Università degli Studi di Padova
Dipartimento di Biologia

Corso di Dottorato di Ricerca in BIOSCIENZE E BIOTECNOLOGIE

Curricolo BIOLOGIA EVOLUZIONISTICA

Ciclo XXIX

GENETIC ANALYSES FOR THE RESTORATION OF THE BELUGA STURGEON (*Huso huso*) IN THE PO RIVER BASIN

Analisi genetiche per la reintroduzione dello storione Beluga (*Huso huso*) nel bacino del fiume Po

Coordinatore: Ch.mo Prof. Paolo Bernardi
Supervisore: Ch.mo Prof. Leonardo Congiu

Dottoranda: Chiara Caruso

CONTENTS

Abstract

The Beluga sturgeon (*Huso huso*) is a locally-extinct species in the Adriatic region, mainly due to anthropic factors. In order to successfully re-introduce this species, strong genetic bases are needed to set up a proper reintroduction plan that will lead to achieve a self-sustaining population as genetically similar as possible to those previously living in the Adriatic basin. On this basis, a phylogeographic investigation was performed at a mitochondrial level (for subsequent comparison with museum samples of the Adriatic population) in 119 Beluga specimens from 5 non-extinct populations from Black Sea and Caspian Sea basins. At a nuclear level, 110 specimens were analyzed at microsatellite markers and 54 at SNP loci through genome wide 2b-RAD approach.

Based on genetic analyses of a 367 bp region of the mitochondrial D-loop, of 16 microsatellite loci and of 4736 polymorphic loci from RAD-seq, a significant geographically-related structure between Black Sea and Caspian Sea was observed. These analyses have shown the lack of a clear-cut differentiation of the population from the Sea of Azov from all the other populations, probably as result of the combined effect of the connection with the Black Sea and of the intensive release of individuals originating from the Caspian Sea basin. In this study, two distinct haplotypes of Adriatic Beluga have been identified from two museum specimens of *H. huso*. These haplotypes differed from all the 39 haplotypes identified in contemporary populations. Although the lack of shared haplotypes is not sufficient to demonstrate the existence of a former genetically distinct population in the Adriatic basin, this hypothesis is supported by the fact that the two museum haplotypes observed in three individuals were observed only in the Adriatic and were very similar one to each other.

Two molecular tools have been developed for both the conservation and management of aquaculture stocks: 1) a species-specific nuclear tool for the identification of the Beluga based on a diagnostic SNP identified within a predicted intron of the Ribosomal Protein S7 (RP2S7) gene; 2) a panel of microsatellite loci developed for relatedness analyses.

The identified RP2S6 marker resulted as highly reliable for routine controls in the Beluga sturgeon, given its high efficiency and simplicity based on its single-locus diagnostic approach with a fast and easy visualization by agarose gel from a single PCR. The nuclear RP2S6 marker, in combination with pairs of species-specific primers previously developed for numerous and different sturgeon species on locus RP1S7, guarantees also the identification of various hybrids that include Beluga sturgeon as a maternal or paternal species.

On the other hand, the selected panel of microsatellite loci will be extremely useful to perform relatedness analysis among individuals reared in aquaculture in Italy, and potentially selectable for the establishment of captive broodstocks.

The present study allowed for the first time to depict the geographical pattern of genetic distribution across the entire geographical range of the Beluga.

For what concerns the comparison of museum samples from the Adriatic Sea, the limited amount of information collected did not allow to unambiguously allocate these animals to any of the main basins. For this reason, the choice of the best source population cannot be based only on genetic evaluations, but should also take into consideration also other factors, such as the geographical proximity or the ecological exchangeability. In any case, the genetic differentiation observed between basins raises the opportunity to avoid admixing individuals from the two main basins for the establishment of captive broodstocks as a prudential strategy for outbreeding avoidance.

For what concerns the reintroduction of the Beluga in the Adriatic area, the need of a careful genetic analysis of all the individuals that will be involved in future rehabilitation programs is a clear outcome of the present work. The availability of a genetic tool for the assessment of species purity as well as of relatedness, lays the bases for a properly conducted management of this critically-endangered species in Italy.

Sommario

Lo storione Beluga è una specie di storione ormai estinta nella regione Adriatica. Per riuscire a reintrodurre efficacemente questa specie, è necessario ideare un adeguato piano che abbia solide basi genetiche al fine di ottenere una popolazione autosufficiente e il più possibile geneticamente simile a quella in precedenza presente nel bacino del mar Adriatico.

A tal fine, è stato effettuato uno studio filogeografico a livello mitocondriale su 119 individui di Beluga appartenenti a 5 popolazioni attualmente esistenti nei bacini del Mar Nero e Caspio per poter successivamente effettuare un confronto con campioni museali appartenenti alla popolazione Adriatica.

A livello nucleare, invece, sono stati analizzati 110 individui tramite analisi a livello di loci microsatellite, e 54 individui tramite analisi a livello di loci SNP tramite un approccio genome-wide (2b-RAD).

Attraverso analisi di una regione di 367 bp del D-loop mitocondriale, di 16 loci microsatellite e di 4736 loci polimorfici ottenuti tramite RAD-seq, è stato possibile individuare la presenza di una struttura genetica fra i bacini geograficamente distinti del Mar Nero e Caspio. Le analisi effettuate hanno inoltre evidenziato, per la popolazione del Mar d'Azov, la mancanza di una netta e chiara differenziazione rispetto alle altre popolazioni; ciò può essere spiegato dal fatto che tale popolazione risulta geograficamente connessa al bacino del Mar Nero e, inoltre, è stata oggetto di azioni di rilascio di individui provenienti dal bacino del Mar Caspio.

In questo studio, sono stati inoltre identificati due diversi aplotipi di Beluga Adriatico a partire da due distinti campioni museali di *H.huso*. Tali aplotipi differiscono dagli altri 39 identificati nelle popolazioni attualmente esistenti. Nonostante la mancanza di aplotipi condivisi non sia sufficiente per dimostrare la presenza nel Mar Adriatico di una popolazione geneticamente distinta, questa ipotesi è supportata dal fatto che i due aplotipi identificati siano circoscritti unicamente agli individui Adriatici e risultino essere altamente simili fra di loro.

Due differenti tecniche molecolari sono state sviluppate per il controllo e la conservazione degli stock di acquacoltura: 1) un approccio a livello nucleare per l'identificazione specie-specifica del Beluga basato sulla presenza di uno SNP diagnostico identificato all'interno dell'introne del gene codificante per la proteina ribosomiale S7 (RP2S7); 2) un pannello di loci microsatellite sviluppato per effettuare analisi di parentela.

Il marcatore RP2S7 identificato, grazie alla sua elevata efficienza e semplicità dovute alla metodologia diagnostica a singolo locus tramite una rapida identificazione di un prodotto di PCR su gel di agarosio, risulta essere altamente affidabile in caso di controlli di routine per il Beluga. Il marcatore RP2S7 inoltre, unitamente all'utilizzo di primer specie-specifici per differenti specie di

storione precedentemente identificati sul locus RP1S7, garantisce l'identificazione di numerosi ibridi che includono il Beluga sia come specie materna che paterna.

Il pannello di loci microsatellite identificato risulterà inoltre utile per effettuare analisi di parentela tra gli individui presenti negli impianti di acquacoltura del territorio italiano, i quali costituiscono una possibile fonte per la selezione di potenziali riproduttori.

Il presente studio ha consentito per la prima volta di delineare il pattern di variabilità genetica tra le diverse popolazione dell'intero areale di distribuzione del Beluga. Per quanto riguarda il confronto dei campioni museali Adriatici, le poche informazioni ottenute non hanno permesso di allocare con certezza tali individui a uno dei principali bacini. Per questo motivo l'identificazione di una popolazione sorgente ideale non può essere unicamente basata su informazioni di tipo genetico, ma deve piuttosto tenere conto di differenti fattori, come la prossimità geografica o la similarità ecologica.

In ogni caso il livello di differenziamento genetico osservato fra bacini pone le basi per evitare il mescolamento di individui appartenenti ai due diversi bacini principali negli stock riproduttivi in cattività, prudente strategia per evitare outbreeding.

Per quanto riguarda la reintroduzione del Beluga nell'area dell'Adriatico, la necessità di un'attenta analisi genetica di tutti gli individui che verranno impiegati nei futuri piani di rilascio è un chiaro risultato del presente lavoro.

La disponibilità di tecniche genetiche per l'accertamento della purezza della specie e per le analisi di parentela getta le basi per una corretta gestione di questa specie in pericolo critico nel territorio italiano.

1. INTRODUCTION

1.1. Preamble

Sturgeons represent an important natural resource with evolutionary, economic and conservationist importance that is nowadays exposed to a critical risk of extinction in most extant natural populations. The Beluga sturgeon (*Huso huso*) is one of three species historically resident in Italy, together with *Acipenser sturio* and *A. naccarii*. The Italian population of Beluga was present only in the Adriatic Sea (with spawning migration in the Po River) (Rossi et al., 1991, Vecsei et al., 2002). Worldwide there are nowadays only three known populations of *H. huso* in nature: in the Black Sea (with migration towards the Danube River), in the Caspian Sea (with migration towards the Volga and Ural rivers) and in the Sea of Azov, where the stock survival is guaranteed only by the continuous reintroductions of juveniles obtained by controlled reproductions (Birstein, 1993; Vecsei et al., 2002; Gesner et al., 2010). In the immediate future, the survival of the Beluga is strictly dependent on restocking actions, effective fisheries management and control of illegal fishing. For what concerns the Italian population, the possibility to re-introduce this species will depend on the ability of setting up a project based on a two-sided approach. On one side, the direct translocation to the Po river of animals from areas where the species is still surviving, on the other side, the establishment of an *ex situ* program aimed to rear local stocks of future breeders. Therefore, from the genetic point of view, it is necessary to perform a preliminary analysis aimed to obtain all possible information relevant to the above actions. In order to establish a reintroduction plan with the highest chance of (long term) success for the operation, it is necessary to follow specific criteria with genetic, geographical and/or bio-ecological basis. According to the guidelines for reintroduction drafted by the International Union for Conservation of Nature (IUCN), it is preferable that reintroduced individuals are as genetically similar as possible to those previously living in the target region. In particular, regarding sturgeons, any future practice aimed at preservation of the species must be supported by detailed genetic analyses (The Ramsar Declaration on Global Sturgeon Conservation; Rosenthal et al., 2006).

1.2. Biological Interest

Sturgeons (order: Acipenseriformes, infraclass: Chondrostei) belongs to a very ancient fish group distributed in the Palearctic hemisphere with about 25 species, most of which are on the brink of extinction (Congiu et al., 2011). From a biological standpoint, sturgeons represent an interesting group of species with archaic traits considered as a survivor group that, in past ages, had a considerable expansion. Often referred to as living fossils, sturgeons underwent a very ancient separation from teleosts, which occurred over 250 Mya (William Bemis, 2001), placing them in a key phylogenetic position for evolutionary studies on vertebrates. The study of genetic variability in sturgeons represents an important step towards the understanding of genetic and evolutionary dynamics for such species. The genetic characterization will give a clearer picture of the biology of this species, laying the ground for future conservation plans of this critically endangered species.

1.3. Acipenseriformes

PHYLUM: CHORDATA
SUBPHYLUM: VERTEBRATA
CLASS: ACTINOPTERYGII
SUPERORDER: CHONDROSTEI
ORDER: ACIPENSERIFORMES
FAMILY: ACIPENSERIDAE
SUBFAMILY:
SCAPHIRHYNCHINAE GENERA: Scaphirhynchus; Pseudoscaphirhynchus
ACIPENSERINAE GENERA: Acipenser; Huso

Acipenseriformes represent the only currently living order of Chondrostei superorder, which is the most ancient taxon of Actinopterygii survived until the present. Sturgeons underwent remarkably little morphological change, indicating that their evolution has been exceptionally slow, earning them the informal status of living fossils (Krieger and Fuerst, 2002). Species belonging to this order share numerous archaic morphological traits including: a prevalently cartilaginous postcranial skeleton, the persistence of a notochord surrounded by acentric vertebrae, a cranial vault, a heterocercal caudal fin and five longitudinal series of bony scutes (one along the spine, two dorso-lateral and two ventro-lateral). The origin of this order is supposed to date back to at least 200 mya, given that fossils of Acipenseriformes have been found in deposits from early Jurassic in Western Europe and from late Jurassic in China (Bemis e Kynard, 1997; Billard and Lecointre, 2001; Krieger et al., 2008).

In its current classification, the order includes two living families: Polyodontidae (commonly known as paddlefish), which comprises two species (*Polyodon spathula* in North America and *Psephurus gladius* in China), and the family of Acipenseridae, including the sub-family of Acipenserinae (genera Acipenser and Huso) and Scaphirhynchinae (genera Scaphirhyncus and Pseudoscaphirhyncus).

A recent molecular dating placed the separation between Polyodontidae and Acipenseridae at 184 mya, in accordance with paleontological data (Peng et al., 2007). Acipenseridae diversification the occurred in four main regions (Ponto-Caspian region, China and western America, Atlantic area, North-eastern America) according to Artyukhin (1995), or alternatively in nine biogeographic regions according to Bemis and Kynard (1997). The Ponto-Caspian region is currently the richest in terms of species abundance, probably because of the formation of different hydrological basins

(Black Sea, Caspian Sea, Aral Sea) 120 mya that allowed the diversification of numerous species (Birstein et al., 1997; Bemis e Kynard, 1997; Peng et al., 2007).

In the present, Acipenseriformes exclusively live in seas or freshwater basins in the northern hemisphere, with spawning areas in freshwaters and with diadromous or potamodromous reproductive migrations (Bemis e Kynard, 1997). Species belonging to Acipenseridae can be distinguished on the basis of the number of chromosomes in two main groups: one with about 100-120 chromosomes and the other 240-270, considered as functional diploid and tetraploid respectively (Fontana et al., 2001; Havelka et al., 2013).

The most accepted phylogenetic hypothesis suggests that the diversification of Acipenseriformes took place from an event of polyploidization (4n = 120 chromosomes), followed by a diploidization prior to the radiation of this order (Birstein et al., 1997; Fontana et al., 2001; Peng et al., 2007; Havelka et al., 2013). The distribution of tetraploid species suggests that different and separated events of polyploidization have happened in the evolutionary history of the group, probably favored by the very frequent hybridization rate between species sharing the same reproductive areas (Birstein et al., 1997; Billard and Lecointre, 2001).

From a conservationist point of view, according to IUCN (International Union for Conservation of Nature), sturgeons represent the animal group at the highest risk of extinction in the Red List, with the 63% of species classified as “critically endangered”, the highest level of risk, and with four species already considered as “possibly extinct”.

1.4. The Beluga Sturgeon

Systematic classification:

PHYLUM: CHORDATA

SUBPHYLUM: VERTEBRATA

CLASS: ACTINOPTERYGII

SUPERORDER: CHONDROSTEI

ORDER: ACIPENSERIFORMES

FAMILY: ACIPENSERIDAE

SUBFAMILY: ACIPENSERINAE

GENUS: HUSO

Figure 1: Representation of an adult Beluga sturgeon (from International Commission for the Protection of the Danube River).

According to Berg (1948), the genus Huso can be distinguished from other sturgeons by possession of branchiostegal membranes that are joined and form a flap. Moreover, the two species in the genus possess a relatively large crescent shaped mouth. In order to distinguish between *Huso* congeners, it is known that in *H. huso* the first dorsal scute is the smallest; barbels have foliate appendages and the dorsal fin usually has not less than 60 rays.

1.4.1. Status of the species in Italy

The Beluga sturgeon (*Huso huso*) is one of the three species of sturgeon historically present in Italy. In particular, the Italian population was present only in the Adriatic Sea (with spawning migration in the Po River) (Rossi et al., 1991, Vecsei et al., 2002). The distribution of the extinct Adriatic Sea population included: Croatia, Greece, Italy and Slovenia.

Unfortunately, accurate data on the frequency of this sturgeon in Italian waters does not exist. The main information on its distribution comes from few historical written reports. It seems that the past populations existed not only at the mouth of the Po River, but perhaps along the river to the height of Papozze or Polesella. Historically, reports in the rivers of the Po-Venetian basin are considered occasional, but during the second half of the nineteenth century and early twentieth century the Beluga sturgeon was common in the Adriatic Sea and several authors reported its presence near to Venice, where the biggest individuals were called "sturiòn disarmà". According to Maglio (1901), some specimens were caught at the beginning of the '900 in the Emilian and Pavia Po River. It is not known whether the Adriatic *H. huso* constituted either a large population or it was separated from the eastern ones. This population has been considered endemically extinct in Italy for at least thirty years (IUCN: Sturgeon Specialist Group. 1996; Bianco, 2014): indeed, since the 80s there have been no more sightings or catches of specimens of this species, but, given that fishing, trade and detention of *H. huso* have been forbidden in the Italian territory since 1980, it can be assumed that any subsequent catch might have been hidden. Since the early 20s the decline of sturgeons in Italy, and of the Beluga sturgeon in particular, is evident; in addition, the very low catches between 1972 and 1975 underline a strong decline of the species in this area (Rossi et al., 1991).

1.4.2. Distribution area and population trend

Nowadays, the Beluga sturgeon is naturally present only in three basins: in the Black Sea (with migration towards the Danube), in the Caspian Sea (with migration towards the Volga and Ural) and in the Sea of Azov (with migration towards the Don and the Kuban) where the stock survival is guaranteed only thanks to the continuous reintroductions of juveniles from breeding plants (Birstein, 1993; Vecsei et al., 2002; Gesner et al., 2010). Currently some *H. huso* are reared in northern Italy, but they are individuals originating from other areas and they do not descend from the Italian extinct population (Porcellotti, 2001; Zerunian, 2003). According to the IUCN Red List, *H. huso* is currently considered critically endangered (CR) and numerically declining in all existing populations.

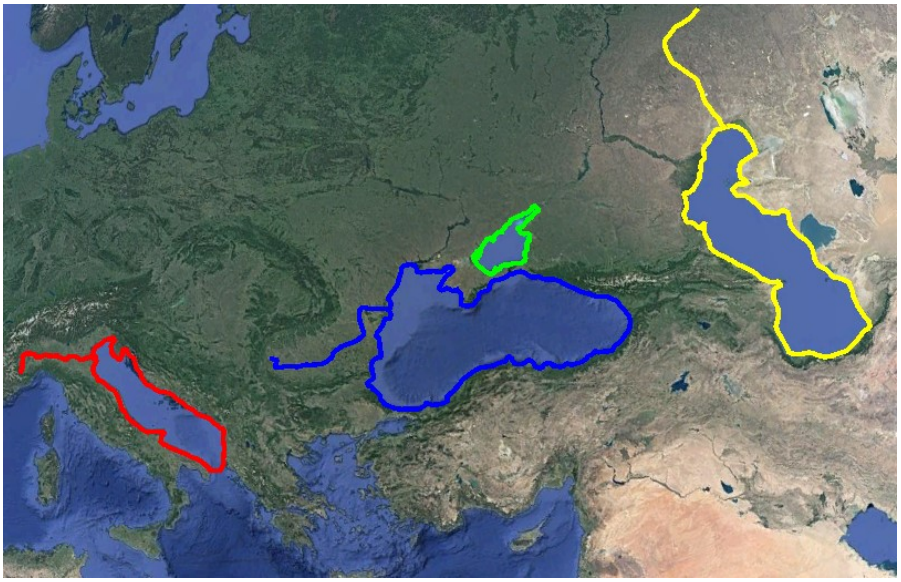


Figure 2- The habitat of extant natural Beluga populations (Blue: Black Sea and Danube River; Green: Sea of Azov; Yellow: Caspian Sea and Ural River), and habitat of the extinct Italian population (Red: Adriatic Sea and Po River).

Unfortunately, information about the total population size are lacking in any basin. The following information about the global trend of populations are taken from the IUCN Red List of Threatened Species by Gesner and colleagues (2010):

The global fisheries statistics available and the other statistics on different basins clearly show a decline in the catch of wild Beluga.

Table 1 - Global catch of Beluga:

Year	***Tonnes***	***Decrease***	***Reference***
1992	520	-93%	FAO 2009
2007	33		

Table 2 - Catches in the Caspian Sea:

Year	***Tonnes***	***Reference***
1945/55	1380	Doukakis et al., 2010
1956/65	1283	
1966/75	1623	
1976/85	849	
1986/95	506	
1996/2003	60.8	

Catches in the Danube River have also declined (Table 3):

Table 3 - Catches in the Danube River (Mid-Danube and Romania):

	Year	*Average tonnes per year*	*Decrease*	*Reference*
Mid-Danube	1972/76	23	-67%	CITES 2000
	1985/89	7.5		

	Year	*Tonnes*	*Decrease*	*Reference*
Romania	2003	21.3	-60%	Paraschiv et al., 2006
	2005	8.4		
	Year	*Quota achieved*		
	2002	85%		
	2003	84%		
	2004	46%		
	2005	34%		
	2006	Banned of catching		

It is thought that nowadays all the Beluga in the Volga River are hatchery-reared, but there is evidence of spawnings elsewhere in its distribution in the Caspian Sea (Khodorevskaya et al., 2009). Despite intensive restocking in the Caspian Sea (91% of each generation is estimated to come from hatchery stock), the annual catch in the northern Caspian Sea has drastically dropped as testified by average annual catch records estimated every 10 years. The official data show that the species was abundant in 1938 and then stable to the late 1980s, with the major decline starting from 1990 to the present, showing over a 90% decline in the past 60 years (Khodorevskaya et al., 2009). The agreed Beluga catch quota for all the Caspian Sea (2007/8 - 28th session of the FAO Commission) was 99.8 tonnes; this quota was not achieved. During trawl surveys, the catch of Beluga in the northern Caspian Sea did not exceed 31 specimens per year (Khodorevskaya et al., 2009).

The number of Beluga annually entering the spawning rivers in the Caspian Sea has decreased as shown by data about migrations (Table 4):

Table 4 – Migration of Beluga individuals in the Caspian Sea (Volga and Ural River).

	Year	*Individuals*	*Decrease*	*Reference*
Volga River	1961/65	26000	-89%	Khodorevskaya et al., 2009 Khodorevskaya et al., 2000
	1996/97	1800		
	1998/2002	2800		
Ural River	2002	2500		Pikitch et al., 2005

Furthermore, it is estimated that, between 1979 and 1981, all the 551.000 individuals living in the Sea of Azov were released from stocks, and consisted of sub-adult and juveniles; despite banning of commercial fishing of Beluga in 1986, in 1988-1993 they dropped to 25.000 and after 1994 they were only caught sporadically. After 1986, the major threat was from bycatch. Since 1994, 98% of recorded individuals in the Sea of Azov have been juveniles. In 2001, the first individuals were produced from captive bred individuals and released (Chebanov and Koziritskaya, 2007).

1.4.3. Biology of the species

Adult Belugas are the biggest among all sturgeons, reaching up to 6 meters of length and 1220 kg of weight; currently, however, specimens that have reached considerable size are very rare in nature. The life span is very long, and the oldest known individual had an estimated age of 110 years (Berg, 1948), but currently the maximum age is 53 years, as reported in 2003 (data from IUCN). Concomitantly, it is estimated that in the Caspian Sea stock there are individuals only up to 280 cm long, weighting up to 650 kg; average length is 240 for females and 220 cm for males, and weight 130 and 65 kg respectively. The life history of this species, characterized by high longevity and advanced age of first reproduction, appears to be quite complex, including several classes of different ages with the potential for migration and reproduction. Fishing and over-exploitation of Beluga, however, led to a drastic reduction of the age structure of the population, with almost all reproductive individuals now represented by first-time spawners (Vecsei et al., 2002).Beluga sturgeon mature later than most other Acipenseriformes: they achieve sexual maturation at the age of 12-16 years for males and 16-22 years for females. Breeding takes place every 4-7 years for males and every 5-7 years for females (Vecsei et al., 2002). Under favorable conditions, a Beluga may spawn up to nine times in its lifetime. Beluga do not spawn every year and females will resorb eggs unless conditions are suitable (Artyukhin et al., 1979). The Beluga is a species with anadromous migration, i.e. adult individuals migrate from salt ponds into rivers to spawn; in particular, the Beluga makes the longest migration among all sturgeon species. Distances for spawning migrations for this species can exceed 2500 km in the Danube and Volga basin (Hensel and Holcik, 1997; Khoderevskaya et al., 1997). There are two biological forms based on the preferred period of migration: one form (vernal form) prefers to migrate during spring, in which case spawning takes place a few weeks after the start of the migration. The second migration form (hiemal form) is more complex, with migration starting in late summer: individuals spend the whole winter in the deeper parts of the river near the site of reproduction, so that in spring it will

be easier for them to reach the site of deposition. Only larger individuals can deal with this type of migration (the so-called two-steps migration) by exploiting their bioenergetic reserves accumulated during the period at the sea (Bemis and Kinard, 1997; Khodorevskaya et al., 2009). These sturgeons typically spawn large numbers of eggs in shallow, gravely sites, and do not provide parental care (Bemis et al. 1997, Lelek 1987). Beluga congregate only during spawning times and seem to be a less gregarious species then other Ponto-Caspian sturgeons (Levin, 1997). During the period of marine life, the adults mainly inhabit the pelagic zone descending at depths of 160-180 m. During both the seaward and the spawning migration, the Beluga usually travels in the deepest parts of the riverbed. Juveniles remain in warmer, shallow habitats during the first year of life (Chebanov et al., 2011). The adults of Beluga are pelagic predators while the young feed mainly on small benthic crustaceans. The main food of juveniles appears to be insect larvae, especially of Ephemeroptera, crustaceans (gammarids, mysids, copepods, and cladocerans). Beluga begins preying on fishes at a very early age (with a length of 24 cm in the lower Danube). Preferred preys such as Alosa spp., *Engraulis encrasicolus*, cyprinids (Cyprinus, Leuciscus, Scardinius, and Aspius), marine fishes, (e.g. *Scomber scombrus*, *Trachurus mediterraneus ponticus* and *Sprattus sprattus*) are important in its diet between May and September, when the Beluga are congregating near the coast prior to entering rivers. During autumn and winter Beluga descend into deep regions of the sea and feed mainly on *Mullus barbatus ponticus*, *Merlangius merlangus euxinus*, *Platichthys flesus flesus* and *Engraulis encrasicolus*. The diet also includes roach (*Rutilus rutilus*), kilka (Clupeonella), crayfish (Astacus), gobies (Gobidae), pike-perch (*Sander lucioperca*), birds, sturgeons and even seal (IUCN Gesner et al., 2010).

1.4.4. Causes of the decline and conservation status

The *H. huso* is considered Critically Endangered (CR) according to the criteria of the IUCN.
The decline of the Beluga turned out to be dramatic, with an overall reduction of 90% of the population in only 3 generations (60 years) (Dudu et al., 2009, Gesner et al., 2010). This severe decline is due to a combination of factors that had a strong impact on the mortality of individuals and on the modification of the habitat of this species. The excessive pollution of fresh waters led to a progressive 'poisoning' of Beluga; for example, reproductive tissues are very sensitive to pollution: indeed, in the gonads of these animals it was found an excessive accumulation of heavy metals with the potential to damage the reproductive potential of individuals (Gesner et al, 2010, Hosseini et al., 2013). Polluted rivers are also unsuitable for reproduction (Rosenthal et al., 1999;

Schram et al., 1999). The reproduction of this species (as of other anadromous species) is blunted by the construction of dams that prevent the access to the highest part of the rivers. An example is the construction of the Danube Iron Gate's dam in the 1970-1984, which completely blocked the access of the sturgeon to the middle and upper river. Furthermore, the construction of the Volgograd's dam in the 60s in the Volga, together with massive fishing activities in the '90s, led to a dramatic reduction of the specimens able to reproduce (Birstein et al., 1997; Ludwig, 2008). In the "Carta ittica del fiume Po" (2008) the construction of the dam of Isola Serafini was explicitly adduced for its role in the extinction of two species of migratory anadromous sturgeon, *A. sturio* and *H. huso*. According to the IUCN, the presence of barriers that prevent the achievement of breeding sites is pointed out as the main cause of the extinction of Beluga in Italy.

Another factor that has severely affected the decrease of adult individuals of Beluga is the over-exploitation caused by fishing (Doukakis et al., 2010). The Beluga, besides being the sturgeon with the maximum tonnage in adulthood, is also the producer of the best quality of caviar (unfertilized eggs) (Fain et al., 2013), for which market demand is very high.

Beluga caviar is among the most sought-after naturally-available commercial products in the world, given its considerable value. Indeed, Beluga caviar can demand up to $12000/kg, and female Beluga can produce up to 12% of their body weight in caviar. Beluga sturgeon is currently at such a high risk of extinction that its survival is supported entirely, or nearly entirely, by hatcheries throughout its limited range (Secor et al., 2000). Nonetheless, population structure, age composition and genetic diversity are all factors negatively affected by the lack of natural recruitment (Secor et al., 2000; Raspopov, 1993).

Historically the larger wild stock of Beluga sturgeon was found in the Caspian Sea, which is fed by over 100 rivers, the most important being the Volga in Russia, which historically supplied 75% of the Caspian Sea's sturgeon catch (Artyukhin 1997, Khodershaya et al. 1997). Illegal catch of sturgeons has been estimated to be ten times greater than the legal catch in the Caspian Sea and Volga River (TRAFFIC 2000). Moreover, reports state that half of the mature Beluga individuals are removed by poachers every year. From December 1995 to December 2000, U.S. Fish and Wildlife Services confiscated 3044.19 kg of illegally-imported eggs of *H. huso* (calculated from USFWS LEMIS system).

Because of the unsustainable demand for caviar, Beluga individuals get caught at ages closer and closer to sexual maturity. At this rate, more than 80% of spawning females are between 17 and 29 years of age, and nearly 90% of males are between 11 and 23 years of age (Raspopov, 1993; Rospopov and Novikova, 1997). The slow maturation rate of sturgeon tied with low numbers of extant populations have placed estimates of population recovery, even with proper management, at more than a century for some species (Secor and Waldman, 1999).

As demonstrated for a wide variety of species (e.g. cod *Gadus morhua*), a strong fishing pressure appears to be unsustainable for the long-term survival of the species (Maroto and Moran, 2014). Simulation studies on demographic trends in relation to overfishing showed that the largest contribution to the increase of the population size is given by the reproductive potential of adult individuals. Given the slow sexual maturation of this species, it would be necessary to extend the period of control on fishing for several decades before achieving a breeding population of individuals composed primarily of adult individuals and not first-time spawners (desirable condition to allow recovery biomass) (Vecsei et al., 2002; Doukakis et al., 2010). Therefore, the survival of the species depends mainly on the management of the stocks, on the control and regulation of fisheries and on the eradication of the widespread illegal fishing.

1.4.5. Population Genetics

Little is known about the genetic diversity of this species. Individuals from different seas differ in morphological, ecological and life-history traits (e.g., age of sexual maturation) (Birstein, 1993). According to these morphological criteria three different subspecies of *H. huso* have been described (Pirogovskii et al., 1989): *H. h. maeoticus* (Sea of Azov), *H. h. ponticus* (Black Sea) and *H. h. caspicus* (Caspian Sea). However, according to recent studies, this distinction is not supported by genetic evidences. In spite of the relevant haplotype diversity across populations, no significant genetic structure among population was observed at mitochondrial level. The taxonomic validity of these subspecies is therefore considered doubtful (Doukakis et al., 2005; Dudu et al., 2014). Some genetic differences between the Black Sea and the Caspian Sea populations were claimed to be detected in the cytochrome b gene (Dudu et al., 2009). However, the very low number of reference individual analyzed (10 Black Sea individuals compared with a single Caspian haplotype from GenBank) makes this observation scarcely reliable. Moreover, at nuclear level, some genetic studies on the Caspian Sea populations variability resulted in the identification of some genetic diversity between the different sampled populations (Pourkazemi, 2008; Ghadirnejad et al., 2008): however, these data referred only to a single basin, which do not allow to infer any conclusion on the population structure of this species from the literature. As a matter of fact, less is known about the population structure. It is assumed even if not verified yet, that females are faithful to the same breeding site, whereas males use to migrate between different breeding sites and populations. This would ultimately lead to a genetic mixing between different populations:

therefore, a single meta-population structure is more plausible than genetically distinct populations.

1.4.6. Past Reintroduction of Beluga sturgeon and Recovery Actions

According to the conservation actions reported in the IUCN Red List of Threatened Species (Gesner et al., 2010), the Beluga sturgeon was listed on CITES Appendix II in 1998. Despite restocking programs are ongoing, natural reproduction are fewer and fewer and populations continue to decline (CITES, 2000); the following information about recovery actions are taken from the IUCN Red List of Threatened Species by Gesner and colleagues (2010):

the annual number of fingerlings released into the Volga is 0.4 million in 1951; 13.1 million between 1966-70 (average per year); 19.4 million between 1981-85; 11.3 million between 1996-2000; and 3 million between 2001-2005 (Khodorevskaya et al. 2009). The 28th Session of the Commission on Aquatic Bioresources of the Caspian Sea reported that the total of Beluga fingerlings released in 2008 was 2.93 million from Russia, Iran, Kazakhstan and the Russian Federation.

Few natural reproductions of this species were recorded in the Volga and Ural Rivers but, at present, the abundance of Beluga is still extremely low. Since 2000 in Russia, catches were allowed only for the purposes of reproduction (for hatcheries) and science. Protective measures at the feeding grounds are necessary to maintain the population of Beluga, as well as the preservation of natural spawning and juveniles breeding at hatcheries.

In the Danube River, the release of recruitment-size individuals (over 15 cm long) from Romania into the Danube has risen from 41.000 during 1998-2005 (information from the Danube Delta National Institute for Research and Development), 12.500 in 2006, to 15.130 in 2007 and 20.000 in 2008 (Suciu, pers. comm.).

It is known that in the past, more than 40% of the released sturgeons (including Beluga sturgeon) from the Sea of Azov hatcheries were of Caspian Sea origin (Tsvetnenko, 1993; Chebanov et al., 2002).

Iran, together with the World Bank, is increasing the release rates of Beluga; in 2003, 6.000 individuals (CWT-tagged and visible-tagged) were released. The size of the released individuals has been increased from 3-5 g to 10-25 g in order to increase survival rates (Pourali et al., 2003). Hatchery experts (which are state-owned) are given financial incentives upon the delivery of

individuals ready to be released. Iran has also developed Beluga farming to supply meat and caviar production to minimize the impact to the wild population (Pourkazemi, pers. comm.).

The extinction of Beluga in the Adriatic region and consequently in the Po river has led to a reduction in the biodiversity of Italian fauna and to a simplification and impoverishment of the fish community of the Po and the Adriatic Sea. The depletion of native species combined with the entry of alien species in Italian reservoirs has led to an imbalance in food webs and biological properties of the Italian territory. In particular, in the Po River (data from 2002 to 2008) there has been a gradual change and depletion of local fish fauna: of the 35 native fish species, 5 has been extirpated, and more than 30 show a general trend towards a decrease in individual densities and are now joined by a large number of non-native species (19), many of which are now naturalized in the ecosystem of the Po river (River Po basin Authority, 2008). The future reintroduction of the Beluga, if successful, could lead to an improvement of the ecosystem and to an increase in biodiversity that is gradually impoverished in this area. It is also conceivable that the Beluga, given its considerable adult size and its nature of pelagic predator (Vecsei et al., 2002), may in some way provide a control for the expansion of non-native species such as wels catfish (*Silurus glanis*). Little is known from the literature, but on the basis of observations made on young individuals released in the Danube, it was seen that this species has a strong predatory attitude towards other fish species, probably including the wels catfish (Suciu R., personal communication).

The Italian population of Beluga, as already mentioned, is totally extinct in the wild and in captivity. Therefore, there are no more representative individuals of the population that lived in the Adriatic Sea. Hence the question arises of which source population to be chosen for reintroduction in this area. Following the directives concerning reintroduction developed by IUCN, it is preferable that reintroduced individuals are as genetically similar as possible to those previously living in the target region. To do this, it is therefore necessary to find samples and/or specimens of Adriatic *H. huso* preserved in museums. From these samples, it is possible to extract DNA and then to make a genetic comparison with now-existing populations of Beluga. Hypothetically, the Adriatic population should be more similar to geographically close populations: based on the reproductive biology of these individuals, it can be speculated that the Italian population descended from individuals migrated from the basin closest to the Adriatic Sea, that is the Black Sea. However, the genetic comparison is essential in order to make a valid choice and in order to maintain the fittest genetic similarity with the original population.

1.5. Conservation Genetics and Reintroduction

The preservation of biodiversity, that is the variety of ecosystems, species, populations within species and genetic diversity within species, is vital for the preservation of options for future evolution (McNeely et al., 1990; Frankham et al., 2010).

Conservation genetics, a scientific discipline originated from the strict contribution of genetics and ecology with conservation biology (Pertoldi et al., 2007; Van Dyke, 2008), applies genetic knowledge to conserve and restore the biodiversity and to reduce the risk of extinction in threatened species.

The understanding of the processes involved in loss of variability, either genetic or environmental, is a crucial point to establish successful strategies in order to reduce the danger of extinction in threatened species.

There are two main factors contributing to loss of variability: i) anthropic factors, that can impact on species survival either directly (fishing, exploitation for commercial purposes and competition with exotic species artificially introduced in the wild) or indirectly (habitat pollution, habitat degradation and fragmentation) and ii) genetic or stochastic factors, involving demographic and environmental fluctuations, in addition to genetic processes that could cause fitness reduction or high mortality (genetic drift, inbreeding) (Shaffer, 1981; Leus, 2011).

Strategies for the reintroduction of species or populations must therefore focused on their preservation, through careful management, as dynamic entities capable to respond to environmental changes and stochastic demographic processes.

In order to monitor and sustain the survival of populations, one of the main barriers in any conservation plan regards the identification and maintenance of distinct populations within species: to this end, genetic data can be useful and efficient in the confirmation process of presumably distinct entities (Avise, 1994; Mortiz, 1994; Vogler and DeSalle, 1994).

Therefore, it is necessary to study the genetic structure of populations, which includes the number of subpopulations in each population, the frequencies of genetic (allelic) variants in each subpopulation and the degree of genetic isolation between subpopulations.

Population structure can be caused by an adaptive response of populations to their environment (local adaptation), or can be a by-product of an event that affected populations (i.e. a vicariance event that divides a species in subpopulations that start to accumulate differences).

The identification of a genetic structure is indicated by a limited genetic flow across subpopulations. Therefore, the degree of differentiation needs to be evaluated on the basis of the

mix-up of genetic information, which is influenced by many factors (social, geographic, ecological, biological barriers or genetic drift), between the distinct entities of a population.

Any modification induced in the genetic flow between two or more subpopulations leads to the beginning of the differentiation process. The differentiation rate depends on mutation rate and breeding size of subpopulations (Nei and Feldman, 1972). The identification of the pattern of genetic variation is therefore essential to evaluate the presence of eventually differentiated units, in order to allow a correct managing and monitoring of these subpopulations from a conservationist point of view and for a correct definition of management units (ESUs, Evolutionary Significant Units) within the species.

When a species, or a population, is almost or is already locally extinct, the only way to reverse this trend is the reintroduction of individuals from different populations or habitats in order to achieve a self-sustaining population. On these bases, reintroduction is defined as the "intentional movement of an organism into a part of its native range from which it has disappeared or become extirpated in historic times" (IUCN, 1987).

Reintroducing a species into its original habitat is not a novel idea (Seddon et al., 2007); however, the concept has become increasingly popular as a conservation strategy only since the 1970s (Stuart, 1991). In the 40+ years of reintroductions as a conservation focus, there have been hundreds of attempts to reintroduce species to their natural habitat. The success of reintroduction programs is still difficult to assess because long-term data, as well as general and accepted success criteria, are required. At present, there is no adequate and shared definition to identify successful reintroduction programs (Doug and Seddon, 2007; Gusset, 2009), but in most empirical surveys on reintroductions, programs are considered successful if they result in self-sustaining populations (Griffith et al., 1989; Fischer and Lindenmayer, 2000). Whereas failures are easy to identify in some cases (e.g., when extinction is documented), the assessment of successful reintroductions may require the use of various criteria (survival or fecundity rates, population trend, spatial expansion, etc.) to determine whether or not the population is self-sustaining.

According to the IUCN (IUCN/SSC, 2013), the first step in planning a successful reintroduction is the implementation of a feasibility study in the preliminary project.

To consider a valid reintroduction program with good prospects of success, it is therefore necessary to consider many factors, as well as planning a long-term project that would provide a viable and self-sustaining population in the future. For example, the effectiveness of conservation plans can be increased through *ex situ* conservation techniques (captive breeding, zoos, aquaria, arboreta, gene and seed banks) prior to the reintroduction phase. Moreover, in addition to the identification of a suitable release site (ideally characterized by: abundant food sources, sufficient area for dispersing individuals, limited or no competition for niche placement) and to the control

and elimination of the causes that led to the extinction of the population (Stuart, 1991), it is necessary to make a careful genetic evaluation in order to choose the most accurate source population from which individuals are selected for reintroduction (IUCN). Therefore, to obtain the highest chance of long-term success for the operation of reintroduction, it is necessary to keep in mind additional factors that could affect the viability of the reintroduced population. On one hand, the stock of individuals chosen as founders for a new reintroduced population must have sufficient genetic variability (Waldman and Wirgin, 1998), this is essential to avoid the occurrence of inbreeding depression (Pikitch et al., 2005; Ludwig, 2006). On the other hand, this variability must nevertheless fall within the range of the population source, in order to avoid the risk of outbreeding depression. Furthermore, in the case that natural populations are diversified among them, the choice of individuals too dissimilar from the original population could be assimilated to the introduction of a non-native species in the territory, and would finally frustrate the objectives and the purpose of the reintroduction plan, with possible consequences on a larger scale within the ecosystem (Holčík, 2000; Edamands, 2007).

In order to maximize the chances of a successful reintroduction with the consequent reduction of the risk related to the movement of animals from different populations, it is therefore possible to take advantage of museum specimens for genetic analyses with the aim of establishing evolutionary relationships between former-living natural populations and candidates for reintroduction (Roy et al., 1994).

«What we're able to do with aDNA is really look at evolution» A. Cooper

Indeed, numerous individuals of known geographic origin are conserved in museum collections. The vast majority of these collections comes from the last two centuries, therefore the available specimens include most species currently at risk of extinction, or recently extinct. This means that in most cases, museum collections preserve a proper patrimony of lost genetic diversity, which is representative of genetic patterns of populations preceding any anthropic influence (Ceballos and Ehrlich, 2002; Paxinos et al., 2002; Walander et al., 2007).

In the case that an ideal source population could not be identified on the bases of genetic criteria, because a clear-cut distinction among populations is lacking, then it should be appropriate to rely on alternative criteria. One of the possible approaches used is to select in any case a putative "best source population" based, for example, on ecological or geographical criteria and to avoid to mix different populations with the aim of preventing the outbreeding depression effects. Alternatively, especially when the environment has changed since the time of the extinction of the former local population, a sound strategy could be to maximize the amount of diversity (genetic, behavioural, ecological, etc.), gathering animals from the different available populations. In this way, if some of the animals will not be pre-adapted to the environment they will be simply negatively selected.

One major problem of outbreeding is usually given by the potential disruption of adaptive traits of the resident population but, in the absence of a resident pre-adapted population, this aspect can be neglected. A second important drawback of outbreeding is that some reproductive incompatibilities might affect the reproductive potential of the population (Svärdson, 1970; Templeton, 1986).

Moreover, it is also necessary to assess the impact that the species could have on the community in which it will be reintroduced, especially in the case of changes occurred in the meantime within the ecosystem (other extinct species, entry of alien species, etc.).

For the purposes of conservation biology, it is also extremely important to distinguish hybrids from pure-breed animals (Wolf et al., 2001). Sturgeons are a perfect example of frequent hybridization between species leading to fertile breeders (Birstein, 2002). The introduction of exotic species or hybrids in a natural environment might have a deleterious impact on natural populations at risk of extinction: for example, it is possible that specific adaptations of the pure species might be lost through hybridization. The deleterious effects of restocking programs have been proven in numerous and different taxa (e.g.: stock enhancement programs for freshwater fishes such as trout or salmon release fry produced in hatcheries - Rhymer and Simberloff, 1996). In addition to the negative ecological consequences, also the loss of genetic peculiarities that makes genetic-based species recognition more difficult must be considered (Ludwig, 2006).

Special attention should be made in case of *ex situ* conservation: hybrids are often accidentally mixed to captive stocks of pure species used for the production of fingerlings to be released, (Congiu et al 2011, Boscari et al 2014 a) as the hatcheries that provide animals for restocking are often also production facilities with commercial purposes (Chebanov et al., 2011). A clear example of careless management is the one reported by Congiu and colleagues (2014), in which interspecific hybrids between the Adriatic sturgeon (*A. naccarii*) and the white sturgeon (*A. transmontanus*) were unexpectedly detected in a captive broodstock of pure Adriatic sturgeons to be used for *ex situ* conservation. Several other examples of accidental introgression of non-native genomes into sturgeon populations are available, e.g. the hybridization of Siberian sturgeon (*A. baerii*) with a native population of Sterlet sturgeon (*A. ruthenus*) in the Danube River (Ludwig et al., 2009). In this case, the hybridization might have taken place through natural hybridization after the careless release of Siberian sturgeon in the Danube basin. Whatever the mechanism of genetic transfer among species, either through the release of alien species followed by hybridization in the wild, or through the artificial hybridization in captivity followed by escape of the produced hybrids, the ability to detect hybrids and to remove them from natural populations is a major concern for conservation.

1.5.1. Sturgeons and reintroductions

Reintroduction is a very sensitive topic, especially regarding sturgeons. Over the years, numerous reintroductions have been performed in several countries, but few of them followed rigorous programs and were based on genetic studies considering the impact on the ecosystem (Ludwig, 2006). Due to the high economic value of these species and their products, over the years non-native species have been released in the wild for the sole purpose of promoting activities such as sport fishing or restocking for the exploitation of caviar. These reintroductions have caused extensive damages to native populations of sturgeons (Fontana et al., 2001): many species of sturgeon may in fact hybridize to each other (Birstein, 2002). The introduction of individuals from a distinct population of the same species, in an ecosystem with a resident population, as well as causing an ecological imbalance, has in many cases led to phenomena of hybridization between different populations. An example is the introduction (between 1960 and 2000) of specimens of *H. huso*, *A. gueldenstaedtii* and *A. stellatus* in the Caspian Sea: because of these erroneous reintroductions, nowadays the genotypic structures of native populations of sturgeon are markedly changed (Ludwig, 2006).

However, there are some examples of reintroductions properly conducted that led to the recovery of the extinct endemic populations. One example concerns *A. fulvescens* (Schram et al., 1999): this species was endemically extinguished by the St. Louis River from the beginning of the twentieth century due to overfishing, habitat alteration and pollution of the river. Since 1983, thanks to the improvement of the ecological conditions and the relaxation of fishing, a reintroduction project was started in the areas where the species disappeared. This project has reached the goal of bringing back the sturgeon in the waters of the St. Louis River, allowing the population to increase in numbers and thus increasing the chances of long-term survival. Based on these results, it was pointed out that the stocking of age groups over 20 years and the maintenance of a reproduction area upstream of the river are essential to the success of the project. In addition, genetic studies on the reintroduced population have shown that the genetic variability appears to be comparable with the source population (stock of Lake Winnebago), thus emphasizing the success of the reintroduction plan (Drauch and Rhodes, 2007). Moreover, the authors defined as valid a reintroduction program if it relies upon two main criteria. Firstly, they recommend implementing a plan that will last several years and, secondly, this plan should minimize the risks of a bottleneck in case of a reintroduction of a small population at a single moment.

1.6. Aim of the work and project goals

With the final purpose of supporting the reintroduction of an important flag species in the Po River basin, an area from which it was extinct due to human pressure since late 70ies, the present project aimed at conducting all the necessary preliminary genetic investigations that will serve as solid foundations in the future re-establishment plan.

More in general, the global long-term endpoint of this pilot study will be to rescue the depletion of native species in the Adriatic basin and to control alien species that have dramatically emerged in the last few decades.

Considering that this work represented the first attempt to approach the conservation of these fish and very few is known about its population genetics and the distribution of the remnant genetic variability, three are the main topics developed in this project:

1) Species identification and control of illegal fishing:
    - Search for a species-specific marker for the identification of pure Beluga sturgeon and hybrids.
2) Identification of a putative source population:
    - Genetic comparison between Italian extinct population (museum samples) and now existing foreign populations through mitochondrial markers.
    - Identification of the best source populations.
3) Genetic characterization through nuclear markers of the extant Beluga populations to lay the bases for future restocking program:
    - Establishment of a microsatellite panel to be used in analyses of relatedness.
    - Genetic characterization of all individuals using microsatellite and SNPs markers
    - Phylogeographical analyses.
    - Final consideration for future management.

All of the above criteria are in line with the international guidelines suggested by FAO and WSCS (World Sturgeon Conservation Society) for sturgeon conservation (Chebanov et al., 2011) by the IUCN for reintroduction (IUCN/SSC, 2013).

2. MATERIALS AND METHODS

2.1. Samples collection and DNA extraction

All the fresh samples were kept in absolute ethanol at -20 °C as tissue samples from the caudal fin or purified genomic DNA. Table 5 shows all sampled individuals, with relative type of sample and geographic origin. A subset of these samples was used within each analysis depending on the quality of the amplification performed. Beluga samples from Black Sea, Sea of Azov, Ural River and Caspian Sea were kindly provided by Dr. Nikolai S. Mugue, from the Russian Federal Research Institute of Fisheries and Oceanography; samples from Danube River were kindly provided by Dr. Radu Suciu from the Danube Delta National Institute for Research and Development).

The extraction of genomic DNA of fresh samples (using a fragment of about 40 mg) was performed using the standard protocol of the EUROGOLD Tissue-DNA Mini Kit (EuroClone).

In order to visually verify the quality of the DNA extract an electrophoretic run was performed on 1% agarose gel (mass/volume) with 0,5 µl of GelRed (BIOTIUM) every 20 ml of gel in TBE 1X running buffer.

DNA extracts have been quantified with spectrophotometer Nanodrop 2000c (Thermo Scientific).

Table 5 – List of sampled individuals.

Species	N. samples	Sample origin
H. huso	5	VIP (Northern Italy fish farm) – U.G.O.
	21	Caspian Sea
	20	Black Sea
	20	Ural River (Russia)
	17	Azov Sea
	273	Danube River (Romania)
	5	IZW Berlin – U.G.O.
	21	Pisani-Dossi (Northern Italy fish farm)
A. baerii	6	Lena River (Russia)
	14	Lake Baikal (Siberia, Russia)
	7	Ob' River (Siberia, Russia)
A. gueldenstaedtii	2	VIP (Northern Italy fish farm) – U.G.O.
	10	Azov Sea
	9	Caspian Sea
	6	Volga River (Russia)
A. naccarii	23	VIP (Northern Italy fish farm)
A. fulvescens	3	Embarras River (Wisconsin)
	3	Lake Winnebago (Wisconsin)
	3	Wolf River (Wisconsin)
	3	Upper Fox River (Wisconsin)
	3	Bad River (Wisconsin)
A. trasnmontanus	5	VIP (Northern Italy fish farm) – U.G.O.
	9	North America
A. sinensis	14	Yangtze River (China)
A. ruthenus	33	Danube River (Romania)
	8	VIP (Northern Italy fish farm) – U.G.O.
A. stellatus	11	Danube River (Romania)
	1	USSR
	1	Kazakhstan
	1	Iran
	5	IZW Berlin – U.G.O.
A. schrenckii	22	Amur River (China)
	3	Yangtze River Fisheries Research Institute (China)
H. dauricus	21	Amur River (China)
H. huso x A. ruthenus	5	University of Padova
A. naccarii x H. huso	5	VIP (Northern Italy fish farm)

Forensic processing of DNA sturgeon samples:

Background:

The molecular analysis of ancient DNA (aDNA) requires a completely different approach. The same techniques used for fresh samples analysis has some limitations for aDNA analysis, due to the highly degraded origin of ancient samples (Cooper, 2000; Hofreiter et al., 2001 a): numerous biological, physical and chemical factors affect the DNA quality of specimens from Natural History Collections (NHC) similar to ancient DNA (strand breaks caused by hydrolytic cleavage, oxidative lesions induced by free radicals, DNA cross-links formed between reducing sugars and amino groups; Shapiro, 1981; Lindahl, 1993; Pääbo et al., 2004). In addition to DNA fragmentation and structural modifications, aDNA suffers of other modifications that, without altering the amplifications of molecules used as templates, may cause an incorrect addition of nitrogen bases (nucleotide misincorporations) during aDNA amplification process (PCR). For example, deaminated C residues that are identical to uracile lead to the incorporation of A residues instead of G residues by Taq DNA polymerase (Hansen et al., 2001). The deamination products are of particular relevance for the amplification of aDNA, since they cause incorrect bases to be inserted when new DNA strands are synthesized by a DNA polymerase. This kind of displacements in aDNA further worsen the correct determination of specific sequences. Notably, these mistakes should not be confused with normal errors of Taq DNA polymerase during any amplification process, independently of sample origin. In order to distinguish these two kinds of misincorporation, multiple amplifications of the same fragment of interest should be performed. Comparing the different sequences, it is possible to identify the presence of nucleotide differences within the same sample. These false polymorphisms occur at a high frequency within the population of replicate sequences since the vast majority of these consistent substitutions are due to mistakes during the first PCR cycle, when the amplification template is constituted by the original DNA extracted from the ancient sample. (Hofreiter et al., 2001 b). In addition, different preservation methods can negatively affect the ability to extract, amplify and sequence DNA (enzyme inhibitors, cross-contamination among samples) from NHC samples (Walander et al., 2007). For example, the process of fixation with formaldehyde has profound effects on the extraction of RNA and DNA from tissues: not only cross-linking does occur with the surrounding histones, but formaldehyde can react directly with nucleotides (Mc Ghee and von Hippel, 1975 a and b). As part, or as a consequence, of this process, DNA extraction and PCR amplifications from formalin-fixed (FF) specimens is particularly difficult (Tang, 2006; Stuart et al., 2006). Furthermore, formaldehyde alter sequences through nucleotide misincorporations (Williams et al., 1999). Therefore, for this kind of samples the age is not the only important factor altering DNA quality (Walander et al., 2007). However, the most frequent DNA damage observed in ancient or museum specimens

remains the elevated degree of degradation with fragments ranging between 100 and 500 bp (Pääbo, 1989; Hofreiter et al., 2001 a). The consequence is that only a very small portion of the original genome is available for analysis, which is the main reason why most studies on aDNA focus on mtDNA rather than nuclear DNA: given the higher numbers of mitochondria the chance to recover a given sequence is increased.

Samples for DNA analysis have been searched throughout museum collections on the Italian territory. A total of 10 specimens was obtained: 6 from FF tissues and 4 from dried tissues or bony scutes. These museum specimens identified as *Huso huso* from the Adriatic Sea have been found in the Natural History Museum of Venice (1307, 1316, 7825, 21443, 20312, 21337) and in the Natural History Museum of Florence (5720, 5920, 5921, 6474); an example of sample is shown in Figure 3.

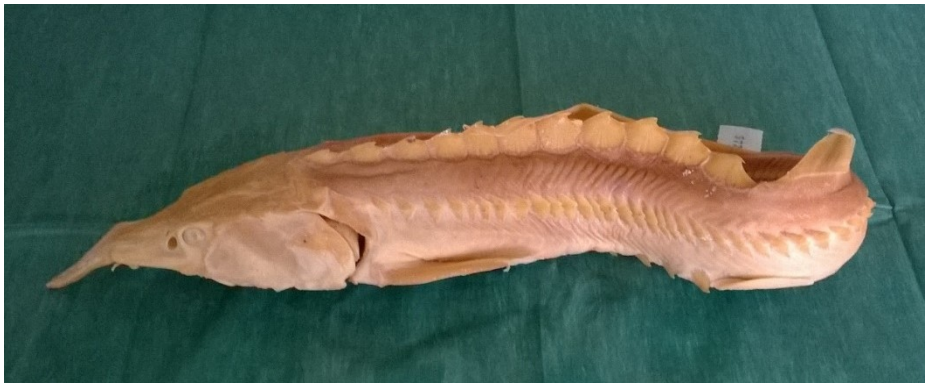


Figure 3 – Adriatic *H. huso* sample (5720) fixed in aqueous formalin from the Natural History Museum of Florence.

Samples identified as 1307, 1316, 5720, 5920, 5921, 6474 were preserved in ethanol, but very likely they underwent a previous process of fixation in formalin.
Samples 7825, 21443, 20312, 21337 were instead kept dry and prepared by a taxidermist. Samples from the Natural History Museum of Florence have been processed in the the laboratory of Anthropology (Molecular Anthropology/Paleogenetics Unit) at the Department of Biology of the University of Florence, whereas samples from the Natural History Museum of Venice have been processed in the taxidermy laboratory of the same museum. Different tissues have been taken from samples: gills and abdominal cavity from specimens in ethanol, dorsal and cranial bony scutes for dry samples, in which the external layer was removed prior to tissue extraction, in order to avoid any previous contamination.

For cleaning and tissue collection of bony tissues, a miniature Dremel drill was used in order to have a bone powder, which was then transferred into a sterile 15-mL conical polypropylene tube (Falcon). After drilling, also tissue samples below bony scutes were available for sampling and have been removed with sterile tweezers and scalpel and then kept in sterile tubes.
Specimens conserved in ethanol have been removed from the liquid and sample collection was performed with sterile tools as above. Sampled tissues were kept in pure ethanol in sterile 2 ml-tubes. To avoid cross-contamination, only one sample at a time was sampled.
Different DNA extraction protocols for Formalin-Fixed (FF) samples were tested (Paireder et al., 2013):

- Using the Qiagen DNeasy Blood and Tissue Kit preceded by a pre-treatment for FF samples recommended by the protocol (samples were initially cut into small pieces; then they were mechanically broken in liquid nitrogen with a pounder).
- Cetyltrimethylammonium bromide (CTAB) standard protocol (with both nitrogen liquid breaking and homogenizing tissue using the ball mill TeSeE PRECESS 24 BIO-RAD with different incubation digestion time).
- Variant of CTAB standard protocol with pre-extraction heat treatment at 98°C following Wu et al., 2002.
- Using QIAamp DNA FFPE Tissue Kit (QIAGEN).

Samples from FF tissues:
DNA from FF tissues was successfully extracted using QIAamp DNA FFPE Tissue Kit (QIAGEN) that is specially designed for purifying DNA from FF, paraffin-embedded tissue sections. The protocol of these commercial kit basically involves column purification of DNA using silica-gel-membrane. Each sample was removed from the conservation liquid (ethanol) and minced in 1-2 mm fragments. The quantity of tissue for each sample are reported on Table 6:

Table 6 – Quantity, origin and type of tissue sampled from the museum specimens.

n°	Sample	Origin	Tissue	Quantity
1	*5720*	Natural History Museum of Florence	Abdominal cavity	20 mg
2	*5920*	Natural History Museum of Florence	Gills	20 mg
3	*5921*	Natural History Museum of Florence	Gills	20 mg
4	*6474*	Natural History Museum of Florence	Abdominal cavity	100 mg
5	*1307*	Natural History Museum of Venice	Abdominal cavity	100 mg
6	*1316*	Natural History Museum of Venice	Abdominal cavity	100 mg

Each sample was placed in a 1.5 ml-sterile microcentrifuge tube. 180 µL of ATL buffer (Qiagen) and 20 µL of Proteinase K (0.25 mg/mL) for samples 1, 2 and 3, whereas for the remaining samples the mix: 380 µL of ATL buffer and 40 µL of Proteinase K.
Samples were vortexed and suspensions were rotated at 37 °C overnight. For each sample 20 µL of Proteinase K were added and samples were rotated at 57 °C for 2 hours (until completely lysed). 200 µL of AL Buffer (Qiagen) were added, and incubated at 56 °C for 10'. The remaining steps were performed according to manufacturer's instructions. Samples were then eluted in 50 µL and incubated with elution buffer (Buffer ATE, Qiagen) at room temperature for 5 minutes.
Dried tissues or bony scutes samples were processed following the optimized extraction protocol presented by Dabney and colleagues (2013). This method includes the basic principles mostly used for ancient DNA purification, ultrafiltration via spin columns, and binding to silica.

Samples from dried tissues and bony scutes:
Each solution and tool was UV-sterilised before using protocol for DNA extraction adapted from Dabney et al., (2013).
Approximately 20 mg of bone powder or dry tissue went into each extraction along with 1 mL of extraction buffer (final concentrations: 0.45 M EDTA, 0.25 mg/mL Proteinase K, pH 8.0). The bone powder/dried tissue was resuspended by vortexing, and the suspension was rotated at 37 °C overnight (~18 h). Remaining bone powder/dried tissue was then pelleted by centrifugation in a bench-top centrifuge for 2' at maximum speed (16,100 × g). The supernatant was added to 10mL of binding buffer containing, in final concentrations, 5 M guanidine hydrochloride, 40% (vol/vol) isopropanol, 0.05% Tween-20, and to 400µL sodium acetate 3 M (pH 5.2). A binding apparatus was constructed by forcefully fitting an extension reservoir removed from a Zymo-Spin V column (Zymo Research), which had been submerged in a bleach bath for 20', rinsed with water, and UV irradiated before use, into a MinElute silica spin column (Qiagen). The extension reservoir-MinElute assembly was then placed into a 50-mL falcon tube. The solution containing binding buffer and the extraction supernatant was then poured into the extension reservoir, and the falcon tube cap was secured. The binding apparatus was centrifuged for 4 min at 1,500 × g, rotated 90°, and centrifuged for an additional 2' at the same speed. The extension reservoir-MinElute column were removed from the falcon tube and placed into a clean 2-mL collection tube. The extension reservoir was removed, and the MinElute column was dry-spun for 1 min at 6,000 rpm (3,300 × g) in a bench-top centrifuge. Two wash steps were performed by adding 750 µL of PE buffer (Qiagen) to the MinElute column, centrifuging at 3,300 × g and discarding the flow-through. The column was dry-spun for 1' at maximum speed (16,100 × g), rotated 180°, and centrifuged for an additional 1' and then placed in a fresh 1.5-mL siliconized low-retention collection tube. For

elution, 50 μL of TET buffer was pipetted onto the silica membrane, and after a 2- to 5-minute incubation at room temperature was collected by centrifugation for 30'' at maximum speed. This step was repeated for a total of 100 μL of DNA extract. Each extract was then stored at -20°C

All the analyses on museum samples were performed in the laboratory of Anthropology (Molecular Anthropology/Paleogenetics Unit) at the Department of Biology of the University of Florence that is specifically dedicated to ancient DNA analyses

2.2. Species identification: intron characterization, cloning and sequencing

Background:

Various molecular markers have been developed for species-specific identification (DeSalle and Birstein 1996; Birstein et al., 2000; Congiu et al., 2001, 2002; Mugue et al., 2008; Doukakis et al., 2012). In general, available molecular tools for the identification of species can fall within two main groups: mitochondrial and nuclear markers, depending on the localization on the genome. In particular, with regard to sturgeons, the available techniques and their effectiveness have been recently discussed during the Status Workshop on the Identification of Acipenseriformes in Trade, hold in Berlin in 2006 by Species Survival Commission (SSC) and Sturgeon Specialist Group of IUCN (Ludwig, 2008). Nonetheless, mtDNA has a disadvantage: given its maternal inheritance, it is useless for the identification of hybrids or species that have previously undergone even ancient introgression events (Ludwig, 2008; Mugue et al., 2008). Nuclear DNA is on the contrary representative of both maternal and paternal inheritance, making it useful for the identification also of hybrids. For this reason, genetic analyses on mtDNA are often supported by nuclear DNA analysis (Ludwig, 2008). Boscari and colleagues (2014 a) have developed a method based on the presence of species-specific SNPs in the ribosomal protein S7 (RPS7) of the gene sequence, which leads to the identification by PCR of different sturgeon species with a 100% specificity.

The main nuclear markers at single locus used in this study to select species-specific SNP were selected from 10 introns previously isolated in silico (Rossi, 2014/2015) and successively used for validation of *A. schrenckii* e *H. dauricus* (Maccatrozzo 2015/2016).

A method based on the presence of species-specific SNPs in the first intron of the gene encoding for the ribosomal protein S7 (RPS7) was developed by Boscari and colleagues (2014 a). This method allows the identification by PCR of different sturgeon species with a 100% specificity. Given the current lack of any single-locus nuclear marker to discriminate *H. huso* from other sturgeon species, the second intron of the nuclear S6 Ribosomal Protein (RP2S6) has been characterized following the above SNPs-based approach.

A total of 386 samples of 11 sturgeon species were analysed: individuals of the different species were opportunely selected from different geographical origins or, when possible, on the basis of pedigree information in order to avoid the selection of related specimens. Sixty-five individuals were used for the characterization of the intronic region, while a second validation phase was implemented on 321 animals.

The intron characterized in this work was predicted by aligning assembled transcriptomes of three sturgeon species (*A. fulvescens*, *A. naccarii* and *A. stellatus*) (Hale et al., 2009; Vidotto et al., 2013, 2015) with three available genomes of teleost fishes (*Danio rerio*, *Latimeria chalumnae* and *Takifugu rubipres*) (Howe et al., 2013; Amemiya et al., 2006; Brenner et al., 1993); the alignment identified splicing sites in order to design specific primers on flanking exonic regions. For 10 introns selected among the 1867 loci isolated *in silico*, in our laboratory the primer pairs built on the corresponding exon-flanking regions were already available. In detail, focusing on one of these loci, the second intron (RP2) of the ribosomal protein S6 (RPS6) gene, the amplification of which was performed using the primer pair RP2S6_F (5'-TTCATGGGGAAACCCTGCTT-3') RP2S6_R (5'-ATCCTCTGGGTGAGGAGTG-3'). The following touch-down thermic profile has been optimised (Don et al., 1991), resulting in a clear band on agarose gel without aspecific products:

94° 2'
94° 30"
63° 30" x 6 cycles 0.5° each cycle
72° 15"
94° 30"
60° 15" x 25 cycles
72° 15"
72° 5'

The concentrations of reagents used in amplification reactions, for a final volume of 20 µl, are as follows: Buffer 1X (RBCBioscience); dNTPs 0,1mM (5prime); forward primer 0,25 µM; reverse primer 0,25 µM; Taq polymerase 0,025 U/µl; 20 µg of genomic DNA.
Amplification reactions were performed with thermocyclers 2720 Thermal Cycler (Applied Biosystem) and GeneAmp PCR system 9700 (Applied Biosystem).
Three µl of PCR products have been evaluated with electrophoretic run on 1.8% agarose gel with GelRed (0.5 µl in 20 ml of gel) in TBE 1X running buffer. For those loci with a successful optimization of PCR as indicated by a clear band in the agarose gel, sequencing and cloning, when necessary, of the amplification products were performed.
For the species-specific marker identification, cloning was used in order to separate the different alleles amplified at a same locus in different individuals for subsequent sequencing following Boscari et al., 2004a.

The control PCR to identify colonies containing the fragment of interest was performed using the universal plasmid primers T7_For (5'-TAATACGACTCACTATAGGG-3') and M13_Rev (5'-CAGGAAACAGCTATGAC-3').
Final concentrations of reagents in a total volume of 20 µl were: Buffer 1X (RBCBioscience), dNTPs 0,1mM, (5prime) forward primer T7 0,25 µM, reverse primer M13 0,25 µM, Taq polimerasi 0,025 U/µl 2 µl of solution containing a lysed colony.
PCRs have been performed with thermocyclers 2720 Thermal Cycler (Applied Biosystem) and GeneAmp PCR system 9700 (Applied Biosystem) with the following standard thermal profile:

94° 2'
94° 30"
56° 1' x 33 cycles
72° 1'
72° 5'

PCR products have been visualized with electrophoretic run on 1.8% agarose gel: PCR products with the fragment of interest, identified based on bp length, have been purified with enzymatic reaction and sequenced. Primer M13_Rev has been used for sequencing.
Chromatograms resulting from sequencing have been analysed with FinchTV software (Geospiza) for quality check. Alignment of sequences was then performed with MEGA6 (Molecular Evolutionary Genetics Analysis) using ClustalW (Tamura et al., 2013) and Clustal Omega online service (http://www.ebi.ac.uk/Tool/msa/clustalo/) (McWilliam et al., 2013), in order to analyse sequence variability and identify single alleles in different species (in the case of the development of the species-specific marker). Moreover, MEGA6 was used to produce, starting from alignments, phylogenetic trees with UPGMA algorithms (Unweighted Pair Group Method with Arithmetic Mean) and NJ (Neighbour Joining), for a more immediate graphic representation of similarity between sequences.
Once the presence of two groups of sequences (group A and group B) was established, as described in details in the results section, the presence of informative polymorphisms potentially suitable to distinguish Beluga animals from the other species were investigated.
Alignments of sequences from the group A have been analysed for polymorphisms fixed within the same species and variable between different species, on which species-specific primer pairs have been designed with OligoExplorer 1.2 software (Gene Link). Primers have been designed to have the 3'-end nucleotide complementary with the diagnostic SNP obtained from the alignment, in order to have a species-specific primer. To reduce the aspecific amplification in other species,

the second to last from 3'-terminal nucleotide was substituted with the same base present in the target site, thus causing a 2 nucleotides-mismatch in aspecific DNA sequences, which should guarantee the amplification only of the species of interest, where the specific pairing of primers happens (Mugue et al., 2008; Boscari et al., 2014 a).
In detail, one potentially diagnostic SNPs for the Beluga sturgeon was identified in the sequence of the group A, and the specific forward and reverse primers designed were:
Forward (RP2S6_huso_F): 5' - CATAACATTGCACTGAATGTTATA - 3'
Reverse (RP2S6_groupA_R): 5' - CTTTCGTTGATTTAGGGAAATGGT - 3'
Amplification conditions for the primer pair were optimized as follows:

94°C 2'
94°C 30''
61°C 30'' x 35 cycles
72°C 15''
72°C 5'.

After the marker validation, all the available samples used for subsequent analyses (both with mitochondrial and nuclear markers) where tested using the Beluga-specific marker in order to exclude the possibility to deal with hybrids. After the detection of some non-Beluga individuals, the samples where checked for the *A. ruthens*, *A. stellatus* and *A. naccarii* using a PCR-based tool based on diagnostic SNPs on the first intron of the S7 gene that allows a reliable identification of sturgeon species and hybrids specifically developed by Boscari and colleagues, (2014 a).

2.3. Mitochondrial DNA analyses

Background:

The vast majority of vertebrates has a unique form of maternally-inherited mtDNA. Sturgeons, like other species of fish such as codfish (Arnason and Rand, 1992), show a specific variability in mtDNA length (Brown et al., 1992; Ludwig et al., 2000; Buroker et al., 1990). This condition, known as heteroplasmy, means that a particular individual may have different forms of mtDNA due to a different length in the D-loop region, where tandem-repeated units (82-84 bp with a sevenfold factor maximum in sturgeon) are present (Çiftci et al., 2013). This heteroplasmic condition can complicate molecular analyses, making it necessary to discard the initial portion of the D-loop, where the repeats are located (Buroker et al., 1990).

Primers for the amplification of mitochondrial D-loop have been designed with Oligo Explorer 1.2. In particular, primers used for the amplification of extracts obtained from museum samples were designed on sequences already available in our laboratory and on GenBank (Accession number: AY442351) in order to have expected amplification products of no more than 170 bp. This is due to the high degradation, as expected by the processing of ancient DNAs, that makes impossible the amplification of larger regions. Therefore, the 9 primers pairs designed were (Table 7):

Table 7 – Primer pairs designed for the amplification of overlapping fragments of the D-loop Control Region of the Beluga sturgeon.

	FORWARD	REVERSE	Length
1	5'-ATACCACAATGTTTACATATAC-3'	5'-GCATGTTAGTACACATAGAG-3'	146 bp
2	5'-CCACATTAATTTCTAGCCACC-3'	5'-ACAAGGATGTTCGGCTCTTA-3'	130 bp
3	5'-GTACACCATTTTCTCTATGTGTAC-3'	5'-GGACTTGAAATAGGAGCCAG-3'	170 bp
4	5'-GTCTAGAACATGAAGTTAATGAG-3'	5'-GAGTTTATGGTACCTCTAAC-3'	159 bp
5	5'-TATTACTAACATCTGGCTCCT-3'	5'-TGAATGGGAACCGAAATAAGAG-3'	161 bp
6	5'-ACCCCACCAACATTTGGTTA-3'	5'-GCGTCCAGGATATGTCATTATATC-3'	145 bp
7	5'-TGGAACATTCAACAGCCTGC-3'	5'-ATAATAAACGTTTATAGTATCTTGGT-3'	163 bp
8	5'-TGTTTCACGTAGACCCAGTC-3'	5'-CACATGTGGGTAATTAGAGTACG-3'	169 bp
9	5'-CTTATATTTCTTGTCAAACCCC-3'	5'-GCTATATATTATTTTGTGTGATTGTG-3'	147 bp

For the validation, each fragment was also amplified twice from control extracts of *H. huso* (fresh samples).

Thermal profiles optimised for ancient specimens were as follows:
PCR profile for primers pairs 1, 2, 4, 5, 7 and 9:

95° 2’
95° 30”
52° 45” x 35 cycles
72° 30”
72° 5’

PCR profile for primers pairs 3, 6 and 8:

95° 2’
95° 30”
55° 45” x 35 cycles
72° 30”
72° 5’

For each sample two different amplification protocols were performed (bi-phase PCR and single-cycle PCR). The protocol of bi-phase PCR consists of two amplification steps: in the first reaction, genomic DNAs were used as template, while in the second one, the PCR products (diluted 1:5) obtained by the previous reaction were used as templates.
Two-step PCRs were performed in a total volume of 20 µl, consisting of 1 µM of each primer, 0.2 mM dNTPs, 1x buffer (Buffer Gold $MgCl_2$ free: 150 mM Tris-HCl, pH 8.0, 500 mM KCl), 2.5 mM of $MgCl_2$, 2 U/ µl of Taq DNA polymerase (AmpliTaq Gold® DNA Polymerase, ThermoFisher SCIENTIFIC) and 2 µl of extracted DNA or eluted amplificate as template.
The thermocycler was set as follows:

95° 10’
94° 20”
52° or 55° 30” x 27 cycles (first step)/ 33 cycles (second step)
72° 60”
72° 10’

One-step PCRs were performed in a total volume of 50 µl, consisting of 1 µM of each primer, 0.18 mM dNTPs, 1x buffer (Buffer Gold $MgCl_2$ free: 150 mM Tris-HCl, pH 8.0, 500 mM KCl), 2 mM of

$MgCl_2$, 1.5 U/µl of Taq DNA polymerase (AmpliTaq Gold® DNA Polymerase, ThermoFisher SCIENTIFIC) and 2 µl of extracted DNA as template.
The thermocycler was set as follows:

95°	10’	
94°	45”	
52° or 55°	60”	x 50 cycles
72°	60”	
72°	10’	

All amplifications were performed on Peqlab Primus 96 advanced®.
Products were checked by 2% agarose gel electrophoresis, purified from gel or using ExoSAP- IT® and directly sequenced on an ABI Prism 3730XL or an ABI 3100 automatic sequencer at BMR Genomics or Eurofins Genomics.

For fresh samples, two standard primers were used for the amplification of the D-loop Control Region (AcipPRO1F: 5’-CAC CCT TAA CTC CCA AAG C-3’; AcipPHE1R: 5’-CCC ATC TTA ACA TCT TCA GTG-3’) (Congiu et al., 2011), with the following profile:

94°	2’	
94°	45”	
56°	30”	x 30/35 cycles
72°	45”	
72°	5’	

For a remarkable fraction of fresh samples, a poly-cytosine stretch (with a 5-10 bp variable length) located in the terminal part of the D-loop was found in all the individuals. For this reason, a smaller portion was considered for further analyses and a second reverse primer was designed (D-loop_INT: 5’-TGG GTC TAC GTG AAA CAC TC-3’) to exclude the poly-cytosine stretch. A total of 369 nucleotides were then used for the analyses.

PCR conditions were optimized as follows:

94° 2’

94° 45”

59° 45” x 35 cycles

72° 45”

72° 5’

For all PCR reactions performed on fresh sample templates, the final concentrations of reagents were the same reported in Boscari et al., (2014 a). All amplifications were performed on an Applied Biosystem GeneAmp®PCR System 9700. After checking the PCR products on 1.8% agarose gel stained with GelRed (BIOTIUM, GelRedTM Nucleic Acid Stain, 10000X in Water), they were purified using ExoSAP-IT® and directly sequenced at the external service BMR Genomics and/or Eurofins Genomics.

2.4. Microsatellite analyses

Background:

Microsatellites are widely used from the late 80s for forensics identifications, kinship analysis (most commonly in paternity testing), population genetics, and various applications for the conservation and management of biodiversity (Balloux and Lugon-Moulin, 2002; Wan et al., 2004) Recently the use of microsatellites has been gradually reduced in favour of SNPs, which occur at higher frequencies in the genome and can then be isolated and analysed in great numbers. Nonetheless, a single SNP (usually diallelic) has a lower detection power than a single multi-allelic microsatellite locus (Guichoux et al., 2011).

A total of 27 loci were tested in all the extant Beluga populations sampled (Table 8). These loci were selected from a wider set of microsatellite markers tested in our laboratory on other sturgeon species and, preliminarily, on a reduced subset of Beluga samples. Forward primer for each pair was tagged at 5' with a fluorophore (FAM, VIC, NED, HEX, PET or TAMRA), to allow genotyping of loci amplified in multiplex.

Table 8 – List of loci tested in all the Beluga populations. The repeat and the size of each locus, the species in which the locus was firstly isolated, the references, fluorophores and the relative final multiplex assignment are reported. In the multiplex assignment, the * refers to loci individually amplified; - refers to loci discarded for the subsequent analyses.

Locus	Motif	Size (bp)	Species	Fluorophore	Multiplex assignment	Reference
Anac_tag71527	(TA)6	199-204	*A. naccarii*	FAM	A	Boscari et al., 2015
Anac_tag83433	(AG)7	299-309	*A. naccarii*	VIC	-	Boscari et al., 2015
Anac_tag112145	(GGA)7	104-113	*A. naccarii*	FAM	-	Boscari et al., 2015
Anac_c1490	(TG)8	314-324	*A. naccarii*	FAM	-	Boscari et al., 2015
Anac_c6179	(TTTA)10	108-102	*A. naccarii*	VIC	A	Boscari et al., 2015
Anac_c33182	(CA)13	141-158	*A. naccarii*	FAM	-	Boscari et al., 2015
Anac_c7314	(TAGA)17	199-293	*A. naccarii*	VIC	B	Boscari et al., 2015
Anac_c22096	(AGGA)7	188-212	*A. naccarii*	VIC	A	Boscari et al., 2015
Anac_c1538	(TTAT)7	248-268	*A. naccarii*	NED	-	Boscari et al., 2015
Anac_c15214	(AC)13	259-285	*A. naccarii*	NED	-	Boscari et al., 2015

Anac_c159	(TAT)19	336	*A. naccarii*	PET	B	Boscari et al., 2015
Anac_c16902	(AAT)7	271-287	*A. naccarii*	PET	A	Boscari et al., 2015
Anac_c6784	(ATTT)8	311-346	*A. naccarii*	FAM	A	Boscari et al., 2015
Anac_c12159	(ATT)11	238-265	*A. naccarii*	PET	A*	Boscari et al., 2015
Anac_c31601	(TGTC)8	154-206	*A. naccarii*	NED	A	Boscari et al., 2015
Anac_c5367	(TTTA)8	206-211	*A. naccarii*	PET	B	Boscari et al., 2015
Anac_c3133	(GA)11	164-178	*A. naccarii*	PET	-	Boscari et al., 2015
An20	(ATCT)10(TG)5	159–213	*A. naccarii*	TAMRA	C	Zane et al., 2002
AnacE4	(CA)20	326-354	*A. naccarii*	FAM	C	Forlani et al., 2008
Spl168	(TATC)18	174-199	*S. platorynchus*	NED	-	McQuown et al., 2000
Spl163	(GATA)17	186-202	*S. platorynchus*	NED	B	McQuown et al., 2000
AfuG41	(GATA)9TA(GATA)3	237-294	*A. fulvescens*	FAM	-	Welsh et al., 2003
AfuG112	(GATA)12GACA(GATA)6	222-253	*A. fulvescens*	HEX	-	Welsh et al., 2003
AoxD241	(TAGA)36	156–198	*A. oxyrinchus*	HEX	C*	Henderson-Arzapalo and King, 2002
AoxD64	(TAGA)16	216–252	*A. oxyrinchus*	FAM	C	Henderson-Arzapalo and King, 2002
AoxD161	(CTAT)15	111–155	*A. oxyrinchus*	FAM	C	Henderson-Arzapalo and King, 2002
AS-034	(CAA)7	170–205	*A. sinensis*	NED	-	Zhu et al., 2005

Final concentrations of reagents in a total volume of 10 µl of multiplex were: Qiagen Mastermix 1X (HotStarTaq DNA Polymerase, Multiplex PCR Buffer, dNTPs Mix) primer mix 0,2 µM (containing forward and reverse primers of all the necessary loci to be amplified), 20 ng of genomic DNA. All the primer pairs were already available in the laboratory.

The criteria of final selection of markers were the high polymorphism rate and the presence of a maximum of two alleles per locus per individual (in order to avoid the presence of duplicate loci); with these criteria, the final set of microsatellite markers selected was composed by 16 loci. Fourteen loci were simultaneously amplified in 3 multiplex PCR reactions (A, B and C; Table 8),

including respectively 6, 4 and 4 loci. The remnant 2 single loci were individually amplified and then assembled together with the previous one for genotyping.
For each multiplex (A, B and C) the amplification conditions were:

95° 15’
94° 30”
60° 90” x 35 cycles
72° 60”
60° 30’

Loci amplified in single following specific profiles are reported below:
Anac_c12159, coupled with Multiplex A:

94° 2’
94° 30”
60° 30” x 37 cycles
72° 30”
72° 5’

AoxD241, coupled with Multiplex C:

95° 2’
95° 45”
56° 45” x 35 cycles
72° 45”
72° 5’

For each performed PCR reaction, the final concentrations of reagents were the same reported in Boscari et al., (2014 a). All amplifications were performed on an Applied Biosystem GeneAmp®PCR System 9700. The PCR products were checked on 1.8% agarose gel stained with GelRed (BIOTIUM, GelRedTM Nucleic Acid Stain, 10000X in Water) and directly sequenced at the external service BMR Genomics using ABI 3730XL and ABI 3100 sequencer.

2.5. 2b-RAD technique for SNPs analyses

Background:

SNPs are considered as bi-allelic molecular markers, as they frequently show only two alternative variants. This low-grade variability makes them less efficient, compared with microsatellites, which can show even a few dozen alleles per locus. Nonetheless, recent technological advances provided the development of fast and cheap techniques to identify genetic markers with the widest coverage of genetic variability. The vast majority of SNPs within the genome widely balances their reduced variability, in addition to their suitability for high-throughput genotyping (e.g. Illumina sequencing) (Freeland et al., 2011). The usage of SNPs as genetic markers is widespread in model organisms, Indeed, in the case of non-model organisms, for whom genome information is scarce, the first step resides in SNP identification. To this end, the most widely used method is restriction site-associated DNA sequencing (RADseq) (Davey and Blaxter, 2010; Etter et al., 2011). For example, the 2bRAD technique (Wang et al., 2012) is a genome-wide approach, based on Next Generation Sequencing (NGS), that allows to obtain an excellent fractioned representation of the genome, (Wang et al., 2012).

Like microsatellite and mitochondrial analyses, only adult Beluga individuals of the different populations were kept for further genetic comparison analyses: 15 individuals from the Sea of Azov (including one replicate from one sample), 19 from the Black Sea (including one replicate of one sample), 21 from the Caspian Sea (including one replicate of one sample), 20 from the Ural River (including one replicate of one sample) and 24 from the Danube River (including two replicates of one sample). In the case of Danube individuals, the low quality of extracts has not allowed the usage of all the available adults. Fifteen individuals (3 adults and 12 juveniles) from Pisani-Dossi livestock were kept for relatedness analyses. The final number of individuals was 114.

The 2b-RAD protocol used in this study is based on the optimization performed on three species of marine organisms (*Leptopsammia pruvoti*, *Cladocora caespitosa* and *Myriapora truncata*) analysed for the PRIN project "Coastal bioconstruction: structure, function, and management" by Boscari and colleagues in the same laboratory. After a preliminary test on a small set of samples of Beluga chosen from different populations, 400/500 ng of high-quality, RNA-free gDNA, were used from each individual (n=114, including replicates). gDNA was processed with 2U of the 2b restriction endonucleases *Csp*CI (New England BioLabs) overnight at 37°C. This enzyme gives a population of fragments with the same length (35bp) with random protruding ends (Marshall and

Halford, 2010). In order to evaluate the efficacy of restriction enzymes, for each sample 1 µl of digested DNA next to a similar quantity of undigested gDNA was loaded on a 1% agarose gel stained with GelRed. The digested product was used as a template for a ligation reaction for which partially double-stranded adaptors with compatible and partially degenerated overhangs (5'-NGNN-3') were used in a 25 µl total volume reaction consisting of 0.4 µM of each adaptor, 0.2 mM ATP and 40 U/µl T4 DNA ligase (SibEnzyme). The obtained fragments were then amplified to add sample-specific barcodes (7 bp) and the annealing sites for Illumina next-generation sequencing. The PCR reactions were performed in a total volume of 50 µl, consisting of 10 µl of ligated DNA, 0.2 µM of each primer 2b-RAD amp, 0.5 µM of each primer F and barcoded-primer R, 0.3 mM dNTPs (EURX), 1X Phusion HF buffer and 0.02 U/µl Taq Phusion high fidelity polymerase (New England BioLabs). After quality check of the PCR products on 1.8% agarose gel stained with GelRed, amplification products from different samples were mixed (thus generating a pool of fragments) by comparing respective concentrations of the expected target band (approximately 165bp, including the restriction fragment of interest, adapters and barcodes) for a 1-10 µl variable volume, in order to standardize the coverage for following sequencing. The normalization of different concentrations was performed by comparing measures with NanoDrop UV-Vis spectrophotometer and the analysis of band intensity from agarose gel was performed using ImageJ (Schneider et al., 2012).

The resulting pool of fragments was then purified from high-molecular weight fragments and primer-dimers in two steps. Firstly, by cutting the specific band from 1% agarose gel. The pool of fragments in each band was then eluted in distilled water overnight. This purified product was then processed with magnetic beads (SPRIselect, BECKMAN COULTER) according to Solid-Phase Reversible Immobilization (SPRI) method (DeAngelis et al., 1995).

The first pool was sequenced on an Illumina NextSeq 500 platform with a single-end SR50 High Output mode by Genomix4Life S.r.l. (Baronissi, SA, Italy), which also performed data demultiplexing and quality-filtering. Following sequencing results of the first run, the relative amount of the target 170 bp band of each sample was precisely estimated based on the number of reads obtained for each sample. A new pool was then assembled for each group of individuals, adjusting volumes to obtain the final normalization. The new pool was purified and sequenced on an Illumina HiSeq 2500. This procedure allowed to increase the depth coverage and ensured to obtain an equal and comparable number of reads from each analysed sample. For the sequencing of the second pool, the Illumina HiSeq platform was used instead the NextSeq platform because the preliminary analyses on the NextSeq data showed the presence of sequencing errors in the restriction site of the enzyme (later explained in chapter 2.6.3). Therefore, two different pools

were sequenced in two different lanes: the first one with the NextSeq platform, the second one on HiSeq platform.

2.6. Statistical analyses

2.6.1. Mitochondrial DNA - D-loop

For what concerns contemporary samples, only adult individuals were further analysed for the genetic comparison with museum samples. Among the Danube River stored samples there were many juveniles (YOY, young of the year, n= 162). Danube YOY were discarded from the mitochondrial analyses for three reasons: first to balance the sample size of all the populations, second to standardize the type of sampled used (other populations are composed only by adult individuals), third to avoid any bias on the analyses. In fact, given the typical gregarious behaviour of this species during its first life stages, the chance of dealing with family groups is very high, which would have compromised further analyses (since full sibs obviously share the same mitochondrial haplotype there is a high probability to deal with an over representation of some haplotypes).

The following contemporary adult samples were used (total 119): 20 individuals from the Black Sea, 15 individuals from the Sea of Azov, 20 individuals from the Caspian Sea, 20 individuals from the Ural River and 44 individuals from the Danube River.

Sequences from online data bank GenBank® were taken from Dudu et al., (2014) and Doukakis et al., (2005). The accession number of the sequences is:

Sea of Azov: AY846666-67

Black Sea: AY846668-79

Caspian Sea: AY846640-65

Danube River: KF431840 - KF431846

For relatedness analyses the D-loop was sequenced and analysed on 36 Danube YOY individuals and on 9 adult individuals from the Pisani-Dossi livestock.

Sequences were aligned with MEGA6 by ClustalW (Tamura et al., 2013) and each polymorphism checked by eye on the corresponding chromatogram using the software FinchTV.

MEGA6 was also used to produce, from the alignments, phylogenetic trees with the NJ or Maximun Likelihood (ML) algorithms, useful to provide a more immediate graphical representation of the similarity between the analyzed sequences: the NJ was based on the number of nucleotide differences among sequences applying the pairwise deletion option with 1000 bootstrap replications (for the cluster analyses for the identification of the species-specific marker); for the ML algorithm (used for analysis of the D-loop sequences of only Beluga samples) the Tamura 3-

Parameters was chosen as the best-fit model of nucleotide evolution following the suggestion of Dudu and colleagues (2104).

Each sequence obtained from museum specimens was compared with online database of sequences (BLAST: http://blast.ncbi.nlm.nih.gov/Blast.cgi) and with sequences of pure Beluga already available in the same laboratory to verify the matching with *H. huso* sequences.

The genetic variability of each extant population, estimated using the software Arlequin ver. 3.5 (Excoffier et al., 2010), was evaluated through haplotype diversity (*h*) and nucleotide diversity (π), relative haplotype frequencies, ϕ_{st} indices and non-hierarchical and hierarchical analysis of molecular variance (AMOVA, Excoffier et al., 1992).

Different haplotypes were identified by TCS software (Clement et al., 2000) on a total of 121 individuals used for the analyses (including all extant populations and two Adriatic museum samples, M1 and M3). The relationships between haplotypes (i.e. representation of gene genealogies based on a maximum parsimony approach), were organized in a network with the PopART software (http://popart.otago.ac.nz) using the TCS network inference methods (Clement et al., 2002). Using the software DNAsp (Rozas and Rozas, 1995) the presence and the frequency of different haplotypes within each population were estimated.

2.6.2. Microsatellite

The analysis of genetic comparison using microsatellite markers was performed on 110 adult specimens (37 from the Danube River, 19 from the Black Sea, 16 from the Sea of Azov, 19 from the Caspian Sea and 19 from the Ural River). The relatedness analyses were performed on 162 juveniles (young of the year, YOY) Danube Beluga individuals and on 21 individuals (9 adults and 12 juveniles) from the Pisani-Dossi livestock.

The scoring was manually checked using GeneMarker v. 2.2.0.

To test for the presence of highly related familiar groups in the YOY, a relatedness analysis was performed using ML-Relate software (Kalinowski et al., 2006). This software calculates the maximum likelihood estimates of relatedness, r (following Blouin, 2003), and then calculates the likelihood of four relationships for each pair of individuals that have the highest likelihood for each pair of individuals. Some of the YOY were also checked using COLONY v. 2.0.6.2, a computer program that implement a likelihood method to infer sibship (full- and half-sibs) among individuals (Jones and Wang, 2010). The analysis was performed with a medium run, using a pairwise-likelihood score method, without updating the allele frequencies, with no sibship size

prior and unknown allele frequencies. To support the relatedness obtained by the analyses, the distribution of pairwise genetic distances among individuals with known degree of relatedness (generated by the software Hybridlab v.1.1; Nielsen et al., 2006) were estimated (Congiu et al., 2011). The individuals with a known degree of relatedness (Full-sibs, n= 1000) were generated starting from adults' genotype of 20 randomly selected individuals chosen between the different populations available. The genotypes were then converted to band-sharing input data and then used to generate a Sorenson distance matrix (performed using the package PopTools 2.7 for Microsoft Excel; Hood, 2010).

Subsequent analyses were conducted only on adult individuals.

The statistical power for detecting genetic differentiation was evaluated using the software POWSIM (Ryman and Palm, 2006). This analysis simulates sampling from a specified number of populations that differentiated from a common base population to predefined levels of expected divergence under a classical Wright-Fisher model without migration or mutation. The power to detect the expected F_{ST} was tested for different effective subpopulation sizes (N_e; low, medium and high N_e were selected in order to test all the N_e range suggested by the software manual) and generations of drift (t), with their interaction, as suggested by the software manual (N_e used: 100, 500, 1.000, 5.000 and 10.000; t used for F_{st} = 0.01: 2, 10, 20, 100 and 201; t used for F_{st} = 0.005: 1, 5, 10, 50 and 100; t used for F_{st} = 0.025: 5, 25, 50, 25 and 50). Significance was checked for both Fisher's exact test and χ^2 test after 1000 simulations.

The number of alleles and the allelic richness for each locus and population was calculated using FSTAT ver. 2.9.3.2 (Goudet, 1995). For each locus, the average observed (Ho) and unbiased expected heterozygosity (He) was calculated using GENTIX ver. 4.05 (Belkir et al., 2005). All loci were tested for the Hardy-Weinberg Equilibrium using GENEPOP ver. online 4.2 (Raymond and Rousset, 1995; Rousset, 2008) and significance for tests involving multiple comparisons was obtained using the SGoF+ method (Carvajal-Rodriguez and de Uña-Alvarez, 2011).

The genetic variability and differentiation of each sampled population was estimated through F_{st} indices and analysis of molecular variance (AMOVA) using the software Arlequin ver. 3.5.

In order to identify the groups of populations that are geographically homogeneous and maximally differentiated, a Spatial Analyses of Molecular Variance (SAMOVA) was performed using the software SAMOVA 2.0 (Dupanloup et al., 2002). This analysis clusters the populations in different combinations, in order to maximize the proportion of total genetic variance due to differences between groups of populations.

To further evaluate the degree of genetic differentiation between populations, population structure was assessed by using a model-based clustering method implemented in the software STRUCTURE ver. 2.3.4, in which individuals in the sample are probabilistically assigned to one or

jointly to two or more K populations, if their genotypes indicate that they are admixed (Prichard et al., 2000; Falush et al., 2003, 2007). Each run was performed with 10.000 length of Burnin period, 1.000.000 of Markov chain Monte Carlo (MCMC) repeats; after Burnin, the Admixture model and the Correlation of the Allele Frequencies were selected as default parameters and without using the sampling location as prior information to assist clustering. Ten runs were performed for each value of K to check the consistency of results between runs. The most likely number of groups (K) inferred by STRUCTURE analysis was estimated using the online software STRUCTURE HARVESTER WEB (Earl and von Holdt, 2012) using the Delta (K) method (Evanno et al., 2005).

In order to summarize the pattern of the distribution of the genetic microsatellite variance among populations, a Principal Coordinates Analysis (PCoA, performed using the package GenAlEx 6.5 for Microsoft Excel -Peakall and Smouse, 2006; 2012) and a Discriminant Analysis of Principal Components (DAPC, performed using the package ADEGENET for R software 3.2.3. Jombart, 2008) were performed. The PCoA allows to find and plot the major patterns within a multivariate dataset of multiple loci and multiple samples; the DAPC is a multivariate method to represent clusters of genetically related individuals providing a useful visual assessment of between-population differentiation. The DAPC analysis was performed by setting the separation of individuals into pre-defined groups (populations) in order to assess how clear-cut or admixed the clusters are.

To estimate the spatial patterns of the populations, the Bayesian clustering algorithm employed in GENELAND 4.0.0 (a package of R software) was applied, which explicitly takes into account the spatial location of sampling sites and estimates the optimal number of population clusters (Guillot et al., 2005). The precise coordinates of sampling localities were known only for a subset of 93 individuals (28 individuals from the Danube River, 19 individuals from the Black Sea, 9 individuals from the Sea of Azov, 18 from the Caspian Sea and 19 from the Ural River). The analysis was conducted by using correlated allele frequency models, with 10 replicates for each K value (from 1 to 6), with the spatial model selected to infer the number of clusters with 100.000 MCMC iterations of which every 100^{th} was retained. The run processed was the one with the highest mean logarithm of posterior probability. The run was processed on a landscape of 100 x 100 cells and with a burn-in of 200 iterations.

2.6.3. SNPs

The quality of demultiplexed raw reads from the two sequenced lane was checked with FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

A custom-made script was then used to filter reads for the presence of *Csp*CI recognition site and to trim adaptors, obtaining sequences of uniform length (32 bp).

Several sequencing errors in the recognition site of the enzyme were identified for the pool sequenced with the NextSeq technology (first lane). This kind of sequencing error can occur at a higher rate on this type of sequencing technology (Reuter et al., 2015; Ambardar et al., 2016): the sequencing chemistry that the NextSeq platform employs probably is not appropriate to sequence 2b-RAD libraries (Hernandez-Castro et al., under review). Within the recognition site, these errors occurred at fixed positions along the reads and the mutation by base substitution was consistent at those positions.

Therefore, in order to increase the number of reads containing the correct *CspCI* recognition site, a screening with another custom-made script that identify and correct the possible sequencing error was performed (a maximum of one mutation in the recognition site was accepted).

Other errors in the sequence of the reads can also occur, but hypothetically their rate is lower and it can be assumed that these errors can be overlooked (the low occurrence rate allow to discard reads with these errors in subsequent analyses).

The trimmed, high-quality reads (mean quality score per base PF > 37) formed the basis for subsequent analysis. The correlations between raw and trimmed reads were estimated using the software SPSS (IBM Corp., 2012).

In silico assembly of loci and genotyping was performed using STACKS software (Catchen et al. 2013), and, specifically, "denovo_map.pl" pipeline was employed., A technical replicate of 6 sample was used to estimate error rates and to optimize de-novo assembly parameters of 2b-RAD data (Mastretta-Yanes et al., 2015). The parameters were set to a minimum stack depth (m) of 5, the number of mismatches allowed between stacks to build a locus in an individual (M) was 2, the maximum number of mismatches allowed to align secondary reads to primary stacks (N) was 4, the maximum distance between loci from distinct individuals to be merged in the population catalogue (n) was 2, the error rate to call SNP (bound) was set between 0 and 0.1, and significance level required to call a heterozygote or homozygote (alpha) was 0.1. The Populations package in STACKS was used to generate output in GENEPOP format for further downstream analyses.

CREATE software was used for conversion of the GENEPOP input file in input formats for different genetic analysis programs. Individuals with an excess of missing data were discarded to avoid possible bias in the analyses: the final dataset included all polymorphic loci present in at least 90%

of the individuals (11 from the Sea of Azov, 15 from the Black Sea, 9 from the Danube River, 19 from the Caspian Sea and 19 from the Ural River), and characterized by the presence of 1 SNPs. All loci were filtered for the Minor Allele Frequency (MAF = 1%). A probability test for each locus in each population was performed using GENEPOP v. 4.6 in order to identify the most polymorphic loci. A Fisher's test was then used in order to correct the threshold values of the probability test. For each population, the average observed (Ho) and unbiased expected heterozygosity (He) was calculated using GENETIX ver. 4.05 (Belkir et al., 2005) and the allelic richness (Ar) was calculated using HP-RARE (Kalinowski, 2005) based on the smallest sample size (n = 6) across all populations. The genetic variability and differentiation of each sampled population was estimated through F_{st} indices and AMOVA (nonhierarchical and hierarchical) using the software Arlequin ver. 3.5.

The package ADEGENET for R 3.2.3. (Jombart, 2008) was used for DAPC using the same strategy performed for the microsatellite analysis and for the STRUCTURE-like plot performed setting the separation of individuals into pre-defined groups (populations).

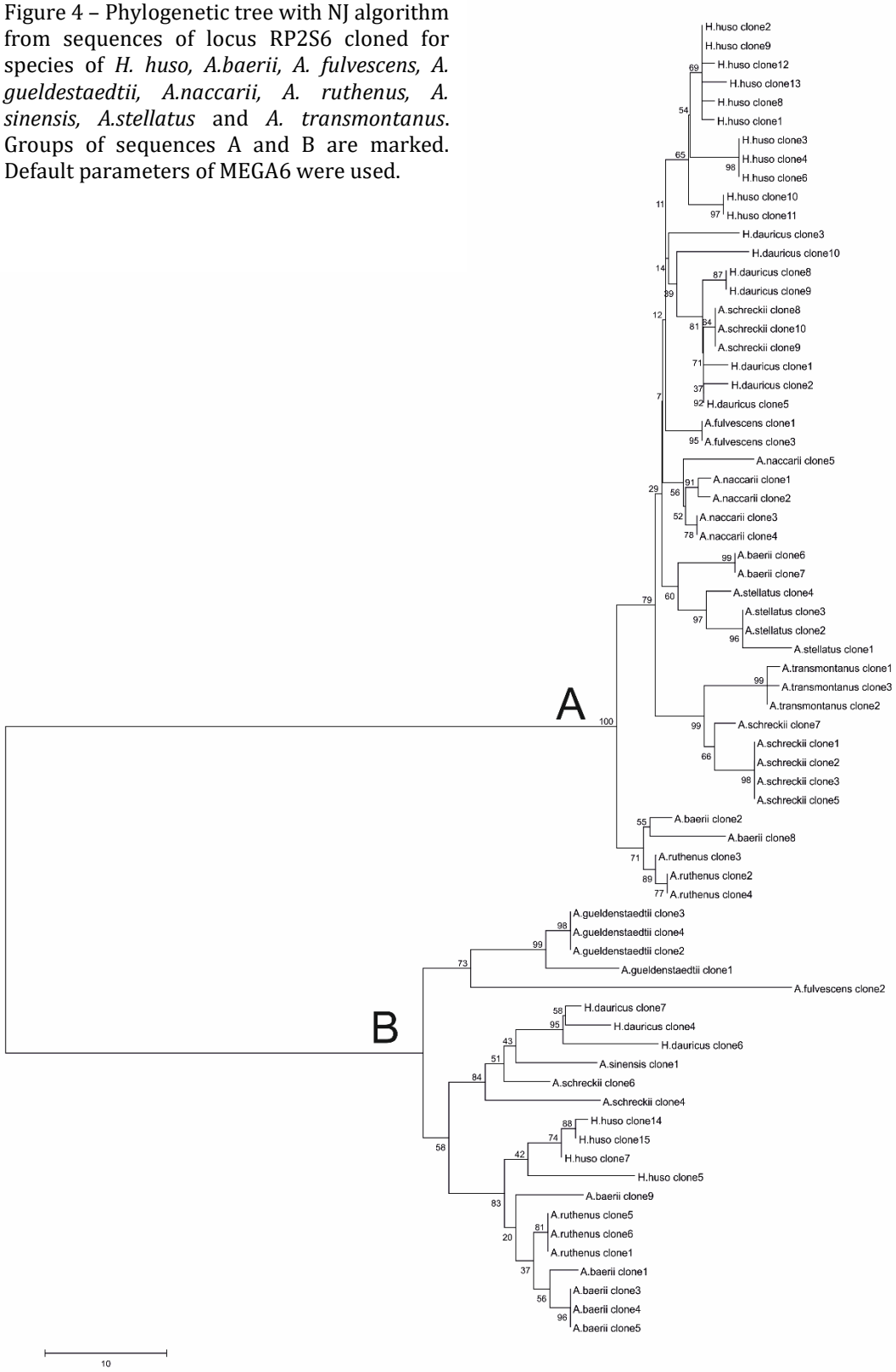
3. RESULTS

3.1. Search for a species-specific marker for the identification of pure Beluga sturgeon and hybrids

Two primers were designed on the exon flanking regions (RP2S6_F 5'-TTCATGGGGAAACCCTGCTT3' and RP2S6_R 5'-ATCCTCTGGGTGAGGAGTG-3') and used to amplify and sequence the predicted second intron (RP2) of the ribosomal protein S6 (RP2S6) from 65 individuals of 11 species. The intron has a length of 326 bp in *T. rubipres*, 1778 bp in *D. rerio* and 1629 bp in *L. chalumnae*. In sturgeons, the amplification with RP2S6_F - RP2S6_R primer pair produced a band about 600 bp long. A total of 71 sequences from 65 individuals of 11 species were obtained after cloning.

From the alignment of sequences (Figure 4), two distinct groups (A and B) have emerged. It can be assumed that these two groups are different loci originated by a duplication event: the retention of duplicated portions of the genome is widely recognized in ribosomal genes (Parenteau et al., 2015).

Figure 4 – Phylogenetic tree with NJ algorithm from sequences of locus RP2S6 cloned for species of *H. huso, A.baerii, A. fulvescens, A. gueldestaedtii, A.naccarii, A. ruthenus, A. sinensis, A.stellatus* and *A. transmontanus*. Groups of sequences A and B are marked. Default parameters of MEGA6 were used.



Several putatively diagnostic SNPs for different species (*A. baerii, A. fulvescens, A. gueldestaedtii, A. naccarii, A. ruthenus, A. sinensis, A. stellatus* and *A. transmontanus*) were available in both groups

A and B, however, for the same species other SNPs are available on the intron1 (RP1) of the RPS7 gene (Boscari et al., 2014 a). Thus, the search for intra-species fixed mutations was focused on the group A for which a higher number of sequences was available. In this locus, two potentially diagnostic SNPs for *H. huso* have emerged from the alignment of sequences of all the analysed species from group A (at positions 96 and 304) (Figure 5). For the SNP in position 304, a specific-primer pair was projected and validated, while the specificity of the SNP at position 96 was not further validated. A specific forward primer for Beluga has been designed in correspondence with the diagnostic SNPs previously identified at position 304, to be coupled with a reverse primer (Table 9):

Table 9- Primer specific for the Beluga species.

Primer	**Sequence**	**Specificity**	**Fragment length (bp)**
RP2S6_huso_F	5' - CATAACATTGCACTGAATGTTATA - 3'	*H. huso*	194
RP2S6_groupA_R	5' - CTTTCGTTGATTTAGGGAAATGGT - 3'		

The forward primer is projected with the 3' end on the diagnostic SNP, while the reverse one is selective for the sequences of the group A. Primers pair RP2S6_huso_F – RP2S6_groupA_R was successfully optimized for the thermic profile.

Finally, this primer pair (RP2S6_huso_F – RP2S6_groupA_R) was validated on a total of 321 samples of 11 species: The Beluga-specific band, 194 bp-long, expected from the amplification with the new diagnostic primer, has been successfully amplified in all the Beluga samples, whereas no amplification was observed for any of the non-target species.

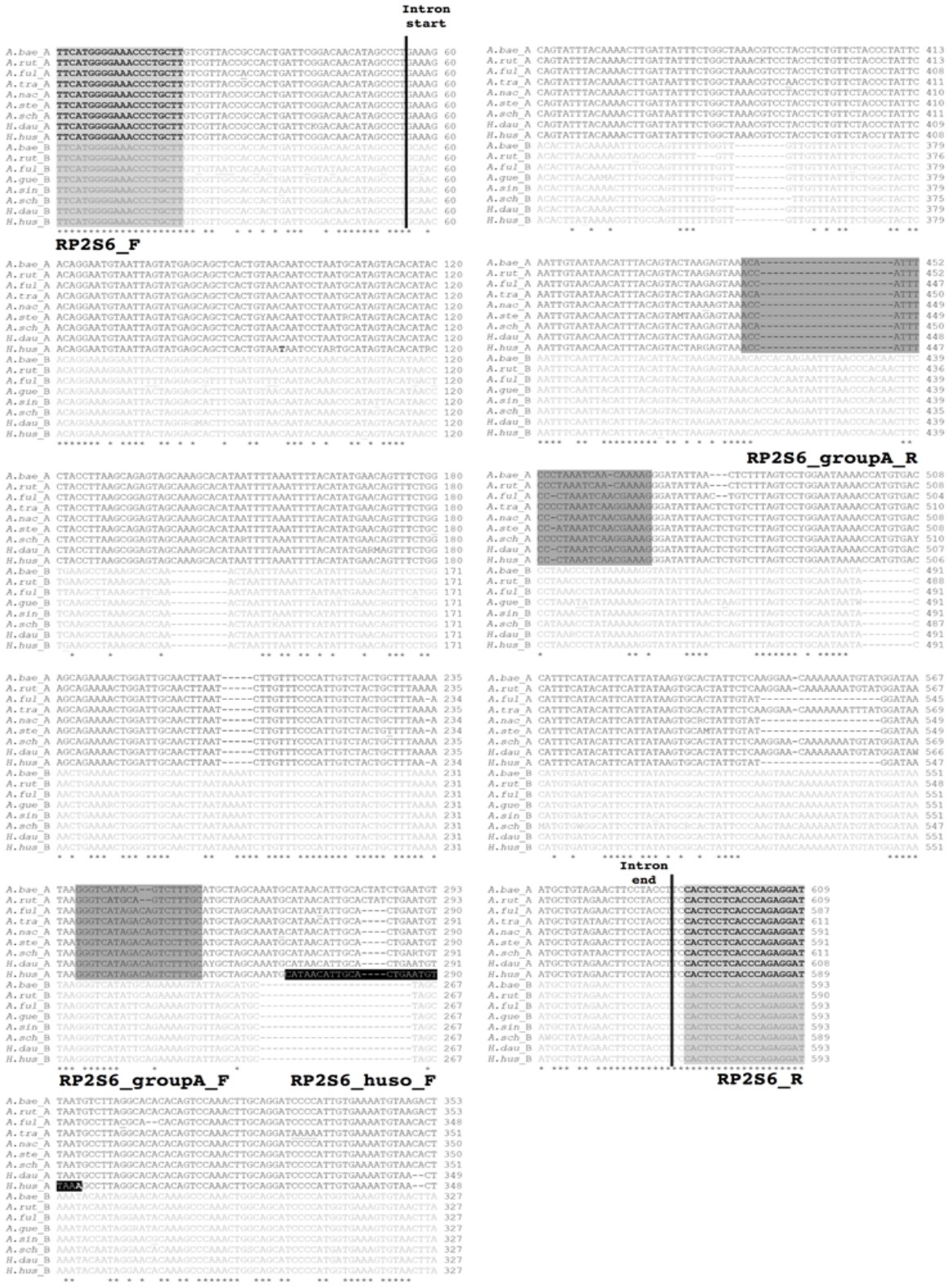


Figure 5 - Multi-species alignment of the second intron of the ribosomal protein S6 (RP2S6, group A and B) including one consensus sequence per species (Accession numbers: KY302717 - KY302733). Shaded regions correspond to primers used. The two SNPs detected for *H. huso* (positions 96 and 294) are marked in bold; only the second one has been validated as a diagnostic marker. Additional putative diagnostic positions for different species, although not validated, are underlined.

All samples were also checked for the amplification of a positive control to confirm the good quality of the DNA extracts. To this end, the primer pair pc_RP1F – RP1_LocusA_R, developed by Boscari et al., (2014 a) was used as a positive control on the RPS7 gene: the expected RP1S7 positive control band, 306 bp-long, was successfully amplified from all the DNA extracts used in this work, thus confirming the good quality of the samples.

These results suggest a 100% identification power for the new diagnostic marker for the identification of the Beluga species. Moreover, given the mendelian heritability of SNP markers, this diagnostic tool can potentially identify the contribution of *H. huso* to all the hybrids generated by the Beluga sturgeon with other sturgeon species (Example in Figure 6).

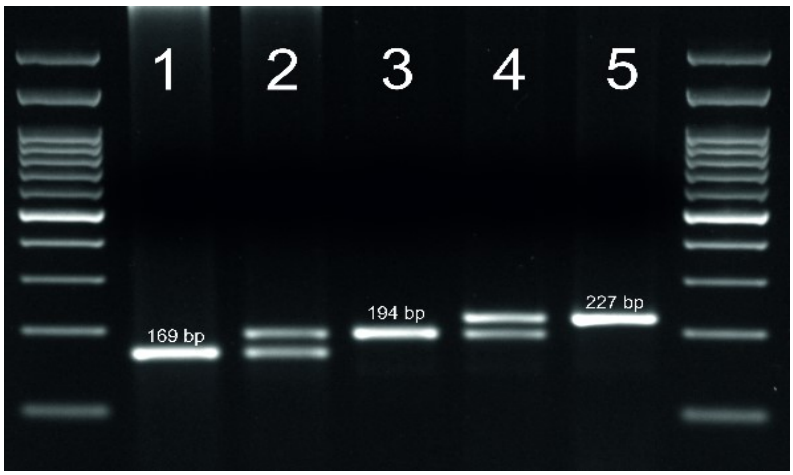


Figure 6- Example of interspecific hybrid identification.
Electrophoresis on agarose gel 1.8% of fragments obtained by the amplification in multiplex with the primer specific for *H. huso* combined with the RP1S7 marker specific for Sterlet (*A. ruthenus*) and for *A. naccarii* (Boscari et al., 2014).
1: *A. ruthenus*;
2: Bester (*A. ruthenus* x *H. huso*);
3: *H. huso*;
4: *A. naccarii* x *H. huso*;
5: *A. naccarii*.

In the light of these results, all the available samples used for subsequent analyses where tested using the Beluga-specific marker in order to exlcude the possibility to deal with hybrids. The markers successfully detected 3 possible not pure Beluga individuals (since no amplification product has appeared as expected in the case of a pure Beluga). These 3 samples where then checked for the *A. ruthens*, *A. stellatus* and *A. naccarii* using a PCR-based tool developed on diagnostic SNPs on the first intron of the S7 gene that allows a reliable identification of sturgeon species and hybrids specifically developed in our laboratory. All the 3 samples showed a positive amplification using the markers for *A. stellatus*. To confirm these results, the D-loop region of those samples was amplified and sequenced. The mitochondrial D-loop showed a matching correspondence with on line sequences of *A. stellauts*, thus confirming the validity of both the species-specific markers.

3.2. Identification of a putative source population: Genetic comparison between Italian extinct population (museum samples) and now existing foreign populations through mitochondrial markers

Several extraction protocols were tested on Formalin-Fixed (FF) samples: two commercial kits, the CTAB standard protocol, and several variants of this standard protocol.

None of the CTAB standard and variants protocols resulted in reliable results. All the successful amplifications obtained with those methods were found to be the result of a contamination with fresh sturgeon DNA. The commercial kit Qiagen DNeasy Blood and Tissue Kit preceded by a pre-treatment for FF samples, recommended by the protocol, also yielded contaminated amplification. The only reliable results were obtained after using the QIAamp DNA FFPE Tissue Kit, performing the procedure in a laboratory specifically dedicated to ancient DNA analyses at the University of Florence.

A total of 5 amplified fragments from FF tissues and 95 fragments from dried tissues/bony scutes were obtained (Table 11). Since replicated fragments obtained from FF tissues showed C → T and G → A transitions, only fragments amplified from dried tissues and bony scutes were considered for further analyses (Figure 7).

The synthesis of successful amplifications obtained from the different samples is reported in Table 10: for each sample, the nine short overlapping regions in which the control region was divided are separately reported (Figure 8).

Table 10- Amplified fragments from DNA extract of dried tissues and bony scutes.

	Primer 2	**Primer 3**	**Primer 4**	**Primer 5**	**Primer 6**	**Primer 7**	**Primer 8**	**Primer 9**
Sample M1 (7828)	6	6	6	5	2	3	6	6
Sample M2 (20312)	-	2	-	-	-	-	-	-
Sample M3 (21443)	5	4	5	3	2	1	5	2
Sample M4 (21337)	2	2	2	-	2	-	1	-

```
consensus ATACCATAATGTTTCATCTACCATTAAATGGTGTACACCATTTTC  45
Ampl.1    .....................................T.....  45
Ampl.2    ......C..................G........G.....  45
Ampl.3    ----------------------------------.....  45

consensus TCTATGTGTACTAACATGTCCCTTCCTAAAACCATAACATGTAAT  90
Ampl.1    ..................T.....................  90
Ampl.2    ..................T.....................  90
Ampl.3    ..................A.....................G.  90
```

Figure 7 - Example of comparison between replicates: DNA sequences from 3 different amplifications from the same FF museum sample.

Table 11- Sequenced fragments from DNA extract of dried tissues and bony scutes. Sample 2 was discarded from the analyses due to the low number of sequences obtained.

Successfully sequenced fragments (including replicates)		Specimes source	Specimes type
Sample M1 (7828)	40	Natural History Museum of Venice	Dried tissues and bony scutes
Sample M2 (20312)	2	Natural History Museum of Venice	Bony scutes
Sample M3 (21443)	37	Natural History Museum of Venice	Dried tissues and bony scutes
Sample M4 (21337)	16	Natural History Museum of Venice	Dried tissues and bony scutes

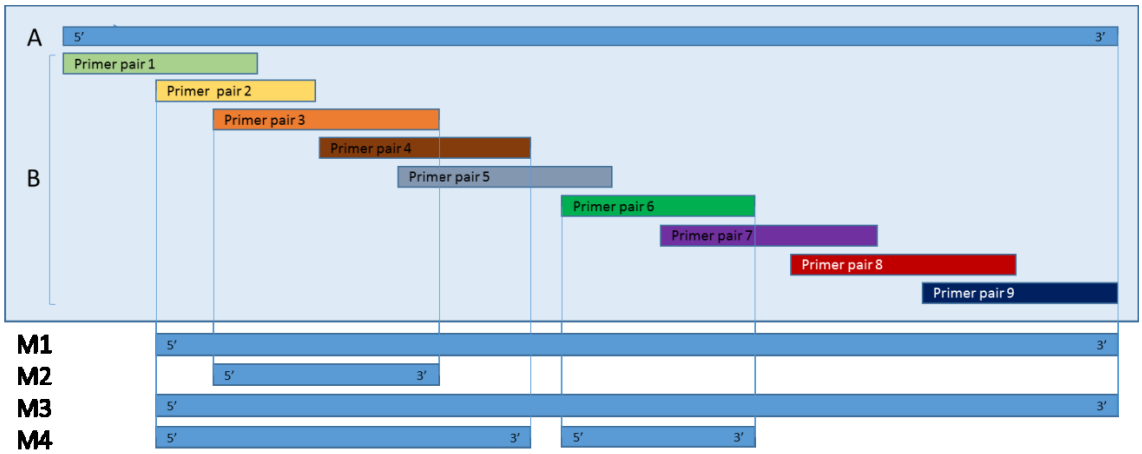


Figure 8- Schematic representation of the consensus fragments obtained for each museum sample:
A, mitochondrial D-loop region (794 bp);
B, fragments obtained from each pair of primers;
M1, museum sample 7825;
M2, museum sample 20312;
M3, museum sample 21443;
M4, museum sample 21337.

For the M1 and M3 samples, the positive amplification of 8 primer pairs out of 9 and the successful sequence of all the PCR products provided a coverage of almost the entire mtDNA Control Region from which a 367 bp consensus sequence was then used for comparison analyses with contemporary samples (Figure 9).

```
M1 ACATGCCCCTTCCTAAAACCATAACATGTAGTAAGAGCCGAACATCCTTGTCTGTCTAGAACATGAA 67
M3 ACATGCCCCTTCCTAAAACCATAACATGTAGTAAGAGCCGAACATCCTTGTCTGTCTAGAACATGAA 67

M1 GTTAATGAGATGAAGGACAATAACCGTGAGATTTCATAACTGAATTATTACTGGCATCTGGCTCCTA 134
M3 GTTAATGAGATGAAGGACAATAACCGTGAGATTCCATAACTGAATTATTACTGGCATCTGGCTCCTA 134

M1 TTTCAAGTCCATCAACAGTTATTTCCCCATAATAAGTTTCCACTGGCATCTGATTAATGTTAGAGGT 201
M3 TTTCAAGTCCATCAACAGTTATTTCCCCATAATAAGTTACCACTGGCATCTGATTAATGTTAGAGGT 201

M1 ACCATAAACTCATGACCCCACATGCCGAGAACCCCACCAACATTTGGTTACTCTTATTTCGGTTTCC 268
M3 ACCATAAACTCATGACCCCACATGCCGAGAACCCCACCAACATTTGGTTACTCTTATTTCGGTTTCC 268

M1 ATTCATTGACATGCAAAGCTCCTTCAGAAAGGATTAATAAGGTGGAACATTCAACAGCCTGCTCGCA 335
M3 ATTCATTGACATGCAAAGCTCCTTCAGAAAGGATTAATAAGGTGGAACATTCAACAGCCTGCTCGCA 335

M1 GGTAATAGATAGTGAATGATATAATGACATAT 367
M3 GGTAATAGATAGTGAATGATATAATGACATAT 367
```

Figure 9 – consensus sequences of the Adriatic museum samples M1 (7825) and M3 (21443). Polymorphic sites are highlighted in bold.

For the M4 (sample 21337) only 4 out of 9 fragments were successfully amplified for a total of 347 sequenced nucleotides (Figure 8). These portions were identical to the one from sample M3. For the sample M2, only the primer pair 2 yielded positive amplification. By blasting the museum samples against online database all museum sequences matched with Beluga sequences.

Given that M2 and M4 sequences were too short to be compared with fragments obtained from the other samples, they were both discarded from subsequent analyses and only sequences obtained from M1 and M3 samples were used for further analyses.

For what concerns the extant populations, since the presence of familiar groups is very likely within YOY (young of the year), these were excluded from mitochondrial analysis, and only adults were included (more in-depth analyses with microsatellite markers will be presented later).

The haplotype diversity was high for all the extant populations ranging from 0.829 ± 0.049 to 0.958 ± 0.028. Regarding the nucleotide diversity, the difference between two sequences is about 2%, on average (overall π = 0.017 ± 0.009) (Table 12).

Table 12- For each basin, the number of individuals (**N**), the number of haplotypes ($\mathbf{N_h}$), the haplotype diversity (***h***) and the nucleotide diversity ($\boldsymbol{\pi}$) with their standard errors are reported.

	N	$\mathbf{N_h}$	***h***	$\boldsymbol{\pi}$
Danube River	44	12	0.863 ± 0.026	0.015 ± 0.008
Black Sea	20	14	0.958 ± 0.028	0.017 ± 0.009
Sea of Azov	15	5	0.829 ± 0.049	0.023 ± 0.013
Ural River	20	11	0.895 ± 0.052	0.018 ± 0.010
Caspian Sea	20	15	0.953 ± 0.035	0.016 ± 0.009

Table 13 shows the values of the ϕ_{st} indices between pairwise populations.

Table 13- Below the diagonal: ϕ_{st} values. Above the diagonal: corresponding values of significance (*p-value* <0.05). Significant values, obtained after multiple test adjustment by the SGoF+ method (Carvajal-Rodriguez and de Uña-Alvarez, 2011), are bold.

	Sea of Azov	***Black Sea***	***Caspian Sea***	***Ural River***	***Danube River***
Sea of Azov	-	**<0.001**	**0.008**	0.090	**<0.001**
Black Sea	**0.11913**	-	0.519	**<0.001**	0.413
Caspian Sea	**0.09238**	-0.00401	-	**0.001**	0.107
Ural River	0.05748	**0.10042**	**0.09300**	-	**<0.001**
Danube River	**0.18447**	0.02157	0.02548	**0.17179**	-

It is evident that genetic similarity between haplotypes does not reflect a geographical proximity between basins. The Sea of Azov population appears to be genetically differentiated from all the other sampled sites with the exception of the Ural River population. Moreover, the Caspian Sea and the Ural River populations are differentiated from each other in contrast with the fact that these two populations are geographically connected. The Black sea population moreover does not appear to be genetically differentiated from the Caspian Sea population, despite the geographical separation of those two basins as shown by Dudu and colleagues, (2014). The Analysis of molecular variance for the five populations of *H. huso* considered separately shows that 92.38% of the variance is explained by intra-population diversity while 7.62% is among populations.

Thirty-nine different haplotypes were identified by TCS software on a total of 121 individuals used for the analyses (including all extant populations and the two Adriatic museum samples). The geographical distribution of the different haplotypes is shown in Figure 10. It seems clear that it is impossible to clearly identify distinct haplogroups related to the sampling locations: the geographical distribution did not affect the haplotype distribution.

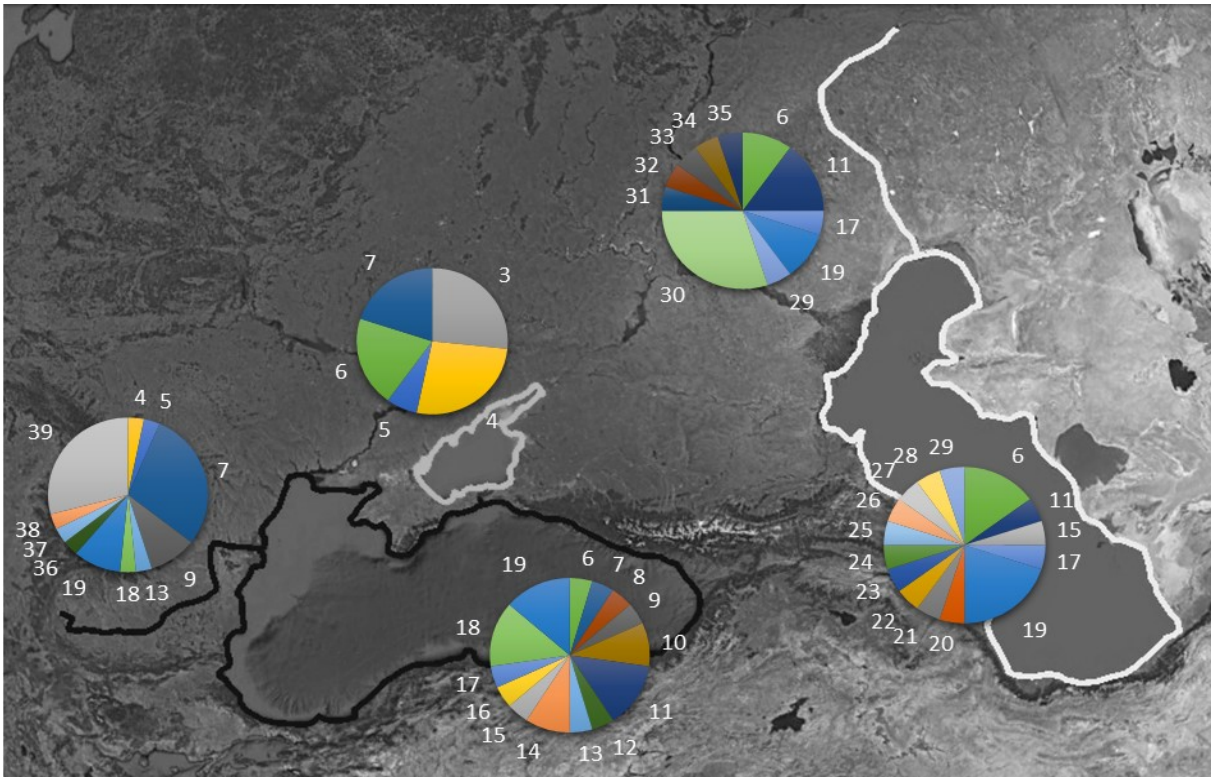


Figure 10 - Map of the distribution of haplotypes in the different sampling locations (Danube River, Black Sea, Sea of Azov, Ural River and Caspian Sea). Pie charts show the frequencies of different haplotypes (referred as different colours and numbers, from1 to 39, corresponding to Table 14) within sampling locations.

The relationships between haplotypes were organized in a network, as shown in Figure 12. The network shows that the clustering of different populations is not resolved: populations of different basins are not unambiguously clustered separately. The network configuration shows a complex structure with many loops (full black circles), many haplotypes are connected by multiple mutational steps and the majority of the haplotypes shows low frequency and are identified in one single basin. Moreover, only 5 haplotypes (out of 39) resulted to be shared between populations of geographically separated basins (Danube River-Black Sea-Azov Sea basin on one side and Caspian Sea-Ural River basin on the other side). The relative frequency of each haplotype is reported on Table 14.

Table 14- Relative haplotype frequencies in populations.

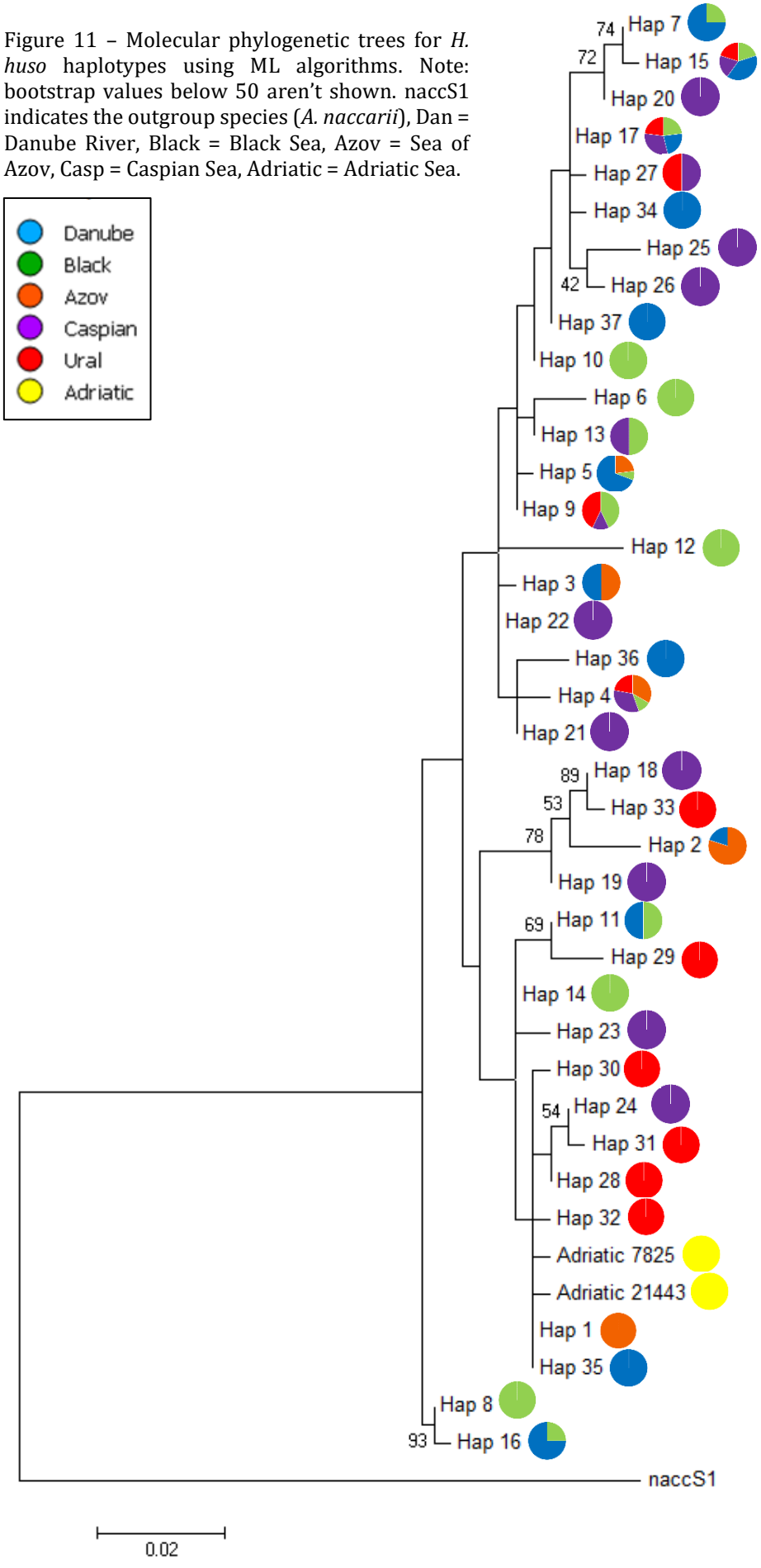
Haplotype:	*Danube River*	*Black Sea*	*Sea of Azov*	*Ural River*	*Caspian Sea*
Hap_1	0	0	0.267	0	0
Hap_2	0.022	0	0.267	0	0
Hap_3	0.022	0	0.067	0	0
Hap_4	0	0.05	0.2	0.1	0.15
Hap_5	0.2	0.05	0.2	0	0
Hap_6	0	0.05	0	0	0
Hap_7	0.067	0.05	0	0	0
Hap_8	0	0.1	0	0	0
Hap_9	0	0.15	0	0.15	0.05
Hap_10	0	0.05	0	0	0
Hap_11	0.022	0.05	0	0	0
Hap_12	0	0.1	0	0	0
Hap_13	0	0.05	0	0	0.05
Hap_14	0	0.05	0	0	0
Hap_15	0.044	0.05	0	0.05	0.05
Hap_16	0.067	0.05	0	0	0
Hap_17	0.067	0.15	0	0.1	0.2
Hap_18	0	0	0	0	0.05
Hap_19	0	0	0	0	0.05
Hap_20	0	0	0	0	0.05
Hap_21	0	0	0	0	0.05
Hap_22	0	0	0	0	0.05
Hap_23	0	0	0	0	0.05
Hap_24	0	0	0	0	0.05
Hap_25	0	0	0	0	0.05
Hap_26	0	0	0	0	0.05
Hap_27	0	0	0	0.05	0.05
Hap_28	0	0	0	0.3	0
Hap_29	0	0	0	0.05	0
Hap_30	0	0	0	0.05	0
Hap_31	0	0	0	0.05	0
Hap_32	0	0	0	0.05	0
Hap_33	0	0	0	0.05	0
Hap_34	0.022	0	0	0	0
Hap_35	0.022	0	0	0	0
Hap_36	0.222	0	0	0	0
Hap_37	0.222	0	0	0	0

The haplotypes characterized in the two museum individuals from the Adriatic Sea were not detected in the contemporary populations. These two haplotypes differed from each other only

for 2 mutations that can be considered highly reliable since supported on average by 4 sequences obtained independently (Table 10, reported above).

Phylogenetic trees based on the D-loop fragments of the different populations (using *A. naccarii* haplotype as outgroups, Figure 11) showed that the haplotypes were distributed randomly along the tree, suggesting that *H. huso* had no definable geographical structure (the extended tree is reported in appendix I).

Figure 11 – Molecular phylogenetic trees for *H. huso* haplotypes using ML algorithms. Note: bootstrap values below 50 aren't shown. naccS1 indicates the outgroup species (*A. naccarii*), Dan = Danube River, Black = Black Sea, Azov = Sea of Azov, Casp = Caspian Sea, Adriatic = Adriatic Sea.



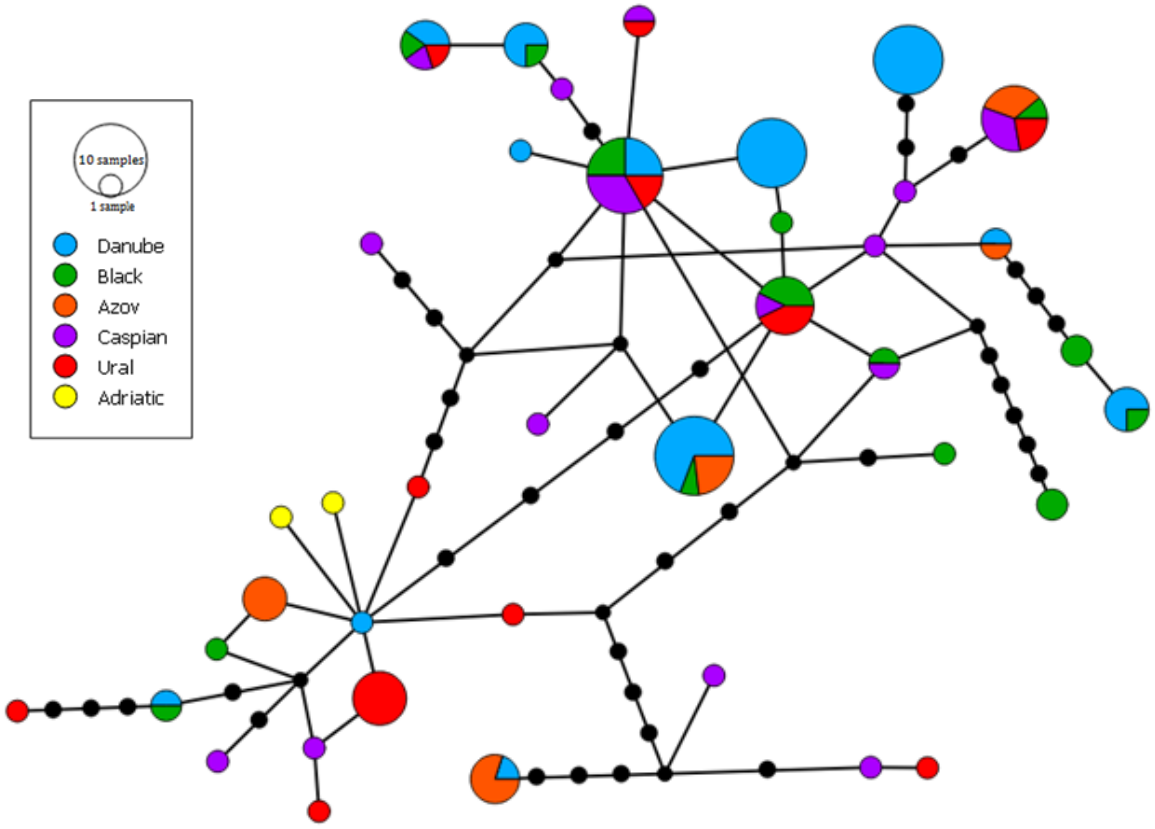


Figure 12- Haplotype network. Size of pie charts is proportional to the corresponding haplotype frequency. Colors of pie charts indicate the corresponding sampling locations.

The number of detected haplotypes was relatively high in relation to the number of individuals used for the analyses, thus in order to have a more complete representation of genetic variability at mitochondrial control region, 42 Beluga D-loop sequences taken from on line databases (used by Dudu et al., 2014) were included. These sequences corresponded to 12 haplotypes shared with the previous data set and to 13 new haplotypes. The new dataset (composed by 367 bp sequences as the previous reduced data set) was then used to compute new ϕ_{st} statistics. The results of the comparison between pairwise populations are reported in Table 15:

Table 15- Below the diagonal: ϕ_{st} distances. Above the diagonal: corresponding values of significance (*p-value* <0.05). Significant values, obtained after multiple test adjustment by the SGoF+ method (Carvajal-Rodriguez and de Uña-Alvarez, 2011), are in bold.

	Sea of Azov	***Black Sea***	***Caspian Sea***	***Ural River***	***Danube River***
Sea of Azov	-	**<0.001**	0.085	0.099	**<0.001**
Black Sea	**0.09237**	-	**0.030**	**<0.001**	0.060
Caspian Sea	0.03075	**0.02601**	-	0.104	**<0.001**
Ural River	0.04734	**0.09249**	0.02476	-	**<0.001**
Danube River	**0.15416**	0.03156	**0.07118**	**0.17511**	-

With the largest dataset, the presence of a phylogeographical pattern is noteworthy, thus confirming the hypothesis of a low detection power of the previous analysis due to a limited sample size. These results are supported by the hierarchical AMOVA analysis (Table 16) that showed two groups as the best grouping composition maximizing the variability, one composed by the Caspian Sea, the Ural River and the Sea of Azov, and the other composed by the Black Sea and the Danube River populations.

Table 16- Hierarchical AMOVA, statistical significance was estimated with 10100 permutations.

Source of variation	**Variance Components**	**Percentage of variation**	***p-value***
Among groups	0.206	5.70	<0.001
Among populations within groups	0.111	3.09	0.014
Within populations	3.290	91.21	<0.001

However, the re-performed network analysis and ML tree again showed that the haplotype corresponding to geographically separated basins are not clustered separately (Figure 13 and 14, the extended ML tree is shown in Appendix II). The unresolved haplotypes network is in accordance with a previous study on this species that highlighted the presence of a complex structure of the D-loop control region. This configuration, together with the low nucleotide diversity and the high haplotype diversity, highlighted by analyses may be due to the presence of admixed populations: the present pattern of mitochondrial variability observed in non-extinct populations, might be the result of genetic drift acting independently on recently separated populations.

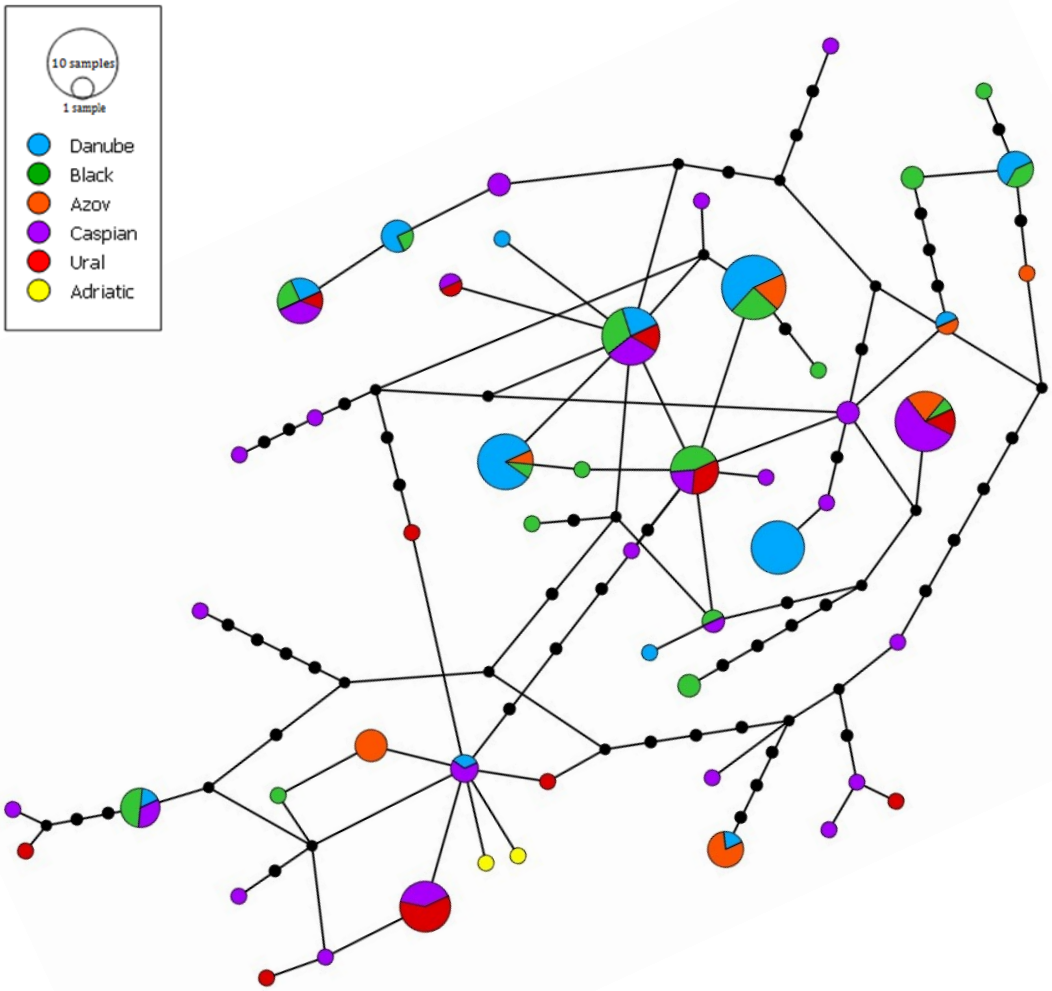


Figure 13- Haplotype network. Size of pie charts is proportional to the corresponding haplotype frequency. Colors of pie charts indicate the corresponding sampling locations.

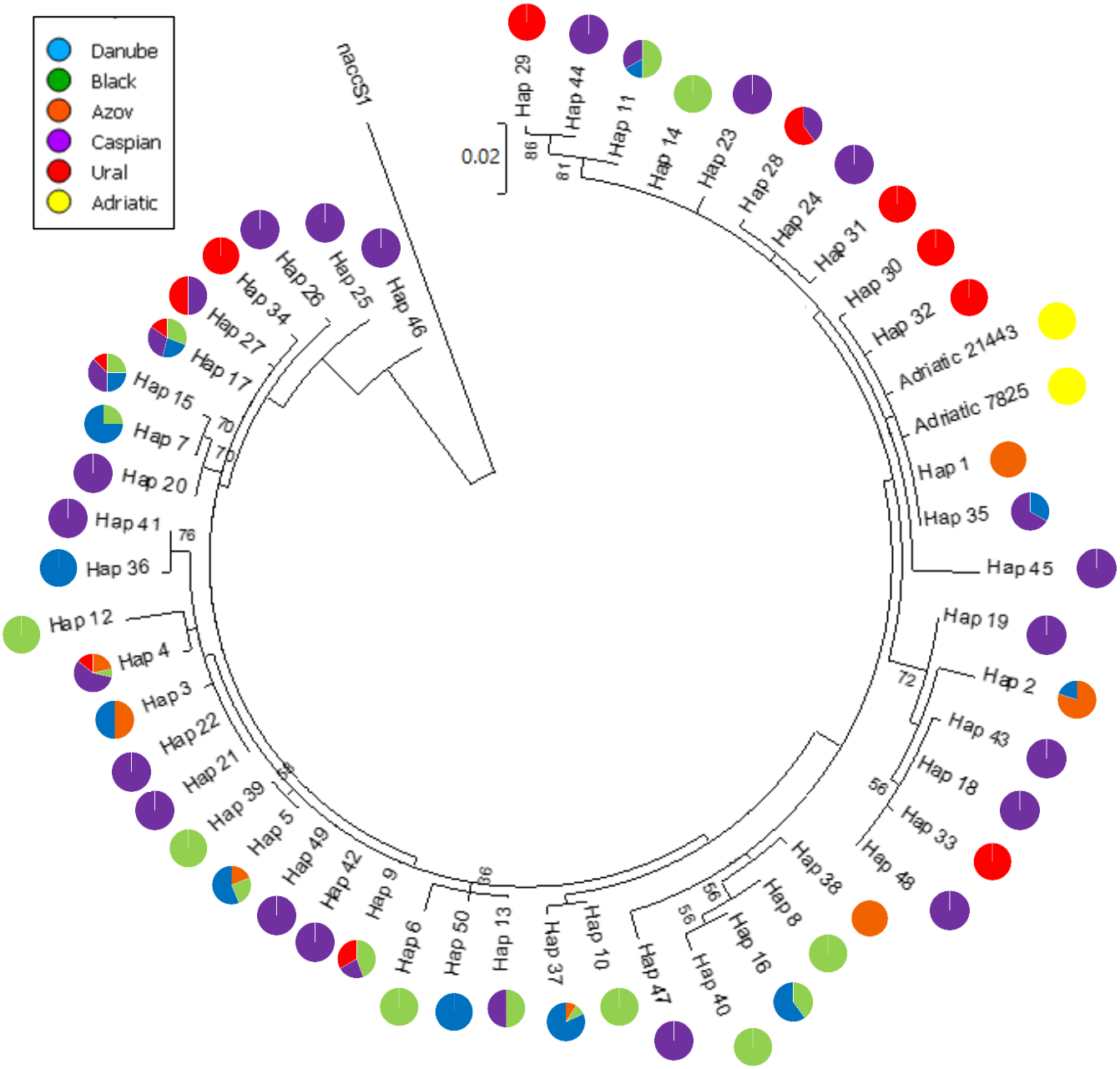


Figure 14 – Molecular phylogenetic trees for *H. huso* haplotypes (including sequences from GenBank) using the ML algorithm. Note: bootstrap values below 50 aren't shown. naccS1 indicates the outgroup species (*A. naccarii*), Dan = Danube River, Black = Black Sea, Azov = Sea of Azov, Casp = Caspian Sea, Adriatic = Adriatic Sea.

It is however interesting to note that the two Adriatic haplotypes were confirmed to be different from all the other haplotypes from the extant populations on a total of 50 different haplotypes identified by TCS software on 163 individuals used for the analyses. The identification of two unique haplotypes in the Adriatic Sea could be a sign of the presence of a genetically distinct population, thus making it impossible to uniquely identify the best source population exclusively on the basis of mitochondrial information.

3.3. Genetic characterization through nuclear markers of the extant Beluga populations to lay the bases for future restocking program

3.3.1. Genetic characterization and phylogeographic analyses of all individuals using microsatellite and SNPs markers

Microsatellite markers:
Ten out of the 27 tested loci have been discarded from subsequent analyses because they showed a multiallelic pattern in place of a maximum of 2 alleles per locus in each individual as expected for a diploid species such as Beluga. One locus (AS-034) was discarded since it showed a monomorphic pattern. The maximum number of alleles per individual per locus for each of the 27 tested loci is shown on Table 17:

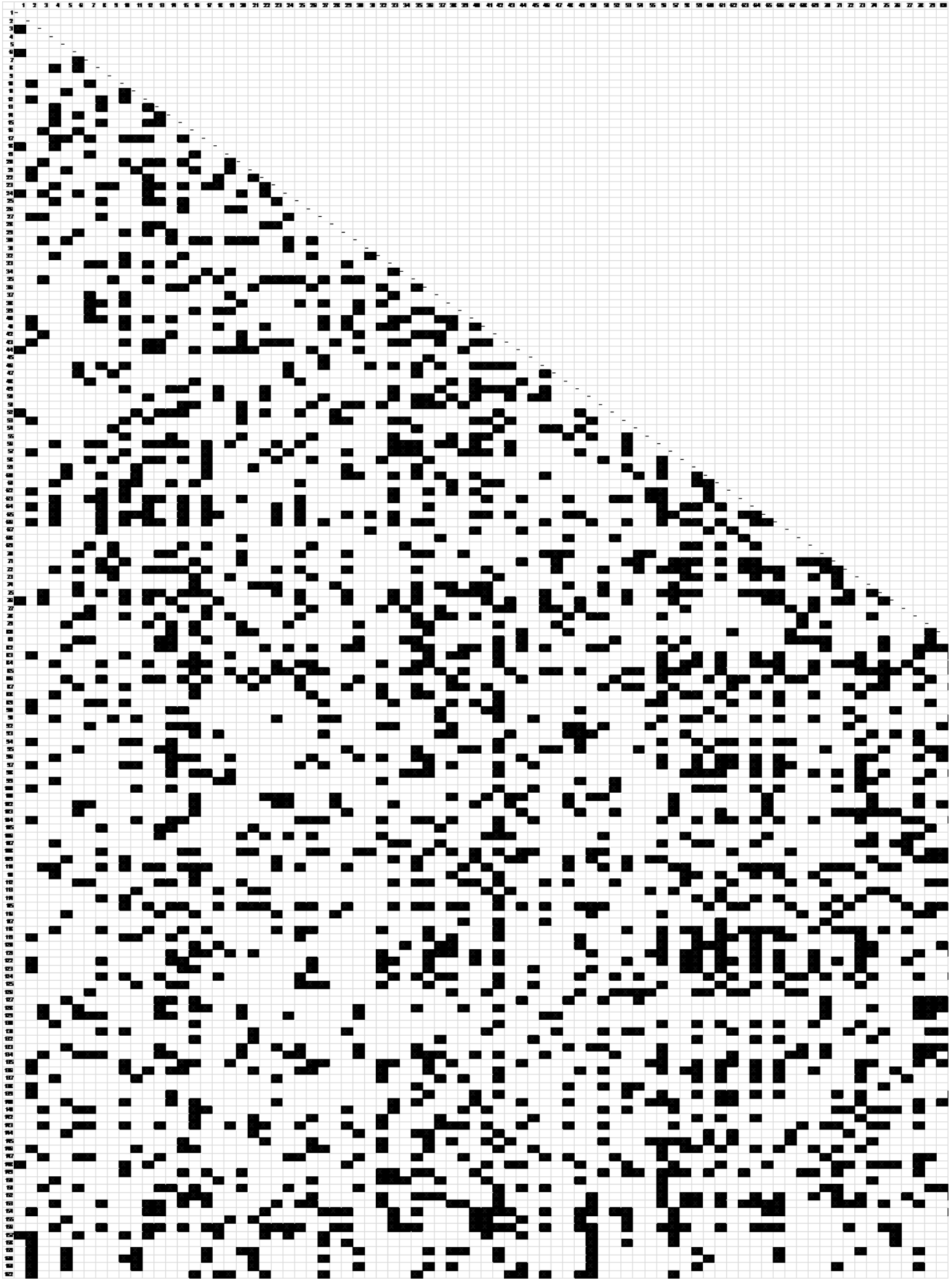
Locus	Maximum number of allele per individual
Anac_tag71527	2
Anac_tag83433	**5**
Anac_tag112145	**6**
Anac_c1490	**4**
Anac_c6179	2
Anac_c33182	**3**
Anac_c7314	2
Anac_c22096	2
Anac_c1538	**3**
Anac_c15214	**3**
Anac_c159	2
Anac_c16902	2
Anac_c6784	2
Anac_c12159	2
Anac_c31601	2
Anac_c5367	2
Anac_c3133	**4**
An20	2
AnacE4	2
Spl168	**6**
Spl163	2
AfuG41	**4**
AfuG112	**3**
AoxD241	2
AoxD64	2
AoxD161	2
AS-034	**1**

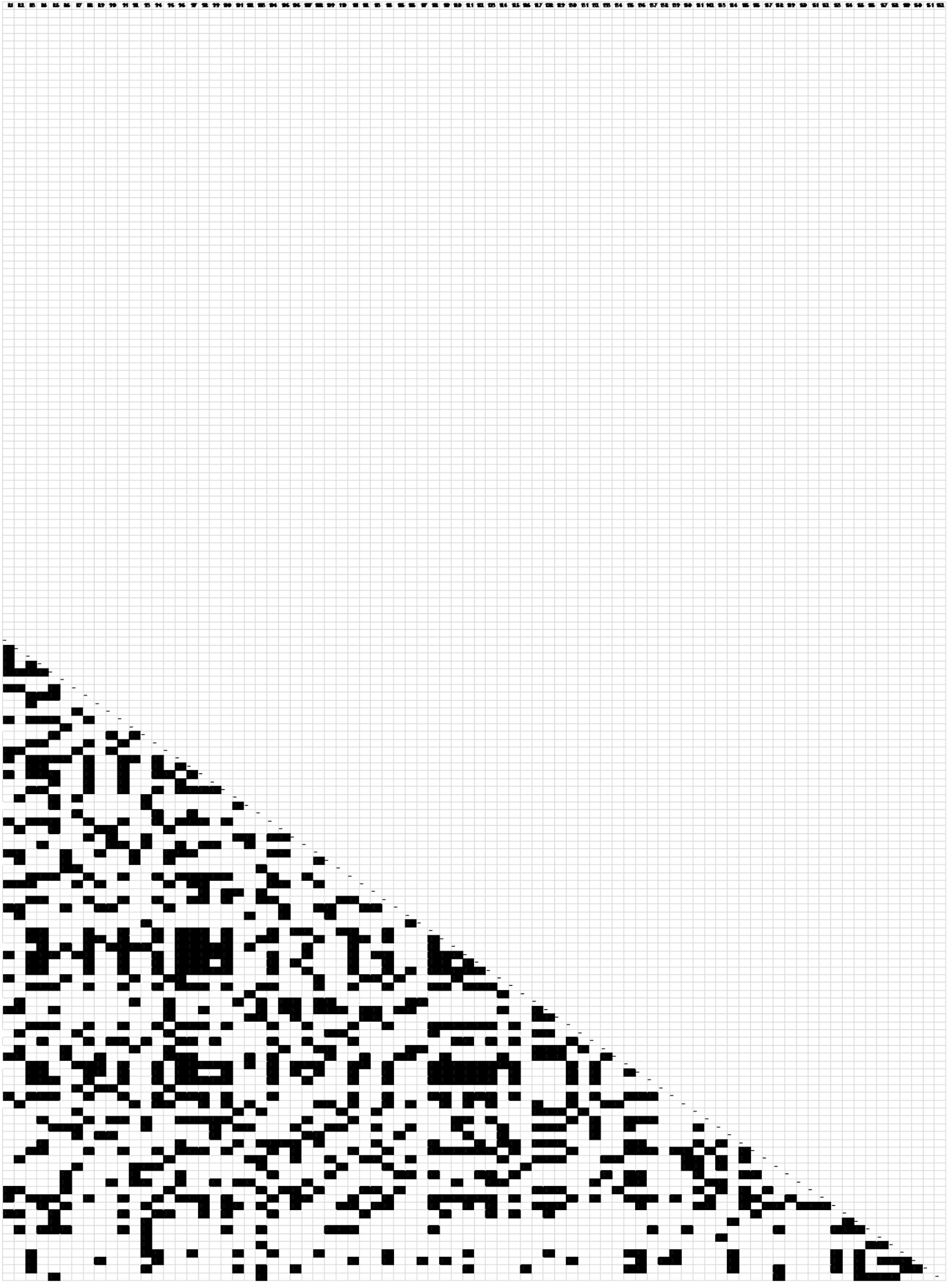
Table 17- Maximum number of alleles per individual per locus for all the microsatellites tested. Discarded loci are in bold.

All the Beluga individuals from the different populations (110 adults and 162 YOY) were then genotyped at the selected 16 loci: a high degree of relatedness (measure of the fraction of alleles shared identical by descent) among several juvenile Danube Beluga individuals has emerged, showing the presence of related individuals (Table 18).

Therefore, in order to eliminate the presence of related juveniles and in order to balance the sample size among the various populations, the YOY have been discarded from further analyses and analysed in detail on Chapter 3.4.

Table 18 (next pages)- Estimated Relatedness (based on r) obtained using ML-Relate: this report shows the relationship between each pair of individuals that has the highest likelihood among four following relationships: unrelated (white), half-siblings, full-siblings and related individuals (Black). For brevity, single samples name on table are reported on hierarchical numbers.





The selected 16 loci genotyped on adult specimens revealed that most loci showed a satisfactory level of polymorphism (minimum mean allelic richness: 4.31; maximum: 13.11), with the exception of 5 loci there were scarcely polymorphic (minimum mean allelic richness: 2.22; maximum: 3.6).

The number of observed alleles (Table 19) and the allelic richness, based on a minimum sample size of 14 diploid individuals (Table 20), are shown below:

Table 19- Number of alleles sampled per locus and populations.

		Danube River	***Black Sea***	***Sea of Azov***	***Caspian Sea***	***Ural River***	***Total***
Loci	*Anac_tag71527*	6	6	6	7	5	7
	Anac_c6784	2	3	2	2	2	3
	Anac_c6179	3	2	2	3	4	5
	Anac_c22096	3	3	3	3	3	3
	Anac_c31601	6	5	5	7	8	10
	Anac_c12159	13	14	10	8	12	17
	Anac_c16902	8	6	8	6	7	9
	Anac_c7314	10	6	6	7	8	11
	Spl163	6	6	6	7	6	8
	Anac_c5367	4	3	3	5	4	6
	Anac_c159	9	9	7	6	7	11
	AoxD16	2	2	3	3	4	4
	AoxD64	9	8	8	10	9	12
	AnacE4	4	5	4	5	5	6
	AoxD24	17	15	10	15	12	25
	An20	7	7	5	5	4	9

Table 20- Allelic Richness per locus and population.

		Danube River	***Black Sea***	***Sea of Azov***	***Caspian Sea***	***Ural River***	***Total***
Loci	*Anac_tag71527*	5.862	5.935	5.874	6.659	4.983	5.925
	Anac_c6784	1.970	2.936	1.999	2.000	1.997	2.216
	Anac_c6179	1.995	1.936	1.999	2.872	3.473	2.452
	Anac_c22096	2.994	3.000	3.000	2.999	2.999	2.996
	Anac_c31601	5.767	4.471	5.000	6.146	7.237	6.254
	Anac_c12159	11.011	12.247	9.588	7.141	10.279	10.739
	Anac_c16902	6.624	5.733	7.613	5.869	6.395	6.266
	Anac_c7314	8.496	5.736	5.873	6.545	7.592	7.512
	Spl163	5.762	5.935	5.738	6.717	5.673	6.285
	Anac_c5367	3.378	3.000	2.875	4.471	3.936	3.600
	Anac_c159	6.762	8.179	6.862	5.673	6.609	7.541
	AoxD16	2.000	2.000	2.875	2.737	3.473	2.478
	AoxD64	7.123	7.381	7.963	9.314	8.651	8.295
	AnacE4	3.409	4.659	3.863	4.737	4.723	4.308
	AoxD24	12.755	13.434	10.000	13.203	11.660	13.112
	An20	5.418	6.595	4.976	4.395	3.673	5.579

Observed-heterozygosity ranged from 0.164 (locus Anac_c6179) to 0.907 (locus Spl163) and expected-heterozygosity ranged from 0.176 (locus Anac_c6179) to 0.909 (locus AoxD241). An excess of homozygotes was found at 3 out of 16 loci (Table 21). At the population level, observed and expected heterozygosity ranged from 0.6523 for the Caspian Sea sampling site to 0.6813 for the Black Sea sampling site and from 0.6105 for the Sea of Azov to 0.6531 for the Danube River, respectively (Table 22).

Table 21- For each locus, the average observed (Ho) and unbiased expected heterozygosity (He) are reported.

Locus	Ho	He
Anac_tag71527	0,784	0,812
Anac_c6784	0,233	0,235
Anac_c6179	0,164	0,176
Anac_c22096	0,522	0,575
Anac_c31601	0,510	0,726
Anac_c12159	0,652	0,861
Anac_c16902	0,744	0,764
Anac_c7314	0,824	0,809
Spl163	0,907	0,805
Anac_c5367	0,642	0,620
Anac_c159	0,865	0,816
AoxD16	0,370	0,370
AoxD64	0,860	0,845
AnacE4	0,624	0,664
AoxD241	0,665	0,909
An20	0,598	0,598

Table 22- Observed (Ho) and unbiased expected heterozygosity (He) across sampling locations.

Population	Ho	He
Danube River	0,6646	0,6531
Black Sea	0,6813	0,6274
Sea of Azov	0,6539	0,6105
Caspian Sea	0,6523	0,6118
Ural River	0,6558	0,611

All loci were tested for Hardy-Weinberg equilibrium and statistic significance was checked for multiple comparisons. Only 3 out of 16 loci were not in equilibrium: one locus (AoxD241) in three populations (Danube River, Caspian Sea and Ural River), one locus (Anac_c12159) in two populations (Danube River and Caspian Sea) and the third locus (Anac_c31601) in only one population (Danube River).

In order to evaluate whether the presence of these loci could bias the results, preliminary analyses (AMOVA and F_{st}) were performed both with the dataset composed only by loci in equilibrium and dataset composed by all the loci, then the results were cross-checked. Since no variation has emerged between the two datasets (with all loci and with loci in equilibrium only), the complete panel of loci was used for further analyses and only results for those data are reported.

To test the power of the microsatellite panel, simulations of statistical power were performed using POWSIM: for all the combinations of generations (t) and population sizes (N_e) the microsatellite data set used in our study is capable to detect significant population differentiation on the basis of FST values as low as 0.005 with a statistical power >93% with both χ^2 and Fisher's exact tests from 1000 replicates.

The genetic differentiation among populations, estimated through F_{st} indices is shown on Table 23. The geographic subdivision between connected and non-connected basins was reflected by the genetic subdivision suggested by the pairwise genetic comparison as shown by F_{st} values.

Table 23- Below the diagonal: F_{st} values among populations. Above the diagonal: corresponding *p-values*. Significant values, obtained after multiple test adjustment by the SGoF+ method, are shown in bold.

	Danube River	***Black Sea***	***Sea of Azov***	***Caspian Sea***	***Ural River***
Danube River	-	0.210	**<0.001**	**<0.001**	**<0.001**
Black Sea	0.004	-	0.086	**<0.001**	**0.011**
Sea of Azov	**0.035**	0.021	-	**<0.001**	**0.001**
Caspian Sea	**0.027**	**0.027**	**0.036**	-	0.525
Ural River	**0.023**	**0.018**	**0.041**	0.001	-

In the comparison between Azov and Black Sea populations, the presence of a differentiation signal has emerged (F_{st} = 0.021; *p-value* = 0.012), which is lost after correction for multiple comparisons (F_{st} = 0.021; *p-value* =0.086).

To test specific biogeographic hypotheses both one level AMOVA and AMOVA with grouped populations were performed: the first analysis showed that 97.75% of the variance was explained by intra-population diversity, whereas 2.25% was among populations.

A first subdivision based on geographic separation of basins (Danube River – Black Sea basin / Sea of Azov basin / Ural River – Caspian Sea basin) was supported by the AMOVA with grouped populations: again, genetic similarity between sampled populations seems to reflect their geographical proximity (highest F_{ct} value = 0.026; *p-value* < 0.001). Using the microsatellite panel developed in this study, the genetic similarity between sampled populations seems to reflect their geographical proximity as shown by F_{st} and AMOVA analyses.

To confirm the concordance between connectivity and genetic similarity, Spatial Analyses of Molecular Variance (SAMOVA) was performed without a constraint on the geographic composition of the groups and was tested on 5 different grouping configurations. The most likely number of groups (k) was 3. These groupings were selected based on the highest significant F_{ct} value (F_{ct} = 0.03; *p-value* < 0.001; Significance tests: 1023 permutations) and again it was

concordant with the geographical origin of the samples (Caspian Sea – Ural River / Sea of Azov / Danube River – Black Sea).

In order to clarify the pattern of genetic variability among the different populations and basins, a structure analysis was performed: assuming no a priori information on sampling origin and using an admixture model. The most probable number of groups (K) inferred by STRUCTURE with microsatellite dataset was 2 (Likelihood = -5220.6). At K = 2, the resulting genetic assemblage (Figure 15), in which the cluster represents the proportional membership (q) of all Beluga adult individuals, showed a higher homogeneity of Caspian and Ural populations compared to the Black Azov and Danube ones. Moreover, by focusing on the genetic composition of the single individuals, many of them seem to be unambiguously ascribable to one of the two identified groups of animals (especially for populations from Azov and Caspian-Ural basins), with only a minor fraction of them showing the mixed contribution of two populations (mainly in individuals from the Danube sea).

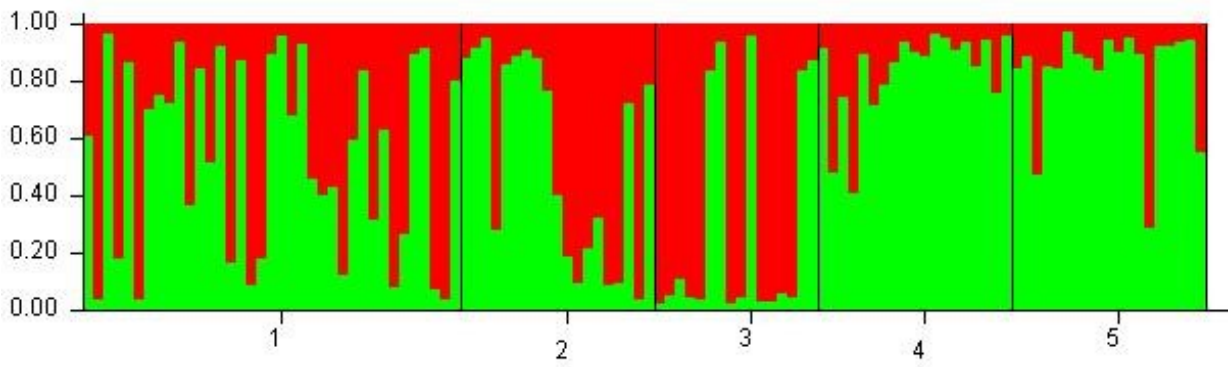


Figure 15- Estimated population structure as inferred by SRUCTURE analysis of microsatellite marker DNA data assuming correlated allele frequencies. Black lines separate sampling sites:
1: Danube River;
2: Black Sea;
3: Sea of Azov:
4: Caspian Sea:
5: Ural River.

This genetic pattern was also confirmed by Principal Coordinates Analysis (PCoA) and DAPC performed on groups composed by the sampling site populations (Figure 16, and 17).

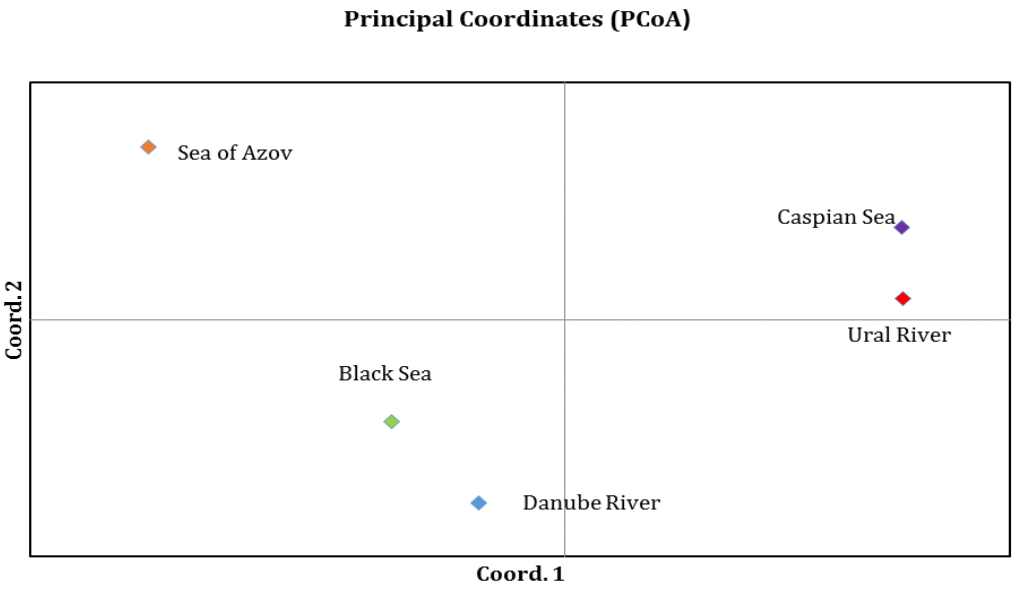


Figure 16- Principal Coordinates Analysis (PCoA) performed with GenAlEx based on F_{st} matrix (Table 23). Distances between points in the plot are close to genetic distances between populations.

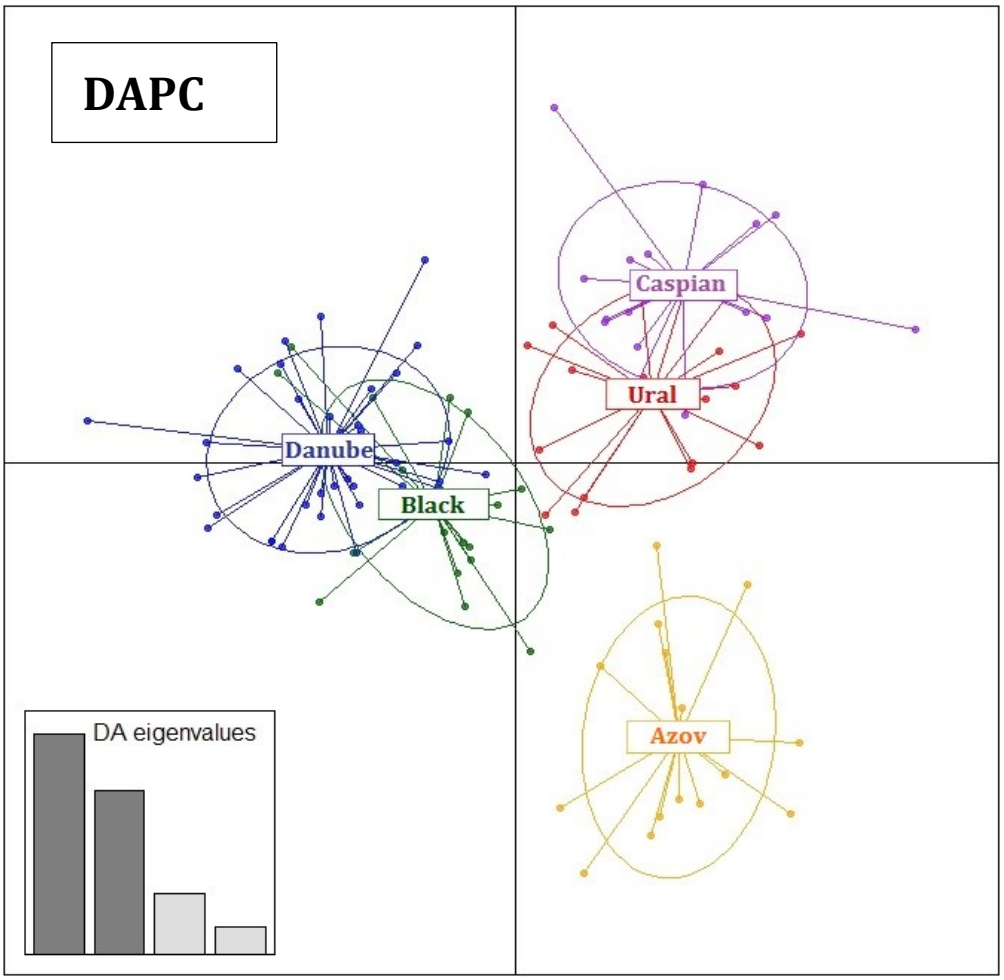


Figure 17 - Discriminant Analysis of Principal Components (DAPC). Clustered dots represent genetically related individuals.

In both cases, individuals from the Sea of Azov cluster separately from the other sampling site populations. Cluster composed by Danube River and Black Sea individuals are genetically as close as the cluster composed by the Ural River and the Caspian Sea populations. These analyses emphasize the correspondence between genetic and geographical distances.

In order to better clarify this correlation, an isolation-by-distance analysis implemented in GENELAND was performed: the Bayesian analysis supported the existence of 3 genetic clusters in the data set. All the runs displayed a clear mode at K=3 which is, hence, the maximum a posterior estimate of K (average log posterior Probability = -1865.9). Spatial patterning indicates the existence of a genetic differentiation between the Black Sea Basin and the Caspian Sea Basin samples. The samples from the Sea of Azov Basin are admixed: the two sampling location were clustered separately with a single Black Sea sampling area (excluded from the cluster containing all the other Black Sea sampling site) on one side and with the Caspian Sea Basin locations on the other (Figure 18).

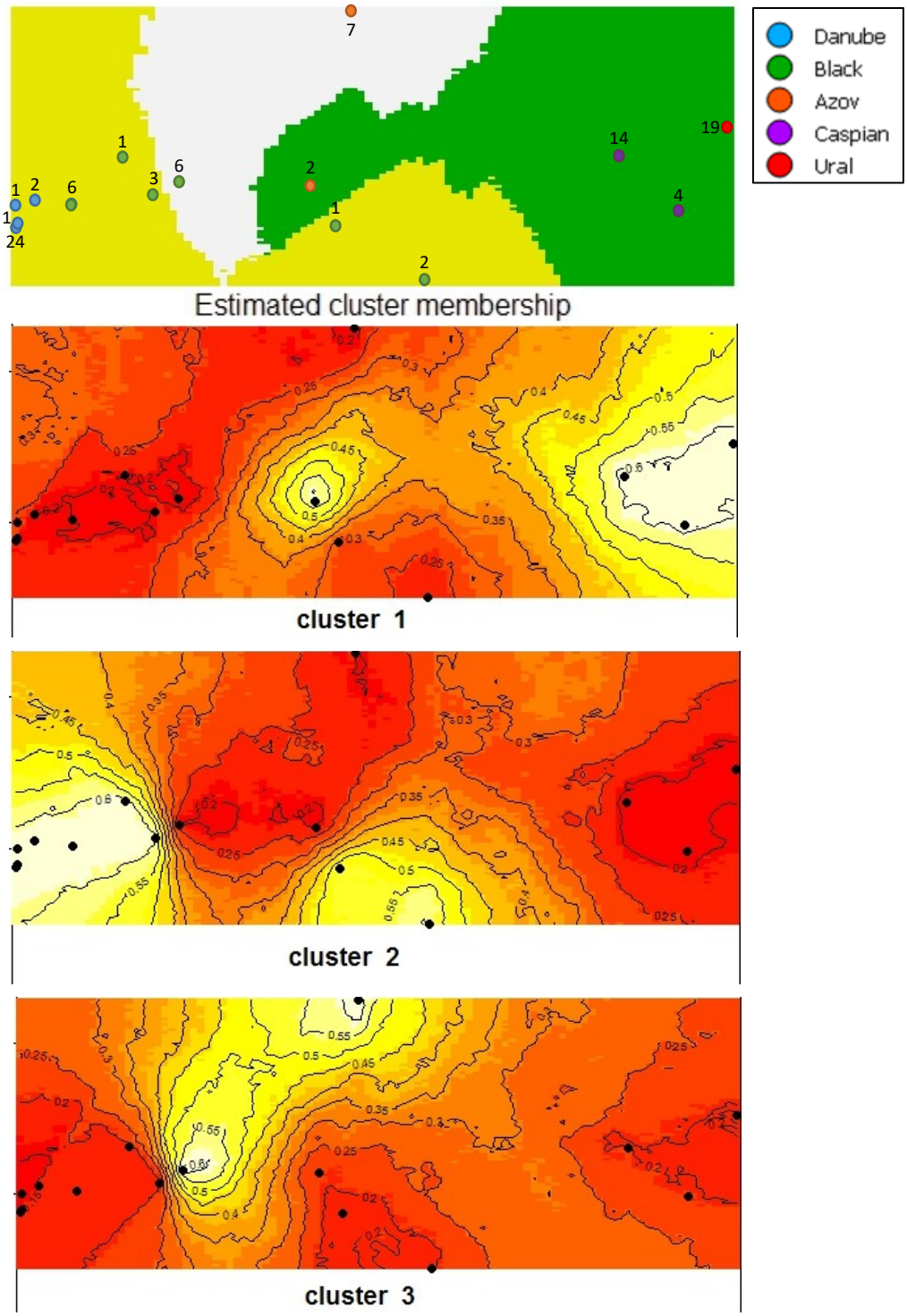


Figure 18 – Maps of posterior probabilities for the three clusters obtained by analysis with GENELAND. Boxes refer to the estimate cluster membership; the first represents the subdivision of all the identified cluster (3) and shows the different sampling location and numerosity; in the others, lighter colour reflects probabilities of belonging to one of the three clusters.

SNPs markers:

In addition to the analyses of microsatellite markers, another approach using other nuclear markers (SNPs) was performed. A total of 485 931 068 de-multiplexed and filtered-quality reads (PF = 37.5) were obtained from the Illumina sequencing (NextSeq and HiSeq) of the 2b-RAD pools. After trimming the adaptors and filtering for restriction site, a total of 335 464 786 reads of 32 bp length were retained (69%), resulting in about 2 942 674 mean reads per individual (Table 24).

Table 24- Population comparison of raw and trimmed reads with the proportion of retained reads.

Population	Total raw reads	Trimmed reads	Retained reads (%)
Sea of Azov	55 488 697	37 762 596	68,1
Black Sea	90 659 830	63 390 974	69,9
Caspian Sea	76 442 308	51 472 908	67,3
Ural River	107 492 795	75 202 818	69,9
Danube River	101 762 281	70 764 741	69,5

The proportion of total reads discarded by trimming is constant in all the individuals: the correlation between the total raw and trimmed reads (r = 0.995; *p-value* < 0.001) is shown in Figure 19.

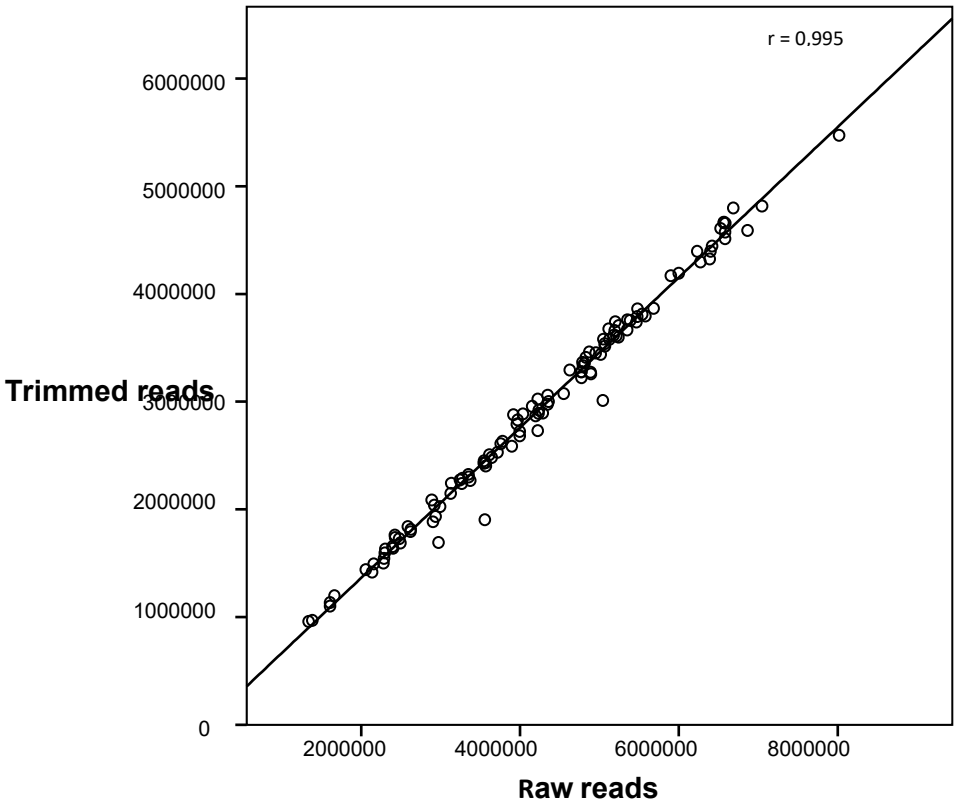


Figure 19 – Correlation between raw (y) and trimmed (x) reads.

A total of 77 082 2b-RAD tags were identified among 114 individuals (of which 51.6% tags were polymorphic and 48.4% tags were monomorphic); for the subsequent analyses, after appropriate filtering, 4 736 polymorphic loci were retained among 73 adult individuals.

The filtered dataset was then used for genetic and population structure analyses. The observed heterozygosity and unbiased expected heterozygosity are generally low. Overall, comparable values of observed heterozygosity, unbiased expected heterozygosity and allelic richness were obtained among all the population samples (Table 25)

Table 25- Observed (Ho), unbiased expected heterozygosity (He) and allelic richness (Ar) across sampling locations.

Population	**Ho**	**He**	**Ar**
Danube River	0,1728	0,1548	1.37
Black Sea	0,1615	0,1436	1.34
Sea of Azov	0,1722	0,1522	1.36
Caspian Sea	0,1597	0,1405	1.34
Ural River	0,1649	0,1449	1.35

The level of genetic differentiation among and within populations was estimated trough one level AMOVA: 99.56% of the variance was explained by intra-population diversity whereas 0.44% was among populations (F_{ST} = 0.00436; p-value < 0.001).

The pairwise F_{st} comparisons between populations and the hierarchical AMOVA reflected the geographical separation between basin (Table 26 and 27).

Table 26- Below the diagonal: F_{st} values among populations. Above the diagonal: corresponding *p-values*. Significant values, obtained after multiple test adjustment by the SGoF method, are shown in bold.

	Danube River	***Black Sea***	***Sea of Azov***	***Caspian Sea***	***Ural River***
Danube River	-	0.826	0.073	**<0.001**	**<0.001**
Black Sea	-0.009	-	**0.011**	**<0.001**	**<0.001**
Sea of Azov	-0.005	**0.03**	-	**<0.001**	**<0.001**
Caspian Sea	**0.001**	**0.006**	**0.012**	-	0.086
Ural River	**0.0004**	**0.005**	**0.010**	-0.002	-

Table 27- Hierarchical AMOVA, statistical significance was estimated with 10,100 permutations. Groups: Black Sea – Danube River / Sea of Azov / Caspian Sea – Ural River

.

Source of variation	Variance Components	Percentage of variation	*p-value*
Among groups	3.456	1.03	<0.001
Among populations within groups	-1.686	-0.50	0.289
Within populations	335.153	99.47	<0.001

The variability pattern emerging from these results is generally low and therefore numerous negative values were obtained from the comparison of populations with no genetic difference. Moreover, all the negative values were associated with no statistical significance supporting the lack of genetic diversification between the compared populations.

In order to support and better visualize the pattern of populations' diversification, a DAPC analysis was performed on groups composed by the sampling site populations (Figure 20).

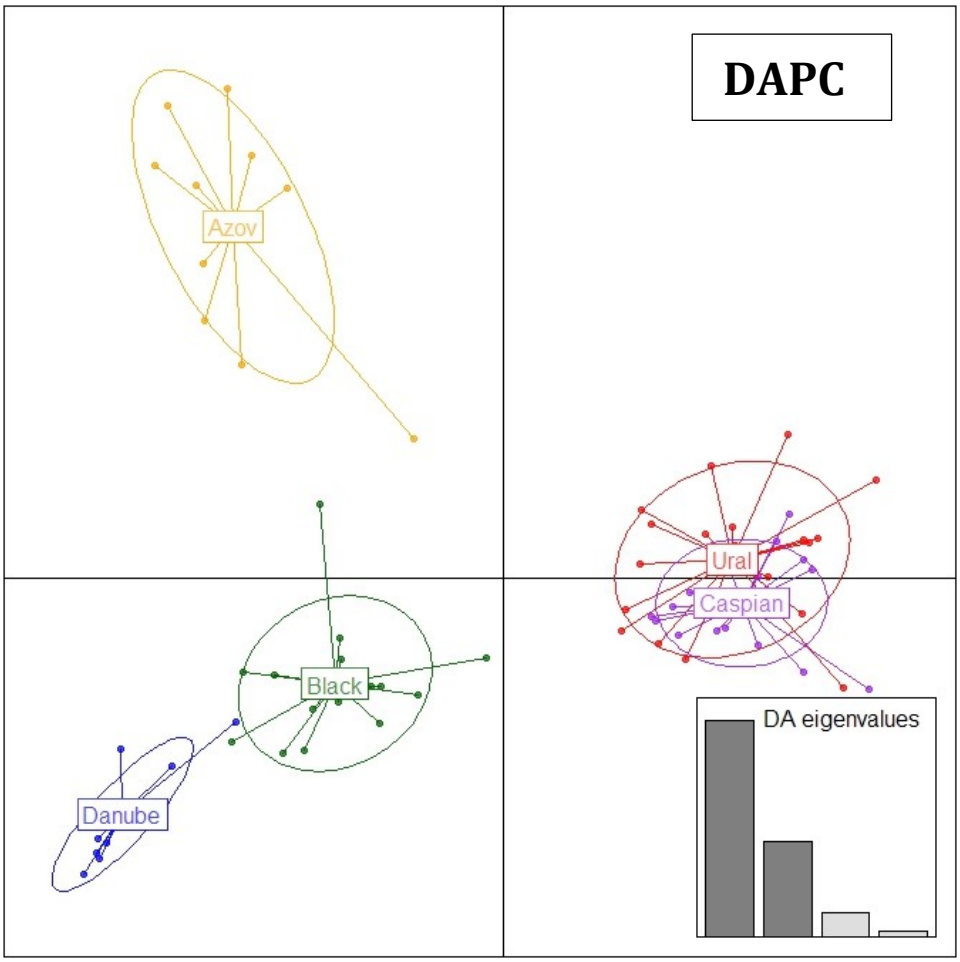


Figure 20 - Discriminant Analysis of Principal Components (DAPC). Clustered dots represent genetically related individuals.

The pattern of diversification and admixture among individuals can be better visualized also with a STRUCTURE-like plot obtained with ADEGENET in which individuals are assigned to clusters (Figure 21).

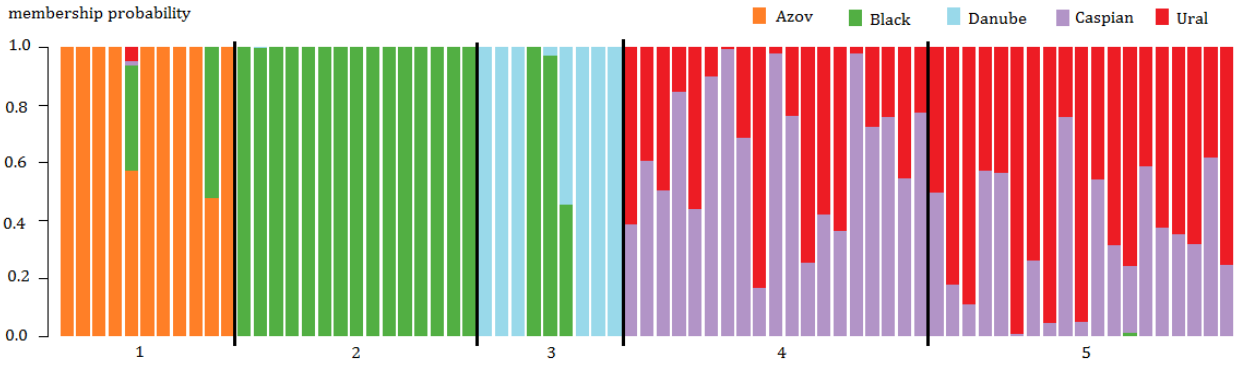


Figure 21 – STRUCTURE-like assignment into the different populations. Black lines separate sampling sites:
1: Sea of Azov;
2: Black Sea;
3: Danube River:
4: Caspian Sea:
5: Ural River.

The structure-like plot shows that the separation between basin is genetically supported: it is interesting to note that the population of the Black Sea/Sea of Azov are mainly composed by individuals assigned to the population of origin, with few individuals assigned to other population of the same basin (the Danube population) or to other basin (the Sea of Azov population). On the other hand, the populations of the Caspian Sea are considered as one mixed group composed by individuals assigned both to the Caspian Sea population and to the Ural River population, suggesting a genetic homogeneity between these two populations.

3.4. Application of the microsatellite panel to relatedness analyses

YOY Danube individuals:

To deeply investigate the genetic relation between YOY individuals a structure analysis including adult individuals was performed in order to identify possible genetically distinct groups in the YOY dataset. The genetic assemblage obtained by STRUCTURE analysis (Figure 22, K = 2, Likelihood = -12507.035) shows that within the Danube River the presence of a group of highly genetically differentiated individuals is noteworthy (later referred as "highly-related" individuals). These individuals are significantly different from all the other populations and are only YOY of the year 2010.

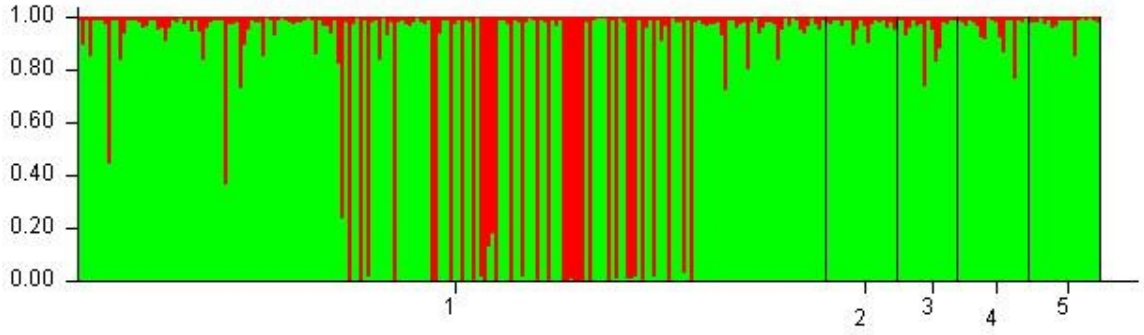


Figure 22- Genetic assemblage as inferred by SRUCTURE analysis of microsatellite marker DNA. Black lines separate sampling sites:
1: Danube River;
2: Black Sea;
3: Sea of Azov:
4: Caspian Sea:
5: Ural River.

To clarify the relationship between those highly-related individuals and to identify the presence of familiar groups, the D-loop of those individuals was sequenced. Two different haplotypes were then identified among the 36 individuals (for 6 individuals the amplification of the D-loop fragment was not successful leading to low quality sequences). A relationship analyses was then performed to validate the grouping of the hypothetical different familiar groups using both ML-relate and COLONY (Figure 23).

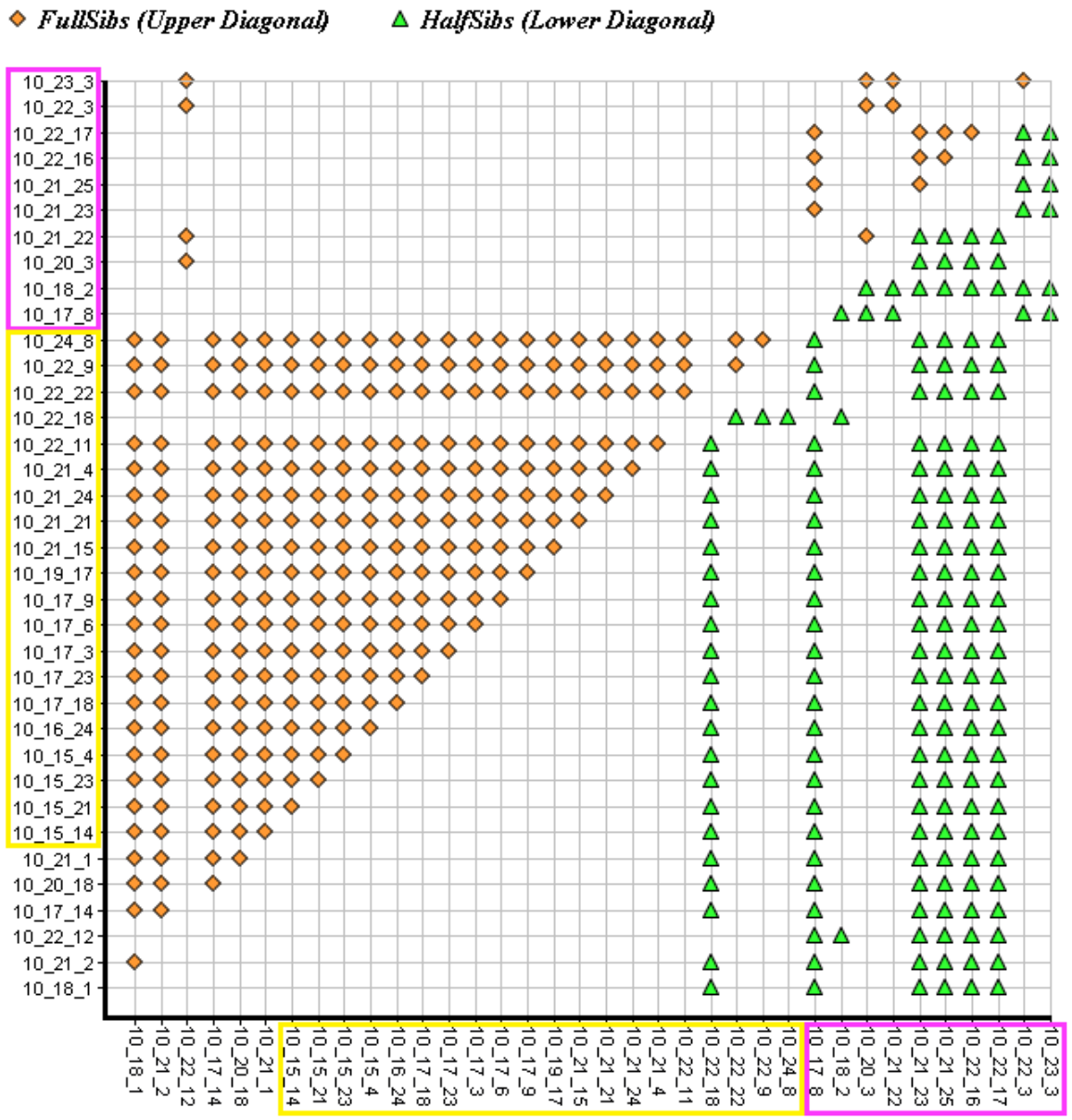


Figure 23 – Best (ML) Sibship Assignment Plot from COLONY of highly related YOY individuals from the Danube population. Individuals corresponding to one of two different haplotypes were grouped (identified with different colours: yellow and purple, no colours indicate that no haplotype was identified for the sample).

It is evident that the majority of the individuals shows a correlation between the identified sibship with the nuclear microsatellite marker and the corresponding haplotype, thus highlighting the good resolving power of the microsatellite panel.

In order to validate the identification of the full-sibs highly-related individuals, a comparison analysis with a data set of full-sibs individuals (n = 1000) generated with Hybridlab was performed: the genetic distribution of highly-related YOY is more similar with the genetic

distribution of full-sibs individuals (the two different plot can be superimposed), thus confirming the related nature of this groups of YOY individuals (Figure 24).

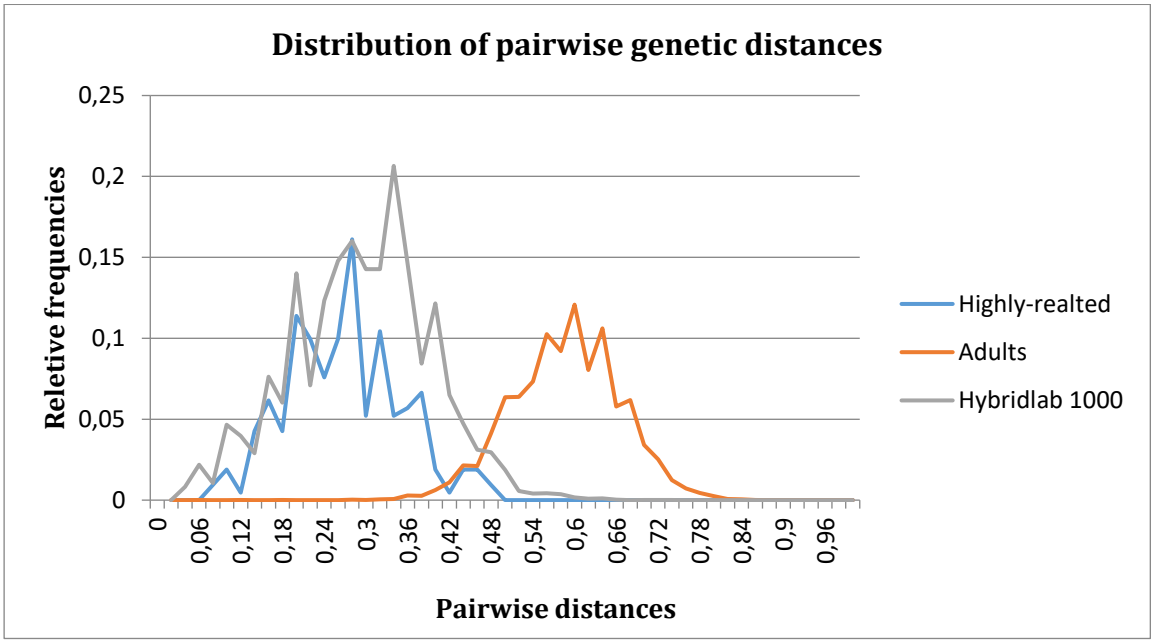


Figure 24 – Distribution of pairwise genetic distances based on band-sharing information for different degree of relatedness.

Individuals from a Livestock:

The 16 selected microsatellite loci for *H. huso* were tested on a sample of individuals from a livestock. The validation process was performed on 9 adults (P), and 12 full-sibs (F1) sired by one female and two males that had to be identified. According to the mtDNA (Table 28) and to information provided by the aquaculture owner that committed the analyses, the adult individuals can be divided in two groups: one composed by fully related individuals (Ind. 3, 5, 6 and 9) and the other composed by the remaining unrelated individuals.

Table 28 - Degree of diversity between sequence pairs (number of polymorphism between the sequences). Comparisons between identical haplotypes (0 differences) are highlighted in bold.

	Ind. 1	***Ind. 2***	***Ind. 3***	***Ind. 4***	***Ind. 5***	***Ind. 6***	***Ind. 7***	***Ind. 8***	***Ind. 9***
Ind. 1	-								
Ind. 2	7	-							
Ind. 3	13	14	-						
Ind. 4	1	8	14	-					
Ind. 5	13	14	**0**	14	-				
Ind. 6	13	14	**0**	14	**0**	-			
Ind. 7	9	2	14	10	14	14	-		
Ind. 8	2	9	15	3	15	15	11	-	
Ind. 9	13	14	**0**	14	**0**	**0**	14	15	-

Using the microsatellite panel for a DAPC, the distribution of genetic variability was consistent with the condition of siblings for individuals 3, 5, 6, 9, as highlighted by mitochondrial data (Figure 25).

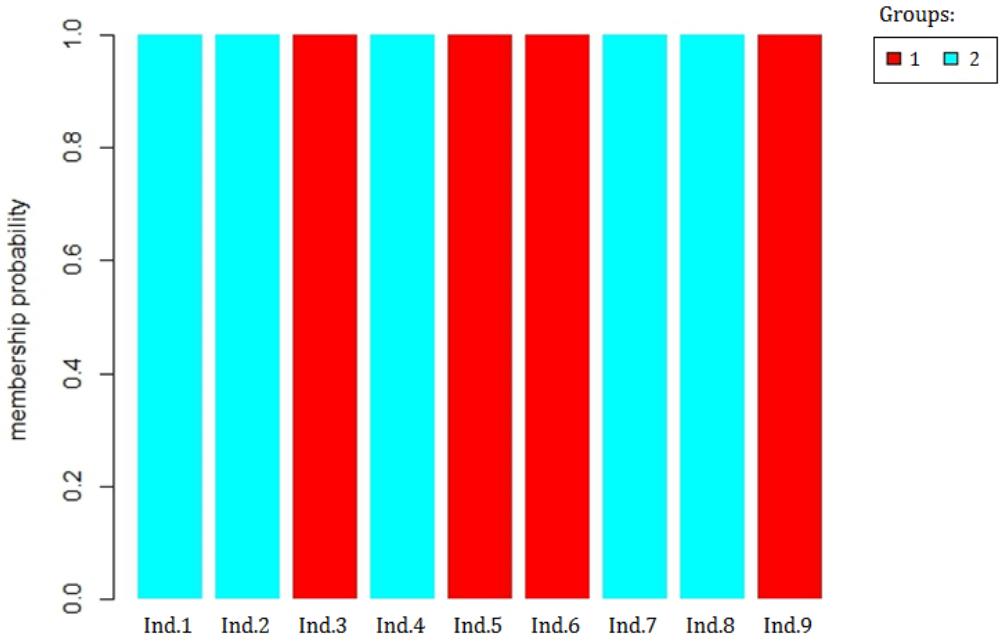


Figure 25 –DAPC: Clusters membership probabilities are represented by different colors.

In addition, the microsatellite panel was successful in the unambiguous identification of the parents of all the individuals, allocating all of them to one dam (as expected) and to one single sire in spite of the two males used for milt collection. As shown on Table 29, it is clear that at least 3 out of the 16 loci are able to unambiguously allocate all the 12 full-sibs juveniles to the respective parents.

Table 29 – Example of allocation with three out of sixteen loci of the panel used. The unambiguous alleles of the putative dam and sire are coloured.

		Locus and Allele													
		Anac_c12159				**Anac_c7314**					**Spl163**				
		241	**248**	**251**	**260**	**179**	**183**	**187**	**191**	**211**	**182**	**186**	**190**	**198**	**202**
Adult	*Ind. 4 ♀*		248	251		179	183				182		190		
	Ind. 5 ♂		248		260			187		211		186			202
	Ind. 9 ♂	241			260			187	191			186		198	
Full-sibs	*Ind. A*			251	260	179			191		182			198	
	Ind. B	241	248			179		187			182			198	
	Ind. C	241		251			183	187				186	190		
	Ind. D	241		251			183	187					190	198	
	Ind. E		248		260		183	187					190	198	
	Ind. F	241	248			179			191			186	190		
	Ind. G	241	248				183		191		182			198	
	Ind. H	241		251		179		187					190	198	
	Ind. I	241	248				183	187			182	186			
	Ind. L	241		251			183		191		182	186			
	Ind. M	241		251			183		191		182	186			
	Ind. N	241		251			183	187					190	198	

In order to support the microsatellite results, another nuclear marker was used to assess the possibility to discriminate between related individuals. The analyses conducted on SNPs markers on adult Beluga individuals for phylogeographic purposes was then applied in order to shed a light on the relatedness of the familiar group of livestock individuals (Ind. 4, 5 and 9 and the 12 full-sibs). Using the DAPC analyses all the individuals fully related (dam, sire and full-sibs) were allocated to the same group, while the unrelated male (Ind.5) was allocated to other groups (Figure 26).

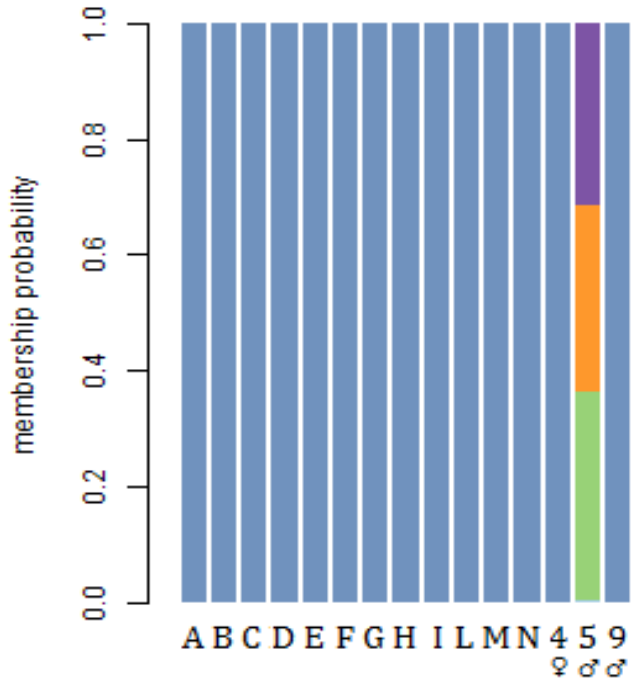


Figure 26 –DAPC: Clusters membership probabilities are represented by different colours. The figure represents the portion of the plot of the entire dataset of Beluga wild samples, this part represents the familiar group including the full-sibs, the dam and the two putative sire. Ind.5 is allocated to other groups (Black Sea, Caspian Sea and Ural River) despite the other individuals are clustered together in a new different group.

4. DISCUSSION

4.1. Search for a species-specific marker for the identification of pure Beluga sturgeon and hybrids

Species of economic relevance that are subject to trade control by CITES, such as sturgeons, are a frequent target of commercial frauds. This aspect has raised the need for reliable, cheap and fast methods for species identification. To this end, various genetic markers have been developed for sturgeons (Congiu et al., 2001, 2002). Generally, the available molecular tools for species identification can be distinguished in two main groups: mitochondrial or nuclear markers, depending on the cellular localization of the analysed genome.

Despite the widespread application of mitochondrial markers, sequencing of mtDNA, such as the D-loop region, is time-consuming and costly, leading to a scarce versatility for routine analysis. However, sequencing of mtDNA is still the most used tool for species identification, since it relies on established methods with evidence-based validation by giving direct checking with public online sequence databases, such as GenBank.

Nonetheless, an alternative approach for the analysis of mtDNA is based on species-specific SNPs on the D-loop control region (Mugue et al., 2008), in which the SNPs are used as target sites for primers used in diagnostic amplifications. However, independently from the approach, it is not possible to identify paternal species in hybrids by using mtDNA, thus strongly limiting its application in diagnostic screening.

To overcome this issue, Boscari and colleagues (2014a) have developed a method based on the identification of species-specific SNPs in the nuclear gene sequence of a ribosomal protein, which leads to the identification by PCR of different sturgeon species with a 100% specificity. This approach has been used for the identification of a specific marker for the Beluga sturgeon in this study. A species-specific SNP was isolated in the second intron (RP2) of the ribosomal protein S6 (RP2S6). The RP2S6 marker resulted as highly reliable for routine controls in the Beluga sturgeon, given its high efficiency and simplicity based on its single-locus diagnostic approach with a fast and easy visualization by agarose gel from a single PCR. The nuclear RP2S6 marker, in combination with pairs of species-specific primers previously developed for numerous and different sturgeon species on locus RP1S7 (Boscari et al., 2014 a), guarantees also the identification of various hybrids that include Beluga sturgeon as a maternal or paternal species. For example, the aforementioned marker RP2S6 has been successfully tested on the hybrid

between Beluga and Sterlet sturgeons (Bester), which is widely bred in aquaculture for the production of meat and caviar of lower economic value than the commercial products from pure Beluga: this makes the Bester a clear example of commercial fraud used to increase profits by deceitful labelling as pure Beluga, thus highlighting the need for unequivocal identification of hybrids.

By using nuclear markers, a sample of Bester is expected to result in a positive amplification in both the specific bands for Beluga and Sterlet parental species, at 194 and 169 bp respectively. A positive outcome at both is considered a sufficient criterion to classify a sample as a Bester hybrid. Recently, Dudu and colleagues (2015) have identified three microsatellite loci (LS54, LS68 e Aox45) with distinct allelic variants in *H. huso* and *Acipenser ruthenus* (Sterlet sturgeon), which have been proposed as diagnostic tools to distinguish the Bester from its parental species (Dudu et al., 2015). However, in many other sturgeon species, the allelic ranges of the above microsatellites are overlapping, thus limiting their application only to two sturgeon species (Beluga and Sterlet), whereas for international trade control all sturgeon species should be unambiguously distinguished.

The nuclear RP2S6 marker is also of particular importance for the identification of inter-specific hybrids in which the Beluga is the paternal species, as it is the case of crossings between a female of *A. naccarii*, *A. baerii* or *A. gueldenstaedtii*, with a male of *H. huso*. Indeed, in these cases the paternal contribution of Beluga is not detectable with mitochondrial markers, given the maternal inheritance of mitochondrial DNA. Therefore, the use of this specific nuclear marker allows to unequivocally identify hybrids in which Beluga is the paternal species. This was confirmed by the positive results obtained from the identification of hybrids with the Adriatic sturgeon (*A. naccarii*) as a maternal species. Recently, another nuclear marker was developed for the discrimination of *H. huso*: this new molecular tool allows the discrimination between pure Beluga and 9 sturgeon species including its hybrids (Havelka et al., 2017).

A further relevant aspect of this study was given by the elevated heterogeneity of Beluga samples used for the validation of RP2S6. Indeed, one of the necessary characteristics for the isolation of a reliable marker for species identification is that all the genetic variability of the target species has to be necessarily covered, otherwise the marker could not be considered universal. For this reason, the main Institutes of Research on the Beluga sturgeon have been involved during the validation step, in order to cover the geographic distribution area of Beluga as wide as possible. Given the correct identification in 100% of cases for such a heterogeneous sample (N = 386), the probability of an incorrect (i.e. false positive or false negative) identification of Beluga in natural populations or aquaculture stocks is reduced at a minimum level.

In addition, from a conservationist point of view, the identification of inter-specific hybrids has a huge impact on the management of stocks of individuals used for breeding in release programs. Unfortunately, hybrids are often accidentally mixed with pure individuals stocks (Congiu et al., 2011; Boscari and Congiu 2014; Boscari et al., 2014 a), since aquaculture plants that usually furnish individuals for reintroduction plans are also involved in commercial activities (Chebanov et al., 2011).

The results of this study were recently published on the journal Food Control (Boscari et al., 2017).

4.2. Identification of a putative source population: Genetic comparison between Italian extinct population (museum samples) and other contemporary populations through mitochondrial markers

Since the Beluga sturgeon population has been considered to be locally extinct in Italy for at least thirty years (IUCN: Sturgeon Specialist Group. 1996; Bianco, 2014), the reintroduction of pure Beluga in the Adriatic would benefit from the genetic comparison of extant populations with museum samples of individuals captured in the Adriatic Sea aimed at identifying the best suitable one as source of individuals to be employed. All criteria for the preliminary investigations required for the reintroduction must be in line with the international guidelines suggested by FAO and WSCS (World Sturgeon Conservation Society) for sturgeon conservation, (Chebanov et al., 2011) and by the IUCN for reintroduction (IUCN/SSC, 2013). To this end, ancient DNA was collected from Beluga specimens available in natural history museums. mtDNA is the only region reliably amplifiable from ancient DNA, being present in thousands copies per cell and, due its high mutation rate, the D-loop can be successfully employed to discriminate between populations of the same species (Doukakis et al., 2005).

For what concerns the extant populations (Danube River, Sea of Azov, Black Sea, Caspian Sea and Ural River), the results from this study show that, on the basis of sampled individuals, the current asset of genetic diversity does not resemble the geographic distribution of populations in the different basins, as indicated my mtDNA analyses. However, despite the lack of haplotypes clusters that characterize a specific geographic region, the shared haplotypes between basins are very scarce, as revealed also by the distribution of haplotypes in the different sampling locations. Therefore, the different Beluga populations have been probably subject to genetic admixture due to migratory events of individuals hypothetically moving across basins.

This result is in agreement with previous studies on Beluga (Doukakis et al., 2004; Dudu et al., 2014) in which this pattern could be ascribed to the relatively recent connection between the Caspian and the Black Sea basins (15 – 11 thousand of years ago) (Osinov, 1984; Zubakov, 1988; Svitoch, 2008; Thom, 2010) that possibly admixed the different Beluga populations. Indeed, after Pleistocene post-glacial flooding, the raise in water levels in these basins led to a connection between them (by the Manych-Kerch Spillway) (Figure 27). Interestingly, the same mixed pattern was observed also in other sturgeon species (Brown et al., 1993; Guenette et al., 1993).

Figure 27- Connection of Ponto-Caspian basin in the epoch of the Late Glacial inundation (~17 to 10 ka BP) by the Manych-Kerch Spillway.

Alternatively, also artificial translocations of animals between different basins due to careless restocking programs (Tsvetnenko, 1993; Secor et al., 2000; Chebanov et al., 2002) have been proposed as a possible cause of genetic admixture (Dudu et al., 2014). Examples of these activities leading to genetic admixture between different populations in other species are widely documented (Rhymer and Simberloff, 1996; Ludwig, 2006).

However, the very high haplotype diversity observed, even though in accordance with previous studies with lower sample sizes (Doukakis et al., 2004; Dudu et al., 2014), could be due on one hand to a low resolving power of the marker, or on the other to a still low sample size, which is not sufficient yet to represent the genetic variability of populations as explained before. In order to verify this hypothesis, sequences available from online databases that were used in previous studies were added to the analysis, thus increasing the sample size.

Apostolidys and colleagues (1997) have shown that the degree of genetic variability in *Salmo trutta* might be very sensitive to potential stochastic effect caused by sample size variations that cannot be ruled out when using mtDNA as a molecular marker. To this end, the addition of mtDNA sequences from previous studies available in GenBank has led to an increase in the sample size of this study, in order to obtain a more accurate resolution of genetic variability in Beluga, as shown by the difference of the results prior to the addition of these sequences to the dataset.

The new dataset showed a completely different distribution of genetic variability with a geographic-related pattern. Indeed, with the exception of Sea of Azov population, in which

massive releases of individuals from the Caspian Sea are documented (Chebanov et al., 2002), resulting in a close similarity to the Caspian Sea basin, the other populations are characterized by a genetic structure concordant with their geographic distribution, based on the pairwise comparisons of ϕ_{st}. However, the ML tree and the haplotype network confirmed previous results with the reduced dataset: it was not possible to identify the presence of exclusive clusters of haplotypes in different basins. The presence of numerous loops in the network is noteworthy: this result is in accordance with a previous study on Beluga and other sturgeon species (*A. stellatus* and *A. gueldenstaedtii*), in which similar patterns of network and phylogenetic tree were identified and described (Dudu et al., 2014).

Although the final sample size is adequate to highlight the presence of a significant genetic differentiation, the high degree of diversity identified between haplotypes may suggest that the level of diversity is not sufficient yet to give a proper mitochondrial representation; this could explain the poor clarity of the network. The complex structure of the network and the present pattern of mitochondrial variability observed in non-extinct populations seems to be the result of genetic drift acting independently on recently separated populations that were probably remixed by post-glacial flooding.

By increasing the sample size, in accordance with the study of Apostolidis and colleagues, the results obtained with mtDNA are consistent with those coming from the other molecular markers (SNPs and microsatellites, discussed below). These results are in contrast with the study of Doukakis and colleagues (2005) and Dudu and colleagues (2014) with lower sample size, which would ultimately lead to a low detection power, such as in other sturgeon species (*A. baerii* and *A. stellatus*, Doukakis et al., 1999). Previous studies based on DNA polymorphisms have demonstrated that events of postglacial colonization have considerably modified and influenced the distribution of genetic variability in numerous species (Avise, 2000; Koskinen et al. 2002). Different freshwater fish show a distribution of genetic variability and a degree of genetic differentiation between sub-populations that are explainable with their geographic distribution (*Thymallus thymallus*, *Barbus barbus*, *Leuciscus cephalus* and *Salmo trutta*; Kotlik and Berrebi, 2001; Durand et al., 1999; Bernatchez, 2001; Froufe et al., 2005; Marić et al., 2014).

In this study, a unique and undebatable source population has not emerged. Nonetheless, the only two specimens of *H. huso* of the extinct Adriatic population (M1 and M3) showed two different haplotypes with respect to all the fresh samples of extant populations from basins in which Beluga is currently and historically present (Black Sea basin, which includes Danube River and Black Sea populations; and Caspian Basin, which includes Ural River and Caspian Sea populations). However, these two haplotypes differ only for two mutations that can be considered as highly reliable since supported by a mean of 4 sequences independently obtained and, interestingly, in

the haplotype network the two museum specimens resulted in two closely-related haplotypes diversified from all the other extant samples. Even though it is not enough to support the presence of an isolated Adriatic population, the high similarity observed between the two Adriatic haplotypes is noteworthy, also considering that the sequences obtained by the fourth museum sample (M4) were identical to M3 sample haplotype, but different from any other sample tested.
Other fish species have shown the same pattern of diversification, for example Apostolidis and colleagues (1997) have used mtDNA to demonstrate the presence of Mediterranean populations of *S. trutta* that have been isolated from the others due to an event of basins separation 2.5-6.0 Myr ago (Messinian or early Pliocene ages).
However, for the Adriatic population, it is necessary to increase the number of specimens for further analysis in order to confirm the data obtained in this study. Unfortunately, being ancient specimens conserved only in Natural History Collections (NHCs), the chance of finding an adequate number of samples is low, since specimens from NHC are often incomplete, inaccurate or missing (Graham et al., 2004; Walander et al., 2007).
For what concerns the preliminary analyses necessary to identify the most suitable source population for future reintroductions in the Adriatic Sea basin, this study suggests that, even though it is not enough to definitively support the presence of an isolated Adriatic population, the high similarity observed between the two Adriatic haplotypes is noteworthy. However, the cluster of haplotypes including the two longest Adriatic sequences, despite being composed by sequences from the Black Sea-Danube and Azov basins, also includes haplotypes detected only in the Caspian Sea and Ural River, thus confirming that the genetic proximity between haplotypes of this portion of the network does not correspond to a shared geographical origin.
Since the results from Beluga genetic structure depicts a species without a strong geographically based haplotype diversification, the identification of the best suitable population for the choice of individuals to be reintroduced is not possible based on mitochondrial information.
However, given the presence of a certain degree of genetic structure between basins, it is necessary to avoid the onset of outbreeding depression in the upcoming reintroduced population (Pikitch et al. 2005; Ludwig, 2006). Populations of the Black Sea basin are clearly distinct from the Caspian Sea basin (from mitochondrial, but also, nuclear point of view, as explained later on Chapter 4.3) and this is something that should be taken into account for the selection of a source population.

4.3. Genetic characterization through nuclear markers of the extant Beluga populations to lay the bases for future restocking program

4.3.1. Genetic characterization of all individuals using microsatellite and SNPs markers and phylogeographical analyses

The information obtained from the analyses of D-loop in all extant populations were confirmed by the application of nuclear information using microsatellite and SNPs markers, with the aim of inferring a more detailed picture of the geographical pattern of genetic variability.

The application of microsatellite loci and SNPs showed an improved resolving power than mitochondrial data that allowed for a straightforward identification of a phylogeographic structure between different populations of Beluga. This structure is in accordance with the one previously described with the use of the extended dataset including GenBank sequences, and resembles the geographic separation of basins: Black Sea Basin (together with the Danube River population) and Caspian Sea basin (together with the Ural River population).

The statistical power of the microsatellite panel has been supported by the analyses performed with POWSIM software, which showed that the number of loci, the number of alleles per locus, their frequency distributions and sample sizes of the different populations had a sufficient power to detect differentiation. This panel has led to the identification of significant differences at a genetic level across different populations, which is further supported by the results obtained from SNPs analyses: data from nuclear markers clearly show the presence of a genetic structure correlated with geographic differentiation between basins. This correlation between genetic diversity and differentiation across basins has been shown also in another sturgeon species with a similar distribution pattern: two studies demonstrated that populations of *A. gueldenstaedtii* from the Sea of Azov and the Caspian Sea are genetically differentiated at both nuclear level, with SNPs, RAPD and microsatellites, and at mitochondrial level (Timoshkina et al., 2009; Rastorguev et al., 2013). In particular, Timoshkina and colleagues (2009) have shown that, based on microsatellites, the highest values of genetic distances were observed between the Caspian and the Azov populations, while the Black Sea population occupied an intermediate position. In addition, Rastorguev and colleagues (2013) have used SNPs to support the differentiation between populations of *A. gueldenstaedtii* from the Sea of Azov and from the Caspian Sea. Interestingly, the differentiation pattern in *A. gueldenstaedtii* did not emerge in a recent study with much lower sample size (28 individuals, Dudu et al. 2014) than previous studies (approximately 500). In the paper from Dudu et al. (2014), among the three analysed Ponto-Caspian sturgeon

species with no significant differentiation, *A. guendenstaedtii* showed the highest genetic structuring, although not statistically significant, thus suggesting a trend towards genetic differentiation that could have been possibly hidden by the low sample size. In another study from Doukakis and colleagues (2005), a genetic structure between populations of *A. guendenstaedtii* was suggested, although the authors concluded that they could not support this hypothesis due to the low sample size (N=33).

The global differentiation pattern emerged in this study was well supported by the different analyses performed with the microsatellite panel, taking into account not only the genetic information (F_{st}, AMOVA, SAMOVA, DAPC, PCoA, STRUCTURE), but also a priori geographic information by performing analysis with GENELAND software. Indeed, the sampling sites formed clusters depending on the geographically-related basins, therefore two different clusters have emerged, including respectively Black Sea sites on one side and Caspian Sea sites on the other.

Results from nuclear markers were also supported by SNPs analyses: the pattern of geographic differentiation was confirmed also with RAD-seq analysis on 298 594 037 trimmed reads that identified the presence of genetic structure between the Black Sea and Caspian Sea basins. However, despite the low F_{st} value obtained from pairwise comparisons of different populations, the degree of genetic distance was statistically significant. The sample size was comparable with other studies on marine species using RAD-seq with the same protocol (Pecoraro et al., 2016; Paterno et al., 2017), thus supporting the reliability of the results from this study.

A peculiar scenario was observed for the population of the Sea of Azov, which showed discordant signs of differentiation from other basins depending on the markers used, on the contrary a uniform pattern was observed for the Black Sea and Caspian Sea basins. According to GENELAND analyses performed on microsatellite loci, the two Azov Sea samples were split in two different clusters: one including all Caspian samples and the other including the northern Azov sample and one population of the Black Sea. Differently, analysis with F_{st} on microsatellites showed that Azov population was genetically isolated from the Caspian basin, but not from the Black Sea after correction for multiple comparisons.

The genetic assemblage performed with the model-based clustering method implemented in the software STRUCTURE on the microsatellites panel has shown the lack of any clear differentiation in clusters of the different populations and the relative basins, but the most supported separation was in two main groups in which individuals were unambiguously allocated, with only a minor fraction of them showing a mixed contribution of two populations. This is especially evident for the Sea of Azov population, showing a clear-cut allocation to either one group or the other.

The analysis on SNPs performed with F_{st} confirms the separation of the Azov population from Caspian basin, but not from Black basin, however by using AMOVA and DAPC analysis on SNPs, a

complete genetic isolation of the Azov population from all the others, including the Black Sea population, was highlighted.

Mitochondrial analysis performed with ϕ_{st} and AMOVA revealed instead a separation of the Sea of Azov population from the Black sea basin, but not from the Caspian basin.

For what concerns this peculiar population, it can be assumed that the well-documented former releases of Beluga specimens from the Caspian Sea into the Sea of Azov (Tsvetnenko, 1993; Chebanov et al., 2002) could have somehow compromised its genetic identity and led to an unresolved pattern of variability distribution. It could be argued that translocations are a possible explanation of the admixture observed (particularly in the analysis with STRUCTURE on microsatellites). Given the long-life cycle of these animals, it is likely that many of these translocated animals are still alive and this might explain the presence of several animals in which the contribution of one single gene pool is detectable.

Moreover, the survival of the extant Sea of Azov population completely relies on continuous reintroductions of juveniles from breeding plants (Birstein, 1993; Vecsei et al., 2002; Gesner et al., 2010). Therefore, the ambiguous structure of this population could have been due to inappropriate release actions in recent times, which have compromised the original genetic structure of this population, similarly to what happened in *A. stellatus*, for which this phenomenon has been documented in the same basin (Chebanov et al., 2002). Moreover, given the connection between the Sea of Azov and the Black Sea basins through the Kerch strait (Dudu et al. 2004), migratory fluxes between basins cannot be excluded.

For what concerns the other basins, various supportive reintroductions of Beluga sturgeon were reported in the past years in the relative countries of the Ponto-Caspian region, in order to deal with the continuous decrease of the population due to anthropic activities (Pourali et al., 2003; Khodorevskaya et al. 2009; Suciu, pers. comm.). Apparently, these restocking actions did not have an impact on the genetic structure of populations of the Caspian and Black Sea basins. Indeed, the different markers tested in this study were concordant in detecting a genetic structure between basins. This result further supports the need of taking into account the geographic origin of animals employed for release actions with conservation purposes (Chebanov et al., 2011; Ludwig, 2006).

In conclusion, the analysis with the 16 microsatellite loci panel and with 4 736 polymorphic loci has highlighted the presence of a defined genetic structure among geographically separated populations.

4.3.2. Application of the microsatellite panel to relatedness analyses

YOY Danube individuals:

For what concerns the juveniles of the Danube population, the presence of numerous related specimens has emerged. This result was not unexpected given the gregarious behaviour of these animals during the early phases of their development (Suciu R., personal communication), nonetheless, in order to confirm this hypothesis, a relatedness analysis on these samples was performed.

Besides the number of individuals with different degrees of relatedness (full-sibs and half-sibs and other relatedness degrees), structure analysis has highlighted the presence of a restricted sample group with a high degree of relatedness belonging to the same familiar group ("highly related individuals") as revealed by the subsequent analyses (Colony and distribution of pairwise genetic distances), with at least two distinct haplotypes and therefore two different dams. On these bases, it was evident the need for phylogeographic analyses in order to avoid the sampling of juveniles from the same spot.

Moreover, two main hypotheses on the possible origin of these "highly related" YOY animals can be advanced and need to be deeper investigated: 1) these animals are the results of recent restocking actions of fingerlings produced from captive breeders. The low number of breeders usually used for artificial reproduction and their frequently high degree of relatedness could justify the genetic homogeneity of the "highly related" group. 2) According to the second hypothesis, these individuals (all classified as YOY of 2010) are juveniles originating from parents of a distinct ecotype of Beluga that, prior to the construction of the Iron Gates Dam (about 800 Km upstream the Danube Delta), were observed to perform reproductive migrations in fall, in order to spend the whole winter in the deeper waters of the river close to the upstream spawning sites (about 2000 Km upstream the delta) (Bemis and Kinard, 1997). These long migrations are nowadays prevented by the presence of the Danube Iron Gates dam I and II (built in the 1970-1984), that completely blocks the access of sturgeons to the middle and any reproduction activity is confined in the lower Danube (Vecsei et al., 2002).

According to information provided by Danube Delta National Institute, during the years 2010, 2011 and 2012 no restocking activity took place in the Danube River downstream of the Iron Gate Dam. Consequently, if the classification of the "highly related" animals as young of the year 2010 is correct, this would allow to exclude the first hypothesis of reintroduction.

However, according to the Sturgeon hatchery practices and management for release–Guidelines (Chebanov et al., 2011), the length of 40-50 days old Beluga is 7-11cm (1.6 – 5.2 g), therefore it seems that the size of the "highly related" Beluga (mean length: 16.58 cm) is too large for YOY

likely born in May and sampled in June. It is also known that young individuals can overwinter in the river (Chebanov et al., 2011). Therefore, it seems plausible that "highly related" individuals were born in 2009 from captive breeders, given that restocking activity is known to take place in 2009.

Individuals from a Livestock:

The microsatellite panel was successful for the parental allocation of a group of juveniles provided by a hatchery after a recent artificial reproduction. All juveniles were allocated to one dam (as expected) and to one single sire in spite of the two males used for milt collection. On one hand, these results confirm the suitability of the microsatellite panel for parental allocation also in the case of different fully-related potential parents and, on the other, suggest a potential key role of sperm competition during controlled reproduction in aquaculture plants (Chebanov et al., 2011). In fact, equal amounts of milt from two males were mixed before fertilization of the eggs, making unlikely that all F1 individuals descend from the same male just for chance. Therefore, the fertilization of eggs with sperm from different males could not be the most suitable strategy to achieve an offspring from different sires, given the presence of uncontrolled factors such as sperm competition and sperm quality of different males.

Specimens previously analysed with the microsatellite panel have been further characterized with 2b-RAD protocol for the identification of SNPs markers. The results have shown that the family previously allocated with microsatellites was successfully identified also with these markers, thus supporting the previous analysis. In addition, the other putative father of this family has been allocated to three different groups of wild individuals, suggesting that the male was not even related with a specific group of individuals belonging to the same basin. These results strengthen the need for accurate genetic screening plans of aquaculture specimens selected for reintroduction (Ludwig, 2006). The selected panel of microsatellite loci represents therefore a precious tool that could be efficiently used to discriminate the degree of genetic diversity and similarity with wild specimens from different basins.

4.3.3. Final consideration for future management and conclusions

This study has revealed a genetic differentiation pattern correlated to the geographic separation between basins in the extant Beluga populations, and hypothetically also in the extinct Adriatic population, despite the possible admixture in the recent post-glacial age (Doukakis et al., 2005; Dudu et al., 2014). This result is of great importance for the evaluation and management of

conservation units of the Beluga and for future reintroduction plans that should necessary consider the genetic variability pattern emerged across different basins.

Mitochondrial data did not allow the allocation of the only two museum samples to a precise source population. Therefore, it is not possible to define a best source population for the Adriatic Sea. However, nuclear marker analysis showed a clear differentiation between Caspian Sea and Black Sea; as a result, the choice of animals for future restocking cannot be random. When choosing between the two basins, it is likely that the best source is the Black Sea, thanks to its geographical proximity. It is indeed likely that, since Beluga is an anadromous species, the population once present in the Adriatic basin was interconnected with the population of the Black Sea. Such natural mixing is not occurring between the Black and Caspian populations. It is therefore essential that the animals for reintroductions in the Po River are subject to an accurate genetic analysis to identify their geographical origin.

By choosing individuals from the Black Sea for reintroduction actions, any admixture between the reintroduced and the native population of the Black Sea would be avoided in the case of bidirectional migration of individuals, given that Adriatic Sea and Black Sea are connected (Dudu et al., 2004). This is crucial to avoid genetic pollution of the original population (Fontana et al., 2001; Ludwig, 2006). Therefore, the genetic-makeup typical of the Black Sea would be preserved by avoiding any genetic damage to different populations due to genetic admixture from introgression or hybridization typical of many fish species (Rhymer and Simberloff, 1996). Moreover, given the ambiguous structure of the Sea of Azov population and the documented release actions with individuals from the Caspian basin, which is geographically separated from the Sea of Azov (Tsvetnenko, 1993; Chebanov et al., 2002), this population should not be considered for any future reintroduction plan.

In order to increase the chances of a successful reintroduction, other factors should be taken into consideration. For example, as suggested by Chebanov and colleagues (2002), it is necessary to consider the different spawning ecotypes eventually present in the selected source population in order to maintain the genetic variability necessary for the long-term survival of the population (Chebanov et al. 2011).

Therefore, further in-depth analysis on ecological aspects of this species are necessary. A further determinant factor to consider is that the opportunity to perform a release programme is limited to the animals available for the establishment of a reintroduction stock. There are various aquaculture plants on the Italian territory with stocks of animals of different origins (for example both from the Black Sea and the Caspian Sea basins). Therefore, the genetic characterization of the Italian Beluga stocks reared in aquaculture will be of primary importance to evaluate the actual availability and diversity of individuals for reintroduction plans (Chebanov et al., 2011). To

this end, the panel of microsatellite loci used in this study has been proven to be sufficiently specific to discriminate both the genetic variability between populations and the genetic relatedness within populations, even between closely-related samples. Therefore, the use of this panel in diagnostic analyses will be necessary to perform the appropriate evaluations in Italian livestocks, in order to obtain a stock of animals as highly genetically diversified as possible, but avoiding at the same time the chance of outbreeding depression in released animals.

5. RECOMMENDATION

The extinction of Beluga in the Adriatic region, combined with the entry of alien species in Italian reservoirs, has led to an imbalance in food webs and biological properties in an important area of the Italian territory (River Po basin Authority, 2008). Moreover, to rescue the depletion of native species in the Adriatic basin and to control alien species that have dramatically emerged in the last few decades, an important step can be made by reintroducing and important flag specie like the Beluga sturgeon. The aim of this study was therefore to give a detailed characterization of the genetic structure in extant Beluga populations from different basins, and to compare them with extinct Adriatic population.

Although in the Beluga the presence of a structured phylogeographic pattern has not been supported by past studies based on mitochondrial data, in the light of the results from this study, the extant populations from different basins show a genetic diversity pattern related with geographic distance. Genetic analysis on mitochondrial data showed the need for an adequate sample size in order to achieve a sufficient resolving power for phylogeographic analyses. On the contrary, the selected microsatellite loci and SNPs have shown a sufficient detection power also with lower sample sizes. The only exception was in the population from the Sea of Azov, for which, however, frequent reintroductions of individuals from other basins are reported (Chebanov et al., 2002). Despite the relatively recent post-glacial admixture across basins and despite the various past restocking actions, the different extant populations have maintained a detectable genetic structure consistent with their geographic location. This result is extremely useful for the establishment of reintroduction plans that should necessarily take into account the genetic diversity subtending extant populations.

Due to fishing restrictions and to the extreme scarcity of wild individuals, the establishment of broodstocks of wild origin is not possible (Resolution Conf. 12.7 -Rev. CoP14- CITES; Paraschiv et al. 2006). The only suitable breeders are those reared in different aquaculture plants, mainly for caviar production (Boscari and Congiu, 2014). These animals should be carefully analysed for what concerns their purity, relatedness and geographical origin, in order to identify the best suitable ones to be employed for broodstock foundation (Boscari et al., 2014 b). Any other initiative in which reintroduction actions are performed or even planned without an adequate genetic characterization should be discouraged as irresponsible actions with potential negative effects on the genetic make-up of Beluga populations.

In this context, the optimal broodstock should possibly represent the maximum amount of diversity of the Black Sea population paying attention to: a) avoiding external contaminations from other populations; b) minimizing the effect of inbreeding.

With particular regard to inbreeding, it should be noted that captive broodstock reared in Italy are mainly composed by fully related animals imported as fingerlings or fertilized eggs as future source of caviar. For this reason, the level of relatedness within captive stock is often very high and the random mating of animals coming from the same plant would likely involve full sibs. Using these animals for reintroduction purposes would contradict every national and international guideline about the best practices to adopt in restocking and reintroduction programs (The Ramsar Declaration on Global Sturgeon Conservation; Chebanov et al., 2011; Linee guida per l'immissione di specie faunistiche).

6. REFERENCES

Adriatic sturgeon (*Acipenser naccarii*). Conservation Genetics. 9. 461–463.

Arnason E. and Rand D.M. 1992. Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. Genetics. 132, 211-220.

Artyukhin E.N. 1995. On biogeography and relationships within the genus Acipenser. The Sturgeon quarterly. 3, 6–8.

Artyukhin E.N. 1997. The current status of commercial sturgeon species in the Black Sea-Caspian Sea basin. Pp. 9-13 in Birstein V.J., Bauer A., Kaiser-Pohlmann A. IUCN Species Survival Commision. Sturgeon stocks and caviar trade workshop. Bonn, Germany: Occasional Papers of the IUCN Species Survival Commission No. 17.

Autorità di bacino del Fiume Po, 2008. Monitoraggio dell'ittiofauna e redazione della Carta ittica del Fiume Po. Qualità dell'ittiofauna e del macrobenthos del fiume Po.

Avise J.C. 1994. Molecular Markers, Natural History, and Evolution. Chapman & Hall, New York, pp. 511.

AA.VV., 2007. Linee guida per l'immissione di specie faunistiche. Ministero dell'Ambiente - Istituto Nazionale Fauna Selvatica. Quaderni di Conservazione della Natura. 27.

Barmintseva A.E. and Mugue N.S. 2013. The use of microsatellite loci for identification of sturgeon species (Acipenseridae) and hybrid forms. Russian Journal of Genetics. 49, 950–961.

Belkir K., Borsa P., Goudet J., Bonhomme F. 2005. Genetix v. 4.05, logiciel sous Windows pour la génétique des populations. Laboratoire Génome et Populations. CNRS UPR 9060, Université Montpellier II.

Bemis, W. E. and Kynard, B. 1997. Sturgeon rivers: an introduction to acipenseriformes biogeography and life history. Environmental Biology of Fishes. 48, 167–183.

Berg L.S. 1948. The freshwater fishes of the USSR and adjacent countries. Vol. 1. English translation published by Israel Program for Scientific Translations, Jerusalem.

Bianco, P.G. 2014. An update on the status of native and exotic freshwater fishes of Italy. Journal of Applied Ichthyology. 30, 62–77.

Billard R., Lecointre G. 2001. Biology and conservation of sturgeon and paddlefish. Reviews in Fish Biology and Fisheries. 10, 355–392.

Birstein V.J., Doukakis P., DeSalle R. 2000. Polyphyly of mtDNA lineages in the Russian sturgeon, *Acipenser gueldenstaedtii*: Forensic and evolutionary implications. Conservation Genetics. 1, 81 – 88.

Birstein, V. J. 2002. Sturgeon species and hybrids: Can Hybrids Produce Caviar? Enviromental Policy Law. 32, 210–214.

Birstein, V.J. 1993. Sturgeons and paddlefishes: threatened fishes in need of conservation. Conservation Biology. 7, 773–787.

Birstein, V.J., Bemis, W. E., Waldman, J. R. 1997. The threatened status of acipenseriformes species: a summary. Environmental Biology of Fishes. 48, 427–435.

Blouin M.S. 2004. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends in Ecology and Evolution. 18. 503-511.

Boscari E., Barmintseva A., Pujolar J. M., Doukakis P., Mugue N.S., Congiu L. 2014 a. Species and hybrid identification of sturgeon caviar: a new molecular approach to detect illegal trade. Molecular Ecology Resources. 14, 489–498.

Boscar E. and Congiu L. 2014. The need for genetic support in restocking activities and ex situ conservation programmes: the case of the Adriatic sturgeon (*Acipenser naccarii* Bonaparte, 1836) in the Ticino River Park. Journal of applied Ichthyology. 30, 1416–1422.

Boscari E., Vidotto M., Martini D., Papetti C., Ogden R. and Congiu L. 2015. Microsatellites from the genome and the transcriptome of the tetraploid Adriatic sturgeon (*Acipenser naccarii*) and cross-species applicability to the diploid Beluga sturgeon (*Huso huso*). Journal of Applied Ichthyology. 31. 977-983.

Boscari, E., Pujolar, J. M., Dupanloup, I., Corradin, R., and Congiu, L. 2014b. Captive breeding programs on family groups in polyploidy sturgeons. PLos ONE, 9(10), e110951. doi: 10.1371/journal.pone.0110951.

Brown J.R., Beckenbach A.T., Smith M.J. 1992. Mitochondrial DNA length variation and heteroplasmy in populations of white sturgeons (*Acipenser trasmontanus*). Genetics. 132, 221-228.

Brumfield R.T., Beerli P., Nickerson D.A., Edwards S.V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology and Evolution. 18, 249-256.

Buroker N.E., Brown J.R., Gilbert T.A., O'Hara P.J., Beckenbach A.T., Thomas W.K., Smith M.J. 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics. 124, 157-163.

Carvajal-Rodriguez A. and de Uña-Alvarez J. 2011. Assessing Significance in High-Throughput Experiments by Sequential Goodness of Fit and q-Value Estimation. PLoS ONE. 6, e24700.

Ceballos G. and Ehrlich P.R. 2002. Mammal population losses and the extinction crisis. Science. 296, 904-907.

Chebanov M. S., Karnaukhov G. I., Galich E. V., Chmir Y. N. 2002. Hatchery stock enhancement and conservation of sturgeon, with an emphasis on Sea of Azov populations. Journal of Applied Ichthyology. 18. 463-469.

Chebanov M.S. and Koziritskaya Yu.E. 2007. Beluga. Red data book of Krasnodar territory (animals). Krasnodar: Center of development, pp. 315-316.

Chebanov M., Rosenthal H., Gesner J., van Anrooy R., Doukakis P., Pourkazemi M., Williot P. 2011. Sturgeon hatchery practices and management for release. Guidelines. FAO Fisheries and Aquaculture Technical Paper, 570.

Çiftci Y., Eroğlu O., Firidin S. 2013. Heteroplasmy and length variation in the tRNApro-D-loop regions of three sturgeon species (*A. stellatus*, *A. gueldenstaedtii* and *H. huso*) from the Turkish coast of the Black Sea. Turkish Journal of Biochemistry. 38, 250-257.

CITES. 2000. Sixteenth Meeting of the CITES Animals Committee Shepherdstown (United States of America) 11-15 December 2000. Implementation of Resolution Conf. 8.9 (Rev.). ACIPENSERIFORMES.

Clement M., Posada D., Crandall K.A. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology. 9, 1657–1659.

Clement M., Snell Q., Walke P., Posada D., Crandall K. 2002. TCS: estimating gene genealogies. Proceedings of the 16th International Parallel and Distributed Processing Symposium. 2, 184.

Congiu L., Dupanloup I., Patarnello T., Fontana F., Rossi R., Arlati G., Zane L. 2001. Identification of interspecific hybrids by amplified fragment length polymorphism: the case of sturgeon. Molecular Ecology. 10, 2355-2359.

Congiu L., Fontana F., Patarnello T., Rossi R., Zane L. 2002. The use of AFLP in sturgeon identification. Journal of Applied Ichthyology. 18, 286-289.

Congiu, L., Pujolar, J.M., Forlani, A., Cenadelli, S., Dupanloup, I., Barbisan, F., Galli, A., and Fontana, F. 2011. Managing polyploidy in ex situ conservation genetics: the case of the critically endangered Adriatic sturgeon (*Acipenser naccarii*). PLoS ONE, 6(3), e18249. doi: 10.1371/journal.pone.0018249.

Cooper A., 2000. Ancient DNA: Do it Right or Not at All. Science letters. 289, 1139.

Dabney J., Knapp M., Glocke I., Gansauge M.T., Weihmann A., Nickel B., Valdiosera C., García N., Pääbo S., Arsuaga J.L., Meyer M. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences of the United States of America. 110, 15758-15763.

Davey J.W., Blaxter M.L. 2010. RADSeq: next generation population genetics. Briefings in Functional Genomics. 9, 416–423.

DeAngelis M.M., Wang D.G., Hawkins T.L. 1995. Solid-Phase Reversible Immobilization for the Isolation of PCR Products. Nucleic Acids Research. 23. 4742–43.

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

Doug P.A. and Seddon P.J. 2007. Directions in reintroduction biology. Trends in Ecology and Evolution. 23, 20–25.

Doukakis P, Birstein V.J., Ruban G.I., Desalle R. 1999 Molecular genetic analysis among subspecies of two Eurasian sturgeon species *Acipenser baerii* and *A. stellatus*. Molecular Ecology. 8, S117–S127.

Doukakis P., Pikitch E., Rothschild A., DeSalle R., Amato G. and Kolokotronis S.O. 2012. Testing the effectiveness of an international conservation agreement: marketplace forensics and CITES caviar trade regulation. Plos ONE. 7, e40907.

Doukakis, P., Babcock, E.A., Pikitch, E.K., Sharov, A.R., Baimukhanov, M., Erbulekov, S., Bokova, Y., Nimatov, A. 2010. Management and recovery options for Ural River Beluga sturgeon. Conservation Biology. 24, 769–777.

Doukakis, P., Birstein, V.J., Salle, R. 2005. Intraspecific structure within three caviar-producing sturgeons (*Acipenser gueldenstaedtii*, *A. stellatus*, and *Huso huso*) based on mitochondrial DNA analysis. Journal of Applied Ichthyology. 21, 457–460.

Drauch, A.M. and Rhodes Jr., O. E. 2007. Genetic Evaluation of the Lake Sturgeon Reintroduction Program in the Mississippi and Missouri Rivers. North American Journal of Fisheries Management. 27, 434–442.

Dudu A., Georgescu S.E., Costache M. 2014. Molecular analysis of phylogeographic subspecies in three Ponto-Caspian sturgeon species. Genetics and Molecular Biology. 37, 587–597.

Dudu A., Macarie R.D., Burcea A., Georgescu S.E. and Costache M. 2015. Identification of Bester hybrid and its parental species (♀ *Huso huso* Linnaeus 1758, and ♂ *Acipenser ruthenus* Linnaeus 1758) by nuclear markers. Scientific Papers Animal Science and Biotechnologies. 48, 81-85.

Dudu, A., Georgescu, S. E., Dinischiotu, A., Costache, M. 2009. Four new polymorphisms of the mitochondrial gene cytochrome b in the Beluga sturgeon (*Huso huso*) from the Black Sea. Roumanian Biotechnological Letters. 14, 4418–4422.

Dupanloup I., Schneider S., Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. Molecular Ecology. 11: 2571-2581.

Earl, Dent A. and vonHoldt, Bridgett M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources vol. 4 (2) pp. 359-361 doi: 10.1007/s12686-011-9548-7.

Edmands S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Molecular Ecology. 16, 463-475.

Etter P.D., Bassham S., Hohenlohe P.A., Johnson E.A., Cresko W.A. 2011. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: Orgogozo V, Rockman MV, editors. Molecular Methods for Evolutionary Genetics. New York: Springer. p. 157 –178.

Excoffier L., and Lischer H.E.L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10, 564–567.

Excoffier L., Smouse P.E., Quattro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotype: application to human mitochondrial DNA restriction data. Genetics. 131, 479–491.

Fain S.R., Straughan D.J., Hamlin B.C., Hoesch R.M. and LeMay J.P. 2013. Forensic genetic identification of sturgeon caviars travelling in world trade. Conservation Genetics. 14, 855-874.

Falush D., Stephens M., Pritchard J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 164, 1567–1587.

Fischer W., Schneider M., Bauchot M. L. 1987. Fiches FAO d´Identification des Especes pour les Besoins de la Pêche. Mediterranee et Mer Noir, Zone de Pêche 37 (Révision 1), Vol. II: Vertèbres. FAO, Rome. Pp. 944-952.

Fischer, J., and Lindenmayer, D. B. 2000. An Assessment of the Published Results of Animal Relocations. Biological Conservation. 96, 1- 11.

Fontana F., Tagliavini J., Congiu L. 2001. Sturgeon genetics and cytogenetics: recent advancements and perspectives. Genetics. 111, 359–373.

Fopp-Bayat D., Jankun M., Woznicki P., Kolman R. 2007. Viability of diploid and triploid larvae of Siberian sturgeon and bester hybrids. Aquaculture Research. 38, 1301-1304.

Forlani A., Fontana F., Congiu L. 2008. Isolation of microsatellite loci from the endemic and endangered Adriatic sturgeon (*Acipenser naccarii*). Conservation Genetics. 9, 461–463.

Frankham R., Ballou J.D., Briscoe D.A. 2010. Introduction to Conservation Genetics (second ed.) Cambridge University Press, Cambridge, U.K.

Freeland J.R., Kirk H., Petersen S. 2011. Molecular marker in ecology. In: Molecular Ecology, second edition, 35–75. John Wiley and Sons, Ltd.

Ghadirnejad H., Zhu B., Chang J. 2008. Study on the Population Genetics of the Great Sturgeon (*Huso huso*) in the Southern Part of the Caspian Sea Using Microsatellite Loci. Journal Of Applied Biological Sciences. 2. 095-099.

Gesner J., Chebanov M., Freyhof, J. 2010. *Huso huso*. In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2.

Goudet J. 1995. FSTAT (Version 1.2): A computer program to calculate F-Statisitics. The Journal of heredity. 86, 485-486.

Graham C.H., Ferrierc S., Huettmand F., Moritzb C., Peterson A.T. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology and Evolution. 19, 497–503.

Griffith B., Scott J. M., Carpenter J. W., Reed C. 1989. Translocation as a Species Conservation Tool: Status and Strategy. Science. 245, 477 -480.

Guichoux E., Lagache L., Wagner S., Chaumeil P., Léger P., Lepais O., Lepoittevin C., Malausa T., Revardel E., Salin F., Petit R.J. 2011 Current trends in microsatellite genotyping. Molecular Ecology Resources. 11, 591-611.

Guillot G., Mortier F., Estoup A. 2005. GENELAND: a computer package for landscape genetics. Molecular Ecology Notes. 5, 712-715.

Gusset, M. 2009. A Framework for Evaluating Reintroduction Success in Carnivores: Lessons from African Wild Dogs. In M.W. Hayward and M. Somers (Eds.). Reintroduction of Top- Order Predators. New York, NY: Wiley -Blackwell.

Hansen A., Willerslev E., Wiuf C., Mourier T., Arctander P. 2001. Statistical evidence for miscoding lesions in ancient DNA templates. Molecular Biology and Evolution. 18, 262–65.

Hayes B., He J., Moen T., Bennewitz J. 2006. Use of molecular markers to maximize diversity of founder populations for aquaculture breeding programs. *Aquaculture* 255, 573-578.

Havelka M., Hulák M., Bailie D.A., Prodöhl P.A., Flajšhans M. 2013. Extensive genome duplications in sturgeons: new evidence from microsatellite data. Journal of Applied Ichthyology. 29, 704–708.

Havelka M., Fujimoto T., Hagihara S., Adachi S., Arai K. 2017. Nuclear DNA markers for identification of Beluga and Sterlet sturgeons and their interspecific Bester hybrid. Scientific Reports. 7. 1694

Henderson-Arzapalo A. and King T.L. 2002. Novel microsatellite markers for Atlantic sturgeon (*Acipenser oxyrinchus*) population delineation and broodstock management. Molecular Ecology Notes. 2. 437–439.

Hensel K., Holcik J. 1997. Past and current status of the sturgeons in the upper and middle Danube River. Environmental Biology of Fishes. 48. 185-200.

Hofreiter M., Jaenicke V., Serre D., Haeseler A.V., Pääbo S. 2001 (b). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Research. 29, 4793–99.

Hofreiter M., Serre D., Poinar H.N., Kuch M., Pääbo S. 2001 (a). Ancient DNA. Nature Reviews Genetics. 2, 353-359.

Holčík, J. 2000. Major problems concerning the conservation and recovery of the Atlantic sturgeon *Acipenser sturio* L., 1758. Boletin - Instituto Espanol de Oceanografia. 16, 139–148.

Hood, G. M. 2010. PopTools version 3.2.5.
Hosseini S.V., Hosseini S.M., Monsef Rad S.F., Mobinifar M., Regenstein J.M. 2013. Heavy metal bioaccumulation and risk assessment for wild and farmed Beluga sturgeon caviar. Environmental Monitoring and Assessment. 185, 9995–9999.

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

IUCN 2016. The IUCN Red List of Threatened Species. Version 2016-1. <http://www.iucnredlist.org>.

IUCN/SSC (2013). Guidelines for Reintroductions and Other Conservation Translocations. Version 1.0.

Jombart T. 2008. Adegenet: A R Package for the Multivariate Analysis of Genetic Markers. Bioinformatics. 24, 1403–1405.

Jones O. and Wang J. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources. 10, 551–555.

Kalinowski, S. T. 2005. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes. 5, 187–189.

Kalinowski S.T., Wagner A.P., Taper M.L. 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes. 6. 576-579.

Khodorevskaya, R., Dovgopol G., Zhuravleva O., Vlasenko A. 1997. Present status of commercial stocks of sturgeon in the Caspian Sea basin. Environmental Biology of Fishes. 48. 209-219.

Khodorevskaya R.P., Krasikov E.V., Dovgopol G.F., Zhuravleva O.L. 2000. Formation of the stock of Caspian Acipenserids under present-day conditions. Journal of Ichthyology. 40, 602–609.

Khodorevskaya R.P., Ruban G.I., Pavlov D.S. 2009. Behaviour, Migrations, Distribution, and Stocks of Sturgeons in the Volga-Caspian Basin. Norderstedt, Germany: Books on Demands GmbH.

Krieger J. and Fuerst P.A. 2002. Evidence for a Slowed Rate of Molecular Evolution in the Order Acipenseriformes. Molecular Biology and Evolution. 19, 891-897.

Krieger J., Hett A.K., Fuerst P.A., Artyukhin E., Ludwig A. 2008. The molecular phylogeny of the order Acipenseriformes revisited. Journal of Applied Ichthyology. 24, 36–45.

Leonard J.A. 2008. Ancient DNA applications for wildlife conservation. Molecular Ecology. 17, 4186-4196.

Leus K. 2011. Captive breeding and conservation. Zoology in the Middle East. 54, 151-158.

Levin, A. 1997. The distribution and migration of sturgeons in the Caspian Sea. Pp. 13-20 in Birstein V., Bauer A., Kaiser-Pohlmann A., eds. Sturgeon Stocks and Caviar Trade Workshop, Proceedings of a Workshop Held 9-10 October in Bonn, Germany. Bonn, Germany: IUCN Occasional Paper No 17.

Li Y., Korol A.B., Fahima T., Beiles A., Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology. 11, 2453-2465

Lindahl T. 1993. Instability and decay of the primary structure of DNA. Nature. 362, 709–715

Ludwig A. 2006. A sturgeon view on conservation genetics. European Journal of Wildlife Research. 52, 3-8.

Ludwig A. and Jenneckens I. 2000. A PCR test for heteroplasmy in sturgeon. Animal Genetics. 31, 153-154.

Ludwig A., Lippold S., Debus L. and Reinartz R. 2009. First evidence of hybridization between endangered sterlets (*Acipenser ruthenus*) and exotic Siberian sturgeon (*Acipenser baerii*) in the Danube River. Biological Invasions. 11, 753-760.

Ludwig, A. 2008. Identification of Acipenseriformes species in trade. Journal of Applied Ichthyology. 24, 2–19.

Machida R.J., Miya M.U., Nishida N., Nishida S. (2004). Large-scale gene rearrangements in the mitochondrial genomes of two calanoid copepods *Eucalanus bungii* and *Neocalanus cristatus* (Crustacea) with notes on new versatile primers for the srRNA and COI genes. Gene. 332, 71–78.

Mardis E.R. 2008. The impact of next-generation sequencing technology on genetics. Trends in Genetics. 24, 133-141.

Maroto, J.M. , Moran, M. 2014. Detecting the presence of depensation in collapsed fisheries: The case of the Northern cod stock. Ecological Economics. 97, 101–109.

Mastretta-Yanes A., Arrigo N., Alvarez N., Jorgensen T.H., Piñero D., Emerson B.C. 2014. Restriction Site-Associated DNA Sequencing, Genotyping Error Estimation and de Novo Assembly Optimization for Population Genetic Inference. Molecular Ecology Resources,.15. 28-41.

McGhee J.D., Von Hippel P.H. 1975 (a). Formaldehyde as a probe of DNA structure. I. Reaction with endocyclic imino groups of DNA bases. Biochemistry. 14 1281–1296

McGhee J.D., Von Hippel P.H. 1975 (b). Formaldehyde as a probe of DNA structure. II. Reaction with endocyclic imino groups of DNA bases. Biochemistry. 14, 1297–1303.

McNeely, J.A., Miller, K.R., Reid, W.V., Mittermeier, R.A., Werner, T.B., 1990. Conserving the World's Biological Diversity. IUCN, World Resources Institute, Conservation International, WWF-US and the World Bank, Washington, DC.

McQuown E.C., Sloss B.L., Sheehan R.J., Rodzen J., Tranah G. and May B. 2000. Microsatellite analysis of genetic variation in sturgeon: new primer sequences for Scaphirhyncus and Acipenser. Canadian Journal of Fisheries and Aquatic Sciences 54. 1542-1547.

Mortiz C. 1994. Defining "evolutionary significant units" for conservation. Trends in ecology and Evolution. 9, 373-375.

Mugue N.S., Barmintseva A.E., Rastorguev S.M., Mugue V.N. and Barmintsev V.A. 2008. Polymorphism of the mitochondrial DNA control region in eight sturgeon species and development of a system for DNA-based species identification. Russian Journal of Genetics. 44, 793–798.

Nei M. and Feldman M.W. 1972. Identity of genes by descent within and between populations under mutation and migration pressures. Theoretical Population Biology. 3, 460-465.

Nielsen E.E.G., L.A. Bach, Kotlicki P. 2006. HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. Molecular Ecology Resources. 6. 971-976

Pääbo S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proceedings of the National Academy of Sciences. 86, 1939–1943.

Pääbo S., Poinar H., Serre D., Jaenicke-Despr´es V., Hebler J., Rohland N., Kuch M., Krause J., Vigilant L., Hofreiter M. 2004. Genetic analyses from ancient DNA. Annual Review of Genetics. 38, 645–79.

Paireder S., Werner B., Bailer J., Werther W., Schmid E., Patzak B., Cichna-Markl M. 2013. Comparison of protocols for DNA extraction from long-term preserved formalin fixed tissues. Analytical Biochemistry. 439. 152-160.

Paraschiv M., Suciu R., Suciu M. 2006. Present state of sturgeon stocks in the lower Danube River, Romania. Proceedings 36th International Conference of IAD. Austrian Committee Danube Research / IAD, Vienna, 152–158.

Paterno M., Schiavina M., Aglieri G., Ben Souissi J., Boscari E., Casagrandi R., Chassanite A., Chaintore M., Congiu L., Guarnieri G., Kruschel C., Maic V., Marino I. A. M., Papetti C., Patarnello T., Zane L, Melià P. 2017. Population genomics meet Lagrangian simulations: Oceanographic patterns and long larval duration ensure connectivity among *Paracentrotus lividus* populations in the Adriatic and Ionian seas. Ecology and Evolution. 1-17.

Paxinos E.E., James H.F., Olson S.L., Ballou J.D., Leonard J.A., Fleischer R.C. 2002. Prehistoric decline of genetic diversity in the nene. Science. 296, 1827.

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

Peakall R. and Smouse P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 28, 2537-2539.

Pecoraro C., Babbucci M., Villamor A., Franch R., Papetti C., Leroy B., Ortega-Garcia S., Muir J., Rooker J., Arocha F., Murua H., Zudaire I., Chassot E., Bodin N., Tinti F., Bargelloni L., Cariani A. 2016. Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (*Thunnus albacares*). Marine Genomics. 25. 43-18.

Peng Z., Ludwig A., Wang D., Diogo R., Wei Q., He S. 2007. Age and biogeography of major clades in sturgeons and paddlefishes (Pisces: Acipenseriformes). Molecular Phylogenetics and Evolution. 42, 854–862.

Pertoldi C., Bijlsma R., Loeschcke V. 2007. Conservation genetics in a globally changing environments: present problems, paradoxes and future challenges. Biodiversity and conservation. 16, 4147-4163.

Pikitch E.K., Doukakis P., Lauck L., Chakrabarty P., Erickson D.L. 2005. Status, trends and management of sturgeon and paddlefish fisheries. Fish and Fisheries. 6, 233–265.

Pirogovskii M.I., Sokolov L.I., Vasil'ev V.P. 1989. *Huso huso* (Linnaeus, 1758). In: The Freshwater Fishes of Europe. Part. II. General Introduction of Fishes, Acipenseriformes. Holcik (Ed.), AULA–Verlag, Wiesbaden. 1, 295–344.

Porcellotti, S. 2001. Web Museum della Fauna Europea. <www.ittiofauna.org>

Pourali H.R., Pourkazemi M., Arshad U., Mohseni M. 2003. Brackish water culture of sturgeons in the coastal regions of south Caspian Sea. Technical Report submitted to the Caspian Environment Programme for the Matching Small Grant Program in the Caspian Region, Round I. International Sturgeon Research Institute Publications.

Pourkazemi M. 2008. The comprehensive study on population genetic structure of sturgeon from the Caspian Sea. Final research report. Iranian Fishery Research Organization, International Sturgeon Research Institute. Pp. 280.

Primrose S.B. and Twyman R.M. 2003. Principles of genome analysis and genomics, 3rd edn. Malden MA, Oxford, Blackwell Pub.

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

Raymond M. and Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity. 86. 248-249.

Raspopov V. 1993. Growth rate of Caspian Sea Beluga. Journal of Ichthyology. 33. 72-84.

Raspopov V., Novikova A. 1997. Size and age composition of larvae and spawners of the great sturgeon, *Huso huso*, migrating in the Volga. Journal of Ichthyology. 37. 166-173.

Rastorguev S. M., Nedoluzhko A V., Mazur A M., Gruzdeva N. M., Volkov A. A., Barminsteva A. E., Mugue N. S., Prokhortchouk E. B. 2013. High-throughput SNP-genotyping analysis of the relationship among Ponto-Caspian sturgeon species. Ecology and Evolution. 3. 2612-2618.

Rhymer J.M. and Simberloff D. 1996. Extinction by hybridization and introgression. Annual Review of Ecology and Systematics. 27, 83-109.

Rosenthal H., Bronzi P., McKenzie D. 1999. Highlights of the 3rd International Symposium on Sturgeon: Conclusions and Recommendations. Journal of Applied Ichthyology. 15, 1–6.

Rosenthal H., Pourkazemi M., Bronzi P., Brunch R. M. 2006. The hth International Symposium on sturgeons: a conference with major emphasis on conservation, environmental mitigation and sustainable use of the sturgeon resources – The Ramsar Declaration on Global Sturgeon Conservation. Journal of Applied Ichthyology. 22, 5–12.

Rossi R., Grandi G., Trisolini P., Franzoi P., Carrieri A., Dezfuli B.S., Vecchietti E. 1991. Osservazioni sulla biologia e la pesca dello storione cobice (*Acipenser naccarii* Bonaparte) nella parte terminale del fiume Po. Atti della Società Italiana di Scienze Naturali e del Museo Civico di Storia Naturale di Milano. 132, 121–142.

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

Rousset F. and Leblois R. 2007. Likelihood and Approximate Likelihood Analyses of Genetic Structure in a Linear Habitat: Performance and Robustness to Model Mis-Specification. Molecular Biology and Evolution. 24. 2730–2745.

Roy M.S., Girman D.J., Taylor A.C. and Wayne R.K. 1994. The use of museum specimens to reconstruct the genetic variability and relationship of extinct populations. Experientia 50, 551-557.

Rozas J. and Rozas R. 1995. DNASP, DNA sequence polymorphism: an interactive program for estimating Population Genetics parameters from DNA sequence data. Computer Applications in the Biosciences. 11, 621–625.

Ryman N. and Palm S. 2006. POWSIM: A Computer Program for Assessing Statistical Power When Testing for Genetic Differentiation. Molecular Ecology Notes. 6. 600–602.

Schneider C.A., Rasband W.S., Eliceiri K.W. 2012. NIH Image to ImageJ: 25 Years of Image Analysis. Nature Methods. 9. 671–75.

Schram S.T., Lindgren J., Lori M. Evrar L.M. 1999. Reintroduction of Lake Sturgeon in the St. Louis River, Western Lake Superior. North American Journal of Fisheries Management. 19, 815–823.

Secor D., Arefiev A., Nikolaev A., Sharov A. 2000. Restoration of sturgeons: lessons from the Caspian Sea sturgeon ranching programme. Fish and Fisheries. 215-230.

Secor D. and Waldman J. 1999. Historic abundance of Delaware Bay Atlantic sturgeon and potential rate of recovery. American Fisheries Society Symposium. 23. 203-216.

Seddon P.J., Doug P.A., Maloney R.F. 2007. Developing the Science of Reintroduction Biology. Conservation Biology. 21, 303-12.

Shaffer M. L. 1981. Minimum Population Sizes for Species Conservation. BioScience. 31, 131-134.

Shapiro R. 1981. Damage to DNA caused by hydrolysis. In Chromosome Damage and Repair, ed. E Seeberg, K Kleppe, pp. 3–12. New York: Plenum.

Stuart, B.L. Dugan K.A., Allard M.W., Keaney M. 2006. Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. Systematics and Biodiversity. 4, 133–136.

Stuart, S.N. 1991. Re-introductions: to what extent are they needed? In Beyond Captive Breeding: Re-introducing Endangered Mammals to the Wild. (ed J.H.W. Gipps), pp.27-37. Symposium of the Zoological Society of London, Oxford: Clarendon Press.

Sturgeon Specialist Group. 1996. *Huso huso* (Adriatic Sea stock). The IUCN Red List of Threatened Species 1996: e.T10273A3188106.

Svärdson G. 1970. Significance of introgression in coregonid evolution. In Biology of Coregonid Fishes,ed. CC Lindsey,CS Woods, pp. 33–59. Winnipeg: Univ.Manitoba Press.

Tang, E.P.Y. 2006. Path to Effective Recovering of DNA from Formalin-Fixed Biological Samples in Natural History Collections: Workshop Summary. The National Academies Press, Washington, D.C.

Templeton A.R. 1986. Coadaptation and outbreeding depression. In Conservation Biology: The Science of Scarcity and Diversity, ed. ME Soulé, pp. 105–16. Sunderland, MA: Sinauer.

Timoshkina N. N., Barminsteva A. E., Usatov A. V., Mugue N. S. 2009. Intraspecific genetic polymorphism of russian sturgeon *Acipenser gueldenstaedtii*. Animal Genetics. 45. 1250-1259.

TRAFFIC. 2000. Review of 10 species of Acipenseriformes, prepared in the Sixteenth Meeting of the CITES Animals Committee.

Van Dyke F. 2008. Conservation Biology: foundations, concepts, applications. Wheaton College, Illinois, USA. Second Editions, Springer Science and Buisiness Media.

Van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes. 4: 535 – 538.

Vecsei P., Suciu R., Peterson D. 2002. Threatened fishes of the world: *Huso huso* (Linnaeus, 1758) (Acipenseridae). Environmental Biology of Fishes. 65, 363–365.

Vidotto M., Grapputo A., Boscari E., Barbisan F., Coppe A., Grandi G., Kumar A. and Congiu L. 2013. Trascriptome sequencing and de novo annotation of the critically endangered Adriatic sturgeon. BMC Genomics. 14. 407.

Vignal A., Milan D., SanCristobal M., Eggen A. 2002. A review of SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution. 34, 275–305.

Vogler A.P., DeSalle R. 1994. Diagnosing units of conservation management. Conservation Biology. 8, 354 –363.

Walander P., Hoeck P.E.A., Keller L.F. 2007. Back to the future: museum specimens in population genetics. Trends in Ecology and Evolution. 22, 634-6442.

Waldman, J.R., Wirgin, I.I. 1998. Status and restoration options for Atlantic sturgeon in North America. Conservation Biology. 12, 631–638.

Wan Q.-H., Wu H., Fujihara T., Fang S.-G. 2004. Which genetic marker for which conservation genetics issue? Electrophoresis. 25, 2165–76.

Wang P., Meyer E., McKay J.K., Matz M.V. 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping. Nature methods. 9, 808-810.

Welsh A.B., Blumberg M. and May B. 2003. Identification of microsatellite loci in lake sturgeon, *Acipenser fulvescens*, and their variability in green sturgeon, *A. medirostris*. Molecular Ecology Notes. 3. 47-55.

William Bemis E.F., 2001. An overview of Acipenseriformes. pp.25–71.

Williams C., Pontén F., Moberg C., Söderkvist P., Uhlén M., Pontén J., Sitbon G., Lundeberg J. 1999. A high frequency of sequence alterations is due to formalin fixation of archival specimens. The American Journal of Pathology. 155, 1467-71.

Wolf D. E., Takebayashi N., Riesberg L. H. 2001. Predicting the risk of extinction through hybridization. Conservation Biology. 15. 1039-1053.

Wu L., Patten N., Yamashiro C.T., Chui B. 2002. Extraction and amplification of DNA from formalin-fixed paraffin-embedded tissues. Applied Immunohistochemistry and Molecular Morphology. 10. 269-274.

Zane L., Patarnello T., Ludwig A., Fontana F., Congi L. 2002. Isolation and characterization of microsatellites in the Adriatic sturgeon (*Acipenser naccarii*). Molecular Ecology Notes. 2. 586-588.

Zerunian, S. 2003. Piano d'azione generale per la conservazione dei pesci d'acqua dolce in Italia. Quaderni di Conservazione della Natura - Ministero dell'Ambiente e della Tutela del Territorio-. 17.

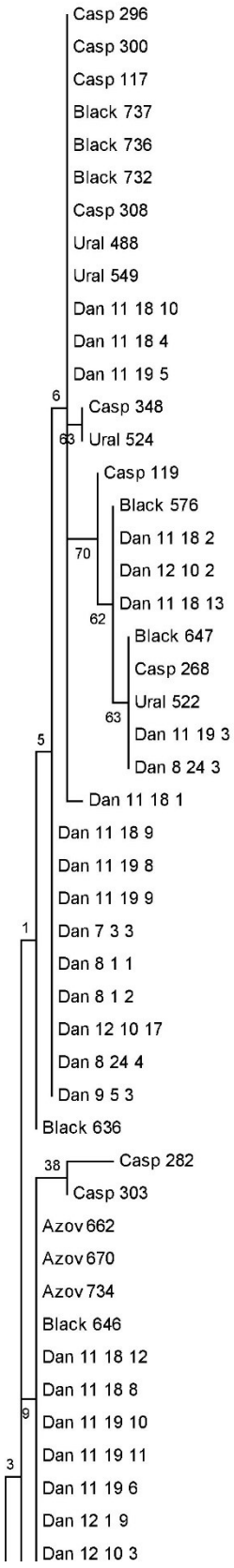
Zhang X., Wu W.H., Li L.M., Ma X.F. and Chen J.P. 2013. Genetic variation and relationships of seven sturgeon species and ten interspecific hybrids. Genetics Selection Evolution. 45, 21–30.

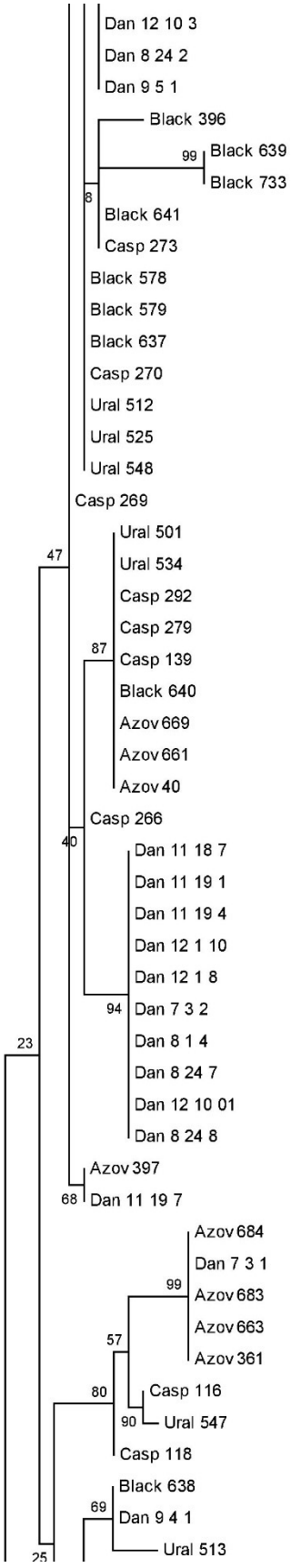
Zhu B., Liao X., Shao Z., Rosenthal H., Chang J. 2005. Isolation and characterization of microsatellites in Chinese sturgeon, *Acipenser sinensis*. Molecular Ecology Resources. 5. 888–892.

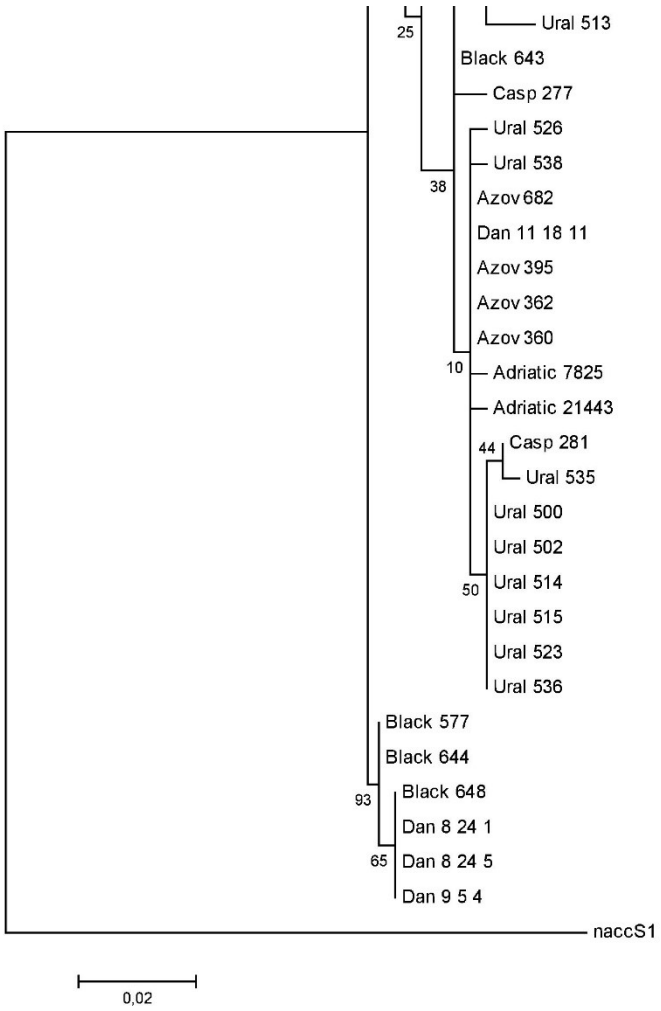
7.1. APPENDIX I

Molecular phylogenetic tree of D-loop fragment for *H. huso* individuals from different sampling locations using ML algorithms.

naccS1 indicates the outgroup species (*A. naccarii*);
Dan = Danube River;
Black = Black Sea;
Azov = Sea of Azov;
Casp = Caspian Sea;
Adriatic = Adriatic Sea.



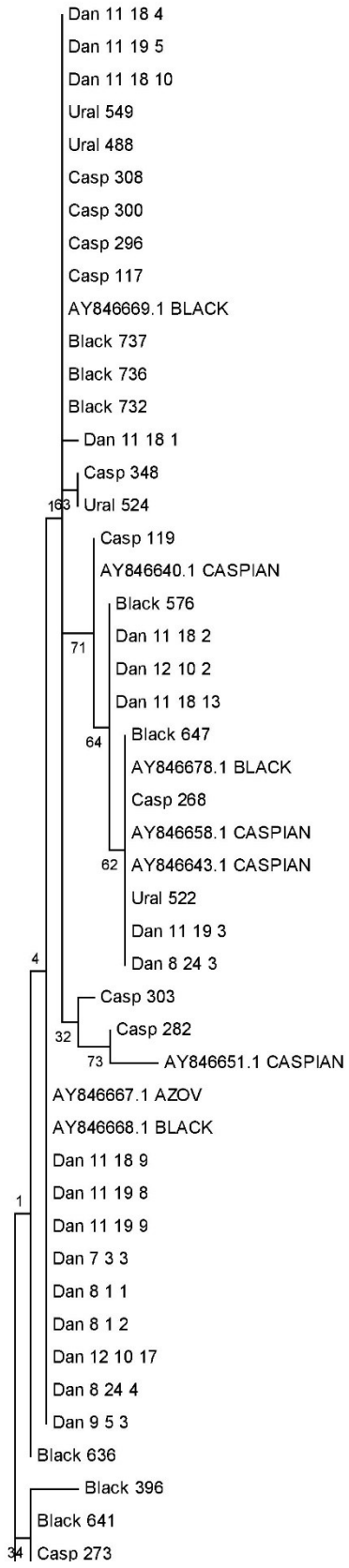

Dan 12 10 3
Dan 8 24 2
Dan 9 5 1
Black 396
99
Black 639
Black 733
6
Black 641
Casp 273
Black 578
Black 579
Black 637
Casp 270
Ural 512
Ural 525
Ural 548
Casp 269
Ural 501
47
Ural 534
Casp 292
Casp 279
87
Casp 139
Black 640
Azov 669
Azov 661
Azov 40
Casp 266
40
Dan 11 18 7
Dan 11 19 1
Dan 11 19 4
Dan 12 1 10
Dan 12 1 8
94
Dan 7 3 2
23
Dan 8 1 4
Dan 8 24 7
Dan 12 10 01
Dan 8 24 8
Azov 397
68
Dan 11 19 7
Azov 684
Dan 7 3 1
99
Azov 683
Azov 663
57
Azov 361
80
Casp 116
90
Ural 547
Casp 118
Black 638
69
Dan 9 4 1
Ural 513
25

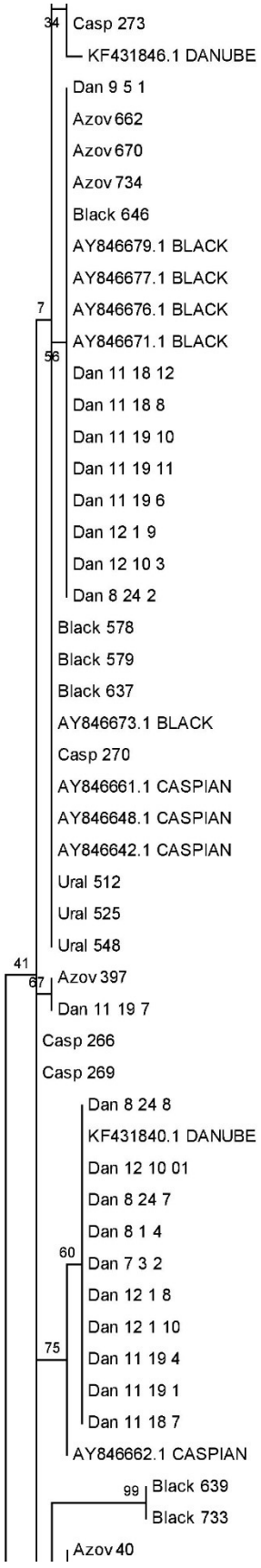

Ural 513
25
Black 643
Casp 277
Ural 526
Ural 538
38
Azov 682
Dan 11 18 11
Azov 395
Azov 362
Azov 360
10
Adriatic 7825
Adriatic 21443
44
Casp 281
Ural 535
Ural 500
Ural 502
50
Ural 514
Ural 515
Ural 523
Ural 536
Black 577
Black 644
93
Black 648
Dan 8 24 1
65
Dan 8 24 5
Dan 9 5 4
naccS1
0,02

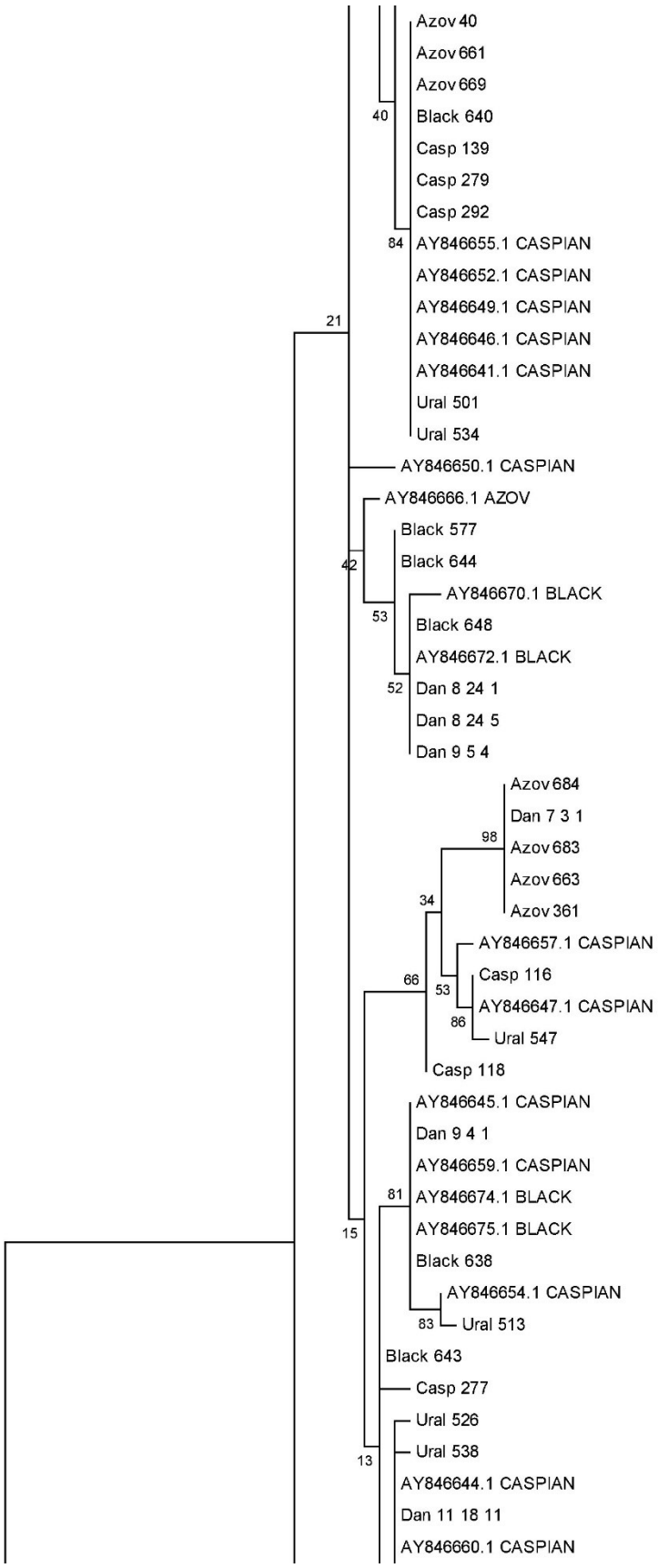
7. 2. APPENDIX II

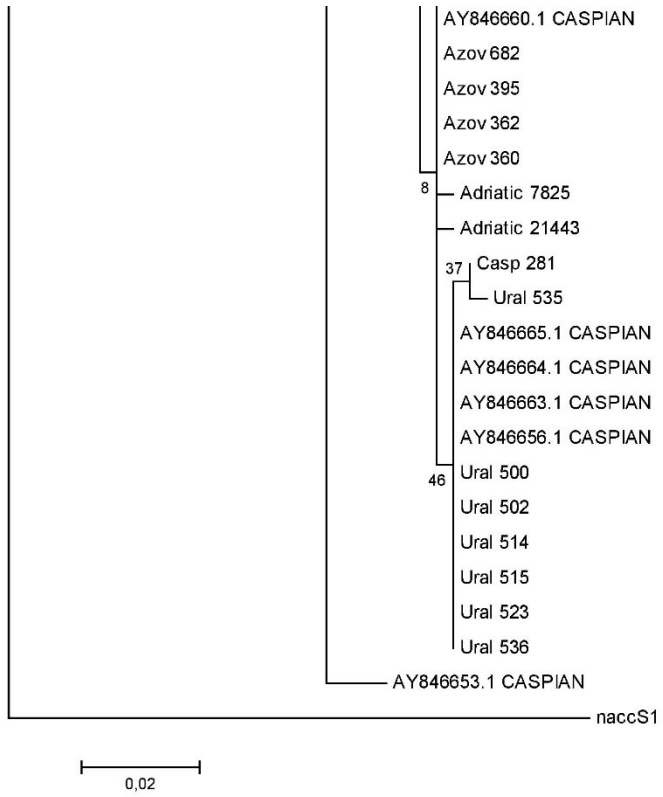
Molecular phylogenetic tree of D-loop fragment for *H. huso* individuals from different sampling locations (including sequences from GenBank) using ML algorithms.

naccS1 indicates the outgroup species (*A. naccarii*);
Dan = Danube River:
Black = Black Sea:
Azov = Sea of Azov:
Casp = Caspian Sea;
Adriatic = Adriatic Sea.




34
Casp 273
KF431846.1 DANUBE
Dan 9 5 1
Azov 662
Azov 670
Azov 734
Black 646
AY846679.1 BLACK
AY846677.1 BLACK
7
AY846676.1 BLACK
AY846671.1 BLACK
56
Dan 11 18 12
Dan 11 18 8
Dan 11 19 10
Dan 11 19 11
Dan 11 19 6
Dan 12 1 9
Dan 12 10 3
Dan 8 24 2
Black 578
Black 579
Black 637
AY846673.1 BLACK
Casp 270
AY846661.1 CASPIAN
AY846648.1 CASPIAN
AY846642.1 CASPIAN
Ural 512
Ural 525
Ural 548
41
Azov 397
67
Dan 11 19 7
Casp 266
Casp 269
Dan 8 24 8
KF431840.1 DANUBE
Dan 12 10 01
Dan 8 24 7
Dan 8 1 4
60
Dan 7 3 2
Dan 12 1 8
Dan 12 1 10
75
Dan 11 19 4
Dan 11 19 1
Dan 11 18 7
AY846662.1 CASPIAN
99
Black 639
Black 733
Azov 40


Azov 40
Azov 661
Azov 669
40
Black 640
Casp 139
Casp 279
Casp 292
84
AY846655.1 CASPIAN
AY846652.1 CASPIAN
AY846649.1 CASPIAN
21
AY846646.1 CASPIAN
AY846641.1 CASPIAN
Ural 501
Ural 534
AY846650.1 CASPIAN
AY846666.1 AZOV
Black 577
42
Black 644
AY846670.1 BLACK
53
Black 648
AY846672.1 BLACK
52
Dan 8 24 1
Dan 8 24 5
Dan 9 5 4
Azov 684
Dan 7 3 1
98
Azov 683
Azov 663
34
Azov 361
AY846657.1 CASPIAN
66
Casp 116
53
AY846647.1 CASPIAN
86
Ural 547
Casp 118
AY846645.1 CASPIAN
Dan 9 4 1
AY846659.1 CASPIAN
81
AY846674.1 BLACK
15
AY846675.1 BLACK
Black 638
AY846654.1 CASPIAN
83
Ural 513
Black 643
Casp 277
Ural 526
13
Ural 538
AY846644.1 CASPIAN
Dan 11 18 11
AY846660.1 CASPIAN


AY846660.1 CASPIAN
Azov 682
Azov 395
Azov 362
Azov 360
8
Adriatic 7825
Adriatic 21443
37
Casp 281
Ural 535
AY846665.1 CASPIAN
AY846664.1 CASPIAN
AY846663.1 CASPIAN
AY846656.1 CASPIAN
46
Ural 500
Ural 502
Ural 514
Ural 515
Ural 523
Ural 536
AY846653.1 CASPIAN
naccS1
0,02