- Wang, C., Bai, Z., He, X., Lin, G., Xia, J., Sun, F., Lo, L., Feng, F., Zhu, Z., Yue, G., 2011. A high-resolution linkage map for comparative genome analysis and QTL fine mapping in Asian seabass, *Lates calcarifer*. *BMC Genomics*. 12, 174.

- Wang, C.M., Lo, L.C., Zhu, Z.Y., Lin, G., Feng, F., Li, J., Yang, W.T., Tan, J., Chou, R., Lim, H.S., Orban, L., Yue, G.H., 2008. Estimating reproductive success of brooders and heritability of growth traits in Asian sea bass (*Lates calcarifer*) using microsatellites. *Aquaculture Research*. 39, 1612-1619.
- Wang, J., Yu, X., Zhao, K., Zhang, Y., Tong, J., Peng, Z., 2012. Microsatellite development for an endangered Bream *Megalobrama pellegrini* (Teleostei, Cyprinidae) using 454 sequencing. *International journal of molecular sciences*. 13, 3009-3021.
- Wang, M., Godfrey, M., 1998. Prenatal and presymptomatic diagnosis of Marfan syndrome using fluorescence PCR and an automated sequencer, pp. 49-54.
- Ward, R.D., Hanner, R., Hebert, P.D.N., 2009. The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*. 74, 329-356.
- Williamson, K.S., Cordes, J.F., May, B., 2002. Characterization of microsatellite loci in chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. *Molecular ecology notes*. 2, 17-19.
- Yue, G.H., Ho, M.Y., Orban, L., Komen, J., 2004. Microsatellites within genes and ESTs of common carp and their applicability in silver crucian carp. *Aquaculture*. 234, 85-98.
- Yue, G.H., Zhu, Z.Y., Lo, L.C., Wang, C.M., Lin, G., Feng, F., Pang, H.Y., Li, J., Gong, P., Liu, H.M., Tan, J., Chou, R., Lim, H., Orban, L., 2009. Genetic variation and population structure of Asian seabass (*Lates calcarifer*) in the Asia-Pacific region. *Aquaculture*. 293, 22-28.
- Zheng, K., Lin, K., Liu, Z., Luo, C., 2007. Comparative microsatellite analysis of grass carp genomes of two gynogenetic groups and the Xiangjiang river group. *Journal of Genetics and Genomics*. 34, 321-330.
- Zheng, X., Kuang, Y., Zhang, X., Lu, C., Cao, D., Li, C., Sun, X., 2011. A genetic linkage map and comparative genome analysis of common carp (*Cyprinus carpio* L.) using microsatellites and SNPs. *Molecular Genetics and Genomics*. 286, 261-277.
- Zilberman, N., Reikhav, S., Hulata, G., Ron, M., 2006. High-throughput genomic DNA extraction protocol from tilapia's fin tissue. *Aquaculture*. 255. 597-599.

# Appendix I

**Table.** Primers sequences for the 144 polymorphic markers.

Marker Name	Forward sequence	Reverse Sequence
ARO1001	CCTAGGTGATCAAGATCTCGCAAAGTGCATGTCACTCACACT	CACCAATGGAGTATTCTTCA
ARO1002	CCTAGGTGATCAAGATCTCGCTGCGGATGTAGATGTCCT	CACATGCAGCTCTCTCCTT
ARO1003	CCTAGGTGATCAAGATCTCGCGTGCAAGGCAAGCTGTGTTA	AGCAGGCATCTTGTTATCTGG
ARO1004	CCTAGGTGATCAAGATCTCGCTGCGATTTAGTCAGACACTC	GCAGACACAACAGTGAAGGTA
ARO1005	CCTAGGTGATCAAGATCTCGCTACTCAATGCGCGCTAGA	TGCTCTCCTGCAGTGGTG
ARO1006	CCTAGGTGATCAAGATCTCGCCCTGTCCCTTCACCTCATC	AAGCTAATGAGGAATGACATC
ARO1007	CCTAGGTGATCAAGATCTCGCCATTCTCATTAGCTTCCTGCT	CATTTTACAGTCTCGTGGTTTCA
ARO1008	CCTAGGTGATCAAGATCTCGCTGACAACTCACTTCCTGTTA	GGTAAATTAACACTGCGTGC
ARO1009	CCTAGGTGATCAAGATCTCGCCATTCTCATTAGCTTCCTGCT	CATTTTACAGTCTCGTGGTTTCA
ARO1010	CCTAGGTGATCAAGATCTCGCCGCGATTTAATTTACCAGC	TGACTAAACAAGTGAAGTGTCA
ARO1011	CCTAGGTGATCAAGATCTCGCCCTGTAGATTATTGTGTGTGGAC	CCTCCATCACACTAGTTGAAAG
ARO1012	CCTAGGTGATCAAGATCTCGCCGTGACAGATGGACAGATCG	AAATCCAATCCAGCAGATCG
ARO1013	CCTAGGTGATCAAGATCTCGCGTAACTCTCCAGGGAA	GAATCTCCTCTGCTGTCAGT
ARO1014	CCTAGGTGATCAAGATCTCGCAGCGATGGAGAGAAGGATGA	CTCTTGAAAAGCAGCGAAGC
ARO1015	CCTAGGTGATCAAGATCTCGCCTGCACCAAAGCACCTAC	GAACAGTTTCATGCCACAA
ARO1016	CCTAGGTGATCAAGATCTCGCGCTTGCTCCACAGACATG	TTGATGAAAGGGAACAATAG
ARO1017	CCTAGGTGATCAAGATCTCGCGCTATGCAGCACATAATT	TGCTGTAAAGGGCTATACTTT
ARO1018	CCTAGGTGATCAAGATCTCGCGCTTCTAAGTTTCAACTGGTT	AATAACGGCTGCATTTCAT
ARO1019	CCTAGGTGATCAAGATCTCGCGCATCTGTATCACGTAGC	CCAACAAATACAACCGTCT
ARO1020	CCTAGGTGATCAAGATCTCGCAATGAGGATCTATTATAGCTGTC	CTGTTACGCTCAGTTTGGA
ARO1021	CCTAGGTGATCAAGATCTCGCTCAAAGAGCTCCTGGAAAC	GTTAGAGGTCATTGAGAGGTG
ARO1022	CCTAGGTGATCAAGATCTCGCTTGATAAGACATTAGAGTTCCTT	GAGGCTAAACAAGCTCAATAG
ARO1023	CCTAGGTGATCAAGATCTCGCCAGTGTGACAGGAAACGT	GTCCTGCACCTTATTGGG
ARO1024	CCTAGGTGATCAAGATCTCGCCACATTTATTGCCTGCTATC	GCGTTTCCATAGGGACTG
ARO1025	CCTAGGTGATCAAGATCTCGCAGACAGAGTGCTAAACATGCT	TGGATTGTAAATACCAGTTGG
ARO1026	CCTAGGTGATCAAGATCTCGCATGTTTCCAATCAAACAAAC	ATTGTGTCCATTACTGGTTG
ARO1027	CCTAGGTGATCAAGATCTCGCTTACCTGAGGCACAGAGG	GTGAAAGGCTTGTTTCGTT
ARO1028	CCTAGGTGATCAAGATCTCGCGCAAAGGAGCTCATCCTC	TTGCAAATTCATGAGTCT
ARO1029	CCTAGGTGATCAAGATCTCGCTGAAATATGTAAGTGATGGGAA	GAGCCCAGTGTGTCAGGA
ARO1030	CCTAGGTGATCAAGATCTCGCAGCCGAGCTCTCTCTCT	CGCTATCAAAGAATCCAGATC
ARO1031	CCTAGGTGATCAAGATCTCGCCCTTAGTCACACTACCGCA	TGTGAAATGTGGAGACCAT
ARO1032	CCTAGGTGATCAAGATCTCGCCGCGCTATATAACAGCGTTACTC	TGACAAAAATGGACACTATCTCTCTC
ARO1033	CCTAGGTGATCAAGATCTCGCAGCAATATCCCTGAGAGGA	TCTCTGGGAACCTTGTAAGC
ARO1034	CCTAGGTGATCAAGATCTCGCCCAAATCCATGAATGTACAGAT	TTTCCACATTGTGCAGGC
ARO1035	CCTAGGTGATCAAGATCTCGCTGTGAGGAGATGCGACAC	CACTTGAGTGAAAGTGAAGA
ARO1036	CCTAGGTGATCAAGATCTCGCGAGGAGAGAGAAGAAAGAAGC	TTGTGTGAGTGTACGCCTT
ARO1037	CCTAGGTGATCAAGATCTCGCAGAGTGAAACACAAGGAATAAA	GGGAGAGTTGTTTCAGGAC
ARO1038	CCTAGGTGATCAAGATCTCGCATCAATAACACCTACAGTTCAGT	CAGGACATCCTTCAGAGTTC

ARO1039	CCTAGGTGATCAAGATCTCGCCCAATAGGATCAGGATTCAC	TGCCTCTGATCTATGGACC
ARO1040	CCTAGGTGATCAAGATCTCGCAAGAATGCAATATGTGGCA	CATCACTGTACGCAGCACT
ARO1041	CCTAGGTGATCAAGATCTCGTGTCTTGTGTTGATTTATGGA	CAAAGAAATAACCCAACATG
ARO1042	CCTAGGTGATCAAGATCTCGCGCGGAGGGATATATTTGG	TGAAGTGATGTATGTTGACAGTT
ARO1043	CCTAGGTGATCAAGATCTCGCGTGGAATCTCACCCGACTGT	TCTGCCAGCCTTACTGGTTC
ARO1044	CCTAGGTGATCAAGATCTCGCAAGTGCTACTCCAGGAACC	CATCTTTCTGGCCTGTAAGTA
ARO1045	CCTAGGTGATCAAGATCTCGCCACGAAATTTTGGCTGAT	GAGAAAGTGGCAATATTTGAC
ARO1046	CCTCGAGGTGGACGGTATCGGCCGACCATGAGAGACAAAGAG	GTTGGACTCATTCTGCC
ARO1047	CCTCGAGGTGGACGGTATCGGCTCTTCCATTCAAGCGTC	GTCTGTATGTCGGCTGCC
ARO1048	CCTCGAGGTGGACGGTATCGGCCAGCTGTCATCAGCTATGTTT	CCAGAAGTAATGTGGTCGTCT
ARO1049	CCTCGAGGTGGACGGTATCGGCCTACCAGCCTGGCGTCAT	GACAAACACCTCCCCTGAAA
ARO1050	CCTCGAGGTGGACGGTATCGGCTGAGTGTGTGCTTGGCTAG	CCTCTCTTCTCTGTATGGT
ARO1051	CCTCGAGGTGGACGGTATCGGCGCAATATGCTGTAGAGCTGTG	CCTGCTGTAACAAACTCTTGT
ARO1052	CCTCGAGGTGGACGGTATCGGCACAGTGTGATTGGTAATTAGATT	CCCTGATGTGTGATCATAATAG
ARO1053	CCTCGAGGTGGACGGTATCGGCTTGAACACACATCTTTCTGC	TGAACAGAGACTCACCGG
ARO1054	CCTCGAGGTGGACGGTATCGGCCACCTCTACCATGAACATA	TCGGATCTCCTCTGGTGT
ARO1055	CCTCGAGGTGGACGGTATCGGCCCTGCAGAGCATCACTACA	TAAGTGTGTGGCCCAAAC
ARO1056	CCTCGAGGTGGACGGTATCGGCCCACTTCATAGCTGTGATAA	TGAGATTTCTGGAGTCCGT
ARO1057	CCTCGAGGTGGACGGTATCGGCAGCATCACATACAGGCAGT	GCTCTACGTATGTGTGAGAAA
ARO1058	CCTCGAGGTGGACGGTATCGGCTTGTGAGTGACTTCTACAGG	AAGGAACAAGTCTCTGTTGAA
ARO1059	CCTCGAGGTGGACGGTATCGGCCAATCAAATCTAAAGCGACA	GTGGACTCTGCTTCTGAGAG
ARO1060	CCTCGAGGTGGACGGTATCGGCTGGGTGGAGAGCCTATCA	GGCCTGTCACTTGCTACA
ARO1061	CCTCGAGGTGGACGGTATCGGCATCTTGGTTTGTCTAGTCTTT	ACTGAACAAGCTGTGATCTCT
ARO1062	CCTCGAGGTGGACGGTATCGGCTGTTTCGCCACAGAAGCTAA	TGTTGGTACATCTGTGTTTGTGG
ARO1063	CCTCGAGGTGGACGGTATCGGCCACTTCTTCTCTCCTTCTT	GGTTGCTGCTGTTTACT
ARO1064	CCTCGAGGTGGACGGTATCGGCAGAGACGTTACATGGCTCA	CAGGATAGCCATAATGGG
ARO1065	CCTCGAGGTGGACGGTATCGGCAACAGAGGAAGCAAGAAAGA	CTTCTCCCTTCGATCACTT
ARO1066	CCTCGAGGTGGACGGTATCGGCTGCAGTTTGTGTTTATGAATAG	TGGAAACATGAACAGTATGTT
ARO1067	CCTCGAGGTGGACGGTATCGGCGAGGGCAGAGATGAACGG	GGATTCTCTCAGCCTCTAT
ARO1068	CCTCGAGGTGGACGGTATCGGCCAACTTAGTGCATGTATAGGTCT	TTGTTGTAATATGACTGTTGAC
ARO1069	CCTCGAGGTGGACGGTATCGGCCACAGAATCGGGATTGGT	GGTTATAGTTATTGGAGAGAAAG
ARO1070	CCTCGAGGTGGACGGTATCGGCTCCCTGTCTGTGTCCTACA	GGTTATTACTTCTTCAACTGTGA
ARO1071	CCTCGAGGTGGACGGTATCGGCGCACCCACAGTAATAAGAA	TTACAAGTGCTGGAGTGTGT
ARO1072	CCTCGAGGTGGACGGTATCGGCTCGGTGGTGTGAGTTTGTGTA	CTGAATCAATGACGTAAAGG
ARO1073	CCTCGAGGTGGACGGTATCGGCCAGCAGCTTCACATCCACAT	CTGGGCTCTCTCATGGTCA
ARO1074	CCTCGAGGTGGACGGTATCGGCGACCTAAGACACACAAAGACTTA	ATTATGCCACAGCTCCAG
ARO1075	CCTCGAGGTGGACGGTATCGGCTGAGTTGAGGGAGGCTTT	TGTAATGTGAAGTGGCTGG
ARO1076	CCTCGAGGTGGACGGTATCGGCTGTTTGTGTGAAATGTAGAGAG	GAATGATATGGATATGGAAGG
ARO1077	CCTCGAGGTGGACGGTATCGGCGGCCTGTGATCTCTGACCTA	CCCTGAACCTGGGTGAAG
ARO1078	CCTCGAGGTGGACGGTATCGGCCGCTTCTTCTCTGAAAGACA	GCAGAAGATCTGCAGGGAG
ARO1079	CCTCGAGGTGGACGGTATCGGCGCTAGCTCTAACCAGAGAAT	TGGGCTGAACCTTGAGGT
ARO1080	CCTCGAGGTGGACGGTATCGGCAGAAAGTGTAGGGTCTCCTA	ACAAACCCTCACTCACAGAT
ARO1081	CCTCGAGGTGGACGGTATCGGCGCCTGAATGAGCTGCTAACC	TCACACTGTGCAAGTCTCTGG
ARO1082	CCTCGAGGTGGACGGTATCGGCCACCCTGCACGTAAGAAAT	CAACACAACCATCGATCAGC
ARO1083	CCTCGAGGTGGACGGTATCGGCCGGTCTTCTTCTCTCCC	TTACTGTTGATTGAGTTGTTGT
ARO1084	CCTCGAGGTGGACGGTATCGGCGGGTTTATTTCAAAGGTCAG	CCCAATGAGGTGTTCAATAT

<i>ARO1085</i>	CCTCGAGGTGGACGGTATCGGCGATGCTGTGTATGACTGTCAC	CAGATTACGATGGCTTATCA
<i>ARO1086</i>	CCTCGAGGTGGACGGTATCGGCCAGGCCGTGTCTGTAATG	GGGAAAAGAGGTGAGGAGAT
<i>ARO1087</i>	CCTCGAGGTGGACGGTATCGGCGCTGATGTGATTTATTCAGATAC	CCCTGCATTAGCCTTAATT
<i>ARO1088</i>	CCTCGAGGTGGACGGTATCGGCGTTCATCATCAGATTTGGAATT	AAATAACCTCACACAGACACAC
<i>ARO1089</i>	CCTCGAGGTGGACGGTATCGGCCACAGGAGCAGAATGTACC	AACAGCACAAAATGCCACTG
<i>ARO1090</i>	CCTCGAGGTGGACGGTATCGGCCAGGTGCCACAGAAAAGA	GGTACGTATCGTATCATTAGTTT
<i>ARO1091</i>	CCTCGAGGTGGACGGTATCGGCCAAGCAGAATAAAGATGAG	CTTACGTCCAACACCAAGT
<i>ARO1092</i>	CCTCGAGGTGGACGGTATCGGCGTATTCAAGGTGTGTGTTTAAA	TGTGCATTTCGTCATGTG
<i>ARO1093</i>	CCTCGAGGTGGACGGTATCGGCTCTCCTACCATTGCTCATC	AGTGGTGACCCCTCTTCGT
<i>ARO1094</i>	CCTCGAGGTGGACGGTATCGGCGAGAGGATGATGTCTATAAGGA	GATTGGTGAGTGTGAGAGGA
<i>ARO1095</i>	CCTCGAGGTGGACGGTATCGGCAACGGAAACAGAATAACCC	TCTTTCAGCTGCTGTGCA
<i>ARO1096</i>	CCTCGAGGTGGACGGTATCGGCAAATCTGTCTAGCGACTACATG	CAAATAAGCACCAGTCTTGTT
<i>ARO1097</i>	CCTCGAGGTGGACGGTATCGGCTTCTCATAAGCAAAGCCTT	CATCACAGATAACAGTTTATTGA
<i>ARO1098</i>	CCTCGAGGTGGACGGTATCGGCGGAAGTGACCTCATCAGCAT	TGAGGGATCATGAAGCAA
<i>ARO1099</i>	CCTCGAGGTGGACGGTATCGGCGAAGAAAGACGGAAGATAGACT	GAATGCAGGGTTGTGCAG
<i>ARO1100</i>	CCTCGAGGTGGACGGTATCGGCCCTACACCAGTTTATTCTGTGA	AGTCGACACAGCTCTCATT
<i>ARO1101</i>	CCTCGAGGTGGACGGTATCGGCCATGTCTAAAGCTGAAGGTGTA	GTAAAGGTGTTAAAGTGTTC
<i>ARO1102</i>	CCTCGAGGTGGACGGTATCGGCGATCCTGCAAGGGCAGAG	GGCTTCCAGTTTGGATGTA
<i>ARO1103</i>	CCTAGGTGATCAAGATCTCGGAATAAATAATGCAAGGCTT	GCACATCAATAATACATCCAG
<i>ARO1104</i>	CCTAGGTGATCAAGATCTCGCAACACACTGAGAAGCACATTC	GTCACCATCACTGACAACG
<i>ARO1105</i>	CCTAGGTGATCAAGATCTCGCTGATAGCTTTACATGCACTCA	CTGAACCTCACCTGAAA
<i>ARO1106</i>	CCTAGGTGATCAAGATCTCGCCTCGTTACCACATCGGGACT	AACACTGGCTGGTTGCACT
<i>ARO1107</i>	CCTAGGTGATCAAGATCTCGCGGGTATAATTGGGACACACTA	AACAGAGGAAAGGAGGGA
<i>ARO1108</i>	CCTAGGTGATCAAGATCTCGCGTTGCCAGGTTACACGA	TGCTGGGAAACATTTACAG
<i>ARO1109</i>	CCTAGGTGATCAAGATCTCGCGACCAGGACGTTGAGTCC	TGTTTGGTGTCTGTTGTA
<i>ARO1110</i>	CCTAGGTGATCAAGATCTCGCCCTCTCGTCTCTGCTGA	TTGTGTTCCGATCACTGAA
<i>ARO1111</i>	CCTAGGTGATCAAGATCTCGCCTTTCTGGGTAGATTTAATTCA	CAAGTCAAAGTGCTGGACC
<i>ARO1112</i>	CCTAGGTGATCAAGATCTCGCATAATATTCGCCACTCG	TTTGTGGGAGTGGGATACT
<i>ARO1113</i>	CCTAGGTGATCAAGATCTCGCTGTTCAGGAGTGGTCAGTG	AATCTACTGGACGCTTACTCTA
<i>ARO1114</i>	CCTAGGTGATCAAGATCTCGCGGCAAAATGTGTTCTCAT	TCTATCGGAAACTGACGAA
<i>ARO1115</i>	CCTAGGTGATCAAGATCTCGCCCGTGTCTTATTGCAGATAGT	CCCTAATGAGTAAGAGGCC
<i>ARO1116</i>	CCTAGGTGATCAAGATCTCGCCAGTGATAAACGTTACTCCT	ACAGTCCTGAGCGGTCTC
<i>ARO1117</i>	CCTAGGTGATCAAGATCTCGCCAAATCCTCCTCAGTCACA	ACTACAACATGTCCACATGC
<i>ARO1118</i>	CCTAGGTGATCAAGATCTCGCACAGGAAGCTGAGCGTCT	AGAGATCATCAGGAAATCGT
<i>ARO1119</i>	CCTAGGTGATCAAGATCTCGGAAATACTGAGCCTGAGCC	GGAAGAGGAGAAGACAACAA
<i>ARO1120</i>	CCTAGGTGATCAAGATCTCGCCTCTGATGCTGTTTACACAAC	TCTCCATCGAAGGTAAAGG
<i>ARO1121</i>	CCTAGGTGATCAAGATCTCGCGGAGAGCTTCTCCACAGCAC	CTTACTGCAGCCCTCTGCTC
<i>ARO1122</i>	CCTAGGTGATCAAGATCTCGCCTATCCTGTCTGGATTAGAATAT	AGGCATTCTGCATTCATG
<i>ARO1123</i>	CCTAGGTGATCAAGATCTCGCAACCTCAATACAGCAGCAC	GCTTCCATTAAACCACACTC
<i>ARO1124</i>	<i>CCTAGGTGATCAAGATCTCGCGTTTAAATGGAAGCCCCAAT</i>	<i>CTGGGGACGTTAAGAACCAA</i>
<i>ARO1125</i>	<i>CCTAGGTGATCAAGATCTCGCACTAAGAGGATCATTATTCAAGA</i>	<i>GAGATGAACATAAATTACATGC</i>
<i>ARO1126</i>	<i>CCTAGGTGATCAAGATCTCGCGTCTTAGGCAGCGAGTTG</i>	<i>TTGTTACGCTGACGAAGAA</i>
<i>ARO1127</i>	<i>CCTAGGTGATCAAGATCTCGCAGCAGTTAACCATTTGGCA</i>	<i>TTTACGTCTGGTGTACAGC</i>
<i>ARO1128</i>	<i>CCTAGGTGATCAAGATCTCGCGCTGTTAGTTAAATTGAAAGAG</i>	<i>CTGGCGGTGTAATTTGA</i>
<i>ARO1129</i>	<i>CCTAGGTGATCAAGATCTCGCTTATGTCAGGAGGGAGA</i>	<i>TCAGAACATACTGTTAGCTGG</i>
<i>ARO1130</i>	<i>CCTAGGTGATCAAGATCTCGCCAAGCACAGAAGCTGAGC</i>	<i>AATACCTCCCTTACAAACCA</i>



<i>ARO1131</i>	<i>CCTAGGTGATCAAGATCTCGCTGTGTGTCAGAGGTGGGT</i>	<i>TGAATTCACTGCATGTTTC</i>
<i>ARO1132</i>	<i>CCTAGGTGATCAAGATCTCGCGACTGAGCCTGAACCCAGAG</i>	<i>GTGCTTCCTTCAGTGCATCA</i>
<i>ARO1133</i>	<i>CCTAGGTGATCAAGATCTCGCGATCCTGGATGTTGGAGAA</i>	<i>AGAGACGGTGCCAGTAAA</i>
<i>ARO1134</i>	<i>CCTAGGTGATCAAGATCTCGCGTCTAGATAACCCAGAAGGG</i>	<i>AAGACTGTTCCATTAACGTGTC</i>
<i>ARO1135</i>	<i>CCTAGGTGATCAAGATCTCGCGCAACCACAATCACCTCAC</i>	<i>AACACTGGCTGGTTTGCA</i>
<i>ARO1136</i>	<i>CCTAGGTGATCAAGATCTCGCTCCAGCACAGCCATAAC</i>	<i>TGATGTCTTACAAATTACAGCTT</i>
<i>ARO1137</i>	<i>CCTAGGTGATCAAGATCTCGCATGGGTATAATTAGGACACACT</i>	<i>AGGAAAGGAGGGAGGAAA</i>
<i>ARO1138</i>	<i>CCTAGGTGATCAAGATCTCGCGTACACGGTGGACTGCAG</i>	<i>GAGACAGGAGATTTGAAATGT</i>
<i>ARO1139</i>	<i>CCTCGAGGTGGACGGTATCGCTGCTCACGGCTCTGAAG</i>	<i>CAGGTTGTGATCGTTAAACA</i>
<i>ARO1140</i>	<i>CCTCGAGGTGGACGGTATCGGGCCTAAGTAACTGTGTTACAG</i>	<i>TGGCTTCACCTGGACAGAC</i>
<i>ARO1141</i>	<i>CCTCGAGGTGGACGGTATCGCTGTGCAAATCAAAGTAATTG</i>	<i>TAATTGCCTCAGAGCCCA</i>
<i>ARO1142</i>	<i>CCTCGAGGTGGACGGTATCGAGGAGCTGCTTCTATTAGAGT</i>	<i>GAATACTTTCAAGTTCATATGT</i>
<i>ARO1143</i>	<i>CCTAGGTGATCAAGATCTCGCCCCTGAGCTGCACTGTATGA</i>	<i>TACCAGGGTGTCTCCTCAGC</i>
<i>ARO1144</i>	<i>CCTAGGTGATCAAGATCTCGCTTCTGTGACTGTGCAAATCA</i>	<i>CCACAGCCAGTCTAGGTATTC</i>

---

ממוצע של 6.2 אללים בטווח של 2 עד 14 אללים לסמן נמצא באוכלוסיית הבר לעומת ממוצע של 3.7 אללים בטווח של 2 עד 7 אללים לסמן שנמצאו באוכלוסיית השבי ( $p = 1.85 \cdot 10^{-20}$ ) מצביע על ירידה בשונות הגנטית בקרב אוכלוסיית השבי לעומת אוכלוסיית הבר.

לשם ניתוח ההורות של אוכלוסיית הצאצאים פותח סט של שמונה סמנים פולימורפים. בקרב שניים עשר ההורים אותרו שני זכרים ושתי נקבות בצירופים שונים אשר תרמו חומר גנטי. עם זאת צמד הורים אחד אחראי ליצירת רב הצאצאים (69.2%). תוצאות אלה מעידות על פוליגמיה בקרב הדקר הלבן.

לאור העובדה שרק מספר מצומצם של הורים תרם חומר גנטי לצאצאים ואבחון הירידה במגוון האללים בקרב אוכלוסיית השבי בהשוואה לאוכלוסיית הבר של הדקר הלבן המסקנה היא שיש לעקוב אחר רמת פולימורפיות בקרב אוכלוסיית השבי בכדי לזהות איבוד שונות ועלייה אפשרית בשיעור ריבוי בשארות. בנוסף, בתוכניות ביות עתידיות יש להרחיב את השונות הגנטית באוכלוסייה זו והרבייה צריכה להיעשות בשילוב עם אבחנה גנטית.

מבחן ההורות משמש לקביעת אוכלוסייה מתפצלת לצורך מיפוי של כל הסמנים הזמינים. בעתיד ניתן יהיה לפתח סמנים נוספים על ידי שימוש בריצוף בשיטת הדור הבא או על ידי הגברה הטרוולוגית של סמנים ידועים ממיני דקר קרובים וכך ליצור מפת תאחיזה לדקר הלבן אשר תהווה תשתית לביות ולרביית מין זה.

## תקציר

הדקר הלבן (*Epinephelus aeneus*) הוא מין פרוטוגיני הרמפרודיט המצוי באזור הים התיכון, לאורך חופה המערבי של אפריקה ודרום אנגולה. ערכו המסחרי של הדקר הלבן גבוה, טעמו משובח, קצב גדילתו מהיר והוא יחסית עמיד בפני מחלות. משום כך מין זה הוא מועמד מבטיח לגידול מסחרי ובשנת 2010 החל פרויקט ביות הדקר הלבן על ידי המחלקה לדיג וחקלאות מים במשרד החקלאות הישראלי. מטרת הפרויקט היא פיתוח טכנולוגיות חדשות לביות, רבייה וגידול מסחרי של הדקר הלבן. הקשיים העיקריים בתהליך הם: זיהוי מהימן של המין המסתמך עד כה על הבדלים מורפולוגיים בלבד, בקרה אחר איבוד השונות הגנטית באוכלוסיית השבי והבנת אסטרטגיית הרבייה של הדקר הלבן, כלומר כמה פרטים בלהקת ההורים תורמים חומר גנטי לצאצאים.

המחקר הנוכחי מתמודד עם נושאים אלה באמצעות מספר כלים גנומיים. תחילה על מנת לזהות את מין הדג יושמה שיטת זיהוי באמצעות DNA המסתמכת על השונות ברצף הגן המיטוכונדריאלי Cytochrome C Oxidase C תת יחידה 1 (COI) אשר שימשה לזיהוי הפרטים המשתתפים בעבודה זו. עצים פילוגנטיים הורכבו על בסיס רצפי DNA וחלבונים של 36 מיני דקר שונים בכדי לבחון את הקשר בין הקרבה הפילוגנטית ליעילות הגברת סמנים הטרוולוגים לדקר הלבן. הסמנים שימשו לקביעת שיעור הפולימורפיות באוכלוסיית השבי ובאוכלוסיית הבר. לבסוף בכדי לקבוע מהי צורת הרבייה של הדקר הלבן בשבי סמנים מיקרוסטליטים פולימורפים שמשו לבדיקת הורות אשר שייכה כל אחת מ-59 פגיות שהוטלו על ידי להקת ההורים בשבי לצמד ההורים שלה. בנוסף, התקבלה אוכלוסיית דגיגים גדולה יותר אשר מתוכה תאותר באמצעות בדיקת ההורות קבוצת אחים גדולה שתשמש לפיתוח מפת תאחיזה מדור ראשון.

במחקר זה אותרו 656 רצפי דקרים המכילים קטעים מיקרוסטליטים, 233 מתוכם התאימו לפיתוח תחלים. סמנים הטרוולוגים פותחו משמונה מינים (*E. awoara*, *E. septemfasciatus*, *E. fuscoguttatus*, *E. itajara*, *E. guttatus*, *E. akaara*, *E. coiodes* and *E. lanceolatus*). ההגברה התבצעה באמצעות שיטת PCR דו שלבית וחסכונית להגברת תוצרי כל זוג תחלים כאשר נעשה שימוש בשני סמנים פלואורסנטים לצורך זיהוי וקביעת הגנוטיפ של כל תוצרי 233 התגובות. סך של 144 מקטעים נמצאו פולימורפים ותנובת הסמנים הפולימורפים לפי מין המקור נעה בין 42 ל 83 אחוזים. האחוז הנמוך ביותר נמצא למין *Epinephelus septemfasciatus* ( $P = 0.001$ ) אשר לפי ניתוח COI היה הכי מרוחק פילוגנטית מהדקר הלבן. עם זאת כל מיני הדקר שנבחרו התאימו לפיתוח סמנים הטרוולוגים.

## תודות

ברצוני להביע את הערכתי ולהודות לכל מי שסייעו לי במהלך עבודה זו:  
לפרופ' מיכה רון ולפרופ' גדעון חולתא על ההדרכה המסורה והאמון שנתתם בי, על  
שדחתם אותי להגיע למיצוי יכולותיי במהלך העבודה ועל חופש הפעולה שהענקתם לי  
במהלך כל התקופה.

תודה מיוחדת לדוקטור אנדריי שיראק על עזרתו ועל רעיונותיו הנפלאים אשר תרמו רבות  
לעבודה זו, וכן על הליווי וההדרכה הסבלנית מהיום הראשון ולכל אורך הדרך.  
לדוקטור סרגיי גרושקוב וצוות מלח"י באילת על תרומת הדוגמאות למחקר ועל העזרה  
המקצועית.

לדוקטור אייל סרוסי על נכונותו לעזור בפתרון שאלות גנומיות.

תודה לגיורא ולאורלי על התמיכה, העזרה והעצות המועילות.

תודה להוריי היקרים על התמיכה ולענת אשתי האהובה שתמיד עומדת מאחורי ומחזקת  
אותי בכל מעשיי.

# פיתוח סמנים גנטיים לייסוד תכנית טיפוח לדג דקר המכמורת (*Epinephelus aeneus*)

עבודת-גמר

מוגשת לפקולטה לחקלאות, מזון וסביבה על שם רוברט ה. סמית  
של האוניברסיטה העברית בירושלים לשם קבלת תואר 'מוסמך  
למדעי החקלאות'

על ידי

ליאור דור

נובמבר 2012

רחובות

חשוון תשע"ג