# Università degli Studi di Padova 

## Laurea Magistrale in Biologia Marina

"Microsatellite genetic analysis of European sea bass (Dicentrarchus labrax) samples from an area subject to a restocking program in the Venice Lagoon"

Relatore: Dott. Lorenzo Zane<br>Dipartimento di Biologia<br>Correlatore: Prof. Luca Bargelloni<br>Dipartimento di Biomedicina Comparata e Alimentazione


#### Abstract

RIASSUNTO Il depauperamento degli stock ittici naturali è uno dei problemi più preoccupanti che l'ecologia sta affrontando negli ultimi anni. In risposta a tale problema, la regolamentazione della pressione di pesca si sta rivelando un approccio efficace. Per le popolazioni ittiche già gravemente colpite dall'overfishing, dei programmi di ripopolamento opportuni possono aiutare a reinstaurare un equilibrio tra prelievo e crescita che permetterebbe alle specie interessate di recuperare la stabilità demografica originale. Trattandosi di azioni delicate, le cui conseguenze possono essere deleterie per la specie e l'ambiente soggetto al restocking, la loro programmazione e esecuzione devono seguire alcune regole che aumentino la probabilità che l'azione vada a buon fine. I punti su cui si sono soffermate maggiormente le ricerche più recenti sono l'origine degli animali liberati (e in particolare la distanza genetica dalla popolazione selvatica target) e l'organizzazione genetica dei gruppi di pesci utilizzati per il ripopolamento (ad esempio in termini di variabilità genetica, eterozigosità...). Lo studio presentato in questa tesi consiste nell'analisi, basata su 9 marcatori microsatellite, di campioni di branzino (Dicentrarchus labrax) provenienti dalla Laguna di Venezia. L'area di campionamento è stata soggetta, alcuni mesi prima dei campionamenti stessi, al rilascio di un gruppo di 30,000 branzini da parte dell' 'Associazione Pescatori della Laguna di Venezia', che lamentava la povertà della zona in animali di questa specie. I giovanili utilizzati per il ripopolamento provenivano dall'allevamento 'Ca' Zuliani' di Pila di Porto Tolle. Due dei campioni analizzati provenivano dall'area della bocca di porto di Malamocco interessata dal ripopolamento, mentre il terzo, commissionato ai membri dell'associazione, è costituito da branzini provenienti da un'area più vasta della laguna e non meglio definita. Per le analisi sono stati utilizzati anche i dati di alcuni campioni della popolazione selvatica della Laguna di Venezia, dei riproduttori utilizzati per generare gli animali liberati e, infine, dati della popolazione selvatica di provenienza di questi riproduttori (area del Delta del Po). Le analisi effettuate comprendono la ricerca di similarità tra i gruppi attraverso l'analisi con il software Structure; un'analisi di parentela tra gli individui dei gruppi campionati e i riproduttori dell allevamento Ca’ Zuliani (per verificare la presenza di individui rilasciati tra i campionati); l’assegnazione, basata sui 9 loci microsatellite, dei campioni alle possibili poplazioni d’origine (Laguna, nel caso di animali selvatici o Delta del Po, nel caso di animali rilasciati). I risultati ottenuti con i tre approcci elencati hanno evidenziato un impatto molto rilevante degli animali liberati, in termini di presenza nell'area soggetta al ripopolamento. Inoltre, grazie alla possibilità di confronto con i data set di Laguna e Delta del Po, è stata riconosciuta una buona somiglianza genetica tra la popolazione selvatica da ripopolare e la popolazione di origine degli animali


rilasciati. Questo ultimo risultato è un indicatore positivo per quanto riguarda l'azione di restocking, che dovrebbe essere effettuata utilizzando come fonte degli animali liberati una popolazione il più possibile vicina a quella salvaguardare. Tuttavia, dai risultati è emerso anche che le varie fasi del programma non hanno seguito le regole suggerite dagli studi più recenti. Infatti l'analisi della struttura genetica dei gruppi di branzini utilizzati ha evidenziato una scarsa variabilità allelica. La causa, desunta dai risultati ottenuti, è il ridotto numero di riproduttori che hanno contribuito a generare i branzini rilasciati. Nell'allevamento di provenienza, i branzini adulti vengono fatti riprodurre attraverso eventi di "mass spawning". L'impossibilità di controllare, in queste condizioni, quali individui effettivamente si incrocino rende difficile ottenere, da un singolo evento riproduttivo, un gruppo geneticamente vario di giovanili, come auspicabile nel caso di animali da liberare per un ripopolamento. I risultati di questo studio sottolineano l'importanza di un approccio scientifico agli interventi umani che possono influire sulla stabilità degli ecosistemi per evitare che azioni che dovrebbero migliorare lo stato dei sistemi naturali impattati dall'uomo siano inefficaci o addirittura dannose.

## INDEX

INTRODUCTION ..... 1
Population analysis using molecular tools ..... 4
Microsatellite markers ..... 6
AIMS ..... 7
STUDY SPECIES: Dicentrarchus labrax ..... 9
Population structure ..... 10
Aquaculture aspects ..... 11
MATERIALS AND METHODS ..... 13
Study area ..... 14
Already available data ..... 14
Sampling methods ..... 15
Individuals Released ..... 15
Samples conservation ..... 16
DNA Extraction ..... 16
Target regions amplification (PCR) ..... 17
Agarose gel electrophoresis ..... 20
Genotyping ..... 21
Data visualization ..... 21
DATA ANALYSIS ..... 25
RESULTS AND DATA DISCUSSION ..... 29
CONCLUSIONS ..... 41
BIBLIOGRAPHY ..... 47
APPENDIX ..... 53

## INTRODUCTION

According to U.S. Census Bureau, at the actual growth rate world population will reach 8.3 billion people in 2025. Together with a necessary reduction of procapite consumption in the richest countries, a massive increase in food production and harvesting will probably be needed. Fish production and fisheries make no exceptions. World fisheries provide more than $15 \%$ of the average animal protein intake for 2.9 billion people (FAO 2009). Aquaculture is growing rapidly, also as a consequence of the wild stock depletion due to overexploitation during the last century (Nomura, 2008; Subasinghe et al., 2009) that, together with habitat destruction, is one of the biggest problem affecting global fisheries and marine ecosystems (Jackson et al., 2001). Nowadays cultured animals account for more than $40 \%$ of the total sea food ( $30 \%$ species). In the literature there are indications that aquaculture might be a solution to the collapse of fisheries stock worldwide. For some species, such as herbivorous or filter feeders, net contribution to global fish supply is great and the impact on the ecosystems is low. However, the production of some other species that are widely cultured request technology or systems that damage the ocean and coastal environment through habitat destruction. This is the case, for example, of salmon and shrimps (Naylor et al., 2000).
Another collateral and less obvious consequence of aquaculture practicing is related to the escapes of cultured individuals. When this event involve species that are new to the environment they come into, these species could become invasive thus the threat to other native species can be very high (Bax et al., 2003). Up to know a lot of cases of alien species invasions are known, as known are their consequences, including loss in biodiversity as well as radical changes in the entire trophic chain (Galil, 2007). Nevertheless, even in the case of escape of a species already present in the natural environment, the consequences can be deleterious. In this case negative effects can be due to disease and parasites carried by the escaped animals (Naylor et al., 2001; Bartley et al., 2006). Moreover, critical issues depend also on genetic parameters of the released batch: origin of the reared animals and their genetic variability, among others. Similar effects on local stocks can be found analyzing the changes to the wild population derived by restocking practices, considering that in this latter case the number of farmed individuals can be larger (except, for example, when a net of a cage breaks and a massive escape happens, comparable with a restocking release).
Together with aquaculture, which can supply fish food that fisheries will probably not be able to supply, appropriate management actions are in fact required to preserve fish stocks from overfishing or, if too late, to restock depleted areas
with animals coming from the wild or with cultured individuals. The same action, in addition to restocking purposes (considered as the restoring of spawning biomass of overexploited fisheries through introduction of juveniles of the same species) can be performed for enhancement, which would permit larger catches or for creating new fisheries (Ward, 2006). These practices have been widely used for salmonid fisheries in the USA and for at least other 90 species of fish and invertebrates in Japan to augment wild stocks (Honma, 1993; Imamura, 1999). Whether this is good or not has been debated for a long and, despite the prevalence of studies indicating negative effects, there is not a unique answer to the question yet, as some positive effects have been underlined either (Araki and Schmid, 2010). Among the most important parameters that should be carefully monitored we can mention the immediate effects due to the increase in animals available for harvesting and the effects on general fitness and genetic diversity of the restocked population. These two features can provide useful information about short term and long term consequences of the restocking, respectively. Combined, these are the variables that make a durable and effective stock enhancement happen or fail.
In the short-term period, release of a considerable quantity of juveniles can provide large net benefits to fisheries. This is mainly due to the availability of the released animals, rather than to a real re-enhancement of the depleted wild stock, which is expected to be stable and durable (Hilborn and Eggers, 2000). To say that a restocking action has been successful, its positive effects should be visible in the long term. In other words the positive effects are not related just to the number of individuals released (e.g. 100 individuals released, then 100 fished), but also to the way these individuals behave in the environment (e.g. they can acclimatize and then reproduce, so that from 100 released individuals it is possible to obtain many more of them). In more technical words we can say that an increase in the census size (or the actual number of individuals in a population) is not necessarily related to a general increase of the target population fitness, which can be described by other indicators such as an increase of genetic variability, evaluated, for example, from the effective population size ( Ne ) of the target population. Ne , loosely speaking, indicates the size of the group of breeding individuals in the given population. Ne is one of the parameters that are fundamental in population studies as it can provide useful information about genetic variation due to drift, inbreeding or recent bottlenecks. Ne, first introduced by Wright in 1931, has proven useful in the design and analysis of artificial breeding program as well as in understanding evolution of natural populations (Lande and Barrowclough, 1987).

Wild stock re-enhancement, which is the main goal of a restocking program, is an action that permits a depleted wild stock to have a number of reproductive individuals (and a genetic variability) that can support the loss due to fishing in the long term. For this reason the effects of a restocking program should be monitored beyond the simple assessment of stock dimension or immediate stock growth (Bell et al., 2005). In fact, the wild population can be damaged by the introduced individuals. For example, an actual reduction of the number of wild individuals could result by competition with released animals, which produce density dependent mortality, or introgression with hatchery genetic background that can disrupt allelic combination relevant for local adaptation. When designing a restocking program these factors should be seriously taken into account before proceeding with release, so that the right source population is chosen, as to avoid negative effects of interbreeding between wild and hatchery stock when this is thought to be deleterious. Among the other possible problems related to a restocking program not properly designed, we can find lower survival, growth rate, and reproductive fitness of the released animals and possibly of the offspring generated by crosses between released and wild individuals.
For these reasons, cultured fishes used for restocking should be very carefully chosen since proper broodstock selection and breeding plans that minimize negative effects on wild stocks. Natural population and reared fishes should be examined before and after the release of juveniles. The former can provide useful information to the choice of the population that is going to provide the juveniles; the latter permits to monitor the effects of the release in terms of efficacy and long-term consequences. Ideally broodstock should be provided by the population to be enhanced or if not possible, the most similar available (Ward, 2006). Nevertheless, some cases have been reported where even the use of local wild broodstock led to a loss in reproductive fitness in the hatchery fish as well as in the wild born, descendant of hatchery-born parents (Araki et al., 2007b, 2009).
The causes of these negative effects can be understood from the analysis of the chain of events that happen before the release. For example broodstock selection for hatcheries and hatchery practices can lead to low effective population size and thus to a loss of genetic variability inside the breeders pool. Inbreeding may occur and, as a consequence, the phenotypic expression of recessive alleles due to increased homozygosity. Attention should be paid to maintain large dimension of broodstock and to reproduce as many breeders as possible. In addition, renewal of breeders in the broodstock population should aim at increasing the genetic diversity of the captive population.

Once released in nature, hatchery stocks can affect the gene pool of the natural population by increasing homogeneity inside the population or even by completely replacing native strain with the introduced one, as reported for brown trout in northern and central Italy (Marzano et al., 2003). This is expected to happen especially when the target wild population is particularly depleted, so that even the introduction of a small number of new animals can make the genetic arrangement of the local population change. Another aspect that should be taken into account is the possibility for hybridization between wild and released stocks. Small effects of hybrid vigor (major fitness for offspring spawned from unrelated parental) has been found for first generation of salmon when distinct strains hybridized (Bryden et al, 2004). But, from the second generation and later, the break-down of co-adapted gene complexes, together with disruption of interaction between the genetic set and the local environment can results in a loss of fitness, that is expected to increase in severity with increasing genetic distances between the parental strains (Edmans and Deimler, 2004).
If the so called 'supportive breeding' (where parents from the population to be restocked constitute the artificial broodstock) can be a solution for this problem, even this procedure requires attention. To maintain high variability (i.e. high Ne ) the number of breeders used for the production of juvenile should be high, especially when the stocked juveniles possibly makeup a large proportion of all offspring. Moreover, it has been reported that a well-planned supportive breeding program can increase Ne , in species with high reproductive ability and high mortality (Wang and Ryman, 2001).

## Population analysis using molecular tools

The use of molecular tools in traceability was one of the earliest recognized applications of non human DNA forensics. The ability to generate genetic data is increasing very quickly, so that many of the barriers that geneticists traditionally face are already being removed. Moreover, progresses in genotyping technology are making genetic monitoring increasingly cost effective (Ogden, 2008). The first application of molecular tools in the fisheries sector was directed toward species identification, in order to discover commercial frauds. Examples of the applications are the investigation on illegally traded caviar (De Salle and Birstein, 1996) and shark fins (Abercrombie et al., 2005)

Another very interesting issue that can be solved using genetic markers is the understanding of the genetic structure of a species in terms of populations and geographical organization. Populations' dynamics, both spatial and temporal, are fundamental information to validate models in fish ecology. A direct consequence of knowing populations boundaries is the possibility to develop
tools to identify the geographic origin of samples. From a forensic genetics perspective, identify the geographical origin of a sample is equivalent to identify its origin population.
Another powerful tool for answering ecological and evolutionary questions is the reconstruction of parental structure. Knowing this structure can be very useful to get information about the population and not only about single individuals. For example genetic variability and Ne are related to the number of animals that reproduce and this number can be calculated on the base of pedigrees analysis. Although being one of the simplest concepts in biology, a pedigree, that is the genealogical relationship between individuals in a population, has become possible to reconstruct only after the development and application of molecular markers. Pedigrees are based on the simple concept that each parent passes one allele per locus to the offspring, that therefore carries one allele from each parent. Anyhow, behind the simplicity of the theory, the practical implementation of parentage assignment is full of difficulties (Hauser et al, 2011). When all parents can be sampled and when using many variable loci (high exclusion power), parents can be assigned by excluding non parents on the base of Mendelian inheritance rules. Nevertheless, much more often geneticists have to deal with open systems (not all parents sampled) and with limited available data or genotyping errors. If this latter is the case exclusion method may fail to assign parents (genotyping errors) or may assign false parents (lack of marker assignment power), and is therefore advisable the use of other assignment methods, such as likelihood, thus determining probabilities of parentage assignment from simulations. Freely available softwares typically offer this statistical approach with different treatment of genotyping errors, estimation of likelihood and many other factors (Hauser et al, 2011).
Among the others, in the last decade microsatellites have been the marker of choice for most parentage and other assignment studies because of their high variability and wide availability. Recently, the interest in Single Nucleotide Polymorphisms (SNPs) is increasing and it is possible that these will be the marker of choice in future assignment studies. This latter marker consists in polymorphisms that involve just one base inside a genome region, that can vary from an organism to the other. In any case, microsatellites are still the most used markers, thanks to its presence among various species genomes and because it's low price and easy to analyze.

The markers of choice for this work have been the microsatellites. Also known as SSRs (simple sequence repeats) or STRs (short tandem repeat), these are DNA regions composed by short sequence of 2 to 6 nucleotides repeated several times. Presence of SSRs has been found both in eukaryotic and prokaryotic genomes (Field and Wills, 1996; Toth et al, 2000). The feature that makes these regions usable as molecular markers is their mutation rate, which is higher than the mutation rate of other genome regions. Indeed its value ranges from about $10^{-3}$ to $10^{-6}$ mutations/locus/generation and the average rate for fish species is widely recognized to be around $5 \times 10^{-4}$ mutations/locus/generation (Lippè et al., 2006). Explanations for this level of mutation rate have been suggested: recombination errors; unequal crossing-over; polymerase slippage during replication or repairing. Together with high variability in repeated regions, the low variability of the sequences that come before and after SSRs (about $10^{-9}-10^{-10}$ mutations/locus/generation) makes these segments easy to amplify using PCR procedure, allowing the development of reliable primers (Hancock, 1999). Presence and density of microsatellites have been demonstrated to be related with non-coding DNA regions, possibly due to the negative effects that SSRs may have on proteins (with an exception for trinucleotide repeats and exanucleotide repeats) (Toth et al., 2000). An interesting feature that characterizes microsatellites technology development is the so called transferability or crossspecies amplification: due to the similarity of the regions surrounding the SSRs, the same pair of primers could be successfully used with species similar to the one the primers were first created for. Nevertheless, the rate of successful amplification gets lower with increasing genetic distance between species (Primmer and Merilä, 2002). Especially for marine organisms, which are characterized by low levels of genetic differentiations when compared to terrestrial organisms (Cooke and Cowx, 2004), microsatellites allowed to identify even small populations differentiation that could not be appreciated using just allozyme or mitochondrial DNA.

The work presented in this thesis consists in the population genetic analysis, based on 9 microsatellites markers, of two European sea bass (Dicentrarchus labrax) samples ( $n=164 ; n=113$ ) collected at the outer side of the Malamocco inlet dams (central Venice Lagoon) during the spring-summer of 2008. A third group of fishes ( $\mathrm{n}=68$ ) was collected in a wider area of the lagoon by local fishermen during their normal fishing sessions. The previous autumn (November 2007) the 'Venice Lagoon Fishermen Association' released 30,000 D. Iabrax juveniles from an enclosed "valle" located near the small village S. Pietro in Volta (see Materials and Methods). The releasing and sampling sites are about one kilometer far from each other. Juveniles released came from a hatchery (Valle Ca' Zuliani located in Pila di Porto Tolle, on the Po river delta) that provided the batch of young sea basses coming from a mass-spawning reproduction event. Other genetic data available, included in the analyses presented in this thesis were (I) the population of origin from which the breeders were collected (II) the actual broodstock that generated the juveniles (i.e. the breeders' genotypes) and (III) a samples from the Venice Lagoon, considered as an example of the wild population in which the reared animals were going to be released. These data were available thanks to previous studies on the same species undertaken in the same department where I attended my graduation training.
For all the data included in the thesis, several analyses have been performed. First of all several basic genetic parameters were calculated to get general information about the populations and groups that are being analyzed. To do this the 'GenAIEx' add-in for Excel was used.
The genetic pattern of the entire data set was then studied, to detect any possible differentiation among the samples that could help in understanding the effects of the introduction. These analyses were performed by the software 'Structure'.
A further analysis performed was the assignment of the animals sampled along the inlet dams to a set of breeders of the hatchery that provided the juveniles for the restocking. This approach is useful to understand the number of animals that has been used to generate the offspring to be released and thus understand the level of variability of the introduced batch. The software used for parentage assignment was 'Colony 2.0'.

Finally the samples were tested with the population assignment. This approach is the only available when genotypes information of the breeders are not available. Here it was used to test its power under the present conditions, and to compare its performance with other assignment methods. 'GeneClass' was the population assignment software chosen.

Comparisons between the various approaches were made to obtain the most accurate results possible. The effects of the restocking action was then discussed in light of what previously mentioned on positive and negative effects of these management actions (such as genetic distance of the broodstock that generated juveniles and the wild population, variability of the released animals, Ne of the samples); some insights on the possible effects of the restocking program undertaken are included.

## THE STUDY SPECIES: Dicentrarchus labrax

The European sea bass, Dicentrarchus labrax (Moronidae, Perciformes) is a highly valuable commercial species broadly distributed in the coastal waters from Norway to Morocco in the eastern Atlantic as well as in the Mediterranean and Black sea (FISHBASE.com), where it inhabits waters up to 100 meters depth, being more common in shallow water (Lloris, 2002). This is a euryhaline and eurythermic species and can bear temperature ranging from $5^{\circ}$ to $28^{\circ}$. Saline thresholds go from $3 \%$ to full strength sea water, allowing this
 species to live in estuaries and brackish water lagoon as well as venture upstream into freshwater. Its behavior changes over age as younger fishes are often found schooling whereas older individuals are solitary. Seasonal migrations are believed to occur starting from when an individual reaches maturity. From that moment sea basses usually over-winter in marine waters, instead of remaining inside lagoons as they do before becoming adults (Pawson and Pickett, 1987). The change to a migrating behavior is related to the fact that osmoregulation and sexual maturation are incompatible (Zanuy and Carrillo, 1984). The return to shallower and warmer water occurs before the reproduction period. The breeding season differs between the Mediterranean and the Atlantic, taking place from December to March and up to June, respectively. Sea bass spawn near to river mouths and estuaries where water salinity is between 30 and 35 (Lloris, 2002). Fecundity is on average 200000 eggs/kg female, with first reproduction event over 2 kg weight and a fecundity that can reach 6 to 7 years in the wild (Froese and Pauly, 2006).


Fig. 1 Computer Generated Native Distribution Map of Dicentrarchus labrax (FISHBASE)
The annual fisheries production for this species has been stable around 10,000 tons/year since 2001 (see Fig. 1) (FAO). In the Mediterranean the main fisheries
are from France and Italy, and together they account for $75 \%$ of the total catch of the area. The most used fishing techniques are bottom trawls, beach seines and hooks and lines. Even if nowadays production from aquaculture is much higher than production from fisheries, wild sea bass maintain an importance on the market, being their price much higher than that of the farmed fish.

In addition to commercial fisheries it
 is important to notice that sea basses are highly important for sport fishermen, both anglers and "spear fishers". For this reason in areas where the recreational fishing effort is high the species could be subject of a strong size-dependent mortality, typical consequence of this sport on many fish species. From a management point of view, the different needs of sport and commercial fishermen (i.e. individual size vs. stock size) imply different actions in order to maintain the population structure as well as the population size (Garcia-Asorey et al., 2011). In the present study restocking was carried out by angling fishermen to overcome a deficiency of sea bass in the Lagoon area. It's thus important to consider that this is the category whose needs required to be matched.

## Population structure

Since European sea bass is exploited by fisheries and widely cultured, the importance of genetic analysis in this species is increasing. Nowadays, the knowledge of the species in the wild is mainly based on genetic studies and also the development of farming technique is getting useful information from these studies. Analysis of wild populations has become more and more easy and affordable with the availability of genetic markers and the large amount of available data collected from all over the distribution area of the species
Three main D. labrax populations have been identified by many studies based on allozymes (Allegrucci et al., 1997; Castilho and McAndrew, 1998), mitochondrial DNA (Patarnello et al., 1993) or microsatellites (Bahri-Sfar et al., 2000; Castilho and Ciftci, 2005): north eastern Atlantic Ocean, western and eastern Mediterranean. Moreover, eastern Mediterranean population has shown differentiations at a smaller geographic scale, with a structure consistent with the existing basins. In particular differentiation has been found between Adriatic, Ionian and Aegean seas, the Libico- Tunisian gulf and the Levantine basin (Katsares et al., 2005).
Beyond this subdivision, it is important to notice that some studies reported that it is possible to find samples genetically distinct from the population of the area
they live (Katsares et al., 2005). Events like this probably began when aquaculture first introduced eggs and larval exchange between different farms, sometimes located far from each other. Escapes of individuals led to interactions between different genetic backgrounds. For example, analysis based on microsatellites data (population analysis) revealed that two samples from Greece and one from Egypt did not cluster according to their expected geographic origin (Katsares et al., 2005). This way, new "populations" might origin from tank or cage escapes, as suggested by a study of a wild population from the gulf of Tunis (Haffray et al., 2007). Lower allelic diversity was observed, probably due to a low Ne , meaning that this population originated from a limited number of parents probably escaped from a local farm.

## Aquaculture aspects

Since the late 60s for about a decade a quick development of the production techniques for the juveniles has taken place, mainly in Italy and France. Before this period sea basses were provided by two systems: from fisheries that harvested animals in the wild; from the several coastal lagoons were also salt was harvested. In this latter case the supply of juveniles that grew in the lagoon came from trapping schools of fish during their migrations. The method is still in use today but is less frequent. European sea bass was the first marine nonsalmonid species commercially cultured and nowadays it is one of the most important commercial fish produced in Mediterranean areas, with Greece, Turkey, Italy, Spain, Croatia and Egypt as the main producers. Production has been growing rapidly since early 90s, reaching over 60,000 tons in 2010 (FAO).


Fig. 2 Left graph representing European Sea Bass fisheries harvests (in thousand tons) from 1950 to 2010; right graph showing European Sea Bass aquaculture production (in thousand tons) for the same period (FAO)

Nowadays the greatest part of global production comes from intensive systems, which, thanks to controlled diet, obtain commercial size fishes (about 400-500 grams for the sea bass) in just 18-24 months, against the 35 months required for extensive and semi-intensive systems. To do this, juveniles are provided by
specialized farmers and kept in tanks or sea cages where food is provided regularly. Breeders for hatcheries production come from either wild stocks or other farms and are kept long-term and selected by farmers. Optimal parents' age is between 5 and 8 for females and 3 to 5 for males. Control of most of the reproduction phases (including induction of ovulation by photoperiod manipulation or hormonal treatment, fertilization in tanks and incubation in an open water circulation system) is required to secure a reliable and sufficient quality of fish eggs and thus juveniles.
Floating sea cages (net pens) are often used for the fattening period as they are for some aspects easier to manage, not requiring particular onshore structures. They can be located close to the land or in the open sea and in this case water exchange is always granted and the quality depends on the site where the netpens are. Gear care is essential. Especially in the hot season frequent net changing, cleaning and removal of moribund or dead organisms are requested to maintain optimal growing conditions. Moreover the constant care of net integrity would reduce the risk of escapes that, in addition of being a loss for the farmers, can lead to several ecological issues, as mentioned in the introduction. If tanks are the choice, then accurate control of the water quality and fish health are fundamental. To control salinity and temperature brackish or sea water can be pumped at occurrence. Temperature is also an efficient way to control growth in hatcheries and production phase. Since stocking densities is usually up to 25$30 \mathrm{~kg} / \mathrm{m}^{3}$ and high water quality is required these techniques are often expensive in terms of filtering, air pumping, general treatment and catabolites removal (FAO).
Improving the harvesting technique by selective breeding has been one of the major aspects of modern harvesting. This could results in loss of heterozygosity (Sola et al., 1998) and alleles that would lower the genetic variability in the reared population. These changes can become a major issue in the case the offspring is used for restocking or in case of escapes of fishes from the farm. The introduction of these animals can be more deleterious for the wild stock if the released/escaped batch had been selected for production (e.g. growth rate, feeding, resistance to antibiotics...) rather than for living in the wild

## MATERIALS AND METHODS

During my work I genotyped 9 microsatellites loci from samples of three European sea bass samples collected near one of the Venice Lagoon inlets, where a release of juveniles of $D$. labrax for restocking purpose had taken place. Then, I analyzed these three samples and three additional ones coming from the North Adriatic Sea area, used as reference. Re-editing of all the 6 groups' genotypes was required in order to get comparable data across data sets. Structure analysis, parentage and population assignments were tested. The following table summarizes used sample sets.

Tab 1. List of the samples analyzed: populations, codes, date of samplings, number of samples

| POPULATIONS | CODE | Date of sampling | $\mathrm{N}^{\circ}$ samples |
| :---: | :---: | :---: | :---: |
| Porto Tolle | DeltaPo | $1998 / 2001 / 2003$ | 161 |
| Venice Lagoon | Laguna | 2005 | 172 |
| $1^{\text {st }}$ sample | VSP1 | spring 2008 | 164 |
| $2^{\text {nd }}$ sample | VSP2 | summer 2008 | 113 |
| $3^{\text {rd }}$ sample | VSP3 | spring-summer 2008 | 68 |
| Breeders from Ca' Zuliani farm | RIPR | 2008 | 178 |

The name 'VSP' for the three samples collected by the fishermen came from 'Valle San Pietro', that is the enclosed water area were hatchery-produced juveniles were grown before being released.

The samples used in this work came from the Northern Adriatic and precisely from the Venice Lagoon and from the area around the Po river delta. In particular the introduction of juveniles for restocking was made in a delimited water area near the Malamocco inlet. This is one of three links the lagoon basin has with open sea and is located between the northern one ('Lido inlet') and the southern one ('Chioggia inlet'). These canals generated from an interruption of the earth line dividing sea from lagoon and are characterized by two parallel artificial rocky dams extending from the earth line toward the open sea, about 500-1000 meters long. The complex constitutes a particular environment since rocky coast traits are new to the interested area that is mainly characterized by sand basins and coasts. Obviously this is a very peculiar area, characterized by many disturbing factors such as intense maritime traffic, water current and in the last years by important works on the basin of the channel as well as on the rocky dams due to the 'Mose' project's dockyard.


Fig. 3 Satellite view of the study area: on the right a wide view of the Venice Lagoon; on the left a zoom on the growth and release site near the small town S. Pietro in Volta (1) and on the sampling sites for the groups VSP1 and VSP2 (2)

## Already available data

The previous work on breeders selection for farm stock improvement, in addition to providing an affordable set of loci for my analysis, made available a large number of already genotyped samples from the Po river delta. From this population the breeders for the farm were collected so it can be regarded as the origin population of the juveniles used for restocking. Moreover, other projects had been carried out to study Mediterranean sea bass populations before I started mine. Thank to these some additional data were available for a wild population sample collected in the Venice Lagoon's from many different sites
inside the Lagoon. These data came as genetic profiles, thus editing of alleles sizes (using Genotyper 3.0) had to be done in order to have values comparable with those of the groups I genotyped.

## Sampling methods

Two fishing techniques were used by the fishermen that provided the samples to collect the sea basses along the inlet dams ('VSP1' and 'VSP2'): the 'Barracuda nets' and the long lines. 'Barracuda' is a set gill net particularly used for fishing small size species in sea or rivers. It is made by thin nylon line 0.20 mm diameter) that form a panel kept vertical by a float line and a weighted ground line. The width of the panel is 1.5 meters while the length is 20 meters. Mesh size is $30 \times 30 \mathrm{~mm}$. The long line is made by a main line with hooks attached at interval through branch lines. Long lines can be used to fish near the bottom as well as at the surface and can either be set by an anchor, left to drift or get dragged. In this case the lines got dragged since the baits were artificial ones simulating small fishes. 'VSP3' samples were collected by local fishermen that used their usual technique for fishing (mainly angling).

## Individuals Released

The batch of juvenile European sea bass came from a local farm (Valle Ca' Zuliani located in Pila di Porto Tolle) and was produced through a "mass spawning" event. This means that animals were let free to reproduce (i.e. there was no egg and sperm collection and no artificial crossing). The breeders list provided by the farmers indicated that the number of fishes kept in each tank ranged from about 20 to about 40, with a female/male ratio that went from 2:1 to 4:1. One of the consequences of a natural mass spawn is that the individuals that actually reproduce depend on the period since, as normal in farming condition, breeders are not ready for reproduction all at the same time. It has been reported, for the Mediterranean sea bass, that a batch from a single spawning event was made up by $95 \%$ of juveniles sharing the same mother. Moreover, $50 \%$ of them shared also the same father (Chatziplis et al., 2007). As a consequence, when a batch is collected to be used for restocking, there is the risk that the greatest part of the juveniles come from a small number of breeders (despite the presence of many more females and males in the tank), thus affecting the restocking success and consequences.

## Samples conservation

Small pieces of each individual (about 200 mg ) from either fins or other body parts were cut and immediately put in 2 ml tube filled with $80 \%$ ethanol. Tubes were then stored inside $10 \times 10$ rack at a temperature of $4^{\circ} \mathrm{C}$ at the 'Department of Comparative Biomedicine and Food Science' (UNIPD) in Legnaro (PD), until DNA extraction. If properly stored samples can last long at this temperature. In the present study DNA extraction was performed three years after collection but no signs of major degradation were detected on tissue pieces. Occasionally, a thin layer of undefined mold on the outside of the tubes was present. In those cases I proceeded by cleaning them with sanitizer solution before opening and taking off the sample pieces. Rarely, a white mucus-like substance around the sample piece was found but no evidences for extraction or amplification problem related to this were found.

All procedures described in the following (extraction, amplification, genotyping and data analysis) have been undertaken at laboratories of the 'Department of Comparative Biomedicine and Food Science' (UNIPD) in Legnaro (PD).

## DNA extraction

The first step for any genetic analysis is the extraction of genetic material from the samples. This means that cellular and nuclear membranes have to be disrupted and eliminated. Among the many methods for DNA extraction, in my case the Chelex resin protocol was used (Promega). Smaller parts of the sample (about 10-15 mg) were cut from the pieces stored in the tubes using clean instruments (between each sample tools were washed with distilled water and ethanol $90 \%$ ) and put into a 96 -wells plate. A water solution containing $5 \%$ water volume Chelex 100 resin and $0.07 \mu \mathrm{~g} / \mu \mathrm{l}$ K proteinase was prepared and each tube was filled with $100-150 \mu$ l of this solution. The K proteinase helps digesting cellular proteins and, more important, nuclease that would otherwise digest the nuclear material. After placing a plastic layer on the plate, to avoid evaporation of solution (problem that anyway occurs sometimes, mainly in the more lateral wells, were layer does not stick properly), digestion proceeds at $55^{\circ} \mathrm{C}$ for about 1 hour. Digestion phase is followed by inactivation of proteinase, necessary to avoid this enzyme to be functional during the amplification phases and inactivate PCR enzymes. Samples are then incubated at $95^{\circ} \mathrm{C}$ for 10 minutes. A light centrifugation follows (2000 rpm for 2 minutes) to permit sedimentation of Chelex in the bottom of the wells and get clear surnatant containing target DNA. If not used immediately, plates were stored at $-20^{\circ} \mathrm{C}$.

## Target regions amplification (PCR)

Amplification of target DNA regions is fundamental to achieve good results when analyzing genetic material. In order to do this the Polymerase Chain Reaction (PCR) is used and, thanks to the modern thermo-cyclers, this procedure is nowadays easy and steady. When approaching to microsatellites analysis for both forensic or management studies, a fundamental step is the identification of an informative and affordable set of loci. Choice of loci has not been necessary in my case, thanks to previous work carried out in the same laboratory and on the same species, which included the selection and application of 9 microsatellite loci to be used for parentage analysis, as part of the process of selection of new breeders for broodstock improvement. The farm was the same that provided the juveniles released for this program. Loci were selected on the base of their high variability (mean number of alleles) and on technical specifics (e.g. dimensions) that allowed a correct and easy development and use of the markers. Among 100 loci available in literature the 9 chosen can be divided in three classes, based on the length of the PCR products ( 3 "small", 3 "medium" and 3 "large"). For each group, the forward primer used in the PCR protocol was marked with either FAM (blue), HEX (green) or TAMRA (yellow) dye. As a result what we got is the following set of loci:

Tab 2. List of loci used divided according to the molecular weight: locus name, repeated sequence, primer sequence, DYE used, annealing temperature ( $T_{a}$ ), size range, number of alleles for each locus, linkage group (LG) (Chistiakov et al., 2005) and multiplex amplification group (Amp group).

LARGE SIZE

| Locus | Repeat | Primer sequence | Dye | $\mathbf{T}_{\mathbf{a}}$ | Size <br> range | No of <br> alleles | LG | Amp <br> group |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| DLA008 | $(\mathrm{AC})_{24}$ | F:AAGCTATCTGATCTCGCTTG <br> R:ACGTGATTAAGTGTTTGTGAG |  | 56 | $214-318$ | 51 | 24 | 4 -plex |
| DLA119 | $(\mathrm{TG})_{10}$ | F:GCAGGTTCAAATTATTTTTGCTC <br> R:TCCTCCTTTGGTTGCTAGG |  | 54 | $221-265$ | 23 | 14 | 4 -plex |
| DLA016 | $(\mathrm{TG})_{24}$ | F:GTGACCGCAGATGAAGAAC <br> R:ACTGTGGGCTCATAAACATC |  | 54 | $220-272$ | 27 | 1 | 4 -plex |

MEDIUM SIZE

| Locus | Repeat | Primer sequence | Dye | $\mathbf{T}_{\mathbf{a}}$ | Size <br> range | No of <br> alleles | LG | Amp <br> group |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| DLA020 | $(\mathrm{TG})_{20}$ | F:GTCTAATGAGCAGTGGAGCAG <br> R:GCATGTTAGATCCACCTCTTTC |  | 56 | $144-180$ | 16 | 12 | 5 -plex |
| DLA105 | $(\mathrm{AC})_{16}$ | F:GAGGCTGTATGCTGTTGCAG <br> R:ACCCATGCATAAGGTCAGTG |  | 56 | $135-181$ | 20 | 8 | 5 -plex |
| DLA145 | $(T C)_{20}$ | F:CCCACAATAGATTCAAATAG <br> R:CACACATGCAATTATACTG |  | 54 | $153-195$ | 20 | 17 | 4 -plex |

SMALL SIZE

| Locus | Repeat | Primer sequence | Dye | $\mathbf{T}_{\mathbf{a}}$ | Size <br> range | No of <br> alleles | LG <br> group |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| DLA248 | $(\mathrm{AC})_{24}$ | F:TGCATGATGATGTGTGAGCA <br> R:TGGCAGGCTAAAACCTCAAG |  | 54 | $120-126$ | 4 | $?$ | 5 -plex |
| DLA228 | $(\mathrm{TG})_{10}$ | F:CCAATGTTTTCATCCCCTCA <br> R:TTGCTGCTTGTGAAGTGACC |  | 54 | $72-104$ | 11 | $?$ | 5 -plex |
| DLA244 | $(T G)_{24}$ | F:ACTGAAAGCACAGCCTGGTT <br> R:CCCCCATCCAATACACTCAC |  | 54 | $94-114$ | 9 | $?$ | 5 -plex |

I then followed a multiplex protocol that allowed the simultaneous amplification of 5 ( 5 -PLEX) and 4 (4-PLEX) loci. The 9 loci set was then split as follow:

| 5-PLEX | Dla248 <br> Dla228 <br> Dla244 <br> Dla020 <br> Dla105 | SMALL |
| :--- | :--- | :--- |
| 4-PLEX | Dla145 | MED.-...- |
|  | Dla008 |  |
| Dla119 |  |  |
| Dla016 | LARGE |  |

In the PCR protocol the reagent mix contains the genomic DNA (with the target region); two specific primers for this region (a forward primer, in this case marked with fluorescent dye and a reverse primer) (MWG Biotech); a thermostable polymerase (normally from the bacteria Thermus acquaticus) (Go Taq Promega), capable of resist at high temperature (in some PCR phases temperature reaches $94^{\circ} \mathrm{C}$ ); dNTPs (dATP, dGTP, dTTP, dCTP) that will be added to the new filaments by the polymerase; $\mathrm{MgCl}_{2}$ ions that help the polymerase
activity; a reaction buffer to create an optimal environment, in terms of pH and ionic strength, for the reaction activities. In my case the mix was as follow:

Tab 3. List of reagents used for the amplification mix: reagent name, volume, final concentration.

| Buffer 10X | $2 \mu \mathrm{l}$ | 1 X |
| :---: | :---: | :---: |
| $\mathrm{MgCl}_{2}(25 \mathrm{mM})$ | $0.8 \mu \mathrm{l}$ | 1 mM |
| dNTPs $(25 \mathrm{mM})$ | $0.056 \mu \mathrm{l}$ | $70 \mu \mathrm{M}$ |
|  | $0.3 \mu \mathrm{l}$ (large) <br> Primer mix $(10 \mu \mathrm{M})$ | $0.15 \mu \mathrm{M}$ <br>  <br>  <br> $0.25 \mu \mathrm{l}$ (medium) <br> $0.15 \mu \mathrm{l}$ (small) |
| $0.125 \mu \mathrm{M}$ |  |  |
| $0.075 \mu \mathrm{M}$ |  |  |
| TAQ polymerase $5 \mathrm{U} / \mu \mathrm{l}$ | $0.16 \mu \mathrm{l}$ | $0.04 \mathrm{U} / \mu \mathrm{L}$ |
| $\mathrm{H}_{2} \mathrm{O}$ | up to $18 \mu \mathrm{l}$ |  |
| gDNA | $2 \mu \mathrm{l}$ |  |
| Total | $20 \mu \mathrm{l}$ |  |

A One Advanced thermocycler (Applied Biosystems) was used to carry out the amplification of the samples contained in the 96 -wells plate. Through cyclic temperature variations a defined number of amplification cycles were performed, each comprising these fundamental phases:

Denaturation: break of hydrogen bonds and separation of DNA strains that yield single stranded DNA molecules;

Annealing: temperature is lowered allowing annealing of the primers to the single stranded DNA template;

Extension: DNA polymerase, adding dNTPs to the template, synthesizes a complementary DNA strand.

In my particular cases, to suit the primers and the polymerase used, the PCR machine followed this protocol:

| Initialization step: | 2 minutes a $94^{\circ} \mathrm{C}$ |
| :--- | :--- |
|  | 30 seconds at $52^{\circ} \mathrm{C}$ |
|  | 2 minutes at $72^{\circ} \mathrm{C}$ |
| 33 cycles of: | 30 seconds at $94^{\circ} \mathrm{C}$ (denaturation) |
|  | 30 seconds at $52^{\circ} \mathrm{C}$ (annealing) |
|  | 40 seconds at $72^{\circ} \mathrm{C}$ (extension) |
| Final extension: | 5 minutes a $72^{\circ} \mathrm{C}$ |

## Agarose gel electrophoresis

To check whether the PCR performed correctly agarose gel electrophoresis is employed. The agarose gel run allows to separate electrically charged fragments thanks to the electric field applied at the side of the gel. Run distance provides information about the charge and the dimension of each fragment. Comparison with a ladder allows to know approximately the dimension of the sample fragment. The ladder is a mix of known length molecules that has to be run next to the sample in the gel. Shorter molecules move faster and migrate farther than longer ones because shorter molecules run more easily through the pores of the gel. In my case this last was prepared by adding 2.7 g of agarose powder to 150 mL of TAE (saline buffer) obtaining a $1.8 \%$ agarose gel, good to separate fragments up to 500 bp (Lewis, 2011). TAE 1X buffer is obtained from the dilution of the initial 50X solution: 242 mM of Tris-base, 18.6 mM of EDTA, $5.7 \%(\mathrm{v} / \mathrm{v})$ of acetic acid. $15 \mu \mathrm{~L}$ of Sybr Safe (Invitrogen) are then added to the solution. Sybr Safe is a dye used to make DNA or RNA bands visible in electrophoresis gel. Indeed, it fluoresces under UV light when intercalated between the major grooves of the DNA. In the last years this dye has substitute the EtBr (Ethidium Bromide) dye being as much as sensitive but safer than EtBr (Invitrogen, 2011). For each sample $5 \mu \mathrm{~L}$ of PCR product were put in the gel wells, together with $5 \mu \mathrm{~L}$ of Loading Dye 2X composed by glycerol $30 \%$, Bromophenol blue $0.25 \%$, Xilene cyanol $0.25 \%$, Orange $0.25 \%$. This mix helps the loading action (thanks to the glycerol that increases the mixture weight and allows a more precise charge into the wells) and monitoring the run since the dye are visible under normal light and co-sediment with DNA (meaning they move at the same speed as DNA). The current applied was 120 V for a time of about 20-25 minutes. At the end the visualization was made under UV light through a trans-illuminator.


Fig. 4. Particular of an electrophoresis agarose gel run. In the upper 8 wells the 5 -plex was loaded; in the lower wells the 4-plex was loaded. 100 bp Ladder on the left

## Genotyping

After extraction and correct amplification what we got was, for each sample, a large number of the target DNA region copies, ready for the genome scan. This last procedure, that will provide a much more precise value for the length of the fragments amplified, is basically another electrophoresis run, this time using acrilamide in place of agarose gel. The analysis was performed by an 'ABI PRISM 3100 Genetic Analyzer' (Applied Biosystems), available for the use at the 'Department of Comparative Biomedicine and Food Science'.
For each group of 96 samples a plate was prepared by adding $4-5 \mu \mathrm{~L}$ of 5 -PLEX PCR product and $4-5 \mu \mathrm{~L}$ of 4-PLEX PCR product. Amount of DNA was chosen for each plate on the base of the agarose gel electrophoresis result: the intensity of DNA bands on the gel provides information about the amount of DNA available after amplification. The risk of adding too much PCR product to the mix used in the genetic analyzer is getting a null response from the machine due to a too strong signal from the samples.
Every well was then filled with $180-200 \mu \mathrm{~L}$ BDH water and $7 \mu \mathrm{~L}$ of a mix of HDF (High Deionized Formamide) and ROX 400 ladder, a mix of fragment 35-400 bp long, regularly spaced, that acts as a reference for the fragment analysis. The mix is composed, for a 96 well plate, by $721 \mu \mathrm{~L}$ HDF and $21.2 \mu \mathrm{~L}$ ROX 400.
Before placing the plate in the machine for the fragment analysis, a 3 minutes step at $95^{\circ} \mathrm{C}$ is required to unfold the double helix, followed by 5 minutes in ice to lower the temperature. After denaturation HDF helps keeping the helix separated to permit the analysis to be performed on the single DNA polymer.

After setting the analyzer, the scan proceeded per 16 -samples runs, each lasting about 90 minutes, and results were made available in format file and then read using the software Genotyper 3.0.

## Data visualization

Output of ABI PRISM 3100 machine is a series of '.fsa' files that can be viewed using one of the many software available. In my case I used Genotyper 3.0. Results can be visualized for dye color and, for each color, a profile of about 400 bp long is displayed, containing 3 loci (one in the small, one in the medium and one in the large range). The software includes an automatic editing of the microsatellites' peaks, giving the range and dye color for all loci. In this case, for each locus the highest two peak are selected for a heterozygote and, in case of a single peak, the sample is considered homozygote for that particular locus. This is the fastest way to edit profiles assuming the entire process, from extraction to analysis, has gone without any problem. Unluckily this was not my case as most runs needed a careful manual editing. Many anomalies could be found in the
profiles and, for most of them, the process phase linked to the problem could be detected.
Generally a good profile consists in a almost flat line disrupted by one (homozygote sample) or two (heterozygote sample) peaks, often preceded by the so called stutter peaks, smaller than the main one (that accounts for the actual value of the allele), consisting in segment of the target region lacking one or more bases due to amplifying variability (variability related to the poly-A final portion of the segment, that can include a different number of $A$ bases).


Fig. 5. Output of the sequencer machine as visualized by Genotyper 3.0. Stutter peaks are visible before the actual alleles' values

When something goes wrong with the extraction or the amplification, dirty or contaminated runs can be obtained. In this case signals coming from something other than the target allele can give peaks higher that the allele itself or, when a contamination occur an unusual number of alleles will be displayed (e.g. up to four alleles when a heterozygote sample is contaminated with another heterozygote with different alleles).


Fig. 6. Example of contaminated samples. Three peaks are visible in both examples
Slippage is a genotyping issue that created some problems during my work. Fragment analysis in a machine like 'ABI 3100 ' is a very delicate electrophoresis, accurate enough to provide a value for the fragment length with 1 bp accuracy. Since microsatellites are more often repetition of a couple or triplets of bases, a slippage of one or two bases forward or backward can result in important bias when analyzing the profiles. Moreover Genotyper assigns the peaks a value number with two decimals. These values have of course no sense when speaking about microsatellites and it's necessary to round the number to an integer one. This feature gains importance when values for the allele range around the . 50
point. In the worst cases a very small slippage results in two different alleles' calls, even if the two individuals carry the same. For this reason it has been necessary, in many cases, to take a direct look to the profile in Genotyper, in order to understand which alleles required manual editing. There are many other sources of bias that can lead to the same errors in genotype data set: one of the most common and known is that related to the use of different analyzing machine or procedures. In this case differences can be really important and it is fundamental to standardize each allele value among the laboratories if a multilaboratory project is the choice (Seeb et al., 2007). However, minor but still problematic slippages can happen also using the same machine, as in my case. This is probably due to differences in the capillaries conditions, sample conditions or temperature that make the sample run further or nearer. During my work I got aware of the problem by looking to the alleles frequencies for some of the 9 loci I used. The frequency patterns differed between one group and the other in terms of 2 or 4 bp . For example, if one group locus showed a frequency of $25 \%$ for allele $143,50 \%$ for allele 145 and $25 \%$ for 147 another group could show $25 \% 141,50 \% 143$ and $25 \% 145$. Such a difference would have been unexpected if it had been due just to a genetic distance between the two groups. In fact, a careful analysis of the runs showed that a slippage occurred, thus a manual calibration of the data was required.


Fig. 7. Example of shifted runs. The same allele (labeled 221) has slightly different values in the upper (rounded 220) and in the lower (rounded 221) samples

After the editing is done using Genotyper, the values for each sample of the data sets have been transposed to an Excel sheet using an ad hoc macro already available in the personal computer I used.

## DATA ANALYSIS

Since the enormous increase in available microsatellite data and the development in laboratories techniques that made this kind of data easy to use, many software packages have been created to work with variable markers and microsatellites in particular. As a consequence, there are also many possible formats in which a set of samples can be displayed and often different software require different organization of data (in terms of fragment length vs. number of repetitions, spaces or commas dividing different loci values...). In order to shift from one format to the others without manually editing every single sample, basic formulas of Microsoft Office Excel have been used together with, when needed, the Excel add-in GenAlex (Peakall and Smouse, 2006). This application for Microsoft Excel can be used to analyze many genetic parameters for group of individuals or populations and, once the sample set is organized according to what the software requests, it is possible to export the data set in many format files, ready to be used with other analyzing programs such as GenePop, GeneClass, Arlequin among the most popular ones. Having the entire sample set organized in an Excel sheet, other actions have been possible. Among these, sorting the set according to the alleles' values (i.e. the number of repetitions) has been very useful to detect a bias in the sampling. Animals from the VSP3 group had been caught by sport fishermen during their routine fishing sessions. Instructions for a correct sampling procedure had been provided so that they should have been able to give us the tissue pieces ready for the extraction. However what they probably did was to put several pieces from the same fish in more than one sample tube, resulting in many identical profiles among the samples for this group. As a consequence the sample number decreased from more than 200 samples to less than 70 "real" individuals. With minor incidence this was found even in the other two 'VSP' samples. Sampling methods and procedures are key phases of a scientific reliable work. Nevertheless, often it requires too much time and a high number of qualified people to collect the necessary amount of samples in the proper way. This is why this work is often delegated to others that, thanks to their work activity or hobbies, have more chances to collect samples. However, the lack of knowledge of scientific method in general and the detachment of these people from the research work can result in bad or biased sampling, compromising the entire work. Luckily in my case the bias has been discovered and eliminated, though this decreased the amount of available data.
Before getting to the conclusion that the identical genotypes were due to the sampling, I also considered the possibility that two or more individuals could have carried the same identical genotype just by chance. This could happen
when the loci set used is not variable enough to allow discerning of closely related individuals (such as full or half-sibs). Anyway simple procedures exist to evaluate the probability to have identical genotypes, given the number and variability of the alleles used. GenAlEx gives the opportunity to calculate this Identity Probability. From its values it could be excluded that identical profiles were due to the set of loci used.
Having checked the data set for sampling errors, it was ready to be analyzed and used with various software packages.

As said before genetic parameters have been calculated using the add-in GenAlex, based on Microsoft Excel. Every population was characterized for their genetic parameters, allele frequencies per locus and tested for HW equilibrium, based on expected and observed heterozygosity.
To calculate the average number of alleles for each group and to make this parameter comparable between the groups a statistic value that takes in account the number of samples in each group is needed. For this purpose the software FSTAT (Goudet, 1995) was used to calculate the 'Allelic Richness'.
To test if there was any genetic pattern beyond the subdivision in 6 sampling groups, the data set has been analyzed using the software Structure (Pritchard and Wen, 2003). The program is used to infer population structure using genotype data consisting of unlinked markers. Briefly the software assumes a model in which there are $k$ populations (with $k$ unknown), each of which is characterized by a set of allele frequencies at each locus. Individuals are then assigned to one population or jointly to two or more population (up to $k$ ) if their genotypes indicate that they are admixed. Assignment method aims to achieve populations whose loci are in Hardy-Weinberg equilibrium and linkage equilibrium. The results of this method are then based on these two assumptions (Pritchard et al., 2000a). This means that, given a data set with populations out of HW equilibrium, Structure will re-arrange the samples in order to have groups with loci in HW equilibrium.
Assignments were performed following two procedures: parental assignment and population assignment. Software used were Colony and GeneClass, respectively. Colony (Jones and Wang, 2009) bases its analysis on a maximum likelihood approach to assign parental and sibship among individuals using their multi-locus genotype. The model assumes a sample of individuals divided in 3 sub-samples: offspring, candidate males and candidate females. Individuals in offspring sub-sample are assigned to maternal and paternal families; candidate males and females are then assigned paternity and maternity to these families. Markers are assumed to be in linkage equilibrium and in Hardy Weinberg equilibrium, otherwise power of analysis may result lowered. Other than the
sub-sets, the program allows to enter information about loci and species mating system (polygamy or monogamy). Loci information includes a genotyping error guess including allele drop out and other processing errors. Known sib-ships, maternity or paternity are other information that the software allows to input. These latter two were unknown in my case and thus left blank. Analysis results include information about full and half sib-ships, paternity, maternity and possible genotype errors. When some offspring are not compatible with any couple of candidate female/male, the software generates a virtual breeder with the proper genotype. As a consequence, the mother and the father of each individual can be two animals from the candidate mothers/fathers sets, a virtual individual and one from the candidates or two virtual individuals. Breeders generated by Colony are indicated with either * or \#.
The approach to reassign individuals using GeneClass (Piry et al., 2004) aimed to test how population assignment performed for the analyzed populations when trying to assign individuals to their origin population or to a genetically close one ('Delta Po' and 'Laguna', respectively). This method could be useful when genotypes of candidate parents are unknown and when the aim is to verify the presence of escaped individuals (e.g. coming from a near farm) in a wild population, using as reference populations (1) the wild one and (2) the population the farmed individuals belong to. In this sense the study of two different events can be studied in the same way because they lead to similar situations from a genetic point of view: the presence of genetically different individuals among a wild population. The variable parameters are the number of "strangers" over the total number of animals in the studied area and the genetic distance between the wild and the escaped/released animals. When the genetic difference between the reference populations (indicated, for example, by the Fst values) is higher the test is expected to perform better in recognizing the exact origin population of both wild and escaped/released individuals. On the contrary, mixed animals from similar populations will be hardly recognized as coming from the wild population or the reared one. It is then important to remember that this situation (i.e. released animals coming from a population similar to the wild one) is the suggested one to better control the restocking negative effects.

For each of the 6 samples the basic genetic parameters were calculated. Observed and expected heterozygosity were calculated using GenAlEx, while FSTAT was used to calculate the Allelic Richness to count for different size of sampled populations. Ne was calculated according to the following formula by Kimura and Ohta (mutation rate $\mu=5 \times 10^{-4}$ ):

$$
\begin{aligned}
& H e=1-\frac{1}{\sqrt{8 N e \mu+1}} \\
& N e=\frac{H e^{2}}{8 \mu\left(1-H e^{2}\right)}
\end{aligned}
$$

Allelic richness appears to be lower in the first two VSP sample, as well as expected heterozygosity and Ne , although these differences are not significant. The other groups show higher and comparable values for these parameters. VSP1 is the only group that shows an observed heterozygosity (Ho) higher than the expected (He), anyway not significant.

Tab. 4 Genetic parameters for each groups: Number of individuals genotyped, Observed heterozygosity ( Ho ) average and range, Expected heterozigosity ( He ) average and range, Allelic Richness, Effective population size ( Ne ) average and range

| Population | No of <br> individuals | Ho | He | Allelic <br> Richness | Ne |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Delta Po | 161 | 0,735 <br> $(0.572-0.886)$ | 0,747 <br> $(0.535-0.955)$ | 13,844 | 316.0 <br> $(100.4-2565.7)$ |
| Laguna | 172 | 0.737 <br> $(0.468-0.948)$ | 0,751 <br> $(0.470-0.927)$ | 14,050 | 324.4 <br> $(70.8-1532.1)$ |
| VSP1 | 164 | 0,734 <br> $(0.543-0.949)$ | 0.695 <br> $(0.508-0.907)$ | 8,125 | 233.5 <br> $(86.7-1158.6)$ |
| VSP2 | 113 | 0.676 <br> $(0.469-0.940)$ | 0.687 <br> $(0.467-0.943)$ | 7,932 | 223.2 <br> $(69.7-1989.3)$ |
| VSP3 | 68 | 0.734 <br> $(0.532-0.926)$ | 0,733 <br> $(0.506-0.953)$ | 13,820 | 290.7 <br> $(86.0-2492.7)$ |
| RIPR | 178 | 0,752 <br> $(0.562-0.854)$ | 0,751 <br> $(0.544-0.873)$ | 14,367 | 323.5 <br> $(104.9-804.7)$ |

Hardy-Weinberg equilibrium probability was calculated for each locus and for each population. 'Laguna', 'VSP1' and 'VSP2' groups have most of their loci out of HW equilibrium. The causes of this deviation could be multiple and related to low Ne , recent bottleneck or inbreeding. Other sources of bias could be
genotyping errors such as null alleles or stuttering. In the case of the 'VSPs' samples this could be somehow due to the presence of released individuals. For the Laguna population it is more difficult to understand but the causes could be inbreeding, mixing between different strains or differences in the time of the samplings. Moreover, the genotyping had been already performed when I used those data and genotyping errors cannot be excluded.

Tab. 5 Hardy Weinberg equilibrium: probability for each locus for each group. In bold significant values ( $p<0.01$ ) after Bonferroni correction ( $N o$ of multiple tests $=54 ; p$ after correction 0.0002)

| Locus | Range | DeltaPo | Laguna | VSP1 | VSP2 | VSP3 | RIPR |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Dla008 | 214-318 | 0.004 | <0.0002 | <0.0002 | <0.0002 | 0.075 | 0.001 |
| Dla020 | 144-180 | 1.000 | <0.0002 | <0.0002 | <0.0002 | <0.0002 | 0.999 |
| Dla105 | 135-181 | 0.918 | <0.0002 | <0.0002 | <0.0002 | 0.015 | 0.886 |
| Dla119 | 221-265 | 0.320 | 0.832 | <0.0002 | <0.0002 | <0.0002 | 0.850 |
| Dla228 | 72-104 | 0.474 | <0.0002 | 0.027 | 0.021 | 0.590 | 0.840 |
| Dla016 | 220-272 | 0.113 | <0.0002 | <0.0002 | <0.0002 | 0.022 | 0.001 |
| Dla244 | 94-114 | 0.908 | 0.994 | 0.858 | 0.558 | 0.257 | 0.983 |
| Dla248 | 120-126 | 0.543 | <0.0002 | 0.898 | 0.011 | 0.309 | 0.676 |
| Dla145 | 153-195 | <0.0002 | 0.921 | <0.0002 | <0.0002 | 0.236 | <0.0002 |

A test for population differentiation was made via the Analysis of Molecular Variance (AMOVA). The Fst values from this analysis are equal to the amount of variation that arises from inter-population differences rather than intrapopulation differences. The pair wise Fst matrix suggests a subdivisions of the six groups into three subgroups: 'VSP1' and 'VSP2' have non-significant, very low Fst and they are both genetically far from the group composed by 'Delta Po', 'Laguna' and 'RIPR'. 'Delta Po' and 'RIPR' are the most similar populations. Indeed, part of the RIPR individuals is the same as some of 'Delta Po' individuals. VSP3 is genetically different from both groups.

Tab. 6 Fst matrix: bold values indicate a level of significance p<0.01 after Bonferroni correction (No of multiple tests $=15 ; p$ after correction 0.0007 )

|  | DeltaPo | Laguna | VSP1 | VSP2 | VSP3 | RIPR |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| DeltaPo |  | 0.1151 | 0.0001 | 0.0001 | 0.0001 | 0.4442 |
| Laguna | 0.001 |  | 0.0001 | 0.0001 | 0.0001 | 0.0173 |
| VSP1 | $\mathbf{0 . 0 2 3}$ | $\mathbf{0 . 0 2 7}$ |  | 0.0260 | 0.0001 | 0.0001 |
| VSP2 | $\mathbf{0 . 0 2 5}$ | $\mathbf{0 . 0 3 1}$ | 0.003 |  | 0.0001 | 0.0001 |
| VSP3 | $\mathbf{0 . 0 2 7}$ | $\mathbf{0 . 0 2 7}$ | $\mathbf{0 . 0 5 4}$ | $\mathbf{0 . 0 5 7}$ |  | 0.0001 |
| RIPR | 0.000 | 0.002 | $\mathbf{0 . 0 2 3}$ | $\mathbf{0 . 0 2 6}$ | $\mathbf{0 . 0 2 8}$ |  |

Principal component analysis (performed by GenAlEx) is based on Fst distances previously calculated and provides a better visualization of the organization, with three distinct groups according with Fst values. Despite the fact that they are all samples from the Venice lagoon, there seem to be some big differences among samples 'VSP1' and 'VSP2', 'Laguna', 'VSP3'. The difference between the 'VSP1VSP2' group and the others could be due to the effect of restocking (i.e. presence of individuals coming from the hatchery that modifies the genetic structure of the group). The difference between 'VSP3' and 'Laguna' is more difficult to explain, but can arise from the time distance of the two sample ( 3 years passed) and from genetic differentiation at smaller scale or, as said before, by genotyping errors in the 'Laguna' group, that cannot be excluded since the data came as already genotyped and were just analyzed. The sample from the Lagoon and the sample from the Po river delta, despite being characterized by a higher geographical distance, appear to be much more genetically similar than the 'VSP3' and the 'Laguna' groups.


Fig 8. Principal Component Analysis of the 6 groups, based on the Fst distances (performed by GenAIEx)

Pie graphs (data not shown) displaying allele frequencies per each locus have been calculated, using GenAIEx, for each population. Such a way of visualizing the most frequent allele and the allelic pattern is very useful to check for genotyping errors such as small slippages or editing errors (e.g. calling the same allele with different "names" in different populations). Thanks to this graphs a bias has been found in locus Dla145 for population 'DeltaPo'. In that case the same alleles were given a length of two base pairs less than in all the other populations, resulting that allele 167 had the same frequency as allele 169 of the others populations, 169 the same as 171 and 171 the same as 173. After correction the pie graphs were much more similar. Nevertheless, thanks to this
method, something can be inferred about the population structure as clear differences can be seen for what regards allelic richness. Some differences in alleles frequencies are present too. The conclusions that can be drawn from this visualization of the genetic arrangement of the groups are in accordance with those that can be drawn from the results collected in Tab.4.

A more accurate analysis of the groups was made using the software Structure, in order to find patterns of differentiation regardless the a priori subdivision in sampling groups. Two different approaches were followed: the first one including all the six populations, in order to get more information about the interpopulation and intra-population structure and to infer something about what makes each population genetically similar or different from the others. The best run (i.e. the run that has the maximum likelihood) was a run that split the whole set in two populations ( $k=2$ ). In the next figure, each line represents an individual. Samples are grouped per population separated by thin black lines in the figure. The proportion of green or red in each line corresponds to the probability of an animal to belong to the first or the other group based on its genotype.



Fig. 9. Output from Structure analysis (K=2) with 6 populations (1-Delta Po, 2-Laguna, 3-VSP1, 4-VSP2, 5-VSP3 and 6-RIPR). Each line corresponds to an individual from the relative group. Amount of green or red color indicates the belonging to green or red group.

The differentiation highlighted by the AMOVA analysis is still clear looking at Structure output. In particular the differentiation between the 'VSP1-VSP2' group (in the graph labeled with 3 and 4) and the 'Delta Po-Laguna-RIPR' group (labeled with 1,2 and 6 , respectively) is here suggested by the high amount of red lines (i.e. individuals assigned to the red population) in VSP1 and VSP2 and a majority of green lines (i.e. individuals assigned to the green population) in the other groups. In this visualization a clear difference between 'VSP3' (labeled with 5) and 'Delta Po - Laguna - RIPR' (1-2-6) group in not visible. In this latter group a further differentiation can be made between the 'Laguna' population and the couple 'Delta Po' and 'RIPR' populations that are characterized from some "mostly red lines" among the numerous "green lines". The individuals corresponding to these red lines are somehow similar to a great part of VSP1's and VSP2's individuals. In particular the table below report the IDs of the four "red lines" from group 'RIPR' and the amount of red lines in the VSP1 and VSP2 groups. Subdivision has been made according to the Q-value for each group (red or green) of each individual.

Tab 7. IDs of the animals assigned to the red group for 'RIPR', percentages of samples assigned to the red group for 'VSP1' and 'VSP2'.

| RIPR | VSP1 | VSP2 |
| :--- | :--- | :--- |
| R26, R30 <br> R50, R67 | $61.6 \%$ <br> (IDs list and Q values in <br> appendix) | $60.2 \%$ <br> (IDs list and Q values in <br> appendix) |

Since 'VSP1', 'VSP2' and 'RIPR' groups appeared to be linked by a small number of individuals in RIPR population, a further Structure run was performed. This time the software was asked to analyze only these three samples, in order to search a more accurate subdivision into groups. The run with maximum likelihood split the individuals in three populations ( $k=3$ ). Again, the output figure shows a series of lines (each corresponding to a single individuals) composed by a red, a green and a blue segment. 'VSP1' is here labeled with 1 , 'VSP2' with 2 and 3 corresponds to 'RIPR' group.


Fig. 10. Output from Structure analysis $(\mathrm{K}=3)$ with 3 populations (1-VSP1, 2-VSP2 and 3-RIPR). Each line corresponds to an individual from the relative group. Amount of green, blue or red indicates that the individual belongs to green, blue or red group.

The differentiation between the samples from the release area and the breeders group is still strong. A link between the groups is still visible as well, this time
including more individuals from the 'RIPR' group. In particular two individuals, marked with red lines, belonging to the 'RIPR' population appear to be somehow similar to a numerous group of individuals of the 'VSP1' and 'VSP2' groups (red lines). In the same way 7 individuals, marked with mostly blue lines, in 'RIPR' group are similar to a large group of individuals from 'VSP1' and 'VSP2' group. The amount of blue, as showed by different $Q$ values, is different among individuals in the 'RIPR' group as well as in the 'VSP' groups. This means that individuals with a larger green proportion of the line share something with the other mostly green individuals. The table below reports the IDs of the nine animals from 'RIPR' that belong either to the "red" or "blue" population and the corresponding amount of individuals from the 'VSP1' and 'VSP2' groups.

Tab 8. IDs of the animals assigned to the red and the blue groups for 'RIPR', percentages of samples assigned to the red and the blue groups for 'VSP1' and 'VSP2'

| Structure group | RIPRODUTTORI | VSP1 | VSP2 |
| :--- | :--- | :--- | :--- |
| Red | R50 | $32.9 \%$ <br> (IDs list and Q values <br> in appendix) | $31.9 \%$ <br> (IDs list and Q <br> values in appendix) |
| Blue | R25, R26, R30, R41 <br> R45, R49, R68 | $45.1 \%$ <br> (IDs list and Q values <br> in appendix) | $43.4 \%$ <br> (IDs list and Q <br> values in appendix) |

Comparing the results from the Structure's runs we can see that the "red group" from the first run (with all 6 populations) has been split in a "red group", that includes R50 and R67, and a "blue" one that, other than R26 and R30 (previously included in the red group), includes 5 more individuals. The amount of individuals from the VSP groups linked to these breeders has increased in 'VSP1' and 'VSP2' from $61.6 \%$ and $60.2 \%$ to $78 \%$ and $75.3 \%$ ("red" + "blue"), respectively.

The pattern highlighted by Structure analysis suggested a possible link between the greatest part of the individuals sampled along the Venice Lagoon inlet and a small number of individuals from the farm that provided the juvenile for the restocking program.
To obtain more information about this situation I tried to assign the animals from 'VSP' groups to their parents through parentage assignment and to their origin population using the population assignment approach.
Parentage assignment was performed using the software Colony 2.0. As candidate male and female parents the individuals from 'RIPR' group have been used. Colony output suggests, for each individual, the pair of parents that have the maximum probability to be the true parents of the tested animal. When one
(or both) parent of a tested individual misses from the given data set, the software generates one (or two) virtual reproducing individual that can provide the alleles the tested animal has. The software then tries to get the maximum parsimony in reproducers number providing a list of real and virtual parents shorter as possible. In the output file real breeders are labeled with the name provided with the data set while program-generated individuals are labeled with either * or \# to differentiate male from female parents. Among the information it is possible to give to the software, a rate of genotyping error can be input. In my case an estimated error rate of $1 \%$ was suggested for all loci. The result from the parentage assignment analysis suggested a subdivision of the greatest part of the offspring in few large families, generated by a small number of parent couples that spawned most of the juveniles. The following table summarizes the major families structure for 'VSP1' and 'VSP2' groups. As mentioned before parents tagged with * or \# are virtual individuals generated by the software to best fit the offspring genotypes.

Tab. 9. Results from Colony 2.0 parentage assignment. Here are indicated the $\mathbf{5}$ most numerous families, the number of individuals from 'VSP1' (left) and 'VSP2' (right) assigned to each family and the percentage on the whole sample. (For a more complete list see the appendix)

|  | VSP1 |  |  | VSP2 |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Fam | No of <br> individuals | $\%$ | Fam | No of <br> individuals | $\%$ |
| R50R67 | 28 | 17.07 | R50R67 | 9 | 7.96 |
| R41R45 | 27 | 16.46 | R41R45 | 23 | 20.35 |
| R26R30 | 22 | 13.41 | R26R30 | 14 | 12.39 |
| *1R67 | 22 | 13.41 | *1R67 | 22 | 19.47 |
| R26R25 | 16 | 9.76 | R26R25 | 11 | 9.73 |

According to the results from the parentage assignment the sum of the offspring generated from just 5 parent pairs accounts for $70 \%$ of the total sampled individuals ( $70.12 \%$ for VSP1 and $69.91 \%$ for VSP2). Moreover, these five pairs are composed by only seven individuals from the 'RIPR' group. This result alone would be unexpected considering the large number of breeders present in the tank where the spawning took place. Moreover the success of a single parent seems to be related to the presence of another precise individual (e.g. 'R41' and 'R45' appear only as a couple, and their success is very low or null when associated to other breeders).
Apparently one male (tagged *1 by the software) was missing from the given data set. This individual, in couple with the female R67, generated a great amount of juveniles: $13.41 \%$ in VSP1 and $19.47 \%$ in VSP2. Important similarities between the results of this analysis and those of Structure can be found looking
at the IDs from 'RIPR' group. The same individuals that appeared "different" from the others in Structure graphs (the red and blue lines) have been found by Colony to be among the most successful parents. In particular the analysis performed with three populations ('VSP1', 'VSP2' and 'RIPR') separated the seven most productive breeders from all the others (plus R49 and R68). A more accurate analysis of the Q values of the animals from 'VSP1' and 'VSP2' and the candidate parents (according to Colony) suggested another conclusion for the individuals that had been assigned to the '* 1 ' male. In fact all the animals assigned to the couple * $1 \times$ R67 are assigned to the "red" population both in the 6 -groups analysis (where "red" included R26, R30, R50 and R67) and in the 3groups analysis (where "red" included only R50 and R67). Moreover the Q values for these individuals are the same as those of the animals assigned to the couple R50 x R67. For these reasons these samples should be assigned to this latter family. Looking at the inferred genotype of the virtual breeder ${ }^{*} 1$ it can be noticed that it differs from the R50's genotype just for two alleles (one in locus Dla008 and one in locus Dla145). In this case genotyping or editing errors in the 'VSPs' individuals are probably the cause of this bias. Adding the individuals assigned to the ${ }^{*} 1 \times$ R67 family to those assigned to the R50 $\times$ R67 family the amount over the total assigned individuals raises to $30.5 \%$ and $27.4 \%$ in 'VSP1' and 'VSP2', respectively. These percentages are much more comparable to those resulting from Structure's 3-groups analysis ( $32.9 \%$ for VSP1 and 31.9\% for VSP2) confirming that this is probably the right assignment pattern.
In the same way an accordance can be found between the "blue" Structure's group and the most productive breeders. Summing the reproductive output of the couples R41 x R45, R25 x R26 and R26 x R30 the resulting percentages ( $39.63 \%$ and $42.47 \%$ for 'VSP1' and 'VSP2', respectively) are very close to the percentages of individuals assigned to the "blue" group ( $45.1 \%$ and $43.4 \%$ ). Less numerous families have not been identified by Structure analysis, that grouped the most numerous family (R50 x R67) in the "red" group and the other three most numerous ( $225 \times$ R26, R26 $\times$ R30 and R42 $\times$ R45) in the "blue" group. Probably this is due to the fact that the minor families are too small to be identified by Structure as separated sub-samples. An alternative explanation to these minor differences between Structure subdivision and parentage assignment results can be the following: Structure is more accurate than Colony in this context and the fact that Colony assigned individuals not assigned by Structure is due to the similarity between wild and released animals present in the VSP1 and VSP2 groups. This means that, indicating a $1 \%$ possible error in the genotyping, the software can assign to a couple of breeders even the wild individuals, just considering as biased one or few alleles.

Nevertheless, the results output by Colony are probably even influenced by the fact that the groups 'VSP1' and 'VSP2' have the most of their loci out of HardyWeinberg equilibrium. As mentioned in the 'Material and Methods' this can lower the power and accuracy of the parental analysis.
Another difference is that Structure analysis indicated R49 and R68 as part of the "blue" group, whereas Colony didn't find individuals in 'VSP1' and 'VSP2' related to these breeders.
The analysis of 'VSP' groups through population assignment has been carried out with the software GeneClass. As reference populations I used the Po River delta population and the Venice lagoon one. In other words the software was asked to assign every individual from the samples either to the wild Lagoon population or to the population the parents came from, on the base of their genotypes. This method is the only one possible when genotyping information of the actual parents are not provided (so that parental assignment cannot be performed). The method used is a Bayesian one, the same used for detecting immigrated individuals among a wild population explained in Rannala and Mountain, 1997 (Rannala and Mountain, 1997).
Results are here reported for all the six populations. This means that the assignment was performed even for the reference populations against themselves. This can provide information about the genetic differentiations between the two reference populations for what regard the population assignment. More detailed results for each individual from 'VSP1' and 'VSP2' group can be viewed in the appendix where, for each individual assigned, a related percentage score is reported. This value indicates the level of probability for the respective assignment results.

Tab. 10. Results from GeneClass population assignment. Results are expressed for each group in terms of percentage of individuals assigned to 'Delta Po' population. The remaining were assigned to the 'Laguna' population.

| Populations | \% assigned to Delta Po |
| :--- | :---: |
| Delta Po | $80.74 \%$ |
| Laguna | $26.74 \%$ |
| VSP1 | $88.41 \%$ |
| VSP2 | $87.61 \%$ |
| VSP3 | $61.76 \%$ |
| RIPR | $71.34 \%$ |

Results reported above reflect the effect of the genetic distances between populations already observed in the Fst values. The 'Laguna' and the 'Delta Po' populations are very similar for the loci used, thus the population assignment
could be inaccurate when analyzing the samples from these populations. Anyway, the percentages of individuals assigned for 'VSP' groups are in accordance with the high prevalence of released individuals among the samples. In both 'VSP1' and 'VSP2' the amount of animals assigned to the 'Delta Po' population is near $90 \%$, even higher than the assignment percentage of the 'Delta Po' population against itself. 'VSP3' seems to be much similar to the 'Venice lagoon' with the second lowest percentage of individuals assigned to the 'Delta Po' population (61.76\%). This was expected considering that these animals had been sampled in the lagoon. Nevertheless, from the Fst values this wasn't so clear, nor from the PCA, since the 'VSP3' group was as far from the Venice lagoon population as from 'Delta Po'. As expected the lowest score is that of the 'Laguna' group. Nevertheless more than a quarter of the individuals of the Venice Lagoon have been assigned to the 'Delta Po' reference population. This can be considered another clue about the low genetic difference between these two groups. The accuracy of the population assignment can be checked comparing the results with the Structure's and Colony's output (see appendix). Among the individuals assigned to the 'Laguna' population there are some of the individuals that both Structure and Colony recognized as released. In particular $12.1 \%$ of them belong to the "blue" Structure group (families R25 x R26 and R41 $\times$ R45) and $30.3 \%$ belong to the "red" group (family R50 x R67). As these results suggest, population assignment, in this case (i.e. when the references populations are genetically very similar), can provide only approximate indications about the composition of samples and the effects of the restocking. Considering the Structure results and parentage assignment as more reliable, GeneClass approach wouldn't be enough accurate to give information about the correct impact of the released animals.

## CONCLUSIONS

In fall 2007 a batch of Mediterranean sea bass juveniles (coming from a local fish farm) was released in the Venice Lagoon (precisely near the Malamocco inlet) by the 'Fishermen association of the Venice Lagoon' to increase the number of fishes available for the sport anglers. The animals had been previously grown in an enclosed area and fed for some months and then freed when their size was 15 cm on average. During the summer of the next year two samplings took place in the area of the inlet dams, near the release site. One more sampling was delegated to the fishermen during their normal fishing sessions. In addition to these, genotyping data from Venice Lagoon wild population and Delta Po wild population were available thanks to previous work carried out in the same laboratories. Breeders' genotypes (coming from the Po river estuary area) were also available for parentage analysis.

This work presents the results of the analysis of 9 microsatellite loci from these groups of Mediterranean Sea Bass. To get information about the population differentiation and the possible effect of the restocking program three approaches were followed: an analysis of the population structure (Structure), a parentage assignment test for the individuals sampled around the release area (Colony 2.0) and a population assignment test (GeneClass).
The Mediterranean sea bass juveniles used for the restocking came from a local farm that uses animals from a wild population of the Po river estuary as adult breeders. Genetic differentiation values (Fst) between the population from the Po river delta and a sample from the Venice Lagoon (available thanks to a previous work on this species) suggests that the genetic distance between them is low and not significant. Such a relation between the depleted population and the population of origin of released individuals should be good for restocking in order to maintain a stock similar to the original one. Anyway, wherever possible, the recommended strategy for releases of juveniles is to use the local broodstock (Ward, 2006). This approach is called 'supportive breeding'. In any case, it's important to remember that it can still have deleterious effect when not planned and carried out carefully. This means that the broodstock should be numerous, avoiding to introduce a large number of closely related individuals that could determine variation in the population structure and could affect seriously the wild stock. The effect that introduced individuals can have on the wild population is also related to the number of released individuals or, more precisely, to their amount over the number of already present fishes. As this ratio gets higher the impact is expected to be more important.

Results from Structure analysis and from parentage assignment indicate that, after the release, the juveniles freed didn't move far from the releasing area, but settled in the surrounding waters taking over the wild fishes for presence and abundance. The distance of the sampling site was indeed in the 1 km range from the enclosed tank where the animals had grown and Colony assigned all the fishes sampled from the inlet dams to at least one of the parents from the 'Ca' Zuliani' farm. Results from Structure analysis suggest a similar but less drastic scenario, as more than $70 \%$ of the samples from VSP1 and VSP2 have been grouped with few breeders. Anyway it's important to remember that the fact that so many released fishes have been caught could be also related to the sizeselective sampling that has been undertaken. In other words, all the fishes caught were of the size that the released fishes were expected to be considering the size at release and an average growth rate. Another possible explanation for the high presence of hatchery fishes in the same sample is the observation for some species that released individuals formed their own school separated from wild stock fishes (Jeong et al., 2007). In any case higher presence of wild individuals was expected, considering the size of the sampling. Maybe, the presence of wild sea bass in the area was actually very low before release (and that's indeed what pushed the fishermen to undertake the restocking). This suggests that the impact of a large number of released individuals ( 30,000 in this case, for the area studied) could be really important and should be considered even more carefully.
The juveniles released in the Lagoon came from a 'mass spawning' event in the farm that provided them. Such a procedure implies the presence of a large number of adults in the same tank that are free to couple. So, no selection for male/female is made by farmers and no control of reproducing success can be made a priori. The result from my analysis highlighted one of the consequences of this method in terms of sib-ship of the offsprings. As already reported by some authors, just few individuals actually reproduce using this procedure (Chatziplis et al., 2007). The greatest part of the animals sampled after the release came from a very small group of breeders. Moreover, these reproducing individuals seem to be arranged in a number of pairs lower than expected considering the number of possible crosses. In other words, a single male/female produces the greatest part of its offspring with just one particular female/male or two. For example, the individuals R41 and R45 are among the parents that generate most of the juveniles. Nevertheless, they appear only together and never alone coupling with other individuals. This fact is very important for aquaculture practicing in general and for the production of individuals for restocking as well. The genetic variability of the released stock is fundamental to grant a long term
effectiveness of a restocking program. Thus, every aspect in the chain of events that go from the juveniles production to the release should be carefully monitored in order to maintain genetic variability as high as possible. In the case reported in this work two facts affected the variability of the batch used: low number of successful breeders and a reduced number of actual crosses. These features explain the lower values of parameters such as Allelic richness and Ne for the group 'VSP1' and 'VSP2'. The result of using a group of juveniles arranged like this (low variability) for a restocking program is the introduction of a group of animals genetically different from the wild one, despite its origin from a population that is similar to the depleted one. Moreover, as said before, the impact of the released animals is expected to be even higher due to the abundance of the hatchery sea basses among the whole sample. This can be seen from the Fst matrix (Tab. 6) and from the Principal Component Analysis graph (Fig. 8): despite being composed by animals either from the farm (thus similar to 'RIPR' group) or from the wild (thus similar to the 'Laguna' group) VSP1 and 'VSP2' are genetically different from both 'RIPR' and 'Laguna'. This feature probably affected the results of population assignment too.

The long term effect that this can have on the effectiveness of the restocking is different from the effect of using a genetically far population (out-breeding depression) but are still critical for the maintenance of the population. Low variability in functional genes (mirrored in this case by low variability in neutral markers) means that a population could not be ready to response to environmental changes. An immediate increase in number of fishes available for fishermen (as reported orally by the 'Fishermen association') should be considered carefully since it is probably just linked to the large number of sea basses released. An actual benefit for the population of the Lagoon would mean a long-term increase in animals fished thanks to the implementation of the released animals to the wild stock, without compromising the genetic structure of the stock itself. Further studies should be undertaken to test whether the restock is still having effects on the studied area or if the situation has changed back to what it was before. The risk in fact is that, after the expected reduction of the released batch due to fishing or natural mortality, the survivors won't be able to sustain the depleted population. Moreover natural mortality for released individuals is expected to be higher than for wild individuals (Brown and Day, 2002), making this issue more critical.

The example of restocking presented here highlights the importance of a scientific approach to the human actions that can lead changes in the environment. Blankenship and Leber (1995) widely described what a "responsible approach to marine stock enhancement" is, identifying ten critical
points of these kind of actions. These comprised the definition of quantitative measures of success; the use of genetic resource management to avoid deleterious genetic effects; the identification of released hatchery fishes to assess stocking effects. The importance of a proper management in stock enhancement is also embodied in articles In the FAO Code of Conduct for Responsible Fisheries (1995).
The juvenile release action was undertaken by a private association without the support of experts in this field. The approach currently suggested comprises, for example, a previous analysis of the group of animals that is going to be released. The information from these analysis are important to choose whether that batch is suitable to be used or not as well as the release "tactics". Moreover consequences of the release can be better forecast and analyzed in the future and appropriate management programs would be easier. Otherwise the risk is to invest money and time for actions that have no appreciable effects on the natural stock or, worse, that have a bad impact on the species. If an analysis based on neutral marker had been undertaken on the batch used in 2007, probably it would have been suggested to add some juveniles from other 'mass spawning' events in order to increase the genetic variability (i.e. allelic richness and Ne ) of the group of animals freed. These parameters are indeed appropriate to test the goodness of the batch used. A low value of these indicators means that the variability in the group to be released could not be sufficient to grant an appreciable re-enhancement of the target population. In this case the values for 'VSP1' and 'VSP2' are in fact lower, even if not significant.
The approach used in this work represents a possible way to monitor an area subjected to recent restocking. The impact of released individuals on the wild population was studied using both parentage and population assignment. The discordance between the results from the two approaches underlines the importance of accurate genetic information about the breeders used to generate juveniles. One of the most important aspects concerning the impact of introduced animals is their number over the total sampled fishes. Without knowing the genotypes of the parents from the Ca' Zuliani hatchery an affordable evaluation of the number of animals coming from the farm would have been very difficult using only the population assignment approach. In my case Structure analysis provided some information about this, as well as parentage assignment. Anyhow, the success of Structure analysis was due to the low number of breeders that generated the batch used for restocking. If a greater number of fishes had reproduced (i.e. if the production of juveniles had been made properly!), probably Structure results would not have been so clarifying about the VSP1 and VSP2 group organization.

At the moment, four years has passed from the releasing. This means that the fishes that have survived are probably mature, considering that fish sexual maturity occurs, in the Mediterranean, at an age of 2-4 years (FISHBASE). The impact of these fishes can thus continue through their progenies or, in a more complicate way, through a mixed progeny coming from crosses between wild and hatchery fishes. The consequences of such a hybridization are multiple and complicated in at least two way: the "biological" way for the complex interaction between genetic strains that, although being similar, can produce actual hybrids, with both positive (e.g. hybrid vigor) (Bryden et al, 2004) and negative (e.g. breakdown of co-adapted genes) consequences (Edmans and Deimler, 2004); the "technical" way since the difficulties of studying and analyzing an admixture of wild, cultured and hybrid subpopulation is a big deal even with the more advanced markers and technologies available.
Nevertheless this kind of analysis could be unnecessary if the restocking program had been carried out properly, to grant a successful re-enhancement. In order to do this some simple rules had to be followed before the actual release, other than after. A scientific approach to this kind of actions performed by humans is fundamental since the consequences of wrong procedures can lead to seriously negative scenarios. When correctly done, anyway, human interventions could provide real benefits to the natural environment, that result in improvements both for "human" populations and for "non-human" populations.

Abercrombie, D.L., Clarke, S.C. and Shivji, M.S. (2005) Global-scale genetic identification of hammerhead sharks: application to assessment of the international fin trade and law enforcement. Conservation Genetics 6, 775788

Allegrucci, G., Fortunato, C., Sbordoni, V. (1997) Genetic structure and allozyme variation of sea bass (Dicentrarchus labrax and D. punctatus) in the Mediterranean Sea. Mar. Biol. 128, 347-358

Araki, H., Cooper, B., Blouin, M.S., (2007b) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318, 100-103

Araki, H., Cooper, B., Blouin, M.S., (2009) Carry-over effect of captive breeding reduces reproductive fitness of wild-born descendants in the wild. Biol. Lett.

Araki, H., Schmid, C. (2010) Is hatchery stocking a help or harm? Evidence, limitations and future direction in ecological and genetic surveys. Aquaculture 308, S2-S11

Bahri-Sfar, L., Lemaire, C., Ben Hassine, O.K., Bonhomme, F. (2000) Fragmentation of sea bass populations $n$ the western and eastern Mediterranean as revealed by microsatellite polymorphism. Proceedings of the Royal Society of London Series B 267, 929-935

Bartley, D.M., Bondad-Reantaso, M.G., Subasinghe, R.P. (2006) A risk analysis framework for aquatic animal health management in marine stock enhancement programmes. Fish. Res. 80, 28-36

Bax, N., Williamson, A., Aguero, M., Gonzalez, E. \& Geeves, W. (2003) Marine invasive alien species: a threat to global biodiversity. Marine Polut. 27, 313323

Bell, J.D., Rothlisberg, P.C., Munro, J.L., Loneragan, N.R., Nash, W.J., Ward, R.D., Andrew, N.L.( 2005) Restocking and stock enhancement of marine invertebrate fisheries. Adv. Mar. Biol. 49, 1-370

Blankenship, H.L., Leber, K.M. (1995) A responsible approach to marine stock enhancement. Am. Fish. Soc. Symp. 15, 167-175

Brown, C., Day, R. L. (2002) The future of stock enhancements: lessons for hatchery practice from conservation biology. Fish and Fisheries 3, 79-94

Bryden, C.A., Heath, J.W., Heath, D.D., (2004) Performance and heterosis in farmed and wild Chinook salmon (Oncorhynchus tshawytscha) hybrid and purebred crosses. Aquaculture 235, 249-261

Castilho, R., Ciftci, Y. (2005) Genetic differentiation between close eastern Mediterranean Dicentrarchus labrax (L.) populations. J. Fish Biol. 67, 17461752

Castilho, R., McAndrew, B.J. (1998) Population structure of seabass in Portugal: evidence from allozymes. J. Fish Biol. 53: 1038-1049

Chatziplis, D., Batargias, C., Tsigenopoulos, C.S. et al. (2007) Mapping quantitative trait loci in European sea bass (Dicentrarchus labrax): the BASSMAP pilot study. Aquaculture 272, S172-S182

Chistiakov, D.A, Hellemans, B., Haley, C.S., Law, A.S., Tsigenopoulos, C.S., Kotoulas, G., Bertotto, D., Libertini, A., Volckaert, F.A. (2005) A microsatellite Linkage Map of the European Seabass Dicentrarchus labrax. Genetics 170, 1821-1826

Cooke, S.J., Cowx, I.G. (2004) The role of recreational fishing in global fish crises. BioScience 54(9), 857-859

DeSalle, R. and Birstein, V.J. (1996) PCR identification of black caviar. Nature 381, 197-198

Edmands, S., Deimler, J.K., (2004) Local adaptation, intrinsic coadaptation and the effects of environmental stress on interpopulation hybrids in the copepod Tigriopus californicus. J. Exp. Mar. Biol. Ecol. 303, 183-196.

Field, D., Wills, C. (1996) Long, Polymorphic Microsatellites in Simple Organisms. Proceedings: Biological Sciences 263, 209-215

Food and Agricolture Organization of the United Nations - www.fao.org
Froese, R., Pauly, D. (2006) Fish base. www.fishbase.org
Galil, B.S. (2007) Loss or gain? Invasive aliens and biodiversity in the Mediterranean Sea. Marine Pollution Bulletin 55, 314-322

García-Asorey, M.I., Escati-Peñaloza, G., Parma, A.M., Pascua, M.A. (2011) Conflicting objectives in trophy trout recreational fisheries: evaluating tradeoffs using an individual-based model. Canadian Journal of Fisheries and Aquatic Sciences 68, 1892-1904

Goudet, J. (1995) Fstat version 1.2: a computer program to calculate Fstatistics. J Heredity 86, 485-486

Haffray, P., Tsigenopoulos, C.S., Bonhomme, F., Chatain, B., Magoulas, A., Rye, M., Triantafyllidis, A., Triantaphyllidis, C., 2007. European sea bassDicentrarchus labrax. Geninpact final scientific report.

Hancock, J.M. (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Microsatellites. Evolution and Applications (eds GoldsteinDB, SchlöttererC), Oxford University Press, New York, 1-9.

Hauser, L., Baird, M., Hilborn, R., Seeb, L., Seeb, J. (2011) An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (Oncorhynchus nerka) population. Molecular Ecology Resources 11 (Suppl. 1), 150-161

Hilborn, R., Eggers, D., (2000) A review of the hatchery programmes for pink salmon in Prince William Sound and Kodiak Island, Alaska. Trans. Am. Fish. Soc. 129, 333-350

Honma, A. (1993) Aquaculture in Japan. Japan FAO Association, Tokyo, Japan, 98 pp

Imamura, K., (1999) The organization and development of sea farming in Japan. In: Howell, B.R., Moskness, E., Svasand, T. (Eds.), Stock Enhancement and Sea Ranching. First International Symposium on Stock Enhancement and Sea Ranching, Bergen, Norway. Blackwell Science, Oxford, 91-101

Invitrogen (2011) SYBR Safe DNA Gel Stain
Jackson, J.B., Kirby,M.X., Berger,W.H., Bjorndal, K.A., Botsford, L.W., Bourque, B.J., Bradbury, R.H., Cooke,R.,Erlandson, J., Estes, J.A.,Hughes, T.P., Kidwell, S., Lange, C.B., Lenihan,H.S., Pandolfi, J.M., Peterson, C.H., Steneck, R.S., Tegner, M.J., Warner, R.R., (2001) Historical overfishing and the recent collapse of coastal ecosystems. Science 293, 629-637

Jeong, D.S., Blanco Gonzalez, E., Morishima, K., Arai, K.E., Umino, T. (2007) Parentage assignment of stocked black sea bream, Acanthopagrus schlegelii in Hiroshima Bay using microsatellite DNA markers. Fisheries Science 73, 823-830

Jones, O., Wang, J. (2009) COLONY: A program for parentage and sibship inference from multilocus genotype data. Mol Ecol Res 10, 551-555

Katsares, V., Triantafyllidis, A., Karaiskou, N., Abatzopoulos, T., Triantaphyllidis, C. (2005) Genetic structure and discrimination of wild and cultured Greek populations of the European sea bass (Dicentrarchus labrax, Linnaeus 1758). 12 ${ }^{\text {th }}$ Panhellenic Congress of Ichthyologists, October 13-16, Drama, Greece, Book of Abstracts, 350-353

Lande, R., Barrowclough GF. (1987) Effective population size, genetic variation, and their use in population management. In: Viable Populations for Conservation, ed. ME Soule, pp. 87-123

Lewis, M. (2011) Agarose gel electrophoresis (basic method). Biological Protocols

Lippé, C., Dumont, P., Bernatchez, I. (2006) High genetic diversity and no inbreeding in the endangered copper redhorse, Moxostoma hubbsi (Catostomidae, Pisces): the positive sides of a long generation time. Molecular Ecology 15, 1769-1780

Lloris, D. (2002) A world overview of species of interest to fisheries. Chapter: Dicentrarchus labrax. Retrieved from www.fao.org. Species Fact Sheets

Marzano, F.N., Corradi, N., Papa, R., Tagliavini, J., Gandolfi, G., (2003) Molecular evidence for introgression and loss of genetic variability in Salmo (trutta) macrostigma as a result of massive restocking of Apennine populations (Northern and Central Italy). Environ. Biol. Fishes 68, 349-356.

Naylor, R., S. L. Williams, and D. R. Strong. (2001) Aquaculture: a gateway for exotic species. Science 294, 1655-1656

Naylor, R.L., Goldburg, R.J., Primavera, J.H. (2000) Effect of aquaculture on world fish supplies. Nature 405, 1017-1024

Nomura, I., (2008) Fisheries management: status and challenges. In: Tsukamoto, K., Kawamura, T., Takeuchi, T., Beard, T.D., Kaiser, M.J. (Eds.), Fisheries for Global Welfare and Environment. TERRAPUB, Tokyo, 1-16

Ogden, R. (2008) Fisheries forensics: the use of DNA tools for improving compliance, traceability and enforcement in the fishing industry. Fish and Fisheries 9, 462-472

Patarnello, T., Bargelloni, L., Caldera, F., Colombo, L. (1993) Mitochondrial DNA sequence variation in the European sea bass Dicentrarchus labrax L. (Serranidae) evidence of differential haplotype distribution in natural and farmed population. Mol. Mar. Biol. Biotechnol. 2, 333-337

Pawson, M.G., Pickett, G.D. (1987) The bass, Dicentrarchus labrax, and management of its fishery in England and Wales. MAFF Directorate of Fisheries Research Lab., leaflet 59, Lowestoft.

Peakall, R., Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295

Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, L., Estoup, A. (2004) Geneclass2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95, 536-539

Primmer, C.R., Merilä, J. (2002) A low rate of cross-species microsatellite amplification success in ranid frogs. Conservation Genetics 3, 445-449

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

Pritchard, J.K., Wen, W. (2003) Documentation for Structure software: Version 2

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

Seeb, L. W., Antonovich, A., Banks, M. A., Beacham, T. D., Bellinger, M. R., Blankenship, S. M., Campbell, M. R., Decovich, N. A., Garza, J. C., Guthrie III, C. M., Lundrigan, T. A., Moran, P., Narum, S. R., Stephenson, J. J., Supernault, K. J., Teel, D. J., Templin, W. D., Wenburg, J. K., Young, S. F., Smith , C. T. (2007) Development of a Standardized DNA Database for Chinook Salmon. Fisheries 32:11, 540-552

Sola, L., De Innocentis, S., Rossi, A. R., Crosetti, D., Scardi, M., Boglione, C., Cataudella S. (1998) Genetic variability and fingerling quality in wild and reared stocks of European sea bass, Dicentrarchus labrax. Cahiers Options Méditerranéennes 34, 273-280

Subasinghe, R., Soto, D., Jia, J., (2009) Global aquaculture and its role in sustainable development. Rev. Aqua. 1, 2-9

Toth, G., Gaspari, Z., Jurka, J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 10, 967-981

Wang, J., Ryman, N., (2001) Genetic effects of multiple generations of supportive breeding. Conserv. Biol. 15, 1619-1631

Ward, R.D. (2006) The importance of identifying spatial population structure in restocking and stock enhancement programmes. Fish. Res. 80, 9-18.

Zanuy, S., Carrillo, M. (1984) La salinite': un moyen pour retarder la ponte du bar. In: Barnabe', G., Billard, R. (Eds.), L'aquaculture du bar. INRA Publication, Paris, 73-80

APPENDIX

VSP1: full list of results for each sample of the group;

| ID | Structure 6 pop. |  | Structure 3 pop. |  |  | Parent. Assign. | Pop. Assign. |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0A_02 | 0,025 | 0,677 | 0,005 | 0,008 | 0,686 | R41xR45 | DeltaPo | 93.10 |
| 0A_03 | 0,017 | 0,683 | 0,004 | 0,008 | 0,6 | R41xR45 | DeltaPo | 92.8 |
| OA_04 | 0,314 | 0,381 | 0,007 | 0,031 | 0, | R26xR25 | DeltaPo | 72.6 |
| 0A_05 | 0,021 | 0,680 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 89.4 |
| 0A_06 | 0,686 | 0,012 | 87 | 0,005 | 0,006 | R50xR67 | Delta | 87.591 |
| 0A_07 |  | 0,013 | 0,688 | 0,00 | 0,0 | R67 | Delta | 91.573 |
| 0A_09 |  | 0,282 | 0, | 0,007 | 0,685 | R41xR25 | Delt | 90.166 |
| 0A_10 |  | 0,094 |  | 0,063 | 0,007 | *1xR67 | DeltaPo | 62.2 |
| 0A_100 | 0, | 0,015 | 0,685 | 0,005 | 0,00 | R50xR67 | DeltaPo | 8. |
| 0A_11 | 0, | 0,016 | 0,685 | 0,006 | 0,00 | 0xR67 | Delt | 64.263 |
| 0A_12 | 0,052 | 0,658 | 0, | 0,00 | 0, | R41xR45 | DeltaPo | 96.748 |
| 0A_13 | 0,084 | 0,610 | 0,014 | 0,656 | 0,041 | R1xR19 | DeltaPo | 58.7 |
| OA_14 | 0, | 0,016 | 0,685 | 0,005 | 0,008 | R50xR67 | DeltaPo | 74.2 |
| 0A_15 | 0, | 0,012 | 0,687 | 0,005 | 0,00 | R5 | DeltaPo | 88.480 |
| OA_16 | 0,014 | 0,685 | 0,003 | 0,005 | 0,6 | R41xR45 | De | 66.607 |
| 0A_17 |  | 0,014 |  | 0,00 | 0,00 | *1xR67 | DeltaPo | 33. |
| 0A_18 |  | 0,037 | 0,018 | 0,007 | 0,677 | R26xR69 | DeltaPo | 91.1 |
| OA_19 | 0, | 0,014 | 0, | 0,005 | 0,00 | *1xR67 | DeltaPo | 70. |
| 0A_20 |  | 0,012 |  | 0,005 | 0,00 | R50xR67 | DeltaPo | 98.413 |
| 0A_21 |  | 0,046 | 0,005 | 0,006 | 0,68 | 6xR30 | De | 2.0 |
| OA_22 |  | 0,015 | 0, | 0,005 | 0,008 | R67 | DeltaPo | 73. |
| 0A_23 | 0,089 | 0,633 | 0,012 | 0,66 | 0,032 | R14xR18 | Laguna | 5.6 |
| OA_24 | 0,0 | 0,680 |  | 0,685 | 0,00 | R12 | Laguna | 64.023 |
| 0A_25 | 0,685 | 0,014 | 0, | 0,006 | 0,007 | 0xR67 | DeltaP | 65.322 |
| 0A_26 | 0,035 | 0,670 | 0, | 0,675 | 0,020 | 3xR12 | DeltaPo | 7.9 |
| 0A_27 | 0, | 0,012 | 0,689 | 0,004 | 0,00 | R50xR67 | DeltaPo | 97.608 |
| 0A_28 |  | 0,096 | 0,007 | 0,018 | 0,6 | R26xR25 | DeltaPo | 91.3 |
| OA_29 | 0,022 | 0,679 | 0,005 | 0,011 | 0,6 | R41xR | Ita | 63.0 |
| 0A_30 | 0,683 | 0,017 | 0, | 0,034 | 0,115 | 67 | DeltaP | 95.745 |
| 0A_31 |  | 0,013 |  | 0,005 | 0,005 | 67 | DeltaPo | 9. |
| OA_33 |  | 0,0 | 0,008 | 0,01 | 0,6 | R26xR30 | DeltaPo | 85.3 |
| 0A_34 |  | 0,02 | 0,013 | 0,00 | 0,6 | R26xR30 | DeltaP | 70.8 |
| 0A_35 | 0,107 | 0,588 | 0,018 | 0,670 | 0,0 | xR19 | Laguna | 7. |
| OA_36 | 0,6 | 0,046 | 0,0 | 0,007 | 0, | R26xR30 | DeltaPo | 90.71 |
| 0A_37 | 0,0 | 0,6 | 0, | 0,005 |  | R41xR45 | DeltaPo | 68.137 |
| OA_38 | 0,041 | 0,666 | 0,008 | 0,12 | 0,5 | R41xR69 | DeltaPo | 68.2 |
| OA_39 | 0,0 | 0,681 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 4.4 |
| 0A_40 |  | 0,0 | 0,00 | 0,008 |  | 6xR3 | DeltaP | 8. |
| 0A_44 |  | 0,0 |  | 0,012 | 0,6 | 6xR30 | DeltaPo | 70.94 |
| 0A_50 | 0,6 | 0,029 | 0,009 | 0,007 | 0,68 | R26xR30 | DeltaPo | 89.19 |
| 0A_51 | 0,046 | 0,663 | 0,011 | 0,667 | 0,027 | 19 | DeltaPo | 58.90 |
| 0A_52 | 0,685 | 0,013 |  | 0,004 | 0,005 | 67 | ItaPo | 88.7 |
| 0A_53 | 0,139 | 0,556 | 0,077 | 0,618 | 0,033 | R2xR12 | DeltaPo | 6.4 |
| OA_54 |  | 0,011 |  | 0,004 | 0,0 | *1xR67 | DeltaPo | 81.76 |
| 0A_57 | 0,6 | 0,012 | 0, | 0,004 | 0,005 | R50xR67 | DeltaPo | 91.145 |
| OA_58 | 0,0 | 0,681 | 0,004 | 0,006 | 0,688 | 1xR4 | DeltaPo | 5.0 |
| 0A_59 | 0, | 0,033 | 0,015 | 0,009 | 0,678 | 26xR25 | DeltaPo | 4.5 |
| 0A_60 |  | 0,185 | 0,064 | 0,636 | 0, | *1xR62 | DeltaPo | 79.18 |
| 0A_61 | 0,6 | 0,059 | 0,006 | 0,011 | 0,68 | R26xR30 | DeltaPo | 59.204 |
| 0A_62 | 0,633 | 0,089 | 0,039 | 0,008 | 0,662 | R41xR25 | DeltaPo | 61.22 |
| 0A_63 | 0,685 | 0,013 | 0,689 | 0,004 | 0,004 | *1xR67 | DeltaPo | 89.03 |
| 0A_64 | 0,363 | 0,331 | 0,089 | 0,625 | 0,011 | *1xR62 | Laguna | 54.75 |
| 0A_65 | 0,56 | 0,134 | 0,656 | 0,043 | 0,013 | R242xR67 | DeltaPo | 75.65 |
| 0A_66 | 0,686 | 0,012 | 0,688 | 0,004 | 0,005 | R50xR67 | DeltaPo | 64.450 |
| 0A_67 | 0,686 | 0,012 | 0,687 | 0,004 | 0,006 | R50xR67 | DeltaPo | 88.59 |
| 0A_68 | 0,019 | 0, | 0,004 | 0,006 | 0,6 | R41xR45 | DeltaPo | 94.53 |


| 69 | 0,547 | 0,148 | 0,008 | 0,024 | 0,672 | R26xR25 | DeltaPo | 94 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0A_70 | 0,674 | 0,030 | 0,012 | 0,008 | 0,681 | R26xR30 | DeltaPo | 85.3 |
| 0A_71 | 0,081 | 0,613 | 0,009 | 0,007 | 0,6 | R41xR45 | DeltaP | 8. |
| OA | 0,672 | 0,033 | 0,027 | 0,038 | 0, | R26xR30 | Delta | 54 |
| 0A_73 | 0,080 | 0,639 | 0,041 | 0,65 | 0,01 | R50xR75 | Delta | 4.1 |
| OA_75 | 0,669 | 0,036 | 0,007 | 0,009 | 0,683 | 26xR3 | DeltaP | 4.9 |
| 0A_76 | 0,059 | 0,653 | 0, | 0,616 | 0, | R1xR5 | Laguna | 51.297 |
| OA | 0, | 0,675 | 0, |  | 0, |  | DeltaPo | 52.751 |
| 0A_78 |  | 0,342 | 0, |  | 0,04 | 6xR1 | DeltaP | 76.0 |
| 0A_79 |  | 0, | 0,073 | 0, | 0,2 | R13xR18 | DeltaPo | 94.485 |
| OA | 0, | 0, | 0, | 0,672 | 0,026 | R1xR5 |  | 67.913 |
| OA |  | 0,0 |  | 0,0 | 0,005 | R50xR67 |  | 71.367 |
| 0A_82 | 0,306 | 0,388 | 0,01 | 0,067 |  | 6xR25 |  | 56.068 |
| 0A_83 | 0,662 | 0,0 | 0,006 | 0,007 | 0,685 | R26xR30 | DeltaPo | 95.6 |
| OA |  | 0,0 | 0, | 0, | 0, | *1xR62 | DeltaPo | 88.298 |
| OA |  |  |  | 0,005 | 0,005 |  |  |  |
| OA | 0,0 | 0,6 |  |  |  | R41xR45 |  | 80.736 |
| OA_87 |  | 0,0 | 0,012 | 0,011 | 0,678 | 6xR25 | DeltaPo | 69.102 |
| OA | 0,0 | 0,6 | 0, | 0,35 | 0,33 | $8 \times R 62$ | Laguna | 7.5 |
| OA |  | 0,0 |  | 0,004 | 0,004 | R50xR67 | DeltaPo | 88.534 |
|  |  | 0,0 |  | 0,0 | 0,0 | *1xR67 |  | 89.072 |
| OA |  | 0,0 | 0, | 0,008 | 0,684 | 6xR3 | DeltaPo | 78.906 |
| OA |  | 0,012 | 0,688 | 0, | 0,0 | 67 | DeltaPo | 51.916 |
| OA |  |  |  |  | 0,004 | R50xR67 |  |  |
| OA |  |  |  | 0,0 | 0,0 | R50xR67 |  |  |
| OA |  |  |  | 0,050 |  | 6xR25 |  | 90.934 |
| OA | 0, | 0,5 |  | 0, | 0, | $3 \times R 1$ | Delta | 8. |
| OA |  |  |  | 0,337 | 0,010 | R242xR67 | DeltaPo | 79.239 |
|  |  |  |  |  |  |  |  |  |
| OA_99 |  | 0,6 | 0, |  | 0,0 | R14xR12 |  | 65.940 |
| 1A | 0, | 0,508 | 0, |  | 0, |  |  | 50.526 |
| 1 A |  |  |  |  | 0, | R13xR12 | Delt | 65.926 |
| 1 A |  |  |  | 0,0 |  |  |  |  |
| 1A |  |  |  | 0,037 |  |  |  |  |
| 1A |  | 0, |  | 0,018 |  | R26xR25 | DeltaPo | 1. |
| 1 A |  | 0, | 0, | 0,006 |  | 1 x | elt | 7.92 |
| 1A_0 |  | 0,0 |  | 0,0 | 0,00 | 0xR6 | elta | 95.370 |
| 1 A |  | 0,645 |  | 0,027 |  | R7xR19 |  | 83.947 |
| 1A_09 |  | 0,090 | 0, | 0, |  | 6xR |  | 4.0 |
|  |  |  |  |  | 0,264 | 19 | DeltaPo | 5. |
| 1A |  | 0,5 | 0,0 | 0, | 0,17 | R19 | Delta | 56.9 |
|  |  | 0,012 |  | 0, |  |  | Delta | 95.907 |
| 1A_13 |  |  |  | 8 |  | xR |  | 90.2 |
|  |  |  |  | 0, |  | R25 |  | 54.540 |
|  | 0,219 | 0, | 0,0 | 0, | 0,0 | xR19 | Laguna | 67. |
|  |  | 0,5 |  | 0,518 | 0,1 | R13xR1 | DeltaPo | 77.461 |
|  |  |  |  |  |  |  |  |  |
| 1A_18 |  | 0,031 |  | 0,008 | 0,683 | 6xR30 | DeltaPo | 4.2 |
| 1A_19 |  | 0, |  | 0,005 | 0,008 |  | eltaPo | 98.327 |
| 1A_20 |  | 0,10 | 0, | 0,0 | 0,6 | 6xR25 | DeltaP | 87.3 |
| 1 A |  | 0,180 |  | 0,519 | 0,0 | R62 | Laguna | 93. |
| 1A_22 |  |  |  | 0,011 |  | 极230 | ItaPo | 6.3 |
| 1A_23 |  | 0,0 |  | 0,010 | 0, | xR30 | Delta | 4. |
| 1A_25 |  | 0,218 |  | 0,072 | 0,018 | 242xR6 | DeltaPo | 51.7 |
| 1A_26 |  | 0,026 | 0,011 | 0,009 | 0, | 6xR25 | DeltaPo | 91.035 |
| 1A_29 |  | 0,036 |  | 0,018 | 0,672 | R30 | aP | 0.8 |
| 1A_30 |  | 0,083 |  | 0,058 | 0, | 0xR67 | DeltaPo | 6.49 |
| 1A_31 | 0,606 | 0,088 | 0,405 | 0,015 | 0,279 | R50xR67 | DeltaPo | 7.08 |
| 1A_32 | 0,022 | 0,679 | 0,004 | 0,006 | 0,688 | $1 \times R 45$ | DeltaPo | 93.66 |
| 1A_33 | 0,044 | 0,664 | 0,012 | 0,631 | 0,079 | R13xR12 | DeltaPo | 8.59 |
| 1A_35 | 0,685 | 0,013 | 0,688 | 0,004 | 0,006 | *1xR67 | Laguna | 54.188 |
| 1A_38 | 0,031 | 0,673 | 0,006 | 0,006 | 0,687 | R41xR45 | DeltaPo | 54.593 |


| 1A_43 | 0,676 | 0,026 | 0,682 | 0,008 | 0,009 | *1xR67 | DeltaPo | 88.697 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1A_44 | 0,683 | 0,016 | 0,685 | 0,006 | 0,008 | *1xR67 | DeltaPo | 99.126 |
| 1A_45 | 0,655 | 0,057 | 0,434 | 0,172 | 0,089 | *1xR62 | DeltaPo | 55.710 |
| 1A_46 | 0,444 | 0,250 | 0,029 | 0,156 | 0,518 | R13xR19 | DeltaPo | 53.624 |
| 1A_47 | 0,035 | 0,670 | 0,019 | 0,670 | 0,016 | R2xR12 | DeltaPo | 56.996 |
| 1A_49 | 0,022 | 0,679 | 0,005 | 0,009 | 0,684 | R41xR45 | DeltaPo | 85.144 |
| 1A_50 | 0,680 | 0,021 | 0,013 | 0,007 | 0,680 | R26xR30 | DeltaPo | 91.732 |
| 1A_51 | 0,035 | 0,670 | 0,007 | 0,007 | 0,685 | R41xR45 | DeltaPo | 86.356 |
| 1A_52 | 0,678 | 0,023 | 0,009 | 0,007 | 0,684 | R26xR30 | DeltaPo | 79.563 |
| 1A_53 | 0,556 | 0,138 | 0,006 | 0,013 | 0,681 | R26xR30 | DeltaPo | 66.698 |
| 1A_54 | 0,682 | 0,018 | 0,684 | 0,007 | 0,009 | R50xR67 | DeltaPo | 92.317 |
| 1A_55 | 0,026 | 0,676 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 68.137 |
| 1A_56 | 0,556 | 0,139 | 0,015 | 0,027 | 0,665 | R26xR69 | DeltaPo | 83.471 |
| 1A_57 | 0,033 | 0,672 | 0,005 | 0,006 | 0,687 | R41xR45 | DeltaPo | 72.006 |
| 1A_58 | 0,685 | 0,014 | 0,688 | 0,005 | 0,005 | *1xR67 | DeltaPo | 87.604 |
| 1A_59 | 0,024 | 0,678 | 0,004 | 0,005 | 0,688 | R41xR45 | DeltaPo | 95.040 |
| 1A_60 | 0,685 | 0,013 | 0,687 | 0,005 | 0,006 | *1xR67 | DeltaPo | 86.618 |
| 1A_61 | 0,037 | 0,669 | 0,006 | 0,007 | 0,686 | R41xR45 | DeltaPo | 61.434 |
| 1A_62 | 0,046 | 0,663 | 0,025 | 0,668 | 0,013 | R7xR19 | Laguna | 60.591 |
| 1A_63 | 0,685 | 0,013 | 0,686 | 0,005 | 0,007 | R50xR67 | DeltaPo | 92.120 |
| 1A_64 | 0,685 | 0,014 | 0,688 | 0,005 | 0,006 | R50xR67 | DeltaPo | 78.604 |
| 1A_66 | 0,052 | 0,658 | 0,005 | 0,006 | 0,687 | R41xR45 | DeltaPo | 91.202 |
| 1A_68 | 0,663 | 0,045 | 0,010 | 0,008 | 0,681 | R26xR30 | DeltaP | 68.379 |
| 1A_69 | 0,093 | 0,630 | 0,007 | 0,181 | 0,509 | R68xR75 | DeltaPo | 63.307 |
| 1A_70 | 0,051 | 0,659 | 0,010 | 0,676 | 0,016 | R1xR19 | Laguna | 63.696 |
| 1A_71 | 0,017 | 0,683 | 0,004 | 0,007 | 0,687 | R41xR45 | DeltaPo | 52.684 |
| 1A_72 | 0,078 | 0,616 | 0,029 | 0,649 | 0,036 | R7xR19 | DeltaPo | 62.726 |
| 1A_73 | 0,032 | 0,672 | 0,010 | 0,619 | 0,098 | R7xR19 | Laguna | 54.003 |
| 1A_74 | 0,685 | 0,013 | 0,688 | 0,005 | 0,005 | *1xR67 | DeltaPo | 70.393 |
| 1A_75 | 0,685 | 0,014 | 0,687 | 0,005 | 0,006 | R50xR67 | DeltaPo | 84.224 |
| 1A_76 | 0,635 | 0,085 | 0,008 | 0,013 | 0,681 | R26xR25 | DeltaPo | 91.989 |
| 1A_77 | 0,683 | 0,016 | 0,684 | 0,006 | 0,009 | R50xR67 | DeltaPo | 96.630 |
| 1A_78 | 0,057 | 0,655 | 0,007 | 0,007 | 0,684 | R41xR45 | DeltaPo | 90.686 |
| 1A_79 | 0,684 | 0,015 | 0,687 | 0,005 | 0,006 | R50xR67 | DeltaPo | 98.226 |
| 1A_80 | 0,684 | 0,015 | 0,687 | 0,005 | 0,006 | R50xR67 | DeltaPo | 99.560 |
| 1A_81 | 0,686 | 0,012 | 0,686 | 0,005 | 0,007 | R50xR67 | DeltaPo | 74.472 |
| 1A_82 | 0,686 | 0,012 | 0,689 | 0,004 | 0,004 | R50xR67 | DeltaPo | 95.370 |
| 1A_83 | 0,685 | 0,014 | 0,686 | 0,005 | 0,008 | *1xR67 | Laguna | 61.963 |
| 1A_84 | 0,093 | 0,630 | 0,007 | 0,005 | 0,686 | R41xR45 | DeltaPo | 79.253 |
| 1A_85 | 0,653 | 0,060 | 0,014 | 0,011 | 0,676 | R26xR25 | DeltaPo | 92.461 |
| 1A_86 | 0,663 | 0,046 | 0,005 | 0,007 | 0,686 | R26xR30 | DeltaPo | 90.711 |
| 1A_87 | 0,081 | 0,614 | 0,036 | 0,611 | 0,084 | R14xR18 | DeltaPo | 58.896 |
| 1A_88 | 0,672 | 0,032 | 0,009 | 0,008 | 0,683 | R26xR25 | DeltaPo | 97.113 |
| 1A_90 | 0,048 | 0,661 | 0,006 | 0,676 | 0,020 | R1xR19 | Laguna | 53.796 |
| 1A_91 | 0,181 | 0,514 | 0,012 | 0,221 | 0,465 | R26xR25 | DeltaPo | 79.522 |
| 1A_92 | 0,085 | 0,635 | 0,005 | 0,667 | 0,036 | R68xR62 | DeltaPo | 81.456 |

VSP2: full list of results for each sample of the group;

| ID | Structure 6 pop. |  | Structure 3 pop. |  |  | Parent. assign. | Pop. assign. |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1B_01 | 0,020 | 0,681 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 90.810 |
| 1B_02 | 0,669 | 0,036 | 0,009 | 0,007 | 0,683 | R26xR30 | DeltaPo | 83.366 |
| 1B_03 | 0,685 | 0,013 | 0,686 | 0,005 | 0,007 | *1xR67 | DeltaPo | 98.508 |
| 1B_04 | 0,246 | 0,449 | 0,025 | 0,354 | 0,323 | R7xR19 | DeltaPo | 55.877 |
| 1B_05 | 0,483 | 0,212 | 0,019 | 0,667 | 0,020 | R13xR18 | DeltaPo | 97.971 |
| 1B_06 | 0,476 | 0,219 | 0,029 | 0,088 | 0,614 | R26xR25 | DeltaPo | 92.687 |
| 1B_07 | 0,675 | 0,028 | 0,010 | 0,009 | 0,681 | R26xR30 | DeltaPo | 81.723 |
| 1B_08 | 0,683 | 0,016 | 0,686 | 0,006 | 0,007 | *1xR67 | DeltaPo | 97.990 |
| 1B_09 | 0,024 | 0,678 | 0,004 | 0,005 | 0,688 | R41xR45 | DeltaPo | 93.387 |
| 1B_12 | 0,557 | 0,138 | 0,019 | 0,667 | 0,020 | R13xR18 | DeltaPo | 98.024 |


| 1B_13 | 0,020 | 0,681 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 71.505 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1B_14 | 0,042 | 0,665 | 0,009 | 0,009 | 0,683 | R41xR45 | Laguna | 54.965 |
| 1B_16 | 0,676 | 0,026 | 0,029 | 0,011 | 0,667 | R26xR25 | DeltaPo | 88.204 |
| 1B_18 | 0,041 | 0,666 | 0,009 | 0,009 | 0,682 | R41xR45 | Laguna | 54.965 |
| 1B_24 | 0,686 | 0,012 | 0,687 | 0,004 | 0,006 | *1xR67 | DeltaPo | 80.032 |
| 1B_25 | 0,677 | 0,025 | 0,030 | 0,011 | 0,666 | R26xR25 | DeltaPo | 88.204 |
| 1B_26 | 0,686 | 0,012 | 0,687 | 0,004 | 0,006 | *1xR67 | DeltaPo | 80.032 |
| 1B_29 | 0,041 | 0,666 | 0,010 | 0,680 | 0,010 | R13xR12 | DeltaPo | 99.043 |
| 1B_33 | 0,020 | 0,681 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 71.505 |
| 1B_42 | 0,226 | 0,469 | 0,025 | 0,109 | 0,568 | R7xR19 | DeltaPo | 91.637 |
| 1B_43 | 0,676 | 0,026 | 0,012 | 0,013 | 0,677 | R26xR30 | DeltaPo | 83.412 |
| 1B_44 | 0,555 | 0,140 | 0,043 | 0,387 | 0,278 | R26xR30 | DeltaPo | 68.790 |
| 1B_46 | 0,684 | 0,015 | 0,686 | 0,005 | 0,007 | *1xR67 | DeltaPo | 97.990 |
| 1B_48 | 0,681 | 0,019 | 0,685 | 0,007 | 0,008 | *1xR67 | DeltaPo | 94.405 |
| 1B_73 | 0,669 | 0,036 | 0,009 | 0,008 | 0,683 | R26xR30 | DeltaPo | 83.366 |
| 1B_81 | 0,681 | 0,020 | 0,681 | 0,007 | 0,012 | *1xR67 | DeltaPo | 96.268 |
| 1B_84 | 0,028 | 0,675 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 95.659 |
| 1B_85 | 0,685 | 0,013 | 0,687 | 0,005 | 0,007 | *1xR67 | DeltaPo | 98.508 |
| 1B_87 | 0,674 | 0,030 | 0,666 | 0,037 | 0,005 | R50xR67 | Laguna | 50.201 |
| 1B_88 | 0,684 | 0,015 | 0,686 | 0,005 | 0,007 | R50xR67 | DeltaPo | 95.734 |
| 1B_94 | 0,677 | 0,025 | 0,028 | 0,011 | 0,667 | R26xR25 | DeltaPo | 88.204 |
| 1B_96 | 0,646 | 0,070 | 0,683 | 0,010 | 0,006 | R242xR67 | DeltaPo | 89.651 |
| 1C_201 | 0,676 | 0,026 | 0,029 | 0,011 | 0,667 | R26xR25 | DeltaPo | 88.204 |
| 1C_202 | 0,042 | 0,665 | 0,005 | 0,005 | 0,688 | R41xR45 | DeltaPo | 71.575 |
| 1C_203 | 0,415 | 0,279 | 0,314 | 0,373 | 0,011 | *1xR75 | DeltaPo | 56.261 |
| 1C_204 | 0,676 | 0,026 | 0,012 | 0,010 | 0,678 | R26xR30 | DeltaPo | 88.707 |
| 1C_205 | 0,577 | 0,117 | 0,123 | 0,567 | 0,007 | R50xR62 | DeltaPo | 83.325 |
| 1C_206 | 0,063 | 0,651 | 0,005 | 0,005 | 0,688 | R41xR45 | DeltaPo | 96.068 |
| 1C_207 | 0,054 | 0,657 | 0,006 | 0,006 | 0,687 | R41xR45 | DeltaPo | 79.430 |
| 1C_209 | 0,032 | 0,672 | 0,005 | 0,685 | 0,009 | R1xR12 | DeltaPo | 61.355 |
| 1C_211 | 0,177 | 0,517 | 0,008 | 0,070 | 0,64 | R26xR25 | DeltaPo | 83.190 |
| 1C_212 | 0,117 | 0,577 | 0,005 | 0,005 | 0,688 | R41xR45 | DeltaPo | 95.660 |
| 1C_213 | 0,615 | 0,079 | 0,084 | 0,599 | 0,017 | *1xR62 | Laguna | 76.436 |
| 1C_214 | 0,059 | 0,653 | 0,017 | 0,666 | 0,024 | R1xR69 | Laguna | 57.917 |
| 1C_215 | 0,101 | 0,594 | 0,007 | 0,017 | 0,678 | R41xR5 | DeltaPo | 82.993 |
| 1C_216 | 0,072 | 0,644 | 0,014 | 0,656 | 0,041 | R2xR25 | DeltaPo | 94.802 |
| 1C_217 | 0,576 | 0,119 | 0,661 | 0,030 | 0,018 | R242xR67 | DeltaPo | 53.197 |
| 1C_218 | 0,032 | 0,672 | 0,007 | 0,674 | 0,023 | R14xR18 | DeltaPo | 78.516 |
| 1C_219 | 0,069 | 0,647 | 0,027 | 0,668 | 0,012 | R1xR12 | Laguna | 73.726 |
| 1C_220 | 0,684 | 0,015 | 0,685 | 0,005 | 0,008 | *1xR67 | DeltaPo | 93.699 |
| 1C_221 | 0,672 | 0,033 | 0,032 | 0,017 | 0,660 | R26xR30 | DeltaPo | 81.316 |
| 1C_222 | 0,653 | 0,060 | 0,636 | 0,057 | 0,027 | *1xR67 | DeltaPo | 60.881 |
| 1C_223 | 0,674 | 0,030 | 0,442 | 0,109 | 0,143 | *1xR67 | DeltaPo | 94.177 |
| 1C_224 | 0,047 | 0,662 | 0,013 | 0,079 | 0,606 | R41xR45 | DeltaPo | 94.171 |
| 1C_225 | 0,684 | 0,015 | 0,687 | 0,006 | 0,005 | *1xR67 | DeltaPo | 54.476 |
| 1C_226 | 0,610 | 0,084 | 0,574 | 0,094 | 0,039 | *1xR61 | DeltaPo | 80.433 |
| 1C_227 | 0,038 | 0,668 | 0,016 | 0,672 | 0,016 | R2xR12 | DeltaPo | 68.151 |
| 1C_228 | 0,678 | 0,024 | 0,680 | 0,007 | 0,014 | *1xR67 | Laguna | 58.984 |
| 1C_229 | 0,679 | 0,022 | 0,023 | 0,010 | 0,672 | R26xR30 | DeltaPo | 87.428 |
| 1C_232 | 0,685 | 0,013 | 0,688 | 0,004 | 0,005 | *1xR67 | Laguna | 51.512 |
| 1C_233 | 0,632 | 0,090 | 0,151 | 0,532 | 0,017 | *1xR62 | Laguna | 78.559 |
| 1C_234 | 0,675 | 0,028 | 0,015 | 0,008 | 0,678 | R26xR30 | DeltaPo | 89.779 |
| 1C_236 | 0,685 | 0,013 | 0,687 | 0,006 | 0,005 | *1xR67 | DeltaPo | 71.996 |
| 1C_238 | 0,030 | 0,674 | 0,008 | 0,668 | 0,030 | R14xR19 | DeltaPo | 72.307 |
| 1C_239 | 0,526 | 0,169 | 0,567 | 0,122 | 0,009 | R242xR67 | DeltaPo | 99.491 |
| 1C_240 | 0,686 | 0,012 | 0,688 | 0,004 | 0,006 | R50xR67 | DeltaPo | 92.482 |
| 1C_241 | 0,684 | 0,015 | 0,686 | 0,006 | 0,007 | R50xR67 | DeltaPo | 80.247 |
| 1C_242 | 0,678 | 0,023 | 0,016 | 0,009 | 0,677 | R26xR30 | DeltaPo | 78.656 |
| 1C_243 | 0,050 | 0,660 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 93.694 |
| 1C_244 | 0,665 | 0,042 | 0,015 | 0,013 | 0,675 | R26xR30 | DeltaPo | 63.686 |
| 1C_245 | 0,606 | 0,089 | 0,011 | 0,018 | 0,674 | R26xR25 | DeltaPo | 88.930 |
| 1C_246 | 0,063 | 0,651 | 0,005 | 0,005 | 0,688 | R41xR45 | DeltaPo | 96.068 |


| 1C_247 | 0,023 | 0,678 | 0,008 | 0,684 | 0,008 | R2xR12 | DeltaPo | 85.184 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1C_248 | 0,676 | 0,026 | 0,011 | 0,008 | 0,681 | R26xR30 | DeltaPo | 92.451 |
| 1C_249 | 0,681 | 0,020 | 0,681 | 0,007 | 0,012 | *1xR67 | DeltaPo | 96.268 |
| 1C_250 | 0,484 | 0,210 | 0,019 | 0,667 | 0,020 | R13xR18 | DeltaPo | 97.971 |
| 1C_251 | 0,053 | 0,658 | 0,006 | 0,006 | 0,687 | R41xR45 | DeltaPo | 79.430 |
| 1C_252 | 0,685 | 0,014 | 0,686 | 0,005 | 0,007 | *1xR67 | Laguna | 54.958 |
| 1C_253 | 0,685 | 0,013 | 0,687 | 0,005 | 0,006 | R50xR67 | DeltaPo | 91.090 |
| 1C_254 | 0,513 | 0,182 | 0,278 | 0,401 | 0,022 | *1xR62 | Laguna | 53.960 |
| 1C_256 | 0,053 | 0,658 | 0,007 | 0,007 | 0,68 | R41xR45 | DeltaPo | 91.787 |
| 1C_258 | 0,663 | 0,046 | 0,020 | 0,014 | 0,671 | R26xR25 | DeltaPo | 76.975 |
| 1C_259 | 0,688 | 0,010 | 0,688 | 0,004 | 0,005 | R50xR67 | DeltaPo | 97.425 |
| 1C_261 | 0,036 | 0,669 | 0,014 | 0,676 | 0,011 | R14xR18 | DeltaPo | 96.237 |
| 1C_262 | 0,683 | 0,016 | 0,687 | 0,006 | 0,006 | *1xR67 | Delta | 86.245 |
| 1C_263 | 0,018 | 0,682 | 0,004 | 0,007 | 0,687 | R41xR45 | DeltaPo | 94.678 |
| 1C_264 | 0,028 | 0,675 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 95.659 |
| 1C_265 | 0,683 | 0,016 | 0,685 | 0,006 | 0,008 | *1xR67 | Laguna | 83.342 |
| 1C_266 | 0,041 | 0,666 | 0,004 | 0,006 | 0,68 | R41xR45 | DeltaPo | 96.100 |
| 1C_267 | 0,636 | 0,084 | 0,679 | 0,013 | 0,009 | R242xR6 | DeltaPo | 75.528 |
| 1C_268 | 0,028 | 0,675 | 0,005 | 0,007 | 0,687 | R41xR45 | DeltaPo | 91.365 |
| 1C_269 | 0,168 | 0,526 | 0,009 | 0,676 | 0,018 | R13xR18 | DeltaPo | 99.965 |
| 1C_271 | 0,683 | 0,017 | 0,684 | 0,006 | 0,009 | R50xR67 | DeltaPo | 94.594 |
| 1C_272 | 0,597 | 0,098 | 0,169 | 0,518 | 0,011 | *1xR62 | DeltaP | 74.658 |
| 1C_273 | 0,235 | 0,459 | 0,009 | 0,110 | 0,579 | R26xR25 | DeltaPo | 55.896 |
| 1C_274 | 0,414 | 0,281 | 0,314 | 0,373 | 0,011 | *1xR75 | DeltaPo | 56.261 |
| 1C_275 | 0,683 | 0,016 | 0,685 | 0,005 | 0,008 | *1xR67 | Laguna | 72.784 |
| 1C_276 | 0,681 | 0,020 | 0,685 | 0,007 | 0,006 | *1xR67 | DeltaPo | 63.054 |
| 1C_278 | 0,176 | 0,518 | 0,008 | 0,068 | 0,641 | R26xR25 | DeltaPo | 83.190 |
| 1C_279 | 0,272 | 0,423 | 0,006 | 0,038 | 0,6 | R26xR25 | DeltaPo | 81.044 |
| 1C_280 | 0,672 | 0,033 | 0,014 | 0,021 | 0,670 | R26xR30 | DeltaPo | 71.887 |
| 1C_283 | 0,683 | 0,017 | 0,685 | 0,005 | 0,007 | *1xR67 | Laguna | 55.005 |
| 1C_284 | 0,598 | 0,097 | 0,014 | 0,006 | 0,680 | R41xR25 | DeltaPo | 88.888 |
| 1C_285 | 0,041 | 0,666 | 0,008 | 0,006 | 0,685 | R41xR45 | DeltaPo | 86.450 |
| 1C_286 | 0,026 | 0,676 | 0,005 | 0,663 | 0,040 | R13xR12 | DeltaPo | 73.861 |
| 1C_288 | 0,686 | 0,012 | 0,688 | 0,004 | 0,005 | R50xR67 | DeltaPo | 71.783 |
| 1C_289 | 0,028 | 0,675 | 0,004 | 0,006 | 0,688 | R41xR45 | DeltaPo | 95.686 |
| 1C_292 | 0,026 | 0,676 | 0,009 | 0,680 | 0,012 | R7xR12 | DeltaPo | 64.570 |
| 1C_294 | 0,032 | 0,672 | 0,005 | 0,685 | 0,009 | R1xR12 | DeltaPo | 61.355 |
| 1C_295 | 0,678 | 0,024 | 0,678 | 0,009 | 0,015 | R50xR67 | DeltaPo | 93.047 |
| 1C_297 | 0,021 | 0,680 | 0,004 | 0,007 | 0,687 | R41xR45 | DeltaPo | 93.721 |
| 1C_299 | 0,681 | 0,019 | 0,015 | 0,007 | 0,679 | R26xR30 | DeltaPo | 95.516 |
| 1C_300 | 0,197 | 0,498 | 0,025 | 0,572 | 0,105 | R26xR69 | DeltaPo | 76.774 |

Colony 2.0: Results from the parentage assignment. Families are sorted in alphabetical order.

|  | VSP1 |  |  | VSP2 |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Fam |  | \% | Fam |  | \% |
| *1R62 | 4 | 2,44 | *1R61 | 1 | 0,88 |
| *1R67 | 22 | 13,41 | *1R62 | 4 | 3,54 |
| R13R12 | 4 | 2,44 | *1R67 | 22 | 19,47 |
| R13R18 | 2 | 1,22 | *1R75 | 2 | 1,77 |
| R13R19 | 1 | 0,61 | R13R12 | 2 | 1,77 |
| R14R12 | 1 | 0,61 | R13R18 | 4 | 3,54 |
| R14R18 | 2 | 1,22 | R14R18 | 2 | 1,77 |
| R1R12 | 1 | 0,61 | R14R19 | 1 | 0,88 |
| R1R19 | 6 | 3,66 | R1R12 | 3 | 2,65 |
| R1R5 | 3 | 1,83 | R1R69 | 1 | 0,88 |
| R242R67 | 3 | 1,83 | R242R67 | 4 | 3,54 |
| R26R18 | 1 | 0,61 | R26R25 | 11 | 9,73 |
| R26R25 | 16 | 9,76 | R26R30 | 14 | 12,39 |
| R26R30 | 22 | 13,41 | R26R69 | 1 | 0,88 |
| R26R69 | 3 | 1,83 | R2R12 | 2 | 1,77 |
| R2R12 | 2 | 1,22 | R2R25 | 1 | 0,88 |
| R41R25 | 2 | 1,22 | R41R25 | 1 | 0,88 |
| R41R45 | 27 | 16,46 | R41R45 | 23 | 20,35 |
| R41R69 | 1 | 0,61 | R41R5 | 1 | 0,88 |
| R50R62 | 1 | 0,61 | R50R62 | 1 | 0,88 |
| R50R67 | 28 | 17,07 | R50R67 | 9 | 7,96 |
| R50R75 | 1 | 0,61 | R7R12 | 1 | 0,88 |
| R68R62 | 2 | 1,22 | R7R19 | 2 | 1,77 |
| R68R75 | 1 | 0,61 |  |  |  |
| R7R19 | 8 | 4,88 |  |  |  |

## RINGRAZIAMENTI

Ringrazio il Prof. Bargelloni e tutti i ragazzi dei laboratori dell' Agripolis per la compagnia e i consigli dati durante il tirocinio. Un grazie anche al Prof. Zane, Ilaria e Cecilia per avermi aiutato nelle analisi e nella stesura della tesi.

Infine, il ringraziamento più grande va ai miei genitori, per tutte le opportunità che mi hanno offerto e mi offriranno, e a tutti gli amici, che fortunatamente è troppo lungo elencare.

