UNISS
UNIVERSITÀ
DEGLI STUDI
DI SASSARI

University of Sassari

Department of Biomedical Sciences

PhD course in Life Sciences and Biotechnologies

XXXVI cycle

PhD thesis

From individuals to populations: molecular analyses, population genetics and modern phylogenetic approach (MPA) applied to different aquatic organisms

Tutors

Prof. Marco Casu
Department of Veterinary Medicine

Prof. Daria Sanna
Department of Biomedical Sciences

PhD student

Dr. Ilenia Azzena

Academic year

2022/2023

Summary

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

Chapter 1

Phylogenetics

The term "phylogeny" finds its roots in ancient Greek, combining "phýlon" (lineage) and "génesis" (development), and it fundamentally refers to the evolutionary development and history of living organisms (Liddell et al. 2023; Roy et al. 2014).
Since Ernst Haeckel's (1866) publication of the animal tree of life, significant progress has been made in understanding the evolutionary relationships among animals. This progress has been driven by decades of intensive phylogenetic research, which initially focused on anatomical and developmental characters (classic phylogeny) and later incorporated molecular data (modern phylogeny) (Dunn et al. 2014).

1. Modern phylogeny and phylogeography

Morphological data have traditionally been significant sources of information in the reconstruction of phylogenetic relationships among different taxa (Bybee et al. 2010). However, with the development of polymerase chain reaction and modern molecular techniques in the field of phylogenetics, DNA has emerged as a major and influential source for inferring these relationships (Bybee et al. 2010).
It all started in 1965 when Zuckerkandl and Pauling made a groundbreaking realization that the primary sequences found in nucleic acids and proteins held the potential to offer valuable insights into the evolutionary history of organisms. They recognized that by aligning and comparing these sequences, researchers could unravel key information about the evolutionary relationships among different species (Zuckerkandl & Pauling 1965).
The late 1960s marked a significant turning point in the field of phylogenetics. During this period, important developments in statistical methods for reconstructing phylogenies from genetic data emerged (Garamszegi et al. 2014). A major milestone came in 1967 when Cavalli-Sforza and Edwards introduced statistical approaches for

this purpose. Importantly, these statistical methods could be applied to various types of data, including continuous characters (Cavalli-Sforza & Edwards 1967).

Another pivotal step was taken by Felsenstein in 1973 when he proposed a method to calculate the likelihood of a tree based on a set of continuous traits. This work was notable because it allowed researchers to quantify the likelihood of different tree structures given continuous data (Felsenstein 1973). The full importance of these innovations became evident years later, when Felsenstein introduced a related method for calculating phylogenetically independent contrasts (PICs) (Felsenstein 1985). One notable advantage of this new method was its practicality, as the calculations could be performed using a hand calculator, making it more accessible to researchers (Felsenstein 1985).

The modern era of animal phylogenetics was initiated with cladistic analyses using partial 18S rRNA sequences (Field et al. 1988) and anatomical traits (Eernisse et al. 1992; Nielsen et al. 1996; Schram 1991). In this field, both DNA (nucleotide) and protein (amino acid) sequences serve as valuable tools for deducing phylogenetic relationships among homologous genes, organelles, or even organisms (Dunn et al. 2014). DNA sequences are particularly informative when studying closely related organisms due to their susceptibility to evolutionary changes. On the other hand, amino acid sequences are more stable, making them preferable for analyzing homologous genes from distantly related organisms (Dunn et al. 2014). Early research mainly focused on the nuclear small ribosomal subunit (SSU or 18S) gene, but contemporary hypotheses have been reinforced by a range of data sources, including the nuclear large ribosomal subunit 28S (Halanych 2004). Some studies integrated both molecular and morphological data, and collectively, these efforts provided substantial support for the "new animal phylogeny" (Adoutte et al. 2000; Halanych 2004). This revised perspective challenged traditional groupings like the annelid-arthropod clade Articulata and introduced new hypotheses, such as Ecdysozoa (Dunn et al. 2014).

Systematic biology plays a pivotal role in establishing ancestral relationships among known species, regardless of whether they are presently extant or extinct. It involves the tracing of branching pathways in the evolutionary history of life. This reconstruction is essential for the field of systematics, which is dedicated to unraveling

the evolutionary relationships and kinship among taxonomic groups of organisms at various systematic levels (Dunn et al. 2014).

In addition to this, systematic biology also delves into the historical dimensions of the current geographic patterns of gene lineages: phylogeography (Avise 1987; Avise 1996). The term "phylogeography" was coined in 1987 (Avise et al. 1987) and its prevalence in the field of evolutionary genetics has significantly increased over the years (Avise 1998). Phylogeography is a relatively young scientific discipline, within the broader field of biogeography, focused on understanding the principles and mechanisms that influence the geographic distribution of genealogical lineages, particularly within and among closely related species (Avise 1998; Avise 2009). The analysis and comprehension of lineage distributions usually require insights from various fields, including molecular genetics, population genetics, phylogenetics, demography, ethology, and historical geography. Consequently, phylogeography is an integrative discipline that draws upon a diverse range of expertise and knowledge (Avise 1996).

The origin of phylogeography can be attributed to the advancements in molecular techniques during the 1980s, which enabled the analysis of DNA sequence variation across the geographical range of a species (Emerson & Hewitt 2005). One of the key drivers of this field was the accessibility of mitochondrial DNA (mtDNA) sequences in animal species (Emerson & Hewitt 2005). Mitochondrial DNA has played a pivotal role in phylogeographic studies, primarily because it undergoes rapid evolution and secondly, it is usually maternally transmitted without recombination (Avise 1998). Consequently, more than 80% of phylogeographic studies conducted to date have heavily relied on mtDNA (Avise 1998).

This, in turn, allowed researchers to reconstruct gene genealogies, whose spatial relationships can be visually represented on maps and analyzed to infer the evolutionary history of populations, subspecies, and species (Emerson & Hewitt 2005).

In recent years, genome and transcriptome analyses have played a pivotal role in further refining our comprehension of animal relationships (Dunn et al. 2014). Initially, these investigations relied on Sanger sequencing techniques (Bourlat et al. 2006, Delsuc et al. 2006, Dunn et al. 2008, Hejnol et al. 2009, Philippe & Telford 2006, Philippe et al. 2005). However, more recently, the field has witnessed the increasing prevalence of

next-generation sequencing (NGS) technologies, which generate extensive datasets and have become a prominent feature of this research (Kocot et al. 2011, Lemmon & Lemmon 2013, Smith et al. 2011).

In conclusion, phylogenetic studies are crucial for addressing a wide range of scientific inquiries, including refining taxonomic classifications, and investigating complex evolutionary processes such as co-evolution and biogeography (Hillis et al. 1994). Animal phylogeny is not only inherently fascinating but also essential for various aspects of animal biology. It offers a framework for making meaningful comparisons among related taxa and helps answer questions related to the evolution of various biological traits, including complexity, at different levels (Dunn et al. 2014).

The primary goal of phylogeny is to construct accurate and detailed representations of the evolutionary history of groups of organisms or genes. To achieve this aim, phylogeny typically involves several key approaches, including molecular phylogenetic analyses, clustering analyses, and phylogeographic analyses (Nixon 2001). These methodologies are essential for uncovering and understanding the intricate relationships and shared ancestry among various entities, providing insights into the evolutionary past of life on Earth (Nixon, 2001).

2. Molecular phylogenetic analysis

Molecular phylogenetic analysis is a powerful tool for investigating the connections between sequences. By exploring these relationships, it allows us to infer the origins and evolutionary history of genes or groups of genes within the same or different populations (Retief, 1999). In this context, several key analyses play a crucial role:

- Phylogenetic signal;
- Evolutionary model;
- Phylogenetic tree;
- Molecular dating.

2.1. Phylogenetic signal

Phylogenetic signal is a fundamental concept in evolutionary biology, denoting the inclination of closely related species to exhibit greater similarity (Münkemüller et al. 2012). This concept is rooted in the expectation that species sharing a recent common ancestor will tend to have more similar traits, whether morphological or genetic, as a result of their shared evolutionary history (Münkemüller et al. 2012).

When working with DNA sequences and other molecular data, it's crucial to differentiate between patterns that reflect genuine phylogenetic signal, indicative of true evolutionary relationships, and patterns that are essentially random noise. The ability to distinguish signal from noise is paramount to ensure that research and analyses are based on data that accurately represent the historical relationships among the taxa being investigated (Hillis & Huelsenbeck 1992).

Quantifying phylogenetic signal is a pivotal aspect of various research areas, but this task is influenced by variables like trait choice, analysis scale, and the evolutionary model employed. A common approach is using multiple indices concurrently to gain a comprehensive understanding of the extent of phylogenetic signal, thereby enhancing research reliability (Münkemüller et al. 2012). Each index has distinct calculation methodologies, capturing various aspects of phylogenetic signal, and their responses can vary depending on factors like inaccuracies in phylogenetic information, sample size, and the availability of branch length data (Hillis & Huelsenbeck, 1992; Revell et al. 2008; Münkemüller et al. 2012; Blomberg et al. 2003; Cavender et al. 2006).

During my doctoral project, I choose to employ one of the widely recognized methods for quantifying phylogenetic signal: the likelihood mapping method (Strimmer & von Haeseler 1997). This approach is particularly useful for assessing the phylogenetic information within a dataset of aligned DNA or amino acid sequences. It revolves around the examination of quartets, which are randomly selected groups of four sequences. The widely used software for quartet analyses is TREE-PUZZLE (Schmidt et al. 2002). For each quartet, there exist three potential unrooted tree topologies, and the likelihood of each topology is estimated through the maximum likelihood method. These likelihood values are then graphically represented as points within an equilateral triangle, forming what is known as a likelihood map (Strimmer & von Haeseler 1996).

The likelihood map is divided into three primary regions, each conveying distinct information about the phylogenetic signal within the dataset:

i. Corners of the Map: represent fully resolved tree topologies, indicating a strong and clear tree-like phylogenetic signal in the data.
ii. Central Area: corresponds to a star-like phylogeny, suggesting a common ancestor with multiple lineages radiating from it.
iii. Sides of the Map: these regions indicate a network-like phylogeny, implying the presence of recombination or conflicting phylogenetic signals.

In silico simulations have revealed that if more than 33% of data points fall within the central area, it signifies a significant star-like signal, indicating the emergence of multiple phylogenetic lineages (Strimmer & von Haeseler, 1997; Schmidt & von Haeseler 2009).

Working with a dataset that has a low phylogenetic signal presents considerable challenges in interpreting results. A weak or obscured phylogenetic signal makes it difficult to accurately reconstruct the true evolutionary relationships among species or organisms. Moreover, it can hinder the detection of crucial evolutionary events like horizontal gene transfer or convergent evolution, which are essential for a comprehensive understanding of evolutionary processes. Hence, when dealing with datasets showing a low phylogenetic signal, it's important to exercise caution and consider alternative approaches or additional data sources to enhance the reliability of phylogenetic inferences (Strimmer & von Haeseler 2009).

2.2. Evolutionary model

Stochastic models have significantly advanced our comprehension of how evolutionary systems function, with a substantial portion of these methods finding a strong foundation in population genetics (Rodríguez et al. 1990; Blythe & McKane 2007).

The idea of utilizing random selection and mutation for optimization purposes finds its origins in the 1950s, with the work of statistician George Box who developed a methodology known as evolutionary operation (Tettamanzi 2005; Box & Draper 1969).

Around the same period, other researchers also explored the idea of simulating evolution through computer-based methods (Tettamanzi 2005). Barricelli and Fraser,

for instance, employed computer simulations to investigate the mechanisms of natural evolution (Bremermann 1962).

Selecting the appropriate evolutionary model is crucial for ensuring precise phylogenetic reconstructions. Indeed, this decision can significantly affect the estimation of essential phylogenetic parameters, such as divergence time and the rate of evolution (Posada & Crandall, 2001). In this context, nucleotide substitutions, which accumulate during the course of evolution, play a pivotal role in the analyses (Yang 1994). These substitutions have a broad range of effects on protein function and can also influence the regulation of gene expression by changing how transcription factors bind to DNA. Furthermore, nucleotide substitutions provide a method for inferring the evolutionary history of sequences (Yang 1994). Consequently, the examination of nucleotide substitutions holds significant importance for gaining an understanding of gene and protein structure and function, as well as for piecing together the evolutionary lineage of organisms (Yang 1994; Rodriguez et al. 1990).

Various evolutionary models are available, each with its level of complexity, and the choice of model depends on the specific dataset and research goals (see Figure 1).

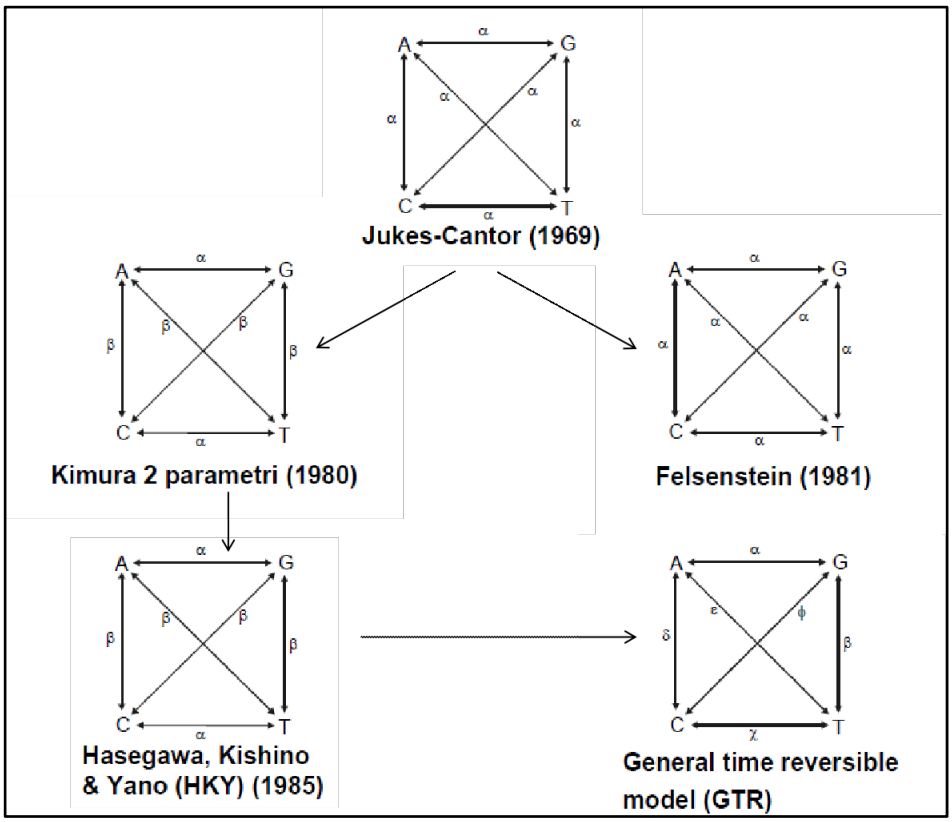


Figure 1. Representation of the substitution schemes of the main evolutionary models.

Researchers should carefully consider the complexity of the model in relation to the dataset in order to select the most appropriate one.

The Jukes-Cantor (JC) model serves as the simplest model, hinging on three fundamental assumptions: (i) each site in DNA or RNA has an equal chance of mutating into any of the four bases; (ii) the mutation rate from one base to another remains constant over time and (iii) mutations occur independently of each other and not influenced by prior mutations (Jukes & Cantor 1969). Although it's a basic model, JC serves as a foundational framework for more intricate models that consider variations in mutation rates between sites (Posada 2003; Posada & Buckley 2004; Posada & Crandall 2001; Raftery 1996).

A more complex model, the Kimura 2-Parameter (K2P), was drew up to address some of JC's limitations and provide a more precise representation of mutation rates in DNA or RNA sequences during evolution (Posada 2003). K2P distinguishes between two types of substitutions, recognizing that transitions and transversions occur at different frequencies in biological reality. This model offers a more nuanced and realistic depiction of nucleotide substitution patterns during evolution (Kimura 1980; Kimura 1981).

The Felsenstein 1981 model is of equivalent complexity to K2P, considering two types of changes as well. However, in this case, the base frequencies differ, and nucleotide substitutions have equal probabilities (Felsenstein 1981).

The HKY model can be seen as a fusion of previous models, incorporating different base frequencies with distinct probabilities for transitions and transversions (Hasegawa et al. 1985).

The General Time-Reversible (GTR) model stands as the most sophisticated model of substitution, offering complexity and flexibility (Hasegawa et al. 1985). It takes into account three critical aspects of sequence evolution: (i) stationary frequencies; (ii) exchangeability rates, and (iii) rate heterogeneity (Hasegawa et al. 1985). Stationary Frequencies allow for variable base frequencies, acknowledging that the four nucleotides may not be equally represented in the sequence. Exchangeability rates permit non-uniform rates of substitution between different nucleotides, considering that the substitution rate from one nucleotide to another and the reverse can differ. This flexibility provides a more accurate representation of sequence evolution (Posada

& Crandall 2001; Raftery 1996; Tavaré 1986). Rate heterogeneity accounts for variable substitution rates at different sites within the sequence, recognizing that some positions may evolve more rapidly or slowly than others (Posada & Crandall 2001). These features make the GTR model particularly valuable in phylogenetic analysis, aiding in the estimation of evolutionary trees and sequence distances (Posada & Crandall 2001).

In order to meet the requirement of identifying the most appropriate model, jModelTest utilizes a maximum likelihood approach to select the best-fitting evolutionary model from a range of options (Darriba et al. 2012). Additionally, the program allows for model comparison using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). These criteria consider the model's complexity and penalize models that prove overly complex for the available data (Darriba et al. 2012).

Furthermore, apart from model selection, jModelTest has the capability to estimate the parameters associated with the chosen model. These parameters encompass variables like the transition/transversion ratio, base frequencies, and the shape parameter for the gamma distribution. These parameters can be used to infer the evolutionary history of the sequences and to test hypotheses about the processes of molecular evolution (Darriba et al. 2012).

2.3. Phylogenetic tree

A phylogenetic tree is a visual representation that illustrates the evolutionary relationships among different species (Russel 2010; Dunn et al. 2014). It features branches connecting nodes, with terminal nodes representing taxa for genetic analysis and internal nodes symbolizing common ancestors. (Russel, 2010).

Phylogenetic trees have different applications and exist in various forms:

i. Phylogram

   These trees use branch lengths to indicate the genetic or evolutionary divergence between connected taxa.

ii. Cladogram

In cladograms, branch lengths are usually uniform and prioritize showing branching patterns and relationships among taxa without revealing the extent of divergence.

iii. Rooted

Rooted trees provide insights into the temporal order of evolutionary events and specify which taxa are more ancient or derived.

iv. Unrooted

Unrooted trees describe relationships between taxa but don't indicate which taxa are more ancient or derived. These can be transformed into rooted trees by establishing a root using an outgroup, a reference taxon known to be distinct from the other taxa (Russel 2010).

In this context, is possible to work with either a gene tree or a species tree. A gene tree primarily represents the evolutionary history of a specific gene, focusing on the gene's evolution. On the other hand, species trees are typically created by analyzing data from multiple genes (Russel 2010).

Methods for constructing phylogenetic trees can be based on various approaches, including distance matrices, maximum parsimony, maximum likelihood, and bayesian methods (Retief 1999; Nixon 2001; Guindon & Gascuel 2003; Douady et al. 2003; Huelsenbeck et al. 2001). It's important to note that methods based on distance matrices have a notable limitation because they result in a loss of information when distance data is extracted (Fitch & Margoliash 1967). An alternative approach to construct phylogenetic trees is maximum parsimony, which operates directly on the sequences (Saitou & Nei 1987; Jones & Pevzner 2004). In this method, the best tree is considered the simplest one, with branch lengths defined as the minimum number of substitutions occurring between connected nodes (Kannan & Wheeler 2012). Otherwise, maximum likelihood (ML) and bayesian methods rely on probabilistic models of sequence evolution to assess the compatibility between the observed data and different phylogenetic trees along with their associated parameters (Guindon & Gascuel 2003; Douady et al. 2003). ML methods seek to find the tree that maximizes this likelihood (Guindon & Gascuel 2003). Bayesian methods use a bayesian framework to infer phylogenies by sampling tree and parameter space (Douady et al. 2003).

Each of these approaches has its strengths and weaknesses and the choice of method should depend on the specific characteristics of the data and the research goals. Notably, maximum likelihood and bayesian methods stand out for their capacity to accommodate diverse evolutionary processes, such as different types of substitutions and site-specific rate variations (Douady et al. 2003).

2.3.1. Maximum likelihood and bayesian methods

These methods provide statistical support for inferred relationships and are considered among the most accurate approaches when sufficient computational power is available (Huelsenbeck et al. 2001; Douady et al. 2003).

Nonetheless, the computational time required often limits the practical application of model-based methods like Maximum Likelihood (ML), particularly when dealing with a large number of taxa (Felsenstein 1985). This time constraint is one of the factors contributing to the growing popularity of bayesian inference methods (e.g., Karol et al. 2001; Lutzoni et al. 2001; Murphy et al. 2001) as implemented in the software MrBayes (Huelsenbeck & Ronquist 2001) used during the doctoral project.

MrBayes applicability goes beyond the conventional methods of constructing trees, as it provides a bayesian framework for deducing phylogenetic trees and evaluating the posterior probabilities associated with these trees (Huelsenbeck & Ronquist 2001). These trees serve as representations of the hypothetical evolutionary history of the species under investigation (Huelsenbeck & Ronquist 2001; Ronquist et al. 2005).

In contrast to older methods like maximum likelihood or distance-based approaches, bayesian inference offers the capability to incorporate prior knowledge, which is expressed through prior probability distributions, alongside the likelihood of observing the data given a specific model and phylogenetic tree. This merging of prior information and observed data is a distinctive feature of bayesian analysis, and it holds particular significance in the field of phylogenetics, where the evolutionary process is shaped by intricate molecular models (Ronquist et al. 2009).

2.4. Molecular dating

In 1965, Zuckerkandl and Pauling introduced the concept of estimating the timing of evolutionary divergences through the examination of calibrated sequence variations (Zuckerkandl & Pauling 1965). Their proposal implied that the degree of dissimilarity observed in the DNA molecules of two species reflects the time that has passed since their evolutionary separation (Zuckerkandl & Pauling 1965). Subsequently, countless studies have been conducted based on this idea (Rutschmann 2006).

2.4.1. Molecular clock

In the specific scenario of a molecular clock, all branches within a phylogenetic tree evolve at a consistent, uniform substitution rate. This results in a "clock-like" tree that is ultrametric, signifying that the overall distance between the root and every tip remains constant (Rutschmann 2006).

Similar to a standard clock that measures time in hours and minutes, the molecular clock quantifies evolutionary time in terms of genetic changes through the application of a mathematical model that has revolutionized our comprehension of the processes through which species evolve and differentiate (Bromham & Penny 2003).

The fundamental concept behind the molecular clock aligns with the principles of the Neutral Theory of Molecular Evolution, which declares that genetic mutations accumulate at a relatively consistent rate over time, serving as the clock's "ticking" (Zuckerkandl & Pauling 1965; Lemey et al. 2012). Although the mutation rate may vary among genes and organisms, the fundamental idea remains constant. Genetic sequences, whether in the form of DNA or proteins, effectively document the changes that have transpired throughout evolutionary history (Lemey et al. 2012).

The accuracy of the molecular clock relies on effective calibration, often achieved through the utilization of fossils (Warnock & Donoghue 2017). Fossils, indeed, provide precise age references for specific species or lineages, and when a genetic sequence can be associated with a well-dated fossil, it serves as a calibration point. This, in turn, allows for the mapping of genetic changes onto the geological timescale and aids in the estimation of divergence times (Warnock & Donoghue 2017).

Nonetheless, fossils may not always be accessible or suitable for calibration purposes. In such situations, alternative methods can be employed, such as relying on geological

events or historical biogeographic data (Knowlton & Weigt 1998). This approach is known as "node dating," where calibration points are determined by the ages of nodes (branching points) in a phylogenetic tree (Knowlton 1993; Knowlton & Weigt 1998). Node dating is particularly prevalent in the field of marine biology, especially when studying soft-bodied species (Knowlton 1993). Within this context, one of the most frequently referenced geographic barriers is the Isthmus of Panama, which closed relatively recently in geological time (approximately 3.1–3.5 million years ago due to the final emergence of the Isthmus) (Knowlton & Weigt 1998).

Several methods can be used to construct a molecular clock but during this project I focused my attention on the bayesian methods as they are considered the forefront of molecular dating methodologies (Dos Reis & Yang 2019).

Essentially, bayesian algorithms utilize robust statistical techniques and incorporate data with prior knowledge, thus offering a flexible and versatile toolkit for exploring the intricate facets of the evolutionary tree of life (Dos Reis et al. 2016).

One of the prominent software applications that implements the molecular algorithm is BEAST (Bayesian Evolutionary Analysis by Sampling Trees) (Drummond & Rambaut 2007). Its primary focus is generating phylogenies that are both rooted and time-measured, allowing for the application of either strict or relaxed molecular clock models (Drummond & Rambaut 2007). BEAST employs Markov Chain Monte Carlo (MCMC) methods to explore the space of potential trees and assigns probabilities to each tree based on its posterior probability. To make the software more user-friendly, BEAST provides an intuitive interface for configuring standard analyses, and it includes a set of companion programs for interpreting and analyzing the data obtained after the analysis (Drummond & Rambaut 2007; Suchard et al. 2018).

2.4.2. Bayesian Skyline Plot and Lineages Through Times

Additional outputs from BEAST analyses can be further explored using software like Tracer (Rambaut & Drummond 2009), which enables the reconstruction of the "Bayesian Skyline Plot" (BSP) and "Lineages Through Times" (LTT) (Suchard et al. 2018). These graphical representations are valuable tools for understanding the historical demographics of species and gaining insights into their evolutionary paths. The BSP provides a visual representation of population dynamics, shedding light on

how effective population sizes have evolved over time and revealing patterns of growth, decline, or stability (Drummond et al. 2011; Ho & Shapiro 2011). On the other hand, the LTT allows for the exploration of diversification rates and the inference of patterns related to speciation and extinction events (Stadler 2008; Harvey et al. 1994).

3. Clustering analysis

Clustering analysis is defined as the process of taking a given set of samples and partitioning it into distinct and homogeneous groups (Lee, 1981; Garcia-Dias, 2020). Typically, this is accomplished by employing unsupervised algorithms that organize a set of n observations (X1, X2, ..., Xn) into K groups (g1, g2, ..., gK) using a similarity criterion. This criterion ensures that observations within the same group exhibit more similarity to each other than they do to observations in different groups (Garcia-Dias 2020). Clustering methods aim to find these natural groupings or patterns within the data without any prior knowledge of the categories (Garcia-Dias 2020).

3.1. Principal Coordinates Analysis (PCoA)

Principal Coordinates Analysis (PCoA) is a clustering analysis method designed to calculate a distance matrix and generate a visual representation in a lower-dimensional Euclidean space, typically in two or three dimensions. In this representation, the distances between points aim to closely mirror the original distances, adhering to the principles of the Pythagorean theorem (Zuur et al. 2007).

In the current doctoral project, the analysis of PCoA is conducted using the GeneAlex software. GeneAlex distinguishes itself with its user-friendly interface, allowing for the analysis of real genetic datasets in a familiar software environment, specifically Microsoft Excel (Peakall & Smouse 2006).

The PCoA analysis can be broken down into three essential steps: (i) computation of a distance matrix, (ii) formatting the matrix by making adjustments to both rows and columns, and (iii) performing eigen-decomposition on the distance matrix (Garcia-Dias, 2020). The rescaled eigenvectors are representative of the principal coordinates, also referred to as principal axes (Peakall & Smouse 2006; Zuur et al. 2007).

In the resulting plot, the percentage value obtained for the horizontal axis holds the greatest significance in terms of results since it reflects the real distances that separate the groups (Zuur et al. 2007).

3.2. Species delimitation methods

Species identification holds significant importance across various biological research domains, including the fields of evolution, conservation, and biodiversity (De Queiroz 2007). However, species identification relies on distinct operational criteria depending on the specific species concept in use. (De Queiroz 2007). Two of the most utilized concepts are the biological species concept, relying on reproductive isolation (Mayr, 1942; Dobzhansky, 1950), and the phylogenetic species concept, which is founded on reciprocal monophyly (Rosen 1979; Baum & Shaw 1995). On the other hand, morphology-based taxonomy often corresponds with the phenetic species concept (Michener 1970; Sokal & Crovello 1970), which continues to serve as a fundamental framework for practical species identification.

The past decade has witnessed the increasing availability of genetic methods for species identification, offering a valuable complement to morphological taxonomy (Luo et al. 2018). In this context, molecular species delimitation methods offer valuable insights into species classification and identification, particularly when conventional taxonomic approaches demonstrate insufficient (Luo et al. 2018). These molecular approaches leverage genetic data to uncover hidden diversity, identify cryptic species, and enhance our understanding of evolutionary relationships (Luo et al. 2018). Molecular species delimitation not only aids in clarifying taxonomic uncertainties but also plays a role in biodiversity conservation and the creation of accurate phylogenetic trees (Luo et al. 2018).

In general, species delimitation methods can be categorized into two groups: those based on the Phylogenetic Species Concept (PSC) and those relying on genetic distances (Scarpa et al. 2016).

3.2.1. Methods based on Phylogenetic Species Concept (PSC)

The two prominent molecular species delimitation methods that operate based on the PSC include the Generalized Mixed Yule Coalescent (GMYC) method, originally

developed by Pons et al. (2006), and the Poisson Tree Processes (PTP) model, along with its Bayesian variant known as bPTP, which was introduced by Zhang et al. (2013).
The GMYC model is designed to detect significant changes in the branching rate within an ultrametric species tree (Pons et al. 2006). To perform this analysis, you need an ultrametric species tree generated from dating analyses. The species delimitation runs were conducted using the SPLITS (SPecies LImits by Threshold Statistics) package developed by Ezard et al. (2009), which is implemented in the R statistical environment. This algorithm enables the identification of species entities using a single threshold option, which determines the transition from between-species to within-species branching (Ezard et al. 2009). It's important to note that although there is a multi-threshold option available, it is generally regarded as less precise and not the recommended choice from a technical standpoint.
In contrast, the PTP and bPTP models evaluate speciation rates based on the number of substitutions (Zhang et al. 2013). These models exhibit versatility as they can be utilized for species delimitation on both ultrametric and conventional trees. The analyses were carried out using the bPTP web server, accessible at http://species.h-its.org/ptp/. Typically, default settings with 500,000 Markov Chain Monte Carlo (MCMC) generations were employed. To ensure the reliability of the results, it's essential to examine each run for convergence by reviewing the likelihood plot. Convergence is considered achieved when the chain consistently remains within high-likelihood regions throughout the run (Zhang et al. 2013).

3.2.2. Methods based on genetic distances

One of the frequently utilized genetic distance-based approaches for species delimitation is the Automatic Barcode Gap Discovery (ABGD) method, introduced by Puillandre et al. (2012). Remarkably, this method does not consider the phylogenetic relationships within the dataset; instead, it relies solely on sequence data. The ABGD method identifies the barcode gap as the initial substantial gap surpassing a predetermined threshold, which is then utilized to partition the dataset (Puillandre et al. 2012). This analysis can be performed on a local PC, or on the ABGD online tool, accessible at http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html. To obtain

accurate species estimates, gene-specific priors are applied to define the maximum allowable intraspecific diversity, typically represented as P = 0.001.

Another genetic distance-based species delimitation method is the Nucleotide Divergence Threshold (NDT), which needs to be applied to individual genes. The analysis was conducted using a custom script developed by Scarpa et al. (2019) within the R statistical environment. This script categorizes taxa into distinct entities by employing specific thresholds for individual genes, using a pairwise Kimura two-parameter model (K2P) (Kimura, 1980) genetic distances matrix. The script is available at https://cran.rproject.org/.

3.2.3. Hybrid method

Another method for species delimitation, known as the K/θ method, offers a hybrid approach that combines elements of the PSC and genetic distance-based methods. This method was introduced by Birky et al. (2010) and is based on a speciation model that relies on the "4x rule." According to this rule, if the ratio between the average pairwise sequence difference between two clades (referred to as K) and the average pairwise sequence difference within a single clade (denoted as θ) is greater than 4, it suggests that the clades represent different species, as proposed by Birky (2013). Therefore, the K/θ method is used to estimate the likelihood that two samples belong to distinct evolutionary species, following the framework outlined by Birky et al. (2010).

4. Phylogeographical analysis: median-joining network

Reconstructing phylogenetic trees from intraspecific data, like variations in mitochondrial DNA, presents a significant challenge due to the combination of a large number of samples and the relatively minor genetic differences between individuals (Bandelt et al. 1999). This complexity leads to the generation of multiple plausible trees (Bandelt et al. 1999). To address this challenge and establish a coherent relationship among lineages, Bandelt et al. (1999) introduced the median-joining (MJ) method for constructing network analysis.

Network analysis is employed to infer the phylogeographic distribution pattern of a species, identify potential discrete genetic clusters, and understand the evolutionary

forces that have influenced its populations (Morrison 2010; Wu et al. 2008; Leigh & Bryant 2015).

Throughout my doctoral project, I utilized MJ network analysis by means of the software Network 10.2.0.0. The MJ method combines principles from Kruskal's algorithm, which seeks to identify minimum spanning trees by favoring shorter connections, and Farris's maximum-parsimony (MP) heuristic algorithm (Bandelt et al. 1999). In the MJ method, new vertices known as "median vectors" are systematically introduced. Biologically, these median vectors can be regarded as potentially existing but unsampled sequences or as ancestral sequences that have gone extinct (Bandelt et al. 1999).

In the graphical output, each lineage or haplotype is depicted as a spot, and the size of each spot is proportional to the number of individuals sharing that specific lineage or haplotype. These spots are linked to one another through branches whose length corresponds to the number of mutations that differentiate two haplotypes.

However, it's important to note that numerous techniques for deducing haplotype networks have been documented in existing literature (Bandelt et al. 1999; Clement et al. 2002; Cassen et al. 2005; Manolopoulou et al. 2011). These methods can used different approach (e.g; maximum parsimony; minimum spanning; median-joining ; integer neighbour-joining) and some of these have been integrated into software applications like TCS (Clement, Posada & Crandall 2000), NETWORK (http://www.fluxus-engineering.com), SPLITSTREE (Huson & Bryant 2006), and PEGAS (Paradis 2010).

Notably, Cassen et al. (2005) suggested that in situations where internal node haplotypes are not part of the sample, the median-joining and maximum parsimony methods offer the most accurate estimation of the actual genealogy, while the minimum-spanning algorithm demonstrates notably inferior performance.

Population genetics

Population genetics is a biological discipline focused on examining the genetic makeup of biological populations and how it evolves due to various factors, including natural selection (Okasha 2006; Amorim 2013).

This discipline emerged from the need to reconcile Charles Darwin's theory of evolution by natural selection with Gregor Mendel's work on heredity (Bowler 1990; Okasha 2006). While Darwin's theory was widely accepted, there was uncertainty about the mechanism of inheritance and whether natural selection alone could explain evolution (Jenkin 1867; Okasha 2006). The rediscovery of Mendel's work in 1900 and the eventual synthesis of Mendelian genetics with Darwinian evolution marked a significant turning point in our understanding of heredity and adaptation (Okasha 2006). Fisher played a pivotal role by showing that Mendelian inheritance could explain the distribution of continuously varying traits (Fisher 1919). This contributed to the reconciliation of Mendelism and Darwinism by illustrating that minor genetic changes could accumulate gradually, resulting in the development of intricate adaptations (Fisher 1919; Okasha 2006).

The 1920s and 30s saw the development of formal mathematical models by Fisher, Haldane, and Wright, which allowed researchers to quantitatively explore how various factors, including natural selection, mutation, and genetic drift, influenced a population's genetic makeup over time (Fisher 1919; Haldane,1927; Wright 1931). Fisher and Haldane emphasized the importance of natural selection, while Wright also recognized the significance of random genetic drift and migration in shaping genetic variation (Fisher 1919; Haldane,1927; Wright 1931).

The study of population genetics has evolved significantly with advances in molecular biology and gene sequencing, nonetheless, many of the foundational models and concepts developed by early population geneticists, including Fisher, Haldane, and Wright, continue to be relevant and influential in the field of genetics and evolutionary biology (Okasha 2006; Charlesworth & Charlesworth 2017). The ongoing exploration of the balance between natural selection and random genetic drift, as well as other fundamental aspects of variation and evolution, continues to be pertinent in contemporary evolutionary genetics (Charlesworth & Charlesworth 2017).

In general, organisms typically give rise to offspring that share strong resemblances with their parents. Nonetheless, having some degree of genetic diversity within a population is crucial for its survival, especially when significant environmental changes occur. A population composed of genetically identical individuals would face challenges in adapting to new conditions. Genetic variability provides the opportunity for certain individuals to adapt and thrive in a changing environment (Pierce 2016).

In 1908, the English mathematician Godfrey Harold Hardy (1877 - 1947) and the German physician Wilhelm Weinberg (1862 - 1937) independently addressed the challenge of preserving genetic diversity, specifically concerning the maintenance of recessive alleles within a population (Pierce 2016). Their mathematical investigations led to the conclusion that the overall genetic composition of a diploid Mendelian population remains constant across generations, even in the presence of genetic recombination. This phenomenon is known as the Hardy-Weinberg equilibrium, as described by Hardy (1908) and Weinberg (1908).

The foundation of this principle lies in the assumptions that the population is sizable, engages in random mating, and remains unaffected by mutations, migration, and natural selection. When these conditions are satisfied, allele frequencies remain stable over time, and genotype frequencies reach equilibrium after a single generation, following the proportions specified by the Hardy-Weinberg equilibrium: p^2, $2pq$, and q^2, where p and q denote the allele frequencies (Hardy 1908; Weinberg 1908).

1. Evolutive forces can shape genetic variability

In line with the mathematical model presented by Hardy-Weinberg, the gene pool of a population is expected to remain stable across generations, with no shifts in the frequencies of different genotypes (Hardy 1908; Weinberg 1908).

In practice, achieving this ideal equilibrium in the real world is a rare occurrence. This is mainly because the five conditions mentioned earlier rarely align simultaneously, making it highly unlikely for a natural population to maintain a state of equilibrium. In fact, if any one of these conditions is not met, the population will undergo evolutionary changes and genetic shifts over time (Pierce 2016).

Evolutionary forces refer to the various mechanisms and factors that drive evolutionary change and influence the genetic composition of populations over time (Pierce 2016). These forces can be categorized into two groups: those that alter allelic frequencies (mutation, migration, genetic drift, and natural selection) and those that affect genotype frequencies (non-random mating).

1.1. Mutation

Genetic variability within a population is essential for evolution to occur. Therefore, all evolutionary processes depend on mechanisms that create genetic diversity. While meiosis and recombination can produce new combinations of existing genes, it's important to note that all genetic variations ultimately originate from mutations (Pierce 2016). In situations where the sole evolutionary force acting on a population is mutation, allele frequencies undergo gradual changes over time as some alleles mutate into others.

Given that mutation rates are generally low for most genes, the change in allele frequency caused by mutations within a single generation is extremely small.

In summary, the impact of mutation rates on Hardy-Weinberg equilibrium is negligible, and it requires many generations for a population to reach mutational equilibrium. Nonetheless, if mutation remains the primary driving force on a population over an extended timeframe, mutation rates will eventually dictate allele frequencies (Pierce 2016).

1.2. Migration

Migration, or gene flow, is another mechanism that can influence changes in allele frequencies, by introducing alleles from other populations (Pierce 2016). Although the Hardy-Weinberg law assumes the absence of migration, many natural populations experience the influx of individuals from other populations. Migration has two primary effects. Firstly, it promotes the convergence of gene pools among populations, maintaining a degree of uniformity in allele frequencies across different populations. Secondly, migration contributes to increased genetic diversity within populations. Furthermore, rare mutational events can give rise to unique alleles within different

populations, and these alleles can disseminate through migration to new populations, consequently enriching the genetic diversity within the recipient population. (Pierce 2016).

1.3. Genetic drift

The Hardy-Weinberg law operates under the assumption of an infinitely large population and random mating. In these ideal conditions, gametes would precisely replicate the parental gene pool. However, real populations are never of infinite size. As a result, in finite and small populations, when organisms reproduce, only a limited number of gametes combine to create the next generation (Pierce 2016). Purely by chance, this subset composition will frequently diverge from that of the parental gene pool, resulting in shifts in allele frequencies (Pierce 2016).

This type of deviation from an expected ratio due to the constraints of a limited sample size is termed "sampling error."

Genetic drift always originates from sampling errors, but there are several ways in which these errors can occur: (i) a population may undergo a gradual reduction in size across multiple generations due to limitations related to factors like space, food, or critical resources; (ii) founder's effect and (iii) bottleneck effect (Pierce 2016).

1.3.1. Founder effect

The founder effect is a source of sampling error and occurs when a population is founded by a small number of individuals. The chance events that determined which genes were present in the founders will significantly shape the genetic makeup of the entire population (Pierce 2016).

1.3.2. Bottleneck effect

The third source of genetic drift is a phenomenon known as a genetic bottleneck. This transpires when a population experiences a dramatic reduction in its numbers. Unlike the founder effect, this reduction in population size is not due to outward migration but rather is typically the result of some form of event, often catastrophic (Pierce 2016; Binelli & Ghisotti 2018).

Due to its random nature, genetic drift can result in both increases and decreases in allele frequencies within populations, leading to a fluctuating pattern over time, hence the term "drift." Another outcome of genetic drift is the reduction of genetic diversity within populations (Pierce 2016).

1.4. Natural selection

The ultimate mechanism responsible for shifts in allele frequencies is natural selection (Pierce 2016). It comes into play when individuals with advantageous traits have a greater number of offspring compared to those without these traits. When these advantageous traits have a genetic basis, they tend to become more prevalent over time. Consequently, traits offering reproductive benefits become more widespread, enhancing the population's adaptation to its environment. Notably, among all the forces of evolution, natural selection is the primary driver of adaptation (Pierce 2016).

The impact of natural selection on a population's gene pool depends on the fitness (relative reproductive success of a genotype) values of the genotypes within that population. Through the process of selection, disparities in fitness among genotypes result in changes in genotype frequencies over time. These changes subsequently lead to variations in the frequencies of alleles found within these genotypes. (Pierce 2016).

1.5. Non-random mating

A fundamental assumption of the Hardy-Weinberg law is that mating occurs randomly concerning genotypes. Non-random mating, indeed, disrupts how alleles combine to produce genotypes and consequently alters the genotype frequencies within a population (Pierce 2016).

Nonrandom mating comes in two forms: positive assortative mating, where individuals prefer those with similar traits (e.g., tall with tall), and negative assortative mating, where individuals prefer those with dissimilar traits (e.g., tall with short). This mating behavior typically influences genes related to the specific trait in question (Pierce 2016).

1.5.1. Inbreeding

Inbreeding is a type of non-random mating and involves the preference for mating with close biological relatives. It's essentially a form of positive assortative mating but distinct from other types because it influences all genes within a population, rather than just those associated with a specific preferred trait (Pierce 2016).

Inbreeding results in deviations from the expected p^2, $2pq$, and q^2 frequencies under Hardy-Weinberg equilibrium, elevating the proportion of homozygotes and reducing the number of heterozygotes (Pierce 2016).

General aim of the doctoral project and species involved

The primary aim of this doctoral project is to harmoniously blend molecular and phylogenetic analyses in the exploration of three case studies centered around aquatic organisms. In particular, I focused my attention on two species, *Pinna nobilis* and *Salariopsis fluviatilis*, which are experiencing different conservational issues, and on third species, *Procambarus virginalis*, which conversely represents in its present range of distribution a threat to the autochthonous biodiversity.

Pinna nobilis, the largest bivalve mollusc native to the Mediterranean Sea, holds a unique and endemic status in this region (Butler et al. 1993). This species plays a pivotal role in marine ecosystems since it serves a dual purpose in conservation efforts, being both a flagship and a keystone species (see Scarpa et al. 2021; Nebot-Colomer et al. 2022a, and references therein). Currently, *Pinna nobilis* is under a severe threat due to recent mass mortality events that have affected its populations throughout the Mediterranean Sea (Scarpa et al. 2020).

Salariopsis fluviatilis is a benthic fish species which inhabits the freshwater ecosystems of various Mediterranean countries (Zander 1986). Presently, the species is experiencing a significant decline in local populations due to its vulnerability to environmental changes and pollution (Laporte et al. 2014). The underlying causes are frequently linked to habitat fragmentation and can be aggravated by factors like eutrophication, water pollution in lakes, and the introduction of invasive predators (Laporte et al. 2014).

Procambarus virginalis, on the other hand, is an alien freshwater crayfish species, whose presence in Europe was first reported in the mid-1990s (Scholtz et al. 2003). This species is considered a perfect invader due to its capability to reproduce through parthenogenesis (Chucholl et al. 2012). Its presence in European freshwater habitats poses numerous threats, including competition with native crayfish, as well as potential adverse effects on amphibians, invertebrates, and the environment. These threats arise from intraspecific and interspecific competition, the potential transmission of crayfish plague, and its ecosystem-altering activities (Oidtmann et al. 1999; Souty-Grosset et al. 2016; Statzner et al. 2003).

References

- Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'Homme, B., & De Rosa, R. (2000). The new animal phylogeny: reliability and implications. Proceedings of the National Academy of Sciences, 97(9), 4453-4456.
- Amorim, A. (2013). Population Genetics, Editor(s): Stanley Maloy, Kelly Hughes, Brenner's Encyclopedia of Genetics (Second Edition), Academic Press, pp 407-411, ISBN 9780080961569.
- Avise, J. C. (1996). Space and time as axes in intraspecific phylogeography. In Past and future rapid environmental changes: the spatial and evolutionary responses of terrestrial biota (pp. 381-388). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. Molecular Ecology, 7(4), 371-379.
- Avise, J. C. (2009). Phylogeography: retrospect and prospect. Journal of biogeography, 36(1), 3-15.
- Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., ... & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual review of ecology and systematics, 18(1), 489-522.
- Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution, 16(1), 37-48.
- Baum, D. A., & Shaw, K. L. (1995). Genealogical perspectives on the species problem. Experimental and molecular approaches to plant biosystematics, 53(289-303), 123-124.
- Binelli, G; Ghisotti, D. (2018). Genetica. Genetica di popolazioni, 22, 602-634.
- Birky Jr, C. W. (2013). Species detection and identification in sexual organisms using population genetic theory and DNA sequences. PLoS One, 8(1), e52544.
- Birky Jr, C. W., Adams, J., Gemmel, M., & Perry, J. (2010). Using population genetic theory and DNA sequences for species detection and identification in asexual organisms. PLoS one, 5(5), e10609.
- Blomberg, S. P., Garland Jr, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution, 57(4), 717-745.
- Blythe, R. A., & McKane, A. J. (2007). Stochastic models of evolution in genetics, ecology and linguistics. Journal of Statistical Mechanics: Theory and Experiment, 2007(07), P07018.
- Bourlat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Aronowicz, J., Kirschner, M., ... & Telford, M. J. (2006). Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. Nature, 444(7115), 85-88.
- Bowler, P.J. (1990). The non-Darwinian revolution. Reinterpreting a historical myth. The Johns Hopkins University Press, Baltimore: 1988. Pp xii, 239, ISBN 0-8018-3678-6.
- Box, G. E., & Draper, N. R. (1969). Evolutionary operation: A statistical method for process improvement. John Wiley & Sons.
- Bremermann, H. J. (1962). Optimization through evolution and recombination. Self-organizing systems, 93, 106.
- Bromham, L., & Penny, D. (2003). The modern molecular clock. Nature Reviews Genetics, 4(3), 216-224.
- Butler, A., Vicente, N., & de Gaulejac, B. (1993). Ecology of the pterioid bivalves *Pinna bicolor* Gmelin. Mar Life, 3(1-2), 37-45.
- Bybee, S. M., Zaspel, J. M., Beucke, K. A., Scott, C. H., Smith, B. W., & Branham, M. A. (2010). Are molecular data supplanting morphological data in modern phylogenetic studies?. Systematic Entomology, 35(1), 2-5.
- Cassens, I., Mardulyn, P., & Milinkovitch, M. C. (2005). Evaluating intraspecific "network" construction methods using simulated sequence data: do existing algorithms outperform the global maximum parsimony approach?. Systematic biology, 54(3), 363-372.

- Cavalli-Sforza, L. L., & Edwards, A. W. (1967). Phylogenetic analysis. Models and estimation procedures. American journal of human genetics, 19(3 Pt 1), 233.
- Cavender-Bares, J., Keen, A., & Miles, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. Ecology, 87(sp7), S109-S122.
- Charlesworth, B., & Charlesworth, D. (2017). Population genetics from 1966 to 2016. Heredity, 118(1), 2-9.
- Chucholl, C., Morawetz, K., & Groß, H. (2012). The clones are coming--strong increase in Marmorkrebs [*Procambarus fallax* (Hagen, 1870) f. *virginalis*] records from Europe. Aquatic invasions, 7(4).
- Clement, M., Posada, D. C. K. A., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. Molecular ecology, 9(10), 1657-1659.
- Clement, M., Snell, Q., Walke, P., Posada, D., & Crandall, K. (2002, January). TCS: estimating gene genealogies. In Parallel and Distributed Processing Symposium, International (Vol. 2, pp. 7-pp). IEEE Computer Society.
- Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. Nature Methods, 9, 772.
- De Queiroz, K. (2007). Species concepts and species delimitation. Systematic biology, 56(6), 879-886.
- Delsuc, F., Brinkmann, H., Chourrout, D., & Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature, 439(7079), 965-968.
- Dobzhansky, T. (1950). Mendelian populations and their evolution. The American Naturalist, 84(819), 401-418.
- Dos Reis, M., & Yang, Z. (2019). Bayesian molecular clock dating using genome-scale datasets. Evolutionary genomics: Statistical and computational methods, 309-330.
- Dos Reis, M., Donoghue, P. C., & Yang, Z. (2016). Bayesian molecular clock dating of species divergences in the genomics era. Nature Reviews Genetics, 17(2), 71-80.
- Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F., & Douzery, E. J. (2003). Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Molecular biology and evolution, 20(2), 248-254.
- Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology, 7(1), 1-8.
- Drummond, A. J., Rambaut, A., & Xie, W. (2011). Bayesian Skyline Plot.
- Dunn, C. W., Giribet, G., Edgecombe, G. D., & Hejnol, A. (2014). Animal phylogeny and its evolutionary implications. Annual review of ecology, evolution, and systematics, 45, 371-395.
- Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. A., ... & Giribet, G. (2008). Broad taxon sampling improves resolution of the Animal Tree of Life in phylogenomic analyses. Nature, 452, 745-750.
- Eernisse, D. J., Albert, J. S., & Anderson, F. E. (1992). Annelida and Arthropoda are not sister taxa: a phylogenetic analysis of spiralian metazoan morphology. Systematic Biology, 41(3), 305-330.
- Ellstrand, N. C. (2014). Is gene flow the most important evolutionary force in plants?. American journal of botany, 101(5), 737-753.
- Emerson, B. C., & Hewitt, G. M. (2005). Phylogeography. Current biology, 15(10), R367-R371.
- Ezard, T., Fujisawa, T., & Barraclough, T. G. (2009). splits: SPecies' LImits by Threshold Statistics. R package version 1.0-14/r31.
- Felsenstein, J. (1973). Maximum-likelihood estimation of evolutionary trees from continuous characters. American journal of human genetics, 25(5), 471.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of molecular evolution, 17, 368-376.

- Felsenstein, J. (1985). Phylogenies and the comparative method. The American Naturalist, 125(1), 1-15.
- Field, K. G., Olsen, G. J., Lane, D. J., Giovannoni, S. J., Ghiselin, M. T., Raff, E. C., ... & Raff, R. A. (1988). Molecular phylogeny of the animal kingdom. Science, 239(4841), 748-753.
- Fisher, R. A. (1919). XV.—The correlation between relatives on the supposition of Mendelian inheritance. Earth and Environmental Science Transactions of the Royal Society of Edinburgh, 52(2), 399-433.
- Fitch, W. M., & Margoliash, E. (1967). Construction of phylogenetic trees: a method based on mutation distances as estimated from cytochrome c sequences is of general applicability. Science, 155(3760), 279-284.
- Garamszegi, L.Z., Gonzalez-Voyer, A. (2014). Working with the Tree of Life in Comparative Studies: How to Build and Tailor Phylogenies to Interspecific Datasets. In: Garamszegi, L. (eds) Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology. Springer, Berlin, Heidelberg.
- Garcia-Dias, R., Vieira, S., Pinaya, W. H. L., & Mechelli, A. (2020). Clustering analysis. In machine learning (pp. 227-247). Academic Press.
- Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology, 52(5), 696-704.
- Haeckel, E. (1866). Generelle Morphologie der Organismen: Allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformierte Descendenz-Theorie. Band 1: Allgemeine Anatomie. Band 2: Allgemeine Entwicklungsgeschichte. de Gruyter.
- Halanych, K. M. (2004). The new view of animal phylogeny. Annual Review of Ecology, Evolution, and Systematics, 35, 229-256.
- Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection. In Mathematical Proceedings of the Cambridge Philosophical Society (Vol. 23, No. 5, pp. 607-615). Cambridge University Press.
- Hardy, G. H. (1908). Mendelian proportions in a mixed population. Science, 28(706), 49-50.
- Harvey, P. H., May, R. M., & Nee, S. (1994). Phylogenies without fossils. Evolution, 48(3), 523-529.
- Hasegawa, M., Kishino, H., & Yano, T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution, 22, 160-174.
- Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G. W., Edgecombe, G. D., ... & Dunn, C. W. (2009). Assessing the root of bilaterian animals with scalable phylogenomic methods. Proceedings of the Royal Society B: Biological Sciences, 276(1677), 4261-4270.
- Hillis, D. M., & Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analyses. Journal of heredity, 83(3), 189-195.
- Hillis, D. M., Huelsenbeck, J. P., & Cunningham, C. W. (1994). Application and accuracy of molecular phylogenies. Science, 264(5159), 671-677.
- Ho, S. Y., & Shapiro, B. (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. Molecular ecology resources, 11(3), 423-434.
- Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17(8), 754-755.
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. Science, 294(5550), 2310-2314.
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution, 23(2), 254-267.
- Jenkin, F. (1867). The origin of species. North British Review, 46(92), 277-318.
- Jones, N. C., & Pevzner, P. A. (2004). An introduction to bioinformatics algorithms. MIT press.
- Jukes, T. H., & Cantor, C. R. (1969). Evolution of protein molecules. 21–132 Munro HN ed Mammalian protein metabolism Academic Press. New York.

- Kannan, L., & Wheeler, W. C. (2012). Maximum parsimony on phylogenetic networks. Algorithms for molecular biology, 7(1), 1-10.
- Karol, K. G., McCourt, R. M., Cimino, M. T., & Delwiche, C. F. (2001). The closest living relatives of land plants. Science, 294(5550), 2351-2353.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of molecular evolution, 16, 111-120.
- Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. Proceedings of the National Academy of Sciences, 78(1), 454-458.
- Knowlton, N. (1993). Sibling species in the sea. Annual review of ecology and systematics, 24(1), 189-216.
- Knowlton, N., & Weigt, L. A. (1998). New dates and new rates for divergence across the Isthmus of Panama. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1412), 2257-2263.
- Kocot, K. M., Cannon, J. T., Todt, C., Citarella, M. R., Kohn, A. B., Meyer, A., ... & Halanych, K. M. (2011). Phylogenomics reveals deep molluscan relationships. Nature, 477(7365), 452-456.
- Laporte, M., Bertolo, A., Berrebi, P., & Magnan, P. (2014). Detecting anthropogenic effects on a vulnerable species, the freshwater blenny (*Salaria fluviatilis*): the importance of considering key ecological variables. Ecological indicators, 36, 386-391.
- Lee, R. C. (1981). Clustering analysis and its applications. In Advances in Information Systems Science: Volume 8 (pp. 169-292). Boston, MA: Springer US.
- Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. Methods in ecology and evolution, 6(9), 1110-1116.
- Lemmon, E. M., & Lemmon, A. R. (2013). High-throughput genomic data in systematics and phylogenetics. Annual Review of Ecology, Evolution, and Systematics, 44, 99-121.
- Liddell, H. G., Scott, R., & Jones, H. S. (2023). A Greek-English Lexicon/compiled by Henry George Liddell and Robert Scott. La Trobe.
- Luo, A., Ling, C., Ho, S. Y., & Zhu, C. D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. Systematic Biology, 67(5), 830-846.
- Lutzoni, F., Pagel, M., & Reeb, V. (2001). Major fungal lineages are derived from lichen symbiotic ancestors. Nature, 411(6840), 937-940.
- Manolopoulou, I., Legarreta, L., Emerson, B. C., Brooks, S., & Tavaré, S. (2011). A Bayesian approach to phylogeographic clustering. Interface focus, 1(6), 909-921.
- Mayr, E. R. N. S. T. (1942). Systematics and theOrigin of Species. Columbia UniversityPress, New York.
- Michener, C. D. (1970). Diverse approaches to systematics. Evolutionary biology, 4, 1.
- Morrison, D. A. (2010). Using data-display networks for exploratory data analysis in phylogenetic studies. Molecular Biology and Evolution, 27(5), 1044-1057.
- Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. Methods in Ecology and Evolution, 3(4), 743-756.
- Murphy, W. J., Eizirik, E., O'Brien, S. J., Madsen, O., Scally, M., Douady, C. J., ... & Springer, M. S. (2001). Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science, 294(5550), 2348-2351.
- Nebot-Colomer, E., Vázquez-Luis, M., Boissin, E., Peyran, C., Deudero, S., & Planes, S. (2022a). Inferred family structure and dispersal patterns of a Critically Endangered species, *Pinna nobilis*, using molecular analyses: implications for conservation. Endangered Species Research, 49, 87-102.
- Nielsen, C., Scharff, N., & Eibye-Jacobsen, D. (1996). Cladistic analyses of the animal kingdom. Biological Journal of the Linnean Society, 57(4), 385-410.

- Nixon, K.C. Phylogeny. In Encyclopedia of Biodiversity, 2nd ed.; Levin, S.A., Ed.; Academic Press: Waltham, MA, USA, 2001; pp. 16–23.
- Oidtmann, B., Cerenius, L., Schmid, I., Hoffmann, R., & Söderhäll, K. (1999). Crayfish plague epizootics in Germany--classification of two German isolates of the crayfish plague fungus *Aphanomyces astaci* by random amplification of polymorphic DNA. Diseases of Aquatic Organisms, 35(3), 235-238.
- Okasha, S. (2006). Population genetics.
- Paradis, E. (2010). pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics, 26(3), 419-420.
- Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular ecology notes, 6(1), 288-295.
- Philippe, H., & Telford, M. J. (2006). Large-scale sequencing and the new animal phylogeny. Trends in Ecology & Evolution, 21(11), 614-620.
- Philippe, H., Lartillot, N., & Brinkmann, H. (2005). Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. Molecular biology and evolution, 22(5), 1246-1253.
- Pierce, B. A. (2016). Genetics. 2nd Edi., Population Genetics. 25, 731-753.
- Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic biology, 55(4), 595-609.
- Posada, D. (2003). Using MODELTEST and PAUP* to select a model of nucleotide substitution. Current protocols in bioinformatics, (1), 6-5.
- Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic biology, 53(5), 793-808.
- Posada, D., & Crandall, K. A. (2001). Selecting the best-fit model of nucleotide substitution. Systematic biology, 50(4), 580-601.
- Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. J. M. E. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular ecology, 21(8), 1864-1877.
- Raftery, A. E. (1996). Hypothesis testing and model selection Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics ed WR Gilks, S Richardson and DJ Spiegelhalter.
- Rambaut, A., & Drummond, A. J. (2009). Tracer version 1.5. 0. http://beast. bio. ed. ac. uk/software/tracer/.
- Retief, J. D. (1999). Phylogenetic analysis using PHYLIP. Bioinformatics methods and protocols, 243-258.
- Revell, L. J., Harmon, L. J., & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. Systematic biology, 57(4), 591-601.
- Rodriguez, F. J. L. O. J., Oliver, J. L., Marín, A., & Medina, J. R. (1990). The general stochastic model of nucleotide substitution. Journal of theoretical biology, 142(4), 485-501.
- Ronquist, F., Huelsenbeck, J. P., & van der Mark, P. (2005). MrBayes 3.1 Manual.
- Ronquist, F., van der Mark, P., & Huelsenbeck, J. P. (2009). Bayesian phylogenetic analysis using MrBayes. The phylogenetic handbook, 2(210-266), 723.
- Rosen, D.E. (1979). Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. Bull Am Mus Nat Hist, 162, 267-376.
- Roy, S. S., Dasgupta, R., & Bagchi, A. (2014). A review on phylogenetic analysis: a journey through modern era. Computational Molecular Bioscience, 4(03), 39.
- Russell 2010 - iGenetics : a molecular approach / Peter J. Russell. -- 3rd ed. pp.713
- Rutschmann, F. (2006). Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. Diversity and Distributions, 12(1), 35-48.
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution, 4(4), 406-425.

- Scarpa, F., Cossu, P., Lai, T., Sanna, D., Curini-Galletti, M., & Casu, M. (2016). Meiofaunal cryptic species challenge species delimitation: the case of the *Monocelis lineata* (Platyhelminthes: Proseriata) species complex. Contributions to Zoology, 85(2), 123-145.
- Scarpa, F., Cossu, P., Sanna, D., Lai, T., Casu, M., & Curini-Galletti, M. (2019). New insights on the genus *Otoplana* Du Plessis, 1889 (Platyhelminthes: Proseriata), with description of two new species from the Canary Islands. Marine Biodiversity, 49, 2075-2087.
- Scarpa, F., Sanna, D., Azzena, I., Cossu, P., & Casu, M. (2021). From dark to light and back again: is *Pinna nobilis*, the largest Mediterranean shellfish, on the brink of extinction? What about *Pinna nobilis*. Veterinaria, 70(1), 1-14.
- Scarpa, F., Sanna, D., Azzena, I., Mugetti, D., Cerruti, F., Hosseini, S., ... & Casu, M. (2020). Multiple non-species-specific pathogens possibly triggered the mass mortality in *Pinna nobilis*. Life, 10(10), 238.
- Schmidt, H. A., & von Haeseler, A. (2009). Phylogenetic inference using maximum likelihood methods. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing, 2.
- Schmidt, H. A., Strimmer, K., Vingron, M., & Von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics, 18(3), 502-504.
- Scholtz, G., Braband, A., Tolley, L., Reimann, A., Mittmann, B., Lukhaup, C., ... & Vogt, G. (2003). Parthenogenesis in an outsider crayfish. Nature, 421(6925), 806-806.
- Schram, F. R. (1991). Cladistic analysis of metazoan phyla and the placement of fossil problematica. In The early evolution of Metazoa and the significance of problematic taxa (pp. 35-46).
- Smith, S. A., Wilson, N. G., Goetz, F. E., Feehery, C., Andrade, S. C., Rouse, G. W., ... & Dunn, C. W. (2013). Correction: Corrigendum: Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature, 493(7434), 708-708.
- Sokal, R. R., & Crovello, T. J. (1970). The biological species concept: a critical evaluation. The American Naturalist, 104(936), 127-153.
- Souty-Grosset, C., Anastacio, P. M., Aquiloni, L., Banha, F., Choquer, J., Chucholl, C., & Tricarico, E. (2016). The red swamp crayfish *Procambarus clarkii* in Europe: impacts on aquatic ecosystems and human well-being. Limnologica, 58, 78-93.
- Stadler, T. (2008). Lineages-through-time plots of neutral models for speciation. Mathematical biosciences, 216(2), 163-171.
- Statzner, B., Peltret, O., & Tomanova, S. (2003). Crayfish as geomorphic agents and ecosystem engineers: effect of a biomass gradient on baseflow and flood-induced transport of gravel and sand in experimental streams. Freshwater Biology, 48(1), 147-163.
- Strimmer, K., & Von Haeseler, A. (1996). Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Molecular biology and evolution, 13(7), 964-969.
- Strimmer, K., & Von Haeseler, A. (1997). Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proceedings of the National Academy of Sciences, 94(13), 6815-6819.
- Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus evolution, 4(1), vey016.
- Tavaré, S. (1986). Some probabilistic and statistical problems on the analysis of DNA sequence. Lecture of Mathematics for Life Science, 17, 57.
- Tettamanzi, A. G. (2005). Algoritmi evolutivi: concetti ed applicazioni. Mondo digitale, 2005(1), 3-17.
- Warnock, R. C., Yang, Z., & Donoghue, P. C. (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. Proceedings of the Royal Society B: Biological Sciences, 284(1857), 20170227.

- Weinberg, W. (1908). Über den Nachweis der Vererbung beim Menschen. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg, Stuttgart 64: 369–382. [On the demonstration of inheritance in humans]. Translation by R. A. Jameson printed in D. L. Jameson (Ed.), (1977). Benchmark papers in genetics, Volume 8: Evolutionary genetics (pp. 115–125). Stroudsburg, PA: Dowden, Hutchinson & Ross.
- Wright, S. (1931). Evolution in Mendelian populations. Genetics, 16(2), 97.
- Wu, Q., James, S. A., Roberts, I. N., Moulton, V., & Huber, K. T. (2008). Exploring contradictory phylogenetic relationships in yeasts. FEMS yeast research, 8(4), 641-650.
- Yang, Z. (1994). Estimating the pattern of nucleotide substitution. Journal of molecular evolution, 39, 105-111.
- Zander, C.D. (1986). Blenniidae. In: Fishes of the North-eastern Atlantic and the Mediterranean= Poissons de l'Atlantique du Nord-est et de la Méditerranée. Whitehead, P. J. P. eds
- Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), 2869-2876.
- Zuckerkandl, E., Pauling, L. (1965). Molecules as documents of evolutionary history. Journal of theoretical biology, 8(2), 357-366.
- Zuur, A. F., Ieno, E. N., & Smith, G. M. (2007). Principal coordinate analysis and non-metric multidimensional scaling. Analysing ecological data, 259-264.

Chapter 2

Pinna nobilis

1. General description

Pinna nobilis (Linnaeus 1758), commonly known as fan mussel or the noble pen shell, is a long-lived species of large marine bivalve mollusc, endemic of the Mediterranean Sea, and belonging to the family Pinnidae (Mollusca: Bivalvia) (Butler et al. 1993; Katsanevakis 2005; Acarli et al. 2011; Sanna et al. 2013; Basso et al. 2015). The taxonomy of the family Pinnidae has undergone under several revision during the last decades. To date, three genera have been described: *Pinna, Atrina* and *Streptopinna* (Lemer et al. 2014). Due to its morphological and ecological characteristic, *Pinna nobilis* is considered a pteriomorphian bivalve, inhabiting the Mediterranean Sea since the Miocene era (Gomez-Alba 1998). This filter-feeder species is commonly found within *Posidonia oceanica* meadows where it lives semi-buried, anchoring itself to the substratum thanks to its byssus threads, which glue to pebbles, sand, small fragments of robust biodetritic material, as well as roots and rhizomes of *Posidonia oceanica* (Garcia-March 2005). The role that *Pinna nobilis* plays in the ecosystem is significant, as it actively filters considerable amounts of detritus contributing to improve the clarity of water (Trigos et al. 2014). Furthermore, it introduces a hard type of substrate in soft bottom offering a surface that can be inhabited by various benthic species, both plant and animal. Due to its economic and ecological value, *Pinna nobilis* has gained the attention of many researchers over time, leading to a rich body of literature (Basso et al. 2015).

1.1. Shell morphology and gaping activity

Pinna nobilis exhibits a distinctive habitus with some slight differences in shell morphology, as variations in curvature or width, that can be noticed comparing populations from different geographical regions (Garcia-March 2005). These differences can be influenced by local environmental conditions, including

hydrodynamics, substrate type, and seagrass species composition (Combelles et al. 1986). However, it's important to note that the overall morphology of *Pinna nobilis* remains relatively consistent throughout its range of distribution (Garcia-March 2005).

In general, the shell consists of two valves, relatively thin and fragile, which are connected by a hinge and characterized by a fan shape that can reach lengths of up to one meter thanks to the significant extensions of their mantle and shell, which likely result from their semi-infaunal habitat in soft substrates (Rosewater 1961; Basso et al. 2015; Donato et al. 2021). Juveniles are characterized by the presence of spines in the external surface of the valves, which erode and disappear with the growth of the individual (Garcia-March 2005). Internally, and only in the anterior part of the shell, is present a smooth and iridescent layer of nacre arranged in staggered rows which starts approximately at the point of insertion of the posterior adductor muscles (Garcia-March et al. 2008a; Basso et al. 2015).

Gaping (valves/shells activity) is linked to important physiological functions, including feeding, respiration, and metabolism (Garcia-March et al. 2008b). *Pinna nobilis* displays distinct patterns of gaping activity. The primary trends involve the shell being open during daylight hours and closed during the night. However, the shell may also open at night when the moon is present, and its illumination exceeds 50% (Garcia-March et al. 2008b). Moreover, storms characterized by intense hydrodynamics can induce changes in the gaping activity, resulting in a decrease in the shell's maximum open duration, encouraging the shell closure. Notably, when individuals are studied simultaneously, their behaviour displays synchronization, highlighting that the entire population responds uniformly to the same stimuli (Garcia-March et al. 2008b).

1.2. Byssus

Pinna nobilis is attached to the bottom in upright position thanks to its byssus threads, which can reach up to 20,000–30,000 filaments and fix to different substrates such as: sand, pebbles, mead, and matte of *Posidonia oceanica* (Garcia-March 2005; Basso et al. 2015; Diana et al. 2017). The high number of byssus filaments produced by the homonymous gland situated in the foot (see Figure 2), together with the attachment strategy, force *Pinna nobilis* to stay sessile and to constantly reinforced the fixation to the substrate (Cerruti 1938; Cerruti 1939). This surrounding, common to all the Pinnids,

has led to the developing of an anysomiarian condition (i.e., the anterior adductor muscle is reduced in respect to the posterior adductor muscle) (Yonge 1953).

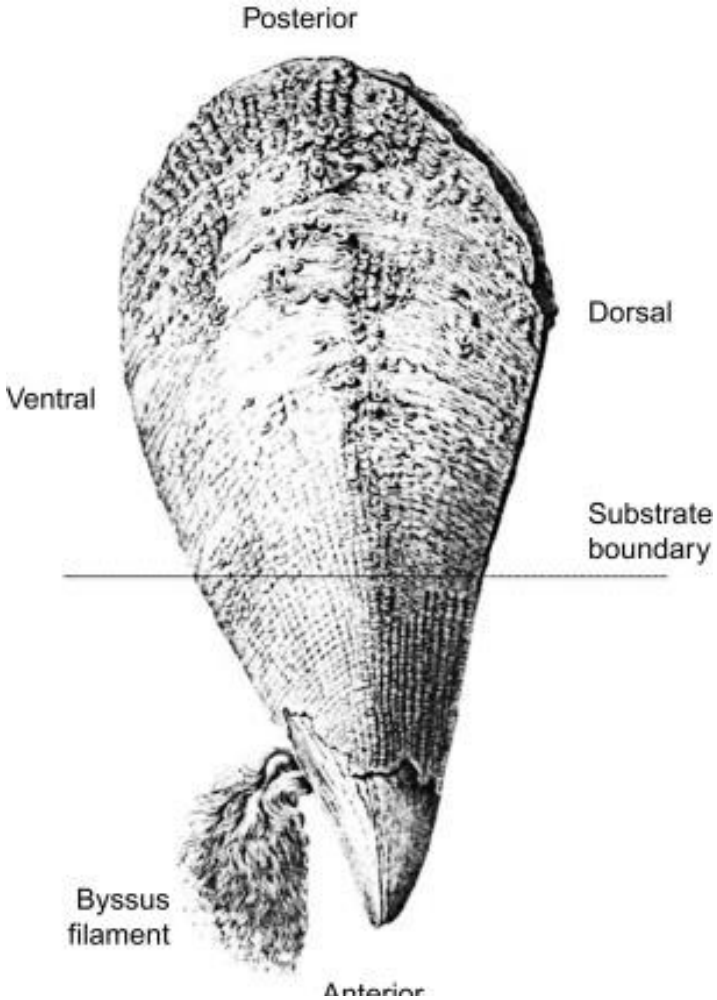


Figure 2. Typical representation of a fully grown *Pinna nobilis*, commonly seen in nature. It depicts the front portion of the shell gradually narrowing and buried in the sediment, firmly connected to the substrate through byssus threads. The illustration has been redrawn from Czihak and Dierl (1961) (Basso et al. 2015).

1.3. Ecology and feeding

Pinna nobilis lives in coastal water between 0.5 up to 60 meters, preferring soft bottom inhabit by mixed seagrass mead like *Posidonia oceanica* (see Figure 3), *Cymodocea nodosa* (see Figure 4), *Zostera marina* or *Zostera noltii* but, occasionally, some individuals have been spotted on bare sandy bottom or even on maërl beds (Zavodnik 1967; Zavodnik et al. 1991; Butler et al. 1993; Katsanevakis 2005; Basso et al. 2015; Scarpa et al. 2021).

Figure 3. Specimens of *Pinna nobilis* in *Posidonia oceanica* meadows. Photos by Fabio Scarpa.

Figure 4. Specimens of *Pinna nobilis* in *Cymodocea nodosa* meadows. Photos by Fabio Scarpa.

Pinna nobilis is considered a long-lived mollusc, with individuals that can reach up to 45-50 years old, as reported by Rouanet et al. (2015) in Port-Cros National Park (Provence, France). Nonetheless, our knowledge concerning the life expectancy of *Pinna nobilis* is still due to the challenging nature of developing an experimental approach that enables the evaluation of the species' maximum lifespan in undisturbed environments (Rouanet et al. 2015).

Pinna nobilis is a filter feeder organism, whose gills are composed by outer and inner layer that both move upwards and downwards. These filaments are intricately interconnected, giving rise to a densely compacted gill structure that create a cavity

which leads into a spacious square chamber and ultimately into the upper mantle cavity. Notably, the outer gill filaments are rich in glandular cells producing mucus, which covers the gill and forms a net for particle capture and filtration (Czihak & Dierl 1961).

Several studies have been conducted on the feeding habits of the species (Alomar et al. 2015; Cabanellas-Reboredo et al. 2009; Cabanellas-Reboredo et al. 2010; Kennedy et al. 2001; Davenport et al. 2011; Najdek et al. 2013; Trigos et al. 2014). *Pinna nobilis* appears to prefer selectively consuming detritus (constituting 95% of its ingested material), along with phytoplankton, micro- and mesozooplankton, and pollen grains. However, this dietary inclination could be influenced by specific geographical areas (Basso et al. 2015). In this context, Alomar et al. (2015) demonstrated that the diet of *Pinna nobilis* is more closely associated with the pelagic environment rather than the benthic habitat, with phytoplankton being the primary food source. Nevertheless, analysis of fatty acids in digestive gland and abductor muscle tissues indicated that smaller *Pinna nobilis* individuals are associated with a detrital food chain, characterized by saturated and branched fatty acids. In contrast, the diets of medium and large individuals contain a higher proportion of polyunsaturated fatty acids (Najdek et al. 2013). This finding suggests that smaller individuals primarily feed within the benthic boundary layer, where detritus concentrations are elevated (Najdek et al. 2013; Trigos et al. 2014).

1.4. Reproduction and recruitment

Pinna nobilis is a successive hermaphrodite with asynchronous maturation to prevent self-fertilization (De Gaulejac 1995). The timing and duration of the reproductive cycle are influenced by a combination of environmental and internal factors (Sastry 1979). However, we have limited information on larval mortality, juvenile stages, and dispersal capacity (Katsanekavis 2007). Sexual maturity is reached at 2 years of age and the sexual cycle is characterized by four main phases: (i) sexual quiescence from October to March; (ii) increased sexual activity and gametogenesis from March to June; (iii) alternation of spawning and rapid gametogenesis from June to August, and (iv) last partial maturation in September before returning to sexual rest (De Gaulejac 1995; Richardson et al. 1999).

Pinna nobilis dispersal phase occurs with the pelagic larvae, transported by the currents, whose life expectancy is about 5-10 days (Butler et al. 1993). This step represents a moment of potential vulnerability for the species, as the first stages of life (eggs and larvae) are particularly sensitive to environmental stressors (Przeslawski et al. 2005). For what concerning the recruitment of species, there is limited accessible data within the Pinnidae family using spat collectors (Narvaez et al. 2000; Beer & Southgate 2006). Notably, only a single study has attempted to model the recruitment impact of *Pinna nobilis* within a small marine protected area (Peharda & Vilibic 2008).

Recruitment is influenced by multiple factors, including rising seawater temperatures which lead to a reduction in reproductive output and an earlier onset of bivalve spawning in the spring (Philippart et al. 2003). Ceballos-Vazquez et al. (2000) highlighted a direct link between the seasonal reproductive cycle of the bivalve *Pinna rugosa*, water temperature and photoperiod. In this context, climate changes might affect *Pinna nobilis* recruitment, challenging the already vulnerable populations (Garcia-March et al. 2007).

1.5. Epibionts and commensals

Research on the epibiontic community of *Pinna nobilis* is limited. However, these few studies were capable to identified different species associated with fan mussel: molluscs dominated the epibiontic community, followed by annelids, crustaceans, ascidians, sponges, cnidarians, echinoderms, and bryozoans (Giacobbe 2002; Cosentino & Giacobbe 2008; Raboui et al. 2009; Addis et al. 2009).

Additionally, two crustaceans can live in association with *Pinna nobilis*: the shrimp *Pontonia pinnophylax* and the crab *Nepinnotheres pinnotheres*, but their relationship is not fully understood (Richardson et al. 1997; Raboui et al. 2008). It can be concluded that the taxonomic community linked to *Pinna nobilis* is synonym of high biodiversity, both in terms of species richness and taxonomic relationships. Moreover, the presence of *Pinna nobilis* enhance local spatial diversity and facilitate the colonization of benthic species in soft-bottom substrate (Basso et al. 2015).

2. Human exploitation over the course of centuries

Pinna nobilis has been exploited by humans for various purposes over the centuries. In the past, the inner layer of shells was used for ornamental and decorative purposes and valued for its aesthetic qualities in crafting jewellery, buttons, and other decorative items (Scarpa et al. 2021). Moreover, the distinct shape of its shell made it desirable for collectors, contributing to their overharvesting in some areas (Butler et al. 1993; Rabaoui et al. 2010). Furthermore, during the mid of XX century, a new widespread trend emerged: the abductor muscle of *Pinna nobilis* was employed for culinary applications, while the remaining body was used as bait for fishing purposes (Greenwald 1996; Scarpa et al. 2020).

Finally, it's worth noting that in the Southern part of Italy, particularly in Apulia (Taranto), and in the island of Sardinia (Sant'Antioco), existed a longstanding tradition of harvesting the byssus produced by *Pinna nobilis* for creating luxurious fabrics and textiles. This tradition dates to ancient times, with records of its existence tracing back to both the Roman and Greek periods (Centoducati et al. 2007; Scarpa et al. 2020).

Apart from these, the populations of *Pinna nobilis* were also severely impacted by indirect human activities, including boat anchoring, pollution, and habitat fragmentation (Vázquez-Luis et al. 2015; Öndes et al. 2020a). Therefore, the populations of *Pinna nobilis* in all the Mediterranean Sea started a demographic decline that escalated significantly by the late 1980s (Öndes et al. 2020b).

2.1. Protection regime and its effects over time

To invert the trend of demographic decline, at the begging of the 1990s *Pinna nobilis* was included in a full protection regime under the Annex IV of the EU Habitats Directive (European Council Directive 92/43/EEC) and Annex II of Barcelona Convention (SPA/BD Protocol 1995). Moreover, many countries have enacted its own legislative measures, establishing conservation protocols to address the historical exploitation and continuous threats facing *Pinna nobilis*. In Italy, within the context of the Marine Strategy Monitoring Program, the noble pen shell was identified as a species of particular interest among the Mediterranean species worthy of attention, as outlined in Article 11 of Legislative Decree 190/210 (Scarpa et al. 2020).

Despite the great commitment of the European community and single Nations in the protection of this species, some areas of the South-Western Mediterranean Sea manifest a poor adherence to these conservation policies and some individuals continue to be illegally harvested for consumption (Katsanevakis et al. 2011) or artisanal purposes (Kersting et al. 2019).

Nevertheless, this pattern of illegal actions does not represent the prevailing trend among the nations surrounding the Mediterranean Sea. Indeed, over the course of a few decades, the protective measures implemented led to a remarkable revival of the species (Scarpa et al. 2020) as testify by the study of Sanna et al. (2013, 2014) whose authors were engaged in an extensive sampling initiative which involved more than 100 individuals in many collection sites.

3. The outbreak of mass mortality events (MME)

Few years after the study of Sanna et al. (2013), in 2016 abnormally mortality events with mortality rates reaching up to 100% started to invest the populations of *Pinna nobilis* located in the centre and southernmost coasts of Spain, together with Balearic Islands (Vasquez-Luis et al. 2017). Histological analysis conducted on affected individuals revealed the presence of a haplosporidian-like parasite within the digestive gland (Vasquez-Luis et al. 2017; Darriba 2017).

Since this first warning, countless events of mass mortality (MME) have been reported, starting from the western coast of Mediterranean Sea involving Spain, France, Italy and reaching the centre and the eastern part affecting Croatia, Bosnia and Herzegovina, Greece, and Turkey (Vasquez-Luis et al. 2017; Catanese et al. 2018; Katsanevakis et al. 2019; Panarese et al. 2019; Cabanellas-Reboredo et al. 2019; Lattos et al. 2020; Šarić et al. 2020; Čelebičić et al. 2020; Künili et al. 2021).

During the initial stages of MME, it was frequently observed the presence of individuals still standing in upright position with a delay in gaping activity after stimulation (sick animal), with tissue collapsed at the bottom of the shell or even with empty shell colonized by other species (see Figure 5). In the first scenario, the delay in gaping activity made the individual vulnerable to predation, while in the second case, death occurred in a few days (Vasquez-Luis et al. 2017; Scarpa et al. 2020).

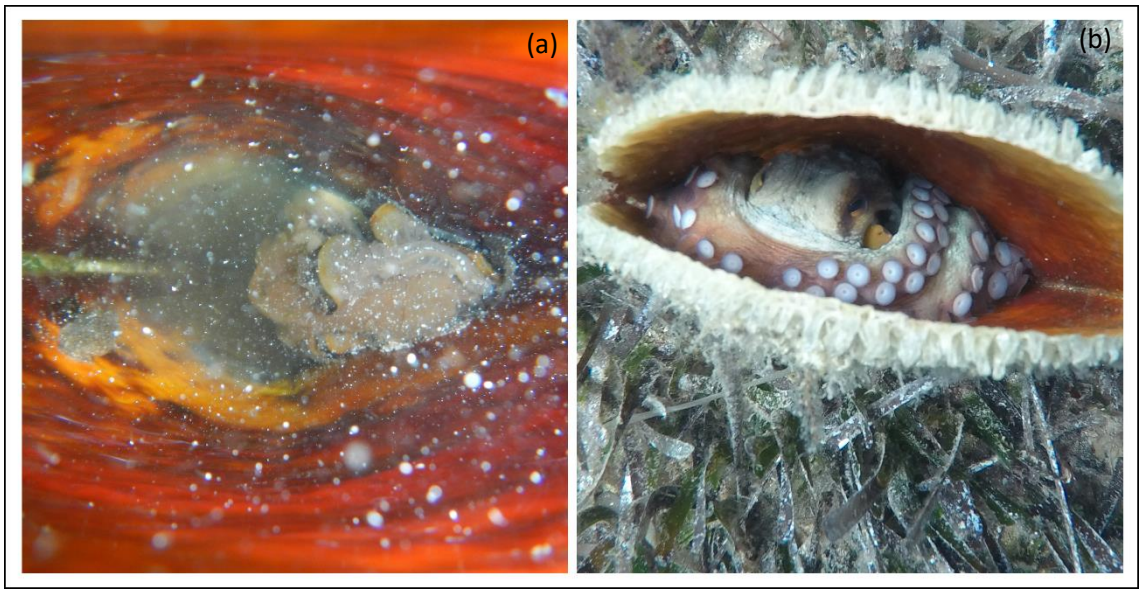


Figure 5. Specimen of *Pinna nobilis* with tissue collapsed at the shell's bottom (a). Empty shell of *Pinna nobilis* colonized by *Octopus vulgaris* (b). Photos by Fabio Scarpa.

Immediately after the report of Vasquez-Luis et al. (2017) several studies were conducted to identify the pathogens responsible for the mortality of *Pinna nobilis*. Protozoans, particularly haplosporidian endoparasites, have been widely acknowledged as the causative agents of various disease that led to mass mortality in numerous bivalve populations. Perhaps for this reason, the initial investigations into the causes of MME of *Pinna nobilis* were focused mainly on the search for protozoa (Vazquez-Luis et al. 2017; Daribba 2017; Catanese et al. 2018; Katsanevakis et al. 2019; López-Sanmartín et al. 2019; Panarese et al. 2019; Scarpa et al. 2020; Tiscar et al. 2022). Haplosporidian parasites have a life cycle consisting of two primary phases: (i) a unimultinucleate plasmodium stage and (ii) a sporulation phase that produces resilient spores which are released into the environment upon the host's death (Darriba 2017; Carella et. 2023). The morphological characteristics of haplosporidian-like parasite found in *Pinna nobilis* by Vasquez-Luis et al. (2017) were firstly reported by Darriba

(2017) who described the infection as systemic, affecting the connective tissue and causing disruptions in the digestive tubules.

Later, Catanese et al. (2018) characterized the new protozoan linked the MME as *Haplosporidium pinnae*, which the same authors initially believed to be host-specific for the fan mussel. The plasmodium stage of *Haplosporidium pinnae* is typically found within the hemocytes and the connective tissues of mantle and digestive gland while, sporulation takes place in the epithelium of the digestive gland tubules or, sporadically in the gut epithelium (Carella et al. 2023).

Although the initial research focused the attention on the identification of haplosporian-like parasites, subsequent studies have identified several bacterial species (*Mycobacterium* spp. and *Vibrio* spp.) alongside *Haplosporidium pinnae* as potential pathogens contributing to the MME of *Pinna nobilis* (Carella et al. 2019; Carella et al. 2020; Prado et al. 2020; Lattos et al. 2020, Šarić et al. 2020; Scarpa et al. 2020; Künili et al. 2021; Tiscar et al. 2022).

Moreover, Scarpa et al. (2020) highlight the finding of *Haplosporidium pinnae* in other bivalve species as *Mytilus galloprovincialis* and *Ruditapes decussatus* (sampled before 2016). These results overturn some scientific thoughts since: (i) *Haplosporidium pinnae* is not species-specific as previously hypothesized by Catanese et al. 2018; (ii) this protozoan was present in the Mediterranean Sea even before MME. Noteworthy, these same authors identified in the tissue of moribund or dead individuals of *Pinna nobilis* three taxonomic entities of *Mycobacterium* spp. never described before.

Furthermore, the presence of *Rhodococcus erythropolis* and *Perkinsus* sp. in association with *Mycobacterium* sp. was reported in few individuals of *Pinna nobilis* which showed signs of disease (Scarpa et al. 2020; Carella et al. 2020).

In the following years, several studies documented the presence of bacterial species in the tissues of dead or moribund individuals of fan mussel but, remarkably, these same tissues were not always infected also by *Haplosporidium pinnae* (Pavlinec et al. 2020; Scarpa et al. 2020; Grau et al. 2022). These results align with the theory firstly exposed by Scarpa et al. (2020), suggesting that MME should be attributed to a multifactorial disease rather than a single pathogen.

4. Genetic structure

Genetic diversity, also known as genetic variability, refers to variations in the genetic make-up of a species due to modifications in DNA sequences which lead to distinctions between individuals. Genetic variability is a fundamental concept in biology and genetics, and it plays a crucial role in the evolution and adaptation of species over time. A lack of genetic variability can make a population more vulnerable to diseases, environmental changes, or other challenges (Fisher 1930).

Several studies were conducted over the genetic variability of *Pinna nobilis* so far (Katsares et al. 2008; Rabaoui et al. 2011; Sanna et al. 2013; Wesselman et al. 2018; Gonzales-Wanguemert et al. 2019; Peyran et al. 2021; Peyran et al. 2022; Nebot-Colomer et al. 2022a).

4.1. *Pinna nobilis* genetic variability before 2016

Only a limited number of molecular studies were carried out over the genetic variability of the specie prior to 2016 (see Katsares et al. 2008; Rabaoui et al. 2011; Sanna et al. 2013), year in which *Pinna nobilis* populations experienced the mass mortality events (MME), with initial impact observed along the Spanish coasts (Vázquez-Luis et al. 2017).

The major aim of these papers was to infer on the genetic variability of *Pinna nobilis* after the strictly regime of protection started in the 90's. In particular, Katsares et al. (2008) explored the genetic variability within four populations situated in the Aegean Sea using two mitochondrial markers: the Cytochrome c Oxidase subunit I (COI) and the 16S ribosomal subunit genes. The study revealed a notable level of haplotypic diversity for the COI gene (except for the population sample from Corinthiakos Gulf), while the more conserved 16S gene exhibited a lower level of variability. Despite the substantial levels of haplotypic diversity observed in COI sequences, the genetic differentiation among the sampled populations of *Pinna nobilis* was minimal, resulting in an absence of genetic structuring for *Pinna nobilis* within the Aegean Sea.

Later, Rabaoui et al. (2011) depicted the genetic variability across five populations of Tunisia along the northern, eastern, and southern coasts. The results obtained indicate a decreasing gradient of genetic variability from Northern to Eastern coast. The authors

suggested that this decreasing gradient could be likely influenced by variations in the hydrodynamic conditions of the studied areas.

In 2013 Sanna et al. (2013) provided a comprehensive Mediterranean-scale assessment of the genetic variability of *Pinna nobilis*, encompassing the western Mediterranean, Ionian Sea, and Adriatic Sea marine ecoregions. To achieve this, they analysed samples from 25 sites, implementing a non-lethal sampling technique for the first time. In order to amplify the geographic area previously investigated, the authors employed the same mitochondrial markers (COI and 16S) used by Katsares et al. (2008) and Rabaoui et al. (2011), combining their obtained sequences with those from the aforementioned studies. The results obtained highlight significant levels of genetic variability across three distinct marine ecoregions: (1) the western Mediterranean and Ionian Sea; (2) the Adriatic Sea; and (3) the Aegean Sea and Tunisian coastal regions. From a conservation perspective, the populations within these three genetically distinct clusters could be considered as distinct management units.

Notably, the sequences of *Pinna nobilis* from studies by Katsares et al. (2008) and Sanna et al. (2011) were utilized by Lemer et al. (2014) in their research to elucidate the phylogeny of the family Pinnidae.

Few years later González-Wangüemert et al. (2015) isolated and developed ten microsatellite loci specifically designed for *Pinna nobilis*. These microsatellite markers, where tested in 76 individuals from two populations located in the Balearic Islands. The microsatellites revealed notable levels of polymorphism, holding significant promise for assessing the genetic diversity and connectivity patterns of *Pinna nobilis*, thereby facilitating the development of new conservation strategies. Three years later, González-Wangüemert et al. (2019) used these microsatellite markers to analysed different populations of *Pinna nobilis* sampled between 2010 and 2011 along the Spanish Mediterranean coast. The aim of the study was to employ a multidisciplinary approach, including population genetics and hydrodynamic modeling. The results revealed high genetic diversity (*i.e.,* variation between individuals belonging to the same population) and significant genetic differentiation (*i.e.,* variation among two or more populations) among post-larvae samples but not among adult populations. Notably, genic, and genotypic differentiation was observed for both adults and post-

larvae. Furthermore, this genetic connectivity pattern was found to align with marine currents and dispersal models.

In the same period, Wesselman et al. (2018), investigated the genetic structure of six populations of *Pinna nobilis* sampled during the summer of 2014 (using microsatellites from González-Wangüemert et al. 2015) in the Western Mediterranean Sea (France and Spain). Moreover, the authors chose to evaluate phylogeographical patterns in both Western and whole Mediterranean basin using the partial sequences of the mitochondrial marker cytochrome oxidase subunit I. To address this large geographical scale study, overlapping sequences from the literature were used. The results based on COI and microsatellite analyses indicated high genetic diversity and limited inter-population differentiation in *Pinna nobilis*.

As previously hypothesized by Sanna et al. (2013), Wesselman et al. (2018) in his research suggests that *Pinna nobilis* comprises a single lineage that experienced recent population expansion starting from a small, original population. Genetic differentiation was observed in the Venetian Lagoon, likely due to the semi-enclosed nature of the Adriatic Sea. However, no significant genetic break was detected between Western and Eastern Mediterranean populations.

4.2. *Pinna nobilis* genetic variability throughout and after MME

After the outbreak of MME, only two study have been performed to study the genetic variability of resilient populations of fan mussels, both involving France coast. In the summer of 2018, Peyran et al. (2021) conducted an analysis over *Pinna nobilis* populations along the Gulf of Lion coastline, stretching from the Spanish border to the Rhône estuary (France). The study involved 960 individuals and the combined use of Peyran et al. (2020) and González-Wangüemert et al. (2015) microsatellite markers.

The results revealed that the genetic diversity within all the surveyed sites was remarkably high but, surprisingly, no significant genetic differentiation was detected. Additionally, the percentages of closely related individuals were consistently low and similar across all locations. Consequently, the fan mussels inhabiting the entire surveyed coastline are part of a genetically uniform and interconnected population throughout the Gulf of Lion.

One year later, Peyran et al. (2022) presented an academic paper with the aim to study the population of *Pinna nobilis* inhabiting the bay of Peyrefite (Gulf of Lion, France), across multiple years: 2011, 2015 and 2018, focusing the attention on the recruitment processes. Only few individuals were resampled, indicating a rapid turnover during the years. The microsatellite's analyses revealed a consistently high level of genetic diversity in the bay of Peyrefite. Slight but statistically significant genetic differences between these years were highlighted. The research also demonstrated that fan mussels in this bay formed family clusters, with one family being notably dominant. This dominance could be due to a larger number of spawners in the parental population or better larval survival. Furthermore, approximately half of the sampled individuals were related to each other, suggesting that a significant portion of new recruits originated from the same source population year after year, consistently. No self-recruitment was detected between the individuals analysed.

5. Relict populations of *Pinna nobilis*

Few populations of *Pinna nobilis* survived to MME (Nebot-Colomer et al. 2022b). They are primarily situated in estuaries and isolated coastal lagoons within France, Italy, and Spain (García-March et al. 2020). These areas serve as the last reservoirs for *Pinna nobilis* in the western Mediterranean, while in the eastern Mediterranean, the Sea of Marmara plays a similar role (Cinar et al. 2021). Unfortunately, it's important to note that in the year 2023, some of these relic populations no longer exist.

The specific reasons why these populations have remained unaffected by recent declines are not yet understood. However, it's possible that differences in salinity and temperature, compared to open waters, play a role in their survival (Cabanellas-Reboredo et al. 2019; García-March et al. 2020; Prado et al. 2021). In this contest, Katsanevakis et al. (2022) suggested that both low temperatures and extreme salinities can restrict infection and mortality in fan mussels. However, it was noted that the precise optimal temperature and salinity ranges for the survival of *Pinna nobilis* need to be precisely determined. Nevertheless, it's important to note that coastal lagoons, despite being refuges for the species, face their own set of challenges: pollution, habitat destruction, and overexploitation (Barbier et al. 1997; Assessmen 2005; Pérez-Ruzafa et

al. 2006). Additionally, the limited water exchange, can make the lagoons vulnerable to issues like eutrophication, which can result in disruptive algal blooms and negative effects on the ecosystem and human health (Sunda et al. 2006).

5.1. Is it possible to recover from *Haplosporidium pinnae*? The case study of Mar Menor lagoon

Pinna nobilis population inhabiting the Mar Menor lagoon (Spain) has been extensively monitored in recent years. Nebot-Colomer et al. (2022b) conducted a large-scale monitoring program during the years 2014, 2017, and 2019. This program involved surveying an average of 32,000 m^2 of the lagoon each year to evaluate population densities, distributions, and size structures.

In 2014, a stable and widely distributed population was observed, with lower densities than expected. The population's shell size structure indicated a lack of juveniles and large individuals, potentially due to predation (mainly by the marine gastropods *Hexaplex trunculus*) and exposure to contaminants. Unfortunately, human activity like agriculture led to increased nutrient loads, causing eutrophication and a disruptive algal bloom in 2015. This event reduced light penetration and resulted in significant mortality of *Pinna nobilis* individuals, particularly below 2.5 meters depth (Nebot-Colomer et al. 2022b).

Despite initial indications of ecosystem recuperation in 2017, subsequent events from 2017 to 2019, placed the population in jeopardy once again:

- A significant proliferation of polychaetes serpulids belonging to the *Hydroides* genus, whose rapid expansion on some *Pinna nobilis* specimens led to a new episode of mortality because they blocked the shells and hindered their filtration capacity (2017) (Sandonnini et al. 2021a; Sandonnini et al. 2021b);
- Some *Pinna nobilis* individuals exhibited sub-lethal impacts, like reduced growth, because of invasive algae growth on their shells (2017) (Kersting & García-March 2017);
- Multiple rainfalls introduced extra nutrients to the ecosystem (2018) (Cortés-Melendreras et al. 2022);

- Another peak in chlorophyll and an extreme rainfall event (2019) (Nebot-Colomer et al. 2022b).

In summary, the combined effects of eutrophication and massive rainfall events have had a profound impact on the population's sustainability. The decrease in salinity and oxygen levels led to the mortality of the remaining benthic organisms situated below depths of 3–4 meters, including the scarce survivors of *Pinna nobilis* from 2017 (Nebot-Colomer et al. 2022b). Furthermore, the individuals inhabiting the shallower areas of the lagoon, close to the watercourses, might have been buried in sediment due to the substantial influx of water and sediments through these watercourses (Nebot-Colomer et al. 2022b). This has resulted in the survival of only a limited number of individuals (Nebot-Colomer et al. 2022b). Moreover, in 2019 Nebot-Colomer et al. (2022b) identified for the first time the presence of *Haplosporidium pinnae* in Mar Menor lagoon. This pathogen was previously found in the Ebro Delta and its presence raised questions about its origin. The authors suggest that the introduction of *Haplosporidium pinnae* to the Mar Menor lagoon might have been facilitated by a significant decrease in salinity and temperature caused by torrential rain in September 2019 (which lasted until June 2020) and that the subsequent environmental conditions facilitated the pathogen's spread and infection, likewise to what happens with *Perkinsus marinus* infection in oysters (Chu & Greene 1989). Whereas, according to Cortés-Melendreras et al. (2022), the reinfection of individuals observed in various areas of the lagoon in November 2019 by Nebot-Colomer et al. (2022b) is likely due to the activation of resistant forms of the parasite. These forms had been present in the Mar Menor since their initial entry in 2017, during the episode of low salinity.

Other scientists, like Cortés-Melendreras et al. (2022), propose that the presence of the parasite *Haplosporidium pinnae* is linked to fluctuations in salinity and temperature. They suggest that a decline in salinity, possibly due to Mediterranean water inflow through Las Encãnizadas and El Estacio channels during storms in January and February 2017, created an opportunity for *Haplosporidium pinnae* to enter the Mar Menor. This could explain the disappearance of *Pinna nobilis* near the entrances in 2017 (Cortés-Melendreras et al. 2022). However, the potential intrusion of the protozoan was seemingly regulated by the rapid rise in salinity within the lagoon, reaching levels (42

psu) that should hinder the pathogen's viability and its spread (Cabanellas-Reboredo et al. 2019; Cortés-Melendreras et al. 2022).

Thankfully, the molecular analysis conducted in June 2020 revealed the recovery of three individuals that had previously been infected in November 2019. However, based on the available data, the salinity values recorded during the sampling days in June 2020 (ranging from 38.6 to 39.35) were still within the infection range of *Haplosporidium pinnae*. Hence, factors beyond salinity, such as the interaction and combined effects of various environmental factors with the parasite, might explain this recovery. Indeed, harmful algal blooms and pollution have been noted to affect not just the host but also the reproduction and growth of several parasites (Soudant et al. 2013). Currently, according to Cortés-Melendreras et al. (2022), the estimated population of *Pinna nobilis* inhabiting the Mar Menor lagoon does not exceed the 1.500 individuals but, the persistent instability of the ecosystem endangers the survival of the species.

6. Hybridization between *Pinna nobilis* and *Pinna rudis*

Interactions between different species through hybridization play a dynamic role in the process of evolution (Anderson & Stebbins 1954). On one hand, it can erode the boundaries that separate species (Kleindorfer et al. 2014), while conversely, it can give rise to entirely novel species (Abbott et al. 2013). Hybrids can gain advantages from distinctive traits that empower them to flourish in new ecological environments (Nice et al. 2012) or through the fusion of traits that offer a notable competitive advantage over their parent species (Eberlein et al. 2019). The frequent instances of interspecific hybridization observed in various groups highlight its significant role in the evolutionary history of animals. Primarily, it enhances their ability to adapt to changes in their environment (Mallet 2005; Vázquez-Luis et al. 2021). In this context, Vázquez-Luis et al. (2021) documented the occurrence of several empty shells in Cabrera National Park, Spain, displaying characteristics of both *Pinna nobilis* and *Pinna rudis*. During their investigation, the researchers encountered three live individuals that displayed a combination of shell morphology and mantle coloration traits intermediate between *Pinna nobilis* and *Pinna rudis*. The use of microsatellite markers confirmed the presence of *Pinna nobilis* × *Pinna rudis* hybrids for the first time. Moreover, these

hybrids exhibited variations among themselves, and in certain instances, it was challenging for the scientists to differentiate them from the parent species.

There is a notable lack of studies focused on the biology and ecology of *Pinna rudis* (Cosentino & Giacobbe 2006; Cosentino & Giacobbe 2008; Wirtz & d'Acoz 2001; Wirtz & d'Acoz 2008; Dietl & Alexander 2005; Gvozdenović et al. 2019). Consequently, a significant knowledge gap exists, particularly regarding its reproductive aspects, which remain largely unexplored (Vázquez-Luis et al. 2021). However, the discovery of hybridization between *Pinna nobilis* and *Pinna rudis* suggests that these two species may share a reproductive window, possibly releasing gametes into the water column simultaneously or with partial overlap (May-July) (Vázquez-Luis et al. 2021). It's important to consider that Cabrera National Park presents a unique study area due to the notable abundance of both *Pinna rudis* and *Pinna nobilis* (prior MME). This high concentration of both species is uncommon in the Mediterranean region (Vázquez-Luis et al. 2021). Nevertheless, the classification of the three hybrids identified by Vázquez-Luis et al. (2021) remains uncertain. While genetic analysis using 28S sequencing suggests they are likely F1 hybrids, microsatellite analysis complicates matters by segregating them into two groups: (i) Hybrids 2 and 3, genetically closer to *Pinna rudis*, and (ii) hybrid 1, showing greater similarity to *Pinna nobilis*. Interestingly, hybrids 2 and 3 were collected in *Posidonia oceanica* meadows, typically dominated by *Pinna nobilis*, while hybrid 1 was found in a cave primarily inhabited by *Pinna rudis*. Given the available data, it was challenging for the researchers to definitively classify these hybrids as either F1 hybrids or recombined hybrids with purebred individuals since the two genetic markers do not provide a clear answer and do not align with habitat preferences. Consequently, whether these hybrids are fertile and capable of producing second-generation hybrids remains unknown (Vázquez-Luis et al. 2021). Noteworthy, Vázquez-Luis et al. (2021) suggested a crucial aspect to consider in the context of the ongoing pandemic: hybrids appear to exhibit resistance to *Haplosporidium pinnae*, like *Pinna rudis*, suggesting that potential genetic markers for resistance may exist within hybrid individuals.

7. Conservation and restocking programs

As a result of the mass mortality events, the conservation status of *Pinna nobilis* was reassessed, leading to its classification being updated from endangered to critically endangered (Kersting et al. 2019). In response to this critical situation, several Mediterranean organizations launched *ex situ* conservation programs, primarily focusing on captive breeding and reintroduction initiatives (Prado et al. 2020; Haberle et al. 2020; Kersting et al. 2020). Furthermore, there is a strong emphasis on safeguarding the remaining, albeit few, wild populations of *Pinna nobilis* in the Mediterranean Sea (Haberle et al. 2020). In such a context, the European Union has provided funding for two ongoing LIFE projects, dedicated to the conservation, and restocking of *Pinna nobilis*: LIFE PINNARCA and LIFE PINNA.

7.1. LIFE PINNARCA

LIFE PINNARCA is a European project dedicated to the conservation and restoration of *Pinna nobilis* populations in the Mediterranean Sea. Its primary objective is to prevent the extinction of *Pinna nobilis* through a comprehensive approach that includes raising public awareness, promoting collaboration among stakeholders, collecting data on the remaining populations, and implementing active recovery measures (https://www.lifepinnarca.com/the-project). The study encompasses various regions across the Mediterranean and involves multiple marine protected areas in Spain and Italy, including Punta Campanella and Cileno. It also covers lagoons and semi-enclosed bays like the Gulf of Kalloni in Greece, the Mar Menor in Spain, Brusc Lagoon in France, and the Ebro Bays in Spain.

7.2. LIFE PINNA

The project LIFE PINNA titled "Conservation and Re-stocking of the *Pinna nobilis* in the Western Mediterranean and Adriatic Sea," has a duration of four years and encompasses the protection and monitoring of the remaining *Pinna nobilis* populations, along with the reestablishment of this species in its native habitats. The main emphasis of the project centres on two key objectives: firstly, the supervision and safeguarding of the remaining *Pinna nobilis* individuals in the western Mediterranean and Adriatic Sea,

and secondly, the development of captive breeding methods for reintroducing disease-resistant individuals into designated areas (https://www.lifepinna.eu/en/the-project/). This ambitious project spans two country: Italy (Sardinia, Liguria, Friuli Venezia Giulia, and Tuscany) and Slovenia. Collaboration is a key aspect of the project, bringing together several organizations, which include:

- ARPAL (Regional Agency for the Protection of the Environment of Liguria);
- University of Sassari.
- Asinara National Park;
- National Institute of Biology (N.I.B.);
- Shoreline Cooperative Society;
- Triton Research S.r.l.;
- University of Genova;

In the initial phase of the project, comprehensive environmental and health assessments have been conducted in selected habitats located in northwestern Liguria, northwestern Sardinia, and the Upper Adriatic. Additionally, the surviving *Pinna nobilis* specimens underwent in-depth analyses, including genetic assessments, to understand the factors contributing to their disease resistance (Coupé et al. 2022). These findings will aid in identifying suitable candidates for breeding in laboratory settings. In this context, the University of Sassari is playing a key role in the project, led by Professor Marco Casu, from the Department of Veterinary Medicine, and under the supervision of Professor Daria Sanna from the Department of Biomedical Science. Indeed, the University of Sassari are overseeing the following actions of the project:

- A2 → "Molecular characterization of sentinel species in the putative pilot sites of restocking, and selection of the best candidates to be reproduced; by way of the most recent and powerful tools for bioinformatic analysis";
- C2 → "Molecular characterization of surviving individuals of *Pinna nobilis"*;
- D1 → "Monitoring of pathogens in restocking sites by using sentinel species".

In the second phase, the larvae of disease-resistant *Pinna nobilis*, primarily sourced from the Upper Adriatic, will be breed in captivity until they reach an appropriate age for transplantation. These young molluscs will then be reintroduced into four designated "pilot areas": three areas in Italy (Marine Protected Areas of Capo Mortola, Asinara, and Miramare), and one in Slovenia (Strunjan). The selection of the most

suitable sites has already been facilitated by molecular analyses conducted at regular intervals on different bivalves, principally *Mytilus galloprovincialis*, referred to as "sentinel" species, to provide early indications of potential threats such as protozoa, bacteria, and viruses.

The project's design allows for replication in different contexts by establishing comprehensive operational protocols that encompass all stages, ranging from monitoring and captive breeding to the ultimate successful reintroduction of the species into its native habitat.

7.2.1. "Sentinel" species

The practice of employing biological indicators to assess water quality has a rich historical background, as noted by Washington (1984) (Beeby 2001). We can consider "sentinel" species all the organisms exposed to environmental pollutants and/or pathogens, from which infection data can be regularly and systematically collected and analysed. These species play a crucial role in detecting and continuously monitoring a wide range of harmful environmental contaminants that pose threats to human health, animal species, and the overall stability of ecosystems. In this context, bivalve molluscs are frequently employed in studies evaluating the presence of contaminants and/or pathogens and their impacts on aquatic ecosystems (Taylor et al. 2017). Frequently, the concentrations of substances within these "sentinel" species are utilized to determine the accessibility of pollutants or pathogens to other organisms (Beeby 2001).

In the ongoing LIFE project, we have elected to employ organisms such as *Mytilus galloprovincialis*, *Ostrea* spp., and related species that, like *Pinna nobilis*, are filter-feeders and belong to the Bivalvia class.

8. Aims of the study

In the context of my doctoral project, the main aims of the following study are fourfold. Firstly, it aims to delineate the phylogeographic patterns and genetic diversity of *Pinna nobilis* before the Mass Mortality Event (MME), with the goal of facilitating the development of effective strategies for the recovery and rejuvenation of populations resembling those that became extinct.

Secondly, the study seeks to corroborate the hypothesis put forth by Sanna et al. (2013), which situates the genetic boundary between the western and eastern Mediterranean for *Pinna nobilis* to the east of the Sicilian strait. This goal will be achieved by incorporating data from three new Adriatic populations into the dataset.

Thirdly, the research aims to gain a deeper understanding of the evolutionary dynamics of *Pinna nobilis*. This understanding will enable us to make informed predictions regarding the species' potential for recovery following the genetic bottleneck resulting from the MME.

Lastly, the following paper aim to elucidate the key patterns in the phylogeny of *Pinna nobilis*, in order to develop a comprehensive understanding of how evolutionary processes have shaped the species' adaptation to the Mediterranean environment.

References

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J., Bierne, N., ... & Zinner, D. (2013). Hybridization and speciation. Journal of evolutionary biology, 26(2), 229-246.
- Acarli, S., Lok, A., Yigitkurt, S., & Palaz, M. (2011). Culture of fan mussel (*Pinna nobilis*, Linnaeus 1758) in relation to size on suspended culture system in Izmir Bay, Aegean Sea, Turkey. Kafkas Üniversitesi Veteriner Fakültesi Dergisi, 17(6).
- Addis, P., Secci, M., Brundu, G., Manunza, A., Corrias, S., & Cau, A. (2009). Density, size structure, shell orientation and epibiontic colonization of the fan mussel *Pinna nobilis* L. 1758 (Mollusca: Bivalvia) in three contrasting habitats in an estuarine area of Sardinia (W Mediterranean). Scientia Marina, 73(1), 143-152.
- Alomar, C., Vázquez-Luis, M., Magraner, K., Lozano, L., & Deudero, S. (2015). Evaluating stable isotopic signals in bivalve *Pinna nobilis* under different human pressures. Journal of experimental marine biology and ecology, 467, 77-86.
- Anderson, E., & Stebbins Jr, G. L. (1954). Hybridization as an evolutionary stimulus. Evolution, 378-388.
- Assessment, M. E. (2005). Ecosystems and human well-being: wetlands and water. World resources institute.
- Barbier, E. B., Acreman, M., & Knowler, D. (1997). Economic valuation of wetlands: a guide for policy makers and planners. Gland, Switzerland: Ramsar Convention Bureau.
- Basso, L., Vázquez-Luis, M., García-March, J. R., Deudero, S., Alvarez, E., Vicente, N., ... & Hendriks, I. E. (2015). The pen shell, *Pinna nobilis*: A review of population status and recommended research priorities in the Mediterranean Sea. Advances in marine biology, 71, 109-160.
- Beeby, A. (2001). What do sentinels stand for?. Environmental pollution, 112(2), 285-298.
- Beer, A. C., & Southgate, P. C. (2006). Spat collection, growth and meat yield of *Pinna bicolor* (Gmelin) in suspended culture in northern Australia. Aquaculture, 258(1-4), 424-429.
- Butler, A., Vicente, N., & de Gaulejac, B. (1993). Ecology of the pterioid bivalves *Pinna bicolor* Gmelin. Mar Life, 3(1-2), 37-45.
- Cabanellas-Reboredo, M., Blanco, A., Deudero, S., & Tejada, S. (2010). Effects of the invasive macroalga Lophocladia lallemandii on the diet and trophism of *Pinna nobilis* (Mollusca: Bivalvia) and its guests *Pontonia pinnophylax* and *Nepinnotheres pinnotheres* (Crustacea: Decapoda). Scientia marina, 74, 101-110.
- Cabanellas-Reboredo, M., Deudero, S., Alós, J., Valencia, J. M., March, D., Hendriks, I. E., & Álvarez, E. (2009). Recruitment of *Pinna nobilis* (Mollusca: Bivalvia) on artificial structures. Marine Biodiversity Records, 2, e126.
- Cabanellas-Reboredo, M., Vázquez-Luis, M., Mourre, B., Álvarez, E., Deudero, S., Amores, Á., ... & Hendriks, I. E. (2019). Tracking a mass mortality outbreak of pen shell *Pinna nobilis* populations: A collaborative effort of scientists and citizens. Scientific reports, 9(1), 13355.
- Carella, F., Aceto, S., Pollaro, F., Miccio, A., Iaria, C., Carrasco, N., ... & De Vico, G. (2019). A mycobacterial disease is associated with the silent mass mortality of the pen shell *Pinna nobilis* along the Tyrrhenian coastline of Italy. Scientific reports, 9(1), 2725.
- Carella, F., Antuofermo, E., Farina, S., Salati, F., Mandas, D., Prado, P., ... & De Vico, G. (2020). In the wake of the ongoing mass mortality events: co-occurrence of *Mycobacterium*, *Haplosporidium* and other pathogens in *Pinna nobilis* collected in Italy and Spain (Mediterranean Sea). Frontiers in marine science, 7, 48.
- Carella, F., Palić, D., Šarić, T., Župan, I., Gorgoglione, B., Prado, P., ... & Vico, G. D. (2023). Multipathogen infections and multifactorial pathogenesis involved in noble pen shell (*Pinna nobilis*) mass mortality events: Background and current pathologic approaches. Veterinary Pathology, 60(5), 560-577.
- Catanese, G., Grau, A., Valencia, J. M., Garcia-March, J. R., Vázquez-Luis, M., Alvarez, E., ... & Villalba, A. (2018). *Haplosporidium pinnae* sp. nov., a haplosporidan parasite associated

with mass mortalities of the fan mussel, *Pinna nobilis*, in the Western Mediterranean Sea. Journal of invertebrate pathology, 157, 9-24.
- Ceballos-Vazquez, B. P., Arellano-Martinez, M., Garcia-Dominguez, F. A., & Villalejo-Fuerte, M. (2000). Reproductive cycle of the rugose pen shell, *Pinna rugosa* Sowerby, 1835 (Mollusca). Journal of Shellfish Research, 19(1), 95-99.
- Čelebičić, M., Gradaščević, N., & Viteškić, V. (2020). Is *Pinna nobilis* l. critically endangered in Neum bay, Bosnia and Herzegovina?. Veterinaria, 69(3), 221-227.
- Centoducati, G., Tarsitano, E., Bottalico, A., Marvulli, M., Lai, O. R., & Crescenzo, G. (2007). Monitoring of the endangered *Pinna nobilis* Linne, 1758 in the Mar Grande of Taranto (Ionian sea, Italy). Environmental Monitoring and Assessment, 131, 339-347.
- Cerruti, A. (1938). Primi esperimenti di allevamento della Pinna "*Pinna nobilis* L." nel Mar Piccolo di Taranto. Tipografia Terme.
- Cerruti, A. (1939). Ulteriori notizie sull'allevamento della *Pinna nobilis* L. nel Mar Piccolo di Taranto. La Ricerca Scientifica, 18, 1110-1120.
- Chu, F. L. E., & Greene, K. H. (1989). Effect of temperature and salinity on in vitro culture of the oyster pathogen, *Perkinsus marinus* (Apicomplexa: Perkinsea). Journal of Invertebrate Pathology, 53(2), 260-268.
- Cinar, M. E., Bilecenoglu, M., YOKEŞ, M. B., & Güçlüsoy, H. A. R. U. N. (2021). The last fortress fell: mass mortality of *Pinna nobilis* in the Sea of Marmara. Mediterranean Marine Science, 22(3), 669-676.
- Combelles, S., Moreteau, J. C., & Vicente, N. (1986). Contribution a la connaissance de l'ecologie de *Pinna nobilis* L.(mollusque eulamellibranche). Sci. Rep. Port--Cros Nat. Park, 12, 29-43.
- Cortés-Melendreras, E., Gomariz-Castillo, F., Alonso-Sarría, F., Martín, F. J. G., Murcia, J., Canales-Cáceres, R., ... & Giménez-Casalduero, F. (2022). The relict population of *Pinna nobilis* in the Mar Menor is facing an uncertain future. Marine Pollution Bulletin, 185, 114376.
- Cosentino, A., & Giacobbe, S. (2006). Shell ornament in *Pinna nobilis* and *Pinna rudis* (Bivalvia: Pteriomorpha). Journal of Conchology, 39(2), 135.
- Cosentino, A., & Giacobbe, S. (2008). Aspects of epizoobiontic mollusc assemblages on *Pinna* shells. II. Does the Mediterranean P. nobilis represent an isle of biodiversity?. Cahiers de biologie marine, 49(2), 161.
- Czihak, G., & Dierl, W. (1961). *Pinna nobilis* L.: eine Präparationsanleitung. Fischer.
- Darriba, S. (2017). First haplosporidan parasite reported infecting a member of the Superfamily Pinnoidea (*Pinna nobilis*) during a mortality event in Alicante (Spain, Western Mediterranean). Journal of invertebrate pathology, 148, 14-19.
- Davenport, J., Ezgeta-Balić, D., Peharda, M., Skejić, S., Ninčević-Gladan, Ž., & Matijević, S. (2011). Size-differential feeding in *Pinna nobilis* L.(Mollusca: Bivalvia): exploitation of detritus, phytoplankton and zooplankton. Estuarine, Coastal and Shelf Science, 92(2), 246-254.
- De Gaulejac, B. (1995). Successive hermaphrodism with asynchronous maturation of *Pinna nobilis* (L.)(Bivalvia: Pterioidea). Comptes Rendus de l'Academie des Sciences-Serie III-Sciences de la Vie, 318(1), 99-104.
- Diana, A., Reguzzoni, M., Congiu, T., Rescigno, A., Sollai, F., & Raspanti, M. (2017). The byssus threads of *Pinna nobilis*: A histochemical and ultrastructural study. European Journal of Histochemistry: EJH, 61(4).
- Dietl, G. P., & Alexander, R. R. (2005). High frequency and severity of breakage-induced shell repair in western Atlantic Pinnidae (Bivalvia). Journal of Molluscan Studies, 71(3), 307-311.
- Donato, G., Vázquez-Luis, M., Nebot-Colomer, E., Lunetta, A., & Giacobbe, S. (2021). Noble fan-shell, *Pinna nobilis*, in Lake Faro (Sicily, Italy): Ineluctable decline or extreme opportunity?. Estuarine, Coastal and Shelf Science, 261, 107536.

- Eberlein, C., Hénault, M., Fijarczyk, A., Charron, G., Bouvier, M., Kohn, L. M., ... & Landry, C. R. (2019). Hybridization is a recurrent evolutionary stimulus in wild yeast speciation. Nature Communications, 10(1), 923.
- Fisher, R. A. (1930). The Natural Selection.
- Garcia-March, J. R. (2005). Aportaciones al conocimiento de la Biología de *Pinna nobilis* Linneo, 1758 (Mollusca: Bivalvia) en el litoral mediterráneo Ibérico. Universitat de Valencia (Spain).
- García-March, J. R., García-Carrascosa, A. M., Peña Cantero, A. L., & Wang, Y. G. (2007). Population structure, mortality and growth of *Pinna nobilis* Linnaeus, 1758 (Mollusca, Bivalvia) at different depths in Moraira bay (Alicante, Western Mediterranean). Marine Biology, 150, 861-871.
- García-March, J. R., Hernandis, S., Vázquez-Luis, M., Prado, P., Deudero, S., Vicente, N., & Tena-Medialdea, J. (2020). Age and growth of the endangered fan mussel *Pinna nobilis* in the western Mediterranean Sea. Marine environmental research, 153, 104795.
- García-March, J. R., Márquez-Aliaga, A., & Carter, J. G. (2008a). The duplivincular ligament of recent *Pinna nobilis* L., 1758: further evidence for pterineid ancestry of the Pinnoidea. Journal of Paleontology, 82(3), 621-627.
- García-March, J. R., Sanchís Solsona, M. Á., & García-Carrascosa, A. M. (2008b). Shell gaping behaviour of *Pinna nobilis* L., 1758: circadian and circalunar rhythms revealed by in situ monitoring. Marine Biology, 153, 689-698.
- Giacobbe, S. (2002). Epibiontic mollusc communities on *Pinna nobilis* L.(Bivalvia, Mollusca). Journal of Natural History, 36(12), 1385-1396.
- Gómez-Alba, J. (1988). Guía de Campo de los Fósiles de España y de Europa. Ediciones Omega.
- González-Wangüemert, M., Basso, L., Balau, A., Costa, J., Renault, L., Serrão, E. A., ... & Hendriks, I. E. (2019). Gene pool and connectivity patterns of *Pinna nobilis* in the Balearic Islands (Spain, Western Mediterranean Sea): Implications for its conservation through restocking. Aquatic Conservation: Marine and Freshwater Ecosystems, 29(2), 175-188.
- Gonzalez-Wanguemert, M., Costa, J., Basso, L., Duarte, C. M., Serrão, E., & Hendriks, I. (2015). Highly polymorphic microsatellite markers for the Mediterranean endemic fan mussel *Pinna nobilis*. Mediterranean Marine Science, 16(1), 31-35.
- Grau, A., Villalba, A., Navas, J. I., Hansjosten, B., Valencia, J. M., García-March, J. R., ... & Catanese, G. (2022). Wide-geographic and long-term analysis of the role of pathogens in the decline of *Pinna nobilis* to critically endangered species. Frontiers in Marine Science, 9, 666640.
- Greenwald, M. (2010). The cruising chef cookbook. Paradise Cay Publications.
- Gvozdenović, S., Mačić, V., Pešić, V., Nikolić, M., Peraš, I., & Mandić, M. (2019). Review on *Pinna rudis* (Linnaeus, 1758)(bivalvia: Pinnidae) presence in the Mediterranean. Poljoprivreda i Sumarstvo, 65(4), 115-126.
- Haberle, I., Marn, N., Geček, S., & Klanjšček, T. (2020). Dynamic energy budget of endemic and critically endangered bivalve *Pinna nobilis*: A mechanistic model for informed conservation. Ecological Modelling, 434, 109207.
- Katsanevakis, S. (2005). Population ecology of the endangered fan mussel *Pinna nobilis* in a marine lake. Endangered species research, 1, 51-59.
- Katsanevakis, S. (2007). Growth and mortality rates of the fan mussel *Pinna nobilis* in Lake Vouliagmeni (Korinthiakos Gulf, Greece): a generalized additive modelling approach. Marine Biology, 152, 1319-1331.
- Katsanevakis, S., Carella, F., Çınar, M. E., Cizmek, H., Jimenez, C., Kersting, D. K., & Moreno, D. (2022). The fan mussel *Pinna nobilis* on the brink of extinction in the Mediterranean. Imperiled: The Encyclopedia of Conservation: Volume 1-3.

- Katsanevakis, S., Poursanidis, D., Issaris, Y., Panou, A., Petza, D., Vassilopoulou, V., ... & Sini, M. (2011). " Protected" marine shelled molluscs: thriving in Greek seafood restaurants. Mediterranean Marine Science, 12(2), 429-438.
- Katsanevakis, S., Tsirintanis, K., Tsaparis, D., Doukas, D., Sini, M., Athanassopoulou, F., ... & Bakopoulos, V. (2019). The cryptogenic parasite *Haplosporidium pinnae* invades the Aegean Sea and causes the collapse of *Pinna nobilis* populations. Aquatic Invasions, 14(2).
- Katsares, V., Tsiora, A., Galinou-Mitsoudi, S., & Imsiridou, A. (2008). Genetic structure of the endangered species *Pinna nobilis* (Mollusca: Bivalvia) inferred from mtDNA sequences. Biologia, 63, 412-417.
- Kennedy, H., Richardson, C. A., Duarte, C. M., & Kennedy, D. P. (2001). Diet and association of *Pontonia pinnophylax* occurring in *Pinna nobilis*: insights from stable isotope analysis. Journal of the Marine Biological Association of the United Kingdom, 81(1), 177-178.
- Kersting, D. K., & García-March, J. R. (2017). Long-term assessment of recruitment, early stages and population dynamics of the endangered Mediterranean fan mussel *Pinna nobilis* in the Columbretes Islands (NW Mediterranean). Marine environmental research, 130, 282-292.
- Kersting, D. K., Vázquez-Luis, M., Mourre, B., Belkhamssa, F. Z., Álvarez, E., Bakran-Petricioli, T., ... & Hendriks, I. E. (2020). Recruitment disruption and the role of unaffected populations for potential recovery after the *Pinna nobilis* mass mortality event. Frontiers in Marine Science, 882.
- Kersting, D., Benabdi, M., Čižmek, H., Grau, A., Jimenez, C., Katsanevakis, S., ... & Otero Villanueva, M. (2019). *Pinna nobilis*. The IUCN red list of threatened species 2019: e. T160075998A160081499.
- Kleindorfer, S., O'Connor, J. A., Dudaniec, R. Y., Myers, S. A., Robertson, J., & Sulloway, F. J. (2014). Species collapse via hybridization in Darwin's tree finches. The American Naturalist, 183(3), 325-341.
- Künili, İ. E., Ertürk Gürkan, S., Aksu, A., Turgay, E., Çakir, F., Gürkan, M., & Altinağaç, U. (2021). Mass mortality in endangered fan mussels *Pinna nobilis* (Linnaeus 1758) caused by co-infection of *Haplosporidium pinnae* and multiple *Vibrio* infection in Çanakkale Strait, Turkey. Biomarkers, 26(5), 450-461.
- Lattos, A., Giantsis, I. A., Karagiannis, D., & Michaelidis, B. (2020). First detection of the invasive Haplosporidian and Mycobacteria parasites hosting the endangered bivalve *Pinna nobilis* in Thermaikos Gulf, North Greece. Marine environmental research, 155, 104889.
- Lemer, S., Buge, B., Bemis, A., & Giribet, G. (2014). First molecular phylogeny of the circumtropical bivalve family Pinnidae (Mollusca, Bivalvia): Evidence for high levels of cryptic species diversity. Molecular Phylogenetics and Evolution, 75, 11-23.
- López-Sanmartín, M., Catanese, G., Grau, A., Valencia, J. M., García-March, J. R., & Navas, J. I. (2019). Real-Time PCR based test for the early diagnosis of *Haplosporidium pinnae* affecting fan mussel *Pinna nobilis*. PloS one, 14(2), e0212028.
- Mallet, J. (2005). Hybridization as an invasion of the genome. Trends in ecology & evolution, 20(5), 229-237.
- Najdek, M., Blažina, M., Ezgeta-Balić, D., & Peharda, M. (2013). Diets of fan shells (*Pinna nobilis*) of different sizes: fatty acid profiling of digestive gland and adductor muscle. Marine Biology, 160, 921-930.
- Narvaez, N., Lodeiros, C., Freites, L., Nunez, M., Pico, D., & Prieto, A. (2000). Abundance and growth of *Pinna carnea* (Mytiloida: Pinnacea) juveniles in suspended-frame culture. Revista de biologia tropical, 48(4), 785-797.
- Nebot-Colomer, E., Álvarez, E., Belando, M. D., Deudero, S., Catanese, G., Bernardeau-Esteller, J., ... & Vázquez-Luis, M. (2022b). Living under threat: Will one of the last *Pinna nobilis* populations be able to survive?. Aquatic Conservation: Marine and Freshwater Ecosystems, 32(1), 1-13.

- Nebot-Colomer, E., Vázquez-Luis, M., Boissin, E., Peyran, C., Deudero, S., & Planes, S. (2022a). Inferred family structure and dispersal patterns of a Critically Endangered species, *Pinna nobilis*, using molecular analyses: implications for conservation. Endangered Species Research, 49, 87-102.
- Nice, C. C., Gompert, Z., Fordyce, J. A., Forister, M. L., Lucas, L. K., & Buerkle, C. A. (2013). Hybrid speciation and independent evolution in lineages of alpine butterflies. Evolution, 67(4), 1055-1068.
- Öndes, F., Alan, V., Akçalı, B., & Güçlüsoy, H. (2020b). Mass mortality of the fan mussel, *Pinna nobilis* in Turkey (eastern Mediterranean). Marine ecology, 41(5), e12607.
- Öndes, F., Kaiser, M. J., & Güçlüsoy, H. (2020a). Human impacts on the endangered fan mussel, *Pinna nobilis*. Aquatic conservation: Marine and freshwater ecosystems, 30(1), 31-41.
- Panarese, R., Tedesco, P., Chimienti, G., Latrofa, M. S., Quaglio, F., Passantino, G., ... & Otranto, D. (2019). *Haplosporidium pinnae* associated with mass mortality in endangered *Pinna nobilis* (Linnaeus 1758) fan mussels. Journal of invertebrate pathology, 164, 32-37.
- Pavlinec, Ž., Zupičić, I. G., Oraić, D., Petani, B., Mustać, B., Mihaljević, Ž., ... & Zrnčić, S. (2020). Assessment of predominant bacteria in noble pen shell (*Pinna nobilis*) collected in the Eastern Adriatic Sea. Environmental Monitoring and Assessment, 192, 1-10.
- Peharda, M., & Vilibić, I. (2008). Modelling the recruitment effect in a small marine protected area: the example of saltwater lakes on the Island of Mljet (Adriatic Sea). Acta Adriatica: International Journal of Marine Sciences, 49(1), 25-35.
- Pérez-Ruzafa, A., Garcıa-Charton, J. A., Barcala, E., & Marcos, C. (2006). Changes in benthic fish assemblages as a consequence of coastal works in a coastal lagoon: the Mar Menor (Spain, Western Mediterranean). Marine pollution bulletin, 53(1-4), 107-120.
- Peyran, C., Boissin, E., Morage, T., Nebot-Colomer, E., Iwankow, G., & Planes, S. (2021). Genetic homogeneity of the critically endangered fan mussel, *Pinna nobilis*, throughout lagoons of the Gulf of Lion (North-Western Mediterranean Sea). Scientific Reports, 11(1), 7805.
- Peyran, C., Boissin, E., Morage, T., Nebot-Colomer, E., Iwankow, G., & Planes, S. (2022). Investigating population dynamics from parentage analysis in the highly endangered fan mussel *Pinna nobilis*. Ecology and Evolution, 12(1), e8482.
- Peyran, C., Planes, S., Tolou, N., Iwankow, G., & Boissin, E. (2020). Development of 26 highly polymorphic microsatellite markers for the highly endangered fan mussel *Pinna nobilis* and cross-species amplification. Molecular Biology Reports, 47, 2551-2559.
- Philippart, C. J., van Aken, H. M., Beukema, J. J., Bos, O. G., Cadée, G. C., & Dekker, R. (2003). Climate-related changes in recruitment of the bivalve *Macoma balthica*. Limnology and Oceanography, 48(6), 2171-2185.
- Prado, P., Carrasco, N., Catanese, G., Grau, A., Cabanes, P., Carella, F., ... & Andree, K. B. (2020). Presence of *Vibrio mediterranei* associated to major mortality in stabled individuals of *Pinna nobilis* L. Aquaculture, 519, 734899.
- Prado, P., Grau, A., Catanese, G., Cabanes, P., Carella, F., Fernández-Tejedor, M., ... & García-March, J. R. (2021). *Pinna nobilis* in suboptimal environments are more tolerant to disease but more vulnerable to severe weather phenomena. Marine Environmental Research, 163, 105220.
- Przeslawski, R., Davis, A. R., & Benkendorff, K. (2005). Synergistic effects associated with climate change and the development of rocky shore molluscs. Global Change Biology, 11(3), 515-522.
- Rabaoui, L., Mejri, R., Tlig-Zouari, S., Bahri, L., Ben Hassine, O. K., & Tsigenopoulos, C. S. (2011). Genetic variation among populations of the endangered fan mussel *Pinna nobilis* (Mollusca: Bivalvia) along the Tunisian coastline. Hydrobiologia, 678, 99-111.
- Rabaoui, L., Tlig-Zouari, S., Cosentino, A., & Hassine, O. K. B. (2009). Associated fauna of the fan shell *Pinna nobilis* (Mollusca: Bivalvia) in the northern and eastern Tunisian coasts. Scientia Marina, 73(1), 129-141.

- Rabaoui, L., Tlig-Zouari, S., Katsanevakis, S., & Ben Hassine, O. K. (2010). Modelling population density of *Pinna nobilis* (Bivalvia) on the eastern and southeastern coast of Tunisia. Journal of Molluscan Studies, 76(4), 340-347.
- Rabaoui, L., Zouari, S. T., & Hassine, O. K. B. (2008). Two species of Crustacea (Decapoda) associated with the fan mussel, *Pinna nobilis* Linnaeus, 1758 (Mollusca, Bivalvia). Crustaceana, 433-446.
- Richardson, C. A., Kennedy, H., Duarte, C. M., & Proud, S. V. (1997). The occurrence of *Pontonia pinnophylax* (Decapoda: Natantia: Pontoninae) in *Pinna nobilis* (Mollusca: Bivalvia: Pinnidae) from the Mediterranean. Journal of the Marine Biological Association of the United Kingdom, 77(4), 1227-1230.
- Richardson, C. A., Kennedy, H., Duarte, C. M., Kennedy, D. P., & Proud, S. V. (1999). Age and growth of the fan mussel *Pinna nobilis* from south-east Spanish Mediterranean seagrass (*Posidonia oceanica*) meadows. Marine Biology, 133, 205-212.
- Rosewater, J. (1961). The family PINNIDAE in the Indo-Pacific. Indo-Pacific Mollusca, 1(4), 175-226.
- Rouanet, E., Trigos, S., & Vicente, N. (2015). From youth to death of old age: the 50-year story of a *Pinna nobilis* fan mussel population at Port-Cros Island (Port-Cros National Park, Provence, Mediterranean Sea). Sci. Rep. Port-Cros Natl. Park, 29, 209-222.
- Sandonnini, J., Del Pilar Ruso, Y., Cortés Melendreras, E., & Gimenez Casalduero, F. (2021b). Massive aggregations of Serpulidae associated with eutrophication of the Mar Menor, southeast Iberian Peninsula. Frontiers in Marine Science, 7, 531726.
- Sandonnini, J., Ruso, Y. D. P., Melendreras, E. C., Barberá, C., Hendriks, I. E., Kersting, D. K., & Casalduero, F. G. (2021a). The emergent fouling population after severe eutrophication in the Mar Menor coastal lagoon. Regional Studies in Marine Science, 44, 101720.
- Sanna, D., Cossu, P., Dedola, G. L., Scarpa, F., Maltagliati, F., Castelli, A., ... & Casu, M. (2013). Mitochondrial DNA reveals genetic structuring of *Pinna nobilis* across the Mediterranean Sea. PLoS One, 8(6), e67372.
- Sanna, D., Dedola, G. L., Scarpa, F., Lai, T., Cossu, P., Caronni, S., ... & Casu, M. (2011). Preliminary data on the genetic variability of the fan mussel *Pinna nobilis* in the Northern Sardinia= Dati preliminari sulla variabilità genetica del mollusco bivalve *Pinna nobilis* nella Sardegna settentrionale. Biologia marina mediterranea, 18(1), 286-287.
- Sanna, D., Dedola, G. L., Scarpa, F., Lai, T., Cossu, P., Curini-Galletti, M., ... & Casu, M. (2014). New mitochondrial and nuclear primers for the Mediterranean marine bivalve *Pinna nobilis*. Mediterranean Marine Science, 15(2), 416-422.
- Šarić, T., Župan, I., Aceto, S., Villari, G., Palić, D., De Vico, G., & Carella, F. (2020). Epidemiology of noble pen shell (*Pinna nobilis* l. 1758) mass mortality events in adriatic sea is characterised with rapid spreading and acute disease progression. Pathogens, 9(10), 776.
- Sastry, A. N., Giese, A. C., & Pearse, J. S. (1979). Pelecypoda (excluding ostreidae). Reproduction of marine invertebrates, 5, 113-292.
- Scarpa, F., Sanna, D., Azzena, I., Cossu, P., & Casu, M. (2021). From dark to light and back again: is *Pinna nobilis*, the largest Mediterranean shellfish, on the brink of extinction? What about *Pinna nobilis*. Veterinaria, 70(1), 1-14.
- Scarpa, F., Sanna, D., Azzena, I., Mugetti, D., Cerruti, F., Hosseini, S., ... & Casu, M. (2020). Multiple non-species-specific pathogens possibly triggered the mass mortality in *Pinna nobilis*. Life, 10(10), 238.
- Soudant, P., Chu, F. L. E., & Volety, A. (2013). Host–parasite interactions: Marine bivalve molluscs and protozoan parasites, *Perkinsus* species. Journal of invertebrate pathology, 114(2), 196-216.
- Sunda, W. G., Graneli, E., & Gobler, C. J. (2006). Positive feedback and the development and persistence of ecosystem disruptive algal blooms 1. Journal of Phycology, 42(5), 963-974.

- Taylor, A. M., Edge, K. J., Ubrihien, R. P., & Maher, W. A. (2017). The freshwater bivalve *Corbicula australis* as a sentinel species for metal toxicity assessment: an in situ case study integrating chemical and biomarker analyses. Environmental toxicology and chemistry, 36(3), 709-719.
- Tiscar, P. G., Rubino, F., Paoletti, B., Di Francesco, C. E., Mosca, F., Della Salda, L., ... & Fanelli, G. (2022). New insights about *Haplosporidium pinnae* and the pen shell *Pinna nobilis* mass mortality events. Journal of Invertebrate Pathology, 190, 107735.
- Trigos, S., García-March, J. R., Vicente, N., Tena, J., & Torres, J. (2014). Utilization of muddy detritus as organic matter source by the fan mussel *Pinna nobilis*. Mediterranean Marine Science, 667-674.
- Vázquez-Luis, M., Álvarez, E., Barrajón, A., García-March, J. R., Grau, A., Hendriks, I. E., ... & Deudero, S. (2017). SOS *Pinna nobilis*: a mass mortality event in western Mediterranean Sea. Frontiers in Marine Science, 4, 220.
- Vázquez-Luis, M., Borg, J. A., Morell, C., Banach-Esteve, G., & Deudero, S. (2015). Influence of boat anchoring on *Pinna nobilis*: a field experiment using mimic units. Marine and Freshwater Research, 66(9), 786-794.
- Vázquez-Luis, M., Nebot-Colomer, E., Deudero, S., Planes, S., & Boissin, E. (2021). Natural hybridization between pen shell species: *Pinna rudis* and the critically endangered *Pinna nobilis* may explain parasite resistance in *P. nobilis*. Molecular Biology Reports, 48(1), 997-1004.
- Wesselmann, M., González-Wangüemert, M., Serrão, E. A., Engelen, A. H., Renault, L., García-March, J. R., ... & Hendriks, I. E. (2018). Genetic and oceanographic tools reveal high population connectivity and diversity in the endangered pen shell *Pinna nobilis*. Scientific reports, 8(1), 4770.
- Wirtz, P., & d'Acoz, U. (2008). Crustaceans associated with Cnidaria, Bivalvia, Echinoidea and pisces at São Tomé and Príncipe islands.
- Wirtz, P., & d'Udekem-d'Acoz, C. (2001). Decapoda from Antipatharia, Gorgonaria and Bivalvia at the Cape Verde Islands. Helgoland Marine Research, 55, 112-115.
- Yonge, M. (1953). Form and habit in *Pinna carnea* Gmelin. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 237(648), 335-374.
- Zavodnik, D. (1967). Contribution to the ecology of *Pinna nobilis* L.(moll. hivalvia) in the northern Adriatic. JAZU.
- Zavodnik, D., Hrs-Brenko, M., & Legac, M. (1991). Synopsis on the fan shell *Pinna nobilis* L. in the eastern Adriatic Sea. Les espèces marines à protéger en Méditerranée, 169-178.

animals

MDPI

Article

Reconstructing the Evolutionary History of *Pinna nobilis*: New Genetic Signals from the Past of a Species on the Brink of Extinction

Daria Sanna [1,*,†], Ilenia Azzena [1,†], Chiara Locci [1,2], Pavel Ankon [3], Petar Kružić [3], Chiara Manfrin [4], Alberto Pallavicini [4], Saul Ciriaco [5,6], Marco Segarich [6], Edoardo Batistini [6], Fabio Scarpa [1] and Marco Casu [2]

[1] Department of Biomedical Sciences, University of Sassari, Viale San Pietro 43b, 07100 Sassari, Italy; iazzena@uniss.it (I.A.); c.locci3@phd.uniss.it (C.L.); fscarpa@uniss.it (F.S.)
[2] Department of Veterinary Medicine, University of Sassari, Via Vienna 2, 07100 Sassari, Italy; marcasu@uniss.it
[3] Department of Biology, Faculty of Science, University of Zagreb, Horvatovac 102a, 10000 Zagreb, Croatia; pavel.ankon@gmail.com (P.A.); petar.kruzic@biol.pmf.hr (P.K.)
[4] Department of Life Sciences, University of Trieste, Via L. Giorgieri 5, 34127 Trieste, Italy; cmanfrin@units.it (C.M.); pallavic@units.it (A.P.)
[5] WWF AMP Miramare, Via Beirut 2/4, 34151 Trieste, Italy; saul@ampmiramare.it
[6] Shoreline Soc. Coop., AREA Science Park, Padriciano 99, 34149 Trieste, Italy; marco.segarich@shoreline.it (M.S.); edoardo.batistini@shoreline.it (E.B.)
* Correspondence: darsanna@uniss.it
† These authors contributed equally to this study.

Simple Summary: *Pinna nobilis*, a species of marine shellfish living in the Mediterranean Sea, is at a high risk of extinction due to a not-entirely-known disease that started affecting its populations in 2016. In this paper, we reported the main traits of its evolutionary history to understand how this species evolved over time and space from the moment its ancestor entered the Mediterranean. To achieve this goal, we analysed a total of 469 sequences from all over the Mediterranean Sea. Our research showed that *P. nobilis* evolved from its ancestor about 2.5 million years ago, following a rapid and catastrophic entry of waters from the Atlantic Ocean that pushed the *P. nobilis* ancestor into the Mediterranean around 5.3 million years ago. Our results also suggest that the central part of the western Mediterranean was the first marine area where this species settled and, later on, it spread to the Adriatic and the eastern part of the basin. This information is of twofold importance, as it helps us to understand how this species adapted to the Mediterranean over time and may be the basis of present and future restocking plans which want to take into consideration the reconstruction of pre-existing genetic variability.

Abstract: *Pinna nobilis*, commonly known as the noble pen shell, is a marine bivalve endemic to the Mediterranean Sea. Unfortunately, due to a multifactorial disease that began affecting its populations in 2016, the species is currently facing the threat of extinction. To gain insights into the evolutionary history of *P. nobilis* before the mass mortality event (MME), and to obtain a comprehensive understanding of how evolutionary processes led to the adaptation of the species into the Mediterranean Sea, phylogenetic and phylogeographic analyses were carried out. The dataset analysed includes 469 sequences of COI gene fragment both from GenBank and the present study (100). The analysis performed evidenced that *P. nobilis* diverged about 2.5 mya, after the entrance of its ancestor into the Mediterranean Sea following the Zanclean flood (5.33 mya). Moreover, our results suggest that the starting point of colonisation was the central part of the western Mediterranean basin, with the eastern basin being populated subsequently. From a conservational viewpoint, these results provide important hints for present and future restocking plans, helping to reconstruct the pre-existing genetic variability in sites where the species became extinct.

Keywords: fan mussel; mtDNA; cytochrome c oxidase subunit I; evolution; molecular dating

Citation: Sanna, D.; Azzena, I.; Locci, C.; Ankon, P.; Kružić, P.; Manfrin, C.; Pallavicini, A.; Ciriaco, S.; Segarich, M.; Batistini, E.; et al. Reconstructing the Evolutionary History of *Pinna nobilis*: New Genetic Signals from the Past of a Species on the Brink of Extinction. *Animals* **2024**, *14*, 114. https://doi.org/10.3390/ani14010114

Academic Editor: Fuhua Li

Received: 15 November 2023
Revised: 20 December 2023
Accepted: 26 December 2023
Published: 28 December 2023

1. Introduction

Pinna nobilis Linnaeus, 1758, commonly known as fan mussel or noble pen shell, is a long-lived, large species of marine bivalve endemic to the Mediterranean Sea and belonging to the family Pinnidae (Mollusca: Bivalvia) [1]. Over the past few decades, the taxonomy of the family Pinnidae Leach, 1819, has undergone several revisions. To date, two still-living genera belong to this taxon: *Pinna* Linnaeus, 1758, and *Atrina* Gray, 1847. Indeed, Lemer et al. [2] made the most recent change to the taxonomic status within this family, proposing a new status for the genus *Streptopinna* E. von Martens, 1880, and now considering it as a subgenus (status nov.) of *Pinna*.

Due to its morphological and ecological characteristics, *P. nobilis* is a pteriomorphian bivalve (infraclass: Pteriomorphia) inhabiting the Mediterranean Sea since the Miocene era [3]. This filter-feeding species is commonly found on the coastal sandy sediment of the infralittoral, from 0.5 to 60 m, often in *Posidonia oceanica* meadows, where it lives semiburied, anchoring to the substratum thanks to its byssus threads, which glue to pebbles, sand, small fragments of robust biodetritic material, as well as the roots and rhizomes of *P. oceanica* [4].

Pinna nobilis has been exploited by humans for various purposes over the centuries [1,5–9]. It is worth noting that, in southern Italy, particularly in Apulia (Taranto), and in the Sardinia Island (Sant'Antioco), a longstanding tradition of harvesting the byssus produced by *P. nobilis* occurred, with the aim of creating luxurious fabrics and textiles. Lastly, populations of *P. nobilis* were severely impacted by indirect human activities, including boat anchoring, pollution, and habitat fragmentation [10,11].

Because of all these direct and indirect activities having impacted the species over time and space, *P. nobilis* experienced a strong demographic decline in all of its range, which accelerated significantly by the late 1980s [12]. To invert this trend, at the beginning of the 1990s, *P. nobilis* was included in a full protection regime under Annex IV of the EU Habitats Directive (European Council Directive 92/43/EEC) and Annex II of the Barcelona Convention (SPA/BD Protocol 1995). Moreover, many countries have enacted their own legislative measures, establishing conservation protocols to address the historical exploitation and continuous threats facing *P. nobilis* (e.g., in Italy, Marine Strategy Monitoring Program, Article 11 of Legislative Decree 190/210; in Slovenia, Annexes 1 and 2 of the Regulations for the protection of wild flora and fauna, Official Gazette of the Republic of Slovenia, nos. 46/04, 109/04, 84/05, 115/07, and 32/08; in Croatia, Croatian Nature Protection Act, Official Gazette 144/2013, 73/2016).

In a few decades, the great commitment of both the European community and single countries to protect the species led to a remarkable revival of its populations ([13] and references therein). This occurrence was, e.g., testified during the sampling activities done for several scientific studies on *P. nobilis* [5,14–28].

Unfortunately, an abnormal mortality of *P. nobilis* started in 2016, initially involving the populations located in the centre and southernmost coasts of Spain, together with the Balearic Islands [29], with death rates reaching up to 100%.

Since this first warning in southern Spain, countless mass mortality events (MMEs) have gradually been reported in an eastward direction in the western Mediterranean basin, involving northern Spain, France, and Italy, further reaching the central and the eastern part of the Mediterranean (Tunisia, Greece, and Turkey), and lastly affecting the Adriatic Sea, e.g., Slovenia, Croatia, Bosnia, and Herzegovina [29–39]. The initial investigations into the causes of the MME of *P. nobilis* primarily concentrated on the search for protozoa [29,30,32,33,40–42], and histological analysis on the first affected individuals of *P. nobilis* revealed the presence of a haplosporidian-like parasite (*Haplosporidium pinnae*) within the digestive gland [29,30,40], initially believed to be host-specific for fan mussels. Subsequent studies identified several bacterial species as potential pathogens contributing to the MME of *P. nobilis* [13,35–37,39,43–46], indicating that the disease is a multifactorial pathology [13,47,48]. Scarpa et al. [13] also reported the finding of *H. pinnae* in other bivalve species collected before 2016. These results highlighted that *H. pinnae* is not species-specific

as previously hypothesised, and that this protozoan was present in the Mediterranean Sea even before the start of the *P. nobilis* MME.

Only a few populations of *P. nobilis* survived the MME based on the most recent publications; however, we are aware that some populations disappear within a few months (e.g., Croatian populations, Čižmek et al. [35]) [49]. These populations are primarily situated in estuaries and isolated coastal lagoons in France, Italy, Spain [50], Greece [51], and the Sea of Marmara [52]. Notably, even if there are no recent surveys concerning the status of *P. nobilis* along the Tunisian coastlines, some unpublished studies indicate thriving populations in the Kerkennah Islands, Monastir Bay, and Bizerte Lagoon [53]. Even if the specific reasons for why these populations have remained unaffected are not yet understood, differences in salinity and temperature, in comparison to open waters, could play a role in their survival [34,50,53,54].

As a result of the MME, the conservation status of *P. nobilis* was reassessed, leading to its classification being updated from endangered to critically endangered [55]. In response to this critical situation, several Mediterranean organisations launched ex situ conservation programs, with an emphasis on captive breeding and reintroduction initiatives [45,55,56], and there is a strong emphasis on safeguarding the remaining, albeit few, wild populations of *P. nobilis* in the Mediterranean Sea [56]. Against this backdrop, the European community funded two ongoing LIFE projects devoted to the conservation and restocking of fan mussels: LIFE PINNARCA (https://www.lifepinnarca.com/the-project/) and LIFE PINNA (https://www.lifepinna.eu/en/the-project/). The second project involves the relocation of individuals from donor areas (sited in the northern Adriatic Sea) to receiving areas (in some sites of the western Mediterranean). Among these actions, there is an in-depth study of the genetic structure of *P. nobilis* to restore populations with genetic backgrounds like that which characterised extinct populations.

In such a context, several molecular studies have been performed on the genetic variability of *P. nobilis* so far [14,22,25,49,51,57–61]. The primary objective of the majority of these papers was to infer the genetic variability of *P. nobilis* after the protection plans started in the 1990s. Among them, Sanna et al. [14] provided a comprehensive Mediterranean-scale assessment of the genetic variability of *P. nobilis* using mitochondrial markers (cytochrome c oxidase subunit I (COI) and 16S ribosomal subunit genes) and combining their sequences with those obtained in previous studies [22,25]. Results highlighted high levels of genetic variability across the following marine ecoregions: (1) the western Mediterranean and the Ionian Sea; (2) the Adriatic Sea; (3) the Aegean Sea and Tunisian coastal regions. Furthermore, authors set the genetic boundary between the western and eastern Mediterranean basins in the Ionian Sea, thus suggesting that, for *P. nobilis*, the Sicilian straits do not represent a boundary for larval dispersal. These results were then corroborated by Sarafidou et al. [51], who analysed the residual genetic variation in the surviving populations after the MME from sites in the Aegean and Ionian Seas. Furthermore, based on the COI analyses, and consistent with what was previously hypothesised by Sanna et al. [14], Wesselman et al. [58] suggested that *P. nobilis* is characterised by a single mitochondrial haplogroup that experienced a recent population expansion starting from a small, original population. González-Wangüemert et al. [59] utilised microsatellite markers [62] to analyse different populations of *P. nobilis* sampled between 2010 and 2011 along the Spanish Mediterranean coast. Results showed a high genetic diversity and significant differentiation among post-larvae samples, but not among adult populations, suggesting that the overall genetic connectivity retrieved was correlated to both marine currents and dispersal models.

In this context, our research group had the unique opportunity to analyse "new" samples of *P. nobilis* collected before the mass mortality began. The possibility of genotyping them, considering that the species is on the brink of extinction and some restoration programs are based on the possibility of translocating individuals from the few remaining refuge areas to other areas where the species is now disappeared, has a relevant conservation importance. Indeed, past genetic variation patterns should be preserved when possible,

and phylogeography studies have demonstrated the capacity to help conceive and address conservation measures (see, e.g., [63–65]).

In light of such a background, the present study aimed to analyse the largest dataset of mitochondrial sequences available to understand the evolutionary patterns of genetic variability in the Mediterranean for *P. nobilis*. Analyses were performed using all data from populations that were sampled prior to the MME, and many of these populations are now becoming extinct. Utilising samples from populations not yet impacted by the MME, we investigated the genetic variation patterns without the influence of evolutionary forces resulting from the severe population collapse observed in *P. nobilis*. The combined effects of genetic drift, including bottleneck and/or founder effects, natural selection, and selective sweep, may have led to the disappearance of informative haplotypes or the amplification of previously uncommon ones. Our goal was to create a detailed portrayal of the historical genetic variability of *P. nobilis*, aiming to comprehend its potential survival in the face of mass mortality. We sought to understand how evolutionary forces might impact the species in the future.

For these reasons, our genetic analyses relied on the mitochondrial marker for which sequences from all the populations investigated to date are available in GenBank, and over a hundred newly obtained sequences from individuals collected before the MME in previously unexplored areas. Mitochondrial DNA (mtDNA) served as a valuable molecular marker for inferring population dynamics, dispersal patterns, and evolutionary history across various species [66].

In summary, the study pursued two main objectives: firstly, to depict the phylogeographic patterns and genetic variation in *P. nobilis* before the MME, aiding the development of strategies for the recovery and revival of populations resembling those that became extinct. Secondly, to validate the hypothesis proposed by Sanna et al. [14], establishing the genetic boundary between the western and eastern Mediterranean for this species eastward of the Sicilian strait through analyses of three new Adriatic populations.

A further aim of the present study was to shed light on the temporal and geographical origin of the species *P. nobilis*, whose unique Mediterranean fossil records were found in a late Pliocene–early Pleistocene site in northern Italy [67]. In this context, the Mediterranean Sea gradually refilled with seawater from the Atlantic Ocean after the Messinian salinity crisis ended, thus opening the door for the rise of new marine species in the re-emerging sea (references in Garcia-Castellanos [68]). Many of these new species were descended from Mediterranean ancestors that survived the desiccation, while other species migrated from the Atlantic Ocean. In the aftermath of the Messinian salinity crisis, the interplay between surviving Mediterranean species and Atlantic colonisers resulted in the emergence of numerous endemic species. In such a remarkable biodiversity mosaic, it would be important to gain insights into the evolutionary patterns of *P. nobilis* which allowed its diffusion in the Mediterranean. Indeed, based on these data, we can provide hints to formulate more informed predictions regarding the species' potential for recovery after the population decline caused by the MME.

2. Materials and Methods

2.1. Sample Collection

The dataset analysed in the present study included all the sequences of *Pinna nobilis* collected before the MME. They were either taken from GenBank (n = 369) [14,22,25,58] or obtained in the present study (n = 100) from the fresh tissues of samples collected in the past few years (see Table 1 for details).

Notably, the sequences from the Gulf of Trieste (the north of the Adriatic Sea) used in the present study belonged to populations sampled in 2018 when the northern Adriatic was not yet reached by the MME. For the collection of tissues from still-living individuals, we used a specific nonlethal sampling method, performed by scuba divers, which was developed by these authors [14] and which did not cause significant damage to the shells and soft tissues of *P. nobilis*. With this method, and with the recently received approval

of the Italian "Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA)" and "Ministero dell'Ambiente e della Tutela del Territorio e del Mare" [13], small fragments of mantle tissue from individuals with a minimally invasive technique were taken. No field studies involving the manipulation, dislocation, or removal of *P. nobilis* individuals were performed. For each location under protection, all necessary permits were obtained for the sampling activities by the authority responsible for each protected area. In Italy, the collection of samples from the Gulf of Trieste (Miramare) was performed according to a waiver to the Presidential Decree 357/97, proposed by the Scientific Directorate of the Marine Protected Area of Miramare (Trieste) (Prot. MATM PNM 0028355 of 10/10/2019). In Croatia, the collection of *P. nobilis* specimens within the national parks of Mljlet and Telašćica was performed under a sampling licence issued by the Ministry of Environmental Protection and Energy (KLASA: UP/I-612-07/16-48/103, URBROJ: 517-07-1-1-1-16-4).

Table 1. The table reports data on the sampling collection. Sampling sites are indicated for the individuals of *Pinna nobilis* collected during the present study. Details on the sampling locations are also provided for the sequences of *P. nobilis* from all over the Mediterranean taken from GenBank. The accession numbers with an asterisk (*) for Tunisian samples represent the cases for which only haplotypes were provided in GenBank. The group from the Iberian coastlines also includes samples from the site of Banyuls-sur-Mer that is situated on the Mediterranean coast in the French region of Languedoc-Roussillon, just north of the border with Spain.

Sample Code	Sampling Year	Sampling Area	Specimens	GenBank Code	Paper
		SARDINIA			
OSR	2013	Ossario (Asinara)	12	OR782596–OR782633	Present Study
ASI	2015	Cala di Scombro di Dentro and Cala Reale (Asinara)	38	OR782634–OR782645	
BPC	2010	Baia di Porto Conte	18	JX854788–JX854805	
POR	2010	Torre del Porticciolo	3	JX854806–JX854808	
LAZ	2010	Lazzareto	2	JX854809–JX854810	
OSM	2010	Ospedale Marino	21	JX854811–JX854831	
MOL	2010	Molara	11	JX854832–JX854842	
CCE	2010	Capo Ceraso	13	JX854843–JX854855	
SAL	2011	Le Saline	5	JX854856–JX854860	
MPE	2010	Monte Petrosu (Sassi piatti and Isola Cava)	4	JX854861–JX854864	Sanna et al. [14]
OTT	2011	Porto Ottiolu	5	JX854865–JX854869	
ORI	2011	Oristano	10	JX854870–JX854879	
MAR	2011	Marceddì	5	JX854880–JX854884	
IMV	2011	Isola di Mal di Ventre	4	JX854885–JX854888	
VMS	2011	Villasimius (Capo Caterina)	4	JX854889–JX854892	
CPA	2011	Costa Paradiso	5	JX854893–JX854897	
MAD	2011	Isola di La Maddalena (Cala Camiciotto)	18	JX854898–JX854915	
		CORSICA			
IPI	2011	Isola Piana	13	JX854916–JX854928	Sanna et al. [14]
CPC	2011	Cala Pesciu Cane	12	JX854929–JX854940	
		ELBA ISLAND			
ELB	2011	Capo Enfola	10	JX854992–JX855001	Sanna et al. [14]
		SICILY			
SVC	2011	San Vito lo Capo (Secca di Cala Rossa)	7	JX854941–JX854947	
MON	2011	Mondello	11	JX854948–JX854958	Sanna et al. [14]
MLZ	2011	Milazzo	10	JX854959–JX854968	
PAC	2011	Pachino (Capo Passero)	8	JX854969–JX854976	
OGN	2011	Ognina di Siracusa	15	JX854977–JX854991	

Table 1. *Cont.*

Sample Code	Sampling Year	Sampling Area	Specimens	GenBank Code	Paper
		ADRIATIC SEA			
VEN	2011	Ottagono Alberoni and Santa Maria del Lago	20	JX855002–JX855021	Sanna et al. [14]
MIR	2018	Miramare (Gulf of Trieste)	18	OR782678–OR782695	Present study
TEL	2015	Telašćica–Island Buč	14	OR782646–OR782659	
MLJ	2015	Mljet–Lake Malo Jezero	18	OR782660–OR782677	
		CYPRUS			
CYP	2011	Karaoglanoglu	2	JX855022–JX855023	Sanna et al. [14]
		AEGEAN SEA			
EPA–EPT	2006-2007	Epanomi	9	DQ448215–DQ448217 EF536827–EF536832	Katsares et al. [22]
AGG	2007	Aggelochori	9	EF536833–EF536841	
XIO	2007	Xios Island	5	EF536842–EF536846	
KOR	2007	Korinthiakos Gulf	3	EF536847–EF536849	
		TUNISIAN COASTLINES			
N	2010	Bizerta Lagoon	7	HM998857–HM998866 *	Rabaoui et al. [25]
M	2010	Monastir (Stah Jaber)	9		
S	2010	Kerkennah Island	7		
B	2010	El Bibane Lagoon	9		
K	2010	El Ketef	17		
BIZ	2013	Bizerta Lagoon	1	KF612603	Sanna et al. [57]
		IBERIAN COASTLINES			
BAN	2014	Banyuls (France)	9	KY321755–KY321811	Wesselmann et al. [58]
EBR	2014	Ebro Delta (Spain)	9		
IBI	2011	Ibiza (Spain)	10		
MUR	2014	Murcia (Spain)	9		
MALL	2011	Mallorca (Spain)	10		
ALI	2014	Alicante (Spain)	10		

Overall, the dataset analysed included a total of 469 sequences, 100 of which were newly obtained for the present study (representing 21% of the whole analysed dataset) and belonging to areas sited in the following 11 Mediterranean biogeographic sectors according to Bianchi et al. [69] (Table 1 and Figure 1): (1) Algerian and north Tunisian coasts; (2) Balearic Sea to Sardinia Sea; (3) Gulf of Lions; (4) southern Tyrrhenian Sea; (5) Straits of Messina; (6) Ionian Sea; (7) northern Adriatic Sea; (8) central Adriatic Sea; (9) northern Aegean Sea; (10) southern Aegean Sea; (11) Levant Sea.

2.2. Molecular Analyses

Total genomic DNA was isolated from a portion of mantle tissue using the Macherey-Nagel Nucleo Spin Tissue Kit (MACHEREY-NAGEL GmbH and Co. KG, Düren, Germany) following the supplier's instructions. DNA solutions were quantified using the Nanodrop™ Lite Spectrophotometer (by Thermo Scientific; Waltham, MA, USA), which showed an average yield of 54 ng/µL. A portion of the mitochondrial cytochrome c oxidase sub. I gene (COI) was amplified by standard PCR with specific primers for COI (L: 5′-GGTTGAACTATHTATCCNCC-3′ and H: 5′-GAAATCATYCCAAAAGC-3′) designed by the authors [14], which allowed us to obtain a COI fragment that was 338 base pairs long. Reactions were carried out in a total volume of 25 µL. On average, 10 ng of total genomic DNA were combined with 0.6 µM of each primer and one pellet of PuReTaq Ready-To-Go PCR beads (GE Healthcare, Wauwatosa, WI, USA) containing stabilizers, 4 ng of bovine serum albumin (BSA), deoxynucleotide triphosphates, 2.5 units of PuReTaq DNA

polymerase, and reaction buffer. When a bead was reconstituted to a 25 µL final volume, the concentration of each dNTP and $MgCl_2$ resulted in 200 µM and 1.5 mM, respectively. PCRs were performed in a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems, Waltham, MA, USA), programmed as follows: 1 cycle of 4 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at 46 °C, and 30 s at 72 °C. At the end, a post-treatment of 10 min at 72 °C and a final cooling at 4 °C were carried out. Both positive (high-quality DNA samples from the same species) and negative controls were used to test the effectiveness of the PCR protocols and the absence of possible contaminations. Electrophoresis was carried out on 2% agarose gels, prepared using 1× TAE buffer (Tris–acetate–EDTA, pH of 8.3) and stained with Gel Red Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). PCR products were purified by ExoSAP-IT (USB Corporation, Cleveland, OH, USA) and sequenced for forward and reverse strands (by means of the same primers used for PCR) using an external Sanger sequencing core service (Macrogen Europe, Amsterdam, The Netherlands and Macrogen Europe, Milano, Italy). Noteworthily, dual peaks of similar height, which could be interpreted as evidence of mitochondrial pseudogenes in the nucleus (Numts) or heteroplasmy, were not observed in any of the electropherograms.

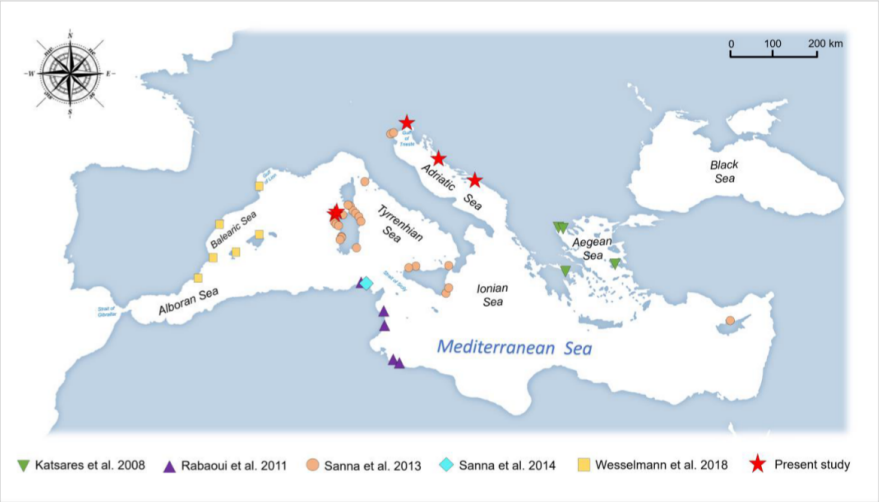


Figure 1. Map of the sampling sites. The map shows the geographical locations for the sequences obtained in the present study along with those from previous research [14,22,25,57,58].

2.3. Phylogenetic and Phylogeographic Analyses

The 100 newly generated sequences (GenBank #OR782596-OR782695) and the 369 already deposited in GenBank were aligned in their overlapping regions using the package Clustal Omega [70] (available at https://www.ebi.ac.uk/Tools/msa/clustalo/ (accessed on 3 August 2023)).

The genetic variation within the datasets was assessed estimating the number of polymorphic sites (S), the number of haplotypes (H), haplotype diversity (h), and nucleotide diversity (π) using the software package DnaSP 6.12.03 [71].

The best probabilistic model of sequence evolution was determined using jModeltest 2.1.3 [72], with a maximum likelihood optimised search, based on the Akaike (AIC) and Bayesian information criterion (BIC). Both criteria selected the GTR + I + G [73] as the best-fitting model for the dataset.

To infer the genetic relationships among haplotypes and to detect the possible occurrence of discrete genetic clusters, a median-joining network [74] was constructed by means of the software Network 10.2.0.0 (www.fluxus-engineering.com (accessed on 5 September

2023)) (Colchester, UK). Transitions and transversions were equally weighted. Due to the absence of information about the possible appearance of retromutation events, the same weight (10) was assigned to each observed polymorphism.

The principal coordinates analysis (PCoA) was performed using GenAlEX 6.5 [75] on a pairwise p-distance matrix, estimated by using the R packages APE (analysis of phylogenetics and evolution) [76]. This analysis allowed us to identify potential subgroups within the genetic clusters and to determine the dissimilarity represented by the genetic variation among the sequences (see Tran Thi et al. [77]).

Phylogenetic relationships among the Mediterranean *Pinna nobilis* populations and other species belonging to the Pinnidae family were investigated on a dataset including 469 sequences (see Table 1 and Figure 1 for details). Analyses were based on Bayesian inference (BI) and performed by means of the software MrBayes 3.2.7 [78]. The BI was performed by setting as the model parameters the following: NST = 6, rates = invgamma, and ngammacat = 4. Two independent runs consisting each of four Metropolis-coupled MCMC chains (one cold and three heated chains) were run simultaneously for 5,000,000 generations, sampling trees every 1000 generations. The first 25% of the 10,000 sampled trees was then discarded as burn-in (see Scarpa et al. [79]). To assess the convergence of the chains, parameters were verified by using the software Tracer 1.7.1 [80]. In addition, it was checked that the average standard deviation of split frequencies (ASDSFs) approached 0 [78] and the potential scale reduction factor (PSRF) was around 1 [81]. Nodes with a percentage of posterior probability lower than 95% were considered not highly supported. The phylogenetic tree was visualised and edited using FigTree 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/ (accessed on 9 October 2023)) (see Scarpa et al. [82]).

2.4. Estimation of the Divergence Time Analyses

The software package Beast 1.10.4 [83] was used to estimate the divergence time for the clades evidenced by the phylogenetic tree, applying the evolutionary rates proposed by Luttikhuizen et al. [84] for marine bivalves with pelagic larval dispersal. Molecular calibration with fossil data was not applicable in this case, as we aimed to set the molecular clock of the species, and fossil findings cannot trace back to the species level since they only allow for the collocation of the species origin within a temporal range. The mutation rates were set in Beauti (Beast package) by using a normal distribution ranging between 0.14% and 0.52% divergence per nucleotide site per million years. Site parameters were set accordingly to the evolutionary models selected by jModeltest: Substitution Model = GTR; Bases Frequencies = Estimated; Site Heterogeneity Model = Gamma + Invariant Sites; Number of Gamma Categories = 4. For the molecular clock rate variation model, the lognormal uncorrelated relaxed clock was selected, as it assumes independent rates on different branches. For the tree prior, the applied demographic model was the Speciation Yule Process [85,86]. The priors for the model parameters and statistics were determined for calibrating the time-tree assuming the mutation rates per million years. Divergence times were estimated using a normal distribution with lower, central, and upper values set according to the mutation rate per million years. Operator parameters were fixed following the instructions of the user manual. Additionally, the application of the lognormal uncorrelated relaxed clock model provided an indication of how clock-like were the data (measured by the ucld.stdev parameter). If the ucld.stdev parameter estimate was close to 0, then the data were quite clock-like; if the estimated value was much greater than 1, then the data exhibited very substantial rate heterogeneity among the lineages. To obtain the effective sample size (ESS) greater than 200 for all the statistic parameters, a run of 200,000,000 generations was performed, sampling a tree every 20,000 generations following Scarpa et al. [87]. The software Tracer 1.7.1 was also used to view the resulting log file, with the aim of ensuring the convergence of the parameter values to verify whether the ESS values exceeded 200 and to estimate the node ages. Tree Annotator (Beast package) and FigTree were used for drawing, visualising, and editing the time-calibrated tree following Scarpa et al. [87].

3. Results

On a total of 469 sequences, a total of 36 polymorphic sites were found, resulting in 49 different haplotypes (see Table 2 and the Figure A1 in Appendix A for the frequency of the distribution of the most common haplotypes in the whole Mediterranean).

Table 2. Sample sizes and genetic diversity estimates obtained for the mitochondrial region analysed for *Pinna nobilis* individuals. N: sample size; S: number of polymorphic sites; H: number of haplotypes; *h*: haplotype diversity; π: nucleotide diversity. Populations are labelled as in Table 1. Sites with gaps were not considered.

Sample	N	S	H	*h*	π
OSR	12	7	6	0.848	0.005
ASI	38	7	6	0.691	0.004
BPC	18	10	7	0.725	0.005
POR	3	3	2	0.667	0.006
LAZ	2	4	2	1.000	0.012
OSM	21	9	8	0.829	0.006
MOL	11	6	6	0.873	0.006
CCE	13	5	5	0.705	0.005
SAL	5	3	3	0.700	0.003
MPE	4	5	3	0.833	0.007
OTT	5	6	4	0.900	0.008
ORI	10	8	7	0.911	0.007
MAR	5	7	4	0.900	0.009
IMV	4	0	1	0.000	0.000
VMS	4	5	4	1.000	0.008
CPA	5	4	3	0.700	0.005
MAD	18	11	10	0.895	0.007
Sardinia	**178**	**28**	**31**	**0.830**	**0.006**
IPI	13	12	9	0.949	0.009
CPC	12	7	6	0.803	0.005
Corsica	**25**	**13**	**11**	**0.890**	**0.007**
Elba Island—ELB	**10**	**7**	**6**	**0.889**	**0.008**
SVC	7	5	4	0.714	0.005
MON	11	6	6	0.836	0.007
MLZ	10	6	5	0.867	0.007
PAC	8	8	7	0.964	0.007
OGN	15	7	9	0.886	0.007
Sicily	**51**	**13**	**16**	**0.882**	**0.007**
VEN	20	10	10	0.895	0.006
MIR	18	5	7	0.791	0.004
TEL	14	9	8	0.890	0.006
MLJ	18	5	5	0.752	0.005
Adriatic Sea	**70**	**16**	**21**	**0.870**	**0.007**
Cyprus—CYP	**2**	**1**	**2**	**1.000**	**0.003**
EP	9	9	6	0.833	0.007
AG	9	2	3	0.667	0.002
XI	5	2	3	0.700	0.002
KO	3	0	1	0.000	0.000
Aegean Sea	**26**	**11**	**8**	**0.720**	**0.004**
N	7	2	3	0.667	0.002
BIZ	1	0	0	0.000	0.000
M	9	4	4	0.694	0.004
S	7	2	3	0.667	0.002
B	9	1	2	0.556	0.002
K	17	1	2	0.382	0.001
Tunisian coastlines	**50**	**7**	**7**	**0.621**	**0.003**

Table 2. *Cont.*

Sample	N	S	H	*h*	π
BAN	9	3	4	0.750	0.004
EBR	9	2	3	0.417	0.002
IBI	10	3	4	0.533	0.002
MUR	9	0	1	0	0
MALL	10	4	4	0.533	0.003
ALI	10	9	7	0.867	0.010
Iberian coastlines	**57**	**15**	**15**	**0.555**	**0.004**
Whole dataset	**469**	**36**	**49**	**0.631**	**0.005**

The highest levels of genetic variation were found for the populations of the western Mediterranean islands (Corsica, Sardinia, Elba, and Sicily) along with the Adriatic Sea (Venetian Lagoon, Gulf of Trieste, Mljet, and Telašćica). On the other hand, lower levels of variation were found for the populations from the eastern Mediterranean (the Aegean Sea and Tunisian coastlines). Interestingly, the lowest rates of genetic variation for the whole Mediterranean were found for the Iberian coastlines, even when considering the relevant number of sequences which were analysed.

In the network analysis performed for this dataset (see Figure 2), the sequences were grouped into three different groups according to the genetic Mediterranean structuring proposed for *Pinna nobilis* by Sanna et al. [14]: the western Mediterranean (including sequences from Iberian coastlines, Corsica, Sardinia, Elba Island, and Sicily), the eastern Mediterranean (including sequences from Tunisian coastlines, the Aegean Sea, and Cyprus), and the Adriatic Sea.

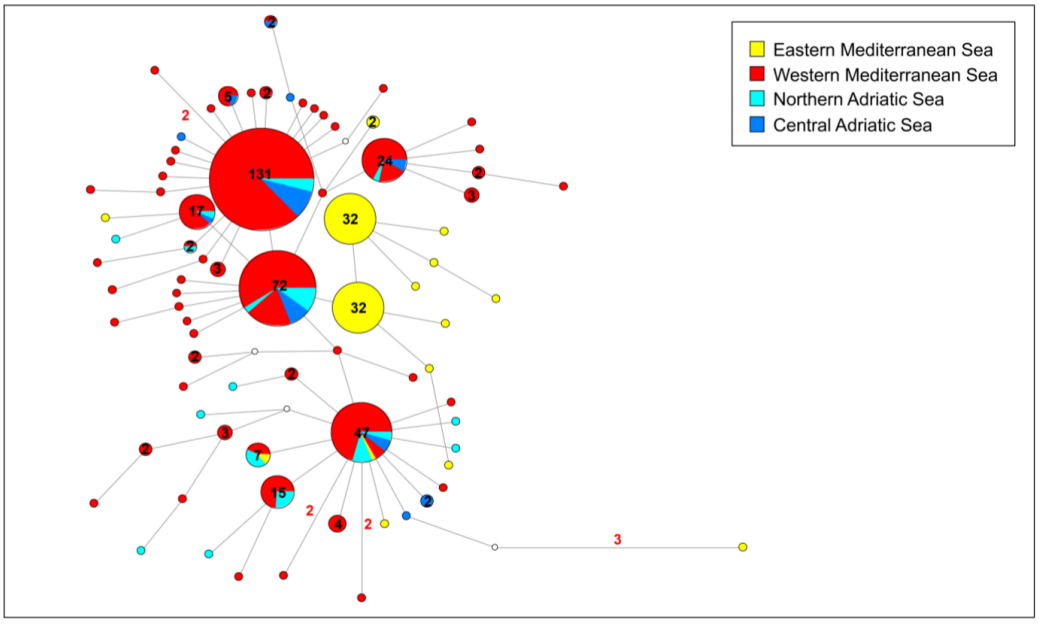


Figure 2. Median-joining network analysis performed on the COI gene fragment. The small white spots on the nodes show median vectors representing the hypothetical sequences that were calculated using the maximum parsimony method. The number of mutations between haplotypes that are greater than n = 1 are reported on the network branches. Additionally, the number of individuals showing the same haplotype that is greater than n = 1 are reported inside the spot.

Please note that, for the Adriatic group, two similar shades of blue were used to differentiate the network graph sequences from the north of the basin (light blue, including the Venetian Lagoon and Gulf of Trieste) from those belonging to the central part of the basin (sky blue, from the national parks of Mljlet and Telašćica). To test the hypothesis provided by Sanna et al. [14], which set the genetic boundary between the western and eastern Mediterranean for *P. nobilis*, eastward of the Strait of Otranto in the Ionian Sea, Adriatic sequences were divided into the northern and central Adriatic only for the network analysis. Certainly, Sanna et al. [14] proposed that the absence of additional Adriatic populations, aside from the Venetian Lagoon, hinders the confirmation of this hypothesis. In the current study, the inclusion of new sequences from four Adriatic populations in various basin locations enables us to offer insights into the accurate placement of the Mediterranean genetic boundary for this species. Results evidence the occurrence of six highly diffused haplotypes, resulting in at least three main typical network star-like shapes. The three most common haplotypes are diffused in the western Mediterranean and northern and central Adriatic, with the only exception being the third most common of them, which has been also found in one individual from Cyprus. Populations from the Adriatic Sea show a high level of haplotype sharing. These three common haplotypes are surrounded by several diverging haplotypes which, in general, differ for a single point mutation and are private to a single individual. Interestingly, two diverging haplotypes are exclusive to the eastern Mediterranean populations and are shared among Tunisian and Aegean individuals. Interestingly, these haplotypes exhibited divergence from those in the western Mediterranean due to two primary single-nucleotide polymorphisms (SNPs) identified within the last 25 nucleotides of the analysed COI fragment. Specifically, only one of these two distinctive polymorphisms is prevalent in nearly all individuals from both eastern and western populations. This constitutes a silent mutation occurring in the third base of a codon encoding for glycine, involving two purines (transition between the bases A and G). Importantly, this mutation does not induce any amino acid alterations in haplotypes.

The second noteworthy polymorphism, which distinguishes eastern and western populations, is present in a smaller proportion of individuals within western populations. Once again, this represents a silent mutation in the third base of a codon encoding for leucine, involving two pyrimidines (transversion between the bases C and T). Similar to the first polymorphism, this mutation does not result in any amino acid changes in haplotypes.

Overall, eastern Mediterranean populations do not share haplotypes with the western Mediterranean, with the exception of only one haplotype, which was found in seven individuals from Cyprus, Sicily, Venetian Lagoon, and Corsica. The network output evidenced a general high level of genetic variation with a huge number of similar haplotypes differentiating for a few mutations, and with a diffused haplotype sharing both among the populations of the western Mediterranean and the Adriatic, and among the populations of the eastern Mediterranean.

In line with the network analysis, the sequences utilised for the principal coordinates analysis (PCoA) were categorised into three distinct groups: the western Mediterranean, eastern Mediterranean, and Adriatic Sea. The overall findings were in harmony with the network analysis, and the cumulative percentage of variation explained by the first two axes (refer to Supplementary Figure S1 and Supplementary Table S1) just exceeded 50% (axis 1: 34.06%; axis 2: 24.29%). This underscores a general genetic uniformity among the sequences encompassed in the dataset. While the percentage of variation only weakly supported it, the results indicated a genetic structuring between two principal groups of sequences (G1 and G2) along axis 1.

The smaller group (G1), comprising 21.11% of the sequences, predominantly included individuals from the western Mediterranean and the Adriatic Sea, with only two exceptions from the island of Cyprus, two from the Aegean Sea, and two from Tunisian coastlines. In contrast, the larger group (G2) encompassed 78.68% of the sequences and grouped individuals from the western Mediterranean, Adriatic Sea, and eastern Mediterranean.

Notably, a single sequence from the Venetian Lagoon (Adriatic Sea) was identified as an outlier, positioned between the two main groups.

The Bayesian phylogenetic tree analysis was drawn based on a dataset including not only the *P. nobilis* COI sequences but also the relatives corresponding to all sequences of the species belonging to the family Pinnidae (*Pinna* and *Atrina* genera) available in GenBank so far (see Figure 3 for the schematic representation of the tree, and Figures 4–7 and the Supplementary Figure S2 for details on the species and GenBank accession numbers).

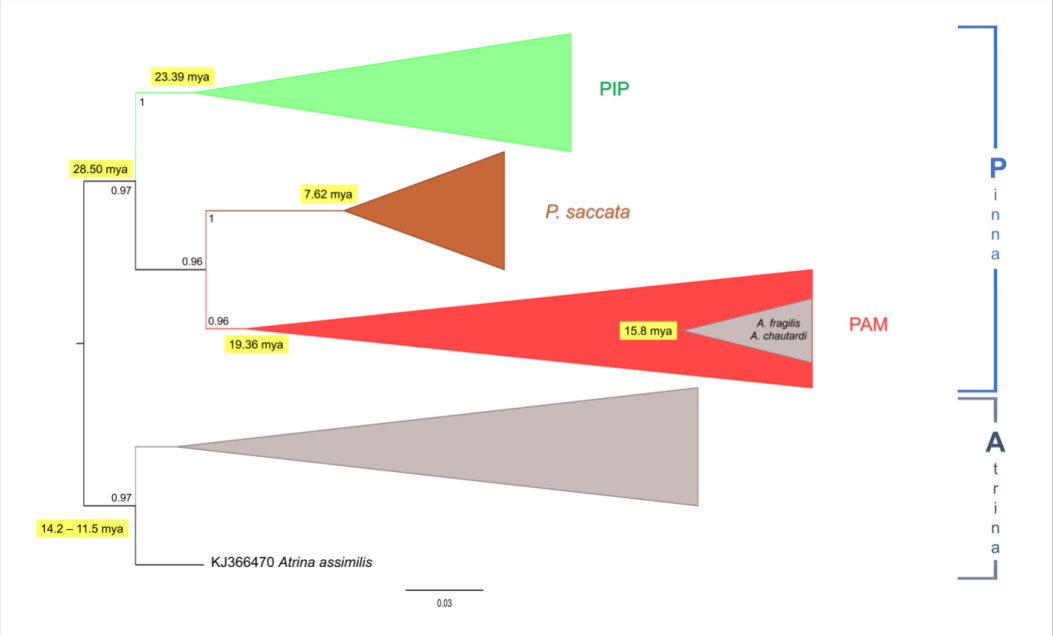


Figure 3. Schematic representation of the Bayesian phylogenetic tree based on the COI gene fragment, which is provided as Figure S2. The values of the node supports are expressed as posterior probabilities.

Bayesian analyses performed with MrBayes and Beast produced two trees with an identical topology at the main nodes. Accordingly, in the present study, only the tree obtained with MrBayes was presented with the graphical integration of the divergence times data from the ultrametric Beast tree, indicated at the main nodes.

It is noteworthy that, since *Streptopinna* is now considered a junior synonym of *Pinna* [2] based on molecular data, the sequences belonging to this subgenus of *Pinna* were included in the dataset as a species of *Pinna*.

The phylogenetic tree was rooted in the clade representative of the genus *Atrina*, with the aim to specifically infer the relationships among species of the genus *Pinna*. All the main nodes of the tree are well-supported, with values of posterior probabilities (pp) higher than 0.95, with the only exception being the internal large cluster including *Pinna carnea* Gmelin, 1791, *Pinna rudis* Linnaeus, 1758, and *P. nobilis* (pp = 0.55). However, it is important to take into consideration that, within this latter cluster, the nodes of *P. rudis* and *P. nobilis* clades (and the large cluster that includes these two clades) were highly supported.

Results highlight the presence of two monophyletic clades (P and A) representative of the genera *Pinna* (10 species) and *Atrina* (11 species), respectively. The clade A (see Figures 3 and 4 and the Supplementary Figure S2) dates back to a temporal range of 14.2–11.5 (HPD 95%: 3.6–36.64) mya and includes the species belonging to the genus *Atrina*, with the only exception being the *Atrina chautardi* (Nicklès, 1953) and *Atrina fragilis*

(Pennant, 1777), which were included in an almost-contemporary (15.84 mya) monophyletic internal subcluster within clade P.

Figure 4. Enlarged detail of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the clade A. The values of the node supports are expressed as posterior probabilities.

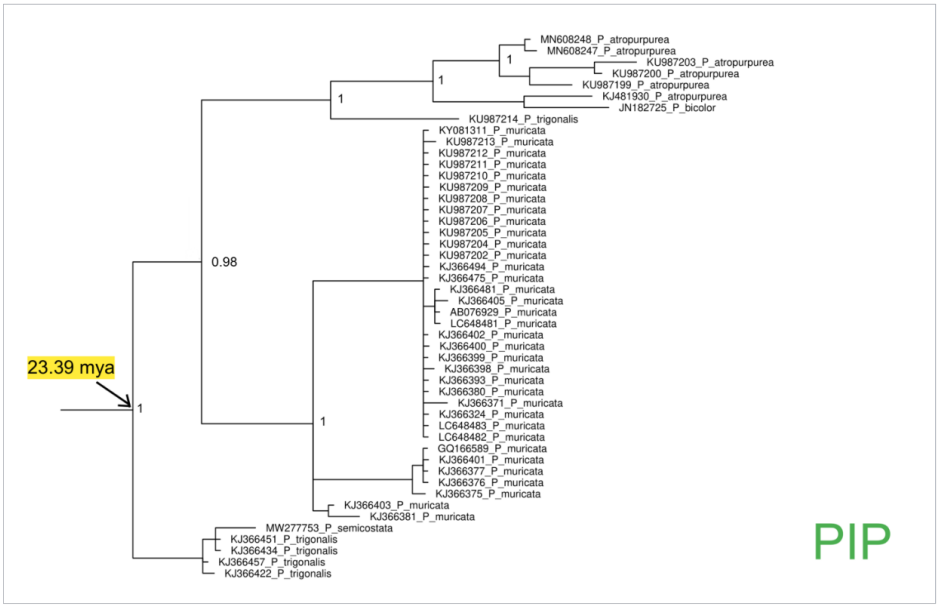


Figure 5. Enlarged detail of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the cluster PIP. The values of the node supports are expressed as posterior probabilities.

The *Atrina* species clustering within the clade A are almost all diffused in the Indian and Pacific Oceans. However, this clade also includes a well-supported monophyletic cluster, grouping sequences belonging to the species *Pinna epica* Jousseaume, 1894 (see Figure 4 and the Supplementary Figure S2). This discrepancy could be explained considering that this species was recently tentatively designed as *Abyssopinna epica* [88] with the genus *Abyssopinna* Schultz and Huber (2013), classified as a subgenus of *Pinna*. For this reason, the taxonomic status of *P. epica* is still puzzling and deserves to be further investigated from a phylogenetic point of view.

On the other hand, the clade P (see Figure 3 and the Supplementary Figure S2), which represents the genus *Pinna*, dates back to 28.50 (HPD 95%: 16.0–29.0) mya and includes two main clusters (PIP and PAM) (see Figures 3–6 and the Supplementary Figure S2) that are representative of Indo-Pacific (PIP) and Atlanto-Mediterranean (PAM) species.

The Indo-Pacific monophyletic cluster PIP (see Figures 3 and 5 and the Supplementary Figure S2) includes species which, in general, are diffused in the Pacific and Indian Oceans and date to 23.39 (HPD 95%: 7.3–69.3) mya.

The Atlanto-Mediterranean cluster PAM (see Figures 3 and 6 and the Supplementary Figure S2) dates back to 19.36 (HPD 95%: 7.4–34.9) mya and includes species from the Atlantic Ocean and Mediterranean Sea.

The sister group of PAM is represented by a well-supported monophyletic cluster that represents the pan-Indo-Pacific species *Pinna saccata* (Linnaeus, 1758) (see Figures 3, 7 and S2) which dates back to 7.62 (HPD 95%: 1.4–17.5) mya.

Within the cluster PAM (see Figure 6 and the Supplementary Figure S2), the monophyletic polytomic clade of *P. rudis* dates back to 1.1 (HPD 95%: 0.2–3.4) mya, while the large polytomy, which grouped the sequences of *P. nobilis*, dates back to a temporal range of 2.25–2.35 (HPD 95%: 1.29–4.47) mya. This latter cluster of *P. nobilis* is also inclusive of a well-supported monophyletic subcluster grouping the two Atlanto-Mediterranean *Atrina* species, *A. fragilis* and *A. chautardi*.

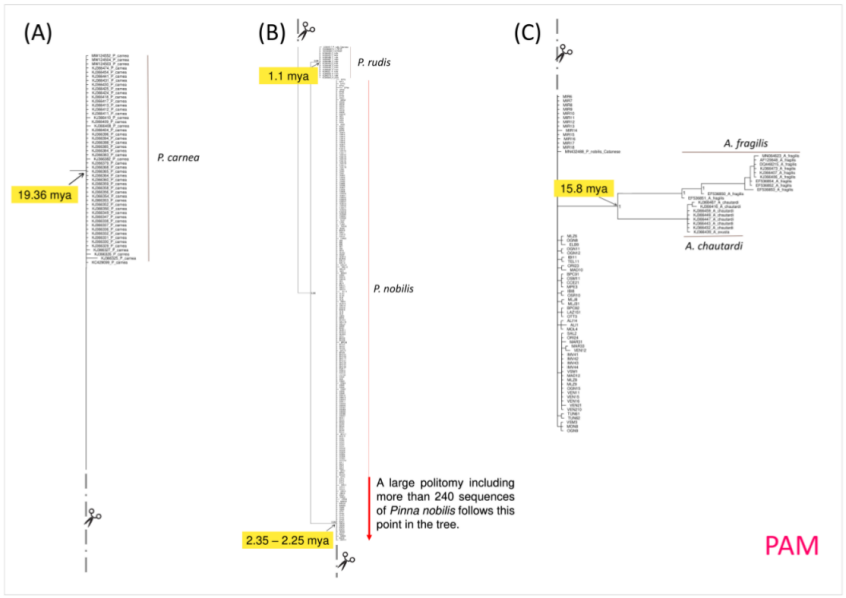


Figure 6. Enlarged details of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the cluster PAM. Due to the large size of the PAM cluster, the corresponding figure was here segmented into three parts (**A**–**C**). The section (**A**) represents all sequences of *Pinna carnea* within the dataset; the sections (**B**,**C**) represent sequences of *Pinna rudis* (**B**) and *Pinna nobilis* (**B**,**C**), along with sequences of *Atrina fragilis* and *Atrina chautardi* (**C**). The values of the node supports are expressed as posterior probabilities. Sequences from the present study are the one without the GenBank accession code.

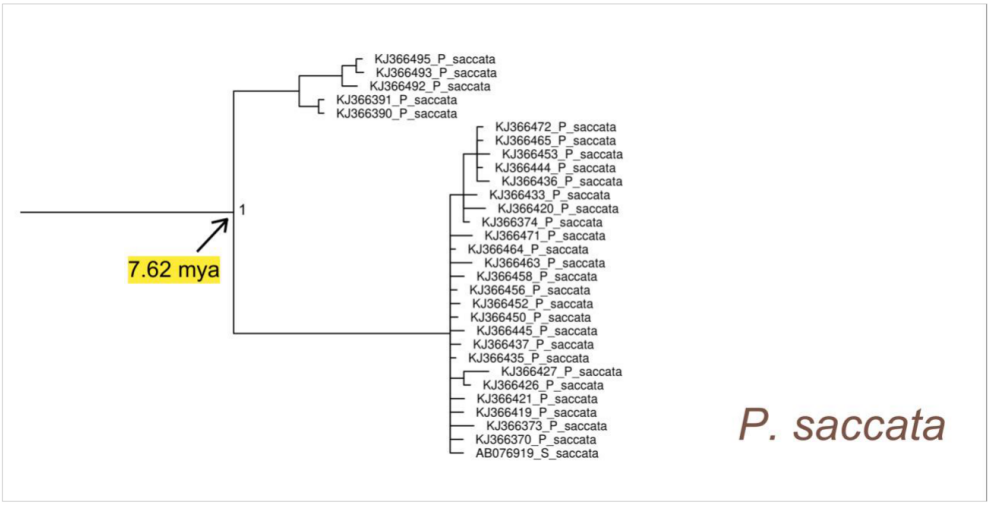


Figure 7. Enlarged detail of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the *Pinna saccata* cluster. The values of the node supports are expressed as posterior probabilities.

Consistent with the geographic areas considered for network and PCoA analyses, the patterns of *P. nobilis* spread were inferred by the BSP (Figures 8–10) for three groups of sequences: the western Mediterranean, eastern Mediterranean, and Adriatic Sea.

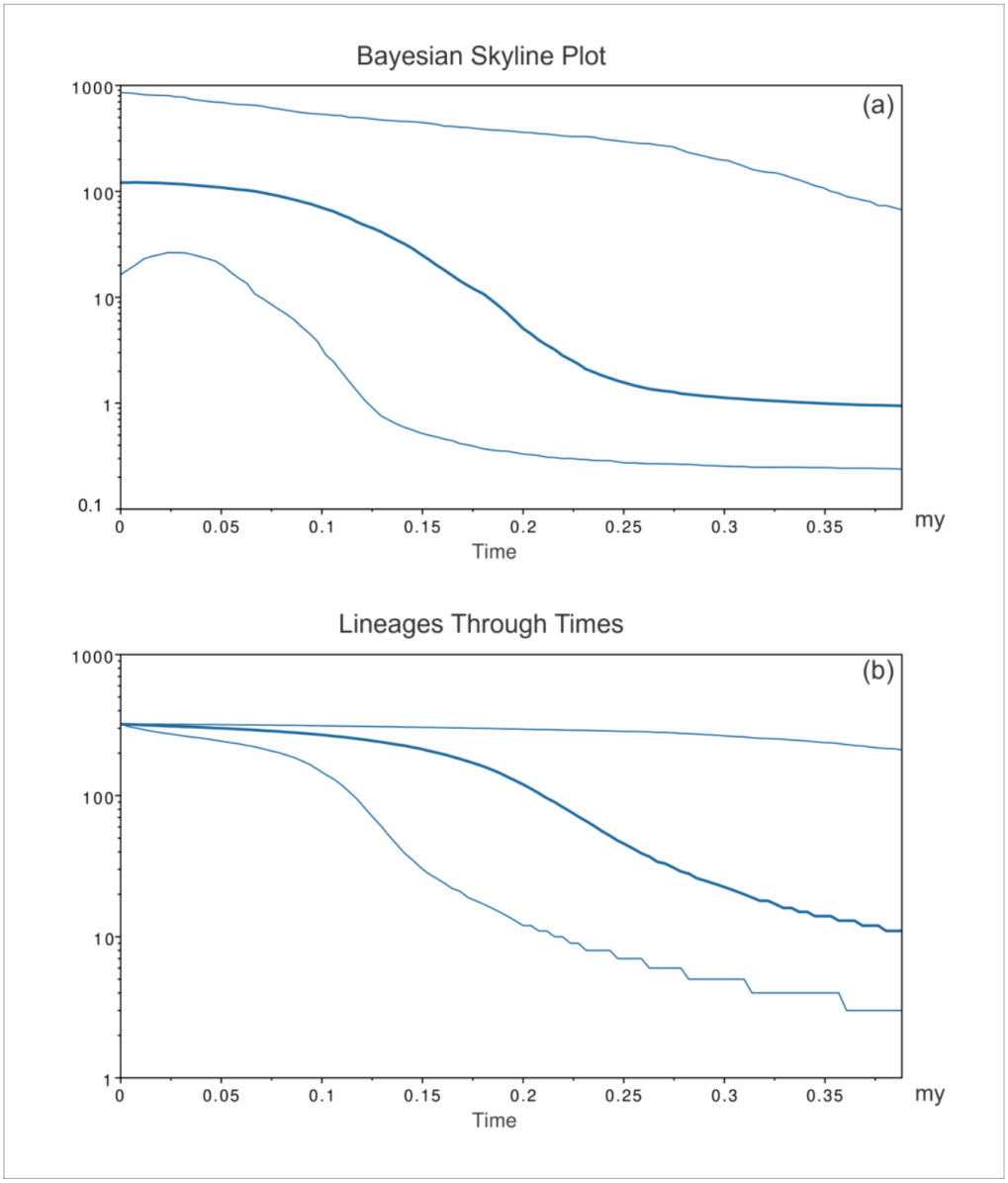


Figure 8. Bayesian skyline plot (**a**) and lineages through time (**b**) for the western Mediterranean populations of *Pinna nobilis*. The effective population size and the number of lineages in the y-axis are shown as a function of million years (my). The thicker central line represents the median value, while the thinnest lines delimit the 95% high posterior density (HPD) region.

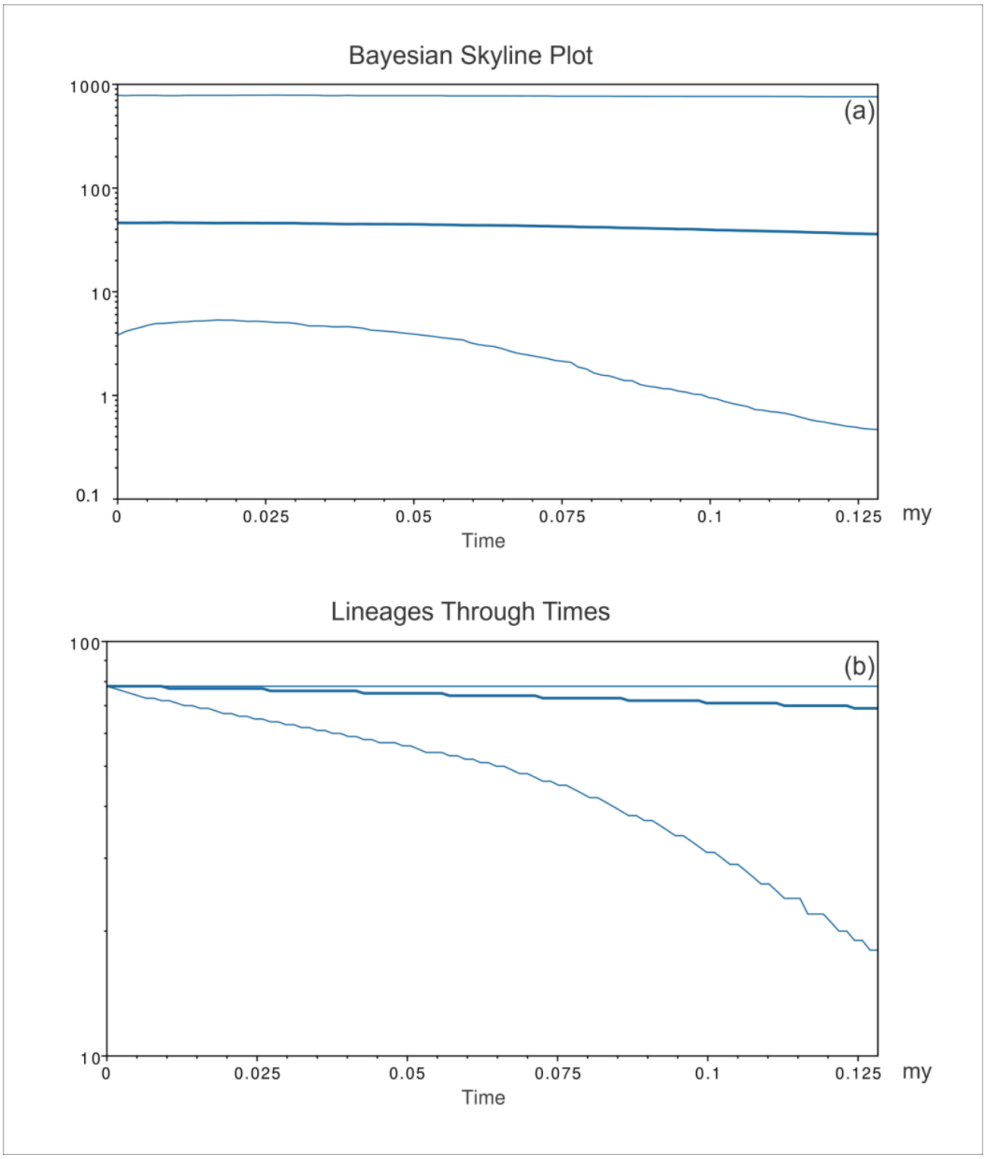


Figure 9. Bayesian skyline plot (**a**) and lineages through time (**b**) for the eastern Mediterranean populations of *Pinna nobilis*. The effective population size and the number of lineages in the y-axis are shown as a function of million years (my). The thicker central line represents the median value, while the thinnest lines delimit the 95% high posterior density (HPD) region.

The western Mediterranean BSP (*P. nobilis*, Figure 8a) showed an initial long-lasting constant size diffusion of the *P. nobilis* ancestor species from its origin up to 2.5 mya. According to the molecular dating based on the phylogenetic tree analysis, this latter moment (2.5 mya) approximately corresponds to the differentiation of the species *P. nobilis*.

The early population of fan mussels experienced a constant exponential expansion of the early population that lasted for about two million years and was followed by a decrease in the population expansion.

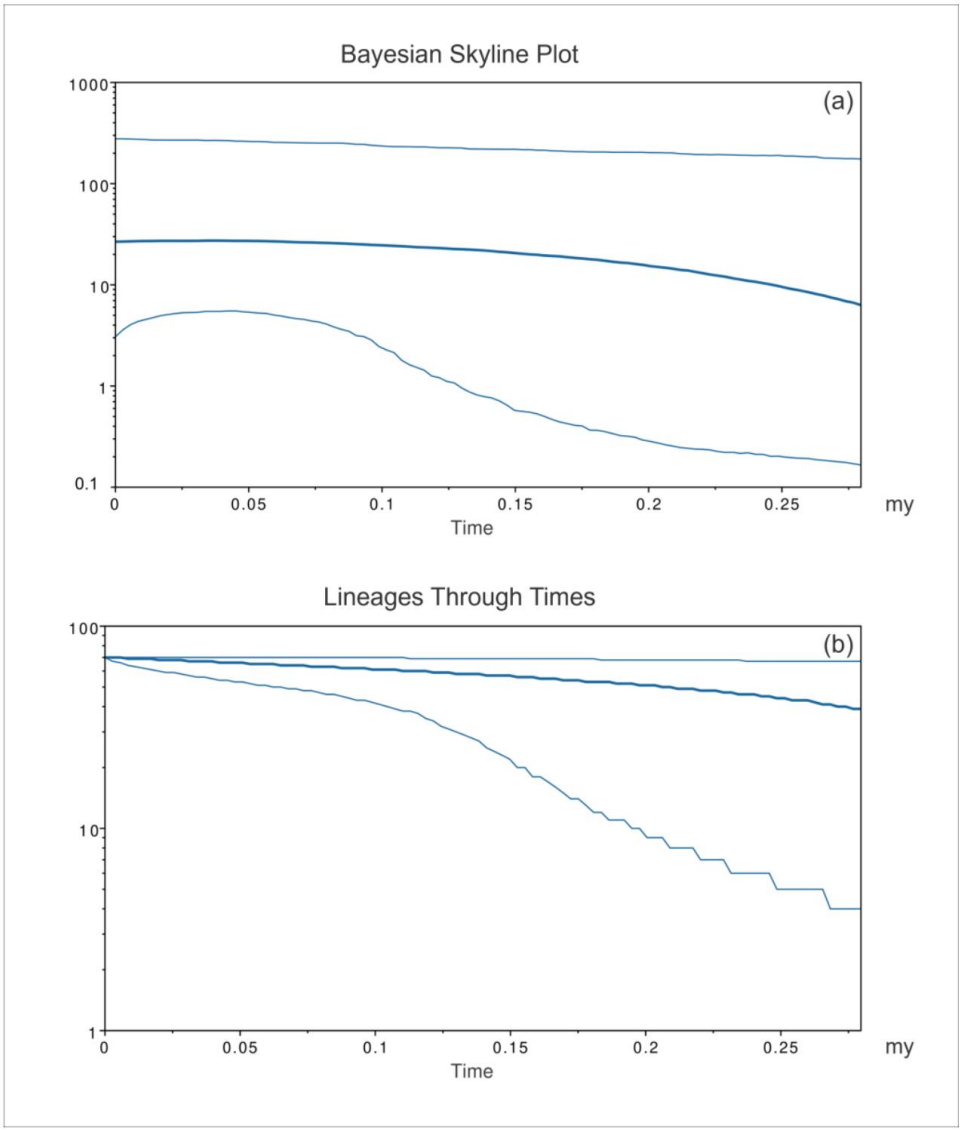


Figure 10. Bayesian skyline plot (**a**) and lineages through time (**b**) for the Adriatic Sea populations of *Pinna nobilis*. The effective population size and the number of lineages in the y-axis are shown as a function of million years (my). The thicker central line represents the median value, while the thinnest lines delimit the 95% high posterior density (HPD) region.

Consistently, the analysis also evidenced (Figure 8b) that, from the first radiation of the *P. nobilis* ancestor species, the amount of its mitochondrial COI lineages in the

western Mediterranean constantly increased. The species *P. nobilis* differentiated during this exponential growth of mitochondrial lineages (about 2.25–2.35 mya according to the molecular dating of the phylogenetic tree), which lasted for about one further million years after the rise of *P. nobilis*. Starting from this moment, the mitochondrial lineages of *P. nobilis* reduced their growth, reaching a plateau about 0.25 mya.

The eastern Mediterranean BSP (*P. nobilis*, Figure 9a) evidenced a general long-lasting constant size diffusion of *P. nobilis*, which started 1.25 mya. From the first radiation of the species in the eastern Mediterranean, the amount of the mitochondrial COI lineages (Figure 9b) increased, with a constant low level of growth whose extent steadily decreased over time until it reached a plateau approximately 0.25 mya.

The Adriatic Sea BSP (*P. nobilis*, Figure 10a) showed a general trend with a long-lasting constant size expansion of the *P. nobilis* early population, which started approximately 2.5 mya (the early expansion might have been produced by the *P. nobilis* ancestor species). A slight population expansion for *P. nobilis* lasted in the Adriatic Sea for about one million years and was followed by a decrease in the population expansion. The analysis also evidenced that, from the first radiation of *P. nobilis* in the Adriatic, the amount of its mitochondrial COI lineages (Figure 10b) increased, with a general constant low level of growth whose extent steadily decreased over time until it reached a plateau approximately 0.25 mya.

4. Discussion

This study, incorporating 21.5% of new *Pinna nobilis* COI sequences from one western Mediterranean and three Adriatic populations not previously investigated, contributes to a more comprehensive understanding of the genetic landscape of *P. nobilis* before the mass mortality event (MME) that pushed the species to the brink of extinction. The choice of the mitochondrial COI marker for the analyses was taken since it was the only one that allowed us to create a dataset of sequences encompassing all the individuals of *P. nobilis* genotyped in the last fifteen years. Importantly, the use of the COI gene enabled the comparison of new molecular data from the present study with populations now extinct, for which it would be impossible to perform analyses with other types of molecular markers.

Consequently, we reconstructed, with the utmost accuracy, the pre-MME phylogeographic patterns of *P. nobilis* in the Mediterranean Sea, shedding light on the origin of the genetic structuring between the western and eastern *P. nobilis* populations hypothesised in Sanna et al. [14] and recently corroborated by Sarafidou et al. [51] based on the analysis of Aegean and Ionian populations surviving after the MME. By utilising samples from populations not yet affected by the mass mortality of *P. nobilis*, we were able to study the evolutionary patterns of the species without the bias of the evolutionary forces that affected the genetic landscape of the populations due to the MME.

The present study infers, with the highest level of resolution possible, the phylogeny of *P. nobilis*, offering, for the first time, insights into the origin of this species. In this context, it must be emphasised that the occurrence of doubly uniparental inheritance (DUI) for mitochondrial DNA, which could impact the reliability of results, has never been reported in *P. nobilis* [14].

4.1. Phylogeography and Evolutionary History of Pinna nobilis

In accordance with Sanna et al. [14] and Wesselman et al. [58], the present study underscores the presence of a general mitochondrial genetic homogeneity among *Pinna nobilis* populations in the Mediterranean before the MME. Notably, the analysis, incorporating a significantly increased number of new sequences (over 20%), confirms the only exception to this trend reported in Sanna et al. [14], wherein populations from Tunisian coastlines and the Aegean Sea exhibit private haplotypes not found in *P. nobilis* populations from the western Mediterranean sampling sites. However, it should be noted that these exclusive haplotypes from the eastern Mediterranean, which were identified by Katsares et al. [22] and Raboui et al. [25], differ from the western Mediterranean and Adriatic haplotypes by

only one or two silent point mutations, which do not impact the amino acid composition of the mitochondrial enzyme cytochrome c oxidase.

Additionally, two sequences from Cyprus showed haplotypes commonly found in the western Mediterranean and Adriatic Sea, and four Tunisian and one Aegean sequence exhibited haplotypes directly derived from those most frequently observed in the western Mediterranean and Adriatic Sea. This suggests that, in the eastern Mediterranean, the prevalent "eastern" haplotypes might have coexisted with several "western" haplotypes that remained undetected due to chance and the limited number of individuals sampled from those areas before the MME. Thus, the entire Mediterranean basin might have been interested by a level of haplotype sharing among western and eastern populations higher than those identified until now. This potential scenario is supported by a recent study on surviving populations from the Aegean and Ionian seas [51], which revealed haplotype sharing between the western Mediterranean and Ionian Sea based on the analysis of a concatenated mitochondrial fragment, including COI and ribosomal 16S genes.

However, it cannot be ruled out that the observed mitochondrial genetic structuring between the western and eastern Mediterranean for *P. nobilis* could result from a selective sweep increasing the frequency of the most adaptive allelic variants in the eastern Mediterranean. The mitochondrial marker used in this study did not reveal any adaptive changes involving protein production for the mutations accounting for the divergence between western and eastern populations. Nevertheless, these haplotypes may be associated with adaptive allelic variants warranting investigation with new nuclear biparental molecular markers in future studies on surviving populations.

The inclusion of new populations from the northern and central Adriatic Sea in this study has provided a better understanding of the level of genetic divergence between this basin and the western Mediterranean. Despite lower genetic variability, our results demonstrate the presence of haplotypes in the Adriatic that also occur in the western Mediterranean. This finding holds significant importance for the LIFE PINNA project (https://www.lifepinna.eu/en/home/), where the transplantation of *P. nobilis* individuals from the northern Adriatic was planned for activities in the western Mediterranean. In light of these results, Adriatic populations of *P. nobilis* have proven to be genetically comparable to those in the western Mediterranean intervention sites (https://www.lifepinna.eu/en/areas-of-intervention/), which have now become extinct.

Furthermore, this study provides a definitive answer to the hypothesis proposed by Sanna et al. [14] regarding the correct position of the genetic boundary spanning the Sicilian Strait. Indeed, this research provides support to the presence of the Mediterranean boundary between western and eastern Mediterranean basins eastward of the Sicilian Strait, in the Ionian Sea.

Several common haplotypes were identified, and their presence and distribution in the network (Figure 2) suggest significant founder effects during the evolutionary history of *P. nobilis*. A similar evolutionary pattern of network star-like shapes has been reported for the mitochondrial COI gene in other species belonging to the family Pinnidae (e.g., *Pinna saccata*, *Pinna muricata*, and *Atrina rigida* (Lightfoot, 1786)) in the Pacific Ocean and the Caribbean Sea [2]. The discovery of this trend in Pinnidae species distributed in vastly different geographic areas suggests a slow mitochondrial mutational substitution rate combined with a high potential for larval dispersal in these species.

The high mitochondrial homogeneity observed in *P. nobilis* is linked to the entry of its ancestor into the Mediterranean from the Atlantic during the Messinian salinity crisis. The Mediterranean became disconnected from the Atlantic Ocean in the late Miocene (5.6 mya). The subsequent Messinian salinity crisis, occurring around this time [68], led to the near-complete desiccation of the basin due to water evaporation. The Zanclean flood event, dated 5.33 mya, marked the return of Atlantic waters to the Mediterranean through the Gibraltar Strait, causing rapid and violent flooding at a rate exceeding ten metres per day [68]. While the return of Atlantic waters might have started weakly and slowly a few thousand years earlier, 90% of the transfer occurred during the Zanclean flood, filling the

Mediterranean basin in a short period, ranging from a few months to two years. This event potentially facilitated the early colonisation of the central Mediterranean seabed by larvae belonging to the ancestor of *P. nobilis*, leading to its adaptation and differentiation into the modern, endemic species. This scenario aligns with the detailed description provided by Bianchi et al. [89], which delineates the phases of the Mediterranean Sea refilling post-Zanclean flood along with the migration of Atlantic species into the Mediterranean.

Furthermore, recent identification of *P. nobilis* fossils in a well-studied late Pliocene–early Pleistocene marine succession along the Stirone riverbanks in northern Italy supports this scenario [67]. This region, rich in Pinnidae fossils [90–92], likely witnessed the early evolution and differentiation of *P. nobilis*.

Consistent with this proposed scenario, the western Mediterranean populations of *P. nobilis* analysed in this study, covering the coastlines of Corsica, Sardinia, Elba, and Sicily, exhibited the highest levels of genetic variability in the entire Mediterranean. In contrast, the lowest level of genetic variation was observed in *P. nobilis* populations from the Iberian coastlines. This discrepancy may be explained by the direction of the Zanclean flooding, initially involving the central part of the Mediterranean and excluding, at least in the initial phase, the Iberian Peninsula area [93]. The Zanclean flood also created a seabed incision (the Zanclean channel) through the erosion produced by the flooding waters [68,94], establishing a direct connection between the Atlantic and the central Mediterranean through the hydrographic constriction produced by the Strait of Gibraltar [93,95]. The areas laterally distant from this dashing flow were initially excluded [89,93].

The rapid adaptation process of the *P. nobilis* ancestor to Mediterranean conditions, following the settlement of larvae from the Atlantic, might have been facilitated by effective recruitment and fast turnover, as evidenced in the modern fan mussel from the Gulf of Lion [96]. According to the proposed model for the evolution of the Mediterranean Sea level after the Zanclean flood [68], the waters reached the Sicily sill at the end of the first phase of filling, leading the western Mediterranean to attain the maximum marine water level. Only later, did water slowly begin to fill the Adriatic Sea and the eastern Mediterranean through the Sicilian Strait. This sequence of events may explain the overall lower level of *P. nobilis* genetic variation in these two basins.

Our results suggest that *P. nobilis* differentiated from its most recent ancestor in the western Mediterranean approximately 2.3 mya. This estimate aligns with the dating (late Pliocene–early Pleistocene) of *P. nobilis* fossils found in deposits along the Stirone river in northwest Italy [67]. Early populations of the species might have rapidly colonised the Adriatic Sea, where a small number of mitochondrial lineages common in the western Mediterranean are present. Subsequently, the species dispersed in the eastern Mediterranean after more than one million years. This region now features exclusive haplotypes, likely derived from the western Mediterranean, following a typical founder effect model possibly associated to the effect of a selective sweep.

Our study reveals that the diffusion of *P. nobilis* in the Adriatic and eastern Mediterranean, along with its evolutionary rate, advanced slowly, possibly due to biological conditions differentiating these basins from the western Mediterranean (e.g., marine currents, temperature, and salinity), where the species likely spread more rapidly. The species, known to be highly responsive to variations in salinity and temperature [34,50,53,54], and indirectly to the hydrodynamic conditions (which strongly influence the patterns of genetic connectivity for this species, as supposed by González-Wangüemert et al. [59] for modern populations of *P. nobilis*), may have experienced a slower evolutionary rate influenced by climatic changes during the Quaternary glaciations.

Indeed, in the whole Mediterranean, the mitochondrial lineages of *P. nobilis* drastically reduced their growth in the middle Pleistocene, reaching a plateau at the beginning of the Riss glaciation. The severe climatic changes that happened during the Quaternary glaciations might have affected the phylodynamic patterns of *P. nobilis*. For example, during the glacial periods, deeper waters (from below 200 m) flowed into the Mediterranean Sea from the Atlantic [97], while near-surface waters flowed out of the Mediterranean into the

Atlantic Ocean across the Straits of Gibraltar [98,99]. In this context, the hydrodynamic of surface waters, where the pelagic larval stage of *P. nobilis* spread, could have been involved in the decreasing of the evolutionary rate of the species.

4.2. *Phylogeny of* Pinna nobilis

As previously mentioned, the phylogenetic analysis results indicate that *Pinna nobilis* is an early Pleistocene species, having differentiated approximately 2.3 million years ago in the central Mediterranean after its ancestor entered the region during the Zanclean flood [68]. *Pinna nobilis* is part of a monophyletic cluster that originated in the Miocene period and, in accordance with Lemer et al. [2], includes species from the same genus found in the Atlanto-Mediterranean area, namely, *Pinna carnea* and *Pinna rudis* (see the PAM cluster in the phylogenetic tree shown in Figure 3).

Remarkably, we recognised *P. rudis* as the sister taxon of *P. nobilis*, whose sequences are included within a monophyletic clade that dates to about one million years ago. This finding partially differs from those of Lemer et al. [2], in which *P. rudis* individuals did not form a monophyletic clade and the sequences from Senegal were genetically divergent from those from the Mediterranean, the Azores, and the Canary Islands, and form a sister group to a clade, including the remaining *P. rudis*, *P. nobilis*, and *P. carnea* sequences. Such a discrepancy could be attributable to the increased number of *P. rudis* sequences from the Mediterranean that were included in the present study, as well as to both the molecular marker and the bioinformatic approach used for constructing the phylogenetic tree. Indeed, regarding the latter issue, the phylogenetic tree proposed by Lemer et al. [2] is based on a concatenated heterogeneous dataset of nuclear and mitochondrial regions (18S rRNA, 28S rRNA, 16S rRNA, and COI) that may have led to a scenario somewhat different to that based on our analyses.

It is noteworthy that, in the current investigation, the clade encompassing *P. rudis* underwent differentiation during the Pleistocene epoch (approximately 1 million years ago), coinciding with a decline in the evolutionary rate of *P. nobilis*. This finding can be explained by the fact that, after a founder-flush effect of speciation, colonisation, and adaptation to the habitat of *P. nobilis* in the Mediterranean, the population then stabilised, and *P. rudis*, which evolved more recently, covered some ecological niches less conducive to the development of *P. nobilis*. Specifically, *P. nobilis* favours coastal areas, predominantly seagrass meadows, at depths ranging from 0.5 to 60 m, as well as rocky bottoms or rhodolith beds. In contrast, *P. rudis* displays a preference for small patches of sand in rocky bottoms and rock crevices. While this species can inhabit depths from the surface to 70 m, it is more commonly found at depths between 20 and 70 m, particularly in areas protected from strong water movements, where *P. nobilis* is generally less abundant due to its susceptibility to currents and fronts [100].

In the present scenario, it is interesting to observe that some ecological niches left vacant by *P. nobilis* following mass mortality are now being partially occupied by *P. rudis*. Kersting and Ballesteros [101] documented a behavioural shift in *P. rudis* populations in the Columbretes Islands Marine Reserve off the Iberian Mediterranean coast. Their study revealed a significant increase in the recruitment rates for *P. rudis* after the mass mortality event of *P. nobilis* in 2017. The proposed hypothesis suggests that the local extinction of *P. nobilis* created an opportunity for *P. rudis* to expand, potentially due to the reduced interspecific competition. Given the rarity of high concentrations of both species in the Mediterranean region [102], this occurrence may signify an ongoing process of *P. nobilis* replacement by *P. rudis*.

Moreover, considering the findings of Coupé et al. [47] that some *P. nobilis* individuals that introgressed with *P. rudis* were resistant to infection by *Haplosporidium pinnae*, the expansion of *P. rudis* in the Mediterranean could lead to an increased frequency of hybridisation and introgression with *P. nobilis*. This, in turn, could confer resistance to *H. pinnae*, a primary etiological agent of mass mortality events, thereby potentially promoting the recovery and survival of the fan mussel.

Our findings align with those of Lemer et al. [2], providing support for the notion that the sister group of the cluster encompassing the Atlanto-Mediterranean species of the genus *Pinna* (see the PAM cluster in the phylogenetic tree of Figure 3) comprises an ancient group containing sequences attributed to *Pinna saccata* originating in the early Pliocene. Originally designated as *Streptopinna*, this species underwent recent revision by Lemer et al. [2], classifying *Streptopinna* as a subgenus (status nov.) of *Pinna* and reinstating *P. saccata* as the type species. Our results further affirm the taxonomic reassignment of *Streptopinna saccata* to its new status as an ancient and divergent species within the genus *Pinna*.

The discovery that *P. saccata*, with a pan-Indo-Pacific distribution, is closely related to other *Pinna* species distributed in the Atlanto-Mediterranean region should not be surprising. It is essential to consider that this species traces back to approximately eight million years ago, when the Pacific and Atlantic were still interconnected at the Isthmus of Panama. Indeed, the complete emergence of this land bridge occurred about 4 million years later (3.1–3.5 million years ago), leading to the total isolation between the Caribbean Sea and the Pacific Ocean [79,103–106].

Interestingly, in addition to *P. saccata*, the species belonging to the genus *Pinna* are divided into two large genetic clusters (PIP and PAM clusters of the phylogenetic tree in Figure 3) on the basis of their geographic origin, the Indo-Pacific and Atlanto-Mediterranean areas, respectively. Having originated in the early Miocene (about 19–23 mya), these clusters are almost contemporary. The beginning of the separation between the Indo-Pacific and the Atlanto-Mediterranean *Pinna* species could be ascribed to the Oligocene adaptive radiation, which gave rise to species very similar to the modern ones. This process might have concluded when the Mediterranean evolved into its current enclosed nature, mostly during Miocene [107], with the closure of the eastern Tethys Sea as Africa and Eurasia (20 mya) [108], and the uplift of the Isthmus of Panama about 3.1–3.5 mya [109,110] that produced the separation of the Atlantic and the Indo-Pacific Oceans.

Overall, our results suggest that the modern genus *Pinna* would have diversified in the Cenozoic era (in the late Oligocene period), which is set earlier than what was previously supposed based on fossils by either Gomez-Alba [3], which fixed its origin in the Miocene period of the Cenozoic era, and Rosewater [111], which fixed its origin in the Jurassic period of the Mesozoic era. This trend has also been observed to be overlapping to that of the genus *Atrina*, since our findings suggest that this taxon is about 15 million years younger than the genus *Pinna*, having originated during the Miocene period of the Cenozoic era. The discrepancy with the previous estimates of its origin based on fossils is relevant because the genus *Atrina* was previously believed to have appeared during the Carboniferous period of the Mesozoic era Rosewater [111].

It is noteworthy that, in our phylogenetic tree, two Atlanto-Mediterranean species of *Atrina* (*Atrina fragilis* and *Atrina chautardi*) do not fall within the clade encompassing all the species of this genus. Instead, they are nested within the clade of *P. nobilis* in a monophyletic cluster that is contemporary to the clade of *Atrina* (see the clade A in the phylogenetic tree in Figure 3). It is crucial to consider that, among *Atrina* species, *A. fragilis* and *A. chautardi* are those predominantly distributed in the Atlantic–Mediterranean region. However, an alternative scenario could be considered to explain the presence of the monophyletic cluster of these two Atlanto-Mediterranean *Atrina* species within the *P. nobilis* clade. This cluster of species, which differentiated in the Miocene, may be representative of the group of ancestor species of *P. nobilis* that originated in the Atlantic Ocean and subsequently migrated into the Mediterranean following the Messinian crisis with the Zanclean flood.

Interestingly, the monophyly of the cluster including *A. fragilis* and *A. chautardi* is partially consistent with Lemer et al. [2], whose maximum clade credibility simplified tree, based on the analysis of concatenated nuclear and mitochondrial genetic data, set the monophyletic cluster of *A. fragilis* and *A. chautardi* in an external position within the clade of *Atrina*.

5. Conclusions

The comprehensive reconstruction of the evolutionary history of *Pinna nobilis* presented in this study stands as the most thorough and detailed depiction of phylogeographical and evolutionary patterns available for populations inhabiting distinct basins of the Mediterranean Sea—specifically, the western and eastern Mediterranean, as well as the Adriatic and Aegean seas—prior to the mass mortality event (MME). To attain this objective, we opted to employ the mitochondrial COI gene as the molecular marker, a choice driven by its capacity to facilitate a broad geographic evolutionary investigation, incorporating molecular data from all the populations studied over the past 15 years. The utilisation of the COI gene played a pivotal role in achieving the remarkable degree of accuracy observed in the present study, a precision that would have been challenging to attain without its inclusion.

The obtained results offer novel insights into the evolutionary history of *P. nobilis* populations preceding the MME. These data serve as the foundational basis for (1) appropriately managing survivor individuals/populations and (2) devising effective restocking/transplantation plans in regions where *P. nobilis* faced extinction, as outlined in the action of the European LIFE PINNA project. Particularly concerning the latter point, understanding the evolutionary dynamics of *P. nobilis* and its distribution patterns in the Mediterranean before the catastrophic mortality event holds paramount significance. This knowledge is crucial for repopulation plans, emphasising the need to restore populations with genetic variation as closely aligned as possible to extinct populations. This approach aims to mitigate deleterious genetic drift phenomena that could lead to maladaptation in the newly established populations [112–114]. In this context, having a comprehensive understanding of the historical genetic variation and evolutionary trajectory of *P. nobilis* can aid in clarifying how this species has coped with genetic drift (bottlenecks in primis) over time, thus adapting to habitat changes. Having also achieved, by phylogeographic inferences in the present study, new hints on the origin and diffusion of *P. nobilis* after its early differentiation makes it possible to identify the better funders to be used during the plans of population restocking so to minimize the effect of genetic drift.

From a phylogeographic viewpoint, our analyses on the pre-MME genetic variability of *P. nobilis* led to the hypothesis that the extreme and rapid climate changes occurred during the Messinian salinity crisis and the spread of new species into the Mediterranean basin from the Atlantic through the Zanclean channel [68,93,95] could be involved in the appearance of this Pleistocene species which differentiated in the central Mediterranean and shaped the genetic variation observed in its modern populations before the MME. Starting from the knowledge of the distribution of *P. nobilis* genetic variability in the past, the present study laid the foundation to shed light on the evolutionary dynamics characterising this species and help to properly address the conservation management of surviving individuals and the restoration plan of extinct populations. In the future, the analysis of surviving populations from the eastern Mediterranean could highlight the occurrence in these refuge areas (where data currently available report a good number of surviving individuals of *P. nobilis*) of never-described haplotypes that were present in the western Mediterranean before the MME. This finding would be suggestive of the past existence of a large pan-Mediterranean population for *P. nobilis* and, if confirmed by further studies, could provide support for the repopulation plans developed in the western Mediterranean for this species. This scenario could facilitate the conservation of this species, whose restored populations would have a high possibility of maintaining a genetic variation near that of the extinct populations. In this context, Sarafidou et al. [51] analysed the genetic variation in the surviving (sampled in 2018–2021) eastern Mediterranean populations of *P. nobilis* from Greek regions not yet investigated. Results were congruent with the studies performed before the MME [14,58] in evidencing the lack of differentiation among the areas within the eastern Mediterranean and the genetic structuring between the western and eastern basins. Interestingly, the "eastern haplotypes" found by Katsares et al. [22] and Rabaoui et al. [25] were not reported in surviving populations, as a possible consequence of the genetic drift-

operated MME, but individuals from the Ionian Sea showed haplotypes that were present in past western and eastern populations, and in modern eastern ones.

In such a context, the Ionian Sea, identified as the genetic boundary between the western and eastern Mediterranean (validated by Sanna et al. [14] and supported by the present study), may serve as a potential overlapping "refuge area". Here, populations with genetic variability spanning the entire Mediterranean could persist. Although there is a lack of genetic information for *P. nobilis* in this geographic region before the mass mortality event, it holds promise as a significant source area for restocking and conservation initiatives. In the future, further genetic analyses focused on Italian surviving populations in the Ionian Sea, if any, have the potential to illuminate the conservation status of *P. nobilis* and provide insights into its capacity for recovery seven years after the onset of the MME.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ani14010114/s1, Figure S1: Principal coordinates analysis performed on the COI gene fragment. Bidimensional plot shows the genetic differentiation among specimens due to the nucleotide substitutions per site found in the dataset. Figure S2: Bayesian phylogenetic tree based on the COI gene fragment analysed in the present study. The values of the node supports are expressed as posterior probabilities. Sequences from the present study are the one without the GenBank accession code. Table S1: Principal coordinates analysis. The table reports the results of the principal coordinates analysis performed on the whole COI fragment dataset.

Author Contributions: Conceptualisation, D.S., I.A. and M.C.; methodology, D.S., F.S. and I.A.; software, F.S.; validation, D.S. and M.C.; formal analysis, D.S., I.A. and F.S.; investigation, D.S., I.A. and F.S.; resources, D.S., F.S. and M.C.; data curation, I.A., F.S., C.L., A.P., C.M., S.C. and P.A.; writing—original draft preparation, D.S., I.A. and M.C.; writing—review and editing, D.S., I.A., F.S., C.L., M.C., P.A., P.K., C.M., A.P., S.C., M.S. and E.B.; visualisation, D.S., I.A., F.S., C.L., M.C., P.A., P.K., C.M., A.P., S.C., M.S. and E.B.; supervision, D.S. and M.C.; project administration, D.S. and M.C.; funding acquisition, D.S., F.S. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the LIFE Program of the European Union: LIFE PINNA LIFE20 NAT/IT/001122 CUP B95F21001880002 "Conservation and re-stocking of Pinna nobilis in the western Mediterranean and Adriatic sea" and by the Italian Ministry of University and Research, PNRR, Missione 4, Componente 2 "Dalla ricerca all'impresa", Investimento 1.4 Project CN00000033 which supported the National Biodiversity Future Center of the University of Sassari.

Institutional Review Board Statement: Ethical review and approval are not applicable for this study, as genotyping was carried out on old genomic DNA samples stored in absolute ethanol at −20 °C. However, for the collection of tissues in the past years, we used a specific nonlethal sampling method that did not cause significant damage to the shells and soft tissues of *P. nobilis*. This method received the approval of the Italian "Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA)" and "Ministero dell'Ambiente e della Tutela del Territorio e del Mare". No field studies involving the manipulation, dislocation, or removal of *P. nobilis* individuals were performed for the present research.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequences obtained in the present study for the mitochondrial cytochrome c oxidase subunit I gene isolated in *Pinna nobilis* from Italy and Croatia were deposited in the GenBank database under the accession numbers OR782596–OR782695.

Conflicts of Interest: Authors Saul Ciriaco, Marco Segarich, Edoardo Batistini were employed by the company Shoreline Soc. Coop that, however, has no conflict of interest. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

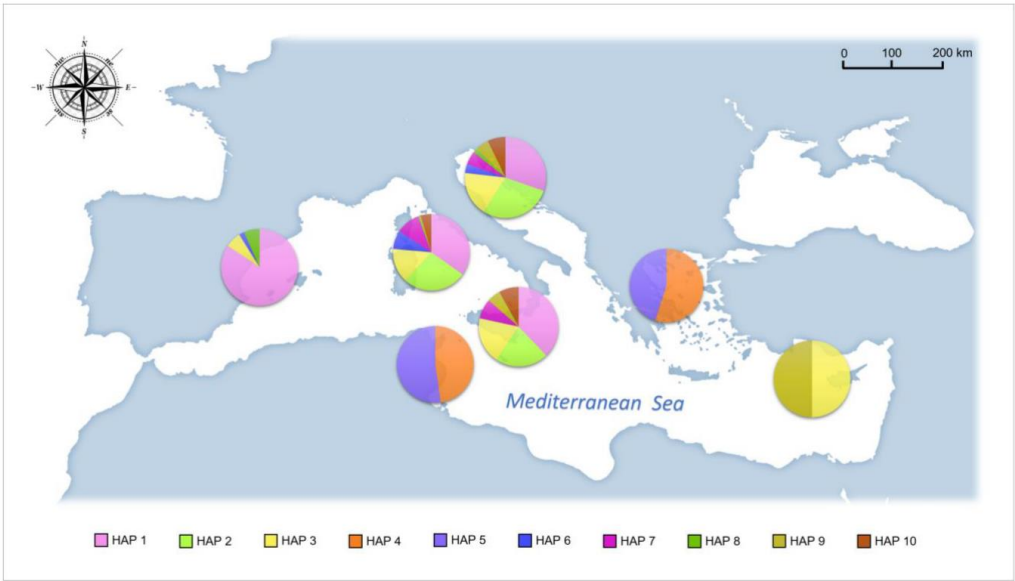


Figure A1. Map showing the distribution of the most common COI haplotypes within the whole Mediterranean basin. In the figure, only haplotypes shared by 5 or more individuals are represented. The number of each haplotype, which is reported in the bottom of the figure, do not refer to a specific nomenclature.

References

1. Butler, A.; Vicente, N.; de Gaulejac, B. Ecology of the pterioid bivalves *Pinna bicolor* Gmelin. *Mar Life* **1993**, *3*, 37–45.
2. Lemer, S.; Buge, B.; Bemis, A.; Giribet, G. First molecular phylogeny of the circumtropical bivalve family Pinnidae (Mollusca, Bivalvia): Evidence for high levels of cryptic species diversity. *Mol. Phylogenet. Evol.* **2014**, *75*, 11–23. [CrossRef] [PubMed]
3. Gómez-Alba, J. *Guía de Campo de los Fósiles de España y de Europa*; Ediciones Omega: Barcelona, Spain, 1988.
4. March, J.R.G. Aportaciones al conocimiento de la Biología de Pinna nobilis Linneo, 1758 (Mollusca: Bivalvia) en el Litoral Mediterráneo Ibérico. Ph.D. Thesis, Universitat de Valencia, Valencia, Spain, 2005.
5. Rabaoui, L.; Tlig-Zouari, S.; Katsanevakis, S.; Ben Hassine, O.K. Modelling population density of *Pinna nobilis* (Bivalvia) on the eastern and southeastern coast of Tunisia. *J. Molluscan Stud.* **2010**, *76*, 340–347. [CrossRef]
6. Basso, L.; Vázquez-Luis, M.; García-March, J.R.; Deudero, S.; Alvarez, E.; Vicente, N.; Duarte, C.M.; Hendriks, I.E. The pen shell, *Pinna nobilis*: A review of population status and recommended research priorities in the Mediterranean Sea. *Adv. Mar. Biol.* **2015**, *71*, 109–160. [CrossRef] [PubMed]
7. Scarpa, F.; Sanna, D.; Azzena, I.; Cossu, P.; Casu, M. From dark to light and back again: Is *Pinna nobilis*, the largest Mediterranean shellfish, on the brink of extinction? What about *Pinna nobilis*. *Veterinaria* **2021**, *70*, 1–14.
8. Greenwald, M. *The Cruising Chef Cookbook*; Paradise Cay Publication: Middletown, CA, USA, 1996.
9. Centoducati, G.; Tarsitano, E.; Bottalico, A.; Marvulli, M.; Lai, O.R.; Crescenzo, G. Monitoring of the endangered *Pinna nobilis* Linne, 1758 in the Mar Grande of Taranto (Ionian sea, Italy). *Environ. Monit. Assess.* **2007**, *131*, 339–347. [CrossRef] [PubMed]
10. Vázquez-Luis, M.; Borg, J.A.; Morell, C.; Banach-Esteve, G.; Deudero, S. Influence of boat anchoring on *Pinna nobilis*: A field experiment using mimic units. *Mar. Freshw. Res.* **2015**, *66*, 786–794. [CrossRef]
11. Öndes, F.; Kaiser, M.J.; Güçlüsoy, H. Human impacts on the endangered fan mussel, *Pinna nobilis*. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2020**, *30*, 31–41. [CrossRef]
12. Öndes, F.; Alan, V.; Akçalı, B.; Güçlüsoy, H. Mass mortality of the fan mussel, *Pinna nobilis* in Turkey (eastern Mediterranean). *Mar. Ecol.* **2020**, *41*, e12607. [CrossRef]
13. Scarpa, F.; Sanna, D.; Azzena, I.; Mugetti, D.; Cerruti, F.; Hosseini, S.; Cossu, P.; Pinna, S.; Grech, D.; Cabana, D.; et al. Multiple Non-Species-Specific Pathogens Possibly Triggered the Mass Mortality in *Pinna nobilis*. *Life* **2020**, *10*, 238. [CrossRef]

14. Sanna, D.; Cossu, P.; Dedola, G.L.; Scarpa, F.; Maltagliati, F.; Castelli, A.; Franzoi, P.; Lai, T.; Cristo, B.; Curini-Galletti, M.; et al. Mitochondrial DNA Reveals Genetic Structuring of *Pinna nobilis* across the Mediterranean Sea. *PLoS ONE* **2013**, *8*, e67372. [CrossRef]
15. Addis, P.; Secci, M.; Brundu, G.; Manunza, A.; Corrias, S.; Cau, A. Density, size structure, shell orientation and epibiontic colonization of the fan mussel *Pinna nobilis* L. 1758 (Mollusca: Bivalvia) in three contrasting habitats in an estuarine area of Sardinia (W Mediterranean). *Sci. Mar.* **2009**, *73*, 143–152. [CrossRef]
16. Combelles, S.; Moreteau, J.C.; Vicente, N. Contribution a la connaissance de l'ecologie de *Pinna nobilis* L. (mollusque eulamellibranche). *Sci. Rep. Port-Cros Nat. Park* **1986**, *12*, 29–43.
17. Coppa, S.; Guala, I.; de Lucia, G.A.; Massaro, G.; Bressan, M. Density and distribution patterns of the endangered species *Pinna nobilis* within a *Posidonia oceanica* meadow in the Gulf of Oristano (Italy). *J. Mar. Biolog. Assoc. U.K.* **2010**, *90*, 885–894. [CrossRef]
18. Šiletić, T.; Peharda, M. Population study of the fan shell *Pinna nobilis* L. in Malo and Veliko Jezero of the Mljet National Park (Adriatic Sea). *Sci. Mar.* **2003**, *67*, 91–98. [CrossRef]
19. Katsanevakis, S. Population ecology of the endangered fan mussel *Pinna nobilis* in a marine lake. *Endanger. Species Res.* **2005**, *1*, 51–59. [CrossRef]
20. Rabaoui, L.; Tlig Zouari, S.; Katsanevakis, S.; Ben Hassine, O.K. Comparison of absolute and relative growth patterns among five *Pinna nobilis* populations along the Tunisian coastline: An information theory approach. *Mar. Biol.* **2007**, *152*, 537–548. [CrossRef]
21. García-March, J.R.; Pérez-Rojas, L.; García-Carrascosa, A.M. Influence of hydrodynamic forces on population structure of *Pinna nobilis* L., 1758 (Mollusca: Bivalvia): The critical combination of drag force, water depth, shell size and orientation. *J. Exp. Mar. Biol. Ecol.* **2007**, *342*, 202–212. [CrossRef]
22. Katsares, V.; Tsiora, A.; Galinou-Mitsoudi, S.; Imsiridou, A. Genetic structure of the endangered species *Pinna nobilis* (Mollusca: Bivalvia) inferred from mtDNA sequences. *Biologia* **2008**, *63*, 412–417. [CrossRef]
23. Katsanevakis, S. Population dynamics of the endangered fan mussel *Pinna nobilis* in a marine lake: A metapopulation matrix modeling approach. *Mar. Biol.* **2009**, *156*, 1715–1732. [CrossRef]
24. Katsanevakis, S.; Thessalou-Legaki, M. Spatial distribution, abundance and habitat use of the protected fan mussel *Pinna nobilis* in Souda Bay, Crete. *Aquat. Biol.* **2009**, *8*, 45–54. [CrossRef]
25. Rabaoui, L.; Mejri, R.; Tlig-Zouari, S.; Bahri, L.; Ben Hassine, O.K.; Tsigenopoulos, C.S. Genetic variation among populations of the endangered fan mussel *Pinna nobilis* (Mollusca: Bivalvia) along the Tunisian coastline. *Hydrobiologia* **2011**, *678*, 99–111. [CrossRef]
26. Prado, P.; Caiola, N.; Ibáñez, C. Habitat use by a large population of *Pinna nobilis* in shallow waters. *Sci. Mar.* **2014**, *78*, 555–565. [CrossRef]
27. Vázquez-Luis, M.; March, D.; Alvarez, E.; Alvarez-Berastegui, D.; Deudero, S. Spatial distribution modelling of the endangered bivalve *Pinna nobilis* in a Marine Protected Area. *Mediterr. Mar. Sci.* **2014**, *15*, 626–634. [CrossRef]
28. Vafidis, D.; Antoniadou, C.; Voultsiadou, E.; Chintiroglou, C. Population structure of the protected fan mussel *Pinna nobilis* in the south Aegean Sea (eastern Mediterranean). *J. Mar. Biolog. Assoc. U. K.* **2014**, *94*, 787–796. [CrossRef]
29. Vázquez-Luis, M.; Álvarez, E.; Barrajón, A.; García-March, J.R.; Grau, A.; Hendriks, I.E.; Jiménez, S.; Kersting, D.; Moreno, D.; Pérez, M.; et al. S.O.S. *Pinna nobilis*: A Mass Mortality Event in Western Mediterranean Sea. *Front. Mar. Sci.* **2017**, *4*, 220. [CrossRef]
30. Catanese, G.; Grau, A.; Valencia, J.M.; Garcia-March, J.R.; Vázquez-Luis, M.; Alvarez, E.; Deudero, S.; Darriba, S.; Carballal, M.J.; Villalba, A. *Haplosporidium pinnae* sp. nov., a haplosporidan parasite associated with mass mortalities of the fan mussel, *Pinna nobilis*, in the Western Mediterranean Sea. *J. Invertebr. Pathol.* **2018**, *157*, 9–24. [CrossRef] [PubMed]
31. Hanžek-Turnšek, H. Ugotavljanje Gostote Populacije Velikega Leščurja (*Pinna nobilis*) z Metodo Podvodnih Videotransektov: Zaključna Naloga. Doctoral dissertation, Univerza na Primorskem, Fakulteta za Matematiko, Naravoslovje in Informacijske Tehnologije, Koper, Slovenia, 2018.
32. Katsanevakis, S.; Tsirintanis, K.; Tsaparis, D.; Doukas, D.; Sini, M.; Athanassopoulou, F.; Kolygas, M.N.; Tontis, D.; Koutsoubas, D.; Bakopoulos, V. The cryptogenic parasite *Haplosporidium pinnae* invades the Aegean Sea and causes the collapse of *Pinna nobilis* populations. *Aquat. Invasions* **2019**, *14*, 150–164. [CrossRef]
33. Panarese, R.; Tedesco, P.; Chimienti, G.; Latrofa, M.S.; Quaglio, F.; Passantino, G.; Buonavoglia, C.; Gustinelli, A.; Tursi, A.; Otranto, D. *Haplosporidium pinnae* associated with mass mortality in endangered *Pinna nobilis* (Linnaeus 1758) fan mussels. *J. Invertebr. Pathol.* **2019**, *164*, 32–37. [CrossRef]
34. Cabanellas-Reboredo, M.; Vázquez-Luis, M.; Mourre, B.; Álvarez, E.; Deudero, S.; Amores, Á.; Addis, P.; Ballesteros, E.; Barrajón, A.; Coppa, S.; et al. Tracking a mass mortality outbreak of pen shell *Pinna nobilis* populations: A collaborative effort of scientists and citizens. *Sci. Rep.* **2019**, *9*, 13355. [CrossRef]
35. Čižmek, H.; Čolić, B.; Gračan, R.; Grau, A.; Catanese, G. An emergency situation for pen shells in the Mediterranean: The Adriatic Sea, one of the last *Pinna nobilis* shelters, is now affected by a mass mortality event. *J. Invertebr. Pathol.* **2020**, *173*, 107388. [CrossRef]
36. Lattos, A.; Giantsis, I.A.; Karagiannis, D.; Michaelidis, B. First detection of the invasive Haplosporidian and Mycobacteria parasites hosting the endangered bivalve *Pinna nobilis* in Thermaikos Gulf, North Greece. *Mar. Environ. Res.* **2020**, *155*, 104889. [CrossRef]

37. Šarić, T.; Župan, I.; Aceto, S.; Villari, G.; Palić, D.; De Vico, G.; Carella, F. Epidemiology of Noble Pen Shell (Pinna *nobilis* L. 1758) Mass Mortality Events in Adriatic Sea Is Characterised with Rapid Spreading and Acute Disease Progression. *Pathogens* **2020**, *9*, 776. [CrossRef] [PubMed]
38. Čelebičić, M.; Gradaščević, N.; Viteškić, V. Is *Pinna nobilis* L. critically endangered in Neum bay, Bosnia and Herzegovina? *Veterinaria* **2020**, *69*, 221–227.
39. Künili, I.E.; Ertürk Gürkan, S.; Aksu, A.; Turgay, E.; Çakir, F.; Gürkan, M.; Altinağaç, U. Mass mortality in endangered fan mussels *Pinna nobilis* (Linnaeus 1758) caused by co-infection of *Haplosporidium pinnae* and multiple *Vibrio* infection in Çanakkale Strait, Turkey. *Biomarkers* **2021**, *26*, 450–461. [CrossRef] [PubMed]
40. Darriba, S. First haplosporidan parasite reported infecting a member of the Superfamily Pinnoidea (*Pinna nobilis*) during a mortality event in Alicante (Spain, Western Mediterranean). *J. Invertebr. Pathol.* **2017**, *148*, 14–19. [CrossRef] [PubMed]
41. López-Sanmartín, M.; López-Sanmartín, M.; Catanese, G.; Grau, A.; Valencia, J.M.; García-March, J.R.; Navas, J.I. Real-Time PCR based test for the early diagnosis of *Haplosporidium pinnae* affecting fan mussel *Pinna nobilis*. *PLoS ONE* **2019**, *14*, e0212028. [CrossRef] [PubMed]
42. Tiscar, P.G.; Rubino, F.; Fanelli, G.; Paoletti, B.; Della Salda, L. Mass mortality of the fan mussel *Pinna nobilis* in Apulia (Ionian Sea) caused by *Haplosporidium pinnae*. *Rapp. Comm. Int. Pour L'exploration Sci. De La Mer Mediterranée* **2019**, *42*, 30.
43. Carella, F.; Aceto, S.; Pollaro, F.; Miccio, A.; Iaria, C.; Carrasco, N.; Prado, P.; De Vico, G. A mycobacterial disease is associated with the silent mass mortality of the pen shell *Pinna nobilis* along the Tyrrhenian coastline of Italy. *Sci. Rep.* **2019**, *9*, 2725. [CrossRef]
44. Carella, F.; Antuofermo, E.; Farina, S.; Salati, F.; Mandas, D.; Prado, P.; Panarese, R.; Marino, F.; Fiocchi, E.; Pretto, T.; et al. In the Wake of the Ongoing Mass Mortality Events: Co-occurrence of *Mycobacterium*, *Haplosporidium* and Other Pathogens in *Pinna nobilis* Collected in Italy and Spain (Mediterranean Sea). *Front. Mar. Sci.* **2020**, *7*, 48. [CrossRef]
45. Prado, P.; Carrasco, N.; Catanese, G.; Grau, A.; Cabanes, P.; Carella, F.; García-March, J.R.; Tena, J.; Roque, A.; Bertomeu, E.; et al. Presence of *Vibrio mediterranei* associated to major mortality in stabled individuals of *Pinna nobilis* L. *Aquaculture* **2020**, *519*, 734899. [CrossRef]
46. Tiscar, P.G.; Rubino, F.; Paoletti, B.; Di Francesco, C.E.; Mosca, F.; Della Salda, L.; Hattab, J.; Smoglica, C.; Morelli, S.; Fanelli, G. New insights about *Haplosporidium pinnae* and the pen shell *Pinna nobilis* mass mortality events. *J. Invertebr. Pathol.* **2022**, *190*, 107735. [CrossRef] [PubMed]
47. Coupé, S.; Giantsis, I.; Vázquez-Luis, M.; Scarpa, F.; Foulquié, M.; Prévot, J.M.; Lattos, A.; Casu, M.; Michaelidis, B.; Sanna, D.; et al. Toll-like receptor polymorphism in fan mussels: Does the resistance of *Pinna nobilis* to *Haplosporidium pinnae* only depends on its hybridization with *Pinna rudis*? *Authorea Prepr.* **2022**. [CrossRef]
48. Carella, F.; Palić, D.; Šarić, T.; Župan, I.; Gorgoglione, B.; Prado, P.; Andree, K.B.; Giantsis, I.A.; Michaelidis, B.; Lattos, A.; et al. Multipathogen infections and multifactorial pathogenesis involved in noble pen shell (*Pinna nobilis*) mass mortality events: Background and current pathologic approaches. *Vet. Pathol.* **2023**, *60*, 560–577. [CrossRef] [PubMed]
49. Nebot-Colomer, E.; Nebot-Colomer, E.; Vázquez-Luis, M.; Boissin, E.; Peyran, C.; Deudero, S.; Planes, S. Inferred family structure and dispersal patterns of a Critically Endangered species, *Pinna nobilis*, using molecular analyses: Implications for conservation. *Endanger. Species Res.* **2022**, *49*, 87–102. [CrossRef]
50. García-March, J.R.; Tena, J.; Henandis, S.; Vázquez-Luis, M.; López, D.; Téllez, C.; Prado, P.; Navas, J.I.; Bernal, J.; Catanese, G.; et al. Can we save a marine species affected by a highly infective, highly lethal, waterborne disease from extinction? *Biol. Conserv.* **2020**, *243*, 108498. [CrossRef]
51. Sarafidou, G.; Tsaparis, D.; Issaris, Y.; Chatzigeorgiou, G.; Grigoriou, P.; Chatzinikolaou, E.; Pavloudi, C. Insights on *Pinna nobilis* population genetic structure in Greece. *PeerJ* **2023**, *11*, e16491. [CrossRef]
52. Cinar, M.E.; Bilecenoglu, M.; Yokeş, M.B.; Güçlüsoy, H.A.R.U.N. The last fortress fell: Mass mortality of *Pinna nobilis* in the Sea of Marmara. *Mediterr. Mar. Sci.* **2021**, *22*, 669–676. [CrossRef]
53. Katsanevakis, S.; Carella, F.; Çınar, M.; Čižmek, H.; Jiménez, C.; Kersting, D.; Moreno, D.; Rabaoui, L.; Vicente, N. The fan mussel *Pinna nobilis* on the brink of extinction in the Mediterranean. In *Imperiled: The Encyclopedia of Conservation*; DellaSala, D.A., Goldstein, M.I., Eds.; Elsevier: Amsterdam, The Netherlands, 2022; Volumes 1–3.
54. Prado, P.; Grau, A.; Catanese, G.; Cabanes, P.; Carella, F.; Fernández-Tejedor, M.; Andree, K.B.; Añón, T.; Hernandis, S.; Tena, J.; et al. *Pinna nobilis* in suboptimal environments are more tolerant to disease but more vulnerable to severe weather phenomena. *Mar. Environ. Res.* **2021**, *163*, 105220. [CrossRef]
55. Kersting, D.; Benabdi, M.; Cizmek, H.; Grau, A.; Jimenez, C.; Katsanevakis, S.; Oztürk, B.; Tuncer, S.; Tunesi, L.; Vázquez-Luis, M.; et al. *Pinna nobilis* . *IUCN* **2019**, e.T160075998A160081499. [CrossRef]
56. Haberle, I.; Marn, N.; Geček, S.; Klanjšček, T. Dynamic energy budget of endemic and critically endangered bivalve *Pinna nobilis*: A mechanistic model for informed conservation. *Ecol. Modell.* **2020**, *434*, 109207. [CrossRef]
57. Sanna, D.; Dedola, G.; Scarpa, F.; Lai, T.; Cossu, P.; Curini-Galletti, M.; Francalacci, P.; Casu, M. New mitochondrial and nuclear primers for the Mediterranean marine bivalve *Pinna nobilis*. *Mediterr. Mar. Sci.* **2014**, *15*, 416–422. [CrossRef]

58. Wesselmann, M.; González-Wangüemert, M.; Serrão, E.A.; Engelen, A.H.; Renault, L.; García-March, J.R.; Duarte, C.M.; Hendriks, I.E. Genetic and oceanographic tools reveal high population connectivity and diversity in the endangered pen shell *Pinna nobilis*. *Sci. Rep.* **2018**, *8*, 4770. [CrossRef] [PubMed]
59. González-Wangüemert, M.; Basso, L.; Balau, A.; Costa, J.; Renault, L.; Serrão, E.A.; Duarte, C.M.; Hendriks, I.E. Gene pool and connectivity patterns of *Pinna nobilis* in the Balearic Islands (Spain, Western Mediterranean Sea): Implications for its conservation through restocking. *Aquatic Conserv. Mar. Freshw. Ecosyst.* **2019**, *29*, 175–188. [CrossRef]
60. Peyran, C.; Boissin, E.; Morage, T.; Nebot-Colomer, E.; Iwankow, G.; Planes, S. Genetic homogeneity of the critically endangered fan mussel, *Pinna nobilis*, throughout lagoons of the Gulf of Lion (North-Western Mediterranean Sea). *Sci. Rep.* **2021**, *11*, 7805. [CrossRef]
61. Peyran, C.; Boissin, E.; Morage, T.; Nebot-Colomer, E.; Iwankow, G.; Planes, S. Investigating population dynamics from parentage analysis in the highly endangered fan mussel *Pinna nobilis*. *Ecol. Evol.* **2022**, *12*, e8482. [CrossRef] [PubMed]
62. Gonzalez-Wanguemert, M.; Costa, J.; Basso, L.; Duarte, C.M.; Serrão, E.; Hendriks, I. Highly polymorphic microsatellite markers for the Mediterranean endemic fan mussel *Pinna nobilis*. *Mediterr. Mar. Sci.* **2014**, *16*, 31–35. [CrossRef]
63. Costedoat, C.; Gilles, A. Quaternary pattern of freshwater fishes in Europe: Comparative phylogeography and conservation perspective. *Open Conserv. Biol. J.* **2009**, *3*, 36–48. [CrossRef]
64. Machordom, A.; Araujo, R.; Erpenbeck, D.; Ramos, M.Á. Phylogeography and conservation genetics of endangered European Margaritiferidae (Bivalvia: Unionoidea). *Biol. J. Linn. Soc.* **2003**, *78*, 235–252. [CrossRef]
65. Walker, K.F.; Jones, H.A.; Klunzinger, M.W. Bivalves in a bottleneck: Taxonomy, phylogeography and conservation of freshwater mussels (Bivalvia: Unionoida) in Australasia. *Hydrobiologia* **2014**, *735*, 61–79. [CrossRef]
66. Ladoukakis, E.D.; Zouros, E. Evolution and inheritance of animal mitochondrial DNA: Rules and exceptions. *J. Biol. Res.* **2017**, *24*, 2. [CrossRef]
67. Bracchi, V.A.; Bulegato, S.; Basso, D. Palaeoecology of the *Pinna nobilis* biofacies along the Stirone River (Early Pleistocene, Northern Italy). *Boll. Soc. Paleontol. Ital.* **2020**, *59*, 41–55. [CrossRef]
68. Garcia-Castellanos, D.; Estrada, F.; Jiménez-Munt, I.; Gorini, C.; Fernández, M.; Vergés, J.; De Vicente, R. Catastrophic flood of the Mediterranean after the Messinian salinity crisis. *Nature* **2009**, *462*, 778–781. [CrossRef] [PubMed]
69. Bianchi, C.N. Biodiversity issues for the forthcoming tropical Mediterranean Sea. *Hydrobiologia* **2007**, *580*, 7–21. [CrossRef]
70. Sievers, F.; Higgins, D.G. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In *Multiple Sequence Alignment Methods*; Russell, D., Ed.; Humana Press: Totowa, NJ, USA, 2014; Volume 1079, pp. 105–116. [CrossRef]
71. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef] [PubMed]
72. Posada, D. jModelTest: Phylogenetic model averaging. *Mol. Biol. Evol.* **2008**, *25*, 1253–1256. [CrossRef] [PubMed]
73. Tavaré, S. Some probabilistic and statistical problems in the analysis of DNA sequences. In *Some Mathematical Questions in Biology—DNA Sequence Analysis*; Miura, R.M., Ed.; American Mathematical Society: Providence, RI, USA, 1986; pp. 57–86.
74. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef] [PubMed]
75. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]
76. Paradis, E.; Claude, J.; Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004**, *20*, 289–290. [CrossRef]
77. Tran Thi, G.; Azzena, I.; Scarpa, F.; Cossu, P.; Danh Le, C.; Ton Nu, P.A.; Chau Ngo, T.M.; Sanna, D.; Casu, M. Molecular Identification and Appraisal of the Genetic Variation of *Taenia saginata* in Central Regions of Vietnam. *Life* **2022**, *12*, 70. [CrossRef]
78. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]
79. Scarpa, F.; Cossu, P.; Sanna, D.; Lai, T.; Norenburg, J.L.; Curini-Galletti, M.; Casu, M. An 18S and 28S-based clock calibration for marine Proseriata (Platyhelminthes). *J. Exp. Mar. Biol. Ecol.* **2015**, *463*, 22–31. [CrossRef]
80. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef] [PubMed]
81. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
82. Scarpa, F.; Sanna, D.; Cossu, P.; Lai, T.; Casu, M.; Curini-Galletti, M. How to achieve internal fertilization without a vagina: The study case of the genus *Archilina* Ax, 1959 (Platyhelminthes, Proseriata) from Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2057–2073. [CrossRef]
83. Drummond, A.J.; Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **2007**, *7*, e214. [CrossRef] [PubMed]
84. Luttikhuizen, P.C.; Drent, J.; Baker, A.J. Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. *Mol. Ecol.* **2003**, *12*, 2215–2229. [CrossRef] [PubMed]

85. Gernhard, T.J. The conditioned reconstructed process. *J. Theor. Biol.* **2008**, *253*, 769–778. [CrossRef] [PubMed]
86. Yule, G.U. II.—A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Phil. Trans. R. Soc. Lond. B. Biol. Sci.* **1925**, *213*, 21–87.
87. Scarpa, F.; Cossu, P.; Lai, T.; Sanna, D.; Curini-Galletti, M.; Casu, M. Meiofaunal cryptic species challenge species delimitation: The case of the *Monocelis lineata* (Platyhelminthes: Proseriata) species complex. *Contrib. Zool.* **2016**, *85*, 123–145. [CrossRef]
88. Schultz, P.W.; Huber, M. *Revision of the Worldwide Recent Pinnidae and Some Remarks on Fossil European Pinnidae*; Acta Conchyliorum 13; ConchBooks: Harxheim Germany, 2013; pp. 1–164.
89. Bianchi, C.N.; Morri, C.; Chiantore, M.; Montefalcone, M.; Parravicini, V.; Rovere, A. Mediterranean Sea biodiversity between the legacy from the past and a future of change. In *Life in the Mediterranean Sea: A Look at Habitat Changes*, 1st ed.; Stambler, N., Ed.; Bar Ilan University: Ramat Gan, Israel, 2012; pp. 1, 56.
90. Crippa, G.; Azzarone, M.; Bottini, C.; Crespi, S.; Felletti, F.; Marini, M.; Petrizzo, M.R.; Scarponi, S.; Raffi, S.; Raineri, G. Bio-and lithostratigraphy of lower Pleistocene marine successions in western Emilia (Italy) and their implications for the first occurrence of *Arctica islandica* in the Mediterranean Sea. *Quat. Res.* **2019**, *92*, 549–569. [CrossRef]
91. Dominici, S. Taphonomy and paleoecology of shallow marine macrofossil assemblages in a collisional setting (late Pliocene–early Pleistocene, western Emilia, Italy). *Palaios* **2001**, *16*, 336–353. [CrossRef]
92. Pervesler, P.; Uchman, A.; Hohenegger, J.; Dominici, S. Ichnological record of environmental changes in early Quaternary (Gelasian–Calabrian) marine deposits of the Stirone Section, northern Italy. *Palaios* **2011**, *26*, 578–593. [CrossRef]
93. Abril Hernández, J.M.; Periáñez Rodríguez, R. Revisiting the time scale and size of the Zanclean flood of the Mediterranean (5.33 Ma) from CFD simulations. *Mar. Geol.* **2016**, *382*, 242–256. [CrossRef]
94. Van Dijk, G.; Maars, J.; Andreetto, F.; Hernández-Molina, F.J.; Rodríguez-Tovar, F.J.; Krijgsman, W. A terminal Messinian flooding of the Mediterranean evidenced by contouritic deposits on Sicily. *Sedimentology* **2023**, *70*, 1195–1223. [CrossRef]
95. Periáñez Rodríguez, R.; Abril Hernández, J.M.; García Castellanos, D.; Estrada, F.; Ercilla, G. An exploratory modelling study on sediment transport during the Zanclean flood of the Mediterranean. *SN Appl. Sci.* **2019**, *1*, 364. [CrossRef]
96. Peyran, C.; Morage, T.; Nebot-Colomer, E.; Iwankow, G.; Planes, S. Unexpected residual habitats raise hope for the survival of the fan mussel *Pinna nobilis* along the Occitan coast (Northwest Mediterranean Sea). *Endanger. Species Res.* **2022**, *48*, 123–137. [CrossRef]
97. Zander, C.D. Evolution of Blennioidel in the Mediterranean Sea. *Rev. Des Trav. De L'institut Des Pêches Marit.* **1973**, *37*, 215–221.
98. Thunell, R.; Rio, D.; Sprovieri, R.; Vergnaud-Grazzini, C. An overview of the post-Messinian paleoenvironmental history of the western Mediterranean. *Paleoceanography* **1991**, *6*, 143–164. [CrossRef]
99. Kaboth, S.; Grunert, P.; Lourens, L. Mediterranean Outflow Water variability during the early Pleistocene. *Clim. Past* **2017**, *13*, 1023–1035. [CrossRef]
100. Gvozdenović, S.; Mačić, V.; Pešić, V.; Nikolić, M.; Peraš, I.; Mandić, M. Review on *pinna rudis* (Linnaeus, 1758) (bivalvia: Pinnidae) presence in the Mediterranean. *Agric. For.* **2019**, *65*, 115–126. [CrossRef]
101. Kersting, D.K.; Ballesteros, E. Is the local extinction of *Pinna nobilis* facilitating *Pinna rudis* recruitment? *Mediterr. Mar. Sci.* **2021**, *22*, 623–626. [CrossRef]
102. Vázquez-Luis, M.; Nebot-Colomer, E.; Deudero, S.; Planes, S.; Boissin, E. Natural hybridization between pen shell species: *Pinna rudis* and the critically endangered *Pinna nobilis* may explain parasite resistance in *P. nobilis*. *Mol. Biol. Rep.* **2021**, *48*, 997–1004. [CrossRef] [PubMed]
103. Allmon, W.D. Nutrients, temperature, disturbance, and evolution: A model for the late Cenozoic marine record of the western Atlantic. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2001**, *166*, 9–26. [CrossRef]
104. Collins, L.S.; Coates, A.G.; Berggren, W.A.; Aubry, M.P.; Zhang, J. The late Miocene Panama Isthmian strait. *Geology* **1996**, *24*, 687–690. [CrossRef]
105. Jackson, J.B.C.; Budd, A.F.; Coates, A.G. (Eds.) *Evolution and Environment in Tropical America*; University of Chicago Press: Chicago, IL, USA, 1996; Volumes 1–20.
106. Knowlton, N.; Weigt, L.A. New dates and new rates for divergence across the Isthmus of Panama. *Proc. R. Soc. B Biol. Sci.* **1998**, *265*, 2257–2263. [CrossRef]
107. Cornacchia, I.; Brandano, M.; Agostini, S. Miocene paleoceanographic evolution of the Mediterranean area and carbonate production changes: A review. *Earth-Sci. Rev.* **2021**, *221*, 103785. [CrossRef]
108. Smith, A.G.; Smith, D.G.; Funnell, B.M. *Atlas of Mesozoic and Cenozoic Coastlines*; Cambridge University Press: Cambridge, UK, 2004; p. 112.
109. Duque-Caro, H. Neogene stratigraphy, paleoceanography and paleobiogeography in northwest South America and the evolution of the Panama Seaway. *Palaeog. Palaeoclim. Palaeoecol.* **1990**, *77*, 203–234. [CrossRef]
110. Stanley, G.D., Jr. Late Paleozoic and early Mesozoic reef-building organisms and paleogeography: The Tethyan-North American connection. *Cour. Forsch. Inst. Senckenb.* **1994**, *172*, 69–75.
111. Rosewater, J. The family Pinnidae in the Indo-Pacific. *Indo-Pacific Moll.* **1961**, *1*, 175–226.
112. Pauls, S.U.; Nowak, C.; Bálint, M.; Pfenninger, M. The impact of global climate change on genetic diversity within populations and species. *Mol. Ecol.* **2013**, *22*, 925–946. [CrossRef]

113. Wetjen, M.; Hübner, D.; Seehausen, O.; Schulz, R. Genetic diversity of endangered *Chondrostoma nasus* in the River Rhine system: Conservation genetics considerations on stocking and reintroduction. *Knowl. Manag. Aquat. Ecosyst.* **2020**, *421*, 25. [CrossRef]
114. Wetjen, M.; Schmidt, T.; Schrimpf, A.; Schulz, R. Genetic diversity and population structure of burbot *Lota lota* in Germany: Implications for conservation and management. *Fish. Manag. Ecol.* **2020**, *27*, 170–184. [CrossRef]

Chapter 3

Salariopsis fluviatilis

1. General description

The freshwater blenny *Salariopsis fluviatilis* (Asso, 1801) (junior synonym *Ichthyocoris fluviatilis* (Duquenne-Delobel et al. 2022), formerly known as *Salaria fluviatilis*, is a small, benthic freshwater fish belonging to the Actinopterygii class, within the Blenniidae family. This species, originally classified under the genus *Salaria* Forsskål 1775, has been recently reclassified into the genus *Salariopsis* by Vecchioni et al. (2022a). This small freshwater blenny is commonly found in Mediterranean river basins and lakes (Kottelat 2004; Zander, 1986). Its populations exhibit morphological and behavioral variations influenced by predation, competition, and habitat differences, raising the possibility that *Salariopsis fluviatilis* represents a complex of species (Belaiba et al. 2019; Wagner et al. 2021 and references therein). Additionally, two other freshwater blenny species, *Salariopsis economidisi* (Kottelat 2004) and *Salariopsis atlantica* Doadrio, Perea & Yahyaoui, 2011, have been formally described as endemic to specific regions: Lake Trichonis in Greece and the Seboui river basin in Morocco, respectively.

Salariopsis fluviatilis is highly susceptible to anthropogenic disturbances and global environmental changes due to its fragmented habitat. While it is considered endangered in several countries and listed in the Bern Convention, the IUCN classifies it as data deficient due to a lack of information on population abundance and trends. Despite its significance, research on its genetic structure and phylogeography has been limited (Perdices et al. 2000; Almada et al. 2005; Almada et al. 2009; Laporte et al. 2015; Laporte et al. 2016(a); Belaiba et al. 2019; Méndez et al. 2019; Wagner et al. 2021).

1.1. Origin of *Salariopsis fluviatilis*

The genus *Salariopsis* currently comprises freshwater species distributed along the eastern Atlantic coasts, particularly in Morocco, and in regions surrounding the Mediterranean Sea and the Black Sea (Kottelat 2004; Zander 1986). Specifically, *Salariopsis fluviatilis* is commonly found in the river basins and lakes located around the Mediterranean Sea (Zander 1973).

Salariopsis fluviatilis shares a close relationship with its marine counterpart, the peacock blenny *Salaria pavo* (Risso, 1810), which inhabits marine and brackish waters in the Mediterranean Sea and along the European Atlantic coast. Both species exhibit euryhaline behavior, allowing them to coexist in various environments (Zander 1973). *Salaria pavo* can tolerate brackish waters with a salinity as low as 5%, while certain populations of *Salariopsis fluviatilis*, such as those in Le Bourget (France), can endure higher salinity levels of up to 15% (Zander 1973).

Considering the paleohistorical events which involved the Mediterranean area especially during the Quaternary period (Ruggieri 1967), it is unnecessary to consider *Salariopsis fluviatilis* as a polytopic derivative of *Salaria pavo*, as suggested by Kosswig (1967). It is more appropriate to hypothesize a shared ancestor with eurythermal and euryhaline characteristics that not only endured the desalination of the Mediterranean Sea during the upper Pliocene period but also exhibited adaptability to changing conditions (Zander 1967; Plaut 1998).

After the marine influx during the upper Miocene, the marine ancestor embarked on a colonization journey, encompassing rivers, lakes, and the entire Mediterranean Sea, primarily due to its remarkable ecological adaptability (Zander 1973). This adaptability facilitated migrations into freshwater environments at different periods, leading to the emergence of distinct morphological characteristics in isolated populations (Sasse, 1972). *Salariopsis fluviatilis'* ability to thrive in a range of temperatures played a pivotal role in its survival during the glacial periods in the Mediterranean region, possibly within warm spring-fed waters (Kosswig 1967). In contrast, the speciation of *Salaria pavo* likely occurred in the West African refuge, with this species subsequently extending into marine and brackish waters of the Mediterranean Sea after the glacial periods (Zander 1973).

1.2. Biology and ecology

Salariopsis fluviatilis is commonly found in freshwater habitats throughout the Mediterranean region, ranging from Portugal to Israel (Bath 2003; Elvira et al. 2021). These habitats are typically characterized by rivers featuring moderate to fast-flowing sections with a combination of gravel and stones, which are of great significance for their reproductive activities (Freeman et al. 1990; Côté et al. 1999). *Salariopsis fluviatilis* primarily preys on aquatic insects as a benthic predator but it also can capture juvenile fish from the water column (Viñolas, 1986; Freeman et al. 1990). The species exhibits sexual dimorphism, with males characterized by a longer body and larger head compared to females (Roché, 2001; Kottelat & Freyhof 2007; Keith et al. 2011; Laporte et al. 2016). During the breeding season, typically spanning from May to July, male undergo specific changes, developing secondary sexual characteristics such as a cephalic crest and two anal glands covering the first spines of the anal fin (Laporte et al. 2018). During this period, males are actively involved in constructing and safeguarding nests beneath rocks, where multiple females deposit their eggs in a single layer on the undersides of these stones (Fabre et al. 2014). After fertilization, male of *Salariopsis fluviatilis* takes on the exclusive responsibility of caring for the eggs, which involves activities such as fanning and ensuring their protection until they hatch (Wickler, 1957; Fabre et al. 2014; Elvira et al. 2021). Approximately a week later, the eggs give birth to planktonic larvae that initially reside in calm waters near riverbanks before transitioning to a benthic lifestyle (Kottelat & Freyhof, 2007; Gil et al. 2010).

The parental care behavior can be influenced by various environmental factors, like nest availability and interactions with other males (Fabre et al. 2014). In species with resource-defense mating systems, competition among males for mates is primarily dictated by the availability of resources (Almada et al. 1994; Fabre et al. 2014). The limited availability of nesting sites can constrain the number of males that can engage in reproduction (Saraiva et al. 2009). Additionally, the presence of nesting sites is linked to the maturation of young male fish (Takahashi, 2008). When there is an abundance of nesting sites, males are more likely to actively court females, leading to a reduction in the frequency of parasitic reproduction by sneaker males (Almada et al. 1995; Kvarnemo & Ahnesjo 1996; Saraiva et al. 2012). Conversely, when nesting sites are scarce, they become highly coveted resources. In such scenarios, only the most

competitive males succeed in securing a nest for reproduction (Almada et al. 1994; Saraiva et al. 2010).

2. *Salariopsis fluviatilis* distribution in Peninsular Italy and Sardinia Island

In Italy, *Salariopsis fluviatilis* populations were reported starting from the northern regions and extending along the Tyrrhenian coast to Campania, Sardinia and Sicilian islands. Additionally, isolated populations have been identified in Calabria and on the Adriatic coasts of the Italian peninsula (Zava & Violani 1991; Gallo et al. 2012).

The earliest documented report of *Salariopsis fluviatilis* in Sardinia dates back to Vinciguerra in (1895). Nowadays *Salariopsis fluviatilis* exhibits a patchy and discontinuous distribution in Sardinia. Rare populations with low densities of 1-2 individuals per 100 meters have been found in the Riu Flumineddu (Central-southern Sardinia), located upstream of the Cedrino Dam, and in the Riu Girasole (Central-eastern Sardinia), which is also situated upstream of an artificial reservoir, the Santa Lucia Dam. Relatively larger populations with 3-10 individuals per 100 meters have been observed in the Riu Mannu di Lodè (Central Sardinia) and the Coghinas River (North-western Sardinia). However, more structured, and common populations have been documented in the stretch of the Flumendosa River (Central-southern Sardinia) between the municipal territories of Seulo and Gadoni and in the lower stretch near the mouth of the Riu Mannu di Scano di Montiferro (Western Sardinia).

Between 1940 and 1968, *Salariopsis fluviatilis* was considered abundant in various basins on the island. However, its distribution has significantly declined, especially since 1968 (Carta Ittica della Sardegna 2022). The recent analysis of samples collected during the construction of the Carta Ittica della Sardegna (2022) has revealed a substantial contraction in the species' distribution area, with a detection percentage dropping from 25% to 4%. The decline in the distribution of *Salariopsis fluviatilis* is attributed to various factors, including habitat fragmentation, pollution, habitat alteration, and the introduction of alien species. These threats have had a significant impact on the species presence and abundance in Sardinian waters.

3. Threats to freshwater biodiversity

Freshwater habitats support over 10,000 fish species, which accounts for approximately 40% of the world's fish diversity (Dudgeon et al. 2006). At the same time, freshwater ecosystems are among the most endangered in the world, and they have experienced more significant declines in biodiversity compared to the most affected terrestrial ecosystems (Sala et al. 2000).

The decline of freshwater biodiversity is primarily attributed to various factors, including overexploitation, water pollution, flow modification, habitat degradation, and the introduction of competitor species (Dudgeon et al. 2006). Among freshwater ecosystems, Mediterranean rivers face challenges due to the long history of human impact on their aquatic species (Maceda-Veiga & Sostoa 2011).

In this context, the freshwater blenny *Salariopsis fluviatilis* is classified as endangered or vulnerable in several countries, including Spain, France, Italy, Croatia and Turkey, due to their sensitivity to factors such as water pollution, the introduction of competitor species, alterations in water velocity, changes in substrate composition, and severe drought (Laporte et al. 2014).

Even in the relatively well-preserved Corsican rivers, where *Salariopsis fluviatilis* populations exhibit greater genetic diversity (Laporte 2012; Roché 2001), this species is currently facing various threats, including recent introduction of multiple freshwater fish species (Gauthier and Berrebi 2007; Roché 2001) and the expansion of animal-based agriculture and tourism activities (Tempier et al. 2012). The study carried out by Laporte et al. (2014) reinforced the association between *Salariopsis fluviatilis* and the presence of higher water velocity, which was consistent with the findings of Freeman et al. (1990). Therefore, one of the significant contributors to the decrease in the abundance of *Salariopsis fluviatilis* populations is the construction of hydroelectric dams (Aparicio et al. 2000; Collares-Pereira et al. 2000; Ferrito & Tigano 1996).

3.1. The case of Sardinia Island

In this context, Sardinia Island deserves particular attention, as *Salariopsis fluviatilis* is one of the few putative native freshwater fish species known for this Mediterranean island, whose presence is threatened by the introduction of numerous alien species and alterations to freshwater habitats (Orrù et al. 2010).

The earliest records of the fish fauna in Sardinia freshwaters date back to the late 18th century (Cetti 1774), where a relatively low diversity of native fish species in Sardinia were indicated, likely due to its insular nature. Even if this report were not highly detailed in terms of taxonomies, the author confirms the presence of two migratory species in the inland waters, *Anguilla anguilla* (European eel) and *Alosa fallax* (twaite shad), together with one species of trout.

The first documented report of *Salariopsis fluviatilis* in Sardinia can be traced back to Vinciguerra in (1895). The author not only identified the presence of the species in Sardinia but also proposed introducing salmonids and cyprinids to enhance wild fish stocks. In subsequent years, particularly in the early 1900s, additional species were recognized as native to the island: *Atherina boyeri* (bigscale sand smelt), *Gasterosteus aculeatus* (three-spined stickleback), *Petromyzon marinus* (sea lamprey), and *Salmo ghigii* (native trout).

The introduction of non-native species to Sardinia's freshwater ecosystems began in 1895. This initial introduction included species such as *Oncorhynchus mykiss* (rainbow trout) and *Salmo trutta* (brown trout). Subsequently, in 1907, *Tinca tinca* (tench) was introduced, followed by *Cyprinus carpio* (common carp) and *Perca fluviatilis* (European perch) in 1956 (Spano 1956). Additionally, other non-autochthonous species like *Gambusia holbrooki* (mosquitofish), *Carassius auratus* (goldfish), *Lepomis gibbosus* (pumpkinseed), *Cobitis taenia* (spined loach), *Alburnus alburnus alborella* (bleak), *Scardinius erythrophthalmus* (rudd), *Micropterus salmoides* (largemouth bass) or *Pseudorasbora parva* (topmouth gudgeon) were introduced intentionally or unintentionally into the inland waters of Sardinia Island during the last century (Ronchetti 1968; Copp et al. 2005; Orrù et al. 2010).

Some of these species were aquarium pets (e.g., *Carassius auratus*) that were released into the wild, while others were used as baitfish by anglers (e.g., *Cobitis taenia*, *Scardinius erythrophthalmus*) (Orrù et al. 2010). The public's limited awareness of the

regulations regarding the movement and release of alien fish species contributed to their continued introduction into Sardinian freshwater environments (Orrù et al. 2010).

Additionally, from the late 1960s to the late 1990s, there was a noticeable surge in the construction of dams. To date, Sardinia hosts a total of 54 dams, with two more recent dams that are currently in the process of construction and have not been fully completed (Podda et al. 2022). The presence of dams, can physically disrupt river connectivity, reduce, or alter river flow, and lead to the loss and degradation of habitats (Podda et al. 2022). Moreover, in Sardinia, the river discharge is strongly influenced by rainfall that typically ranges from moderate to high during the winter but decreases significantly in the summer. Consequently, in late spring, summer, and early autumn, many riverbeds dry up (Orrù et al. 2010).

Summarizing, the environmental alterations, and hydro-morphological pressures on Sardinian aquatic habitats (e.g., the construction of dams), have led to a situation where freshwaters habitats are significantly different from their natural forms. These alterations have created highly disturbed aquatic systems, making it challenging for native species to survive (Orrù et al. 2010). In contrast, ecological generalist alien fish species introduced to Sardinia have adapted well to these highly disturbed and altered habitats, which are also characterized by poor native fish communities under stress (Massidda & Orrù 2004; Massidda et al. 2006; Orrù et al. 2007; Massidda et al. 2008; Massidda et al. 2008).

4. Taxonomic review of the genus: from *Salaria* to *Salariopsis*

The combtooth blennies (Blenniidae Rafinesque 1810), represent a highly diverse group of fish, with more than 400 species falling under this family (Fricke et al. 2018).

The reorganization of the family's taxonomy resulted in the creation of six tribes based on morphological characteristics: Salariini, Omobranchini, Phenablenniini, Nemophini, Parablenniini, and Blenniini (Vecchioni et al. 2022 (a)). These divisions have remained mostly stable, except for the Parablenniini and Blenniini tribes, which have sparked significant disagreements, especially after Zander rejected the newly proposed genera introduced by Bath (Bath, 1977; Zander, 1978).

The advent of molecular systematics has brought about significant changes to our understanding of the taxonomy within this group. Notably, the Almadablennius clade encompassing Parablenniini and Blenniini, has received considerable attention (Hundt et al. 2014; Vecchioni et al, 2022(a)). This attention intensified after that Almada et al. (2005) published a phylogeny in which Blenniini sensu Williams was nested within Parablennini sensu Williams. Despite these advancements and multiple lines of evidence pointing to issues, the taxonomy of the Almadablennius clade remains unresolved (Vecchioni et al. 2022 (a)).

In this context, in 2022 Vecchioni and colleagues conducted a re-examination of the phylogenetic relationships within the Almadablennius utilizing partial sequences of two nuclear and two mitochondrial loci, with the primary aim to assess the monophyly of the genus *Salaria* and propose revisions to its taxonomy (Vecchioni et al. 2022 (a)). The results obtained by Vecchioni et al. (2022a) strongly supports previous findings indicating a significant differentiation between marine and freshwater species currently categorized within the genus *Salaria* (Hundt et al. 2014; Vecchioni et al. 2019). This distinction in the genetic makeup of freshwater blennies compared to their marine counterparts clearly highlights the inadequacy of their current generic classification and emphasizes the need for a thorough reassessment (Vecchioni et al. 2022 (a)).

Given that the type species of the genus *Salaria* is *Salaria basilisca* (Valenciennes, 1836), marine species rightfully remain under *Salaria* s.s. However, because there is no existing genus-level name available for the distinct freshwater clade, Vecchioni et al. (2022a) proposes a practical solution by introducing a new genus, *Salariopsis*. This newly established genus encompasses the species *Salariopsis fluviatilis*, *Salariopsis economidisi*, and *Salariopsis atlantica* (Vecchioni et al. 2022 (a)).

Thereafter, Duquenne-Delobel et al. (2022) concurred with the idea of segregating the freshwater and marine species of *Salaria* into two distinct genera (Vecchioni et al. 2022(b)). They also autonomously introduced additional morphological diagnostic characteristics for these genera (Duquenne-Delobel et al. 2022). However, they suggest the revalidation of the genus *Ichthyocoris* Bonaparte, 1840 for the freshwater species, citing precedence over the name *Salariopsis*, based on Jordan (1919) and subsequent literature (e.g., Norman 1943; Springer 1968; Bath 1977).

Nevertheless, Vecchioni et al. (2022b) explain that Jordan (1919) listed "*Blennius varus* Risso" as the type species of *Ichthyocoris*, and Bonaparte (1840) explicitly stated that the type species of *Ichthyocoris* genus is *Blennius pavo*. Therefore, *Ichthyocoris* is considered a junior synonym of *Salaria*, and *Salariopsis* remains the valid name for the freshwater blennies previously classified under *Salaria*, thus including *Salariopsis fluviatilis*, *Salariopsis economidisi*, and *Salariopsis atlantica* (Vecchioni et al. 2022(b)).

5. Is *Salariopsis fluviatilis* a species complex?

Several studies, including those conducted by Perdices et al. (2000), Almada et al. (2009), Belaiba et al. (2019), Wagner et al. (2021) consistently report the presence of genetic structuring within *Salariopsis fluviatilis* populations across Mediterranean regions, also evidenced through distinct distribution patterns (Belaiba et al. 2019; Wagner et al. 2021). Specifically, Belaiba et al. (2019) performed a comprehensive phylogenetic study of the genus *Salaria*. It's important to note that at the time of this study, the genus *Salaria* also included the freshwater species that are now classified under the new genus *Salariopsis*. The study was based on both mitochondrial and nuclear markers and played a significant role in clarifying the taxonomy and evolutionary history of these species. According to their findings, the freshwater species within the *Salaria* clade are considered vicariant taxa originating from a common ancestor. Belaiba et al. (2019) proposed that these freshwater species likely spread throughout the circum-Mediterranean inland waters during the late Miocene Messinian salinity crisis, and, after the re-flooding of the Mediterranean basin, these freshwater species underwent a process of allopatric differentiation (Belaiba et al. 2019). Furthermore, based on two different methods of species delimitation, they suggested referring to the population in the Middle East Mediterranean region as *Salaria* cf *fluviatilis*.

Two years later, Wagner et al. (2021) provided support for the existence of at least six distinct genetic lineages within *Salariopsis fluviatilis* populations: (i) Middle East (encompassing Israel, Syria, and southern Turkey); (ii) Guadiana (including the Guadiana basin with the Zùjar and Esteras rivers on the European Atlantic coast); (iii) Algeria-Verde (covering southern Spain and Algeria, with the Verde River and Oued

Boughzazene); (iv) Occidental basin (spanning Algeria, Spain, mainland France, Corsica, and Switzerland); (v) North Oriental basin (encompassing Italy, Croatia, Albania, Greece, and western Turkey) and; (vi) The Island of Crete (Greece).

These results have led to the hypothesis that *Salariopsis fluviatilis* may constitute a complex of species (Belaiba et al. 2019, Wagner et al. 2021, and related references). In scenarios where different freshwater populations inhabit separate watersheds, their dispersal into different habitats and subsequent evolution can follow three main patterns: (i) single colonization/dispersal; (ii) multiple colonization from freshwater ancestor and, (iii) multiple colonization from marine ancestor (Perdices et al. 2000).

First, a single invasion of a marine ancestor into freshwater may form a monophyletic group constituted by various freshwater populations or species. If these fish retain a salinity tolerance, they could continue to disperse among river systems via the sea, leading to the development of genetically distinct populations or species (Perdices et al. 2000). For example, this is observed in some euryhaline diadromous *Galaxiidae* species, where certain members have become resident in freshwater (Bănărescu 1990; Allibone et al. 1996).

Second, monophyletic populations might originate from a freshwater ancestor, potentially due to a temporary drop in the salinity of a marine water body, such as eustatic sea level changes (Perdices et al. 2000). This would result in isolation from the sea, allowing the freshwater species to colonize rivers and lakes. However, they would be unable to disperse via the sea once marine conditions are restored, leading to isolation from the sea (Perdices et al. 2000). Here, the geographic distance between populations would not correlate with their genetic distance (Perdices et al. 2000).

Finally, the divergence of freshwater forms may involve a marine ancestor that has colonized freshwater habitats on multiple occasions, a scenario well-supported by evidence in various species, such as sticklebacks (*Gasterosteus aculeatus*), lampreys, and several salmonids, including *Oncorhynchus nerka* (Vladykov & Kott 1979; Withler & McPhail 1985; Bănărescu 1990).

6. Aims of the study

Based on the above background, during my PhD project, I focused my attention on *Salariopsis fluviatilis* populations belonging to Italian regions previously overlooked to address the knowledge gap in the central Mediterranean region. This area encompassed Sardinia Island and the north-western regions of Italy, which included Liguria, Piedmont, and Lombardy.

The aims of the study were different. First, I focused to determine the phylogeographic relationship of our samples in comparison to those from other Mediterranean border regions, including countries like France and Spain, as well as other Mediterranean regions. Second, I compared the genetic structure of the present samples with the ones belonging to other populations present in GenBank, to place my dataset into a broader geographic context. Last, I sought to contribute to the ongoing debate regarding the possible occurrence of a complex of species for *Salariopsis fluviatilis* by employing different methods of species delimitation to the enlarged dataset.

References

- Allibone, R. M., Crowl, T. A., Holmes, J. M., King, T. M., McDowall, R. M., Townsend, C. R., & Wallis, G. P. (1996). Isozyme analysis of *Galaxias* species (Teleostei: Galaxiidae) from the Taieri River, South Island, New Zealand: a species complex revealed. Biological journal of the Linnean Society, 57(2), 107-127.
- Almada, F., Almada, V. C., Guillemaud, T., & Wirtz, P. (2005). Phylogenetic relationships of the north-eastern Atlantic and Mediterranean blenniids. Biological Journal of the Linnean Society, 86(3), 283-295.
- Almada, V. C., Gonçalves, E. J., Oliveira, R. F., & Santos, A. J. (1995). Courting females: ecological constraints affect sex roles in a natural population of the blenniid fish *Salaria pavo*. Animal Behaviour, 1125-1127.
- Almada, V. C., Gonçlalves, E. J., Santos, A. J., & Baptista, C. (1994). Breeding ecology and nest aggregations in a population of *Salaria pavo* (Pisces: Blenniidae) in an area where nest sites are very scarce. Journal of Fish Biology, 45(5), 819-830.
- Almada, V. C., Robalo, J. I., Levy, A., Freyhof, J., Bernardi, G., & Doadrio, I. (2009). Phylogenetic analysis of Peri-Mediterranean blennies of the genus *Salaria*: Molecular insights on the colonization of freshwaters. Molecular Phylogenetics and Evolution, 52(2), 424-431.
- Aparicio, E., Vargas, M. J., Olmo, J. M., & de Sostoa, A. (2000). Decline of native freshwater fishes in a Mediterranean watershed on the Iberian Peninsula: a quantitative assessment. Environmental Biology of Fishes, 59, 11-19.
- Bănărescu P. (1990). Croizat's biogeographical principles, Pangea and freshwater zoogeography. Crisia (Oradea), 15, 565-577.
- Bath, H. (1977). Revision der Blenniini (Pisces: Blenniidae). Senckenbergiana biologica, 57, 167–234.
- Bath, H. (2003). *Salaria fluviatilis* (Asso, 1784). The freshwater fishes of Europe, 8, 1.
- Belaiba, E., Marrone, F., Vecchioni, L., Bahri-Sfar, L., & Arculeo, M. (2019). An exhaustive phylogeny of the combtooth blenny genus *Salaria* (Pisces, Blenniidae) shows introgressive hybridization and lack of reciprocal mtDNA monophyly between the marine species *Salaria basilisca* and *Salaria pavo*. Molecular phylogenetics and evolution, 135, 210-221.
- Carta Ittica della Sardegna (2022). Volume I – Parte Generale, Acqua dolci. Regione autonoma della Sardegna. pp. 44
- Cetti, F. (1774). I Quadrupedi, gli Uccelli, Anfibi e Pesci di Sardegna. Ed. Piattoli, Sassari, pp. 208.
- Collares-Pereira, M. J., Cowx, I. G., Ribeiro, F., Rodrigues, J. A., & Rogado, L. (2000). Threats imposed by water resource development schemes on the conservation of endangered fish species in the Guadiana River Basin in Portugal. Fisheries Management and Ecology, 7(1-2), 167-178.
- Copp, G. H., Bianco, P. G., Bogutskaya, N. G., Erős, T., Falka, I., Ferreira, M. T., ... & Wiesner, C. (2005). To be, or not to be, a non-native freshwater fish?. Journal of Applied Ichthyology, 21(4), 242-262.
- Côté, I. M., Vinyoles, D., Reynolds, J. D., Doadrio, I., & Perdices, A. (1999). Potential impacts of gravel extraction on Spanish populations of river blennies *Salaria fluviatilis* (Pisces, Blenniidae). Biological Conservation, 87(3), 359-367.
- Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Lévêque, C., ... & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. Biological reviews, 81(2), 163-182.
- Duquenne-Delobel, E., Doadrio, I., & Denys, G. P. (2022). Revalidation of the genus *Ichthyocoris* Bonaparte, 1840 (Actinopterygii: Blenniiformes: Blenniidae). Acta Ichthyologica et Piscatoria, 52(1), 35-41.

- Elvira, B., Nicola, G. G., Ayllon, D., & Almodovar, A. (2021). Determinants of large-scale spatial distribution and seasonal microhabitat selection patterns of the endangered freshwater blenny *Salaria fluviatilis* in the Ebro River basin, Spain. Aquatic Conservation: Marine and Freshwater Ecosystems, 31(11), 3261-3275.
- Fabre, N., Oliva, F., García-Galea, E., & Vinyoles, D. (2014). Plasticity in secondary sexual characteristics in male freshwater blennies (*Salaria fluviatilis*). Canadian Journal of Zoology, 92(6), 537-543.
- Ferrito, V., & Tigano, C. (1996). Decline of *Aphanius fasciatus* (Cyprinodontidae) and *Salaria fluviatilis* (Blenniidae) populations in freshwaters of eastern Sicily. Ichthyological Exploration of Freshwaters, 7(2), 181-184.
- Freeman, M. C., Viñolas, D., Grossman, G. D., & De Sostoa, A. D. O. L. F. O. (1990). Microhabitat use by *Blennius fluviatilis* in the Rio Matarrana, Spain. Freshwater Biology, 24(2), 335-345.
- Fricke, R., Eschmeyer, W. N., & Van der Laan, R. (2018). Catalog of fishes: genera, species, references. California Academy of Sciences, San Francisco, CA, USA http://researcharchive. calacademy. org/research/ichthyology/catalog/fishcatmain. asp.
- Gallo, L., Lucadamo, L., Mezzotero, A., Morisi, A., Battegazzore, M., & Fenoglio, S. (2012). First data on the freshwater fish fauna of Calabria (southern Italy). Italian Journal of Zoology, 79(2), 246-251.
- Gauthier, A., & Berrebi, P. (2007). La Colonisation de l'Ile par les différentes Souches de Truite.
- Gil, F., Faria, C., & Carvalho Almada, V. (2010). An efficient technique for the captive breeding of an endangered freshwater fish *Salaria fluviatilis* (Pisces: Blenniidae), with a description of its ontogeny. Journal of the World Aquaculture Society, 41, 49-56.
- Hundt, P. J., Iglésias, S. P., Hoey, A. S., & Simons, A. M. (2014). A multilocus molecular phylogeny of combtooth blennies (Percomorpha: Blennioidei: Blenniidae): multiple invasions of intertidal habitats. Molecular Phylogenetics and Evolution, 70, 47-56.
- Jordan, D. S. (1919). The genera of fishes. Part II. Leland Stanford Junior Univ. Pub. Univ Ser., ix. pp. 163-284.
- Keith, P., Persat, H., Feunteun, E., Adam, B., & Geniez, M. (2011). Les Poissons d'eau douce de France (Muséum National d'Histoire Naturelle and Publications Biotope). Paris and Mèze.
- Kosswig, C. U. R. T. (1967). Tethys and its relation to the peri-Mediterranean faunas of freshwater fishes. Syst. Assoc. Publ, 7, 313-324.
- Kottelat, M. (2004). *Salaria economidisi*, a new species of freshwater fish from Lake Trichonis, Greece, with comments on variation in *S. fluviatilis* (Teleostei: Blenniidae). Revue Suisse de Zoologie, 111(1), 121-137.
- Kottelat, M., & Freyhof, J. (2007). Handbook of European freshwater fishes. Cornol, Switzerland: Publications Kottelat.
- Kvarnemo, C., & Ahnesjo, I. (1996). The dynamics of operational sex ratios and competition for mates. Trends in Ecology & Evolution, 11(10), 404-408.
- Laporte, M. (2012). Histoire évolutive et phylogéographie de la colonisation des eaux douces par une espèce d'origine marine, la blennie fluviatile (*Salaria fluviatilis*) (Doctoral dissertation, Doctoral dissertation, Université Montpellier, Montpellier, France).
- Laporte, M., Berrebi, P., Claude, J., Vinyoles, D., Pou-Rovira, Q., Raymond, J. C., & Magnan, P. (2018). The ecology of sexual dimorphism in size and shape of the freshwater blenny *Salaria fluviatilis*. Current zoology, 64(2), 183-191.
- Laporte, M., Bertolo, A., Berrebi, P., & Magnan, P. (2014). Detecting anthropogenic effects on a vulnerable species, the freshwater blenny (*Salaria fluviatilis*): the importance of considering key ecological variables. Ecological indicators, 36, 386-391.

- Laporte, M., Claude, J., Berrebi, P., Perret, P., & Magnan, P. (2016). Shape plasticity in response to water velocity in the freshwater blenny *Salaria fluviatilis*. Journal of Fish Biology, 88(3), 1191-1203.
- Laporte, M., Claude, J., Berrebi, P., Perret, P., & Magnan, P. (2016b). Shape plasticity in response to water velocity in the freshwater blenny *Salaria fluviatilis*. Journal of Fish Biology, 88(3), 1191-1203.
- Laporte, M., Leblois, R., Coulon, A., Bonhomme, F., Magnan, P., & Berrebi, P. (2015). Genetic structure of a vulnerable species, the freshwater blenny (*Salaria fluviatilis*). Conservation Genetics, 16, 571-581.
- Laporte, M., Perrier, C., Magnan, P., & Berrebi, P. (2016a). Genetic evidence of recent migration among isolated-by-sea populations of the freshwater blenny (*Salaria fluviatilis*). Conservation genetics, 17, 389-399.
- Maceda-Veiga, A., & De Sostoa, A. (2011). Observational evidence of the sensitivity of some fish species to environmental stressors in Mediterranean rivers. Ecological Indicators, 11(2), 311-317.
- Magnan, P., Proulx, R., Berrebi, P., Blondel, J., Perret, P., & Roché, B. (2014). Morphological variation in the freshwater blenny *Salaria fluviatilis* from Corsican rivers: adaptive divergence, phenotypic plasticity or both?. Journal of Fish Biology, 84(1), 31-44.
- Massidda P., Orrù F., Davini M.A., Pirastu A., Cau A. (2006). Introductions of non-native fishes in Sardinia: first occurrence of *Carassius auratus* (L. 1758) in Lake Coghinas (Sardinia, Italy). Journal of Freshwater Biology, Quaderni ETP/34/, 299-303.
- Massidda, P., Conti, G., Loddo, G., & Cau, A. (2008). Pesci d'acqua dolce della Sardegna. Aìsara. Cagliari, Italia, pp. 93.
- Massidda, P., Orrù, F. (2004). Bacini idrici della Sardegna: laghi artificiali e traverse a maggiore impatto sulla libera circolazione della fauna ittica. In: Passaggi per Pesci Atti del Seminario Tecnico di Modena. E. Pini Prato, M. Gianaroli. (Eds). Provincia di Modena, Modena, pp. 67–74.
- Méndez, L., Perdices, A., & Machordom, A. (2019). Genetic structure and diversity of the Iberian populations of the freshwater blenny *Salaria fluviatilis* (Asso, 1801) and its conservation implications. Conservation Genetics, 20(6), 1223-1236.
- Neat, F. C., Lengkeek, W., Westerbeek, E. P., Laarhoven, B., & Videler, J. J. (2003). Behavioural and morphological differences between lake and river populations of *Salaria fluviatilis*. Journal of Fish Biology, 63(2), 374-387.
- Norman, J. R. (1943). LXVIII.—Notes on the Blennioid Fishes.—I. A provisional Synopsis of the Genera of the Family Blenniidæ. Annals and Magazine of natural History, 10(72), 793-812.
- Orrù, F., Davini, M. A., Mura, M., Cau, A. (2007). First record of the North American sunfish, pumpkinseed *Lepomis gibbosus*, in Sardinian freshwaters. In: Biological Invasions – from Ecology to Conservation, Neobiota 7, W. Rabitsch, F. Essl, F. Klingenstein (Eds). Berlin, Germany, pp. 139–142.
- Orrù, F., Deiana, A. M., & Cau, A. (2010). Introduction and distribution of alien freshwater fishes on the island of Sardinia (Italy): an assessment on the basis of existing data sources. Journal of Applied Ichthyology, 26, 46-52.
- Perdices, A., Doadrio, I., Côté, I. M., Machordom, A., Economidis, P., & Reynolds, J. D. (2000). Genetic divergence and origin of Mediterranean populations of the river blenny *Salaria fluviatilis* (Teleostei: Blenniidae). Copeia, 2000(3), 723-731.
- Plaut, I. (1998). Comparison of salinity tolerance and osmoregulation in two closely related species of blennies from different habitats. Fish Physiology and Biochemistry, 19(2), 181-188.
- Podda, C., Palmas, F., Pusceddu, A., & Sabatini, A. (2022). When the eel meets dams: Larger dams' long-term impacts on Anguilla anguilla (L., 1758). Frontiers in Environmental Science, 10, 876369.

- Roché, B. (2001). Atlas des poissons d'eau douce de Corse. Direction régionale de l'environnement Corse.
- Roché, B. (2001). Atlas des poissons d'eau douce de Corse. Direction régionale de l'environnement Corse. Imprimerie Bastiaise, Bastia, Corse.
- Ronchetti, G. (1968). L'azione antianofelica dei pesci del genere *Gambusia*, utilizzati per la lotta biologica contro la malaria. Natura, 58, 25-41.
- Ruggieri, G. (1967). The Miocene and later evolution of the Mediterranean Sea. Aspects of Tethyan biogeography, 7, 283-290.
- Sala, O. E., Stuart Chapin, F. I. I. I., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., ... & Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. science, 287(5459), 1770-1774.
- Saraiva, J. L., Barata, E. N., Canário, A. V., & Oliveira, R. F. (2009). The effect of nest aggregation on the reproductive behaviour of the peacock blenny *Salaria pavo*. Journal of Fish Biology, 74(4), 754-762.
- Saraiva, J. L., Gonçalves, D. M., & Oliveira, R. F. (2010). Environmental modulation of androgen levels and secondary sex characters in two populations of the peacock blenny *Salaria pavo*. Hormones and Behavior, 57(2), 192-197.
- Saraiva, J. L., Pignolo, G., Gonçalves, D., & Oliveira, R. F. (2012). Interpopulational variation of the mating system in the peacock blenny *Salaria pavo*. Acta ethologica, 15, 25-31.
- Sasse, H. (1972). Untersuchungen zur Morphologie und Zoogeographie verschiedener Populationen des *Blennius fluuiatilis* Asso 1784 im Mittelmeerraum (Pisces, Perciformes, Blennioidei). — (Unpublished).
- Spano, B. (1956). La pesca fluvio-lacustre in Sardegna. Bollettino della Società Geografica Italiana 7, 5–8.
- Springer, V. G. (1968). Osteology and classification of the fishes of the family Blenniidae. Bulletin of the United States National Museum. pp- 1-97.
- Takahashi, D. (2008). Life-history variation in relation to nest site abundance in males of the freshwater goby *Tridentiger brevispinis*. Ecology of Freshwater Fish, 17(1), 71-77.
- Tempier, A., Tirroloni, A., Apicella, L., Bourgeois, A., Gallard, E., Lepidi, A. (2012). L'impact économique du port de Bastia. In: Les dossiers de l'économie Corse.Direction régionale de Corse, France.
- Vecchioni, L., Arculeo, M., Hundt, P. J., & Marrone, F. (2022b). The valid genus name of the European freshwater blennies, *Ichthyocoris* or *Salariopsis* (Teleostei: Blenniidae). Zootaxa, 5162(1), 99-100.
- Vecchioni, L., Ching, A. C., Marrone, F., Arculeo, M., Hundt, P. J., & Simons, A. M. (2022a). Multi-locus phylogenetic analyses of the Almadablennius clade reveals inconsistencies with the present taxonomy of blenniid fishes. Diversity, 14(1), 53.
- Vecchioni, L., Marrone, F., Belaiba, E., Tiralongo, F., Bahri-Sfar, L., & Arculeo, M. (2019). The DNA barcoding of Mediterranean combtooth blennies suggests the paraphyly of some taxa (Perciformes, Blenniidae). Journal of Fish Biology, 94(2), 339-344.
- Vinciguerra, D. (1895). Indagini praticate in Sardegna per determinare la possibilità di ripopolarne le acque dolci e di esercitarvi l'ostricultura. Bollettino delle Notizie Agrarie, Ministero dell'Agricoltura e dell'Industria Commerciale, 17(1), 22.
- Viñolas, D. (1986). Biologia i ecologia de *Blennius fluviatilis* (Asso, 1801) en el riu Matarranya. Tesina, Department de Zoologia (Vertebrats), Universitat de Barcelona.
- Vinyoles, D., Cote, I. M., & De Sostoa, A. (2002). Nest orientation patterns in *Salaria fluviatilis*. Journal of Fish Biology, 61(2), 405-416.
- Vladykov, V. D., & Kott, E. (1979). Satellite species among the holarctic lampreys (Petromyzonidae). Canadian Journal of Zoology, 57(4), 860-867.
- Wagner, M., Zogaris, S., Berrebi, P., Freyhof, J., Koblmüller, S., Magnan, P., & Laporte, M. (2021). Diversity and biogeography of Mediterranean freshwater blennies (Blenniidae, Salaria). Diversity and Distributions, 27(9), 1832-1847.

- Wickler, W. (1957). Beiträge zur Biologie, besonders zur Ethologie von *Blennius fluviatilis* Asso im Vergleich zu einigen anderen Bodenfischen.
- Withler, R. E., & McPhail, J. D. (1985). Genetic variability in freshwater and anadromous sticklebacks (*Gasterosteus aculeatus*) of southern British Columbia. Canadian Journal of Zoology, 63(3), 528-533.
- Zander, C. D. (1973). Evolution of Blennioidel in the Mediterranean Sea. Revue des Travaux de L'institut des Pêches Maritimes, 37(2), 215-221.
- Zander, C. D. (1978). Kritische Anmerkungen zur "Revision der Blenniini (Pisces: Blenniidae)" von H. Bath (1977). Journal of Zoological Systematics and Evolutionary Research, 16(4), 290-296.
- Zander, C.D. (1986). Blenniidae. In: Fishes of the North-eastern Atlantic and the Mediterranean= Poissons de l'Atlantique du Nord-est et de la Méditerranée. Whitehead, P. J. P. eds.
- Zava, B., & Violani, C. (1991). Contributi alla conoscenza dell'ittiofauna delle acque interne siciliane. I. Sulla presenza in Sicilia di *Salaria fluviatilis* (Asso, 1801)(Pisces, Blenniidae). Bollettino del Museo Regionale di Scienze Naturali Torino, 9, 313-324.

Research article

animals

MDPI

Article

Mitochondrial DNA of Sardinian and North-West Italian Populations Revealed a New Piece in the Mosaic of Phylogeography and Phylogeny of *Salariopsis fluviatilis* (Blenniidae)

Ilenia Azzena [1,2,†], Fabio Scarpa [1,†], Chiara Locci [1,2], Piero Cossu [2], Alessio Niffoi [2], Flavio Orrù [3], Stefano Bovero [4], Giuseppe Sotgiu [4], Daria Sanna [1,*,‡] and Marco Casu [2,‡]

[1] Dipartimento di Scienze Biomediche, Università di Sassari, Viale San Pietro 43/B, 07100 Sassari, Italy
[2] Dipartimento di Medicina Veterinaria, Università di Sassari, Via Vienna 2, 07100 Sassari, Italy
[3] ENAS Ente Acqua della Sardegna, Servizio Qualità Acqua Erogata, 09123 Cagliari, Italy
[4] Independent researchers "Zirichiltaggi" Sardinia Wildlife Conservation NGO, 07100 Sassari, Italy
* Correspondence: darsanna@uniss.it
† These authors equally contributed to this work.
‡ These researchers are co-last authors of this manuscript.

Simple Summary: The present study provides new molecular data on populations of the freshwater blenny *Salariopsis fluviatilis* from Italian geographic areas (north-western regions and Sardinia Island) so far never investigated and uses five species delimitation methods in order to shed further light on the Mediterranean phylogeography of this fish and give a more comprehensive scenario of its taxonomic status. Our findings clarified the phylogeographical patterns of the northern Italian populations of *Salariopsis fluviatilis sensu stricto*, considering the Po River basin and some Tyrrhenian basins of the Liguria region on the other side of the Ligurian Alps. The dispersal pathways of the island of Sardinia were also investigated suggesting that patterns of genetic structuring for this species are probably linked to Pleistocene glacial and interglacial periods. Results obtained supported previous studies in evidencing the occurrence of a species complex for *Salariopsis fluviatilis* with at least three taxonomic units: *Salariopsis fluviatilis sensu stricto*, *Salariopsis* cf. *fluviatilis* diffused in the Middle East and a further taxonomic entity from the Iberian Guadiana River basin whose tributaries flow in the Atlantic Ocean. For what concerns *Salariopsis fluviatilis sensu stricto*, two divergent groups of populations were reported in the present study, being the first diffused in western Mediterranean areas and the second in western Adriatic and eastern Mediterranean areas.

Abstract: The genus *Salariopsis* (Blenniidae) comprises freshwater blenny fish that inhabits Mediterranean Sea, Black Sea, and north-east Atlantic areas. Three species were formally described to date: *Salariopsis fluviatilis*. *S. economidisi*, and *S. atlantica*. In this study, 103 individuals were collected from different Italian regions (Sardinia, Liguria, Piedmont, Lombardy) and analyzed using the mtDNA Control Region and the ribosomal 16s gene. We aimed (i) to depict the phylogeographic patterns of *S. fluviatilis* in northern Italy and Sardinia and (ii) to compare the genetic structure of Italian samples with those from other Mediterranean regions. Results obtained showed the presence of a well-supported genetic structuring among Italian *S. fluviatilis* populations, shedding new light on the phylogeographic patterns of northern Italian populations of *S. fluviatilis sensu stricto* across the Ligurian Alpine ridge and the Sardinia Island-mainland dispersal patterns. Furthermore, our species delimitation analysis was consistent in supporting results of previous research about the presence of genetic differentiation among *S. fluviatilis*, evidencing: (i) a large group of *S. fluviatilis sensu stricto* that includes two sub-groups (Occidental and Oriental), (ii) one group comprising populations from the Middle East of a taxonomic entity corresponding to *Salariopsis* cf. *fluviatilis*, and (iii) one group of Iberian individuals from the Guadiana River.

Keywords: freshwater blenny; mtDNA; control region; phylogeny; *Salariopsis fluviatilis*

Citation: Azzena, I.; Scarpa, F.; Locci, C.; Cossu, P.; Niffoi, A.; Orrù, F.; Bovero, S.; Sotgiu, G.; Sanna, D.; Casu, M. Mitochondrial DNA of Sardinian and North-West Italian Populations Revealed a New Piece in the Mosaic of Phylogeography and Phylogeny of *Salariopsis fluviatilis* (Blenniidae). *Animals* **2022**, *12*, 3403. https://doi.org/10.3390/ani12233403

Academic Editor: Francesca Maradonna

Received: 20 October 2022
Accepted: 30 November 2022
Published: 2 December 2022

1. Introduction

The freshwater blenny, *Salariopsis fluviatilis* (Asso, 1801) (Actinopterygii: Blenniiformes: Blenniidae), is a small benthic fish that inhabits river basins and lakes around the Mediterranean Sea. It belongs to the recently described genus *Salariopsis* Vecchioni, Ching, Marrone, Arculeo, Hundt & Simons, 2022, proposed by Vecchioni and co-authors [1] in order to differentiate freshwater species previously referred to the genus *Salaria* Forsskål, 1775. Nowadays, the genus *Salariopsis* includes freshwater species whose distribution is across the eastern Atlantic coasts (Morocco), the Mediterranean Sea, and the Black Sea [2,3].

Worth noting, in the same year Duquenne-Delobel and co-authors [4] revalidated the genus *Ichthyocoris* Bonaparte, 1840 for the freshwater blennies. Although a debate is still ongoing [5], at the current state of the art we chose to use the genus *Salariopsis* as senior synomin, based on both the prior date of publication and the validation of the genus given by Vecchioni and co-authors [5].

Salariopsis fluviatilis—whose origin has been enrolled to a marine ancestor [6–10]—can show both morphological and behavioral differences in its populations, along with diversity in the adults/juveniles ratio, in response to the extended range of distribution and the different environmental factors, such as levels of predation, intraspecific competition, water flow, and lacustrine vs. river habitats [6,7,11–14]. For these reasons, different studies (see e.g., Belaiba et al. [15], Wagner et al. [16] and references therein) give rise to the suspicion that *S. fluviatilis* could be composed of an articulated species complex.

In this context, in addition to *S. fluviatilis*, two other freshwater blennies species have been formally described to date, characterized by a very limited distribution: *Salariopsis economidisi* (Kottelat, 2004), endemic to Lake Trichonis in Greece [2], and *Salariopsis atlantica* Doadrio, Perea & Yahyaoui, 2011, endemic to the Seboui river basin in Morocco [17].

Although *S. fluviatilis* shows, within the genus, the widest range of distribution, its populations are generally small and highly localized [18–23], suggesting that this species might have undergone high fragmentation phenomena that may have also produced relevant genetic divergences, at least in the most isolated populations, as likely happened to the other two freshwater *Salariopsis* species, which show confined (i.e., *S. atlantica*) or even point-like (i.e., *S. economidisi*) distributions.

Similar to other freshwater fish living in fragmented habitats, *S. fluviatilis* is highly susceptible to several anthropomorphic disturbances and to global environmental change effects. Nevertheless, even though it has been considered vulnerable or endangered in several countries and reported in Appendix III of the Bern Convention, *S. fluviatilis* has been evaluated by the International Union for Conservation of Nature (IUCN) as data deficient (IUCN 2013), due to a lack of data on its abundance and population trend.

Despite the potential interest of this species from phylogeographic, taxonomic and conservation perspectives, not many studies have been performed to depict the genetic structure and phylogeography of *S. fluviatilis* [7–9,15,16,20,24,25]. Most of these studies focused on the mitochondrial phylogenetic and phylogeographic patterns of the whole genus *Salariopsis* in the Mediterranean area, with inferences reported also for *S. fluviatilis*. Wagner et al. [16], in particular, supported the presence of at least six divergent genetic lineages within the *S. fluviatilis* populations: (1) Middle East (Israel, Syria and southern Turkey); (2) Guadiana (Guadiana basin with the Zùjar and Esteras rivers that flow on the European Atlantic coast); (3) Algeria-Verde (in southern Spain and Algeria with the Verde River and Oued Boughzazene); (4) Occidental basin (Algeria, Spain, mainland France, Corsica and Switzerland), (5) North Oriental basin (Italy, Croatia, Albania, Greece and western Turkey), and (6) the Island of Crete (Greece). The results of Wagner et al. [16] also support the results of other studies (see e.g., Belaiba et al. [15]), suggesting the presence of more than one taxonomic unit under the binomen *Salariopsis fluviatilis*. Remarkably, based on the results of two methods of species delimitation, Belaiba et al. [15] suggest that the populations of *S. fluviatilis* from the Middle East Mediterranean region should be regarded as a differentiated species, i.e., *Salaria* cf. *fluviatilis*. However, in light of the description of the new genus *Salariopsis* [1,5], we refer to it as *Salariopsis* cf. *fluviatilis* hereafter.

Furthermore, two main studies focusing specifically on the nuclear genetic variation of *S. fluviatilis* in the western Mediterranean [20,25] suggest that Corsica Island may have served as a glacial refuge during the last glacial maximum [25], as well as the presence of specific genetic lineages in the Iberian Peninsula [20]. In particular, Laporte et al. [25] argued that the genetic structuring retrieved between western and eastern populations of *S. fluviatilis* in the Mediterranean would be the product of the recent spreading of individuals from the two glacial refuges of the Iberian Peninsula and Corsica Island. However, in this study, the lack of samples from Liguria (i.e., north-west Italy) probably affected the possibility to fully reconstruct the dynamics of colonization which involved Corsica Island. Indeed, Wagner et al. [16], analyzing *S. fluviatilis* from Corsica, highlighted the importance of focusing studies on island populations due to the peculiarity and fragility of the island's fish fauna, to infer the biogeographical dynamics of the populations that live there. These same authors also expressed hope for extending the geographical sampling to provide a more comprehensive scenario of *S. fluviatilis* in the Mediterranean area.

Focusing on Italy, *S. fluviatilis* is distributed in a discontinuous way, starting from the northern regions and continuing through the Tyrrhenian side, down to Campania, Sardinia, and the Sicily islands. Isolated populations have also been found in Calabria and on the Adriatic coasts of the peninsula [26,27]. Although the Italian peninsula hosts different populations of *S. fluviatilis* and the country is characterized by different (hydro)geographic boundaries that strongly shaped its ichthyofauna with the presence of several endemic species—64 taxa are considered native to inland waters of Italy [28] and among these at least 15 species are endemic [29]—no wide-range molecular studies have been carried out on this species in Italy to date. Populations of *S. fluviatilis* living on Sardinia Island deserve particular attention, as this species is one of the few putative native freshwater fish known for the island [30]. The following two problems are particularly serious for freshwater Sardinian fauna: the introduction of a high number of alien species, most of them perpetrated in the past three decades [30–34] and the freshwater habitat alteration, consisting mainly of the construction of numerous dams starting from the beginning of the last century. In such a context, the present study was aimed at filling a gap in genetic information on populations of *S. fluviatilis* from Italian areas that are not yet under investigation. Indeed, knowledge of the genetic structure of populations from north-west Italy and the Sardinian Island could help to fill in some gaps in the mosaic of *S. fluviatilis* genetic distribution in the central-western Mediterranean.

As a results, the goals of our paper were: (1) to depict the phylogeographic patterns of *S. fluviatilis* populations, sampled in Sardinia and north-west Italy (Liguria, Piedmont, and Lombardy) in order to shed light on the phylogeography of the populations now living in the main central-western Mediterranean islands and the north-west Italian mainland; (2) to compare the genetic structure of our samples with that of other Mediterranean regions to put them in a wider geographic framework; and (3) to make an attempt to further clarify the taxonomy of the species across its whole range of distribution by using different methods of species delimitation to evidence, if any, distinct evolutionary lineages within *S. fluviatilis* that deserve to be deeply investigated in the future.

Two portions of the mitochondrial Control Region and of the 16s ribosomal gene were used as markers. The use of the mtDNA, which is maternally inherited, is considered an appropriate choice to solve intraspecific taxonomic and phylogenetic issues and to depict evolutionary lineages [35–37]. Furthermore, because of its high mutational rate, mtDNA represents a practical tool to obtain preliminary inferences about the level of genetic variability among populations [38]. The high number of mitochondrial Control Region and 16s sequences that are present in the GenBank database for *S. fluviatilis* allowed us to compare the sequences obtained in the present study with many others from different geographic areas. In previous studies, the combined use of Control Region and 16s sequences has been demonstrated to be helpful in providing insights into the genetic structuring and phylogenetic history of *S. fluviatilis* populations [1,8,15,17].

2. Materials and Methods

2.1. Sample Collection

A total of 103 individuals of *Salariopsis fluviatilis* were caught using an electric stunner from the freshwaters of 4 different Italian regions (Sardinia, Liguria, Piedmont and Lombardy) between February 2019 and March 2022 (Table S1 and Figure 1).

The protocol of sampling and analysis of the fish fauna of wadable lotic systems, provided by the Italian Higher Institute for Environmental Protection and Research (ISPRA) [39], was followed for the sampling collection in the present study. In accordance with this document (whose guidelines are compulsory in Italy), which requires that all electrically stunned fish must be collected, recorded and returned to the water, the individuals of *S. fluviatilis* analyzed in this study were caught using an electric stunner from freshwaters, subjected to a non-lethal sampling method by means of small tissue portion removal (fin-clips) and immediately transferred to a recovery tank before being released. Tissues were preserved in absolute ethanol and used to perform DNA extraction. The above reported sampling method was approved by the ethics committee of the University of Sassari (Prot n. 122 770 of 7 November 2022), and its researchers led the sample collection activities during the present study.

Sardinian samples were collected in three tributaries of the Flumendosa river (Accu terrale creek, Riu Pale creek, Sicaderba creek), all confluent in the Alto Flumendosa man-made reservoir (originated in the middle of the last century by a dam) (central-east Sardinia), the Rio Mannu di Scano Montiferro river (central-west Sardinia), and the Rio Mannu di Posada river (north-east Sardinia); Ligurian samples were collected in the Entella river (east Liguria) and in the Roja river (west Liguria); samples from Piedmont were collected in the San Giovanni river, in the San Bernardino river, and in the Strona di Omegna river; the sample from Lombardy was collected in the Tartaro Fuga creek, tributary of the Oglio river.

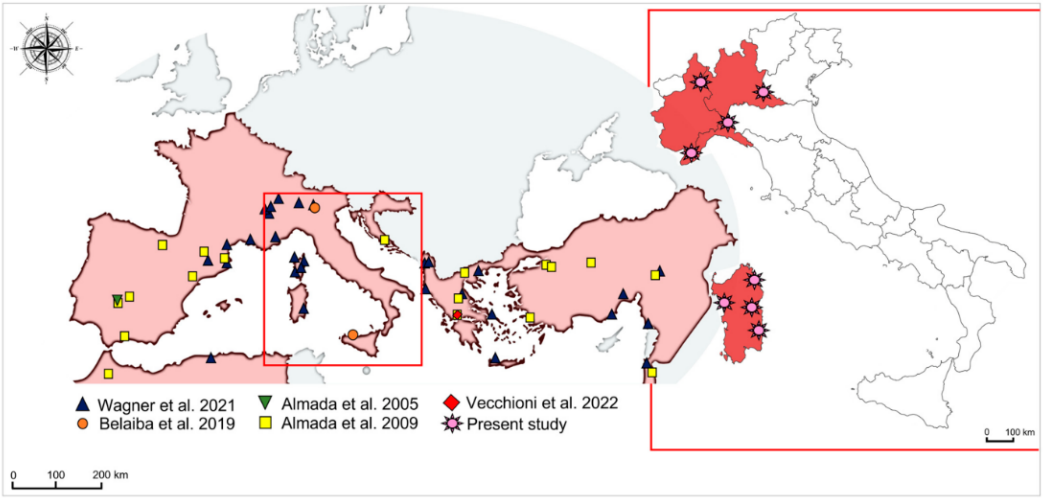


Figure 1. Map of samples' collection sites. The map shows the geographical origin of the sequences isolated in the present study along with those from previous research [1,8,15,16,24].

2.2. Implemented Molecular Analysis

Total genomic DNA was isolated from a portion of fin tissue using the Macherey-Nagel Nucleo Spin Tissue Kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany), following the supplier's instructions. DNA solutions were quantified using the Nanodrop™ Lite Spectrophotometer (by Thermo Scientific; Waltham, MA, USA), which showed an average yield of 29 ng/µL. A portion of the mitochondrial Control Region and rRNA gene 16s were amplified by standard

PCR using the following primers: CR-F (forward) (5′-CCACTAGCTCCCAAAGCTA-3′) and CR-R (reverse) (5′-CAGGACCAAGCTTTTGTGC-3′) [40]; 16s For (5′-CGCCTGTTTATCAAAAACAT-3′) and 16s Rev (5′-CCGGTCTGAACTCAGATCACGT-3′) [41]. Reactions were carried out in a total volume of 25 µL. On average, 10 ng of total genomic DNA were combined with 0.6 µM of each primer and one pellet of PuReTaq Ready-To-Go PCR beads (GE Healthcare, Wauwatosa, WI, USA) containing stabilizers, 4 ng of bovine serum albumin (BSA), deoxynucleotide triphosphates, 2.5 units of PuReTaq DNA polymerase, and reaction buffer. When a bead was reconstituted to a 25 µL final volume, the concentration of each dNTP and $MgCl_2$ was set at 200 µM and 1.5 mM, respectively. PCRs were performed in a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems, Waltham, MA, USA), programmed as follows: 1 cycle of 4 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at 51 °C (for both Control Region and 16s gene), and 30 s at 72 °C. At the end, a post-treatment of 10 min at 72 °C and a final cooling at 4 °C were carried out. Both positive (high-quality DNA samples from the species *Micropterus salmoides*) and negative controls were used to test the effectiveness of the PCR protocols and the absence of possible contaminations. Electrophoresis was carried out on 2% agarose gels, prepared using 1× TAE buffer (Tris-Acetate-EDTA, pH 8.3) and stained with Gel Red Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). PCR products were purified by ExoSAP-IT (USB Corporation, Cleveland, OH, USA) and sequenced for forward and reverse strands (by means of the same primers used for PCR), using an external sequencing core service (Macrogen, Europe, Amsterdam, The Netherlands).

2.3. *Phylogenetic and Phylogeographic Analyses*

A total of 103 newly generated sequences was aligned using the package Clustal Omega [42] (available at https://www.ebi.ac.uk/Tools/msa/clustalo/ (accessed on 16 September 2022)) and deposited in GenBank (OP675744-OP675846 for the Control Region and OP653611-OP653713 for 16s).

The Control Region and 16s *S. fluviatilis* sequences obtained in the present study were concatenated to obtain a dataset as polymorphic as possible to infer on the phylogeographic relationships among populations from Sardinia and mainland northern Italy. Furthermore, with the aim of performing phylogenetic and molecular taxonomy analyses on the species on a wide Mediterranean context, the Control Region and the 16s sequences that were collected in the present study were also merged and aligned with those from other localities (Italy, Switzerland, Croatia, Albania, Greece, Israel, Morocco, Portugal, Spain, Turkey, Syria and Algeria) available on GenBank to date (last update 30 September 2022) (see Tables S2 and S3).

In the Control Region dataset, all the sequences available for the Mediterranean freshwater blennies were also included, i.e., *S. economidisi* (FJ465540, FJ465541, MZ026042, MZ026043, MZ026044) and *S. atlantica* (FJ465527, FJ465526). Furthermore, the marine species *Salaria basilisca* (MW555061) and *Salaria pavo* (MW555062) were used as outgroups. As well, in the 16s dataset, the sequences available for the Mediterranean freshwater blenny species were also included, i.e., *S. economidisi* (FJ465733, FJ465735) and *S. atlantica* (FJ465736, FJ465737). The marine species—*S. basilisca* (MH724820) and *S. pavo* (FJ465707)—were used as outgroups (see Table S3).

The genetic variation within the datasets was assessed estimating the number of polymorphic sites (S), number of haplotypes (H), haplotype diversity (h), and nucleotide diversity (π) using the software package DnaSP 6.12.03 (Barcelona, Spain) [43].

To test the reliability of the three datasets (i.e., Control Region, 16s and concatenated dataset) for taxonomic and phylogenetic analyses, the phylogenetic signal was checked by the software TREEPUZZLE (Wien, Austria) [44]: the likelihood-mapping analysis of 10,000 random quartets, performed for both genes was used and the general time reversible (GTR) was selected as the model of substitution (following Scarpa et al. [45,46]).

The best probabilistic model of sequence evolution was achieved using jModelTest 2.1.3 (Ballwin, MO, USA) [47], with a maximum likelihood optimized search by the Akaike (AIC) and Bayesian Information Criterion (BIC). The TPM3uf+I+G was suggested by the

AIC and the HKY+G by the BIC as the best-fit model for the Control Region dataset, while the TPM2uf+G by AIC and the TPM2+G by BIC were selected for the 16s dataset. The pairwise genetic distances were estimated between populations using the software Mega 6.06 [48] with 1000 bootstrap replications. The correction according to the Kimura two-parameter model (K2P) [49] was applied (see Scarpa et al. [50] for methods).

To infer the genetic relationships among haplotypes and to detect the possible occurrence of discrete genetic clusters, a median-joining network [51] was constructed by means of the software Network 10.2.0.0 (www.fluxus-engineering.com) (accessed on 12 October 2022) (Colchester, UK). Transitions and transversions were equally weighted. Due to the absence of information about the possible appearance of retro-mutation events, the same weight (10) was assigned to each observed polymorphism.

Phylogenetic relationships were investigated on two datasets including all the sequences of Mediterranean blennies belonging to the genus *Salaria* from GenBank: they grouped 280 sequences for the Control Region (see Table S2) and 136 sequences for the 16s (see Table S3). Analyses were based on Bayesian Inference (BI) and performed by means of the software MrBayes 3.2.7 [52]. The BI was implemented specifying a partitioned model and setting as model parameters: NST = 6, rates = invgamma, ngammacat = 4. Two independent runs each consisting of four Metropolis-coupled Markov chain Monte Carlo (MCMC) chains (one cold and three heated chains) were run simultaneously for 5,000,000 generations, sampling trees every 1000 generations. The first 25% of the 10,000 sampled trees was then discarded as burn-in (see Scarpa et al. [50]). In order to assess the convergence of chains, it was checked that the Average Standard Deviation of Split Frequencies (ASDSF) approached 0 [52] and the Potential Scale Reduction Factor (PSRF) was approximately 1 [53]. Nodes with a percentage of posterior probability lower than 95% were considered not highly supported. The phylogenetic tree was visualized and edited using FigTree 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/) (accessed on 18 September 2022) (see Scarpa et al. [50]).

The taxonomic identity of each *S. fluviatilis* sequence was checked using five different methods of species delimitation. The GMYC (Generalized Mixed Yule Coalescent) model tests for a significant shift in the branching rate along an ultrametric species tree. The analysis was performed on the ultrametric species tree obtained from the Bayesian dating analyses by means of the SPLITS (SPecies LImits by Threshold Statistics) package [54] implemented in the R statistical environment (available at http://r-forge.rproject.org/projects/splits/)(accessed on 18 September 2022). Species entities were identified by means of the single threshold option, which uses a single threshold in order to specify the transition from between- to within-species branching (see Scarpa et al. [55]).

The PTP (Poisson Tree Processes) [56]) uses the number of substitutions to assess the speciation rate. Species delimitation was achieved by means of the PTP web server (available at http://species.h-its.org/ptp/) (accessed on 18 September 2022) on the phylogenetic species trees using default options and 500,000 MCMC generations. To test the reliability of results, each run was examined for convergence by visualizing the likelihood plot: if convergence occurred, the chain should stay at high likelihood locations during the run (see Scarpa et al. [55]).

The ABGD (Automatic Barcode Gap Discovery) [57] method is based on the K2P genetic distances [49]. The ABGD does not consider phylogenetic relationships within the dataset and works on sequences, identifying the barcode gap as the first significant gap beyond this limit and using it to partition the data [57]. Species were assessed by means of the ABGD online tool (available at http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) (accessed on 18 September 2022) using the default settings. The correct species estimate was selected, according to Puillandre et al. [57], using the gene specific priors for maximum divergence of intraspecific diversity, corresponding to $p = 0.001$ (see Scarpa et al. [55]).

The NDT (Nucleotide Divergence Threshold) method was used on each gene separately, by means of a script developed by Scarpa et al. [58], written in the R statistical

environment (available at https://cran.rproject.org/) (accessed on 18 September 2022). The script partitions taxa into entities applying the fixed threshold of 2% obtained by Hebert et al. [59], using a pairwise Kimura two-parameter model (K2P) [49] genetic distances matrix (see Scarpa et al. [45]).

The principal coordinates analysis (PCoA) was performed using GenAlEX 6.5 [60] on a matrix of pairwise genetic distances corrected according to the Kimura two-parameter model (K2P), with the aim of identifying potential subgroups within the genetic clusters and to determine the dissimilarity represented by the genetic variation among sequences (see Tran Thi et al. [61]).

2.4. Estimation of Divergence Time Analyses

The software package Beast 1.10.4 [62] was used to estimate the divergence time for the clades evidenced by the Control Region phylogenetic tree, applying the substitution rate estimated on another blennioid species, *Tripterygium delaisi*, in a previous study by Koblmüller et al. [63]. Specifically, a 95% HPD (high density probability) interval of 1.1–6.67% per million years was fixed [16]. Site parameters Substitution Model = GTR; Bases Frequencies = Estimated; Site Heterogeneity Model = Gamma + Invariant Sites; Number of Gamma Categories = 4 have been set accordingly to the evolutionary models selected by jModeltest with the GTR model. For the molecular clock rate variation model, the lognormal uncorrelated relaxed clock was selected as it assumes independent rates on different branches. For the tree prior, the Yule prior process was applied to the speciation model. The priors for model parameters and statistics have been determined for calibrating the time-tree assuming the mutation rates per million years. Divergence times were estimated using a normal distribution with lower, central and upper values set according to the mutation rate per million years. Operator parameters have been fixed following the instructions of the user manual. Additionally, the application of the lognormal uncorrelated relaxed clock model gives an indication of the state of the clock-like data (measured by the ucld.stdev parameter). If the ucld.stdev parameter-estimate is close to 0, then the data is quite clock-like, and if it has an estimated value much greater than 1, then data exhibits very substantial rate heterogeneity among lineages. To obtain an effective sample size (ESS) greater than 200 for all the statistic parameters, a run of 200,000,000 generations was performed, sampling a tree every 20,000 generations following Scarpa et al. [50]. The software Tracer 1.6 (©2022 BEAST Developers. All rights reserved.) was used to view the resulting log file, with the aim of ensuring convergence of parameter values, to verify whether ESS values exceeded 200 and to estimate node ages [64]. Tree Annotator and FigTree were used for drawing and visualizing the time calibrated tree, following Scarpa et al. [50].

3. Results

A total of 103 sequences for both Control Region (302 bp) and 16s gene (547 bp) were obtained in the present study. Both datasets were merged to Genbank sequences (280 for the Control Region and 136 the 16s gene) and used to implement phylogeographic analyses.

All the newly generated sequences were identified as belonging to the species *Salariopsis fluviatilis* (sub *Salaria fluviatilis*) through Basic Local Alignment Search Tool (BLAST) analysis implemented in the GenBank nucleotide database (www.ncbi.nlm.nih.gov) (accessed on 24 August 2022) that showed a percentage of identity ranging from 97% to 100% for the Control Region and from 90% to 100% for the 16s gene.

In accordance, all methods of species delimitation suggest that the sequences isolated from Italian samples in the present study belong to the taxonomic entity of *S. fluviatilis* (see Table S4 for details). After reaching the correct taxonomic identification, further analyses were performed to infer the levels of genetic variation and genetic structuring among populations in Italy and other geographic areas.

3.1. Phylogeographic Relationships among New Samples from North-West Italy and Sardinia Island

Among the 103 sequences obtained in the present study for the Control Region in samples from Liguria, Piedmont, Lombardy and Sardinia, 16 polymorphic sites resulting in 14 haplotypes were found (see Table 1).

Table 1. Indices of genetic variation. The table reports the estimates of genetic variation for the mitochondrial datasets analyzed in the present study. N: sample sizes; bp: fragment size; S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; π: nucleotide diversity.

Control Region Dataset						
	N	bp	H	S	h	π
Present study	103	304	14	16	0.554	0.00461
Sardinia	47	304	4	4	0.272	0.00046
Liguria	44	304	8	7	0.581	0.00298
Piedmont	10	304	4	4	0.644	0.00317
Lombardy	2	304	1	0	0.000	0.00000
16s Dataset						
	N	bp	H	S	h	π
Present study	103	599	8	7	0.620	0.00073
Sardinia	47	599	3	2	0.581	0.00029
Liguria	44	599	5	4	0.216	0.00041
Piedmont	10	599	2	1	0.200	0.00037
Lombardy	2	599	1	0	0.000	0.00000
Concatenated Dataset						
	N	bp	H	S	h	π
Present study	103	849	20	23	0.797	0.00210
Sardinia	47	849	7	6	0.692	0.00038
Liguria	44	849	12	11	0.644	0.00132
Piedmont	10	849	4	5	0.644	0.00136
Lombardy	2	849	1	0	0.000	0.00000

The overall level of haplotype diversity found for samples from Sardinia was very low with only 4 similar haplotypes found out of 47 sequences and a reduced nucleotide diversity (π = 0.00046). For the two samples from Lombardy only one haplotype was retrieved, while higher levels of divergence were found for samples from Liguria and Piedmont with 8 and 4 haplotypes, respectively.

Regarding the 103 sequences obtained in the present study for the 16s gene, low levels of divergence were found for each region with a total of 7 polymorphic sites resulting in 8 haplotypes (see Table 1). Samples from Liguria, Piedmont and Lombardy showed the lowest levels of diversity.

With the aim to obtain a dataset including the level of polymorphism and informativeness as high as possible, these two datasets of sequences (Control Region and 16s) were concatenated to obtain a longer mitochondrial fragment, which was 849 bp long, to perform phylogeographic analyses. In this concatenated dataset, 23 polymorphic sites were found, resulting in 20 haplotypes and moderately low levels of haplotype and nucleotide diversity (see Table 1). A similar level of diversity was found for all the regions analyzed

(with the exception of Lombardy). In accordance with the trend previously reported for the 16s datasets, the highest level of divergence was found in Sardinia; however, in this island 6 polymorphic sites were found, resulting in 7 very similar haplotypes, as suggested by the lowest (with the exception of the two samples from Lombardy) rate of nucleotide diversity (π = 0.00038) (see Table 1).

The network analysis was performed on a concatenated (Control Region and 16s) dataset which included samples from northern Italy and Sardinia (see Figure 2) from the present study. In order to obtain a scenario as complete as possible of the phylogeography of *S. fluviatilis* in Italy, the two Italian sequences (from the Lombard side of Lake Garda and Sicily) which had been previous published [15] were included in the dataset. The analysis took into account gaps and redundant polymorphic bases and evidenced the occurrence of a genetic structuring between two main groups of sequences (Groups A and B) which diverged for 6 point mutations among each other.

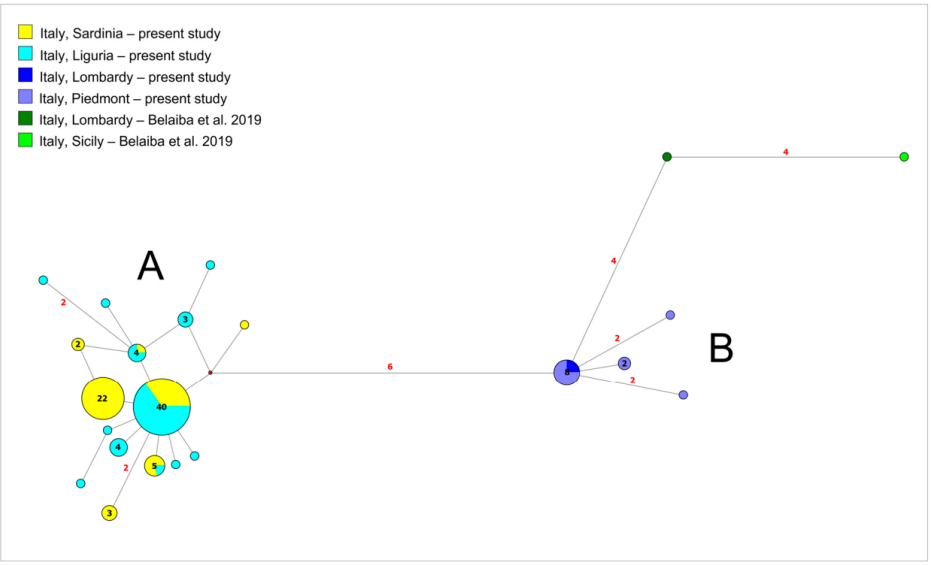


Figure 2. Median-joining network analysis. The network includes all the Italian sequences (from the present study and also from a previous study [15]) for the Control Region and 16s (merged in a concatenated dataset). Clusters A and B are described in the text. The number of mutations between sequences that are greater than $n = 1$ are reported on network branches. The number of individuals showing the same haplotype that is greater than $n = 1$ is reported inside the spots. Belaiba et al., 2019 in the legend corresponds to the reference [15].

Group A included individuals from Sardinia and Liguria with the occurrence of two main star-like shapes involving the two areas. In particular, among samples from Liguria, a high level of variability was witnessed by the number of different haplotypes that are present in this area with a strong founder effect, as suggested by the occurrence of a very common central haplotype (that is also shared by 29.8% of Sardinian individuals), which is surrounded by several derived haplotypes diverging for a few mutations. The haplotypes that likely originated from the lineages most common in Liguria, are private to single Ligurian individuals (with one exception for one haplotype shared by three individuals from Sardinia) or shared among Sardinian and Ligurian individuals. In only two cases, haplotypes likely derived from the Ligurian ancestor lineage are exclusive to Sardinian individuals. In particular, one of these two latter cases corresponds to the most common haplotype among Sardinian individuals, which was found in 46.8% of specimens collected

and is private to Sardinia. In particular, only one haplotype isolated in the north-east of the island (Rio Mannu di Lodè river), which is exclusive to a single individual, could have not necessarily originated from the most common Ligurian and Sardinian founder lineages.

Group B included samples of Piedmont and Lombardy from the present study with a marked founder effect suggested by the occurrence of a star-like shape. The two identical sequences found in Lombardy correspond to the most common haplotype occurring in Piedmont. Two out of three derived haplotypes were private of single Piedmontese individuals. The sequence from Lake Garda (Lombardy) from a previous study [15] was included within this group diverging for four points mutations from the most common haplotype. Additionally, the sequences from Sicily from a previous study [15] were included within this group, diverging for four points mutations from the sequence from Lake Garda.

In the PCoA graph (see Figure 3 and Table S5) the 61.38% of variability was explained by *x*-axis and *y*-axis and accounted for only 13.49%.

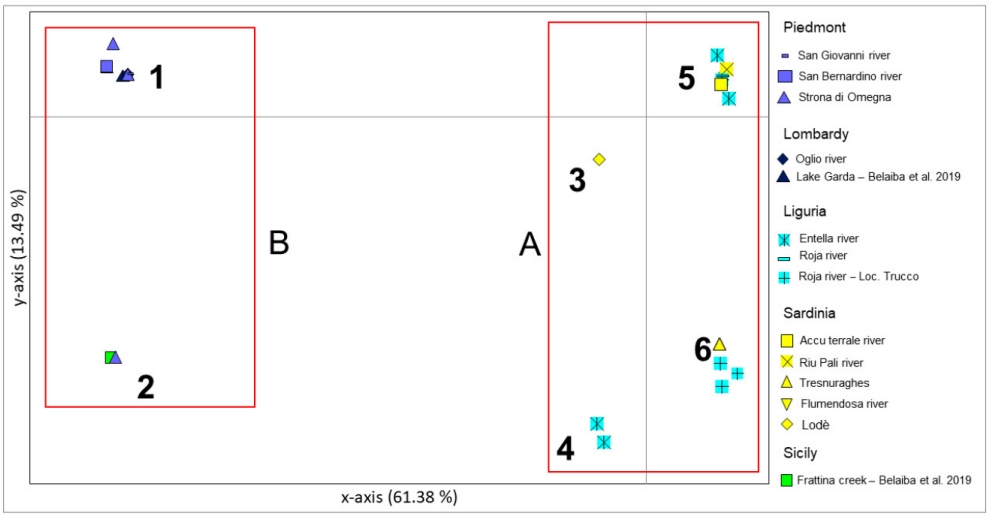


Figure 3. Principal coordinates analysis performed on Italian sequences from the present study and also from a previous study [15]. Bidimensional plot shows the genetic differentiation among specimens due to the nucleotide substitutions per site found in the dataset. Belaiba et al., 2019 in the legend corresponds to the reference [15].

Results were consistent with the genetic structuring evidenced by network analysis, suggesting a divergence along the *x*-axis between two main groups (Groups A and B). Group A included individuals from Liguria and Sardinia with an internal, although poorly supported, sub-structuring. Indeed, an eastern Sardinian sample from the Rio Mannu di Lodè river (subgroup 3 in Figure 3) and a group of four samples from east Liguria (subgroup 4 in Figure 3) collected in the Entella river, slightly diverged from the remaining samples represented by two different clusters of sequences. The first of these two clusters (subgroup 6 in Figure 3) included only few samples from west Liguria (five individuals from Roja river), three individuals from central-west Sardinia: two from Rio Mannu di Scano Montiferrro river and a further sample from the Rio Mannu di Lodè river. The second cluster (subgroup 5 in Figure 3) of sequences was the larger in the graph and included individuals from all the Sardinian and Ligurian sampling sites.

Group B of PCoA included all samples from Piedmont and Lombardy in a unique cluster with the only exception of one Piedmontese sample (subgroup 2 in Figure 3) from

the Strona di Omegna River in the north-east of Piedmont, which diverged from the main group (subgroup 1 in Figure 3) on the *y*-axis.

3.2. Phylogenetic Reconstruction and Species Delimitation Based on Control Region

Both phylogenetic and species delimitation analyses were based on the Control Region dataset that showed a strong phylogenetic signal (see Figure S1a). In particular, either the reduced number of specimens available on GenBank for which sequences of Control Region and 16s were present, and the low phylogenetic signal obtained for 16s (see Figure S1b) prevented us from applying those analyses to a reliable concatenated dataset in terms of a phylogenetic signal (see Figure S1c). Indeed, the test of the Likelihood Map disassembled the dataset in quartets, that represent the smallest set of taxa for which more than one unrooted tree topology exists [65]. The quartet puzzling works on groups of four sequences, in order to obtain a map that allows for understanding whether data are reliable for phylogenetic and taxonomic inferences. The most important information in the map is given by the percentage of star-like trees (which represent the area of the unsolved trees), which, when higher than 30%, suggests that the dataset is not reliable for analyses, due to noisy data; alignment errors; recombination events; not enough informative sites [66]; or inadequate taxonomic coverage (see Scarpa and co-authors [67]). The Control Region dataset showed a percentage of points in the network-like areas of 23.5% (see Figure S1a) and the 16s dataset 92.2% (see Figure S1b). As expected, the high level of noisy data and the low level of information sites in the 16s dataset also affected the concatenated dataset, which showed in the network-like areas a percentage of 79.7% (see Figure S1c). For these reasons, both 16s and concatenated datasets were considered as not reliable for phylogenetic and taxonomic purposes and were not used for phylogenetic analyses in the present study.

The network analysis performed on the Control Region dataset, including all the sequences available for the Mediterranean freshwater blennies (see Figure 4 and Table S2), showed the presence of three main clusters: cluster A, cluster B, and cluster C.

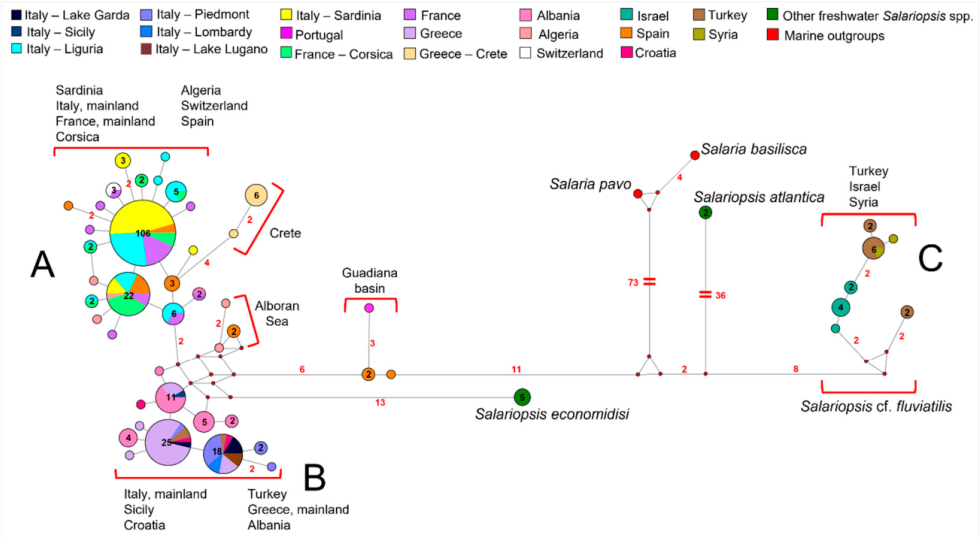


Figure 4. Median-joining network analysis performed on the whole Control Region dataset. Clusters A, B and C are described in the text. The number of mutations between sequences that are greater than $n = 1$ are reported on network branches. Additionally, the number of individuals showing the same haplotype that is greater than $n = 1$ is reported inside the spot.

Cluster A included all the Sardinian and Ligurian samples, together with all the sequences from mainland France, Corsica, Algeria, Switzerland, and almost all from Spain. Interestingly, sequences from Crete showed only two haplotypes which resulted endemic of the island and quite divergent from the others (4 to 6 mutational steps apart from cluster A).

Cluster B was representative of northern Italy (Lake Garda, Piedmont, and Lombardy, excluding Liguria) and Sicily, along with the eastern part of continental Europe (Croatia, Albania, Greece mainland, and western Turkey coasts on the Aegean and the Marmara seas). Furthermore, three sequences from Spain and one from Portugal from the basin of the Guadiana River, which flows into Atlantic Ocean, set 6 to 9 point mutations apart from clusters A and B. Interestingly, a group of two Spanish and two Algerian sequences, all collected throughout a geographic area facing the Alboran Sea, sets in the network on a separate position between clusters A and B diverging for 4 point-mutations from each of them.

Cluster C grouped the sequences from southern Turkey (but also one from Ilica River in the north) that Belaiba et al. [15] identified as *S.* cf. *fluviatilis*, along with sequences from Syria and Israel. Within this cluster, all the haplotypes found in Israel were exclusive to this area, while a common haplotype was shared between Turkey and Syria. Cluster C diverged from the previous reported Guadiana River group for 22 to 29 point mutations.

The sequences of *S. economidisi*, endemic of Lake Trichonis in Greece, sets close to cluster B, 13 point-mutations apart. The two sequences representing *S. atlantica*, which was described in Morocco, were found 44 mutational steps far from cluster C, 50 from the Guadiana River group, and approximately 60 from clusters A and B.

The phylogenetic tree (see Supplementary Figure S2) showed the presence of a large monophyletic clade, including all the Italian sequences belonging to *S. fluviatilis* together with the Mediterranean sequences taken from GenBank, and it split into a dichotomy characterized by two main clusters (clusters 1 and 2). The first cluster (cluster 1) showed a polytomic clade, including almost all *S. fluviatilis* sequences, and its sister clade representative of *S. economidisi* (represented by two sequences).

The large polytomy of *S. fluviatilis* grouped all the Italian sequences along with those from eastern Europe, mainland France, Corsica, Spain, Portugal, Switzerland, and Algeria. It includes some internal well-supported sub-clusters; among them, the three most relevant groups: (i) three Spanish and one Portuguese sequences corresponding to the Guadiana River group evidenced by network analysis, (ii) two Spanish and two Algerian sequences, and (iii) seven sequences from Crete. These three well-supported internal subgroups were also found by network analysis. The second cluster (cluster 2) of the tree included the Turkish, Syrian, and Israeli sequences, which grouped together in the cluster C of network analysis. Accordingly, with network analysis, an internal sub-structuring, consistent with the geographic distribution of samples, was also found in cluster 2 of the phylogenetic tree.

Species delimitation methods outcomes (see Table S4) showed some discrepancies among each other, consistent with the characteristics of methods (see the Section 2 for details). Focusing on *S. fluviatilis*/*S.* cf. *fluviatilis*, the NDT and the ABGD methods converged in recognizing the same three entities. The first entity comprises Italian, French, Swiss, Algerian, Balkan and almost all Spanish sequences. The second entity includes the Turkish, Syrian and Israeli sequences corresponding to the group (including *S.* cf. *fluviatilis*) already retrieved by previous analyses. The third includes the Portuguese and Spanish sequences from the Guadiana River basin (in accordance with results already obtained by Wagner et al. [16]).

On the other hand, the PTP method detects four entities, partially converging with the NDT and ABGD but grouping the Guadiana River entity within the "European" entity, and splitting the group of Turkish, Syrian, and Israeli sequences into three separate entities. The GMYC methods found a unique entity for all the sequences belonging to the genus *Salariopsis* included in the datasets.

In the PCoA analysis (Figure 5 and Table S6), the 41.43% of variability was explained by *x*-axis while *y*-axis accounted for 18.47%.

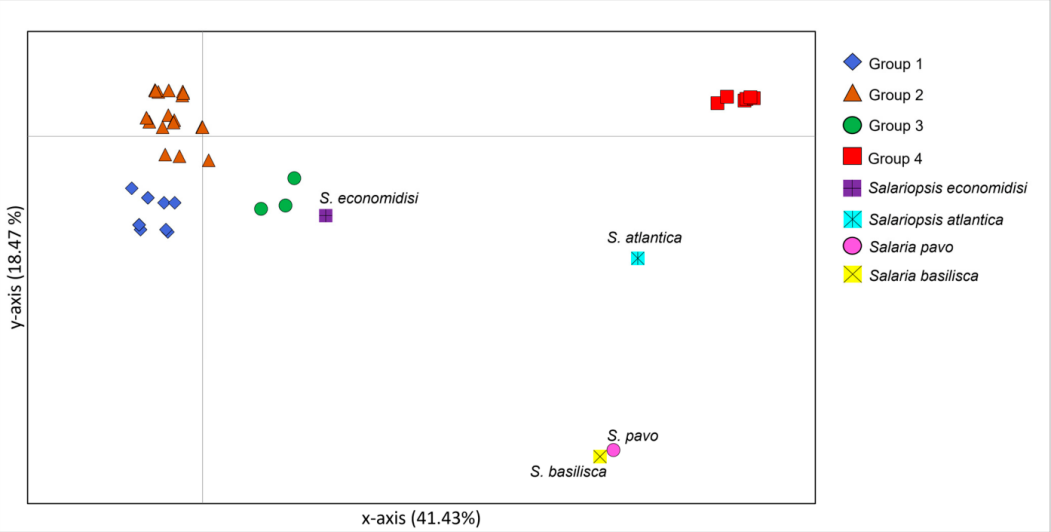


Figure 5. Principal coordinates analysis performed on the whole Control Region dataset. Bidimensional plot shows the genetic differentiation among specimens due to the nucleotide substitutions per site found in the dataset.

Results are consistent with the differences among *S. fluviatilis* Mediterranean populations evidenced by previous analyses, distinguishing four groups: (a) Group 1, which includes sequences from Greece, northern Italy (apart from Liguria), Sicily, Croatia, Albania, and some sequences from the western coasts of Turkey facing the Aegean and the Marmara seas; (b) Group 2, which includes sequences from Sardinia, Liguria, Spain, mainland France, Corsica, Switzerland, and Algeria; (c) Group 3, which includes the Guadiana River sequences from Spain and Portugal; and (d) Group 4, which includes sequences from Turkey that Belaiba et al. [15] identified as *S.* cf. *fluviatilis*, Israel and Syria. It is important to note that group 1 and group 2 diverged only on the *x*-axis which accounts for only 18.47% of variation suggesting that, according to Network analysis, these two groups are structured among each other but with a weak genetic divergence.

Salariopsis economidisi and *S. atlantica*, and the two marine outgroups, *S. pavo* and *S. basilisca*, set in the plot apart from the four *S. fluviatilis*/*S.* cf. *fluviatilis* groups.

The time calibrated tree (Figure S3), obtained by the software Beast, was consistent with the Bayesian phylogenetic tree (see Supplementary Figure S2). The common ancestor to all Mediterranean *Salariopsis* species dates back to 990 kya. The common ancestor of the group of sequences, including *S. fluviatilis*, *S.* cf. *fluviatilis* and *S. economidisi* (see SEGR2_FJ465541_1 and SEGR1_FJ465540_1 in the tree) dates back to approximately 800 kya. *Salariopsis atlantica* (see SAMA2_FJ465526 and SAMA1_FJ465527 in the tree), which originated in 390 kya, sets in the phylogenetic tree as an external clade to the cluster including *S. fluviatilis*, *S.* cf. *fluviatilis* and *S. economidisi*.

The large group of sequences of *S. fluviatilis*, which includes Italian, Eastern European, French, Corsican, Spanish, Portuguese, Swiss, and Algerian individuals, along with the sequences belonging *S. economidisi*, differentiated approximately 610 kya. Within this group, the common ancestor to *S. fluviatilis* diverged approximately 575 kya, while the common ancestor of *S. economidisi* dates back to approximately 130 kya. Within the large clade of *S. fluviatilis*, it was possible to obtain a molecular dating for the three well-supported groups of sequences reported in the previous analyses. One group corresponds to the sequences from the Guadiana River basin and dates back to approximately 180 kya (with two internal

subgroups that originated approximately 40 and 20 kya, respectively), another group of Spanish and Algerian sequences from the Alboran Sea area, dates back to approximately 150 kya (with two internal subgroups that originated approximately 120 and 30 kya, respectively), and the last group, including Cretan sequences, dates back to approximately 210 kya (with an internal subgroup that originated approximately 80 kya).

The group of sequences, including the Turkish, Syrian and Israeli *S.* cf. *fluviatilis* sequences, date back to approximately 630 kya.

3.3. Phylogenetic Reconstruction Based on 16s Gene

The low level of genetic variation found for the 16s dataset (see Table 1) may have affected the values obtained from the phylogenetic signal test that suggested a lower resolution for phylogenetic and taxonomic analyses if performed. The scarce informativeness of the 16s fragment in this study may be explained considering that it is a mitochondrial gene highly conserved, therefore it generally shows lower levels of genetic differentiation [67] and a slower evolutionary rate if compared to the Control Region.

However, even taking into account the low level of polymorphism and the reduced number of available sequences for the 16s dataset, the network analysis (Figure S4) was consistent and corroborated results obtained from the Control Region, with a few discrepancies. Indeed, results obtained evidenced the occurrence of two main clusters of sequences (A and B) for *S. fluviatilis* and two further small clusters of sequences for (1) the Guadiana River basin and (2) western Turkey (*S.* cf. *fluviatilis*) with Israel, respectively. Cluster A includes only sequences from the Italian regions of Piedmont, Lombardy, Liguria, and Sardinia, while cluster B includes sequences from Italy (only from Lake Garda in the north of the peninsula and Sicily Island), Croatia, Spain, Greece, and Turkey (in the West of the country). All the 16s sequences obtained in the present study were included within cluster A.

4. Discussion

The present study represents the first insight into the genetic variability of the Italian populations of *Salariopsis fluviatilis*: phylogeographic and phylogenetic analyses were performed including a high number of Italian specimens from the north-west regions of the peninsula (Liguria, Piedmont, and Lombardy) and the Mediterranean island of Sardinia, which represents an area so far neglected by molecular studies focused on this species. Moreover, molecular dating based on the present dataset, which covers most of the range of distribution of this species, gave us the possibility to put in an updated time range of the evolutionary patterns previously described for the new genus, *Salariopsis* [15,16]. Finally, the results obtained in the present study, which are fully supported by the application of five methods of species delimitation, corroborated previous hypotheses on the taxonomic status of some populations of Mediterranean freshwater blennies [15,16].

The mitochondrial DNA showed to be a suitable tool to illustrate phylogenetic and phylogeographic relationships among *S. fluviatilis* populations and to infer the taxonomic status of this species. In particular, the use of the mitochondrial Control Region has been effective in identifying genetically divergent groups within the species.

4.1. Genetic Structuring and Phylogeographic Patterns of Salariopsis fluviatilis in Italy

Our research evidenced a well-supported genetic structuring among Italian *S. fluviatilis* populations. Indeed, the group that includes Ligurian and Sardinian samples is strongly divergent from the group formed by samples from Piedmont and individuals from Lombardy and Sicily. The genetic structuring between *S. fluviatilis* from Liguria and Piedmont, which do not share common haplotypes among each other, may be ascribed to the geographic isolation between these flanking geographical areas. Indeed, the genetic pattern observed here can be explained considering that the presence in the Ligurian region of boundaries, represented by the Alps ridge in continuity with the Apennines, hindered the Ligurian individuals' colonization of the northernmost Italian regions. This scenario is known for other species of Italian freshwater fishes as well, for instance, within the genus

Telestes, Stefani et al. [68] reported the presence of two main mitochondrial clades within the Italian evolutionary lineage, which can be strongly traced geographically to the two biogeographical districts, Tuscan-Lazio, and Po Valley-Venetian, which are actually separated by the Ligurian Alps-Northern central Apennines system. As well, regarding the genus *Squalius*, the same barrier secretes *S. squalus* in the basins afferent to the Po River drainage network and *S. lucumonis* in the Tyrrhenian [69–72]. The pivotal role of the Appennines as biogeographic barrier for freshwater taxa has been evidenced in other vertebrates other than fish, such as in the European pond turtle (*Emys orbicularis*) (Testudines: Emydidae) [73] and also in invertebrate species, such as in the *Unio* spp. (Bivalvia: Unionidae) [74].

Worth noting, the sole Sicilian individual present in our dataset, from Belaiba et al. [15], grouped with the genetic cluster that also includes *S. fluviatilis* from Piedmont and Lombardy. This is an interesting finding, as it does not support the role ascribed to *S. fluviatilis* as a dispersal vector for the glochidia of the mussel of the genus *Unio* from Sicily to Sardinia or *vice versa*, suggested by Marrone and co-authors [74]. At the present state of the art, the absence of a genetic similarity between Sardinian and Sicilian populations would lead to rejection of this hypothesis, at least until samples that come from watercourses that flow into the Tyrrhenian Sea are analyzed. Indeed, the Sicilian individual was collected in the Frattina creek, a tributary of the Belice river, which flows into the Sicilian Sea, at the edge of the northern Siculo-Tunisian Strait. It is a biogeographic area separating the west and east Mediterranean in which different marine animals, with different dispersal capabilities, are anyway genetically separated from their Tyrrhenian counterpart (see [75–78] and references therein). Thus, to get a correct sight of the biogeographic position of Sicilian *S. fluviatilis*, a higher number of individuals, from different sites in Sicily, should be analyzed. A wider sampling campaign would also allow for enlarging knowledge on the genetic distribution of this island populations.

The genetic relationship that was found in the present study between samples of *S. fluviatilis* from Piedmont and Lombardy can be the result of movements of individuals from the Balkans through the Po river paleo-drainage that during the last glacial maximum included some Balkan tributaries. Indeed, the Balkan Peninsula was one of the main glacial refugia during the Pleistocene period, probably acting as a crossroad of different developmental processes [79]. Furthermore, our results demonstrated that, among Italian samples, northern populations showed the highest level of genetic variation. In contrast, Sardinian individuals showed several shared haplotypes with Ligurian populations, along with a few further private linages, which likely derived from Ligurian ancestors/founders.

4.2. Reconstruction of Phylogeography of Salariopsis fluviatilis in the Sardinia Island

Overall, a generally low level of genetic variation among the *S. fluviatilis* samples analyzed here has been found in populations from Sardinia. As expected, this finding reflects the typical evolutionary model of species dispersal on an island. Taking as a model the migration patterns reported for the *Salmo trutta* lineages [80–83], the following scenario could be invoked to explain our results concerning Sardinian populations of the species: the current distribution pattern of *S. fluviatilis* may be linked to the glaciation episodes that occurred during the Quaternary. Indeed, the climatic oscillations throughout this geological event may have promoted the migration of some individuals of *S. fluviatilis* (see Sanz et al. [84]) from an ancient dispersal center, identifiable in the Iberian Peninsula, as suggested by both the high level of genetic variability and the occurrence of several divergent mitochondrial lineages in that geographic area. Therefore, the spreading of *S. fluviatilis* might have reached Sardinia, following a stepping-stone migration model, starting from the Iberian Peninsula, and passing through Liguria and Corsica. In this context, it is worth mentioning that the island of Sardinia was alternately connected by land with the Corsica Island from the Miocene up to Pleistocene, and during the glaciation acmes, the distance between facing coasts of Corsica and the mainland substantially reduced, probably favoring some adults dispersal by sea. This possibility should be taken into

consideration as Plaut [10] demonstrated that *S. fluviatilis* can survive and osmoregulate in seawater for a period of at least three months.

Due to low sea levels during Quaternary climatic oscillations, Sardinia may have maintained contact with the mainland via the Island of Elba, and during the Last Glacial Maximum (21 kya), connections between Corsica (a possible refuge area for *S. fluviatilis* during the late Pleistocene) and the Sardinian islands were renewed [16,25]. Consequently, according to a model proposed for the Sardinian populations of the brackish water fish *Syngnathus abaster* [41], the land bridges naturally created in the Quaternary between Tuscany and the islands of Elba, Corsica, and Sardinia might have facilitated the occurrence of gene flow between the continental populations of *S. fluviatilis* and those of the islands [85].

The generally low level of genetic differentiation among the Sardinian samples evidenced in the present study would represent the typical effect of genetic drift, which acts as an evolutionary force by means of the founder effect. Indeed a few lineages, also common in Liguria, might have reached Sardinia, becoming the ancestors of modern populations. According to this model, the frequency of founder lineages grows quickly in the areas recently colonized and new, poorly divergent and private haplotypes arise from the first strains that arrived.

A model of a stepping-stone for the Sardinian freshwater species and the genetic structuring between flanking Italian regions separated by the Alps could be taken into account to explain the results obtained by Wagner et al. [16], which demonstrated that haplotypes from Corsican rivers diverged from those found in the Italian Alpine lakes, instead grouping with those from rivers of various localities in France and northern Spain. These authors reported a genetic similarity of Corsican samples not only with French and Spanish populations but also with the very few individuals from Sardinia they analyzed. This suggests that a brackish bridge between Italy and Corsica during the ice melting following a glacial period could have prompted gene flow between the island and mainland populations [25,86]. Genetic similarity between samples from Corsica and those from the southern part of mainland France was also found by Laporte et al. [25] based on nuclear markers. In this context, Wagner et al. [16], claim that a limited taxon and geographical sampling prevented previous studies from providing a comprehensive picture of genetic and biogeographic relationships among and within major freshwater *Salariopsis* lineages. The first data from Ligurian populations obtained in the present study shed further light on the populations of Sardinia and Corsica Islands.

4.3. Uprising of Distinct Taxonomic Entities within Salariopsis fluviatilis

The large number of *S. fluviatilis* sequences obtained in this study from previously unexplored geographic regions provided new insights into the taxonomic status of this freshwater blenny in its whole distribution area. Indeed, the analysis of a large set of Control Region sequences from the whole range of distribution of the Mediterranean *Salariopsis* species allowed us to evidence the occurrence, and the temporal origin, of different taxonomic units within the genus *Salariopsis* in the Mediterranean freshwaters.

Remarkably, the molecular dating estimates obtained in the present study predate those provided by Wagner et al. [16] for the same (when possible) groups of populations. This finding may be a consequence of both the higher numbers of Italian sequences used in the present study and the different datasets used for molecular dating that might account for the discrepancies between the two studies. Furthermore, although we used the same substitution rate as Wagner et al. [16], which was based on the mitochondrial Control Region of the blennioid *Tripterygium delaisi* [63], such a substitution rate was here applied to a dataset including only mitochondrial sequences (Control Region), whereas Wagner et al. [16] used this substitution rate to perform a molecular dating based on a concatenated dataset, including both the mitochondrial Control Region and the nuclear first intron of the S7 ribosomal protein gene sequences, which has a notably slower evolutionary rate.

Wagner et al. [16] and Belaiba et al. [15] evidenced the presence of genetic divergence of *S. fluviatilis* populations from the eastern Mediterranean region and suggested that this

group of individuals should be considered as *Salariopsis* cf. *fluviatilis*. Accordingly, results obtained in the present study show the occurrence of (1) genetic differentiation between two main groups of *S. fluviatilis* populations within the Mediterranean basin; (2) a group of populations from the Middle East represented by individuals from southern Turkey (but also one from the Ilica river in the north of the country), Israel and Syria which corresponds to *S.* cf. *fluviatilis* in accordance with Belaiba et al. [15]; and (3) a group of sequences of Iberian individuals from the basin of the Guadiana River which flows into the Atlantic Ocean.

Within the Mediterranean clade of *S. fluviatilis*, sequences from Italy (with the regions of Piedmont, Lombardy, and Sicily Island), Greece, Croatia, Albania, and part of Turkey (western coasts areas on the Aegean and Marmara seas) are representative of the North Oriental genetic cluster, whereas sequences from Spain, Italy (with Liguria and the Sardinia Island), Algeria, France mainland, and the Corsica Island represent the Occidental genetic cluster. Interestingly, individuals from the basin of the Guadiana River, which is the only investigated freshwater basin whose tributaries flow in the Atlantic, are slightly differentiated by the Occidental genetic cluster, suggesting an incipient genetic divergence of this population from the Mediterranean clade of *S. fluviatilis*.

The tolerance of freshwater blennies, and *S. fluviatilis* too [10], to a wide range of salinities might have favored the evolution from marine to freshwater forms, going through brackish, oligosaline and freshwater conditions. This might have encouraged the appearance of multiple discernible forms [16]. In particular, taking into consideration that the Guadiana genetic cluster is quite recent, according to the hypothesis proposed by Perdices et al. [7], Almada et al. [8] and Laporte et al. [9], it could have originated during the Pleistocene post-glacial event, between the end of Mindel and the beginning of Würm glaciations, from ancestors that were able to disperse via marine environments through adults dispersal. As an alternative hypothesis, we should take into consideration that *S. fluviatilis* colonized the upper Guadiana basin starting from river catches of the Júcar basin, as likely happened to the strictly freshwater fish *Luciobarbus guiraonis* (Cypriniformes: Cyprinidae: Barbinae) [87,88]. Nonetheless, at the present state of knowledge, the possibility of dispersal through marine environments is equally plausible, due to the high tolerance to the seawater osmolality of *S. fluviatilis*, at least for some months [10].

Furthermore, in accordance with Wagner et al. [16], within the genetic variation of *S. fluviatilis*, our results also evidenced the genetic divergence of the group of a few Spanish and Algerian sequences from the area of the Alboran Sea, and the group of sequences from Crete Island, which likely originated in the interglacial period between the Mindel and the Würm glaciations.

Interestingly, the *S. fluviatilis* population of Crete may be regarded as a peripheral isolate (sensu Frey, [89]), namely, a genetically isolated yet persistent population on the margin of the species' existing range. Indeed, populations of *S. fluviatilis* may have diverged in Crete more than would be predicted from the degree of variation of the other Mediterranean populations, due to the very far distance of this island from the Mediterranean mainland coastline.

A similar phenomenon of both partial geographic isolation and local adaptation that prompt high genetic differentiation may also be invoked for the westernmost populations of *S. fluviatilis*, which live at the edge of the distribution of the species. The Alboran Sea can be considered a geographic area where surface marine circulation patterns may directly shape the genetic variation of marine organisms, and indirectly act on the genetic variation of brackish and freshwater species, thus producing a genetic divergence with other Mediterranean populations (e.g., Casu et al. [76]).

Interestingly, within the taxonomic entity *S. fluviatilis sensu stricto*, the geographic isolation among populations and the reduced gene flow prompted the genetic divergence among Occidental and Oriental groups of populations. In such a context, natural fragmentation or secondary contact between populations that were separated in the past but are currently present in adjacent areas may have further helped to generate the genetic

differentiation observed between Liguria and Piedmont in northern Italy. However, such a divergence is not enough to be completely retrieved by species delimitation methods as a trace of incipient speciation, but it is suggestive of a relevant trend of ongoing genetic divergence.

Based on the divergence times estimated for the genetic clusters retrieved in the present study, it is possibly suggesting that the common ancestor to the Mediterranean species of *Salariopsis* originated approximately 1 million of years ago, before the beginning of Pleistocene glaciations. Then, repeated glacial fluctuations may have contributed to shaping the genetic divergence among fragmented populations of freshwater blennies, which were substantiated by speciation events in the genus *Salariopsis*. In such a context, the current distribution of *S. fluviatilis* and *S.* cf. *fluviatilis* lineages is the consequence of the Pleistocene post-glacial events that occurred between the end of Günz and the beginning of Mindel glaciations and that copious events of deglaciation may have led the structure of *S. fluviatilis* populations to be further shaped by the passage through different levels of salinity across river networks [25] and Pleistocene glacial fluctuations [16].

It is reasonable to consider the possibility that the evolutionary stages evidenced in the present study led *S. fluviatilis* to develop into a species complex as a consequence of past geological phenomena and present geographical boundaries. In this context, it is interesting to note that *S. fluviatilis* is separated from *S.* cf. *fluviatilis* by a number of mutations (23) that is greater than the number of mutational steps that separates it from the species *S. economidisi* (15), which is endemic to the western Greek Lake Trichonis [2]. Indeed, according to our results, which are based on the reduced number of sequences available on GenBank for this latter species, *S. economidisi* originated at the end of the interglacial period between the Mindel and Würm glaciations being likely contemporaneous with the Guadiana River taxonomic unit and to Alboran Sea and Crete Island genetic clusters.

Salariopsis fluviatilis, *S. economidisi* and the Guadiana River taxonomic unit share a unique common ancestor that likely originated during the Günz glaciation and that is contemporaneous with the ancestor of *S.* cf. *fluviatilis*. In accordance with Perdices et al. [7] and Neat et al. [6], these two ancient and contemporaneous ancestors might have belonged to one of the marine blennies species that differentiated in freshwater courses during Pleistocene glacial fluctuations and colonized the surrounding areas, which are interconnected by a large network of tributaries.

For this reason, at the current state of knowledge, the *S.* cf. *fluviatilis* taxonomic entity could be representative of a new species, distinct from *S. fluviatilis*, as *S. economidisi* and *S. atlantica*, which is pending a formal morphological description and new genetic analyses. In this context, it is interesting to note that *S.* cf. *fluviatilis* shows a genetic similarity higher with the species *S. atlantica*, which is endemic to the Seboui river basin in Morocco [17], rather than with the Algerian sequences of *S. fluviatilis sensu strictu*. This finding suggests that the Mediterranean area of North Africa could be inhabited by different freshwater species belonging to the genus *Salariopsis*, which may be characterized by relevant levels of genetic structuring so far undetected. What is reported here is in accordance with the phylogeographical patterns observed in other Maghrebian freshwater species (see, e.g., [90] and references therein). A similar trend was also hypothesized for the North African populations of the brackish water fish *Syngnathus abaster*, whose northern Tunisian individuals showed mitochondrial lineages belonging to either the westernmost Mediterranean genetic cluster or the Tunisian [41].

Interestingly, the genetic divergence found for the Guadiana River taxonomic entity and also for the two divergent clusters, found within *S. fluviatilis*, for the Alboran Sea area and Crete Island, may be the result of the genesis and consequent isolation of glacial refugia during the last glacial maximum (Würm), which has promoted genetic exchanges among river basins.

Overall, in the present study, the similar temporal estimates obtained by molecular dating for the origin of (1) *S. fluviatilis* and *S.* cf. *fluviatilis* (approximately 600 kya each) and of (2) the clusters of *S. economidisi*, Guadiana River, Alboran Sea, and Crete Island (on

average 135 kya) suggest the contemporaneous origin of these taxonomic entities. This finding supports the occurrence of similar evolutionary rates for these groups produced by the perfect timing of the mitochondrial DNA molecular clock for all Mediterranean freshwater blennies whose common ancestors likely underwent the same selective pressure prompted by similar geological conditions and geographic barriers. On the other hand, our results point out that the origin of *S. atlantica* is more recent in respect of the differentiation of *S. fluviatilis* and *S.* cf. *fluviatilis*, dating back to the interglacial period (approximately 390 kya) between the end of Mindel and the beginning of Riss glaciations.

5. Conclusions

In the present study, the hypervariable mitochondrial Control Region was shown to be highly informative in depicting the phylogeographic patterns of Italian populations of *Salariopsis fluviatilis sensu stricto* across the Ligurian Alpine ridge and the Sardinia Island-mainland dispersal patterns, with an origin and evolutionary trajectory of founders. In particular, this research contributes to put another piece in the mosaic of the molecular data of the Italian populations and helped to understand the processes of genetic differentiation involving Mediterranean island populations whose limited effective size and high degree of fragmentation make them ideal reference models for evolutionary processes.

Furthermore, the new sequences we obtained and the use of five species delimitation methods helped to obtain a simultaneous uprising of distinct taxonomic entities during the Pleistocene glaciations within the Mediterranean freshwater fish belonging to the genus *Salariopsis*. Indeed, our results support the possible occurrence of a complex of species for *S. fluviatilis*, namely *S.* cf. *fluviatilis* (as already suggested in Belaiba et al. [15]) in the Middle East regions and the Guadiana River basin taxonomic entity. In the future, the use of a combined approach with mitochondrial and nuclear markers (e.g., microsatellites, SNPs) to expand genetic variation analyses may be considered [36,37]. Indeed, previous studies on *S. fluviatilis* [9,10,20,25] evidenced that microsatellites are efficient tools in genetic monitoring to evidence genetic differences among and within populations and to highlight recent speciation processes [91].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ani12233403/s1, Figure S1. Likelihood mapping of the Control Region (A), 16s (B), and concatenated (C) dataset. The three trapezoids at the corners represent the areas supporting strictly bifurcating trees (the occurrence of a tree-like phylogenetic signal). The three rectangles on the sides represent regions where the decision between two trees is not obvious. The center of the triangle represents the region where all three unrooted trees are equally supported. Figure S2. Bayesian phylogenetic tree based on whole Control Region dataset. Values of node supports are expressed as posterior probabilities. Sequences from the present study are in red font. Figure S3. Ultrametric tree obtained by the software Beast on the whole Control Region dataset. It shows divergence time among taxa. Values of divergence time at the main node of the tree are rounded to the second decimal place. Figure S4. Median-joining network analysis performed on the whole 16s dataset. The number of mutations between sequences that are greater than $n = 1$ are reported on network branches. The number of individuals showing the same haplotype that is greater than $n = 1$ is reported inside the spots. Table S1. *Salariopsis fluviatilis* sampled specimens. The table reports data on the sampling collection and the GenBank accession numbers of the sequences obtained in the present study. Table S2. Control Region dataset. The table reports the Control Region sequences used in the present study that were taken from the GenBank database. Table S3. 16s dataset. Whole dataset including the 16s sequences taken from GenBank. Table S4. Species delimitations analyses. The table reports the results obtained from the species delimitation methods for the analysis of the whole Control Region dataset. Specimens with identical numbers within the same column belong to the same taxonomic entity. Table S5. Principal coordinates analysis. The table reports the results of the principal coordinates analysis performed on the dataset including the Control Region sequences obtained in the present study. Table S6. Principal coordinates analysis. The table reports results obtained from the principal coordinates analysis performed on the whole Control Region dataset.

Author Contributions: Conceptualization, I.A., F.S., M.C. and D.S.; methodology, I.A. and F.S.; software, I.A. and F.S.; validation, D.S. and M.C.; formal analysis, I.A. and F.S.; investigation, I.A. and F.S.; resources, D.S. and M.C.; data curation, I.A., F.S., S.B., A.N., F.O. and G.S.; writing—original draft preparation, I.A., F.S., D.S. and M.C.; writing—review and editing, I.A., F.S., D.S., P.C., C.L., S.B. and M.C.; visualization, I.A., F.S., D.S., P.C., C.L., S.B., A.N., F.O. and M.C.; supervision, D.S. and M.C.; project administration, D.S. and M.C.; funding acquisition, D.S. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FONDAZIONE DI SARDEGNA bando 2022–2023 for the Dipartimento di Scienze Biomediche—UNISS (to Daria Sanna and Fabio Scarpa).

Institutional Review Board Statement: The protocol of sampling and analysis of the fish fauna of wadable lotic systems, provided by the Italian Higher Institute for Environmental Protection and Research (ISPRA) [39], was followed for the sampling collection in the present study. In accordance with this document (the guidelines of which are compulsory in Italy), which requires that all electrically stunned fish be collected, recorded and returned to the water, the individuals of *S. fluviatilis* analyzed in this study were caught using an electric stunner from freshwaters, subjected to a non-lethal sampling method by means of small tissue portion removal (fin-clips) and immediately transferred to a recovery tank before being released. Tissues were preserved in absolute ethanol and used to perform DNA extraction. The above reported sampling method was approved by the ethics committee of the University of Sassari (Prot n. 122 770 of 7 November 2022) and its researchers led the sample collection activities during the present study.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequences obtained in the present study for the mitochondrial Control Region and 16s gene isolated in the present study were deposited in the GenBank database under the accession numbers OP675744-OP675846 and OP653611-OP653713, respectively.

Acknowledgments: We are truly indebted to Alessandro Candiotto (NaturaStaff Hydrobiologist, Mongardino, Italy), Luca Ciuffardi, Matteo Capurro, Fabrizio Oneto (Ce.S.Bi.N. srl, Via San Vincenzo n. 2, 16121 Genova) and Dario Ottonello (ARPAL, Via Bombrini n. 8, 16149 Genova) for their help in the sampling collection and their precious and invaluable support during the preliminary phases of this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vecchioni, L.; Ching, A.C.; Marrone, F.; Arculeo, M.; Hundt, P.J.; Simons, A.M. Multi-locus phylogenetic analyses of the Almadablennius clade reveals inconsistencies with the present taxonomy of blenniid fishes. *Diversity* **2022**, *14*, 53. [CrossRef]
2. Kottelat, M. *Salaria economidisi*, a new species of freshwater fish from Lake Trichonis, Greece, with comments on variation in *S. fluviatilis* (Teleostei: Blenniidae). *Rev. Suisse Zool.* **2004**, *111*, 121–137. [CrossRef]
3. Zander, C.D. Blenniidae. In *Fishes of the North-Eastern Atlantic and the Mediterranean*; Whitehead, P.J.P., Bauchot, M.-L., Hureau, J.-C., Nielsen, J., Tortonese, E., Eds.; UNESCO: Paris, France, 1986; pp. 1096–1112.
4. Duquenne-Delobel, E.; Doadrio, I.; Denys, G.P. Revalidation of the genus *Ichthyocoris* Bonaparte, 1840 (Actinopterygii: Blenniiformes: Blenniidae). *Acta Ichthyol. Piscat.* **2022**, *52*, 35–41. [CrossRef]
5. Vecchioni, L.; Arculeo, M.; Hundt, P.J.; Marrone, F. The valid genus name of the European freshwater blennies, *Ichthyocoris* or *Salariopsis* (Teleostei: Blenniidae)? *Zootaxa* **2022**, *5162*, 99–100. [CrossRef] [PubMed]
6. Neat, F.C.; Lengkeek, W.; Westerbeek, E.P.; Laarhoven, B.; Videler, J.J. Behavioural and morphological differences between lake and river populations of *Salaria fluviatilis*. *J. Fish Biol.* **2003**, *63*, 374–387. [CrossRef]
7. Perdices, A.; Doadrio, I.; Côté, I.M.; Machordom, A.; Economidis, P.; Reynolds, J.D. Genetic divergence and origin of Mediterranean populations of the river blenny *Salaria fluviatilis* (Teleostei: Blenniidae). *Copeia* **2000**, *2000*, 723–731. [CrossRef]
8. Almada, V.C.; Robalo, J.I.; Levy, A.; Freyhof, J.; Bernardi, G.; Doadrio, I. Phylogenetic analysis of peri-Mediterranean blennies of the genus *Salaria*: Molecular insights on the colonization of freshwaters. *Mol. Phylogenet. Evol.* **2009**, *52*, 424–431. [CrossRef]
9. Laporte, M.; Perrier, C.; Magnan, P.; Berrebi, P. Genetic evidence of recent migration among isolated-by-sea populations of the freshwater blenny (*Salaria fluviatilis*). *Conserv. Genet.* **2016**, *17*, 389–399. [CrossRef]
10. Plaut, I. Comparison of salinity tolerance and osmoregulation in two closely related species of blennies from different habitats. *Fish Physiol. Biochem.* **1998**, *19*, 181–188. [CrossRef]
11. Vinyoles, D.; Côté, I.M.; De Sostoa, A. Nest orientation patterns in *Salaria fluviatilis*. *J. Fish Biol.* **2002**, *61*, 405–416. [CrossRef]

12. Magnan, P.; Proulx, R.; Berrebi, P.; Blondel, J.; Perret, P.; Roché, B. Morphological variation in the freshwater blenny *Salaria fluviatilis* from Corsican rivers: Adaptive divergence, phenotypic plasticity or both? *J. Fish Biol.* **2014**, *84*, 31–44. [CrossRef] [PubMed]
13. Laporte, M.; Claude, J.; Berrebi, P.; Perret, P.; Magnan, P. Shape plasticity in response to water velocity in the freshwater blenny *Salaria fluviatilis*. *J. Fish Biol.* **2016**, *88*, 1191–1203. [CrossRef] [PubMed]
14. Elvira, B.; Nicola, G.G.; Ayllón, D.; Almodóvar, A. Determinants of large-scale spatial distribution and seasonal microhabitat selection patterns of the endangered freshwater blenny *Salaria fluviatilis* in the Ebro River basin, Spain. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2021**, *31*, 3261–3275. [CrossRef]
15. Belaiba, E.; Marrone, F.; Vecchioni, L.; Bahri-Sfar, L.; Arculeo, M. An exhaustive phylogeny of the combtooth blenny genus *Salaria* (Pisces, Blenniidae) shows introgressive hybridization and lack of reciprocal mtDNA monophyly between the marine species *Salaria basilisca* and *Salaria pavo*. *Mol. Phylogenet. Evol.* **2019**, *135*, 210–221. [CrossRef] [PubMed]
16. Wagner, M.; Zogaris, S.; Berrebi, P.; Freyhof, J.; Koblmüller, S.; Magnan, P.; Laporte, M. Diversity and biogeography of Mediterranean freshwater blennies (Blenniidae, *Salaria*). *Divers. Distrib.* **2021**, *27*, 1832–1847. [CrossRef]
17. Doadrio, I.; Perea, S.; Yahyaoui, A. Morphological and molecular analyses of freshwater blennids: A new species of the genus *Salaria* Forsskål, 1775 (Actinopterygii, Blennidae) in Morocco. *Graellsia* **2011**, *67*, 151–173. [CrossRef]
18. Cote, I.M.; Vinyoles, D.; Reynolds, J.D.; Doadrio, I.; Perdices, A. Potential impacts of gravel extraction on Spanish populations of river blennies *Salaria fluviatilis* (Pisces, Blenniidae). *Biol. Conserv.* **1999**, *87*, 359–367. [CrossRef]
19. Hernandez, R.; Lacomba, R.T.; Uviñtas, Y.N.; Oltra, R. Distribution pattern of river blennies in the Júcar River basin (eastern Spain). *J. Fish Biol.* **2000**, *57*, 250–254. [CrossRef]
20. Méndez, L.; Perdices, A.; Machordom, A. Genetic structure and diversity of the Iberian populations of the freshwater blenny *Salaria fluviatilis* (Asso, 1801) and its conservation implications. *Conserv. Genet.* **2019**, *20*, 1223–1236. [CrossRef]
21. Freeman, M.C.; Viñolas, D.; Grossman, G.D.; De Sostoa, A. Microhabitat use by *Blennius fluviatilis* in the Río Matarraña, Spain. *Freshw. Biol.* **1990**, *24*, 335–345. [CrossRef]
22. Alp, A.; Kara, C. Distribution pattern and morphological differences between the sexes of river Blenny, *Salaria fluviatilis* (Asso, 1801), in the Ceyhan river basin, Turkey. *Turk. J. Zool.* **2007**, *31*, 113–120.
23. Tutman, P.; Glamuzina, B. Detection and Confirmation of Freshwater Blenny *Salaria fluviatilis* (Actinopterygii: Blenniidae) in Bosnia and Herzegovina. *Croat. J. Fish.* **2021**, *79*, 75–82. [CrossRef]
24. Almada, F.; Almada, V.C.; Guillemaud, T.; Wirtz, P. Phylogenetic relationships of the north-eastern Atlantic and Mediterranean blenniids. *Biol. J. Linn. Soc.* **2005**, *86*, 283–295. [CrossRef]
25. Laporte, M.; Leblois, R.; Coulon, A.; Bonhomme, F.; Magnan, P.; Berrebi, P. Genetic structure of a vulnerable species, the freshwater blenny (*Salaria fluviatilis*). *Conserv. Genet.* **2015**, *16*, 571–581. [CrossRef]
26. Zava, B.; Violani, C. Contributi alla conoscenza dell'ittiofauna delle acque interne siciliane. I. Sulla presenza in Sicilia di *Salaria fluviatilis* (Asso, 1801) (Pisces, Blenniidae). *Boll. Mus. Reg. Sci. Nat. Torino* **1991**, *9*, 313–324.
27. Gallo, L.; Lucadamo, L.; Mezzotero, A.; Morisi, A.; Battegazzore, M.; Fenoglio, S. First data on the freshwater fish fauna of Calabria (southern Italy). *Ital. J. Zool.* **2012**, *79*, 246–251. [CrossRef]
28. Lorenzoni, M.; Borghesan, F.; Carosi, A.; Ciuffardi, L.; De Curtis, O.; Delmastro, G.; Di Tizio, L.; Franzoi, P.; Maio, G.; Mojetta, A.; et al. Check-list dell'ittiofauna delle acque dolci italiane. *Italy J. Freshw. Ichthyol.* **2019**, *5*, 239–254.
29. Fricke, R.; Eschmeyer, W.N.; Van der Laan, R. *Catalog of Fishes: Genera, Species, References*; California Academy of Sciences: San Francisco, CA, USA, 2018.
30. Orrù, F.; Deiana, A.M.; Cau, A. Introduction and distribution of alien freshwater fishes on the island of Sardinia (Italy): An assessment on the basis of existing data sources. *J. Appl. Ichthyol.* **2010**, *26*, 46–52. [CrossRef]
31. Orrù, F.; Deiana, A.M.; Cau, A. Conservazione della biodiversità e specie alloctone invasive: Il gambero *Procambarus clarkii* nelle acque interne della Sardegna. *Studi Trentini Sci. Nat.* **2009**, *86*, 162.
32. Orrù, F.; Niffoi, A.; Sanna, D.; Varcasia, A.; Buscarinu, P.; Casu, M. Prima segnalazione di *Leucos aula* (Teleostei: Cyprinidae) in Sardegna basata su analisi morfologiche e DNA barcoding. *Italy J. Freshw. Ichthyol.* **2019**, *5*, 166–172.
33. Casu, M.; Scarpa, F.; Cossu, P.; Lai, T.; Curini-Galletti, M.; Varcasia, A.; Sanna, D. First record of *Esox cisalpinus* (Teleostea: Esocidae) in Sardinia with insight on its mitochondrial DNA genetic variability. *Ital. J. Zool.* **2016**, *83*, 514–523. [CrossRef]
34. Sanna, D.; Azzena, I.; Scarpa, F.; Cossu, P.; Pira, A.; Gagliardi, F.; Casu, M. First Record of the Alien Species *Procambarus virginalis* Lyko, 2017 in FreshWaters of Sardinia and Insight into Its Genetic Variability. *Life* **2021**, *11*, 606. [CrossRef] [PubMed]
35. Nabholz, B.; Glémin, S.; Galtier, N. Strong variations of mitochondrial mutation rate across mammals—The longevity hypothesis. *Mol. Biol. Evol.* **2008**, *25*, 120–130. [CrossRef]
36. Wan, Q.H.; Wu, H.; Fujihara, T.; Fang, S.G. Which genetic marker for which conservation genetics issue? *Electrophoresis* **2004**, *25*, 2165–2176. [CrossRef] [PubMed]
37. Primmer, C.R. From conservation genetics to conservation genomics. *Ann. N. Y. Acad. Sci.* **2009**, *1162*, 357–368. [CrossRef] [PubMed]
38. Galtier, N.; Nabholz, B.; Glémin, S.; Hurst, G.D. Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Mol. Ecol.* **2009**, *18*, 4541–4550. [CrossRef] [PubMed]
39. Biologici, I.G.D.L.M. Metodi Biologici per Le Acque Superficiali Interne. In *Manuali e Linee Guida*; ISPRA-Settore Editoria: Roma, Italy, 2014; ISBN 978-88-448-0651.

40. Sanna, D.; Scarpa, F.; Lai, T.; Cossu, P.; Falautano, M.; Castriota, L.; Andaloro, F.; Follesa, M.C.; Francalacci, P.; Curini-Galletti, M.; et al. *Fistularia commersonii* (Teleostea: Fistulariidae): Walking through the Lessepsian paradox of mitochondrial DNA. *Ital. J. Zool.* **2015**, *82*, 499–512. [CrossRef]
41. Sanna, D.; Biagi, F.; Alaya, H.B.; Maltagliati, F.; Addis, A.; Romero, A.; De Juan, J.; Quignard, J.P.; Castelli, A.; Franzoi, P.; et al. Mitochondrial DNA variability of the pipefish *Syngnathus abaster*. *J. Fish Biol.* **2013**, *82*, 856–876. [CrossRef]
42. Sievers, F.; Higgins, D.G. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In *Multiple Sequence Alignment Methods*; Russell, D., Ed.; Humana Press: Totowa, NJ, USA, 2014; Volume 1079, pp. 105–116. [CrossRef]
43. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
44. Schmidt, H.A.; Strimmer, K.; Vingron, M.; Von Haeseler, A. TREE-PUZZLE: Maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **2002**, *18*, 502–504. [CrossRef]
45. Scarpa, F.; Cossu, P.; Sanna, D.; Lai, T.; Casu, M.; Curini-Galletti, M. New insights on the genus *Otoplana* Du Plessis, 1889 (Platyhelminthes: Proseriata), with description of two new species from the Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2075–2087. [CrossRef]
46. Scarpa, F.; Sanna, D.; Cossu, P.; Lai, T.; Casu, M.; Curini-Galletti, M. How to achieve internal fertilization without a vagina: The study case of the genus *Archilina* Ax, 1959 (Platyhelminthes, Proseriata) from Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2057–2073. [CrossRef]
47. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef] [PubMed]
48. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [CrossRef]
49. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **1980**, *16*, 111–120. [CrossRef]
50. Scarpa, F.; Cossu, P.; Sanna, D.; Lai, T.; Norenburg, J.L.; Curini-Galletti, M.; Casu, M. An 18S and 28S-based clock calibration for marine Proseriata (Platyhelminthes). *J. Exp. Mar. Biol. Ecol.* **2015**, *463*, 22–31. [CrossRef]
51. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef]
52. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]
53. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
54. Ezard, T.H.; Pearson, P.N.; Purvis, A. Algorithmic approaches to aid species' delimitation in multidimensional morphospace. *BMC Evol. Biol.* **2010**, *10*, 175. [CrossRef]
55. Scarpa, F.; Cossu, P.; Lai, T.; Sanna, D.; Curini-Galletti, M.; Casu, M. Meiofaunal cryptic species challenge species delimitation: The case of the *Monocelis lineata* (Platyhelminthes: Proseriata) species complex. *Contrib. Zool.* **2016**, *85*, 123–145. [CrossRef]
56. Zhang, J.; Kapli, P.; Pavlidis, P.; Stamatakis, A. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **2013**, *29*, 2869–2876. [CrossRef]
57. Puillandre, N.; Brouillet, S.; Achaz, G. ASAP: Assemble species by automatic partitioning. *Mol. Ecol. Resour.* **2020**, *21*, 609–620. [CrossRef] [PubMed]
58. Scarpa, F.; Cossu, P.; Delogu, V.; Lai, T.; Sanna, D.; Leasi, F.; Norenburg, J.L.; Curini-Galletti, M.; Casu, M. Molecular support for morphology-based family-rank taxa: The contrasting cases of two families of Proseriata (Platyhelminthes). *Zool. Scr.* **2017**, *46*, 753–766. [CrossRef]
59. Hebert, P.D.; Cywinska, A.; Ball, S.L.; Dewaard, J.R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* **2003**, *270*, 313–321. [CrossRef]
60. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]
61. Tran Thi, G.; Azzena, I.; Scarpa, F.; Cossu, P.; Danh Le, C.; Ton Nu, P.A.; Chau Ngo, T.M.; Sanna, D.; Casu, M. Molecular Identification and Appraisal of the Genetic Variation of *Taenia saginata* in Central Regions of Vietnam. *Life* **2022**, *12*, 70. [CrossRef]
62. Drummond, A.J.; Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **2007**, *7*, e214. [CrossRef]
63. Koblmüller, S.; Steinwender, B.; Weiß, S.; Sefc, K.M. Gene flow, population growth and a novel substitution rate estimate in a subtidal rock specialist, the black-faced blenny *Tripterygion delaisi* (Perciformes, Blennioidei, Tripterygiidae) from the Adriatic Sea. *J. Zool. Syst. Evol.* **2015**, *53*, 291–299. [CrossRef]
64. Rambaut, A.; Suchard, M.A.; Xie, D.; Drummond, A.J. Tracer v1.6, 2014. Available online: http://tree.bio.ed.ac.uk/software/tracer/ (accessed on 19 October 2022).
65. Strimmer, K.; von Haeseler, A. Quartet Puzzling: A quartet Maximum-Likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **1996**, *13*, 964–969. [CrossRef]
66. Schmidt, H.A.; von Haeseler, A. Phylogenetic inference using maximum likelihood methods. In *The Phylogenetic Handbook*, 5th ed.; Lemey, P., Salemi, M., Vandamme, A.M., Eds.; Cambridge University Press: Cambridge, UK, 2009; pp. 181–209.

67. Hwang, U.W.; Kim, W. General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *KJP* **1999**, *37*, 215. [CrossRef] [PubMed]
68. Stefani, F.; Galli, P.; Zaccara, S.; Crosa, G. Genetic variability and phylogeography of the cyprinid *Telestes muticellus* within the Italian peninsula as revealed by mitochondrial DNA. *J. Zool. Syst. Evol. Res.* **2004**, *42*, 323–331. [CrossRef]
69. Ciuffardi, L.; Maldini, M.; Balduzzi, A.; Arillo, A. Caratterizzazione meristica e genetica di Ciprinidi atipici presenti nel bacino spezzino del Magra-Vara-Atti XII Conv. naz. A.I.I.A.D., San Michele all'Adige (TN). *Studi Trentini Sci. Nat.-Acta Biol.* **2010**, *87*, 171–174.
70. Ciuffardi, L.; Balduzzi, A.; Nonnis Marzano, F.; Arillo, A. *Analisi Morfometrica di Leuciscini Atipici Presenti Nel Bacino del Vara-Magra (La Spezia, Italia Nord-Occidentale) (Pisces, Osteichthyes, Cyprinidae)*; Doriana: Genova, Italy, 2011; Volume 8, pp. 1–10.
71. Ciuffardi, L.; Monaci, E.; Civitella, D.; Balduzzi, A.; Palanti, E.; Nonnis Marzano, F.; Arillo, A. Indagini Sulla Genetica e Sulla Biologia Riproduttiva di Ciprinidi Atipici, Appartenenti Alla Sottofamiglia Leuciscinae, Presenti in Provincia Della Spezia-Università Degli Studi di Genova e Polizia Provinciale Della Spezia Sez. Faunistica 2011.
72. Nonnis Marzano, F. Caratterizzazione Genetica di Ciprinidi Atipici Della Provincia di La Spezia-Spin off "Gen-Tech" Università degli Studi di Parma 2010.
73. Vamberger, M.; Stuckas, H.; Sacco, F.; D'Angelo, S.; Arculeo, M.; Cheylan, M.; Corti, C.; Lo Valvo, M.; Marrone, F.; Wink, M.; et al. Differences in gene flow in a twofold secondary contact zone of pond turtles in southern Italy (Testudines: Emydidae: *Emys orbicularis galloitalica*, *E. o. hellenica*, *E. trinacris*). *Zool. Scr.* **2015**, *44*, 233–249. [CrossRef]
74. Marrone, F.; Nardi, G.; Cianfanelli, S.; GovediČ, M.; Barra, S.A.; Arculeo, M.; Bodon, M. Diversity and taxonomy of the genus *Unio Philipsson* in Italy, with the designation of a neotype for *Unio elongatulus* C. Pfeiffer (Mollusca, Bivalvia, Unionidae). *Zootaxa* **2019**, *4545*, 339–374. [CrossRef]
75. Cossu, P.; Dedola, G.L.; Scarpa, F.; Sanna, D.; Lai, T.; Maltagliati, F.; Curini-Galletti, M. Casu, M. Patterns of spatial genetic variation in *Patella ulyssiponensis*: Insights from the western Mediterranean marine ecoregion. *Hydrobiologia* **2015**, *755*, 39–55. [CrossRef]
76. Casu, M.; Rivera-Ingraham, G.A.; Cossu, P.; Lai, T.; Sanna, D.; Dedola, G.L.; Sussarellu, R.; Sella, G.; Cristo, B.; Curini-Galletti, M.; et al. Patterns of spatial genetic structuring in the endangered limpet *Patella ferruginea*: Implications for the conservation of a Mediterranean endemic. *Genetica* **2011**, *139*, 1293–1308. [CrossRef]
77. Maltagliati, F.; Lupi, L.; Castelli, A.; Pannacciulli, F.G. The genetic structure of the exotic ascidian *Styela plicata* (Tunicata) from Italian ports, with a re-appraisal of its worldwide genetic pattern. *Mar. Ecol.* **2016**, *37*, 492–502. [CrossRef]
78. Drábková, M.; Jachníková, N.; Tyml, T.; Sehadová, H.; Ditrich, O.; Myšková, E.; Hypša, V.; Štefka, J. Population codivergence in common cuttlefish (*Sepia officinalis*) and its dicyemid parasite in the Mediterranean Sea. *Sci. Rep.* **2019**, *9*, 14300. [CrossRef] [PubMed]
79. Marić, S.; Sušnik, S.; Simonović, P.; Snoj, A. Phylogeographic study of brown trout from Serbia, based on mitochondrial DNA control region analysis. *Genet. Sel. Evol.* **2006**, *38*, 411. [CrossRef] [PubMed]
80. Bernatchez, L. The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution* **2001**, *55*, 351–379. [CrossRef] [PubMed]
81. Splendiani, A.; Fioravanti, T.; Giovannotti, M.; Negri, A.; Ruggeri, P.; Olivieri, L.; Cerioni, P.N.; Lorenzoni, M.; Caputo Barucchi, V. The effects of paleoclimatic events on Mediterranean trout: Preliminary evidences from ancient DNA. *PLoS ONE* **2016**, *11*, e0157975. [CrossRef] [PubMed]
82. Splendiani, A.; Palmas, F.; Sabatini, A.; Caputo Barucchi, V. The name of the trout: Considerations on the taxonomic status of the *Salmo trutta* L., 1758 complex (Osteichthyes: Salmonidae) in Italy. *Eur. Zool. J.* **2019**, *86*, 432–442. [CrossRef]
83. Splendiani, A.; Berrebi, P.; Tougard, C.; Righi, T.; Reynaud, N.; Fioravanti, T.; Lo Conte, P.; Delmastro, G.B.; Baltieri, M.; Ciuffardi, L.; et al. The role of the south-western Alps as a unidirectional corridor for Mediterranean brown trout (*Salmo trutta* complex) lineages. *Biol. J. Linn. Soc.* **2020**, *131*, 909–926. [CrossRef]
84. Sanz, N.; García-Marín, J.L. Pla, Divergence of brown trout (*Salmo trutta*) within glacial refuge. *Can. J. Fish. Aquat.* **2000**, *57*, 2201–2210. [CrossRef]
85. Grill, A.; Casula, P.; Lecis, R.; Menken, S. Endemism in Sardinia. In *Phylogeography of Southern European Refugia*; Weiss, S., Ferrand, N., Eds.; Springer: Dordrecht, The Netherland, 2007; pp. 273–296. [CrossRef]
86. Gauthier, A.; Berrebi, P. La Colonisation de l'Ile par les différentes Souches de Truite. In *Contribution à la Gestion des Populations de Truites en Corse*; Muracciole, S., Ed.; Fédération Corse pour la Pêche et la Protection du Milieu Aquatique: Corte, France, 2007; pp. 4–10.
87. Doadrio, I.; Perea, S.; Garzón-Heydt, P.; González, Y.J.L. Ictiofauna continental española. In *Bases Para su Seguimiento. DG Medio Natural y Política Forestal*; MARM: Madrid, Spain, 2011; 616p.
88. Doadrio, I. *Atlas y Libro Rojo de Los Peces Continentales de España*; MNCN-CSIC: Madrid, Spain, 2001; p. 364.
89. Frey, J.K. Modes of peripheral isolate formation and speciation. *Syst. Biol.* **1993**, *42*, 373–381. [CrossRef]
90. Marrone, F.; Vecchioni, L.; Deidun, A.; Mabrouki, Y.; Arab, A.; Arculeo, M. DNA taxonomy of the potamid freshwater crabs from Northern Africa (Decapoda, Potamidae). *Zool. Scr.* **2020**, *49*, 473–487. [CrossRef]
91. Segherloo, I.H.; Freyhof, J.; Berrebi, P.; Ferchaud, A.L.; Geiger, M.; Laroche, J.; Levin, B.A.; Normandeau, E.; Bernatchez, L. A genomic perspective on an old question: *Salmo* trouts or *Salmo trutta* (Teleostei: Salmonidae)? *Mol. Phylogenet. Evol.* **2021**, *162*, 107204. [CrossRef]

Chapter 4

Procambarus virginalis

1. General description

Procambarus virginalis Lyko, 2017 (Crustacea, Decapoda, Astacida), commonly known as the marbled crayfish or marmorkrebs, is a unique species among freshwaters crayfish due to its ability to reproduce via apomictic parthenogenesis (females produce offspring without mating) (Martin et al. 2007; Martin et al. 2010; Chucholl et al. 2012; Lyko 2017). The asexual reproduction capability of *Procambarus virginalis* enables it to quickly establish and expand its population, posing a significant challenge for other freshwater species (Chucholl et al. 2012).

1.1. Possible origin of *Procambarus virginalis*

Procambarus virginalis was initially identified in the German aquarium pet trade during the mid-1990s (Lukhaup 2001). However, determining its origin remained a mystery during that period because the trader's unreliable information made it challenging to determine the crayfish's geographical source (Lukhaup in 2003).

Following its discovery, the marbled crayfish gained popularity among aquarists thanks to its attractive marbled coloration, rapid reproductive capabilities, and ease care (Vogt et al. 2008). Aquarium enthusiasts initially speculated that the marbled crayfish might be a hermaphrodite or possibly a parthenogenetic species. However, during that period, parthenogenesis was not yet known to occur in crayfish or other decapods (Vogt et al. 2008).

It was initially thought that this species derived from the congeneric *Procambarus fallax* (Hagen, 1870), native to Florida, due to morphological similarities. In this context, Scholtz et al. (2003) conducted a comparison of sequences from two mitochondrial genes of closely related species, with special attention to *Procambarus fallax.* The genetic

sequences obtained, and the phylogenetic analysis performed, indicated a particularly close relationship with *Procambarus fallax* (Scholtz et al. 2003). Furthermore, this research also presented evidence of the marbled crayfish's parthenogenetic capability.

Subsequent investigations involving molecular analysis, morphological comparisons, and the study of postembryonic development strongly supported the conclusion that the marbled crayfish belonged to the broader North American Cambaridae group (Vogt et al. 2004). However, given that the marbled crayfish is a thelytokous parthenogenetic species, it was not possible to precisely classify this species since, conventional methods of identification based on male gonopod characteristics (Hobbs in 1972; Hobbs 1989), were not applicable.

Otherwise, following these initial studies, many other researchers have put forth various theories regarding the origins of the marbled crayfish. Some have suggested that it may be a parthenogenetic form of *Procambarus alleni* (Faxon, 1884) (see Vogt 2008; Martin et al. in 2010). Moreover, Martin et al. (2010), considered the possibility of designating the marbled crayfish as a new species, potentially naming it *Procambarus fallax* forma *virginalis*. Others have proposed that it could have originated from hybridization between *Procambarus fallax* and other coexisting *Procambarus* species (Martin 2015; Scholtz 2015; Faulkes 2016).

Subsequently, in a distinct study that focused on crossbreeding and parentage analysis, Vogt et al. (2015) determined that the marbled crayfish and *Procambarus fallax* exhibit distinct reproductive strategies. According to the authors, the marbled crayfish derives from *Procambarus fallax* through triploidization and simultaneous epigenetic modifications (Vogt et al. 2015). Furthermore, morphological traits and microsatellite patterns strongly suggest that the marbled crayfish origin is associated with autopolyploidization rather than hybridization with a closely related species, which is the more common cause of triploidy in animals (Vogt et al. 2015). Indeed, the marbled crayfish lacks the morphological characteristics of hybrids (Kawai et al. 2009; Martin et al. 2010; Vogt et al. 2015), and there is no evidence of hybridization at the genetic level. Additionally, there is no significant bias toward heterozygosity in the microsatellite patterns, a typical feature of hybrids (Soltis & Soltis 2000; Alves et al. 2001). Moreover, among the seven microsatellite loci investigated in marbled crayfish, three were homozygous, and four were heterozygous (Martin et al. 2007; Vogt et al. 2008; Vogt et

al. 2015), which largely dismisses the possibility of allopolyploidization for marbled crayfish and further corroborates the notion that marbled crayfish originated through autopolyploidization (Vogt et al. 2015).

Subsequently, in accordance with Martin et al. (2010) recommendation to designate the marbled crayfish as a new species, Lyko formally described and named it in 2017, changing the proposed name from *Procambarus fallax* forma *virginalis* to *Procambarus virginalis* (Lyko 2017). It's worth noting that under the International Code of Zoological Nomenclature (ICZN, 1999), the use of "forma" for names published after 1960 is not permitted (Lyko 2017).

To date, despite the various plausible theories regarding the species development, the exact origin of the *Procambarus virginalis*, whether in captivity or in the wild, continues to elude our understanding (Vogt et al. 2015).

Being the marbled crayfish not native to Europe (Lyko 2017), it falls under the category of Non-Indigenous Crayfish Species (referred to as NICS, following Holdich 2002). Its presence in Europe carries the potential to negatively impact a wide variety of freshwater species and disrupt local ecosystems due to its ability for parthenogenetic reproduction, remarkable adaptability to new environments, and resilience in challenging ecological conditions (Kaldre et al. 2015; Veselý et al. 2017). The presence of these characteristics, combined with irresponsible practices within the aquarium pet trade that increased the risk of invasive spread (Chucholl 2014; Patoka et al. 2017), designates this crayfish species as one of the most potentially hazardous freshwater invaders (Nentwig et al. 2018).

1.2. Non-Indigenous Crayfish Species (NICS) in Europe

Approximately one-third to one-half of crayfish species worldwide are estimated to be at risk of declining populations or extinction (Taylor 2002) and, one of the major causes is the presence of NICS (Holdich et al. 2009). Non-native species in Europe currently exceed indigenous crayfish species (ICS) by a ratio of 2:1 and, it has been projected that NICS could entirely supplant ICS in the coming decades if protective measures are not taken (Holdich et al. 2009).

In this context, a series of freshwater crayfish species have been introduced in Europe, most of them from North America, while only two from Australia (Holdich et al. 2009).

Among these crayfish species, three North American NICS were originally introduced to Europe with the intent of restocking crayfish populations. The first documented introduction took place in the 1890s when the spiny-cheek crayfish, *Faxonius limosus* (Rafinesque, 1817), native to Pennsylvania (USA), was transported to Barnowko village in Poland (Gherardi & Holdich 1999).

The other two species that were introduced at that time were *Pacifastacus leniusculus* (Dana, 1852), introduced to Sweden in 1959 (Abrahamsson 1973), and *Procambarus clarkii* (Girard, 1852), introduced to Spain in 1973 (Habsburgo-Lorena 1979). These three species are often referred to as "Old NICS" since they were introduced before 1975. In contrast, the "New NICS" have been introduced since 1980, primarily through the aquarium trade and for aquaculture (Holdich et al. 2009). *Procambarus virginalis* stands within this latter group of NICS.

1.3. Distribution of *Procambarus virginalis*

The popularity of *Procambarus virginalis* among hobbyists often lead to the rapid overpopulation of aquariums (Kouba et al. 2014; Kouba et al. 2021). The overpopulation not only results in an excess of crayfish being supplied to other aquarium enthusiasts and pet shops but also contributes to an increased number of releases into natural environments (Holdich et al. 2009; Chucholl et al. 2012). In this context, the first report regarding the presence of *Procambarus virginalis* in the wild can be traced back to 2003 in Germany (Marten et al. 2004). Subsequently, in 2004, a few specimens were found on land after a canal cleaning operation in the Netherlands (Soes & van Eekelen 2006). Following these first cases, numerous other reports began to be documented.

In 2005 the species was reported even in Madagascar by researchers from the University of Antananarivo who observed a distinct decapod species being sold in local markets. Later, in 2007, a team led by Tadashi Kawai from the Wakkanai Fisheries Experiment Station in Japan tentatively identified these specimens as *Procambarus virginalis* (Jones et al. 2009).

In 2008, a single marbled crayfish specimen was reported within an established population of *Procambarus clarkii* in a slow-flowing canal in Tuscany, Italy (Marzano et al. 2009). In 2010, the situation underwent a significant change as established

populations of marbled crayfish were identified not only in Germany but also in Slovakia (Janský & Mutkovic 2010; Chucholl et al. 2012). Further reports of isolated specimens were subsequently made from Germany, and since the overview provided by Chucholl et al. (2012), at least three more records have been reported from Germany. Some of these reports may indicate the presence of established populations (C. Chucholl, pers. comm.).

In December 2012, a concerning report came from Sweden, where 13 marbled crayfish specimens were discovered in the River Märstaån (central Sweden) (Bohman et al. 2013). However, subsequent attempts to confirm the presence of this species were unsuccessful, leaving uncertainty regarding its ability to establish a reproducing population under Scandinavian climatic conditions (Bohman et al. 2013).

Over the subsequent years, numerous additional reports emerged, and summarily, to date, the presences of *Procambarus virginalis* has been reported in 22 countries: Italy (Marzano et al. 2009; Vojkovská et al. 2014; Deidun et al. 2018), France (Collas 2019), Belgium (Scheers et al. 2021), Germany and Netherlands (Souty-Grosset et al. 2006; Holdich et al. 2007), Denmark (see Vogt 2023 and reference therein), Sweden (Bohman et al. 2013), Austria (Aigner 2022), Croatia (Cvitanić et al. 2016), Romania (Pârvulescu et al. 2017), Slovakia (Lipták et al. 2017), Estonia (Ercoli et al. 2019), Czech Republic (Patoka et al. 2016), Hungary (Lőkkös et al. 2016), Poland (Maciaszek et al. 2022), Ukraine (Novitsky et al. 2016), Malta (Deidun et al. 2018), Israel (Carneiro et al. 2023), Madagascar (Jones et al. 2009), China (see Vogt 2023 and reference therein), Japan (Faulkes et al. 2012) and Taiwan (see Vogt 2023 and reference therein).

Noteworthy, in 2023, the first reports of *Procambarus virginalis* for the North America (Canada) were documented (refer to https://sites.google.com/view/marmorkrebs/ for more information).

1.4. Is it possible to control the spread of *Procambarus virginalis*?

NICS, such as *Procambarus virginalis*, are typically difficult to manage and highly resistant to eradication once they have established self-sustaining populations (Holdich et al. 1999; Peay et al. 2006; Aquiloni et al. 2009). Furthermore, eradicating *Procambarus virginalis* is especially challenging due to its parthenogenetic nature. The only possible avenue for restraining the spread of *Procambarus virginalis* might be associated with its

remarkably low levels of genetic diversity, making its populations highly susceptible to unforeseen environmental stresses (personal observation).

2. Implication related to the spread of *Procambarus virginalis* in freshwater habitats

From the very first articles about the discovery of *Procambarus virginalis*, the potential risks associated with its release into the natural environment have consistently been emphasized (Scholtz et al. 2003). This species possesses several traits that make it a significant threat to native European freshwater crayfish, as well as other freshwater species and their habitats (Chucholl et al. 2012).

First, being a parthenogenetic species, it can be considered as the perfect invader since, the introduction of even a solitary individual into the natural environment could result in the establishment of a population that may potentially compete with native crayfish in terms of food and shelter (Scholtz et al. 2003; Marten et al. 2004; Jones et al. 2009; Jimenez & Faulkes 2010; Martin et al. 2010).

Second, as an American species, it has the potential to act as a carrier of the so-called crayfish plague (Culas 2003), a disease caused by the fungus *Aphanomyces astaci* Schikora, 1906, that nearly led to the extinction of native European crayfish and continues to endanger both wild and farmed populations (Oidtmann et al. 1999).

Furthermore, there is supporting evidence indicating that crayfish could serve as carriers of the fungus *Batrachochytrium dendrobatidis*, which has the potential to cause diseases in native amphibian populations (see Battisti & Scalici 2020 and reference therein). Moreover, the competition for food resources, the grazing activity carried out by crayfish on periphyton and their predation on amphibian eggs and tadpoles may indirectly contribute to a decline in populations of amphibians and invertebrates (Klose & Cooper 2012; Souty-Grosset et al. 2016; Oficialdegui et al. 2019).

Additionally, like other freshwater crayfish species, their role as ecosystem engineers can have a significant impact on their habitat. They influence sediment transport through their continuous search for new refuges, which is linked to the changing characteristics of their habitat from riffles to pools and vice versa, as observed in species like *Orconectes limosus* (Statzner et al. 2003). Furthermore, their activities can

alter the roughness of the riverbed, change the proportion of sand in gravel interstices, modify sand coverage, or impact the extent of filamentous algae cover (Statzner et al. 2003). These different activities can result in complex effects, making it challenging to predict an overall outcome during subsequent floods (Statzner et al. 2003).

2.1. The crayfish plague

The introduction of the crayfish plague fungus *Aphanomyces astaci* into Europe is strictly linked to the introduction of North American crayfish (Söderhäll & Cerenius 1992). *Aphanomyces astaci*, which belongs to the *Saprolegniales*, is a mycosis that poses serious threat to both wild and farmed European freshwater crayfish species (Oidtmann et al. 1999). In the absence of a secure water supply, there remains an ongoing risk of contamination from *Aphanomyces astaci* from upstream sources. This risk primarily stems from the illegal release of non-native crayfish carrying the disease or through contaminated fishing equipment (Oidtmann et al. 1999).

During the past century, the distribution of NICS has significantly increased due to several factors (e.g., stocking activities, natural spread of stocked populations, and the release of exotic crayfish by private aquarists or pond owners). Therefore, in area where NICS have been introduced, there have been frequent records of mass mortalities among European crayfish (Alderman 1996).

North American crayfish species exhibit a low susceptibility to this disease (Söderhäll & Cerenius 1992). Indeed, they can carry the fungus on their cuticles for extended periods of time as a benign infection. However, they might succumb to the disease when subjected to stress (Söderhäll & Cerenius 1992). In contrast, European crayfish species are highly susceptible to the disease and often suffer from infections (Oidtmann et al. 1999).

The remarkable resistance of NICS to *Aphanomyces astaci* can be attributed to their shared North American origin, and their coevolution history which has played a crucial role in the observed resistance within North American freshwater crayfish species (Unestam 1969; Unestam 1972; Unestam 1975). Consequently, their interaction appears to establish a relatively stable host-pathogen relationship, with infrequent severe outbreaks, in contrast to the severe threats that *Aphanomyces astaci* can caused in European freshwater habitat (Alderman 1996).

3. Aims of the study

Based on the above background, during my PhD project, I focused my attention on the examination of *Procambarus virginalis* individuals collected in Sardinia. The basis for this research stemmed from the preliminary reports of specimens found in the central and southern regions of the island, which were tentatively identified as *Procambarus virginalis*. In this context, the following paper aims, as its primary objective, to perform a taxonomic identification of the individuals collected in Sardinia by using molecular tools and a species delimitation approach. Second, to evaluate the genetic variability of *Procambarus virginalis* in Sardinia, and third to carry out phylogenetic and phylogeographic analyses on the population established in Sardinia, which was the only known population in Italy. To provide a broader geographical context for the dataset, the analyses included the comparison of sequences from Sardinian with those from other countries available in GenBank.

References

- Abrahamsson, S. (1973). Methods for restoration of crayfish waters in Europe–the development of an industry for production of young *Pacifastacus leniusculus*. Freshwater Crayfish, 1, 203-210.
- Aigner, K. (2022). Ecology of marbled crayfish and possible management measures at fishing ponds in Salzburg. Master's thesis, Department of Environment and Biodiversity, University of Salzburg.
- Alderman, D. J. (1996). Geographical spread of bacterial and fungal diseases of crustaceans. Revue scientifique et technique (International Office of Epizootics), 15(2), 603-632.
- Alves, M. J., Coelho, M. M., & Collares-Pereira, M. J. (2001). Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. Genetica, 111, 375-385.
- Aquiloni, L., Becciolini, A., Berti, R., Porciani, S., Trunfio, C., & Gherardi, F. (2009). Managing invasive crayfish: use of X-ray sterilisation of males. Freshwater Biology, 54(7), 1510-1519.
- Battisti, C., & Scalici, M. (2020). First records of the red swamp crayfish *Procambarus clarkii* (Girard, 1852) (Decapoda Cambaridae) from a small circum-Sardinian island (central Mediterranean Sea). BioInvasions Records, 9(2), 333-339.
- Bohman, P., Edsman, L., Martin, P., & Scholtz, G. (2013). The first Marmorkrebs (Decapoda: Astacida: Cambaridae) in Scandinavia. BioInvasions Record, 2(3).
- Carneiro, V. C., & Galil, B. (2023). A voyage into the Levant: the first record of a marbled crayfish *Procambarus virginalis* (Lyko, 2017) population in Israel. BioInvasions Records, 12(3), 829-836.
- Chucholl, C. (2014). Predicting the risk of introduction and establishment of an exotic aquarium animal in Europe: insights from one decade of Marmorkrebs (Crustacea, Astacida, Cambaridae) releases. Management of Biological Invasions, 5(4), 309.
- Chucholl, C., Morawetz, K., & Groß, H. (2012). The clones are coming--strong increase in Marmorkrebs [*Procambarus fallax* (Hagen, 1870) f. *virginalis*] records from Europe. Aquatic invasions, 7(4).
- Collas, M. 2019. Premier signalement de l'Écrevisse marbrée (Procambarus virginalis) en France.
- Culas, A. (2003). Entwicklung einer molekularbiologischen Methode zum Nachweis des Krebspesterregers Aphanomyces astaci SCHIKORA in nordamerikanischen Flusskrebsen (*Pacifastacus leniusculus*; *Orconectes limosus*; *Procambarus clarkii*) (Doctoral dissertation, lmu).
- Cvitanić, M., Hudina, S., & Maguire, I. (2016, September). Reproductive cycle of the marble crayfish from an established population in Croatia. In Proceedings of the 21st Symposium of the International Association of Astacology-Program and Book of Abstracts (p. 48).
- Deidun, A., Sciberras, A., Formosa, J., Zava, B., Insacco, G., Corsini-Foka, M., & Crandall, K. A. (2018). Invasion by non-indigenous freshwater decapods of Malta and Sicily, central Mediterranean Sea. Journal of Crustacean Biology, 38(6), 748-753.
- Ercoli, F., Kaldre, K., Paaver, T., & Gross, R. (2019). First record of an established marbled crayfish *Procambarus virginalis* (Lyko, 2017) population in Estonia. BioInvasions Records, 8(3).
- Faulkes, Z., Feria, T. P., & Muñoz, J. (2012). Do Marmorkrebs, *Procambarus fallax* f. *virginalis*, threaten freshwater Japanese ecosystems?. Aquatic biosystems, 8(1), 1-12.
- Faulkes, Z., Kawai, T., & Scholtz, G. (2016). Marble crayfish as a new model organism and a new threat to native crayfish conservation (pp. 31-53). CRC Press, Boca Raton, FL, USA.
- Gherardi, F., & Holdich, D. M. (Eds.). (1999). Crayfish in Europe as alien species. CRC Press.
- Habsburgo, L. Present situation of exotic species of crayfish introduced into Spanish continental waters. Present situation of exotic species of crayfish introduced into Spanish continental waters, 175-184.

- Hobbs Jr, H. H. (1989). An illustrated checklist of the American crayfishes (Decapoda, Astacidae, Cambaridae, Parastacidae). pp 1-236.
- Hobbs, H. H. (1972). Crayfishes (Astacidae) of North and Middle America. Biota of Freshwater Ecosystems. Identification Manual 9. Environmental Protection Agency, Washington, DC, 115. pp 1-173.
- Holdich, D. M. (2002). Distribution of crayfish in Europe and some adjoining countries. Bulletin Français de la Pêche et de la Pisciculture, (367), 611-650.
- Holdich, D. M., & Pöckl, M. (2007). Invasive crustaceans in European inland waters. In Biological invaders in inland waters: Profiles, distribution, and threats (pp. 29-75). Dordrecht: Springer Netherlands.
- Holdich, D. M., Gydemo, R., & Rogers, W. D. (1999). Possible methods for controlling alien crayfish populations. CRUSTACEAN ISSUES, 11, 245-270.
- Holdich, D. M., Reynolds, J. D., Souty-Grosset, C., & Sibley, P. J. (2009). A review of the ever increasing threat to European crayfish from non-indigenous crayfish species. Knowledge and management of aquatic ecosystems, (394-395), 11.
- Janský, V., & Mutkovič, A. (2010). Rak *Procambarus* sp. (Crustacea: Decapoda: Cambaridae)–Prvý nález na Slovensku. Acta Rerum Naturalium Musei Nationalis Slovaci, 56, 64-67.
- Jimenez, S. A., & Faulkes, Z. (2011). Can the parthenogenetic marbled crayfish Marmorkrebs compete with other crayfish species in fights?. Journal of Ethology, 29, 115-120.
- Jones, J. P., Rasamy, J. R., Harvey, A., Toon, A., Oidtmann, B., Randrianarison, M. H., ... & Ravoahangimalala, O. R. (2009). The perfect invader: a parthenogenic crayfish poses a new threat to Madagascar's freshwater biodiversity. Biological Invasions, 11, 1475-1482.
- Kaldre, K., Meženin, A., Paaver, T., & Kawai, T. (2016). A preliminary study on the tolerance of marble crayfish *Procambarus fallax* f. *virginalis* to low temperature in nordic climate. In: Freshwater crayfish: global overview, 54-62.
- Kawai, T., Scholtz, G., Morioka, S., Ramanamandimby, F., Lukhaup, C., & Hanamura, Y. (2009). Parthenogenetic alien crayfish (Decapoda: Cambaridae) spreading in Madagascar. Journal of Crustacean Biology, 29(4), 562-567.
- Klose, K., & Cooper, S. D. (2012). Contrasting effects of an invasive crayfish (*Procambarus clarkii*) on two temperate stream communities. Freshwater Biology, 57(3), 526-540.
- Kouba, A., Lipták, B., Kubec, J., Bláha, M., Veselý, L., Haubrock, P. J., ... & Buřič, M. (2021). Survival, growth, and reproduction: comparison of marbled crayfish with four prominent crayfish invaders. Biology, 10(5), 422.
- Lipták, B., Mojžišová, M., Gruľa, D., Christophoryová, J., Jablonski, D., Bláha, M., ... & Kouba, A. (2017). Slovak section of the Danube has its well-established breeding ground of marbled crayfish *Procambarus fallax* f. *virginalis*. Knowledge & Management of Aquatic Ecosystems, (418), 40.
- Lőkkös, A., Müller, T., Kovács, K., Várkonyi, L., Specziár, A., & Martin, P. (2016). The alien, parthenogenetic marbled crayfish (Decapoda: Cambaridae) is entering Kis-Balaton (Hungary), one of Europe's most important wetland biotopes. Knowledge and Management of Aquatic Ecosystems, (417), 16.
- Lukhaup, C. (2001). *Procambarus* sp.–der Marmorkrebs. Aquaristik aktuell, 7(8), 48-51.
- Lukhaup, C. (2003). Süwasserkrebse aus aller Welt. Dähne, Ettlingen.
- Lyko, F. (2017). The marbled crayfish (Decapoda: Cambaridae) represents an independent new species. Zootaxa, 4363(4), 544-552.
- Maciaszek, R., Jabłońska, A., Prati, S., Wróblewski, P., Gruszczyńska, J., & Świderek, W. (2022). Marbled crayfish *Procambarus virginalis* invades a nature reserve: how to stop further introductions?. The European Zoological Journal, 89(1), 888-901.
- Marten, M., Werth, C., & Marten, D. (2004). Der Marmorkrebs (Cambaridae, Decapoda) in Deutschland–ein weiteres Neozoon im Einzugsgebiet des Rheins. Lauterbornia, 50, 17-23.

- Martin, P. (2015). Parthenogenesis: mechanisms, evolution, and its relevance to the role of marbled crayfish as model organism and potential invader. Freshwater crayfish: a global overview. pp 63-82.
- Martin, P., Dorn, N. J., Kawai, T., van der Heiden, C., & Scholtz, G. (2010). The enigmatic Marmorkrebs (marbled crayfish) is the parthenogenetic form of *Procambarus fallax* (Hagen, 1870). Contributions to Zoology, 79(3), 107-118.
- Martin, P., Kohlmann, K., & Scholtz, G. (2007). The parthenogenetic Marmorkrebs (marbled crayfish) produces genetically uniform offspring. Naturwissenschaften, 94, 843-846.
- Marzano, F. N., Scalici, M., Chiesa, S., Gherardi, F., Piccinini, A., & Gibertini, G. (2009). The first record of the marbled crayfish adds further threats to fresh waters in Italy. Aquatic Invasions, 4(2), 401-404.
- Nentwig, W., Bacher, S., Kumschick, S., Pyšek, P., & Vilà, M. (2018). More than "100 worst" alien species in Europe. Biological Invasions, 20(6), 1611-1621.
- Novitsky, R. A., Novitsky, R. O., Novitskiy, R. A., Novitskiy, R. O., Novickii, R. A., Novitskij, R. A., ... & Son, M. (2016). The first records of Marmorkrebs [*Procambarus fallax* (Hagen, 1870) f. *virginalis*](Crustacea, Decapoda, Cambaridae) in Ukraine.
- Oficialdegui, F. J., Sánchez, M. I., Monsalve-Carcaño, C., Boyero, L., & Bosch, J. (2019). The invasive red swamp crayfish (*Procambarus clarkii*) increases infection of the amphibian chytrid fungus (*Batrachochytrium dendrobatidis*). Biological Invasions, 21, 3221-3231.
- Oidtmann, B., Cerenius, L., Schmid, I., Hoffmann, R., & Söderhäll, K. (1999). Crayfish plague epizootics in Germany--classification of two German isolates of the crayfish plague fungus *Aphanomyces astaci* by random amplification of polymorphic DNA. Diseases of Aquatic Organisms, 35(3), 235-238.
- Pârvulescu, L., Togor, A., Lele, S. F., Scheu, S., Șinca, D., & Panteleit, J. (2017). First established population of marbled crayfish *Procambarus fallax* (Hagen, 1870) f. *virginalis* (Decapoda, Cambaridae) in Romania. BioInvasions Record, 6(4).
- Patoka, J., Bláha, M., Kalous, L., & Kouba, A. (2017). Irresponsible vendors: Non-native, invasive and threatened animals offered for garden pond stocking. Aquatic Conservation: Marine and Freshwater Ecosystems, 27(3), 692-697.
- Patoka, J., Buřič, M., Kolář, V., Bláha, M., Petrtýl, M., Franta, P., ... & Kouba, A. (2016). Predictions of marbled crayfish establishment in conurbations fulfilled: Evidences from the Czech Republic. Biologia, 71(12), 1380-1385.
- Peay, S., Hiley, P. D., Collen, P., & Martin, I. (2006). Biocide treatment of ponds in Scotland to eradicate signal crayfish. Bulletin Français de la Pêche et de la Pisciculture, (380-381), 1363-1379.
- Scheers, K., Brys, R., Abeel, T., Halfmaerten, D., Neyrinck, S., & Adriaens, T. (2021). The invasive parthenogenetic marbled crayfish *Procambarus virginalis* Lyko, 2017 gets foothold in Belgium. BioInvasions Records, 10(2), 326-340.
- Scholtz, G. (2015). Happy birthday! The first decade of Marmorkrebs research – results and perspectives. In Freshwater Crayfish: A Global Overview. Ed. T. Kawai, Z. Faulkes and G. Scholtz, pp. 3-12. Boca Raton: CRC Press.
- Scholtz, G., Braband, A., Tolley, L., Reimann, A., Mittmann, B., Lukhaup, C., ... & Vogt, G. (2003). Parthenogenesis in an outsider crayfish. Nature, 421(6925), 806-806.
- Soderhall, K. (1987). Controlled growth and development in filamentous Oomycetes with emphasis on *Aphanomyces* spp. Zoosporic fungi in teaching and research, 264-267.
- Soes, D. M., & van Eekelen, R. (2006). Rivierkreeften, een oprukkend probleem?. De Levende Natuur, 107(2), 56-59.
- Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences, 97(13), 7051-7057.
- Souty-Grosset, C., Anastacio, P. M., Aquiloni, L., Banha, F., Choquer, J., Chucholl, C., & Tricarico, E. (2016). The red swamp crayfish *Procambarus clarkii* in Europe: impacts on aquatic ecosystems and human well-being. Limnologica, 58, 78-93.

- Statzner, B., Peltret, O., & Tomanova, S. (2003). Crayfish as geomorphic agents and ecosystem engineers: effect of a biomass gradient on baseflow and flood-induced transport of gravel and sand in experimental streams. Freshwater Biology, 48(1), 147-163.
- Unestam, T. (1969). Resistance to the crayfish plague in some American, Japanese and European crayfishes. Report of the Institute of Freshwater Research, Drottningholm, 49, 202-209.
- Unestam, T. (1975). Defence reactions in and susceptibility of Australian and New Guinean freshwater crayfish to European-crayfish-plague fungus. Australian Journal of Experimental Biology and Medical Science, 53(5), 349-359.
- Unestem, T. (1972). On the host range and origin of the crayfish plague fungus. Report / Institute of Freshwater Research, Drottningholm, 52, 192-198.
- Veselý, L., Hrbek, V., Kozák, P., Buřič, M., Sousa, R., & Kouba, A. (2017). Salinity tolerance of marbled crayfish *Procambarus fallax* f. *virginalis*. Knowledge & Management of Aquatic Ecosystems, (418), 21.
- Vogt, G. (2008). The marbled crayfish: a new model organism for research on development, epigenetics and evolutionary biology. Journal of Zoology, 276(1), 1-13.
- Vogt, G. (2023). Phenotypic plasticity in the monoclonal marbled crayfish is associated with very low genetic diversity but pronounced epigenetic diversity. Current Zoology, 69(4), 426-441.
- Vogt, G., Tolley, L., & Scholtz, G. (2004). Life stages and reproductive components of the Marmorkrebs (marbled crayfish), the first parthenogenetic decapod crustacean. Journal of morphology, 261(3), 286-311.
- Vojkovská, R., Horká, I., Tricarico, E., & Ďuriš, Z. (2014). New record of the parthenogenetic marbled crayfish *Procambarus fallax* f. *virginalis* from Italy. Crustaceana, 87(11-12), 1386-1392.

life

MDPI

Article

First Record of the Alien Species *Procambarus virginalis* Lyko, 2017 in Fresh Waters of Sardinia and Insight into Its Genetic Variability

Daria Sanna [1,*,†], Ilenia Azzena [1,2,†], Fabio Scarpa [2], Piero Cossu [2], Angela Pira [3], Flavio Gagliardi [3] and Marco Casu [2]

[1] Dipartimento di Scienze Biomediche, Università di Sassari, Viale San Pietro 43/B, 07100 Sassari, Italy; iazzena@uniss.it
[2] Dipartimento di Medicina Veterinaria, Università di Sassari, Via Vienna 2, 07100 Sassari, Italy; fscarpa@uniss.it (F.S.); picossu@uniss.it (P.C.); marcasu@uniss.it (M.C.)
[3] Acquario di Cala Gonone, Via La Favorita, 08020 Cala Gonone, Italy; acquariologia@acquariocalagonone.it (A.P.); flaviogagliardi@panaque.com (F.G.)
* Correspondence: darsanna@uniss.it
† These authors contributed equally to this work.

Abstract: In the fresh waters of Sardinia (Italy), the non-indigenous crayfish species *Procambarus clarkii* has been reported from 2005, but, starting from 2019, there have been several reports of a new non-indigenous crayfish in southern and central areas of this Mediterranean island, and its morphology suggests that this species may be the marbled crayfish *Procambarus virginalis*. Forty-seven individuals of this putative species were analyzed, using the mitochondrial gene Cytochrome c Oxidase subunit I as molecular marker to identify this crayfish and investigate the level of genetic variability within the recently established population. Phylogenetic and phylogeographic analyses were carried out on a dataset including sequences from the Sardinian individuals and from all congenerics available in GenBank. Results showed that the new Sardinian crayfish belong to the species *P. virginalis*. All the sequences belonging to *P. virginalis* from European countries are identical, with only few exceptions found among Sardinian individuals. In conclusion, this paper highlights the occurrence of a new further alien species in the Sardinian fresh waters, which are already characterized by the high presence of non-indigenous species.

Keywords: alien species; invasive species; non-indigenous crayfish; biological invasion; marbled crayfish; Sardinian fresh waters; mtDNA

Citation: Sanna, D.; Azzena, I.; Scarpa, F.; Cossu, P.; Pira, A.; Gagliardi, F.; Casu, M. First Record of the Alien Species *Procambarus virginalis* Lyko, 2017 in Fresh Waters of Sardinia and Insight into Its Genetic Variability. *Life* **2021**, *11*, 606. https://doi.org/10.3390/life11070606

Academic Editors: Carl Beierkuhnlein and Anna Walentowitz

Received: 24 May 2021
Accepted: 22 June 2021
Published: 24 June 2021

1. Introduction

The so-called freshwater crayfish (Malacostraca, Decapoda) are a monophyletic group of crustaceans with challenging taxonomy and phylogeny [1–3], which are present in each continent, except for Antarctica and mainland Africa [4,5], and occur in almost every type of freshwater habitat, both lentic and lotic [5–7].

As they are among the largest invertebrate predators in their habitats, freshwater crayfish are an important component in the structuring of the aquatic fauna [8,9]. Indeed, many studies showed a clear impact as a keystone species, due to their feeding activity, which mainly involves vegetation, invertebrates, and vertebrates such as fish. Furthermore, as ecosystem engineers, they can create major impacts on their habitat, affecting the sediment transport as a consequence of their constant search for new refugia, from riffles to pools and vice versa [7,8].

Over the years, multiple crayfish species were translocated, either deliberately or accidentally [10–12]. For this reason, at least one non-indigenous crayfish species (hereafter NICS, following Holdich [13]) has been reported in most European countries, and the general number of NICS is rapidly growing [14–16].

The presence of NICS can have serious consequences on native ecosystems [17], not only due to interspecific competition with indigenous species or habitat modification, but also by carrying diseases, such as the "crayfish plague", i.e., *Aphanomyces astaci* [17]. This water mould is mainly transmitted by the largely resistant introduced North American crayfish species, and can lead to high levels of mortality, causing declines of indigenous fauna [18–20]. Furthermore, as this oomycete can survive on fish skin and use these animals as vectors, another way of transmission is represented by fish moving away from areas where the crayfish plague is present, which can thus spread the infection among drainage basins [21].

In Europe, there are few native species of crayfish [22–24], and the first documented introduction of a NICS dates back to 1890 in Barnowko village (Poland), where the spiny-cheek crayfish, *Faxonius limosus* (Rafinesque, 1817), was introduced for commercial purposes [25] from Pennsylvania (USA). It was followed by other NICS introductions, mainly *Procambarus clarkii* Girard, 1852 (firstly introduced in Spain in 1973), which is considered as a very harmful problematic NICS due to the plastic life cycle, capability of naturalization, and rapid potential for dispersal [26,27].

Later, in the 90s, in a German pet trade, a new species was identified for the first time, it was the marbled crayfish, Procambarus virginalis Lyko, 2017 [28].Preliminary studies hypothesized that *P. virginalis* might represent the product of hybridization between the sexually reproducing *Procambarus fallax* (Hagen, 1870) (from FL, USA) and its congeneric *P. clarkii* (from LA, USA) [29]. However, recent studies have highlighted the autopolyploid nature of the marbled crayfish, which is a triploid organism that differentiated from its mother species *P. fallax* [30,31]. All marbled crayfish show common phenotypic, genetic, and epigenetic characteristics, despite their broad geographical distribution [31] and reproduce by parthenogenesis, being able to quickly create "wild" populations in the temperate zones of the planet [32,33].

Considering the high interest in the marbled crayfish from aquarists, their presence in European ecosystems is probably due to the voluntary release into the wild carried out by some owners and traders [32–37]. To date, stable populations of *P. virginalis* in Europe are known from: Germany and Netherlands [10], Croatia [38], Romania [39], Slovakia [40], Estonia [41], Czech Republic [42], Hungary [43], Ukraine [44], and Sweden [45]. The spread of this species is not limited to Europe, in fact its presence has been also reported in Japan [46] and Madagascar [29,47]. Furthermore, although not yet mentioned in scientific publications, several individuals of marbled crayfish have been discovered in different freshwater areas of Poland, Taiwan, and Macau, based on recent local reports. In Italy, (where four autochthonous crayfish species occur [48]), the first report of *P. virginalis* dates back to 2008 [11], but, at present, stable populations are not reported.

In 2019, some individuals morphologically attributable to the marbled crayfish were found in freshwater habitats of southern Sardinia (Italy). Just one year later, individuals that can be identified as marbled crayfish have been found in many other areas of Sardinia (pers. obs.). In this Mediterranean island, two other species of crayfish are known so far: *Austropotamobius pallipes* Lereboullet, 1858, which is indigenous to Italy, and is rarely found in Sardinian freshwaters, where, likely, its occurrence is not natural but due to translocation events [49]; and the NICS *P. clarkii*, which was recorded for the first time in 2005 [50] and it is now widely distributed across the island.

This paper aimed to: (1) identify the new freshwater crayfish found in Sardinia, applying a species delimitation approach; (2) investigate the levels of the species' genetic variability in Sardinia; and (3) perform phylogenetic and phylogeographic analyses on the newly established population (the only known in Italy), comparing sequences of Sardinian individuals with those already available in the literature, in order to place the data obtained in a wider geographic background.

2. Materials and Methods

2.1. Sample Collection

Forty-seven individuals were analysed (Table 1), whose morphology was consistent with marbled crayfish or the closely related slough crayfish, *Procambarus fallax* following Hobbs [51].

Table 1. The table reports data on the sampling collection. Sampling sites for each geographic area are indicated for the individuals of *Procambarus virginalis* collected in Sardinia during the present study. Details are also provided for the GenBank sequences of *P. virginalis* and *P. fallax* from all over the world used for the analyses.

Sample ID	Sampling Site	Species	GB #	Coordinates		Habitat	Sampling Date
				Long	Lat		
PFMO1	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097549	8,68540286	39,6536863	Creek	May 2019
PFMO2	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097550	8,68540286	39,6536863	Creek	May 2019
PFMO3	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097551	8,68540286	39,6536863	Creek	May 2019
PFMO7	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097555	8,68540286	39,6536863	Creek	May 2019
PFOR1	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097559	8,571233244	39,72729907	Creek	May 2019
PFMO4	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097552	8,68540286	39,6536863	Creek	June 2019
PFMO5	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097553	8,68540286	39,6536863	Creek	June 2019
PFMO6	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097554	8,68540286	39,6536863	Creek	June 2019
PFMO8	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097556	8,68540286	39,6536863	Creek	June 2019
PFMO9	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097557	8,68540286	39,6536863	Creek	June 2019
PFMO10	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097558	8,68540286	39,6536863	Creek	June 2019
PFMO11	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097560	8,68540286	39,6536863	Creek	June 2019
PFMO12	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097561	8,68540286	39,6536863	Creek	June 2019
PFMO13	Italy (Sardinia)—Morimenta	*Procambarus virginalis*	MZ097562	8,68540286	39,6536863	Creek	June 2019
PFOR2	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097563	8,571233244	39,72729907	Creek	November 2019
PFOR5	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097564	8,571233244	39,72729907	Creek	November 2019
PFOR6	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097565	8,571233244	39,72729907	Creek	November 2019
PFOR7	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097566	8,571233244	39,72729907	Creek	November 2019
PFOR8	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097567	8,571233244	39,72729907	Creek	November 2019

Table 1. *Cont.*

Sample ID	Sampling Site	Species	GB #	Coordinates		Habitat	Sampling Date
				Long	Lat		
PFOR9	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097568	8,571233244	39,72729907	Creek	November 2019
PFOR10	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097569	8,571233244	39,72729907	Creek	November 2019
PFOR11	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097570	8,621508481	39,82035494	Creek	November 2019
PFOR12	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097571	8,621508481	39,82035494	Creek	November 2019
PFOR13	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097572	8,621508481	39,82035494	Creek	November 2019
PFOR14	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097573	8,621508481	39,82035494	Creek	November 2019
PFOR15	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097574	8,621508481	39,82035494	Creek	November 2019
PFOR16	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097575	8,578457118	39,82507761	Creek	November 2019
PFOR17	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097576	8,578457118	39,82507761	Creek	November 2019
PFOR18	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097577	8,549284943	39,74396601	Creek	November 2019
PFOR19	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097578	8,60873508	39,8203551	Creek	November 2019
PFOR20	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097579	8,60873508	39,8203551	Creek	November 2019
PFOR21	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097580	8,60873508	39,8203551	Creek	November 2019
PFOR22	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097581	8,60873508	39,8203551	Creek	November 2019
PFOR23	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097582	8,60873508	39,8203551	Creek	November 2019
PFOR24	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097583	8,60873508	39,8203551	Creek	November 2019
PFOR25	Italy (Sardinia)—Arborea	*Procambarus virginalis*	MZ097584	8,60873508	39,8203551	Creek	November 2019
PFSL1	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097585	8,842628592	39,51785049	Creek	December 2019
PFSL2	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097586	8,842628592	39,51785049	Creek	December 2019
PFSL3	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097587	8,842628592	39,51785049	Creek	December 2019
PFSL4	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097588	8,842628592	39,51785049	Creek	December 2019
PFSL5	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097589	8,842628592	39,51785049	Creek	December 2019
PFSL6	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097590	8,844295005	39,51785048	Creek	December 2019

Table 1. *Cont.*

Sample ID	Sampling Site	Species	GB #	Coordinates		Habitat	Sampling Date
				Long	Lat		
PFSL7	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097591	8,844295005	39,51785048	Creek	December 2019
PFSL8	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097592	8,844295005	39,51785048	Creek	December 2019
PFSL9	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097593	8,844295005	39,51785048	Creek	December 2019
PFSL10	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097594	8,844295005	39,51785048	Creek	December 2019
PFSL11	Italy (Sardinia)—Sanluri	*Procambarus virginalis*	MZ097595	8,844295005	39,51785048	Creek	December 2019
PFBE8	Belgium	*Procambarus virginalis*	LR884227	-	-	Pond	May 2020
PFBE9	Belgium	*Procambarus virginalis*	LR884226	-	-	Pond	May 2020
PFBE10	Belgium	*Procambarus virginalis*	LR884225	-	-	Moat	May 2020
PFBE5	Belgium	*Procambarus virginalis*	LR884230	-	-	Pond	June 2020
PFBE6	Belgium	*Procambarus virginalis*	LR884229	-	-	Pond	June 2020
PFBE7	Belgium	*Procambarus virginalis*	LR884228	-	-	Basin	June 2020
PFBE1	Belgium	*Procambarus virginalis*	LR884234	-	-	Moat	July 2020
PFBE2	Belgium	*Procambarus virginalis*	LR884233	-	-	Moat	July 2020
PFBE3	Belgium	*Procambarus virginalis*	LR884232	-	-	Pond	July 2020
PFBE4	Belgium	*Procambarus virginalis*	LR884231	-	-	Pond	July 2020
PFCR1	Czech Republic	*Procambarus virginalis*	MK439899	-	-	-	-
PFIT1	Italy (Veneto)	*Procambarus virginalis*	KJ690261	-	-	Channel	April 2009
PFJA1	Japan	*Procambarus virginalis*	LC228303	-	-	-	November 2016
PFGE2	Germany	*Procambarus virginalis*	HM358011	-	-	Stream	October 2009
PFGE5	Germany	*Procambarus virginalis*	KC107813	-	-	-	-
PFGE6	Germany	*Procambarus virginalis*	KT074364	-	-	-	-
PFGE4	Germany	*Procambarus virginalis*	HM358010	-	-	Stream	October 2009
PFGE1	Germany	*Procambarus fallax*	HM358012	-	-	-	October 2009

Table 1. *Cont.*

Sample ID	Sampling Site	Species	GB #	Coordinates		Habitat	Sampling Date
				Long	Lat		
PFGE3	Germany	*Procambarus fallax*	JF438007	-	-	Lake	-
PFGE8	Germany	*Procambarus fallax*	KT074365	-	-	-	-
PFGE7	Germany	*Procambarus fallax*	NC_020021	-	-	-	-
PFFL1	Florida—USA	*Procambarus fallax*	HQ171459	-	-	-	-
PFFL2	Florida—USA	*Procambarus fallax*	HQ171458	-	-	-	-
PFFL3	Florida—USA	*Procambarus fallax*	HQ171457	-	-	-	-
PFFL4	Florida—USA	*Procambarus fallax*	HQ171456	-	-	-	-
PFFL5	Florida—USA	*Procambarus fallax*	HQ171455	-	-	-	-
PFFL6	Florida—USA	*Procambarus fallax*	HQ171453	-	-	-	-
PFFL7	Florida—USA	*Procambarus fallax*	HQ171454	-	-	-	-
PFSW1	Sweden	*Procambarus fallax*	KF033123	-	-	River	December 2012

stands for accession number.

Forty-one adult individuals, which were tentatively identified as the marbled crayfish (*Procambarus virginalis*), were collected from May to December 2019 in the central and southern areas of Sardinia which are the alleged center of the first dispersion of this crayfish in the island (see Figure 1 and Table 1 for details). Six eggs (specimens PFMO8–PFMO13 in Table 1), taken from one of the adult individuals, were added to our molecular analyses. In particular, we performed a non-standardized survey in the three areas of the island where the possible occurrence of this species was reported by local communities (see Figure 1 and Table 1 for details on sampling stations). The waterbodies of these sampling areas are characterized by slow flowing, irregular streams and temporary natural or artificial ponds, mainly with muddy and poorly vegetated riverbeds. Natural streams are often connected one another by drainage canals, which greatly facilitate crayfish spreading. Live crayfish, collected using fish traps, were delivered to the laboratory to be sacrificed by an anaesthetic overdose, and then stored at −20 °C until the DNA extraction.

2.2. Diagnostic Molecular Analysis

Total genomic DNA was isolated from a portion of muscle tissue using the Macherey-Nagel Nucleo Spin Tissue Kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany) following the supplier's instructions. DNA solutions were quantified using the Nanodrop™ Lite Spectrophotometer (by Thermo Scientific; Waltham, MA, USA), which showed an average yield of approximately 30 ng/µL.

A fragment of the subunit I of the mitochondrial Cytochrome c Oxidase gene (COI) was amplified by standard PCR using universal primers [52]. Reactions were carried out in a total volume of 25 µL. On average, 10 ng of total genomic DNA were combined with 0.6 µM of each primer and one pellet of PuReTaq Ready-To-Go PCR beads (GE Healthcare, Wauwatosa, WI, USA) containing stabilizers, bovine serum albumin (BSA), deoxynucleotide triphosphates, 2.5 units of PuReTaq DNA polymerase, and reaction buffer. When a bead was reconstituted to a 25 µL final volume, the concentration of each dNTP and $MgCl_2$ was set at 200 µM and 1.5 mM, respectively. PCRs were performed in a GeneAmp

PCR System 9700 Thermal Cycler (Applied Biosystems, Waltham, MA, USA), programmed as follows: 1 cycle of 4 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at 48 °C, and 30 s at 72 °C. At the end, a post-treatment of 10 min at 72 °C and a final cooling at 4 °C were carried out. Both positive (high-quality DNA samples from the congeneric *P. clarkii*) and negative controls were used to test the effectiveness of the PCR protocols, and the absence of possible contaminations. Electrophoresis was carried out on 2% agarose gels, prepared using 1× TAE buffer (Tris-Acetate-EDTA, pH 8.3) and stained with Gel Red Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). PCR products were purified by ExoSAP-IT (USB Corporation, Cleveland, OH, USA) and sequenced for forward and reverse strands (by means of the same primers used for PCR), using an external sequencing core service (Macrogen, Europe, Amsterdam, The Netherlands).

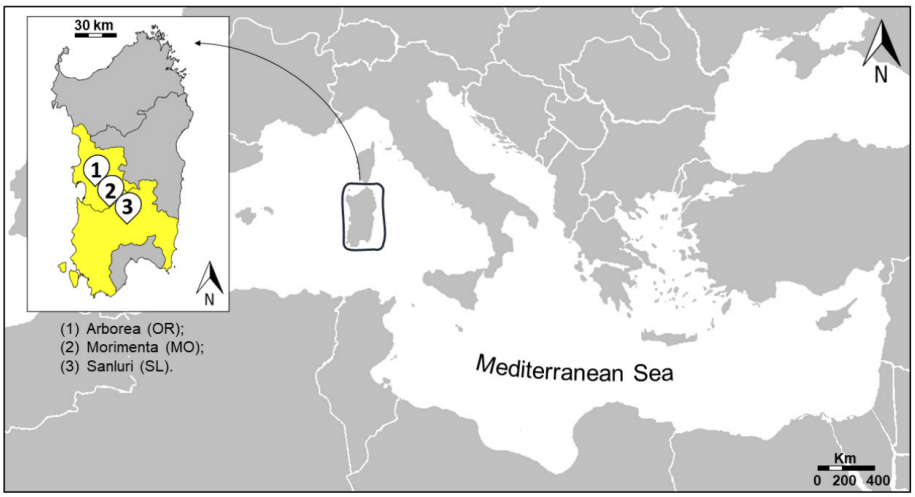


Figure 1. The three geographic areas of the sites where samples were collected are indicated in the map of the Mediterranean island of Sardinia.

2.3. Phylogenetic and Phylogeographic Analyses

All the sequences obtained from crayfish adults and eggs were identified as belonging to the species *Procambarus virginalis* through BLAST analysis implemented in the GenBank nucleotide database (www.ncbi.nlm.nih.gov (accessed on 21 July 2020)) that showed a 100% identity.

Sequences were aligned using the package Clustal Omega [53] (available at https://www.ebi.ac.uk/Tools/msa/clustalo/ (accessed on 4 May 2021)) after a manual checking and editing by means of Unipro UGENE v.35 (by the Unipro Center for Information Technologies, Novosibirsk, Russia) [54].

COI sequences of the marbled crayfish collected in Sardinia were aligned with the sequences belonging to *P. virginalis* and *P. fallax* from other localities (Belgium, Czech Republic, Germany, Italy, Sweden, Japan, and Florida—USA) so far available on GenBank (last update 5 April 2021) (see Table 1 for details). As outgroups, three sequences belonging to the species *P. clarkii* from two different localities were chosen: two from Sardinia (obtained in the present study) (Genbank accession numbers: MZ099652, MZ099653), and one from Alabama (USA) (Genbank accession number: KX417114).

Levels of genetic variation among sequences were assessed estimating the number of polymorphic sites (S), number of haplotypes (H), nucleotide diversity (π), and haplotype diversity (h), using the software package DnaSP 6.12.03 (by Universitat de Barcelona, Barcelona, Spain) [55].

To assess the taxonomic status of crayfish collected in Sardinia, the nucleotide divergence threshold (NDT) method of species delimitation [56] was also performed. The NDT method is based on genetic distances and does not consider the phylogenetic relationships within the dataset. It works on sequences to rank taxa into taxonomic entities applying the fixed threshold of 2% given by Hebert et al. [56] for DNA barcodes, using the pairwise Kimura (1980) two-parameter model (K2P) to compute the matrix of genetic distances [57]. The analysis was performed by means of a script [58] written in the R statistical environment (available at https://cran.r-project.org/ (accessed on 5 May 2021)).

The probabilistic model of sequence evolution that better fit the sequence data was detected using the software JmodelTest 2.1.7 [59]. Based on the best-fitting model, a Bayesian phylogenetic species tree was obtained using the software MrBayes 3.2.7 [60] setting as model parameters: NST = 6, rates = invgamma, ngammacat = 4. Two independent runs, each consisting of four Metropolis-Coupled MCMC chains (one cold and three heated chains), were run simultaneously for 5,000,000 generations, sampling trees every 1000 generations. The first 25% of the 10,000 sampled trees was discarded as burn-in. To assess the convergence of chains, we checked that the Average Standard Deviation of Split Frequencies (ASDSF) approached 0 [60], and the Potential Scale Reduction Factor (PSRF) was around 1 [61], following Scarpa et al. [62]. The phylogenetic tree was visualized and edited using FigTree 1.4.0 (available at http://tree.bio.ed.ac.uk/software/figtree/).

To distinguish genetic clusters, a Principal Coordinate Analysis (PCoA) was also performed on a K2P pairwise genetic distance matrix using *GenAlEX* 6.5 [63]. Thevariation rate among sites was modelled with a gamma distribution and all ambiguous positions were removed for each sequence pair.

A median-joining network [64] was constructed using the software package Network 10.0.0.0 (www.fluxus-engineering.com) to infer the genetic relationships among haplotypes and to detect the occurrence (if any) of discrete genetic clusters. The transitions and transversions were equally weighted. Due to the lack of knowledge on the possible occurrence of retromutation events, the same weight (10) was assigned to each observed polymorphism.

3. Results

In the present study, 47 COI sequences of 617 bp length were obtained for *Procambarus virginalis* and deposited in GenBank (see Table 1 for accession numbers). Among the sequences belonging to Sardinian *P. virginalis*, a very low level of genetic divergence was found, with three polymorphic sites resulting in four haplotypes (see Table 2 for details on genetic divergence estimates). A high percentage (93.62%) of Sardinian *P. virginalis* share the same haplotype, while 6.38% of the Sardinian individuals showed a private lineage.

Table 2. Genetic divergence estimates for *Procambarus virginalis* and *Procambarus fallax* based on COI gene sequences.

	N	**bp**	***S***	***H***	***hd***	π
Sardinia	47	617	3	4	0.125 ± 0.065	0.00021
Whole dataset	76	617	14	9	0.312 ± 0.069	0.00178

The genetic analysis performed on the dataset including all the Sardinian COI sequences obtained in the present study, and those belonging to *P. virginalis* and *P. fallax* deposited in GenBank to date, evidenced 14 polymorphic sites and 9 haplotypes resulting in a low level of genetic variability (see Table 2 for details on genetic divergence estimates) for *P. virginalis* and *P. fallax*. In particular, three haplotypes were found in Sardinian individuals. The NDT species delimitation method ranked all the sequences belonging to *P. virginalis* (also including the Sardinian sequences) and *P. fallax*, into a unique taxonomic entity.

Accordingly, in the phylogenetic tree (Figure 2), all the sequences belonging to *P. virginalis* and *P. fallax* from Sardinia, other European countries, and Japan, were included in a

well-supported monophyletic clade, with an internal sub-cluster that included a sequence of *P. fallax* from Florida (USA) and two sequences of *P. fallax* isolated in Germany. The other sequences belonging to *P. fallax* from Florida were set outside the main clade.

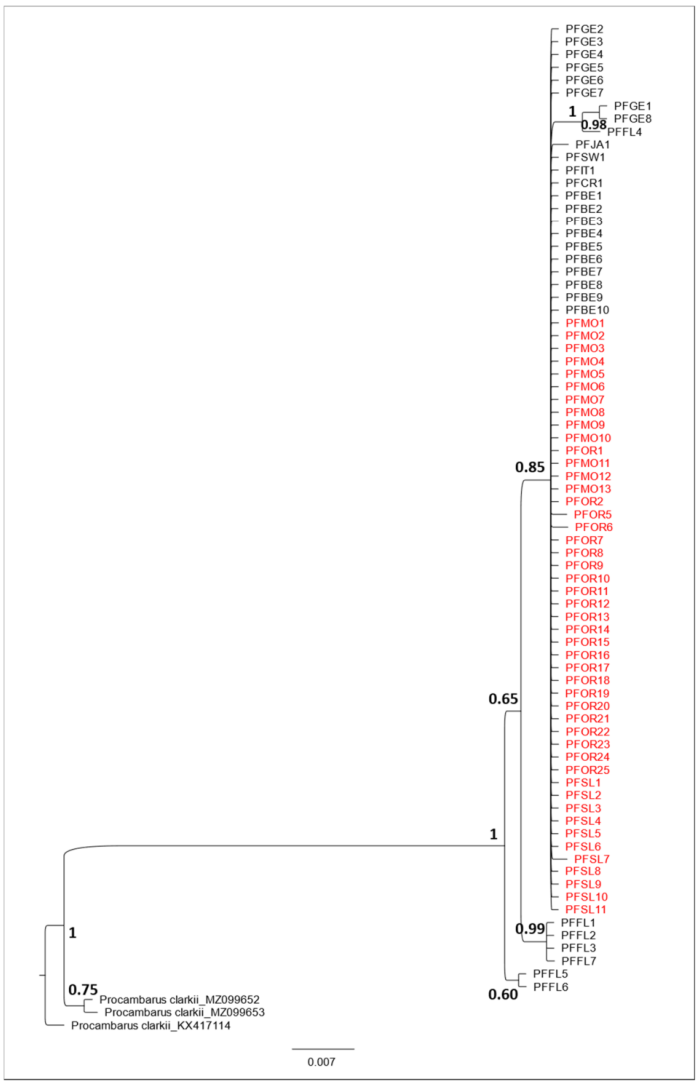


Figure 2. Bayesian phylogenetic tree showing the relationships among *Procambarus virginalis* from Sardinia (indicated with a red font) and marble crayfish from all over the world. Values of support at each node are expressed in Posterior Probabilities. The samples codes are as reported in Table 1.

The PCoA plot (Figure 3) showed the occurrence of four groups of sequences (groups 1–4). A genetic similarity was found along the *x*-axis between the group 2—including all the Sardinian sequences and almost all the sequences of *P. virginalis* and *P. fallax* from European countries and Japan—and the groups 1 and 3 which included sequences of *P. fallax* isolated

in Florida (see Appendix A Table A1 for details). A further divergent group (group 4) of three sequences, already evidenced in the phylogenetic tree (Figure 2), was separated along the *x*-axis, and included two sequences of *P. fallax* isolated in Germany and one sequence of *P. fallax* isolated in Florida.

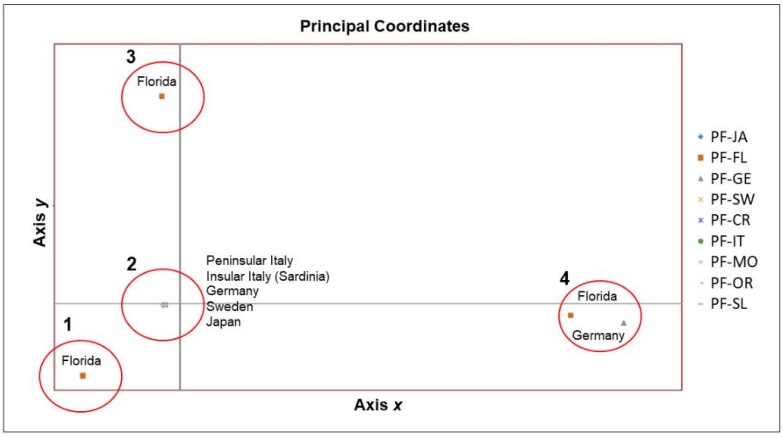


Figure 3. The plot of the Principal Coordinate Analysis (PCoA) evidences the genetic relationships among sequences (and sample sites) based on a matrix of genetic distances. The codes of sequences indicated on the right are as reported in Table 1.

The network analysis (Figure 4) showed a typical star-like shape, with the occurrence of a common haplotype shared by 93.75% and 25.00% of individuals belonging to *P. virginalis* and *P. fallax* (all collected in Germany and Sweden), respectively. Four sequences of *P. virginalis* diverged from the common haplotype by one- or two-point mutations; three of them were isolated in Sardinia (two from Arborea and one from Sanluri) and one in Japan. Further sequences belonging to *P. fallax* collected in Florida and Germany diverged by 3- to 4-point mutations.

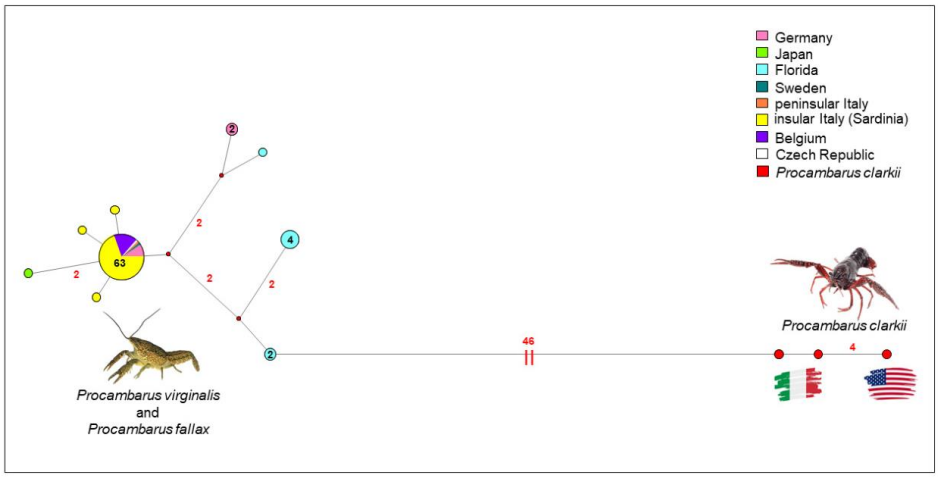


Figure 4. The median-joining network shows the relationships among COI gene haplotypes. The small red spots on the nodes show median vectors representing the hypothetical connecting sequences that were calculated using the maximum parsimony method. The numbers of mutations between sequences that are greater than 1 are reported on network branches.

4. Discussion

This paper reports the first record of *Procambarus virginalis* in Sardinia, supported by molecular identification. This study also represents the first insight into the genetic variability of the first Italian population of *P. virginalis*, as only one individual from Tuscany [10] and three from Veneto [65] are known from Italy to date. New inferences are also provided on the global genetic variation of this species.

Overall, the Sardinian population of the marble crayfish is characterized by a very low level of genetic variability, with a high percentage of specimens sharing the same haplotype, apart from three individuals which show new, slightly divergent haplotypes. Two possible scenarios may be invoked to explain this remarkable finding: (1) the parthenogenetic marble crayfish was introduced in Sardinia with a highly variable group of individuals, and further studies, involving a greater number of specimens from areas outside Sardinia, are needed to check for the occurrence of these never reported haplotypes in the mainland; or (2) based on the current knowledge, we cannot rule out that the three private haplotypes found in Sardinia might stem as the result of strong selective pressures acting on the population of this recently introduced invasive species.

Indeed, all the sequences belonging to *P. virginalis* from European countries are identical, with only few exceptions. The most common haplotype of *P. virginalis* also occurs in *P. fallax* from Germany and Sweden, which may represent individuals directly imported from the USA. Interestingly, no one of the COI sequences isolated in *P. fallax* from the USA correspond to the most common lineage found in *P. virginalis* from Europe and Japan. The lack of the most common, worldwide spread haplotype of *P. virginalis* among the sequences of *P. fallax* from the USA might reflect their low number so far available on GenBank. However, this pattern may also depend on the genetic drift resulting from the small number of founders selected in the USA to be imported in Europe for the pet trade. In this latter case, only a small and poorly representative portion of the source population genetic variability might have been transferred to the European populations, thus resulting in the lack of shared haplotypes between the two continents.

Furthermore, when adaptive evolution takes place, the molecular lineages of introduced populations can be rare, or not expressed in the source population but become common and distinctive among individuals in the new colonized habitats as a consequence of founder events and natural selection [66]. Although representing an uncommon trend, selective pressure on mtDNA promoted by new environmental conditions was already reported for the highly invasive Lessepsian migrant *Fistularia commersonii* [67], whose recently established Mediterranean population does not include any mitochondrial lineages present among individuals of the source population. For this reason, in the case of *P. virginalis*, mitochondrial variants, arrived in Europe from the USA, may have undergone a selective sweep which might have favoured the spread in European populations of rare alleles that likely allow a better adaptation of the species to new environmental conditions.

5. Conclusions

Although there is no evidence of the presence of native crayfish in Sardinia, the autochthonous nature of *Austrapotamobius pallipes*, included in Annexes II and V of the European Union Habitats Directive 2000 (92/43/EEC), is still controversial. Nonetheless, there are endemic species of amphibians that may suffer due to the presence of this new invasive alien species in their fresh waters. For instance, Souty-Grosset et al. [68] and Oficialdegui et al. [69] showed that crayfish eat amphibian eggs and tadpoles. Additionally, competition for food could also indirectly reduce amphibian populations, as the grazing activity that crayfish exert on periphyton may lead periphyton-associated invertebrates to disappear [70]. Lastly, there is evidence that crayfish could carry the chytrid *Batrachochytrium dendrobatidis* (see Battisti and Scalici [26] and references therein), which may cause diseases in populations of Sardinian endemic amphibians.

Therefore, future studies that investigate the stomach content of these crayfish would be useful to provide a clear idea of their impact on Sardinian native fauna. At the same

time, further studies are needed to investigate the presence in the island of the oomycete *Aphanomyces astaci*, which is generally associated to crayfish. Indeed, the presence of *A. astaci* in other crustaceans, such as *Eriocheir sinensis* [71,72], suggests the occurrence of spill over events.

Author Contributions: Conceptualization, D.S., F.S. and M.C.; methodology, D.S. and I.A.; software, F.S.; validation, D.S., F.S. and M.C.; formal analysis, D.S., I.A. and F.S.; investigation, D.S., I.A. and F.S.; resources, D.S. and M.C.; data curation, D.S., I.A. and F.S.; writing—original draft preparation, D.S., I.A. and M.C.; writing—review and editing, D.S., I.A., F.S., P.C. and M.C.; visualization, D.S., I.A., F.S., P.C., I.A., A.P., F.G. and M.C.; supervision, D.S. and M.C.; project administration, D.S. and M.C.; and funding acquisition, D.S. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FONDO DI ATENEO PER LA RICERCA 2019 (to Daria Sanna and Marco Casu) of the University of Sassari (Italy), grant numbers FAR2019SANNAD and FAR2019CASUM.

Institutional Review Board Statement: Scientific research and collecting of specimens of *P. fallax* f. *virginalis* in Sardinia were authorised by Regione Autonoma della Sardegna during the activities for management of invasive alien species (Regulation (EU) no. 1143/2014 and Legislative Decree 230 of 15 December 2017): *Procambarus fallax* f. *virginalis*—presentation of regional guidelines for rapid eradication. Project protocol n. 18164—17 September 2020.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequences obtained in the present study for the mitochondrial Cytochrome c Oxidase subunit I gene isolated in Sardinian crayfish were deposited in the GenBank database under the accession numbers MZ097549-95, MZ099652-53.

Acknowledgments: We are truly indebted with the veterinary technical officer Davide Brugnone, of the Assessorato della Difesa dell'Ambiente of the Regione Autonoma della Sardegna for his precious and invaluable support during the preliminary phases of this study. Furthermore, we would like to thank Commissioner Maria Tiziana Pinna and Senior Inspector Giuseppe Santucciu of the Servizio Ispettorato Ripartimentale di Oristano—Stazione Forestale and of the Vigilanza Ambientale di Marrubiu; Paolo Briguglio of the Duemari veterinary clinic of Oristano; and Angelo Ruiu of the Istituto Zooprofilattivo Sperimentale G. Pegreffi di Oristano, for their help in the sampling collection.

Conflicts of Interest: The authors and the founders declare no conflicts of interest.

Appendix A

Table A1. Principal Coordinate Analysis (PCoA) results. The table shows the specific composition of the four groups of sequences.

Group 1		
NORTH AMERICA		
Sample ID	**Country**	**Species**
PFFL1	Florida—USA	*Procambarus fallax*
PFFL2	Florida—USA	*Procambarus fallax*
PFFL3	Florida—USA	*Procambarus fallax*
PFFL7	Florida—USA	*Procambarus fallax*
Group 2		
EUROPE		
Sample ID	**Country**	**Species**
PFGE2	Germany	*Procambarus fallax*
PFGE3	Germany	*Procambarus fallax*
PFGE4	Germany	*Procambarus virginalis*
PFGE5	Germany	*Procambarus virginalis*

Table A1. *Cont.*

Group 2		
EUROPE		
Sample ID	**Country**	**Species**
PFGE4	Germany	*Procambarus virginalis*
PFGE5	Germany	*Procambarus virginalis*
PFGE6	Germany	*Procambarus virginalis*
PFGE7	Germany	*Procambarus virginalis*
PFCR1	Czech Republic	*Procambarus virginalis*
PFSW1	Sweden	*Procambarus virginalis*
PFBE1	Belgium	*Procambarus virginalis*
PFBE2	Belgium	*Procambarus virginalis*
PFBE3	Belgium	*Procambarus virginalis*
PFBE4	Belgium	*Procambarus virginalis*
PFBE5	Belgium	*Procambarus virginalis*
PFBE6	Belgium	*Procambarus virginalis*
PFBE7	Belgium	*Procambarus virginalis*
PFBE8	Belgium	*Procambarus virginalis*
PFBE9	Belgium	*Procambarus virginalis*
PFBE10	Belgium	*Procambarus virginalis*
PFIT1	Peninsular Italy—Veneto	*Procambarus virginalis*
PFMO1	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO2	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO3	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO4	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO5	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO6	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO7	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO8	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO9	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO10	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO11	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO12	Insular Italy—Sardinia	*Procambarus virginalis*
PFMO13	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR1	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR2	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR5	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR6	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR7	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR8	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR9	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR10	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR11	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR12	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR13	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR14	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR15	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR16	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR17	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR18	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR19	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR20	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR21	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR22	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR23	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR24	Insular Italy—Sardinia	*Procambarus virginalis*
PFOR25	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL1	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL2	Insular Italy—Sardinia	*Procambarus virginalis*

Table A1. *Cont.*

Group 2		
EUROPE		
Sample ID	**Country**	**Species**
PFSL3	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL4	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL5	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL6	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL7	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL8	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL9	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL10	Insular Italy—Sardinia	*Procambarus virginalis*
PFSL11	Insular Italy—Sardinia	*Procambarus virginalis*
ASIA		
Sample ID	**Country**	**Species**
PFJA1	Japan	*Procambarus virginalis*
Group 3		
NORTH AMERICA		
Sample ID	**Country**	**Species**
PFFL5	Florida—USA	*Procambarus fallax*
PFFL6	Florida—USA	*Procambarus fallax*
Group 4		
NORTH AMERICA		
Sample ID	**Country**	**Species**
PFFL4	Florida—USA	*Procambarus fallax*
EUROPE		
Sample ID	**Country**	**Species**
PFGE1	Germany	*Procambarus fallax*
PFGE8	Germany	*Procambarus fallax*

References

1. De Grave, S.; Pentcheff, N.D.; Ahyong, S.T.; Chan, T.Y.; Crandall, K.A.; Dworschak, P.C.; Felder, D.L.; Feldmann, R.M.; Fransen, C.H.J.M.; Goulding, L.Y.D.; et al. A classification of living and fossil genera of decapod crustaceans. *Raffles Bull. Zool.* **2009**, *21*, 1–109.
2. Crandall, K.A.; Harris, D.J.; Fetzner, J.W. The monophyletic origin of freshwater crayfish estimated from nuclear and mitochondrial DNA sequences. *Proc. R. Soc. Lond. B* **2000**, *267*, 1679–1686. [CrossRef]
3. Crandall, K.A.; De Grave, S. An updated classification of the freshwater crayfishes (Decapoda: Astacidea) of the world, with a complete species list. *J. Crust. Biol.* **2017**, *37*, 615–653. [CrossRef]
4. Gherardi, F. Invasive crayfish and freshwater fishes of the world. *Rev. Sci. Tech.* **2010**, *29*, 241–254. [CrossRef]
5. Hobbs, H.H.; Lodge, D.M. Decapoda. In *Ecology and Classification of North American Freshwater Invertebrates*, 3rd ed.; Thorp, J.H., Covich, A.P., Eds.; Academic Press: Cambridge, MA, USA, 2010; pp. 901–967. [CrossRef]
6. Antón, A.; Serrano, T.; Angulo, E.; Ferrero, G.; Rallo, A. The use of two species of crayfish as environmental quality sentinels: The relationship between heavy metal content, cell and tissue biomarkers and physico-chemical characteristics of the environment. *Sci. Total Environ.* **2000**, *247*, 239–251. [CrossRef]
7. Reynolds, J.; Souty-Grosset, C.; Richardson, A. Ecological roles of crayfish in freshwater and terrestrial habitats. *Freshw. Crayfish* **2013**, *19*, 197–218. [CrossRef]
8. Statzner, B.; Peltret, O.; Tomanova, S. Crayfish as geomorphic agents and ecosystem engineers: Effect of a biomass gradient on baseflow and flood-induced transport of gravel and sand in experimental streams. *Freshw. Biol.* **2003**, *48*, 147–163. [CrossRef]
9. Creed, R.P.; Reed, J.M. Ecosystem engineering by crayfish in a headwater stream community. *J. N. Am. Benthol. Soc.* **2004**, *23*, 224–236. [CrossRef]
10. Souty-Grosset, C.; Holdich, D.M.; Noël, P.Y.; Reynolds, J.D.; Haffner, P. *Atlas of Crayfish in Europe*; Muséum National d'Histoire Naturelle: Paris, France, 2006; p. 87.

11. Marzano, F.N.; Scalici, M.; Chiesa, S.; Gherardi, F.; Piccinini, A.; Gibertini, G. The first record of the marbled crayfish adds further threats to freshwaters in Italy. *Aquat. Invasions* **2009**, *4*, 401–404. [CrossRef]
12. Lodge, D.M.; Taylor, C.A.; Holdich, D.M.; Skurdal, J. Nonindigenous crayfishes threaten North American freshwater biodiversity: Lessons from Europe. *Fisheries* **2000**, *25*, 7–20. [CrossRef]
13. Holdich, D.M. Distribution of crayfish in Europe and some adjoining countries. *Bull. Fr. Pêche Piscic.* **2002**, *367*, 611–650. [CrossRef]
14. Weiperth, A.; Bláha, M.; Szajbert, B.; Seprős, R.; Bányai, Z.; Patoka, J.; Kouba, A. Hungary: A European hotspot of non-native crayfish biodiversity. *Knowl. Manag. Aquat. Ecosyst.* **2020**, *421*, 43. [CrossRef]
15. Weiperth, A.; Gál, B.; Kuříková, P.; Bláha, M.; Kouba, A.; Patoka, J. Cambarellus patzcuarensis in Hungary: The first dwarf crayfish established outside of North America. *Biologia* **2017**, *72*, 1529–1532. [CrossRef]
16. Kotovska, G.; Khrystenko, D.; Patoka, J.; Kouba, A. East European crayfish stocks at risk: Arrival of non-indigenous crayfish species. *Knowl. Manag. Aquat. Ecosyst.* **2016**, *417*, 37. [CrossRef]
17. Martin, P.; Shen, H.; Füllner, G.; Scholtz, G. The first record of the parthenogenetic Marmorkrebs (Decapoda, Astacida, Cambaridae) in the wild in Saxony (Germany) raises the question of its actual threat to European freshwater ecosystems. *Aquat. Invasions* **2010**, *5*, 397–403. [CrossRef]
18. Oidtmann, B.; Cerenius, L.; Schmid, I.; Hoffmann, R.; Söderhäll, K. Crayfish plague epizootics in Germany–classification of two German isolates of the crayfish plague fungus *Aphanomyces astaci* by random amplification of polymorphic DNA. *Dis. Aquat. Org.* **1999**, *35*, 235–238. [CrossRef]
19. Keller, N.S.; Pfeiffer, M.; Roessink, I.; Schulz, R.; Schrimpf, A. First evidence of crayfish plague agent in populations of the marbled crayfish (*Procambarus fallax* forma *virginalis*). *Knowl. Manag. Aquat. Ecosyst.* **2014**, *414*, 15. [CrossRef]
20. Putra, M.D.; Bláha, M.; Wardiatno, Y.; Krisanti, M.; Jerikho, R.; Kamal, M.M.; Mojžišová, M.; Bystřický, P.K.; Kouba, A.; Kalous, L.; et al. *Procambarus clarkii* (Girard, 1852) and crayfish plague as new threats for biodiversity in Indonesia. *Aquatic Conserv. Mar. Freshw. Ecosyst.* **2018**, *28*, 1434–1440. [CrossRef]
21. International Office of Epizootics. Crayfish plague (*Aphanomyces astaci*). In *Manual of Diagnostic Tests for Aquatic Animals*, 4th ed.; Aquatic Animal Health Standards Commission: Paris, France, 2006.
22. Holdich, D.M. Crayfish in Europe–an overview of taxonomy, legislation, distribution, and crayfish plague outbreaks. In *Management & Conservation of Crayfish, Proceedings of the Management & Conservation of Crayfish, Nottingham Forest Football Club, Nottingham, UK, 7 November 2002*; Environment Agency: Bristol, UK, 2003; pp. 15–34.
23. Blaha, M.; Patoka, J.; Japoshvili, B.; Let, M.; Buřič, M.; Kouba, A.; Mumladze, L. Genetic diversity, phylogenetic position and morphometric analysis of *Astacus colchicus* (Decapoda, Astacidae): A new insight into Eastern European crayfish fauna. *Integ. Zool.* **2021**, *16*, 368–378. [CrossRef]
24. Pârvulescu, L. Introducing a new *Austropotamobius* crayfish species (Crustacea, Decapoda, Astacidae): A miocene endemism of the Apuseni Mountains, Romania. *Zool. Anz.* **2019**, *279*, 94–102. [CrossRef]
25. Gherardi, F.; Holdich, D.M. *Crayfish in Europe as Alien Species. How to Make the Best of a Bad Situation?* 1st ed.; A.A. Balkema: Rotterdam, The Netherlands, 1999; p. 14.
26. Battisti, C.; Scalici, M. First records of the red swamp crayfish *Procambarus clarkii* (Girard, 1852) (Decapoda Cambaridae) from a small circum-Sardinian island (central Mediterranean Sea). *BioInvasions Rec.* **2020**, *9*, 333–339. [CrossRef]
27. Gherardi, F. Crayfish invading Europe: The case study of *Procambarus clarkii*. *Mar. Freshw. Behav. Physiol.* **2006**, *39*, 175–191. [CrossRef]
28. Lyko, F. The marbled crayfish (Decapoda: Cambaridae) represents an independent new species. *Zootaxa* **2017**, *4363*, 544–552. [CrossRef]
29. Jones, J.P.; Rasamy, J.R.; Harvey, A.; Toon, A.; Oidtmann, B.; Randrianarison, M.H.; Raminosoa, N.; Ravoahangimalala, O.R. The perfect invader: A parthenogenic crayfish poses a new threat to Madagascar's freshwater biodiversity. *Biol. Invasions* **2009**, *11*, 1475–1482. [CrossRef]
30. Gutekunst, J.; Andriantsoa, R.; Falckenhayn, C.; Hanna, K.; Stein, W.; Rasamy, J.; Lyko, F. Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nat. Ecol. Evol.* **2018**, *2*, 567. [CrossRef] [PubMed]
31. Vogt, G.; Falckenhayn, C.; Schrimpf, A.; Schmid, K.; Hanna, K.; Panteleit, J.; Helm, M.; Schulz, R.; Lyko, F. The marbled crayfish as a paradigm for saltational speciation by autopolyploidy and parthenogenesis in animals. *Biol. Open* **2015**, *4*, 1583–1594. [CrossRef]
32. Chucholl, C.; Morawetz, K.; Groß, H. The clones are coming–strong increase in Marmorkrebs [*Procambarus fallax* (Hagen, 1870) f. *virginalis*] records from Europe. *Aquat. Invasions* **2012**, *7*, 511–519. [CrossRef]
33. Hossain, M.S.; Patoka, J.; Kouba, A.; Buřič, M. Clonal crayfish as biological model: A review on marbled crayfish. *Biologia* **2018**, *73*, 841–855. [CrossRef]
34. Patoka, J.; Kalous, L.; Kopecký, O. Risk assessment of the crayfish pet trade based on data from the Czech Republic. *Biol. Invasions* **2014**, *16*, 2489–2494. [CrossRef]
35. Faulkes, Z.; Kawai, T.; Scholtz, G. Marble crayfish as a new model organism and a new threat to native crayfish conservation. In *Freshwater Crayfish: Global Overview*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2015; pp. 31–53. [CrossRef]
36. Faulkes, Z. The spread of the parthenogenetic marbled crayfish, Marmorkrebs (Procambarus sp.), in the North American pet trade. *Aquat. Invasions* **2010**, *5*, 447. [CrossRef]

37. Patoka, J.; Bláha, M.; Kalous, L.; Kouba, A. Irresponsible vendors: Non-native, invasive and threatened animals offered for garden pond stocking. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2017**, *27*, 692–697. [CrossRef]
38. Cvitanić, M.; Hudina, S.; Maguire, I. Reproductive cycle of the marble crayfish from an established population in Croatia. In Proceedings of the 21st Symposium of the International Association of Astacology-Program and Book of Abstracts, Madrid, Spain, 5–8 September 2016; p. 48.
39. Pârvulescu, L.; Togor, A.; Lele, S.F.; Scheu, S.; Șinca, D.; Panteleit, J. First established population of marbled crayfish *Procambarus fallax* (Hagen, 1870) f. *virginalis* (Decapoda, Cambaridae) in Romania. *BioInvasions Rec.* **2017**, *6*, 357–362. [CrossRef]
40. Lipták, B.; Mojžišová, M.; Gruľa, D.; Christophoryová, J.; Jablonski, D.; Bláha, M.; Petrusek, A.; Kouba, A. Slovak section of the Danube has its well-established breeding ground of marbled crayfish *Procambarus fallax* f. *virginalis*. *Knowl. Manag. Aquat. Ecosyst.* **2017**, *418*, 40. [CrossRef]
41. Ercoli, F.; Kaldre, K.; Paaver, T.; Gross, R. First record of an established marbled crayfish *Procambarus virginalis* (Lyko, 2017) population in Estonia. *BioInvasions Rec.* **2019**, *8*, 675–683. [CrossRef]
42. Patoka, J.; Buřič, M.; Kolář, V.; Bláha, M.; Petrtýl, M.; Franta, P.; Tropek, R.; Kalous, L.; Petrusek, A.; Kouba, A. Predictions of marbled crayfish establishment in conurbations fulfilled: Evidences from the Czech Republic. *Biologia* **2016**, *71*, 1380–1385. [CrossRef]
43. Lőkkös, A.; Müller, T.; Kovács, K.; Várkonyi, L.; Specziár, A.; Martin, P. The alien, parthenogenetic marbled crayfish (Decapoda: Cambaridae) is entering Kis-Balaton (Hungary), one of Europe's most important wetland biotopes. *Knowl. Manag. Aquat. Ecosyst.* **2016**, *417*, 16. [CrossRef]
44. Novitsky, R.A.; Son, M.O. The first records of Marmorkrebs [*Procambarus fallax* (Hagen, 1870) f. *virginalis*] (Crustacea, Decapoda, Cambaridae) in Ukraine. *Ecol. Montenegrina* **2016**, *5*, 44–46. [CrossRef]
45. Bohman, P.; Edsman, L.; Martin, P.; Scholtz, G. The first Marmorkrebs (Decapoda: Astacida: Cambaridae) in Scandinavia. *BioInvasions Rec.* **2013**, *2*, 227–232. [CrossRef]
46. Faulkes, Z.; Feria, T.P.; Muñoz, J. Do Marmorkrebs, *Procambarus fallax* f. *virginalis*, threaten freshwater Japanese ecosystems? *Aquat. Biosyst.* **2012**, *8*, 13. [CrossRef] [PubMed]
47. Kawai, T.; Takahata, M. (Eds.) *The Biology of Freshwater Crayfish*, 1st ed.; Hokkaido University Press: Sapporo, Japan, 2010.
48. Morpurgo, M.; Aquiloni, L.; Bertocchi, S.; Brusconi, S.; Tricarico, E.; Gherardi, F.; Romana, V. Distribuzione dei gamberi d'acqua dolce in Italia. *Studi Trent. Sci. Nat.* **2010**, *87*, 125–132.
49. Amouret, J.; Bertocchi, S.; Brusconi, S.; Fondi, M.; Gherardi, F.; Grandjean, F.; Chessa, L.; Tricarico, E.; Souty-Grosset, C. The first record of translocated white-clawed crayfish from the *Austropotamobius pallipes* complex in Sardinia (Italy). *J. Limnol.* **2015**, *74*, 491–500. [CrossRef]
50. Orrù, F.; Deiana, A.M.; Cau, A. Conservazione della biodiversità e specie alloctone invasive: Il gambero *Procambarus clarkii* nelle acque interne della Sardegna. *Studi Trent. Sci. Nat.* **2009**, *86*, 162.
51. Hobbs, H.H. The crayfishes of Georgia. *Smithson. Contrib. Zool.* **1981**, *318*, 1–549. [CrossRef]
52. Folmer, O.; Black, M.; Hoeh, W.; Lutz, R.; Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **1994**, *3*, 294–299.
53. Sievers, F.; Higgins, D.G. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In *Multiple Sequence Alignment Methods*; Russell, D., Ed.; Humana Press: Totowa, NJ, USA, 2014; Volume 1079, pp. 105–116. [CrossRef]
54. Okonechnikov, K.; Golosova, O.; Fursov, M.; Ugene Team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [CrossRef] [PubMed]
55. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
56. Hebert, P.D.; Cywinska, A.; Ball, S.L.; Dewaard, J.R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* **2003**, *270*, 313–321. [CrossRef]
57. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **1980**, *16*, 111–120. [CrossRef]
58. Scarpa, F.; Cossu, P.; Sanna, D.; Lai, T.; Casu, M.; Curini-Galletti, M. New insights on the genus *Otoplana* Du Plessis, 1889 (Platyhelminthes: Proseriata), with description of two new species from the Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2075–2087. [CrossRef]
59. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelT- est 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]
60. Ronquist, F.; Teslenko, M.; Van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]
61. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
62. Scarpa, F.; Sanna, D.; Cossu, P.; Lai, T.; Casu, M.; Curini-Galletti, M. How to achieve internal fertilization without a vagina: The study case of the genus *Archilina* Ax, 1959 (Platyhelminthes, Proseriata) from Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2057–2073. [CrossRef]
63. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]

64. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef]
65. Vojkovská, R.; Horká, I.; Tricarico, E.; Ďuriš, Z. New record of the parthenogenetic marbled crayfish *Procambarus virginalis* from Italy. *Crustaceana* **2014**, *87*, 1386–1392. [CrossRef]
66. Willis, J.H.; Orr, H.A. Increased heritable variation following population bottlenecks—The role of dominance. *Evolution* **1993**, *47*, 949–957. [CrossRef] [PubMed]
67. Sanna, D.; Scarpa, F.; Lai, T.; Cossu, P.; Falautano, M.; Castriota, L.; Andaloro, F.; Follesa, M.C.; Francalacci, P.; Curini-Galletti, M.; et al. *Fistularia commersonii* (Teleostea: Fistulariidae): Walking through the Lessepsian paradox of mitochondrial DNA. *Ital. J. Zool.* **2015**, *82*, 499–512. [CrossRef]
68. Souty-Grosset, C.; Anastacio, P.M.; Aquiloni, L.; Banha, F.; Choquer, J.; Chucholl, C.; Tricarico, E. The red swamp crayfish *Procambarus clarkii* in Europe: Impacts on aquatic ecosystems and human well-being. *Limnologica* **2016**, *58*, 78–93. [CrossRef]
69. Oficialdegui, F.J.; Sánchez, M.I.; Monsalve-Carcaño, C.; Boyero, L.; Bosch, J. The invasive red swamp crayfish (*Procambarus clarkii*) increases infection of the amphibian chytrid fungus (*Batrachochytrium dendrobatidis*). *Biol. Invasions* **2019**, *21*, 3221–3231. [CrossRef]
70. Klose, K.; Cooper, S.D. Contrasting effects of an invasive crayfish (*Procambarus clarkii*) on two temperate stream communities. *Freshw. Biol.* **2012**, *57*, 526–540. [CrossRef]
71. Schrimpf, A.; Schmidt, T.; Schulz, R. Invasive Chinese mitten crab (*Eriocheir sinensis*) transmits crayfish plague pathogen (*Aphanomyces astaci*). *Aquat. Invasions* **2014**, *9*, 203–209. [CrossRef]
72. Tilmans, M.; Mrugała, A.; Svoboda, J.; Engelsma, M.Y.; Petie, M.; Soes, D.M.; Nutbeam-Tuffsfg, S.; Oidtmanng, B.; Roessinka, I.; Petrusek, A. Survey of the crayfish plague pathogen presence in the Netherlands reveals a new *Aphanomyces astaci* carrier. *J. Invertebr. Pathol.* **2014**, *120*, 74–79. [CrossRef] [PubMed]

Chapter 5

Appendix

Throughout the doctoral years, the analytical approach employed in the study of aquatic organisms was consistently applied to additional projects focusing on terrestrial organisms. Specifically, the studies conducted during this period focus the attention on: *Taenia saginata*, *Fasciola hepatica* and, *Canis lupus familiaris*. The undertaken work has resulted in the publication of two articles in internationally peer-reviewed journals and one internationally peer-reviewed book chapter.

Taenia saginata

This research was conducted in collaboration with the Department of Parasitology at the University of Medicine and Pharmacy in Hue, Vietnam, as part of the Master's degree program in Medical Biotechnologies. The program was organized by the Department of Biomedical Sciences at the University of Sassari in collaboration with Hue University of Medicine and Pharmacy (Hue, Vietnam).

This collaboration resulted in the publication of the subsequent article: Molecular Identification and Appraisal of the Genetic Variation of *Taenia saginata* in Central Regions of Vietnam. doi:10.3390life12010070

The present work focused on 38 specimens of *Taenia* spp. from central Vietnam, using the mitochondrial gene Cytochrome c Oxidase subunit I (COI) to identify the species and examine genetic variation across different geographic scales. The phylogenetic and phylogeographic analyses carried out, including COI sequences from Vietnamese specimens and conspecifics from the whole world, revealed that the specimens were indeed *Taenia saginata*. In Southeast Asia, a potential founder effect was observed, with common haplotypes widespread across countries, except for Lao PDR, which shared its most common haplotype only with Thailand. Surprisingly, a globally unique taxonomic entity was identified.

Research article

life

MDPI

Article

Molecular Identification and Appraisal of the Genetic Variation of *Taenia saginata* in Central Regions of Vietnam

Giang Tran Thi [1,†], Ilenia Azzena [2,3,†], Fabio Scarpa [2], Piero Cossu [2], Cong Danh Le [4], Phuong Anh Ton Nu [1], Thi Minh Chau Ngo [1], Daria Sanna [3,*,‡] and Marco Casu [2,‡]

[1] Department of Parasitology, Hue University of Medicine and Pharmacy, Hue University, 06 Ngo Quyen Street, Hue City 4200, Vietnam; ttgiang.med@hueuni.edu.vn (G.T.T.); tnphuonganh@hueuni.edu.vn (P.A.T.N.); ntmchau@hueuni.edu.vn (T.M.C.N.)
[2] Dipartimento di Medicina Veterinaria, Università di Sassari, Via Vienna 2, 07100 Sassari, Italy; iazzena@uniss.it (I.A.); fscarpa@uniss.it (F.S.); picossu@uniss.it (P.C.); marcasu@uniss.it (M.C.)
[3] Dipartimento di Scienze Biomediche, Università di Sassari, Viale San Pietro 43/B, 07100 Sassari, Italy
[4] Traditional Medicine Hospital of Thua Thien Hue Province, Hue City 4200, Vietnam; congdanhbvyhcthue@gmail.com
* Correspondence: darsanna@uniss.it
† These authors contributed equally to this study.
‡ These researchers are co-last authors of this manuscript.

Abstract: *Taenia saginata* is a globally distributed tapeworm responsible for human taeniasis due to the ingestion of raw or undercooked beef. *T. saginata* is present in several Asian countries, including China, Thailand, Lao PDR, Cambodia, and Vietnam, but little is known about its genetic variation. Studying the tapeworm's phylogeographic patterns is crucial to better understanding their association with the geographic distribution of taeniasis/cysticercosis in human populations. In the present study, 38 specimens of this putative species were collected in central regions of Vietnam and analysed using the mitochondrial gene Cytochrome c Oxidase subunit I (COI) as a molecular marker to assess the correct species identification and investigate the level of genetic variation at different geographic scales. Phylogenetic and phylogeographic analyses were carried out on a dataset that included COI sequences from Vietnamese specimens and from all conspecifics available in GenBank to date. The results showed that the collected Vietnamese specimens belonged to the species *T. saginata*. In Southeast Asia, signs of a possible founder effect were discovered, with the most common haplotypes frequent and present in many countries, except Lao PDR, which shares its most common haplotype only with individuals from Thailand. Remarkably, a unique taxonomic entity was found worldwide, even though the available COI sequences of *T. saginata* belonging to non-Asiatic countries are, at present, limited. Therefore, future studies including more COI sequences from a higher number of countries and the use of a combined molecular approach with multiple genetic markers would be useful to provide deeper insight into the global genetic variation of this species.

Keywords: cestodes; *Taenia* spp.; beef tapeworm; Cytochrome c Oxidase subunit I; Asia

Citation: Tran Thi, G.; Azzena, I.; Scarpa, F.; Cossu, P.; Danh Le, C.; Ton Nu, P.A.; Chau Ngo, T.M.; Sanna, D.; Casu, M. Molecular Identification and Appraisal of the Genetic Variation of *Taenia saginata* in Central Regions of Vietnam. *Life* **2022**, *12*, 70. https://doi.org/10.3390/life12010070

Academic Editor: Haichun Gao

Received: 19 November 2021
Accepted: 1 January 2022
Published: 4 January 2022

1. Introduction

Human taeniasis is a parasitic zoonosis caused by infection of helminth agents belonging to the genus *Taenia* (Linneus 1758) (Cyclophyllidea, *Taeniidae*). *Taenia saginata*, *Taenia solium*, and *Taenia asiatica* are responsible for human taeniasis. Among them, *T. saginata* and *T. solium* are distributed globally, while *T. asiatica* is mostly distributed in Asian countries [1–5]. The symptomatology of infection is broad, but the most frequent and distinctive symptoms are the discharge of proglottids, abdominal pain, weight loss, nausea, and fever [6]. Interestingly, *T. saginata*, *T. solium*, and *T. asiatica* are unique among tapeworms for their life cycle since humans act as their only definitive host.

In general, the *Taenia* life cycle is based on predator-prey interactions. Faeces, containing helminth eggs, are released into the environment by definitive hosts (carnivorous

or omnivorous predators) and ingested by intermediate hosts (herbivorous prey) which become infected. Specifically, when a predator ingests infected meat, the tapeworm completes its life cycle in the intestinal tract of the host and produces egg-filled proglottids, which are released into the environment inside faeces [5,7], where they can be viable for several weeks or even months [2,8]. The eggs released contain oncospheres (a larval form of *Taenia*) which passively infect intermediate hosts when ingested, migrating from the small intestine to skeletal muscle and developing into cysticerci (an intermediate stage that evolves into adult tapeworms inside the human intestinal tract) [2,3,9].

For *T. saginata*, human infection occurs through the ingestion of raw or undercooked beef infected by cysticerci, while *T. solium* and *T. asiatica* human infections occur as a consequence of ingesting infected pork [2,9,10].

Stringent socio-economic conditions and the high level of meat inspections allowed European countries to completely eradicate *T. solium* and partially eradicate *T. saginata*. In contrast, the three species are endemic on the Asian mainland, representing a problem in the economy not only in terms of livestock and food production losses, but also for human health (e.g., cysticercosis and neurocysticercosis) [9]. These species coexist in Korea, Taiwan, China, Lao PDR, Thailand [1,5,11], and Vietnam [12], where human taeniasis cases were reported in more than 50 of the 63 provinces.

As it is difficult to correctly distinguish between *Taenia* species, based on adult tapeworm morphology (isolated in hospitalized patients), molecular techniques are used to ensure a correct taxonomic identification [11]. This method provides a solution to solve the issue of identifying tapeworms, which is essential for diagnosis, treating and controlling taeniasis/cysticercosis. In the present study, we performed a molecular species identification on tapeworm specimens collected in central Vietnam which were tentatively morphologically attributable to *Taenia* sp. after collection.

We identified tapeworm specimens based on the Cytochrome c Oxidase subunit I (COI) gene sequence, which is also used as a common marker to infer phylogenetic relationships among cestode species [4,5]. However, other molecular markers, such as 28s, 18s and ITS, could be used to study phylogenetic relationships between cestodes [13–16].

The present study aimed to: (i) identify the species of cestodes found in Vietnamese patients, (ii) investigate the levels of genetic variation in the specimens isolated in Vietnam, and (iii) investigate, if present, the genetic structure of this species among populations, thus further comparing our data to those from other worldwide and Asian countries using the mitochondrial COI gene as a molecular marker.

2. Materials and Methods

2.1. Sample Collection

Thirty-eight adult specimens (gravid proglottids) of *Taenia* spp. were collected from six different regions of central Vietnam between August 2019 and June 2020 (see Table 1 and Figure 1) from hospitalized patients from the Traditional Medicine Hospital of Thua Thien Hue Province (n = 8) and Hue University of Medicine and Pharmacy Hospital (n = 30). Among them, 16 were males, between 5 and 77 years old, and 22 patients were females, between 5 and 71 years old (see Table 2). After treatment, patients were monitored for a period of 3–4 months before being declared completely cured.

The 38 specimens obtained from patients were firstly identified as *Taenia* sp. on the basis of their morphologic traits: white and flat, approximately 1.5–2 cm long and 0.5 cm wide. Proglottids were not stained to estimate the number of uterine branches as it is difficult to perform a correct species attribution using this method, especially when it is necessary to discriminate between *T. saginata* (14–32 branches) and *T. asiatica* (12–26 branches).

Table 1. Sampling plan. The table reports data on the sampling collection and the GenBank accession numbers of the sequences obtained in the present study.

Sample ID	Sampling Site	Species	Host	Genbank Accession #	Sampling Date
TSVN01	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459841	August 2019
TSVN02	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459842	August 2019
TSVN04	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459844	September 2019
TSVN05	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459845	September 2019
TSVN06	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459846	September 2019
TSVN07	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459847	September 2019
TSVN08	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459848	September 2019
TSVN09	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459849	September 2019
TSVN10	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459850	November 2019
TSVN11	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459851	November 2019
TSVN12	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459852	November 2019
TSVN15	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459855	November 2019
TSVN24	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459864	January 2020
TSVN25	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459865	January 2020
TSVN26	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459866	February 2020
TSVN27	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459867	February 2020
TSVN28	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459868	February 2020
TSVN29	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459869	March 2020
TSVN30	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459870	May 2020
TSVN31	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459871	May 2020
TSVN32	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459872	May 2020
TSVN33	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459873	May 2020
TSVN34	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459874	May 2020
TSVN36	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459875	May 2020
TSVN37	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459876	June 2020
TSVN38	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459877	June 2020
TSVN39	Vietnam: Thua Thien Hue	*Taenia saginata*	*Homo sapiens*	OL459878	June 2020
TSVN03	Vietnam: Da Nang	*Taenia saginata*	*Homo sapiens*	OL459843	September 2019
TSVN13	Vietnam: Da Nang	*Taenia saginata*	*Homo sapiens*	OL459853	November 2019
TSVN14	Vietnam: Quang Nam	*Taenia saginata*	*Homo sapiens*	OL459854	November 2019
TSVN16	Vietnam: Kon Tum	*Taenia saginata*	*Homo sapiens*	OL459856	December 2019
TSVN19	Vietnam: Kon Tum	*Taenia saginata*	*Homo sapiens*	OL459859	December 2019
TSVN23	Vietnam: Kon Tum	*Taenia saginata*	*Homo sapiens*	OL459863	December 2019
TSVN17	Vietnam: Binh Dinh	*Taenia saginata*	*Homo sapiens*	OL459857	December 2019
TSVN20	Vietnam: Binh Dinh	*Taenia saginata*	*Homo sapiens*	OL459860	December 2019
TSVN18	Vietnam: Quang Tri	*Taenia saginata*	*Homo sapiens*	OL459858	December 2019
TSVN21	Vietnam: Quang Tri	*Taenia saginata*	*Homo sapiens*	OL459861	December 2019
TSVN22	Vietnam: Quang Tri	*Taenia saginata*	*Homo sapiens*	OL459862	December 2019

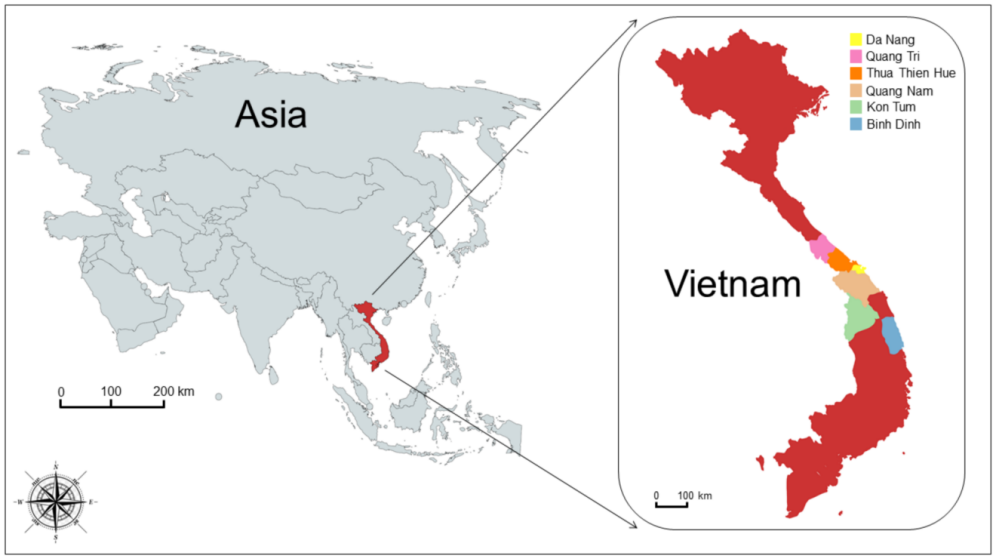


Figure 1. Map of samples' collection sites. The map shows the geographical origin of Vietnamese sequences which were isolated in the present study.

Table 2. Sampling plan. The table reports data of the patients who had taeniasis infections, and were cured between August 2019 and June 2020, from which the *Taenia* sp. specimens used in the present study were collected.

Sample ID	Age	Sex	Sampling Region	Hospital	Collection Year
TSVN01	5	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN02	41	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN04	71	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN05	57	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN06	32	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN07	48	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN08	67	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN09	60	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN10	27	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN11	57	F	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2019
TSVN12	51	F	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2019
TSVN15	55	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2019
TSVN24	77	M	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2020
TSVN25	52	F	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2020
TSVN26	46	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN27	66	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN28	53	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN29	53	F	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2020

Table 2. *Cont.*

Sample ID	Age	Sex	Sampling Region	Hospital	Collection Year
TSVN30	32	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN31	24	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN32	31	M	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2020
TSVN33	47	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN34	47	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN36	56	F	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2020
TSVN37	27	M	Thua Thien Hue	Traditional Medicine Hospital of Thua Thien Hue Province	2020
TSVN38	36	M	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN39	52	F	Thua Thien Hue	Hue University of Medicine and Pharmacy Hospital	2020
TSVN03	19	F	Da Nang	Hue University of Medicine and Pharmacy Hospital	2019
TSVN13	50	F	Da Nang	Hue University of Medicine and Pharmacy Hospital	2019
TSVN14	38	F	Quang Nam	Hue University of Medicine and Pharmacy Hospital	2019
TSVN16	47	F	Kon Tum	Hue University of Medicine and Pharmacy Hospital	2019
TSVN19	63	F	Kon Tum	Hue University of Medicine and Pharmacy Hospital	2019
TSVN23	46	F	Kon Tum	Hue University of Medicine and Pharmacy Hospital	2019
TSVN17	48	F	Binh Dinh	Hue University of Medicine and Pharmacy Hospital	2019
TSVN20	51	F	Binh Dinh	Hue University of Medicine and Pharmacy Hospital	2019
TSVN18	69	F	Quang Tri	Hue University of Medicine and Pharmacy Hospital	2019
TSVN21	52	F	Quang Tri	Hue University of Medicine and Pharmacy Hospital	2019
TSVN22	5	M	Quang Tri	Hue University of Medicine and Pharmacy Hospital	2019

2.2. Diagnostic Molecular Analysis

Total genomic DNA was isolated from a portion of muscle tissue using the Qiagen DNeasy Blood & Tissue Kits (Qiagen, Hilden, Germany) following the supplier's instructions. After extraction, DNA was stored as a solution at 4 °C. Sample quality and DNA concentration were quantified using the NanoDrop™ 2000 Spectrophotometers (by Thermo Scientific; Waltham, MA, USA), which showed an average yield of approximately 248 ng/µL. A fragment of subunit I of the mitochondrial Cytochrome c Oxidase gene (COI) was amplified by standard PCR using primers, cox1 (forward) (5′-CATGGAATAATAATGATTTTC-3′) and cox1 (reverse) (5′-ACAGTACACACAATTTTAAC-3′) [13]. All PCRs were carried out in a total volume of 25 µL. On average, 10 ng of total genomic DNA were combined with 0.6 µM of each primer and one pellet of PuReTaq Ready-To-Go PCR beads (GE Healthcare, Wauwatosa, WI, USA) containing stabilizers, bovine serum albumin (BSA), deoxynucleotide triphosphates (dNTPs), 2.5 units of PuReTaq DNA polymerase, and reaction buffer.

When a bead was reconstituted to a 25 µL final volume, the concentration of each dNTP and $MgCl_2$ was set at 200 µM and 1.5 mM, respectively. PCRs were performed in a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems, Waltham, MA, USA), programmed as follows: 1 cycle of 4 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s 56 °C and, 30 s at 72 °C. At the end, a post-treatment of 10 min at 72 °C and a final cooling at 4 °C were carried out. Both positive and negative controls were used to test the effectiveness of the PCR protocols, and the absence of possible contaminations. Electrophoresis was carried out on 2% agarose gels, prepared using 1x TAE buffer (Tris-Acetate-EDTA, pH 8.3) and stained with Gel Red Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). PCR products were purified by ExoSAP-IT (USB Corporation, Cleveland, OH, USA) and sequenced

for forward and reverse strands (by means of the same primers used for PCR), using an external sequencing core service (Macrogen Europe, Amsterdam, The Netherlands).

2.3. Phylogeographic and Phylogenetic Analysis

Thirty-eight newly-generated sequences of COI fragments were aligned using the package Clustal Omega [17], available at https://www.ebi.ac.uk/Tools/msa/clustalo/ (last access: 10 October 2021) and deposited in GenBank (see Table 1 for GenBank accession numbers). To perform molecular analyses to consider our data in a wider geographic context, we constructed two datasets: one including all the species of *Taenia* globally distributed, available on GenBank to date, and a second one including only the sequences of *T. saginata* from Asian countries, available on GenBank to date (last access: 24 August 2021) (See Table S1 for GenBank accession number). Both of these two datasets include a sequence of *Echinococcus granulosus* (MN787534) as an outgroup.

Levels of genetic variation among sequences were assessed estimating the number of polymorphic sites (*S*), number of haplotypes (*H*), nucleotide diversity (π), and haplotype diversity (*hd*), using the software package DnaSP 6.12.03 [18].

Two median-joining networks [19] were constructed using the software package Network 10.2.0.0 (www.fluxus-engineering.com) (Colchester, UK) to infer the genetic relationships among Vietnamese and Asian haplotypes. The transitions and transversions were equally weighed. Due to the lack of knowledge regarding the possible occurrence of retromutation events, the same weight (10) was assigned to all the observed polymorphisms.

Phylogenetic relationships among specimens were investigated using Bayesian inference (BI) by means of the software MrBayes 3.2.7 [20].

The simplest evolutionary model that best fits the sequence data was detected using the software JmodelTest 2.1.7 [21] and PartitionFinder 2.1.1 [22]. These softwares provided the same result and, in accordance with the best-fitting model, the runs in MrBayes were performed by setting the following model parameters: NST = 6, rates = invgamma, ngammacat = 4. Two independent runs, each consisting of four Metropolis-Coupled MCMC chains (one cold and three heated chains), were run simultaneously for 5 million generations, sampling trees every 1000 generations.

In order to test the convergence of chains, we checked that the average standard deviation of split frequencies (ASDSF) approached 0 [20], and the potential scale reduction factor (PSRF) was around 1 [23], following Scarpa et al. [24].

The phylogenetic tree was visualized and edited using FigTree 1.4.0 (available at http://tree.bio.ed.ac.uk/software/figtree/) (last access: 10 October 2021) (Edinburgh, UK).

In order to verify the taxonomic assessment of each sequence in the dataset, three different methods of species delimitation were used. The use of different methods, based on different criteria and algorithms, is crucial for a conservative approach, which avoids attribution of sequences to erroneous taxonomic entities. In the present study, only results corroborated by all the used species delimitation methods were considered as consistent and well-supported. The first method used is the PTP (Poisson Tree Processes) model and its Bayesian implementation bPTP [25]. PTP/bPTP work on the phylogenetic species concept (PSC) using the number of substitutions in tree branches to assess the speciation rate. The method tests for a significant shift in the substitution rate, which is indicative of the switch from between-species to within-species processes. Species delimitation was performed by means of the bPTP web server (available at http://species.h-its.org/ptp/) (last access: 10 October 2021) (Heidelberg and Karlsruhe, Germany) by using the Bayesian phylogenetic species tree as input file, with default options and 500,000 MCMC generations. Chain convergence was verified by visualizing the likelihood plot. If convergence occurred, the chain should stay at high likelihood locations most of the time during the run.

The second method was the Nucleotide Divergence Threshold (NDT), which was implemented by means of a customized script written in the R statistical environment proposed by Scarpa et al. [26]. The script ranks specimens into taxonomic entities applying the fixed threshold of 2% given by Hebert et al. [27] for DNA barcodes, using a pairwise

Kimura two-parameter model (K2P) genetic distances matrix [28]. In the present study the K2P was chosen as it is recommended to estimate genetic distances that will be used for taxonomic purposes.

The last used method was the Assemble Species by Automatic Partitioning (ASAP) [29] which was performed by using the *p*-distance model (as substitution model to calculate the distances matrix), selecting default options. The ASAP is a fully exploratory method, and it does not require any kind of a priori knowledge. The species hypothesis was accepted by Puillandre et al. [29], within the list of the best partitions valuating their gap-width score, *p*-value and threshold distance.

To identify potential subgroups within the genetic clusters and to determine the dissimilarity represented by the genetic variation among sequences, a Principal Coordinates Analysis (PCoA) was performed using GenAlEX 6.5 [30] on a matrix of pairwise genetic distances corrected according to the Kimura two-parameter model (K2P) [28].

Furthermore, in order to verify the occurrence of genetic association between genetic variability and host age, the PCoA was also performed grouping patients according to their age, thus using three age classes (0–30; 30–50; $\geq$50).

3. Results

Thirty-eight sequences of the initial portion of the COI gene were obtained from Vietnamese specimens (Table 1). This dataset showed 18 polymorphic sites which defined 15 haplotypes (see Table 3 for details on genetic divergence estimates).

Table 3. Indices of genetic variation. The table reports the estimates of genetic variation for the mitochondrial COI gene dataset. N: sample sizes; bp: fragment size; *S*: number of polymorphic sites; *H*: number of haplotypes; *hd*: haplotype diversity; π: nucleotide diversity.

	N	**bp**	***H***	***S***	***hd***	π
Vietnamese COI dataset	38	1013	15	18	0.667	0.00142
Asian COI dataset	182	1013	50	51	0.789	0.00211
Whole world COI dataset	202	1013	57	60	0.794	0.00209

All the sequences obtained from the Vietnamese specimens showed a 100% of genetic identity match with the species *T. saginata* with the Basic Local Alignment Search Tool (BLAST) analysis implemented in the GenBank nucleotide database (www.ncbi.nlm.nih.gov, accessed on 21 July 2020).

To correctly assess the species of samples collected during the present study, before proceeding with phylogeographic inferences, two analyses were performed on a dataset that included the COI sequences of all the species belonging to the genus *Taenia* available in the GenBank database to date. A total of 80 sequences, representative of 21 *Taenia* species (at least two individuals per species when possible) and one outgroup (*Echinococcus granulosus*), were used (see in Appendix A, Table A1 and Figure A1 for details on species and GenBank accession numbers).

First a phylogenetic tree analysis (see in Appendix A, Figure A1) was performed, which showed a unique, well-supported monophyletic cluster, that included the sequences obtained in the present study, along with those of *T. saginata* from GenBank. This genetic clade is characterized by an extended polytomy with some internal well-supported sub-clusters including only a few sequences, and it represents the sister taxon of *T. asiatica*. No relevant structuring based on the geographic distribution of hosts was found among the sequences.

For the second analysis performed on the *Taenia* species dataset, three species delimitation methods (see in Appendix A, Table A1 for details) were used. In general, all the methods were consistent, evidencing a number of taxonomic entities corresponding to the number of species present in the dataset. The only exceptions were represented by *T. hydatigena*, *T. omissa*, *T. polyacantha, T. serialis* and *T. taeniaeformis,* whose sequences split

into different taxonomic entities by some of the methods (see in Appendix A, Table A1 for details) as a possible consequence of the high discrimination capacity of these methods. In accordance with BLAST and phylogenetic tree analyses, all methods of species delimitation suggest that the sequences isolated from Vietnamese samples in the present study belong to the taxonomic entity of *T. saginata*.

After reaching the correct taxonomic identification for Vietnamese cestodes, further analyses were performed to infer the levels of genetic variation among populations in Vietnam and other countries.

The network analysis performed on the Vietnamese samples (Figure 2) showed the presence of two clusters. One cluster is more common and characterized by a well-defined star-like shape with a central highly diffused haplotype including about 60.5% of the sequences analysed.

Thirteen haplotypes diverged from the most common haplotype, by accumulation of one-to-four-point mutations. The second cluster is less widespread, including 15.8 % (six individuals) of the sequences analysed and shows an almost star-like shape, with one common haplotype and three derivatives. Specimens coming from the Thua Thien Hue (4), Quang Nam (1) and Quang Tri (1) provinces belong to this cluster.

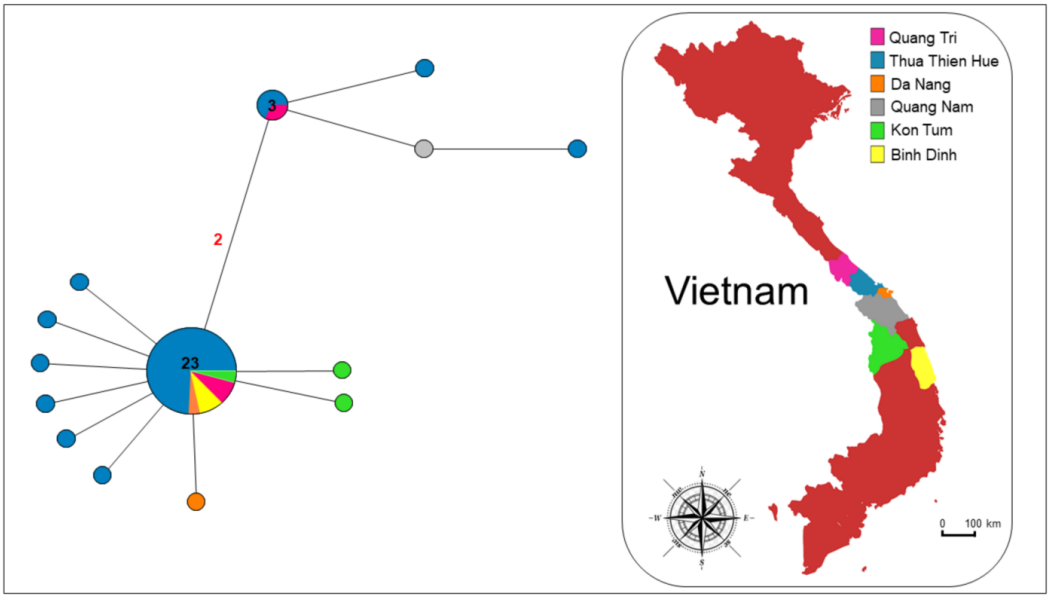


Figure 2. Median-joining network analysis. The network includes COI Vietnamese sequences from the present study. The number of mutations between sequences that are greater than n = 1 are reported on network branches. Additionally, the number of individuals showing the same haplotype that is greater than n = 1 is reported inside the spot.

The same Vietnamese dataset was used to evaluate whether genetic structuring was present among samples based on the age of the hosts (patient age are provided in Table 2) by performing a PCoA (see in the supplementary material, Figure S1 and Table S2). Samples were considered depending on the age (in years) of patients, three age groups were used (0–30, 30–50, ≥50). Results evidenced the absence of genetic structuring among sequences, also evidenced by the low percentage of variation explained by the first two axes (PCoA 1: 41.99%; PCoA 2: 10.97%), thus suggesting a lack of association between genetic variation and the age of infected patients.

Another dataset, including the Vietnamese COI sequences obtained in the present study, along with those belonging to *T. saginata* from other Asian isolates deposited in GenBank, was constructed to perform a second network analysis (see Figure 3 and the supplementary Table S1 for details).

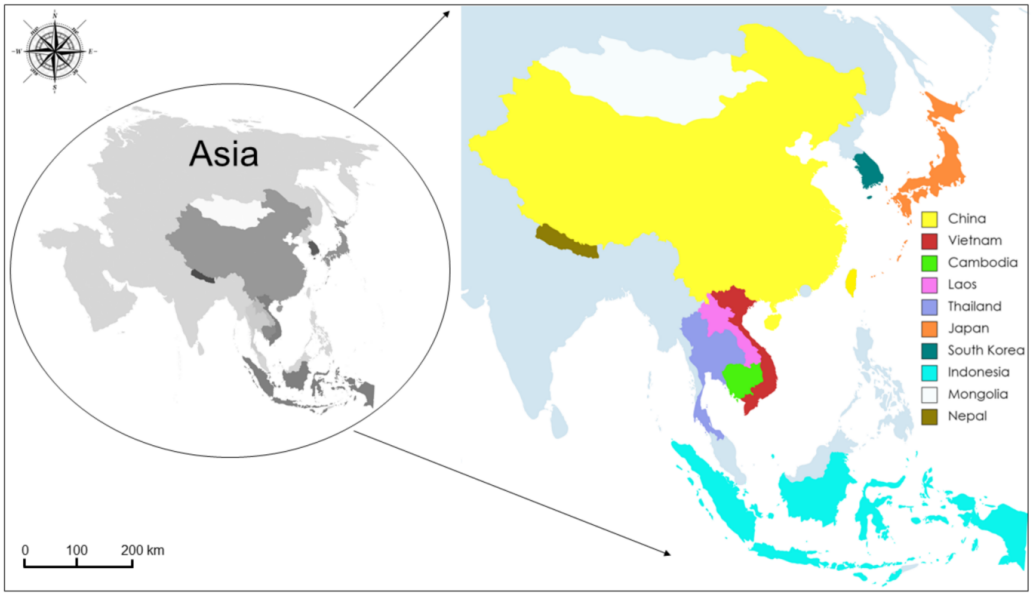


Figure 3. Map of the Asian sequences' distribution. The map shows the geographical sites where the Asian sequences from GenBank (the accession numbers are reported in the supplementary Table S1), which were used in the present study, were collected.

This dataset includes 182 sequences and comprises 51 polymorphic sites, which define 50 haplotypes. In contrast, a very low nucleotide diversity (π) was found (see Table 3 for details on genetic divergence estimates).

The analyses revealed the presence of two clusters (Figure 4): cluster A, which was more common and spread in all the Asian countries analysed, except Lao PDR, and cluster B, which was only present in Thailand and Lao PDR. Cluster A is characterized by a well-defined star-like shape with a common haplotype shared by 45.6% of the total sequences. Thai samples show the principal and derivative haplotypes of both clusters, and Lao PDR samples share only one haplotype with eight sequences from Thailand. New derivative haplotypes in the network diverged from the most common ones by the accumulation of one point mutation. Ten of them had never been previously reported and were isolated in four of the six Vietnamese provinces which were analysed in the present study (six from Thua Thien Hue, two from Kon Tum, one from Da Nang, and one from Quang Nam). Two not-resolved reticulations are present in the network, but negligible, as the relationships between the haplotypes are very clear.

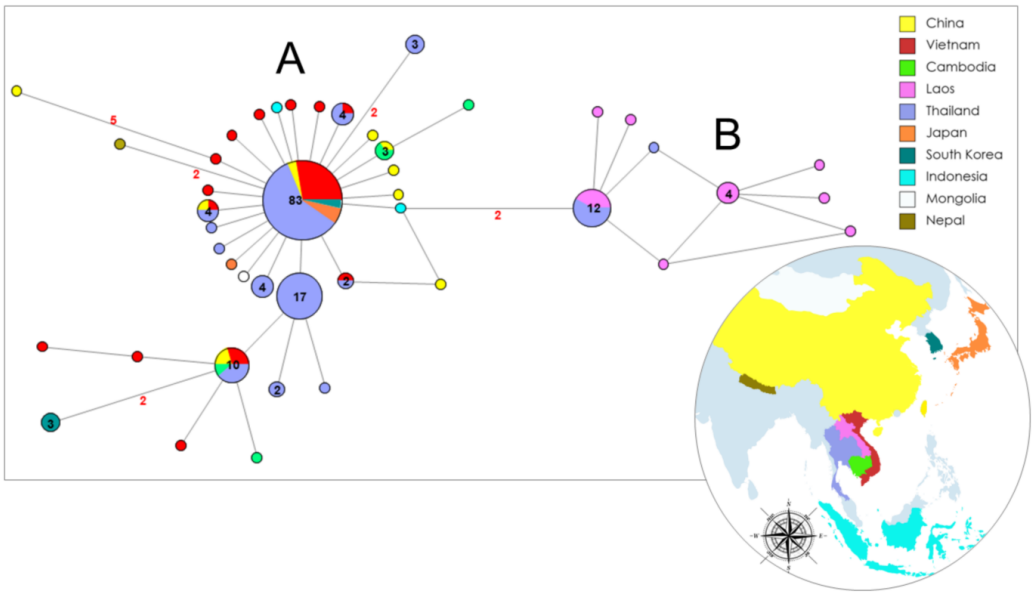


Figure 4. Median-joining network analysis. The network includes COI Vietnamese sequences from the present study along with those belonging to *Taenia saginata* from other Asian isolates which are available in GenBank (the accession numbers are reported in the supplementary Table S1). The number of mutations between sequences that are greater than n = 1 are reported on network branches. Additionally, the number of individuals showing the same haplotype that is greater than n = 1 is reported inside the spot.

A third dataset was also constructed upon which to perform phylogenetic analyses and evaluate the data obtained in the present study in a wider geographic context. It included the Vietnamese sequences obtained in the present study and those corresponding to the same portion of the COI gene of *T. saginata* isolated worldwide and deposited in GenBank (see Table S1 for accession numbers).

This dataset included 202, 1013 bp-long, sequences of *T. saginata* from 18 countries (see supplementary Table S1 for details). Among them, 60 polymorphic sites were retrieved, corresponding to 57 haplotypes. Moderate high levels of polymorphic sites (*S*) and haplotype diversity (*hd*) were found, while a very low values of nucleotide diversity (π) was found (see Table 3 for further details on genetic divergence estimates).

Accordingly, the PTP/bPTP, NDT, and ASAP species delimitation methods were consistent in grouping all *T. saginata* COI sequences in a unique worldwide distributed taxonomic entity. Indeed, only one taxonomic entity for all the sequences was evidenced by all the methods of species delimitation used (see in Appendix A, Table A1).

To further verify whether the few internal well-supported clusters evidenced in the phylogenetic tree (see Figure A1 in Appendix A) were consistent with the possible occurrence of different taxonomic entities within *T. saginata*, a PCoA (Figure 5) was performed, including in the analysed dataset the outgroup that was also used for the phylogenetic tree.

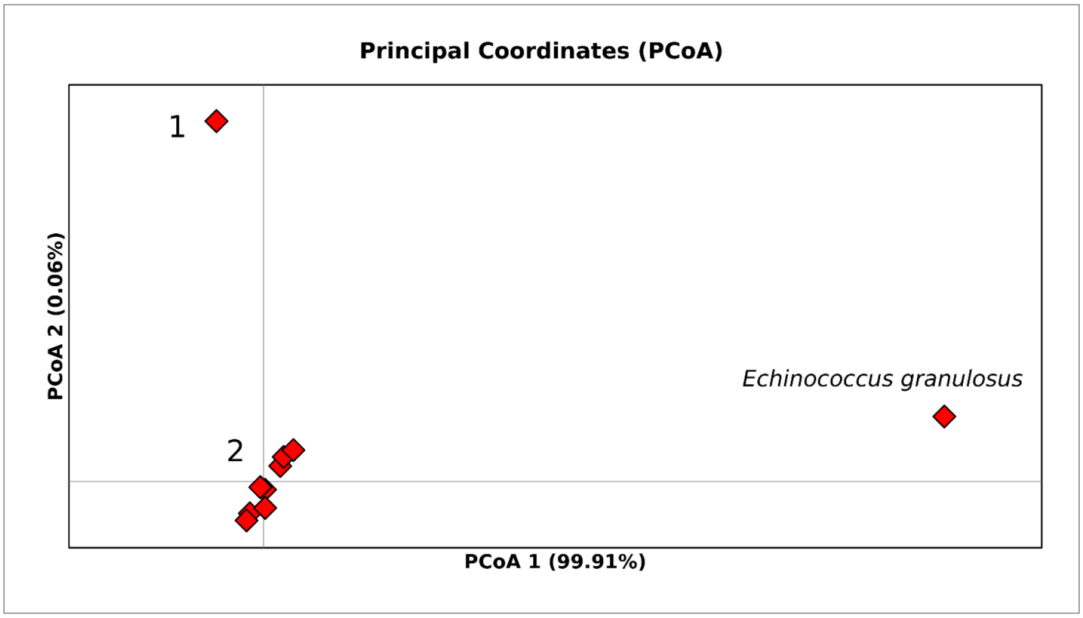


Figure 5. PCoA performed on the whole world *Taenia saginata* COI dataset. Bi-dimensional plot shows the genetic differentiation among specimens due to the nucleotide substitutions per site found in the dataset. The number 1 inside the figure stands for Group 1 and the number 2 stands for Group 2 (see the supplementary Table S3 for details). PCoA1 (axis X) accounts for 99.91% of variation, while PCoA2 (axis Y) accounts for 0.06% of variation.

PCoA results showed the occurrence of two closely related genetic groups (1 and 2 in Figure 5) along the axis X (PCoA1), for *T. saginata* which were equally divergent from the outgroup *E. granulosus*, thus corroborating the general homogeneity evidenced within *T. saginata* by previous analyses. In particular, the first axis (PCoA 1) accounted for 99.91% of the variation, while the second axis (PCoA 2) accounted for 0.06% of the variation. Group 1 was distributed worldwide and included 98.5% of the sequences, while Group 2 included only 1.5% of the total sequences (three strains isolated in Thailand in 2010 and 2016; see the supplementary Table S3 for details).

4. Discussion

This study found that the 38 *Taenia* specimens collected in central Vietnam belong to the species *T. saginata*. Inferences on the genetic variation of this species at different geographic scales—within Vietnam, through Asia, and worldwide—are also provided, using the mitochondrial COI sequences currently available in GenBank.

Correct taxonomic attribution of *Taenia* spp. specimens, based only on morphological characteristics, is very difficult to be performed, especially when it is necessary to differentiate between *T. saginata* and *T. asiatica*. Indeed, the main morphological feature taken into account to distinguish between *Taenia* species is the proglottids, and the aforementioned species are closely related due to the similarity of their proglottids [31–35]. Therefore, the use of the mitochondrial COI gene in this study allowed an accurate taxonomic identification of *Taenia* sp. specimens to be conducted. In addition, the use of several molecular species delimitation methods based on different criteria and algorithms allowed us to perform a conservative and robust approach. All the methods used agreed in evidencing

that the sequences obtained from Vietnamese samples in the present study belonged to the same taxonomic entity: *Taenia saginata*.

Extant knowledge about genetic variation in Vietnamese *T. saginata* is limited, and few studies based on mitochondrial markers [2,13] are available. However, analyses of genetic variation based on sequences of *T. saginata* COI gene in Asia were performed for populations from Thailand and Lao PDR [2,13], where a high level of variation was found. This suggests that *T. saginata* might have spread through different routes in Southeast Asia [2]. In accordance with these data, the present study's results suggest the occurrence of a possible founder effect (as evidenced by star-like shapes in the Network analyses) in Asian countries. Thailand in particular seems to be a hotspot of biodiversity from which *T. saginata* expands into neighboring areas, according to the typical trend of expansion reported for the populations recently originated from a few founders. In fact, either the most common and several derivatives mitochondrial haplotypes were found among the Thai samples, suggesting that genetic variation of tapeworms in Thailand may be representative of one of the first Asian areas colonized by *T. saginata*. Accordingly, the most common COI haplotype found in Thailand and other Asian countries may correspond to one of the oldest mitochondrial variants and it is likely representative of the first lineages introduced to the region.

Even considering a possible bias due to the high number of Thai samples present in the dataset, this occurrence could be explained in the context of the introduction of *T. saginata* to Asia, where the Northeast area of Thailand and Lao PDR may represent the first centres of infection, according to Sanpool et al. [2].

This finding could be explained by the human mediated movements of cattle. Thailand was known as a source of cattle export [36] where transhumance of animals started, passed across the Southern provinces of Laos, and arrived in Vietnam [36,37]. However, it is interesting to note that Lao PDR *Taenia* populations do not seem to have expanded beyond the borders of the country; this may be partly due to the economic conditions of the area, with its predominantly rural-based agriculture and limited cattle export [10]. The few Lao PDR sequences shared with Thailand belong to specimens collected in the Northwest, near the border between Lao PDR and Thailand. In addition, no haplotype is shared between the Vietnamese and Lao PDR *T. saginata* populations. These results could be a consequence of: (i) few collection sites and sequences available from Lao PDR on GenBank; (ii) excessively rapid animal transit during transhumance in Lao PDR; and (iii) genetic drift, leading to the spread of haplotypes in Lao PDR that are uncommon in Thailand.

The presence in Vietnam of several haplotypes shared with other populations from Asia, may be a consequence of the economic importance of commercial routes between Vietnam and other Southeast Asian countries. Indeed, most of the haplotypes shared between Thailand and Vietnam could be closely related to the commercial movements of cattle for grazing or slaughtering [38]. For a long time, informal cross-border trade between Thailand and Vietnam was difficult to be controlled, especially in the border of transit shared between Lao PDR and the Vietnamese province of Quang Tri [38]. Furthermore, a constant flow of human workers from Vietnam to Thailand has lasted for many centuries. There is no official estimate for the number of Vietnamese migrants in Thailand, as most of them are illegal immigrants, but unofficial estimates suggest the presence of approximately 50,000 Vietnamese workers between 2012 and 2014 [39]. In this context, it is very likely that workers became infected during their tenure in Thailand, but because they were illegally present in Thai territory, they could only be cured after returning to their hometowns in Vietnam. On the other hand, although labor migration, legal or not, also occurs between Vietnam and Laos PDR [37], the results obtained in the present study do not show any haplotype shared between these two countries. In the future, molecular studies with more sequences from Lao PDR could be useful to confirm or refute this trend and to better understand the dynamics of the *T. saginata* genetic flow between Lao PDR and Vietnam. In general, human activities could explain the homogeneity of haplotypes found in Southeast Asia in the present study.

A bias in the frequencies of sequences from outside Asia is present in the worldwide dataset analysed here, which is composed of only 11.58% (n = 19) of non-Asian country sequences. This low number may have influenced the results obtained for the whole world population in the present study, which evidenced a unique monophyletic taxonomic entity for *T. saginata* globally. However, even if this finding is not conclusive for worldwide populations, it could be representative of the real taxonomic status of *T. saginata* in Asia, whose common ancestor seems to have differentiated in 2001.

Interestingly, no trace of the occurrence of *T. solium* was found among the Vietnamese specimens analysed here. This pathogen is responsible for neurocysticercosis which is an infection belonging to the group of Neglected Tropical Diseases (NTDs). This result is in accordance with a previous study by Ng-Nguyen et al. [40] which showed very low *T. solium* infection rates in central Vietnam. Although these findings should be further corroborated by a deeper sampling campaign, they are nonetheless suggestive of a low risk for cysticercosis/neurocysticercosis in the human communities of central Vietnam and also provide evidence that the introduction of proper prevention and management strategies in these Vietnamese areas may have yielded optimal results in controlling the spread of tapeworm infection.

In conclusion, the present study sheds further light on the origin and spread of *T. saginata* in Asia as well as in central Vietnam. In the future, further molecular surveys with a wider sampling plan (including sequences from Vietnam and neighboring countries), and the combined use of mitochondrial and nuclear markers will be needed to corroborate the genetic trends evidenced here.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/life12010070/s1, Table S1: Sequences from GenBank. The table reports data on the whole world *Taenia saginata* COI dataset of sequences which were used in the present study and are available on GenBank. Table S2: Sequences used for PCoA. The table reports data on Vietnamese *Taenia saginata* COI sequences (from the present study) which were used for PCoA, they are categorized according to the host age. We set three different age groups: 0–30 years identified by the code 30, 30–50 years identified by the code 50 and, $\geq$50 years identified by the code 50+. Table S3: Sequences for PCoA. The table reports data on the *Taenia saginata* COI sequences which were used for PCoA in the present study and are available on GenBank.

Author Contributions: Conceptualization, I.A. and D.S.; methodology, G.T.T., I.A. and F.S.; software, I.A. and F.S.; validation, D.S., P.A.T.N., T.M.C.N. and M.C.; formal analysis, I.A. and F.S.; investigation, G.T.T. and I.A.; resources, D.S. and M.C.; data curation G.T.T., I.A., P.A.T.N., T.M.C.N. and F.S.; writing—original draft preparation, I.A., D.S. and M.C.; writing—review and editing, I.A., D.S., M.C., G.T.T., P.A.T.N., T.M.C.N., P.C. and C.D.L.; visualization I.A., D.S., M.C., G.T.T., P.A.T.N., T.M.C.N., P.C. and C.D.L.; supervision, D.S. and M.C.; project administration, D.S. and M.C.; and funding acquisition, D.S. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FONDO DI ATENEO PER LA RICERCA 2020 (to Daria Sanna and Marco Casu) of the University of Sassari (Italy), grant numbers FAR2020SANNAD and FAR2020CASUM.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Hue University of Medicine and Pharmacy (protocol code H2020/014 approved in date 15 January 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Sequences obtained in the present study for the mitochondrial Cytochrome c Oxidase subunit I gene isolated in Vietnamese *Taenia saginata* were deposited in the GenBank database under the accession numbers OL459841-OL459878.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

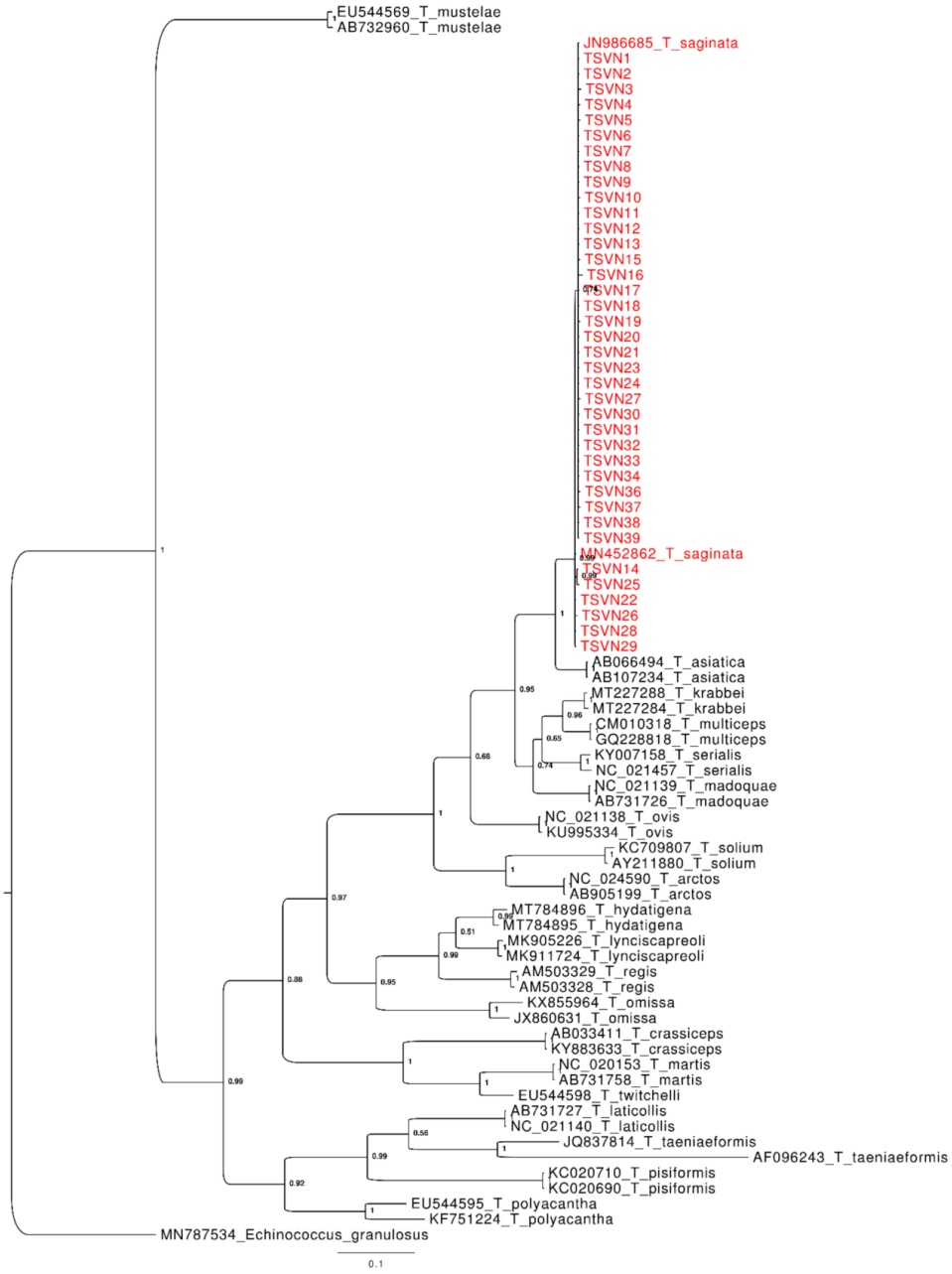


Figure A1. Bayesian phylogenetic tree based on the *Taenia* spp. COI sequences dataset. Posterior probability values are reported at each node. Vietnamese sample codes are as reported in the Table 1 of the main text.

Table A1. Species delimitation methods results based on the analysis of the COI sequences belonging to different *Taenia* species. Specimens with identical numbers within the same column belong to the same taxonomic unit.

Genbank #	Species	NDT	ASAP	PTP/bPTP
MN787534	*Echinococcus granulosus*	1	1	1
JN986685	*Taenia saginata*	2	2	20
MN452862	*Taenia saginata*	2	2	20
AB066494	*Taenia asiatica*	3	3	21
AB107234	*Taenia asiatica*	3	3	21
KC709807	*Taenia solium*	4	4	8
AY211880	*Taenia solium*	4	4	8
NC_024590	*Taenia arctos*	5	5	9
AB905199	*Taenia arctos*	5	5	9
AB033411	*Taenia crassiceps*	6	6	7
KY883633	*Taenia crassiceps*	6	6	7
MT784896	*Taenia hydatigena*	7	7	24
MT784895	*Taenia hydatigena*	8	7	25
MT227288	*Taenia krabbei*	9	8	22
MT227284	*Taenia krabbei*	9	8	22
AB731727	*Taenia laticollis*	10	9	4
NC_021140	*Taenia laticollis*	10	9	4
MK905226	*Taenia lynciscapreoli*	11	10	17
MK911724	*Taenia lynciscapreoli*	11	10	17
NC_021139	*Taenia madoquae*	12	11	16
AB731726	*Taenia madoquae*	12	11	16
NC_020153	*Taenia martis*	13	12	10
AB731758	*Taenia martis*	13	12	10
CM010318	*Taenia multiceps*	14	13	23
GQ228818	*Taenia multiceps*	14	13	23
EU544569	*Taenia mustalae*	15	14	6
AB732960	*Taenia mustalae*	15	14	6
KX855964	*Taenia omissa*	16	15	18
JX860631	*Taenia omissa*	17	16	19
NC_021138	*Taenia ovis*	18	17	13
KU995334	*Taenia ovis*	18	17	13
KC020710	*Taenia pisiformis*	19	18	5
KC020690	*Taenia pisiformis*	19	18	5
EU544595	*Taenia polyacantha*	20	19	14
KF751224	*Taenia polyacantha*	21	20	15
AM503329	*Taenia regis*	22	21	12
AM503328	*Taenia regis*	22	21	12
KY007158	*Taenia serialis*	23	22	26
NC_021457	*Taenia serialis*	23	22	27

Table A1. *Cont.*

Genbank #	Species	NDT	ASAP	PTP/bPTP
EU544598	*Taenia twitchelli*	24	23	11
JQ837814	*Taenia taeniaeformis*	25	24	2
AF096243	*Taenia taeniaeformis*	26	25	3
OL459841	*Taenia* sp. from the present study (TSVN1)	2	2	20
OL459842	*Taenia* sp. from the present study (TSVN2)	2	2	20
OL459843	*Taenia* sp. from the present study (TSVN3)	2	2	20
OL459844	*Taenia* sp. from the present study (TSVN4)	2	2	20
OL459845	*Taenia* sp. from the present study (TSVN5)	2	2	20
OL459846	*Taenia* sp. from the present study (TSVN6)	2	2	20
OL459847	*Taenia* sp. from the present study (TSVN7)	2	2	20
OL459848	*Taenia* sp. from the present study (TSVN8)	2	2	20
OL459849	*Taenia* sp. from the present study (TSVN9)	2	2	20
OL459850	*Taenia* sp. from the present study (TSVN10)	2	2	20
OL459851	*Taenia* sp. from the present study (TSVN11)	2	2	20
OL459852	*Taenia* sp. from the present study (TSVN12)	2	2	20
OL459853	*Taenia* sp. from the present study (TSVN13)	2	2	20
OL459854	*Taenia* sp. from the present study (TSVN14)	2	2	20
OL459855	*Taenia* sp. from the present study (TSVN15)	2	2	20
OL459856	*Taenia* sp. from the present study (TSVN16)	2	2	20
OL459857	*Taenia* sp. from the present study (TSVN17)	2	2	20
OL459858	*Taenia* sp. from the present study (TSVN18)	2	2	20
OL459859	*Taenia* sp. from the present study (TSVN19)	2	2	20
OL459860	*Taenia* sp. from the present study (TSVN20)	2	2	20
OL459861	*Taenia* sp. from the present study (TSVN21)	2	2	20
OL459862	*Taenia* sp. from the present study (TSVN22)	2	2	20
OL459863	*Taenia* sp. from the present study (TSVN23)	2	2	20
OL459864	*Taenia* sp. from the present study (TSVN24)	2	2	20
OL459865	*Taenia* sp. from the present study (TSVN25)	2	2	20
OL459866	*Taenia* sp. from the present study (TSVN26)	2	2	20
OL459867	*Taenia* sp. from the present study (TSVN27)	2	2	20
OL459868	*Taenia* sp. from the present study (TSVN28)	2	2	20
OL459869	*Taenia* sp. from the present study (TSVN29)	2	2	20
OL459870	*Taenia* sp. from the present study (TSVN30)	2	2	20
OL459871	*Taenia* sp. from the present study (TSVN31)	2	2	20
OL459872	*Taenia* sp. from the present study (TSVN32)	2	2	20
OL459873	*Taenia* sp. from the present study (TSVN33)	2	2	20
OL459874	*Taenia* sp. from the present study (TSVN34)	2	2	20
OL459875	*Taenia* sp. from the present study (TSVN36)	2	2	20
OL459876	*Taenia* sp. from the present study (TSVN37)	2	2	20
OL459877	*Taenia* sp. from the present study (TSVN38)	2	2	20

Table A1. *Cont.*

Genbank #	Species	NDT	ASAP	PTP/bPTP
OL459878	*Taenia* sp. from the present study (TSVN39)	2	2	20
TOTAL		26	25	27

References

1. Ito, A.; Nakao, M.; Wandra, T. Human taeniasis and cysticercosis in Asia. *Lancet* **2003**, *362*, 1918–1920. [CrossRef]
2. Sanpool, O.; Rodpai, R.; Intapan, P.M.; Sadaow, L.; Thanchomnang, T.; Laymanivong, S.; Maleewong, W.; Yamasaki, H. Genetic diversity of *Taenia saginata* (Cestoda: Cyclophyllidea) from Lao People's Democratic Republic and northeastern Thailand based on mitochondrial DNA. *Parasites Vectors* **2017**, *10*, 141. [CrossRef]
3. Xiao, L.; Ryan, U.; Feng, Y. Introduction and Public Health, Importance of Foodborne Parasite. In *Biology of Foodborne Parasites*, 1st ed.; Xiao, L., Ryan, U., Feng, Y., Eds.; CRC Press: Boca Raton, FL, USA, 2015; Volume 1, pp. 463–480. [CrossRef]
4. Chang, T.; Jung, B.K.; Hong, S.; Shin, H.; Ryoo, S.; Lee, J.; Lee, K.H.; Park, H.; Eom, K.S.; Khieu, V.; et al. Occurrence of a Hybrid Between *Taenia saginata* and *Taenia asiatica* Tapeworms in Cambodia. *Korean J. Parasitol.* **2021**, *59*, 179. [CrossRef]
5. Eom, K.S.; Jeon, H.K.; Rim, H.J. Geographical distribution of *Taenia asiatica* and related species. *Korean J. Parasitol.* **2009**, *47*, S115. [CrossRef]
6. Pawlowski, Z.; Schultz, M.G. Taeniasis and cysticercosis (*Taenia saginata*). *Adv. Parasitol.* **1972**, *10*, 269–343. [CrossRef]
7. Hoberg, E.P. *Taenia* tapeworms: Their biology, evolution and socioeconomic significance. *Microbes Infect.* **2002**, *4*, 859–866. [CrossRef]
8. Ilsøe, B.; Kyvsgaard, N.C.; Nansen, P.; Henriksen, S.A. A Study on the Survival of *Taenia saginata* Eggs on Soil in Denmark. *Acta Vet. Scand* **1990**, *31*, 153–158. [CrossRef]
9. Dorny, P.; Praet, N. *Taenia saginata* in Europe. *Vet. Parasit.* **2007**, *149*, 22–24. [CrossRef]
10. Eichenberger, R.M.; Thomas, L.F.; Gabriël, S.; Bobić, B.; Devleesschauwer, B.; Robertson, L.J.; Saratsis, A.; Torgerson, P.R.; Braae, U.C.; Dermauw, V.; et al. Epidemiology of *Taenia saginata* taeniosis/cysticercosis: A systematic review of the distribution in East, Southeast and South Asia. *Parasites Vectors* **2020**, *13*, 234. [CrossRef]
11. Somers, R.; Dorny, P.; Geysen, D.; Nguyen, L.A.; Thach, D.C.; Vercruysse, J.; Nguyen, V.K. Human tapeworms in north Vietnam. *Trans. R. Soc. Trop. Med. Hyg.* **2006**, *101*, 275–277. [CrossRef]
12. Van De, N.; Le, T.H.; Lien, P.T.H.; Eom, K.S. Current status of taeniasis and cysticercosis in Vietnam. *Korean J. Parasitol.* **2014**, *52*, 125. [CrossRef]
13. Anantaphruti, M.; Thaenkham, U.; Kusolsuk, T.; Maipanich, W.; Saguankiat, S.; Pubampen, S.; Phuphisut, O. Genetic variation and population genetics of *Taenia saginata* in north and northeast Thailand in relation to *Taenia asiatica*. *J. Parasitol. Res.* **2013**, *2013*. [CrossRef]
14. Wickström, L.M.; Haukisalmi, V.; Varis, S.; Hantula, J.; Henttonen, H. Molecular phylogeny and systematics of anoplocephaline cestodes in rodents and lagomorphs. *Syst. Parasitol.* **2005**, *62*, 83–99. [CrossRef]
15. Mariaux, J. A molecular phylogeny of the Eucestoda. *J. Parasitol.* **1998**, 114–124. [CrossRef]
16. Neov, B.; Vasileva, G.; Radoslavov, G.; Hristov, P.; Littlewood, D.; Georgiev, B. Phylogeny of hymenolepidids (Cestoda: Cyclophyllidea) from mammals: Sequences of 18S rRNA and COI genes confirm major clades revealed by the 28S rRNA analyses. *J. Helminthol.* **2021**, *95*, E23. [CrossRef]
17. Sievers, F.; Higgins, D.G. Clustal Omega. *Curr. Protoc. Bioinform.* **2014**, *48*, 1.25.1–1.25.33. [CrossRef]
18. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
19. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef]
20. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]
21. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]
22. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol.* **2017**, *34*, 772–773. [CrossRef]
23. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
24. Scarpa, F.; Sanna, D.; Cossu, P.; Lai, T.; Casu, M.; Curini-Galletti, M. How to achieve internal fertilization without a vagina: The study case of the genus *Archilina* Ax, 1959 (Platyhelminthes, Proseriata) from Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2057–2073. [CrossRef]
25. Zhang, J.; Kapli, P.; Pavlidis, P.; Stamatakis, A. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **2013**, *29*, 2869–2876. [CrossRef]

26. Scarpa, F.; Cossu, P.; Sanna, D.; Lai, T.; Casu, M.; Curini-Galletti, M. New insights on the genus *Otoplana* Du Plessis, 1889 (Platyhelminthes: Proseriata), with description of two new species from the Canary Islands. *Mar. Biodivers.* **2019**, *49*, 2075–2087. [CrossRef]
27. Hebert, P.D.N.; Cywinska, A.; Ball, S.L.; de Ward, J.R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B.* **2003**, *270*, 313–322. [CrossRef]
28. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **1980**, *16*, 111–120. [CrossRef]
29. Puillandre, N.; Brouillet, S.; Achaz, G. ASAP: Assemble species by automatic partitioning. *Mol. Ecol. Resour.* **2020**, *21*, 609–620. [CrossRef]
30. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research-an update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]
31. Flisser, A.; Viniegra, A.E.; Aguilar-Vega, L.; Garza-Rodriguez, A.; Maravilla, P.; Avila, G. Portrait of human tapeworms. *J. Parasitol.* **2004**, *90*, 914–916. [CrossRef]
32. Galán-Puchades, M.T.; Fuentes, M.V. *Taenia asiatica*: The most neglected human *Taenia* and the possibility of cysticercosis. *Korean J. Parasitol.* **2013**, *51*, 51–54. [CrossRef]
33. Flisser, A. State of the art of *Taenia solium* as compared to *Taenia asiatica*. *Korean J. Parasitol.* **2013**, *51*, 43. [CrossRef]
34. Okello, A.L.; Thomas, L.F. Human taeniasis: Current insights into prevention and management strategies in endemic countries. *Risk Manag. Healthc.* **2017**, *10*, 107. [CrossRef]
35. Mayta, H.; Talley, A.; Gilman, R.H.; Jimenez, J.; Verastegui, M.; Ruiz, M.; Garcia, H.H.; Gonzalez, A. Differentiating *Taenia solium* and *Taenia saginata* infections by simple hematoxylin-eosin staining and PCR-restriction enzyme analysis. *J. Clin. Microbiol.* **2000**, *38*, 133–137. [CrossRef]
36. Bunmee, T.; Chaiwang, N.; Kaewkot, C.; Jaturasitha, S. Current situation and future prospects for beef production in Thailand—A review. *Asian-australas. J. Anim. Sci.* **2018**, *31*, 968. [CrossRef]
37. Phonivsay, A.; Phuthasone, O.; Phonepaseuth, S. *Assessment of Cattle Trade Development in Lao PDR: Study on Potential Impacts of Trade Liberalization under AFTA on Cattle Trade and Its Implication for the Cattle Development Policy in Viet Nam Migration Profile*; International Organization for Migration. National Agriculture and Forestry Research Institute (NAFRI): Vientiane, Laos, 2016.
38. Smith, P.; Lüthi, N.B.; Huachun, L.; Oo, K.N.; Phonvisay, A.; Premashthira, S.; Abila, R.; Widders, P.; Kukreja, K.; Miller, C. *Movement Pathways and Market Chains of Large Ruminants in the Greater Mekong Sub-region*; World Organisation of Animal Health: Bangkok, Thailand, 2015; 60p. [CrossRef]
39. Anh, N.T. Labor Migration flows from Vietnam to Thailand in the context of ASEAN regional integration. *RUDN J. Economics* **2017**, *25*, 275–282. [CrossRef]
40. Ng-Nguyen, D.; Stevenson, M.A.; Breen, K.; Van Phan, T.; Thi Nguyen, V.A.; Vo, T.V.; Traub, R.J. The epidemiology of *Taenia* spp. infection and *Taenia solium* cysticerci exposure in humans in the Central Highlands of Vietnam. *BMC Infect. Dis.* **2018**, *18*, 527. [CrossRef]

Fasciola hepatica

This study was conducted in collaboration with the researchers from the Laboratory of Biodiversity and Environment: Interactions and Genomes, Faculty of Biological Sciences, University of Sciences and Technology Houari Boumediene, Algiers, Algeria and resulted in the publication of the subsequent article: New insights into the genetic variability of *Fasciola hepatica* (Trematoda) in Algeria and relationships with other geographic regions revealed by mitochondrial DNA. doi: 10.2478/helm-2022-0021

This work aimed to explore the genetic variability of *Fasciola hepatica* isolated from cattle in Algeria and to understand their phylogenetic and phylogeographic relationships with sequences from all continents using mitochondrial (Cytochrome c Oxidase subunit I gene - COI) and nuclear markers (Internal Transcribed Spacers of nuclear ribosomal DNA - ITS). The results obtained indicate overall low levels of genetic variation throughout, suggesting a genetic similarity among *Fasciola hepatica* isolates worldwide. Additionally, the study found genetic similarity among *Fasciola hepatica* specimens in different hosts, indicating that the parasite's genetic structure is not influenced by the host species.

Research article

sciendo

DOI 10.2478/helm-2022-0021

HELMINTHOLOGIA, 59, 2: 152 – 164, 2022

New insights into the genetic variability of *Fasciola hepatica* (Trematoda) in Algeria and relationships with other geographic regions revealed by mitochondrial DNA

M. CHAOUADI[1,#], F. SCARPA[2,#], I. AZZENA[2,3,#], P. COSSU[3], K. HARHOURA[4], M. AISSI[4], F. TAZEROUTI[1], G. GARIPPA[3], P. MERELLA[3], M. CASU[3,*], D. SANNA[2,*]

[1]Laboratory of Biodiversity and Environment: Interactions and Genomes, Faculty of Biological Sciences, University of Sciences and Technology Houari Boumediene, BP 32, El Alia Bab Ezzouar, Algiers, Algeria, E-mail: *chaouadi.melissa@gmail.com*, *fadilatazerouti_parasitol@yahoo.com*; [2]Department of Biomedical Sciences, University of Sassari, 07100 Sassari, Italy, E-mail: **darsanna@uniss.it*, *fscarpa@uniss.it*, *iazzena@uniss.it*; [3]Department of Veterinary Medicine, University of Sassari, 07100 Sassari, Italy, E-mail: *picossu@uniss.it*, **marcasu@uniss.it*, *garippa@uniss.it*, *paolomerella@uniss.it*; [4]Laboratory of Animal Health and Production, Superior National Veterinary School Rabie Bouchama, Street Issad Abbes, Oued Smar, Algiers, Algeria, E-mail: *harhourakhaled@yahoo.fr*, *aissimiriem@yahoo.fr*

Article info

Received January 19, 2022
Accepted July 14, 2022

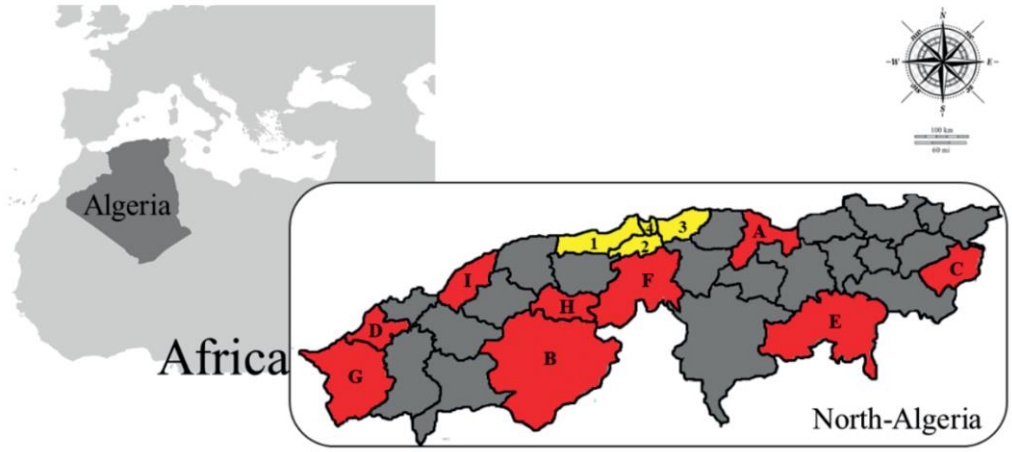
Summary

This study aims to investigate the level of genetic variability of *Fasciola hepatica* flukes isolated from cattle in Algeria and to determine the phylogenetic and phylogeographic relationships with sequences isolated worldwide. Mitochondrial (Cytochrome c Oxidase subunit I gene - COI) and nuclear markers (Internal Transcribed Spacers of nuclear ribosomal DNA - ITS) for 24 *F. hepatica* flukes isolated from 12 cattle in North Algeria were characterised. Only two haplotypes were obtained for the COI gene, resulting in a low level of genetic variation. The analysis of variation among the COI sequences isolated from around the world did not show high levels of genetic divergence, and the phylogenetic analysis revealed a genetic similarity among *F. hepatica* isolates from different areas of the world. The analysis of the ITS region showed a low level of variability, which prevented obtaining informative phylogenetic and phylogeographic results. The present study also revealed that specimens of *F. hepatica* are genetically similar in different hosts, indicating that the genetic structure among populations of this parasite is not influenced by the host species. The low levels of genetic variation for COI and ITS regions among fluke isolates from all continents are consistent with a common origin for the flukes' worldwide distribution.
Keywords: *Fasciola hepatica*; molecular characterization; COI; ITS; phylogeography

Introduction

Fasciolosis is a parasitic disease caused by two trematode species of the genus *Fasciola, F. hepatica* Linnaeus, 1758 and *F. gigantica* Cobbold, 1856, which affect the liver of herbivores and humans. The *Fasciola* infection has a negative impact on public health (Mas-Coma *et al*., 2009), and it is also responsible for significant economic losses in livestock production due to the reduction of meat and milk production, slowing growth and reducing fertility. Moreover, infected livers are condemned at meat inspection and removed from sale (Kaplan, 2001). Fasciolosis has a worldwide distribution due to the occurrence of *F. hepatica* in the temperate zone and *F. gigantica* in the tropical zone (Mehmood *et al*., 2017). Both species coexist in subtropical areas (Mas-Coma *et al*., 2005). *Fasciola hepatica* is the only endemic species causing fasciolo-

* – corresponding author
# – These authors equally contributed to the paper.



From GenBank → Bejaïa (A), Tiaret (B), Souk-Ahras (C), Ain-Temouchent (D), Batna (E), Médéa (F), Tlemcen (G), Tissemsilt (H) and Mostaganem (I).
Present study → Tipaza (1), Blida (2), Boumerdès (3) and Algiers (4).

Fig. 1. Map of the Algerian sample collection sites. The map shows the geographical origin of the sequences from Algeria isolated in the present study and those from GenBank and BOLD.

sis in Algeria and represents one of the most common helminths infecting ruminants in this country. However, the occurrence of *F. gigantica* was recently reported for the first time in sheep from southern Algeria, close to the border with Mali (Chougar *et al.*, 2020). The hosts had a trans-Saharan geographical origin, with introduction from Ghana, through the Sahel countries of Burkina Faso and Mali into Algeria (Chougar *et al.*, 2020).

The prevalence of *F. hepatica* infection in cattle varies from one Algerian region to another. Infection levels of up to 27.6 % (Ouchene-Khelifi *et al.*, 2018) and 52.4 % (Boucheikhchoukh *et al.*, 2012) were in general reported for the eastern area but in this same region, Chougar *et al.* (2019) reported prevalence reaching 22.3 %. A lower prevalence (from 6 % to 18 %) was reported in the north-central area (Aissi *et al.*, 2009; Chaouadi *et al.*, 2019), while bovine fasciolosis is less frequent in western and southern Algeria, where prevalences from 2.3 % to 4.4 % (Chougar *et al.*, 2019) and 1.7 % (Ouchene-Khelifi *et al.*, 2018) were reported.

The morphological identification of the two species of *Fasciola* is based on key characteristics (Periago *et al.*, 2006; Valero *et al.*, 2018): leaf-shaped body, evident shoulders, and oblique body angle for *F. hepatica*, and indistinct shoulders, parallel margins, and rounded posterior end for *F. gigantica*. However, several studies have shown the characteristics of the two species overlapping (Ashrafi *et al.*, 2006; Periago *et al.*, 2008) and the existence of intermediated forms in localities where both species coexist, due to phenomenon of hybridisation (Alasaad *et al.*, 2007; Amer *et al.*, 2016). Using molecular tools is considered the best method to provide a clear differentiation between *F. hepatica* and *F. gigantica* (Mas-Coma *et al.*, 2009; Itagaki *et al.*, 2005; Le *et al.*, 2008; Shoriki *et al.*, 2016) and contribute to a better understanding of genetic diversity, origin, evolution, and expansion of this parasite (Semyenova *et al.*, 2006; Itagaki *et al.*, 2009; Ai *et al.*, 2011; Thang *et al.*, 2020). Internal transcribed spacers of nuclear ribosomal DNA (ITS) have proven to be useful in performing a proper taxonomic discrimination among *Fasciola* species. Mitochondrial genes may be highly polymorphic even at an intraspecific level and represent powerful markers to infer the evolutionary processes and phylogenetic relationships within populations of *Fasciola* species (Itagaki *et al.*, 2005; Semyenova *et al.*, 2006; Cwiklinski *et al.*, 2015).

Some studies have explored the genetic variability of *F. hepatica* in Algeria (Farjallah *et al.*, 2009; Chougar *et al.*, 2019; Laatamna *et al.*, 2021); and until now, the occurrence of two ITS haplotypes (Chougar *et al.*, 2019), from 7 to 20 Cytochrome c Oxidase subunit I (COI) haplotypes (Chougar *et al.*, 2019; Laatamna *et al.*, 2021) and from 5 to 24 NADH dehydrogenase subunit I (NADH) haplotypes (Chougar *et al.*, 2019; Laatamna *et al.*, 2021) revealed a weak population structuring worldwide for *F. hepatica*. Haplotypes from Algeria are closely related to those from Europe and Africa (Chougar *et al.*, 2019; Laatamna *et al.*, 2021).

A recent study on *Fasciola* flukes from several populations located within the Tunisian-Algerian border on combined novel (pepck and pold) and common molecular markers (ITS, COI, NADH and COI-

Table 1. Sampling plan. The table reports data on the sampling collection, the GenBank accession numbers of the sequences obtained in the present study, and the haplotypes (Hap. type) found among individuals. The unique allelic variant isolated for the ITS fragment in all the samples analysed in the present study was deposited in GenBank under the accession number MZ292402. The presence of an identical host code for different samples indicates that flukes were isolated from the same host.

Sample code	Host code	Area	Site	Host	Sampling date	GenBank COI #	Hap. type
C1_142	142	Algeria	Tipaza	Cattle	03-29-2016	MT920965	2
C2_142	142	Algeria	Tipaza	Cattle	03-29-2016	MT920966	1
C1_170	170	Algeria	Tipaza	Cattle	04-06-2016	MT920980	1
C1_858	858	Algeria	Blida	Cattle	10-03-2016	MT920969	1
C2_858	858	Algeria	Blida	Cattle	10-03-2016	MT920970	1
C1_903	903	Algeria	Boumerdes	Cattle	10-10-2016	MT920978	1
C1_995	995	Algeria	Algiers	Cattle	11-02-2016	MT920967	1
C2_995	995	Algeria	Algiers	Cattle	11-02-2016	MT920968	2
C1_1000	1000	Algeria	Algiers	Cattle	11-02-2016	MT920974	1
C2_1000	1000	Algeria	Algiers	Cattle	11-02-2016	MT920975	1
C1_1110	1110	Algeria	Boumerdes	Cattle	11-26-2016	MT920982	1
C1_1211	1211	Algeria	Algiers	Cattle	12-31-2016	MT920976	1
C2_1211	1211	Algeria	Algiers	Cattle	12-31-2016	MT920977	1
C1_1215	1215	Algeria	Algiers	Cattle	12-31-2016	MT920981	1
C1_1230	1230	Algeria	Algiers	Cattle	12-31-2016	MT920983	1
C2_1230	1230	Algeria	Algiers	Cattle	12-31-2016	MT920984	1
C3_1230	1230	Algeria	Algiers	Cattle	12-31-2016	MT920985	1
C4_1230	1230	Algeria	Algiers	Cattle	12-31-2016	MT920986	1
C5_1230	1230	Algeria	Algiers	Cattle	12-31-2016	MT920987	1
C6_1230	1230	Algeria	Algiers	Cattle	12-31-2016	MT920988	1
C1_1279	1279	Algeria	Algiers	Cattle	01-18-2017	MT920979	1
C1_1342	1342	Algeria	Algiers	Cattle	02-04-2017	MT920971	1
C2_1342	1342	Algeria	Algiers	Cattle	02-04-2017	MT920972	1
C4_1342	1342	Algeria	Algiers	Cattle	02-04-2017	MT920973	1
FHLAC1	LAC	Italy	Laconi	Cattle	12-09-2013	MT920989	1
FHLAC2	LAC	Italy	Laconi	Cattle	12-09-2013	MT920990	1
FHLAC3	LAC	Italy	Laconi	Cattle	12-09-2013	MT920991	1
FHLAC4	LAC	Italy	Laconi	Cattle	12-09-2013	MT920992	1
FHLAC5	LAC	Italy	Laconi	Cattle	12-09-2013	MT920993	1
FHLAC6	LAC	Italy	Laconi	Cattle	12-09-2013	MT920994	1
FHLAC7	LAC	Italy	Laconi	Cattle	12-09-2013	MT920995	1
FHLAC8	LAC	Italy	Laconi	Cattle	12-09-2013	MT920996	1
FHGIA1	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT920997	1
FHGIA2	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT920998	1
FHGIA3	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT920999	1
FHGIA4	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT921000	1
FHGIA5	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT921001	1
FHGIA6	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT921002	1
FHGIA7	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT921003	1
FHGIA8	GIA	Italy	Giara di Genoni	Goat	11-28-2013	MT921004	1

trnT-rrnL) revealed a relevant gene flow between Tunisian and Algerian populations of *F. hepatica* (Amor *et al*. 2020; Chougar *et al*., 2020).

The present study aims to investigate the levels of genetic variability among specimens of *F. hepatica* in Algeria, and it uses two molecular markers with different systems of transmission to the offspring. The first marker is a mitochondrial encoding gene (COI) that does not recombine, is uniparentally inherited, and it is extensively used to depict the phylogeographic patterns of the distribution of species. The second marker is a nuclear noncoding region (ITS) that undergoes a rapid concerted evolution via unequal crossing-over and gene conversion, and it could be highly variable and effective in depicting genetic structuring among groups of species.

Materials and Methods

Sampling

In the present study, 24 individuals of *F. hepatica* were collected from 12 cattle slaughtered at the Mitidja area in the North-center of Algeria (1 to 6 flukes *per* each host were isolated) between March 2016 and February 2017 (Table 1 and Fig. 1). Furthermore, 16 individuals of *F. hepatica*, from one cattle (8 flukes) and one goat (8 flukes) in the Mediterranean island of Sardinia (Italy), in September and November of 2013 (see Table 1 for details), were collected with the aim to enlarge the dataset of isolates used for comparison with Algerian isolates.

DNA extraction, PCR and sequencing

Genomic DNA of the specimens was extracted using the kit Macherey-Nagel NucleoSpin Tissue (MACHEREY-NAGEL GmbH & Co. KG), following the protocol used by Cossu *et al.* (2015). Sample quality and DNA concentration were determined via spectrophotometry using a NanoDrop™ Lite (NanoDrop Technologies, Thermo Fisher Scientific Inc., Wilmington, DE). The DNA mean concentration obtained for the samples was 75 ng/µL.

PCR amplification of a partial fragment of the COI gene (441 bp) for samples from Algeria and Sardinia were performed using the primers, Ita 8 (forward; 5'-ACGTTGGATCATAAGCGTGT-3') and Ita 9 (reverse: 5'-CCTCATCCAACATAACCTCT-3') (Itagaki *et al*., 2005). Furthermore, PCRs were also performed for a fragment (900 bp) of the nuclear region including ITS-1, 5.8S rDNA, and ITS-2 (ITS) using the primers BD1 (forward: 5'-GTCGTAACAAGGTTTCCGTA-3') and BD2 (reverse: 5'-TATGCTTAAATTCAGCGGGT-3') (Luton *et al*., 1992). All PCRs were carried out in a total volume of 25 µl containing 10 ng of total genomic DNA on average which was combined with 0.6 µM of each primer and one pellet of PuReTaq Ready-To-Go PCR beads (GE Healthcare; 9900 West Innovation Drive, Wauwatosa, WI, USA). Each pellet of PuReTaq Ready-To-Go PCR beads contained reaction buffer, 2.5 units of PuReTaq DNA polymerase, bovine serum albumin (BSA), deoxynucleotide triphosphates (dNTPs) and stabilizers. For each bead reconstituted to a 25 µl final volume, the concentration of each dNTP was 200 µM and of $MgCl_2$ was 1.5 mM. The PCR conditions were 4 min at 94 °C as an initial step, followed by 35 cycles of 30 sec at 94 °C, 30 sec at the annealing temperature (56 °C for COI and 57 °C for ITS), and 30 sec at 72 °C, with a final post-treatment of 5 min at 72 °C. Both positive and negative controls were used to test the effectiveness of the PCR protocols, and the absence of possible contamination. The PCR products were visualized on 2 % agarose gels (TAE 1×) and purified by ExoSAP-IT (USB Corporation). Sequencing was performed for both strands using the PCR primers by an external sequencing core service (Macrogen Europe).

Phylogeographic and phylogenetic analyses

The sequences obtained for specimens of *F. hepatica* from Algeria in the present study were merged with those available for this species in GenBank and Barcode of Life Data system (BOLD) from all over the world (see Fig. S1 in Supplementary Materials for GenBank accession numbers), with the scope to perform a broader phylogeographic analysis of *F. hepatica* in Algeria and the rest of the world.

The sequences of *F. hepatica* isolated in Sardinia (Italy) were also included in the analysis, to involve data also from this poorly investigated western Mediterranean island.

Forty contiguous sequences for COI and 32 for the ITS, were aligned and inspected for errors using the package Clustal Omega (Sievers & Higgins, 2014) available at https://www.ebi.ac.uk/Tools/msa/clustalo/) and the data were deposited in the GenBank (see Table 1 for GenBank accession numbers).

The genetic variation within the datasets was assessed estimating the number of polymorphic sites (*S*), number of haplotypes (*H*), haplotype diversity (*hd*), and nucleotide diversity (π) using the software package DnaSP 6.12.03 (Librado & Rozas, 2009).

Median-joining networks (Bandelt *et al*., 1999) were constructed using the software package Network 10.0.0.0 (www.fluxus-engineering.com) to infer the genetic relationships among haplotypes and allelic variants, thus detecting the possible occurrence of evolutionary forces acting on populations. The transitions and transversions were equally weighed. Due to the lack of knowledge regarding the possible occurrence of retromutation events, the same weight (10) was assigned to all of the observed polymorphisms.

The Tajima's *D* (Tajima, 1989) and Fu's Fs (Fu, 1997) neutrality tests were performed using the software package DnaSP 6.12.03 (Librado & Rozas, 2009) to infer departures from equilibrium models of the Algerian population. Combining different neutrality tests can help to distinguish among the different evolutionary processes responsible for departures from equilibrium; Fu's Fs can better detect demographic expansions, whereas Tajima's D can better detect bottlenecks and populations contractions (Soriano *et al*., 2008).

The best probabilistic model of sequence evolution was determined using jModeltest 2.1.1 (Posada, 2008), with a maximum

likelihood optimised search. The Akaike Information Criterion (AIC) found "TPM3uf+I+G" as the best-fitting model, while the Bayesian Information Criterion (BIC) found the "HKY+G" model. The parameters of the more sophisticated model between the two which were detected were used for input files (i.e. TPM3uf+I+G). Phylogenetic relationships among different taxa (if any) were investigated using a species tree based on Bayesian Inference (BI) by means of the software MrBayes 3.2.7 specifying setting as model parameters: NST = 3, rates = invgamma, ngammacat = 4. Two independent runs, each consisting of four metropolis-coupled MCMC chains (one cold and three heated chains), were run simultaneously for 5,000,000 generations, sampling trees every 1,000 generations. The first 25 % of the 10,000 sampled trees was discarded as burn-in. Runs were executed by means of the CIPRES Phylogenetic Portal (Miller *et al*., 2010). In order to verify the convergence of chains, it was checked that the average standard deviation of split frequencies (ASDSF), approached 0 (Ronquist *et al*., 2012), and the Potential Scale Reduction Factor (PSRF) was around 1 (Gelman & Rubin, 1992) following Scarpa *et al*. (2019a).

Phylogenetic trees were visualized and edited using FigTree 1.4.1 (available at http://tree.bio.ed.ac.uk/software/figtree/).

To verify the taxonomic assessment of every sequence in the dataset, four different methods of species delimitation, which are listed below, were used.

The ST-GMYC (Single Threshold-Generalized Mixed Yule Coalescent) method (Pons *et al*., 2006), which follows the phylogenetic species concept to delimit species, was applied by means of the SPLITS (SPecies LImits by Threshold Statistics) package (Ezard *et al*., 2009) implemented in the R statistical environment (available at http://r-forge.r-project.org/projects/splits/) on the ultrametric species tree which was obtained by the software Beast 1.10.4 (Drummond & Rambaut, 2007) following Scarpa *et al*. (2018). The nucleotide divergence threshold (NDT) method was implemented by means of a script (Scarpa *et al*., 2019b) written in the R statistical environment. For the K/θ method (Birky *et al*., 2010), used with the corrected formula for sexual organism showed in Birky (2013), clades were selected on the topology of a mid-point rooted Neighbour-joining (NJ) tree (Saitou & Nei, 1987) obtained using the R package APE (Analysis of Phylogenetics and Evolution) (Paradis *et al*., 2004). Clades showing values of K/θ ≥ 4 should be considered as well-defined entities with a 95 % probability of having an independent evolutionary history. The ASAP (Assemble Species by Automatic Partitioning) method (Puillandre *et al*., 2020), which is fully exploratory (it does not require any kind of a priori knowledge), was performed using the p-distance model (as substitution model to calculate the distances matrix), selecting default options. Within the list of the best partitions, the species hypothesis, valuating their gap-width score, p-value and threshold distance following Puillandre *et al*. (2020) were chosen.

On the datasets obtained the principal coordinate analysis (PCoA) was performed using GenAlEX 6.5 (Peakall & Smouse, 2012). This analysis allows to distinguish genetic clusters running on a pairwise genetic distance matrix corrected with K2P (Kimura, 1980) model. The rate of variation among sites was modelled with a gamma distribution and all ambiguous positions were removed for each sequence pair.

Ethical Approval

The manuscript does not contain clinical studies or patient data. Sampling of parasites was not performed on live animals but only on tissues collected post-mortem in a slaughterhouse.

Results

COI

Twenty-four sequences of the central portion of the COI gene (441 bp) were obtained for the samples from Algeria in the present study (Table 1). Among them, only one polymorphic site was found that defined two haplotypes (type 1 and type 2, see Table 1 for details) that were shared by 92 % and 8 % of the samples, respectively (see Table 2 for details on the genetic divergence estimates). The two haplotypes diverged from one another for one neutral point-mutation (transition A → G), which does not affect the protein structure since it produces a change between two non-polar aliphatic amino acids (isoleucine → valine). The mutation occurred at position 799 of the COI gene nucleotide sequence (reference sequence used for the *F. hepatica* COI gene: NC_002546).

A COI dataset, which included the sequences from Algeria obtained in the present study and those of *F. hepatica* from Algeria recorded in GenBank (see Fig. 1 for details on the geographic origin of the sequences), was constructed to infer a set of sequences that could likely represents the Algerian population. The dataset showed low levels of genetic variation and it included 32 sequences (24 from the present study and 8 from GenBank), with 7 polymorphic sites that defined 7 haplotypes (see Table 2 for details on genetic divergence estimates).

A further COI dataset, including the sequences from Algeria (24)

Table 2. Indices of genetic variation. The table reports the estimates of genetic variation for the mitochondrial COI gene dataset. N: sample sizes; bp: fragment size; *S*: number of polymorphic sites; *H*: number of haplotypes; *hd*: haplotype diversity; *π*: nucleotide diversity.

	N	bp	*S*	*H*	*hd*	*π*
Samples from Algeria – present study	24	441	1	2	0.159	0.00036
Samples from Algeria – whole dataset	32	441	7	7	0.393	0.00271
Total COI dataset	187	441	42	32	0.753	0.00664

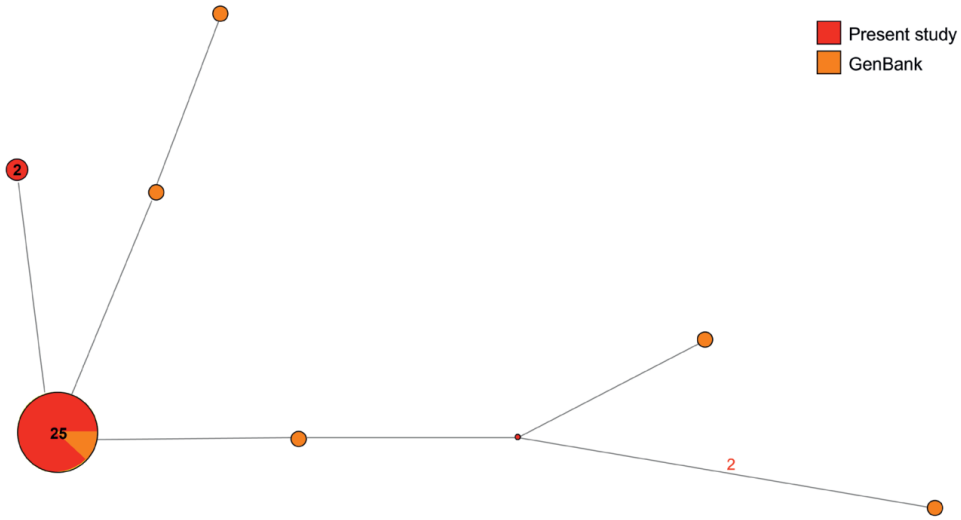


Fig. 2. Median-joining network analysis. The network includes COI sequences from Algeria obtained in the present study along with those from GenBank. The small red plot on one node shows a median vector representing the hypothetical connecting sequence that was calculated using the maximum parsimony method. The number of mutations between sequences that are greater than 1 are reported on network branches. As well, the number of individuals showing the same haplotype that is greater than 1 is reported inside the spot.

and Sardinia (16) obtained in the present study and those (147) corresponding to the same portion of the COI gene (441 base pairs of the central fragment of the gene) from *F. hepatica* strains isolated worldwide and deposited in GenBank and BOLD (see Table S1 for the accession numbers), was also constructed to include the data obtained in the present study on a wider geographic context. This dataset included 187 COI sequences of *F. hepatica* belonging to 15 countries from every continent except South America (see Supplementary Table S1 for details). Among the sequences, 42 polymorphic sites were retrieved, resulting in a good level of genetic divergence that corresponds to 32 haplotypes (see Table 2 for further details on the genetic divergence estimates). In particular, the sequences from Algeria and Sardinia obtained in the present study belong to the most frequent worldwide diffused haplotypes, except for two never-reported haplotypes found in Algeria (in two flukes from two different cattle, C1_142 from Tipaza and C2_995 from Algiers in Table 1) and one fluke in Sardinia (FHLAC 5 from Laconi in Table 1). Interestingly, these two uncommon lineages were isolated from flukes infecting cattle that were also infected with flukes characterised by the most common COI haplotype.

Haplotype network analysis and neutrality tests

The network analysis based on sequences from Algeria (Fig. 2) showed a well-defined, star-like shape with a major and highly diffused haplotype that was found in 78 % of sequences and 6 derived haplotypes diverging for 1 to 4 point-mutations. Almost all the derived haplotypes were exclusive to single individuals, except for one lineage that was found in two individuals from Algiers (sample C2_995 in Table 1) and Tipaza (sample C1_142 in Table 1). Overall, 22 of the sequences from Algeria obtained in the present study belonged to the most frequent haplotype of the network.

The neutrality tests performed on the same dataset of sequences from Algeria showed non-significant negative values of *D* (-1.604 with $0.10 \geq P \geq 0.05$) and Fu's Fs (-3.567 with $P \geq 0.10$).

The network analysis performed on the COI dataset, including sequences from all over the world (Fig. 3), showed evidence of the occurrence of three highly diffused haplotypes, which are surrounded by many derived lineages that diverged for a single point mutation from the central ancestor and are generally exclusive to single individuals. Two of the three most frequent haplotypes of the network were diffused across nearly every country included in the analysis, while the third most frequent haplotype was exclusive to sequences from Spain, aside from one sequence from Austria. Furthermore, 10.7 % of the haplotypes included in the dataset were exclusive to single individuals, and one haplotype found in three flukes from China was highly divergent (more than 20 point-mutations) from the others.

The neutrality tests performed on this dataset of sequences showed a significant departure from the equilibrium, with a significant negative value for Tajima's *D* test (D=-2.065 with $P \leq 0.05$) that is consistent with population expansions and a non-significant negative value for Fu's Fs (-13.864 with $P \geq 0.10$).

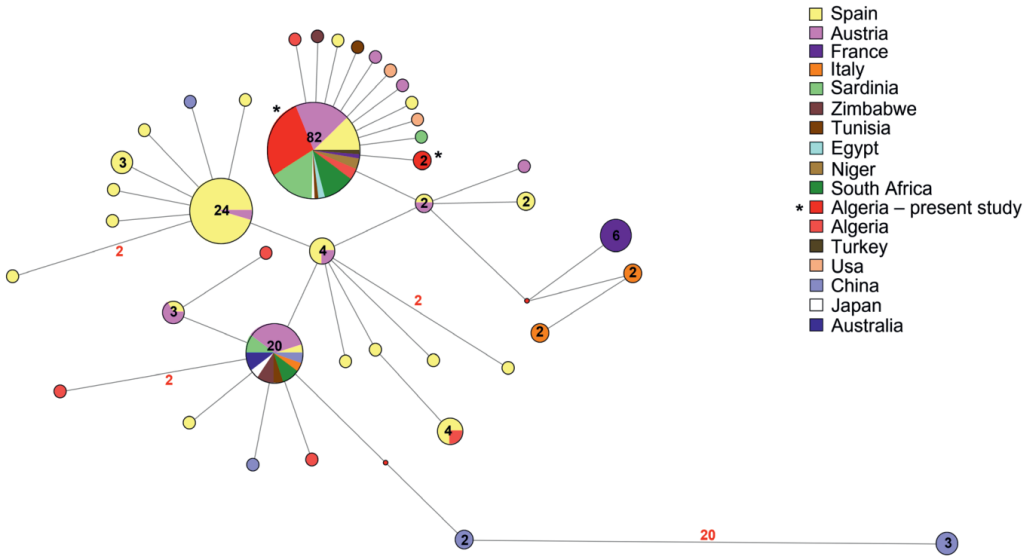


Fig. 3. Median-joining network analysis. The network includes all COI sequences from the present study along with those from GenBank. The small red plots on one node show a median vector representing the hypothetical connecting sequence that was calculated using the maximum parsimony method. The number of mutations between sequences that are greater than 1 are reported on network branches. As well, the number of individuals showing the same haplotype that is greater than 1 is reported inside the spot. The MP calculation post-processing option, that uses only the shortest trees sufficient to generate the graphic output, has been applied for drawing the network. This option allows to obtain a network without showing the reticulations. All Sardinian sequences in the network are form the present study.

Phylogenetic analyses and PCoA

For phylogenetic inferences, one sequence of *F. jacksonii* and three sequences of *F. gigantica* were included within the COI dataset of sequences from around the world as outgroups (see Fig. S1 in Supplementary Materials for GenBank accession numbers).

The phylogenetic tree (Fig. S1 in Supplementary Materials) showed a unique, well-supported monophyletic cluster that included all *F. hepatica* sequences, except for five sequences from one province in central China (Gansu), which were deposited in GenBank in 2016 and isolated from goats. In particular, two of these sequences set within the large *F. hepatica* clade in a poorly supported internal sub-cluster, while the three remaining sequences set in an internal well-supported sub-cluster within the *F. gigantica* clade.

Consistently, every species delimitation method that was used showed that all the COI sequences of *F. hepatica* included in the dataset belong to a unique, worldwide-distributed taxonomic entity. In accordance with results of the phylogenetic tree, the only exception was represented by three Chinese sequences from Gansu that belonged to the same taxonomic unit found for *F. gigantica*.

Principal coordinate analysis (PCoA) was performed on all the COI sequences of *F. hepatica* (n=184), except for the three sequences from China that belong to the *F. gigantica* variability. PCoA explained 54.76 % of the variability (PCoA1/x-axis: 40.96 %, PCoA2/y-axis: 13.80 %). Results (see Fig. 4 and Table S2) showed the occurrence of three genetic groups, including 59.24 % (Group A), 25.54 % (Group B), and 15.22 % (Group C) of the sequences, respectively, along with one sequence isolated in Algeria from one cattle that set as an outlier outside the three groups.

Group A mainly spread in western Europe and Africa, with only a few sequences isolated on other continents; individuals of *F. hepatica* included in this group were isolated from cattle, sheep, and goats without a specific structuring pattern related to the hosts. Group B was almost exclusive to sequences isolated from sheep in Spain, with a few sequences generally isolated in cattle from China, Italy, Austria, and Algeria. Group C was less frequent and scattered across all continents, particularly in flukes isolated from Chinese goats, European cattle and sheep, and African cattle. No association was found between genetic groups and sample collection dates or host species.

ITS

For both Algerian and Sardinian flukes, 32 identical sequences (see Table 1 for details) were obtained for the nuclear ITS fragment in the present study (GenBank accession number: MZ292402). It was not possible to obtain good and scorable sequences for 8 samples, which were not included in the analyses.

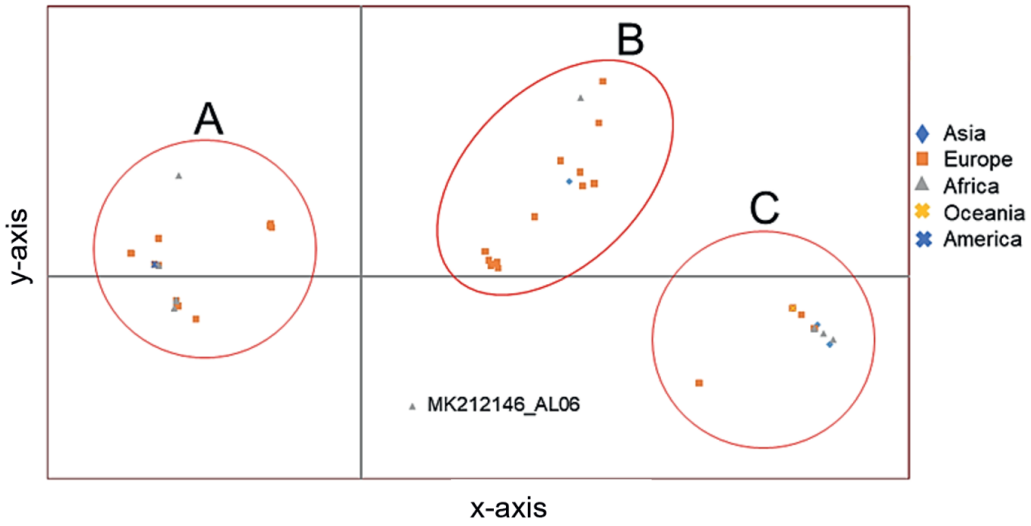


Fig. 4. Principal coordinates analysis performed on the COI gene dataset. Bi-dimensional plots show the genetic differentiation among populations due to the base differences per site found in the dataset. Percentage of variation explained by the first three axes for the COI dataset: 1st = 40.96, 2nd = 13.80, 3rd=9.31.

A dataset was constructed that included all 32 identical ITS sequences obtained in the present study, along with those (137) from GenBank that exactly matched the ITS fragment used in the present research (see supplementary Table S3 for the GenBank accession numbers and further details). The dataset included 169 sequences of *F. hepatica* (905 bp) belonging to 14 countries from every continent, except Oceania. Within the dataset, 6 polymorphic sites were retrieved, resulting in a very low level of genetic divergence (*hd*: 0.058, π: 0.00044) that corresponded to 5 allelic variants. All but 5 sequences (97 %) of the dataset belonged to the most frequent worldwide diffused allelic variant. The only exceptions were represented by 5 sequences from Asia (4 from China and 1 from Iran).

Because of the low level of genetic variability found among ITS variants, neither the network nor phylogenetic tree and neutrality tests analyses were informative (data not shown).

One sequence of *F. jacksonii* (GB# MN970006) and three of *F. gigantica* (GB# MW793531, JF432073, MW793533) were included in the ITS dataset for species delimitation analyses. Every species delimitation method showed that all ITS sequences of *F. hepatica* belonged to a unique, worldwide-distributed taxonomic entity. The only exceptions were represented by two sequences from China included in the dataset, which showed a divergent haplotype belonging to the same taxonomic unit found for *F. gigantica*.

PCoA was performed on 167 sequences of *F. hepatica*, excluding the two outlier sequences from China, which likely fell within the variability of the species *F. gigantica*. PCoA explained 60.08 % of the variability (PCoA1/x-axis: 36.71 %, PCoA2/y-axis: 23.37 %). The results (Fig. 5 and Table S4) showed the occurrence of three genetic groups, including 91 % (Group A), 6 % (Group B), and 3 % (Group C) of the sequences, respectively. A genetic similarity was found along the x-axis between Groups A and B. Group A was the most common and it is present on all continents; individuals of *F. hepatica* included in this group were isolated from different species of ruminants without a specific pattern of structuring related to the hosts. Group B included sequences isolated from several hosts that spread in the Iberian Peninsula (Spain and Andorra), North Africa (Tunisia), Central America (Mexico), and South America (Bolivia). Group C was exclusive to flukes isolated in cattle from China. No evidence of relations was found between genetic groups and sample collection dates, or host species.

Discussion

The mitochondrial and nuclear markers used in this study identified all flukes from Algeria as *F. hepatica*. Although Chougar *et al.* (2020) recently revealed the presence of *F. gigantica* in Algerian sheep, present results confirmed the dominance of *F. hepatica* in Algeria, in accordance with previous molecular studies that found only *F. hepatica* in this country (Chougar *et al.*, 2019; Farjallah *et al.*, 2009; Laatamna *et al.*, 2021; Farjallah *et al.*, 2013). In accordance with previous studies focused on the genetic variation of

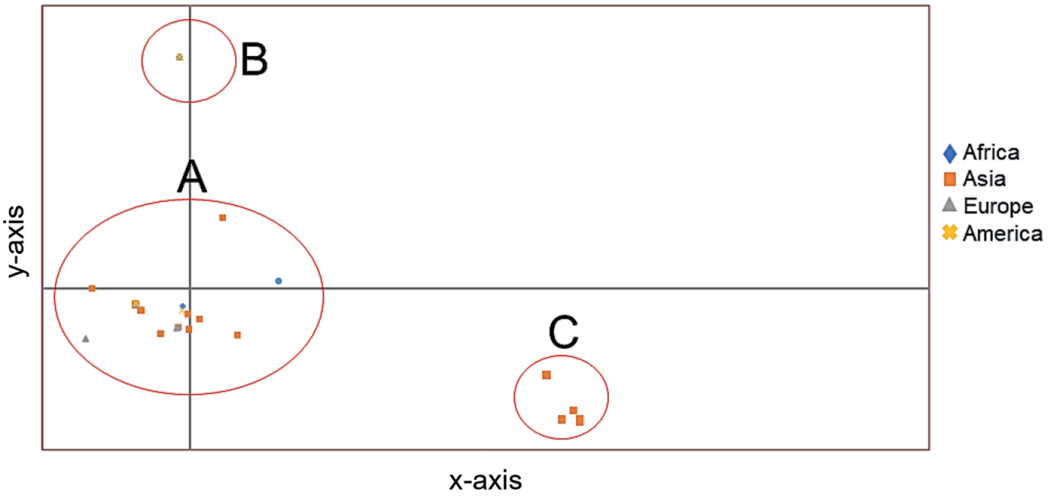


Fig. 5. Principal coordinates analysis performed on the ITS fragment dataset. Bi-dimensional plots show the genetic differentiation among populations due to the base differences per site found in the dataset. Percentage of variation explained by the first three axes for the ITS dataset: 1st = 36.71, 2nd = 23.37, 3rd=11.22.

this parasite in Algeria (Chougar *et al*., 2019), the analyses of the COI sequences' variability showed a generally low level of genetic divergence among individuals in Algeria, with traces of a recent population expansion, as suggested by the Tajima's *D* neutrality test results (Fu, 1997). This trend is supported by the low level of genetic variability that was found among sequences that is usually observed in areas recently colonised by these parasites (Mas-Coma *et al*., 2009; Hewitt, 2000; Robinson & Dalton, 2009). The possible recent introduction of *F. hepatica* in the Algerian sites is consistent with the fact that this species is native to Europe and expanded its geographic distribution quite recently after Europeans operated a global colonisation with livestock movement over the past five centuries (Mas-Coma *et al*., 2009).

In this context, the most common COI haplotype found in Algerian isolates (from GenBank and the present study) may correspond to one of the oldest mitochondrial variants present in this country and may be representative of the first lineages introduced from Europe (Walker *et al*., 2007). The few haplotypes found to be exclusive to single individuals may have recently derived *in situ* from European founders. They might have differentiated because of silent or neutral nucleotide mutations that originated synonymous codons or amino acids with similar chemical structures.

Interestingly, according to what other authors have already reported (Walker *et al*., 2007; Elliot *et al*., 2014), the two new COI haplotypes found in the present study for *F. hepatica* in Algeria and Sardinia co-occurred in hosts where other different mitochondrial lineages were also present. Consistently, Walker *et al*. (2007) found several mitochondrial composite PCR–restriction fragment length polymorphism haplotypes in the same cattle, and Elliot *et al*. (2014) found sheep and cattle with up to ten different mtDNA genotypes. These authors explained their findings by considering the possible occurrence of host infections with diverse fluke metacercariae coexisting in the geographical area where animals usually live or otherwise considering that livestock may have occasionally moved and grazed towards areas where *F. hepatica* individuals are also characterised by rare mitochondrial haplotypes. For these reasons, plants growing on the edges of wades, rivers, marshes, and irrigation canals could be sources of infection in Algeria (Massot & Senouci-Horr, 1983).

Since a low level of genetic divergence was found for the COI fragment analyzed among isolates on every continent, the present study suggests a common origin of flukes sharing the same haplotype, as it was also suggested by other authors (Le *et al*., 2000; Lotfy *et al*., 2008; Amor *et al*., 2011; Simsek *et al*., 2011; Martinez-Valladares & Rojo-Vazquez, 2014; Mucheka *et al*., 2015). Accordingly, a unique taxonomic entity corresponding to the monophyletic clade of *F. hepatica,* evidenced by phylogenetic tree, further supports the genetic affinity among *F. hepatica* isolates from different parts of the world. Consistently, Semyenova *et al*. (2006) also reported low levels of genetic variability of the COI gene in *F. hepatica* in several countries, with 10 haplotypes found and only 2.3 % of polymorphic sites. Although these authors

considered a different fragment of the COI gene from the one the present study analysed, they demonstrated reduced levels of genetic structuring among hosts and geographic regions in Russia, Belarus, Ukraine, Bulgaria, Armenia, Azerbaijan, Georgia, Turkey, Turkmenistan, and China.

The results obtained in the present study also suggest a lack of association between the genetic structuring of the COI gene and host species. Similarly, Santos (2012), who analysed a different and more variable fragment of the COI gene than the present study examined, reported the same haplotype diffused in different hosts (cattle and sheep) from the same geographic Portuguese region. Additionally, Elliot *et al*. (2014) revealed that many haplotypes are shared between cattle and sheep from Australia, indicating that there is no host selection. Furthermore, present results are also consistent with Bozorgomid *et al*. (2019), who used the mitochondrial NADH gene to demonstrate low levels of gene flow between *Fasciola* species isolated from different hosts in Iran (cattle, sheep, and goats), thus suggesting that differences in host species cannot influence the genetic structure of *F. hepatica*.

The intergenic spacers (ITS1 and ITS2) located between the 18S, 5.8S and 28S rRNA regions generally showed a low level of genetic variability among *F. hepatica* isolates from almost every continent, confirming that - despite some reported differences in restricted geographical localities - this molecular marker is usually monomorphic within each trematode species (Nolan & Cribb, 2005). For this reason, the ITS molecular marker is rarely used in phylogeographic studies of trematodes, but it is useful and effective in the taxonomic attribution of these parasites (Nolan & Cribb, 2005), being a reliable genetic tool for identifying and differentiating species belonging to the genus *Fasciola* (Itagaki *et al*., 2009; Amor *et al*., 2011; Amer *et al*., 2016). For both COI and ITS fragments, some *F. hepatica* isolated from China diverged from the others in the present study, clustering with sequences of *F. gigantica*. Although the occurrence of a new cryptic species for the genus *Fasciola* in China cannot be ruled out, this finding more likely suggests the occurrence of hybrids between these species. More generally, introgressed forms of *Fasciola* are frequently reported in temperate and subtropical regions of eastern Asia based on mitochondrial and nuclear-ribosomal marker identifications (see Le *et al*. (2008) and references therein).

In conclusion, the present study reports two new mitochondrial COI lineages for *F. hepatica* identified in cattle from Algeria and Sardinia. The presence of a low number of COI haplotypes among Algerian samples may be the consequence of the recent introduction of a few founders from Europe and the possible occurrence of a high number of clonal parasites, as already reported for other geographic areas (Beesley *et al*., 2017).

The general low level of genetic variation retrieved for COI and ITS fragments is a frequent genetic pattern of *F. hepatica* (Beesley *et al*., 2017) and may suggest a common worldwide origin for this species. Even considering that the short length or the uniparental inheritance (for COI) of the molecular markers might have partially hindered the actual level of genetic variation, this trend can be also explained by taking into account that *F. hepatica* is a hermaphrodite capable of both cross- and self-fertilisation (although cross-fertilisation is most common), and the occurrence of self-fertilisation may have prompted the loss of genetic diversity in nuclear regions (Cwiklinski *et al*., 2015). Furthermore, it should be also taken into account that the limited number of sequences analysed for Algeria in the present study might have affected outputs thus introducing a bias due to the occurrence of genetic drift among samples. To solve this matter and better understand the level of genetic variation of *F. hepatica* in Algeria, additional nuclear (microsatellites) and mitochondrial (whole genomes) genetic data based on a larger sample set would be necessary to depict higher levels of polymorphism and shed new light on the phylogeographic patterns of this species.

Conflict of Interest

Authors state no conflict of interest.

Acknowledgements

Funding: This research was funded by Fondazione di Sardegna, bando 2022-2023 for the Dipartimento di Scienze Biomediche - UNISS (to Daria Sanna and Fabio Scarpa).

Data Availability Statement: The sequences isolated in the present study are available on GenBank under the accession numbers MT920965 – MT921004 (for COI gene); MZ292402 (for ITS fragment).

Supplementary Materials: Table S1: COI gene whole dataset sampling. The table reports data on the sequences from GenBank, isolated worldwide, which are included in the COI gene dataset. Table S2: PCoA groups. The table reports details on the sequences included in the groups evidenced by PCoA for the COI gene dataset. Table S3: PCoA groups. The table reports details on the sequences included in the groups evidenced by PCoA for the ITS region dataset. Table S4: ITS region whole dataset sampling. The table reports details on the sequences from GenBank, isolated worldwide, that were included in the ITS dataset. Fig. S1: Bayesian phylogenetic tree. The phylogenetic tree analysis is based on a portion of the mitochondrial COI gene. All the nodes of the tree are fully supported by high values of posterior probabilities with only few exceptions. The sequences from Algeria obtained in the present study are indicated with a red font, while the sequences from the island of Sardinia obtained in the present study are indicated with a blue font. The samples codes of the sequences obtained in the present study are reported as in Table 1.

References

Ai, L., Weng, Y.B., Elsheikha, H.M., Zhao, G.H., Alasaad, S., Chen,

J.X., Li, J., Li, H.L., Wang, C.R., Chen, M.X., Lin, R.Q., Zhu, X.Q. (2011): Genetic diversity and relatedness of *Fasciola* spp. isolates from different hosts and geographic regions revealed by analysis of mitochondrial DNA sequences. *Vet Parasitol*, 181: 329 – 334. DOI: 10.1016/j.vetpar.2011.03.057

Aissi, M., Harhoura, K.H., Gaid, S., Hamrioui, B. (2009): Étude préliminaire sur la prévalence de la fasciolose due à *Fasciola hepatica* dans quelques élevages bovins du nord centre algérien (la Mitidja). [Preliminary study on the prevalence of fasciolosis due to *Fasciola hepatica* in some cattle farms in northern central Algeria (Mitidja).] *Bull Soc Pathol Exot*, 102: 177 – 178

Alasaad, S., Huang, C.Q., Li, Q.Y., Granados, J.E., Garcia-Romero, C., Perez, J.M., Zhu, X.Q. (2007): Characterization of *Fasciola* samples from different host species and geographical localities in Spain by sequences of internal transcribed spacers of rDNA. *Parasitol Res*, 101: 1245 – 1250. DOI: 10.1007/s00436-007-0628-2

Amer, S., Elkhatam, A., Zidan, S., Feng, Y., Xiao, L. (2016): Identity of *Fasciola* spp. in sheep in Egypt. *Parasit Vectors*, 9: 623. DOI: 10.1186/s13071-016-1898-2

Amor, N., Farjallah, S., Merella, P., Alagaili, A. N., Mohammed, O.B. (2020): Multilocus approach reveals discordant molecular markers and corridors for gene flow between North African populations of *Fasciola hepatica*. *Vet Parasitol,* 278: 109035. DOI: 10.1016/j.vetpar.2020.109035

Amor, N., Halajian, A., Farjallah, S., Merella, P., Said, K., Slimane, B.B. (2011): Molecular characterization of *Fasciola* spp. from the endemic area of northern Iran based on nuclear ribosomal DNA sequences. *Exp Para*, 128: 196 – 204. DOI: 10.1016/j.exppara.2011.03.011

Ashrafi, K., Valero, M.A., Panova, M., Periago, M.V., Massoud, J., Mas-Coma, S. (2006): Phenotypic analysis of adults of *Fasciola hepatica*, *Fasciola gigantica* and intermediate forms from the endemic region of Gilan. *Iran Parasitol Int*, 55: 249 – 260. DOI: 10.1016/j.parint.2006.06.003

Bandelt, H.J., Forster, P., Rohl, A. (1999): Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, 16: 37 – 48. DOI: 10.1093/oxfordjournals.molbev.a026036

Beesley, N.J., Williams, D.J., Paterson, S., Hodgkinson, J. (2017): *Fasciola hepatica* demonstrates high levels of genetic diversity, a lack of population structure and high gene flow: possible implications for drug resistance. *Int. J. Parasitol.*, 47(1): 11 – 20. DOI: 10.1016/j.ijpara.2016.09.007

Birky, C.W. (2013): Species detection and identification in sexual organisms using population genetic theory and DNA sequences. *PLoS One*, 8: e52544. DOI: 10.1371/journal.pone.0052544

Birky, C.W., Adams, J., Gemmel, M., Perry, J. (2010): Using population genetic theory and DNA sequences for species detection and identification in asexual organisms. *PLoS One*, 5: e10609. DOI: 10.1371/journal.pone.0010609

Bozorgomid, A., Rouhani, S., Harandi, M.f., Ichikawa-Seki, M., Raeghi, S. (2019): Genetic diversity and distribution of *Fasciola hepatica* haplotypes in Iran: Molecular and phylogenetic studies. *Vet Parasitol Reg Stud Reports*, 19: 100359. DOI: 10.1016/j.vprsr.2019.100359

Chaouadi, M., Harhoura, K., Aissi, M., Zait, H., Zenia, S., Tazerouti, F. (2019): A post-mortem study of bovine fasciolosis in the Mitidja (north center of Algeria): Prevalence, risk factors, and comparison of diagnostic methods. *Trop Anim Health Prod*, 51: 2315 – 2321. DOI: 10.1007/s11250-019-01951-w

Chougar, L., Mas-Coma, S., Artigas, P., Harhoura, K., Aissi, M., Agramunt, V.H., Bargues, M.D. (2020): Genetically "pure" *Fasciola gigantica* discovered in Algeria: DNA multimarker characterization, trans-Saharan introduction from a Sahel origin and spreading risk into northwestern Maghreb countries. *Transbound Emerg Dis*, 67: 2190 – 2205. DOI: 10.1111/tbed.13572

Chougar, L., Amor, N., Farjallah, S., Harhoura, K., Aissi, M., Alagaili, A.N., Merella, P. (2019): New insight into genetic variation and haplotype diversity of *Fasciola hepatica* from Algeria. *Parasitol Res*, 118: 1179 – 1192. DOI: 10.1007/s00436-019-06270-5

Cossu, P., Dedola, G.L., Scarpa, F., Sanna, D., Lai, T., Maltagliati, F., Curini-Galletti, M., Casu, M. (2015): Patterns of spatial genetic variation in *Patella ulyssiponensis*: Insights from the western Mediterranean marine ecoregion. *Hydrobiologia*, 755: 39 – 55. DOI: 10.1007/s10750-015-2216-2

Cwiklinski, K., Dalton, J.P., Dufresne, P.J., La Course, J., Williams, D.Jl., Hodgkinson, J., Paterson, S. (2015): The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol*, 16: 71. DOI: 10.1186/s13059-015-0632-2

Cwiklinski, K., Allen, K., Lacourse, J., Williams, D.J., Paterson, S., Hodgkinson, J.E. (2015): Characterization of a novel panel of polymorphic microsatellite loci for the liver fluke, *Fasciola hepatica*, using a next generation sequencing approach. *Infect Genet Evol*, 32: 298 – 304. DOI: 10.1186/s13059-015-0632-2

Drummond, A.J., Rambaut, A. (2007): BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*, 7: e214. DOI: 10.1186/1471-2148-7-214

Elliott, T., Muller, A., Brockwell, Y., Murphy, N., Grillo, V., Toet, H.M., Andersone, G., Sangster, N., Spithill, T.W. (2014): Evidence for high genetic diversity of NAD1 and COX1 mitochondrial haplotypes among triclabendazole resistant and susceptible populations and field isolates of *Fasciola hepatica* (liver fluke) in Australia. *Vet Parasitol*, 200: 90 – 96. DOI: 10.1016/j.vetpar.2013.11.019

Ezard, T., Fujisawa, T., Barraclough, T.G. (2009): SPLITS: SPecies' LImits by Threshold Statistics. R package version 1.0-11, Available at http://R-Forge.R-project.org/projects/splits.

Farjallah, S., Sanna, D., Amor, N., Ben Mehel, B., Piras, M.C., Merella, P., Casu, M., Curini-Galletti, M., Said, K., Garippa, G. (2009): Genetic characterization of *Fasciola hepatica* from Tunisia and Algeria based on mitochondrial and nuclear DNA sequences. *Parasitol Res*, 105: 1617 – 1621. DOI: 10.1007/s00436-009-1601-z

Farjallah, S., Slimane, B., Piras, C.M., Amor, N., Garippa, G., Merella, P. (2013): Molecular characterization of *Fasciola hepat-*

ica from Sardinia based on sequence analysis of genomic and mitochondrial gene markers. *Exp Parasitol*, 135: 471 – 478. DOI: 10.1016/j.exppara.2013.08.006

Fu, Y.X. (1997): Statistical tests of neutrality of mutations against population growth, hitchhiking and backgroud selection. *Genetics*, 147: 915 – 925. DOI: 10.1093/genetics/147.2.915

Gelman, A., Rubin, D.B. (1992): Inference from iterative simulation using multiple sequences. *Stat Sci*, 7: 457 – 472. DOI: 10.1214/ss/1177011136

Hewitt, G. (2000): The genetic legacy of the Quarternary ice ages. *Nature*, 405: 907 – 913. DOI: 10.1038/35016000

Itagaki, T., Kikawa, M., Sakaguchi, K., Shimo, J., Terasaki, K., Shibahara, T., Fukuda, K. (2005): Genetic characterization of parthenogenic *Fasciola* sp. in Japan on the basis of the sequences of ribosomal and mitochondrial DNA. *Parasitol*, 131: 679 – 685. DOI: 10.1017/S0031182005008292

Itagaki, T., Sakaguchi, K., Terasaki, K., Sasaki, O., Yoshihara, S., Van Dung, T. (2009): Occurrence of spermic diploid and aspermic triploid forms of *Fasciola* in Vietnam and their molecular characterization based on nuclear and mitochondrial DNA. *Parasitol Int*, 58: 81 – 85. DOI: 10.1016/j.parint.2008.11.003

Kimura, M. (1980): A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*, 16: 111 – 120. DOI: 10.1007/BF01731581

Laatamna, A.E., Tashiro, M., Zokbi, Z., Chibout, Y., Megrane, S., Mebarka, F., Ichikawa-Seki, M. (2021): Molecular characterization and phylogenetic analysis of *Fasciola hepatica* from high-plateau and steppe areas in Algeria. *Parasitol Int*, 80: 102234. DOI: 10.1016/j.parint.2020.102234

Le, T., De, N., Agatsuma, T., Thi, Nguyen, T., Nguyen, Q., Mcmanus, D., Blair, D. (2008): Human fascioliasis and the presence of hybrid/introgressed forms of *Fasciola hepatica* and *Fasciola gigantica* in Vietnam. *Int J Parasitol*, 38: 725 – 730. DOI: 10.1016/j.ijpara.2007.10.003

Le, T.H., Blair, D., Agatsuma, T., Humair, P.F., Campbell, N.J., Iwagami, M., Littlewood, D.T., Peacock, B., Johnston, D.A., Bartley, J., Rollinson, D., Herniou, E.A., Zarlenga, D.S., Mcmanus, D.P. (2000): Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. *Mol Biol Evol*, 17: 1123 – 1125. DOI: 10.1093/oxfordjournals.molbev.a026393

Librado, P., Rozas, J. (2009): DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451 – 1452. DOI: 10.1093/bioinformatics/btp187

Lotfy, W.M., Brant, S.V., Dejong, R.J., Le, T.H., Demiaszkiewicz, A., Rajapakse, R.P., Perera, V.B., Laursen, J.R., Loker, E.S. (2008): Evolutionary origins, diversification, and biogeography of Liver Flukes. *Am J Trop Med Hyg*, 79: 248 – 255. DOI: 10.4269/ajtmh.2008.79.248

Luton, K., Walker, D., Blair, D. (1992): Comparisons of ribosomal internal transcribed spacers from two congeneric species of flukes (Platyhelminthes: Trematoda: Digenea). *Mol Biochem Parasitol*, 56: 323 – 327. DOI: 10.1016/0166-6851(92)90181-I

Martinez-Valladares, M., Rojo-Vazquez, F.A. (2014): Intraspecific mitochondrial DNA variation of *Fasciola hepatica* eggs from sheep with different level of anthelmintic resistance. *Parasitol Res*, 113: 2733 – 2741. DOI: 10.1007/s00436-014-3934-5

Mas-Coma, S., Bargues, M.D., Valero, M.A. (2005): Fascioliasis and other plant-borne trematode zoonoses. *Int J Parasitol*, 35: 1255 – 1278. DOI: 10.1016/j.ijpara.2005.07.010

Mas-Coma, S., Valero, M.A., Bargues, M.D. (2009): *Fasciola*, lymnaeids and human fascioliasis, with a global overview on disease transmission, epidemiology, evolutionary genetics, molecular epidemiology and control. *Adv Parasitol*, 69: 41 – 146. DOI: 10.1016/S0065-308X(09)69002-3

Massot, M., Senouci-Horr, K. (1983): Étude de la répartition de Lymnaea truncatola dans le nord-ouest algérien et de sa réceptivité a *Fasciola hepatica*. [Study of the distribution of *Lymnaea truncatola* in northwestern Algeria and its receptivity to *Fasciola hepatica*.] *Ann. Parasitol. Hum. Comp.*, 58: 19 – 25. DOI: 10.1051/parasite/1983581019

Mehmood, K., Hui, Z., Sabir, A.J., Abbas, R.Z., Ijaz, M., Durrani, A.Z., Saleem, M.H., Rehman, M.U., Iqbal, M.K., Wang, Y., Ahmad, H.I., Abbas, T., Hussain, R., Ghori, M.T., Ali, S., Khan, A.U., Li, J. (2017): A review on epidemiology, global prevalence and economical losses of fasciolosis in ruminants. *Microb Pathog*, 109: 253 – 262. DOI: 10.1016/j.micpath.2017.06.006.

Mekroud, A., Benakhla, A., Vignoles, P., Rondelaud, D., Dreyfus, J. (2004): Preliminary studies on the prevalences of natural fasciolosis in cattle, sheep, and the host snail (*Galba truncatula*) in north-eastern Algeria. *Parasitol Res*, 92: 502 – 505. DOI: 10.1007/s00436-004-1072-1

Miller, M.A., Pfeiffer, W., Schwartz, T. (2010): Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the 2010 gateway computing environments workshop (GCE), New Orleans, LA, USA, 14 November 2010*.

Mucheka, V.T., Lamb, J.M., Pfukenyi, D.M., Mukaratirwa, S. (2015): DNA sequence analyses reveal co-occurrence of novel genotypes of *Fasciola gigantica* (Fasciolidae, Echinostomida, Trematoda) with *F. hepatica* in South Africa and Zimbabwe. *Vet Parasitol*, 214: 144 – 151. DOI: 10.1016/j.vetpar.2015.09.024

Nolan, M.J., Cribb, T.H. (2005): The use and implications of ribosomal DNA sequencing for the discrimination of digenean species. *Adv Parasitol*, 60: 102 – 163. DOI: 10.1016/S0065-308X(05)60002-4

Ouchene-Khelifi, N.A., Ouchene, N., Dahmani, H., Dahmani, A., Said, M., Douifi, M. (2018): Fasciolosis due to *Fasciola hepatica* in ruminants in abattoirs and its economic impact in two regions in Algeria. *Trop Biomed*, 35: 181 – 187

Paradis, E., Claude, J., Strimmer, K. (2004): APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20: 289 – 290. DOI: 10.1093/bioinformatics/btg412

Peakall, R., Smouse, P.E. (2012): GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research - an

update. *Bioinformatics*, 28: 2537 – 2539. DOI: 10.1093/bioinformatics/bts460

Periago, M., Valero, M., Sayed, M., Ashrafi, K., Wakeel, A., Mohamed, M., Desquesnes, M., Curtale, F., Mas-Coma, S. (2008): First phyenotypic description of *Fasciola hepatica*/*Fasciola gigantica* intermediate forms from human endemic area of the Nile Delta Egypt. *Infect Genet Evol*, 8: 51 – 58. DOI: 10.1016/j.meegid.2007.10.001

Periago, M.V., Valero, M.A., Panova, M., Mas-Coma, S. (2006): Phenotypic comparison of allopatric populations of *Fasciola hepatica* and *Fasciola gigantica* from European and African bovines using a computer image analysis system (CIAS). *Parasitol Res*, 99: 368 – 378. DOI: 10.1007/s00436-006-0174-3

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P. (2006): Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol*, 55: 595 – 609. DOI: 10.1080/10635150600852011

Posada, D. (2008): jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol*, 25, 1253 – 1256. DOI: 10.1093/molbev/msn083

Puillandre, N., Brouillet, S., Achaz, G. (2020): ASAP: assemble species by automatic partitioning. *Mol Ecol Resour*, 21: 609 – 620. DOI: 10.1111/1755-0998.13281

Robinson, M.W., Dalton, J.P. (2009): Zoonotic helminth infections with particular emphasis on fasciolosis and other trematodiases. *Philos Trans R Soc Lond B Biol Sci*, 364: 2763 – 2776. DOI: 10.1098/rstb.2009.0089

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P. (2012): MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*, 61: 539 – 542. DOI: 10.1093/sysbio/sys029

Saitou, N., Nei, M. (1987): The neighbor-joining method: A new method for reconstruction of phylogenetic trees. *Mol Biol Evol*, 4: 406 – 425

Santos, R. (2012): *Genetic characterization of Portuguese* Fasciola hepatica *isolates.* Master's degree in molecular, Genetics and Biomedicine. Universidade Nova de Lisboa, Portugal, pp. 80

Scarpa, F., Sanna, D., Cossu, P., Lai, T., Casu, M., Curini-Galletti, M. (2019a): How to achieve internal fertilization without a vagina: the study case of the genus *Archilina* Ax, 1959 (Platyhelminthes, Proseriata) from Canary Islands. *Mar Biodivers*, 49: 2057 – 2073. DOI: 10.1007/s12526-018-0890-9

Scarpa, F., Cossu, P., Sanna, D., Lai, T., Casu, M., Curini-Galletti, M. (2019b): New insights on the genus *Otoplana* Du Plessis, 1889 (Platyhelminthes: Proseriata), with description of two new species from the Canary Islands. *Mar Biodivers*, 49: 2075 – 2087. DOI: 10.1007%2Fs12526-017-0785-1

Scarpa, F., Sanna, D., Cossu, P., Lai, T., Curini-Galletti, M., Casu, M. (2018): A molecular approach to the reconstruction of the pre-Lessepsian fauna of the Isthmus of Suez: the case of the interstitial flatworm *Monocelis lineata* sensu lato (Platyhelminthes: Proseriata). *J Exp Mar Biol Ecol*, 502: 174 – 181. DOI: 10.1016/j.jembe.2017.08.011

Semyenova, S.K., Morozova, E.V., Chrisanfova, G.G., Gorokhov, V.V., Arkhipov, I.A., Moskvin, A.S., Movsessyan, S.O., Ryskov, A.P. (2006): Genetic differentiation in eastern European and western Asian populations of the liver fluke, *Fasciola hepatica*, as revealed by mitochondrial nad1 and cox1 genes. *J Parasitol*, 92: 525 – 530. DOI: 10.1645/GE-673R.1

Shoriki, T., Ichikawa-Seki, M., Suganuma, K., Naito, I., Hayashi, K., Nakao, M., Aita, J., Mohanta, U.K., Inoue, N., Murakami, K., Itagaki, T. (2016): Novel methods for the molecular discrimination of *Fasciola* spp. on the basis of nuclear protein-coding genes. *Parasitol Int*, 65: 180 – 183. DOI: 10.1016/j.parint.2015.12.002

Sievers, F., Higgins, D.G. (2014): Clustal Omega. *Curr Protoc Bioinformatics*, 48: 1.25.1 – 1.25.33. DOI: 10.1002/0471250953.bi0313s48

Simsek, S., Utuk, A.E., Balkaya, I. (2011): Molecular differentiation of Turkey cattle isolates of *Fasciola hepatica* and *Fasciola gigantica*. *Helmintologia*, 48(1): 3 – 7. DOI: 10.2478/s11687-011-0001-y

Soriano A.R., Ramos-Onsins S.E., Rozas J., Calafell F., Navarro A. (2008): Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics,* 179: 555 – 567. DOI: 10.1534/genetics.107.083006

Tajima, F. (1989): Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585 – 589

Thang, T.N., Vázquez-Prieto, S., Vilas, R., Paniagua, E., Ubeira, F.M., Ichikawa-Seki, M. (2020): Genetic diversity of *Fasciola hepatica* in Spain and Peru. *Parasitol Int*, 76: 102100. DOI: 10.1016/j.parint.2020.102100

Valero, M.A., Bargues, M.D., Calderon, L., Artigas, P., Mas-Coma, S. (2018): First phenotypic and genotypic description of *Fasciola hepatica* infecting highland cattle in the state of Mexico. *Infect Genet Evol*, 64: 231 – 240. DOI: 10.1016/j.meegid.2018.06.032

Walker, S.M., Prodohl, P.A., Fletcher, H.L., Hanna, R.E., Kantzoura, V., Hoey, E.M., Trudgett, A. (2007): Evidence for multiple mitochondrial lineages of *Fasciola hepatica* (liver fluke) within infrapopulations from cattle and sheep. *Parasitol Res*, 101: 117 – 125. DOI: 10.1007/s00436-006-0440-4

Supplementary Materials are available online at https://doi.org/10.2478/helm-2022-0021.

Canis lupus familiaris

This work was conducted within the biennial project entitled: "Camps of Mongolian nomads – an ethnoarchaeological perspective" which was carried out thanks to the collaboration between Associazione Italiana di Etnoarcheologia (Rome, Italy) and the Department of Biomedical Sciences of the University of Sassari (Sassari, Italy). This collaboration resulted in the publication of one book chapter entitled: Mitochondrial DNA Variation among Dogs of Mongolian, Tuvinian and Altaic Nomads (ISBN 978-1-80327-354-9).

The study delves into the genetic variability of dogs in nomadic camps across Mongolia and the Republics of Tuva and Altai in the Russian Federation. The findings reveal distinctive signs of expanding populations with diverse origins. This pattern suggests genetic exchanges among dogs from various camps, likely facilitated by the nomadic communities. Furthermore, results evidenced the possible occurrence of hybrids between *Canis lupus familiaris* and *Canis lupus*.

Research article

1.2 Mitochondrial DNA Variation among Dogs of Mongolian, Tuvinian and Altaic Nomads

Daria Sanna[1], Ilenia Azzena[1,2], Piero Cossu[2], Fabio Scarpa[1], Massimo Scandura[2], Marco Apollonio[2], Francesca Lugli[3], Paolo Francalacci[4], Paolo Mereu[1], Marco Casu[2]

[1] Dipartimento di Scienze Biomediche, University of Sassari, Viale San Pietro 43/b, 07100 Sassari, Italy, darsanna@uniss.it, fscarpa@uniss.it, pmereu@uniss.it

[2] Dipartimento di Medicina Veterinaria, University of Sassari, Via Vienna 2, 07100 Sassari, Italy, iazzena@uniss.it, picossu@uniss.it, scandura@uniss.it, marcoapo@uniss.it, marcasu@uniss.it

[3] Associazione Italiana di Etnoarcheologia, Via Principe Umberto 41, 00185 Roma, Italy, luglifrance@gmail.com

[4] Dipartimento di Scienze della Vita e dell'Ambiente, University of Cagliari, Via Ing. Tommaso Fiorelli 1, 09126 Cagliari, Italy, paolo.francalacci@unica.it

Corresponding author: Daria Sanna·darsanna@uniss.it

Abstract

Dogs originated from the domestication of Eurasian grey wolves. From a genetic viewpoint, they can be grouped into two main clusters: the first is represented by several breeds obtained by artificial selection, whereas the second is of dogs that adapted to a human commensal lifestyle. Here we have provided a molecular survey aimed to infer on the genetic variability of dogs from nomadic camps in Mongolia, and the Republics of Tuva and Altai belonging to the Russian Federation. The results provided evidence of typical marks of expanding populations with multiple origins. Such a scenario could be the result of genetic exchanges among dogs from different camps, that were likely mediated by nomads.

Keywords: *Canis familiaris*, mtDNA, control region, genetic variability, Mongolia, Siberia.

1 Introduction

The dog (*Canis familiaris*) was the first domesticated species, which likely originated in the upper Paleolithic from the domestication of the Eurasian grey wolf (*Canis lupus*) (*i.a.* Shannon *et al.* 2015). From a genetic viewpoint, domestic dogs can be ranked into two main, highly divergent, groups: the first is represented by a large variety of pure breeds obtained by means of human-mediated artificial selection; the second encompasses large and strongly diversified groups of free-ranging animals adapted to a human commensal lifestyle (the so-called village dogs). Genetic data collected worldwide support a single geographical origin for domesticated dogs. In this context, the supposed first centre of domestication is located in Central Asia, as suggested by the highest levels of genetic variation that are generally reported in populations from this region.

Mitochondrial DNA (mtDNA) molecular markers were extensively used to infer on the phylogenetic relationships among canine populations distributed throughout the world. The mtDNA is a separate genome located inside cytoplasmatic organelles (the mitochondria) in all eukaryotic cells (Anderson *et al.* 1981). It is a small circular molecule, which is present in multiple copies per cell and is inherited maternally. Savolainen *et al.* (1997) described the occurrence of two highly informative, hypervariable regions (HVS-I and HVS-II) in the canine mtDNA. Pang *et al.* (2009) used these mtDNA regions to analyse 1,543 dogs spread across the Old World, evidencing the presence of six phylogenetic mitochondrial haplogroups (*i.e.* groups of similar sequences that share a common ancestor), named as clades A-F. Clades A, B and C occur at high frequencies in every canine population, suggesting the hypothesis of a possible common origin of these groups from a single domestication event. Conversely, the clades D, E and F showed a limited geographical dispersal and low frequencies of distribution.

The dogs of nomadic populations that live in areas near to the first centre of wolf domestication are generally poorly influenced by foreign gene flows and might show peculiar genetic traits that deserve to be investigated (Irion *et al.* 2005; Boyko *et al.* 2009; Pedersen *et al.* 2013; Shannon *et al.* 2015). In nomad camps, dogs are fundamental to protect livestock against wolves and predators. Therefore, nomads usually pick up dogs when they are puppies, preferring the bravest cubs with the strongest physical structures and peculiar morphological features, such as specific coat colours (Lugli 2016).

Current Mongolian and Siberian pastoralism can be considered the result of a multi-millenary process which started in prehistoric times. Current nomadism has had to face modernity and its market and social

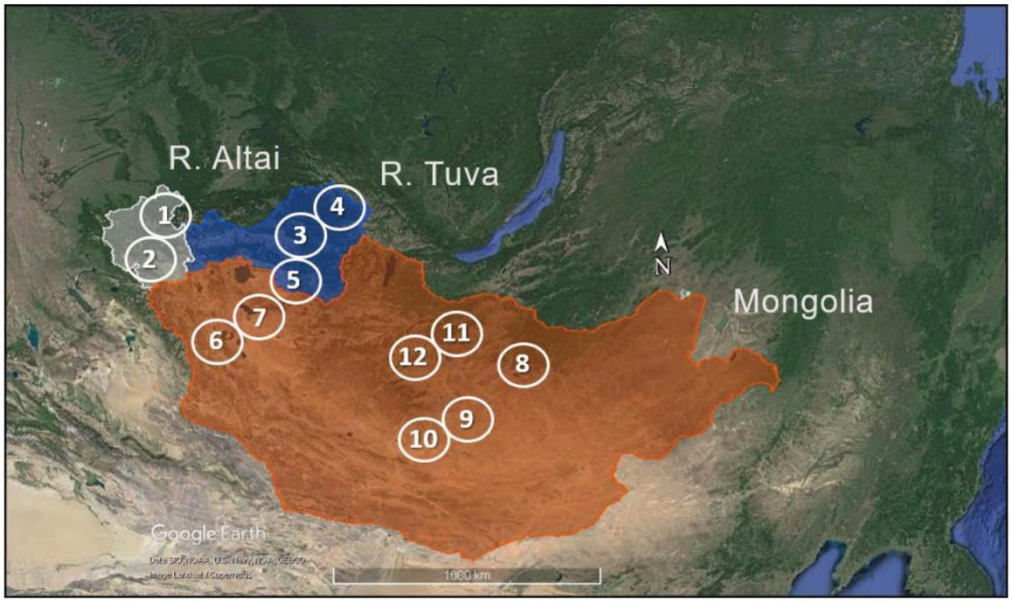


Figure 1. Sampling map showing the countries and the areas where dog hair samples were collected for DNA extraction. Arabic numbers inside white circles indicate the sites where sampling was performed. Republic of Altai:(1) Kurmach Baygol and (2) Kulada. Republic of Tuva: (3) Kyzyl, (4) Systig-Hem, Ador Kezhik, Toora Hem, (5) Erzin region. Mongolia: (6) Khovd aimag, (7) Zavkhan aimag, (8) Ulan Bator, (9) Delgerkhangai (Dundgovi aimag), (10) Övörkhangai aimag, (11) Mogod, Ulziit, Tal Bulag, Tsagaan Khust (Bulgan aimag), (12) Jargalant, Erdenet Mandal, UndurUlaan, BayyanUul (Arkhangai aimag).

models which push young people to abandon their traditional lifestyle.

In such a context, here we have provided a preliminary molecular survey, performed by means of the mitochondrial hypervariable region I (HVS-I) marker, aimed to infer for the first time on the genetic variability and the phylogeographic relationships among village dogs from nomadic camps of rural areas of Mongolia, and the Republics of Tuva and Altai (the Russian Federation).

2 Material and methods

In 2005 the Italian Association for Ethnoarchaeology (AIE) with the sponsorship of the Ministry of Foreign Affairs and International Cooperation – Italy (MFA) started the research project 'The camps of Mongolian nomads: an ethnoarchaeological perspective', which aimed to study the various crucial points of steppe pastoralism in different Mongolian ecosystems.[1]

This project also included research on dogs, which started in Mongolia in 2012. Consequently, the Italian Association for Ethnoarchaeology (AIE) with the sponsorship of Ministry of Foreign Affairs and International Cooperation – Italy MFA, started the mission 'Siberian nomads and their dogs' in 2013, which was conducted in the Republics of Tuva (2013, 2017), and Altai (2014, 2016) (the Russian Federation), and in the Kemerovo region (2015) in collaboration with the Novosibirsk State Conservatory, the Institute of Philology (SB RAS, Novosibirsk), the Institute for Humanities and Kyzyl College of Arts (RT), and the Institute of Altaistics (RA).[2]

The Mongolian and Siberian projects aimed to study and document the relationship between nomads, hunters and dogs in various socio-historical contexts and different ecosystems. Villages and camps of various regions were visited in order to document traditional situations. Hair samples from dogs owned by the families that were studied and interviewed were taken during the missions both in Mongolia and in Siberia.

[1] The project was conducted in collaboration with Tserenkhand (2002–2006) (National University of Ulaanbaatar; Academy of Sciences) and Dulam Sedenjav (after 2007) (National University of Ulaanbaatar). The missions were conducted by F. Lugli and G. Capitini and students and graduated of the National University. The research was conducted in various regions to register differences and variabilities.

[2] The project was conducted in Tuva (2013, 2017), Republic of Altai (2014, 2016), and Kemerovo region (2015) by F. Lugli and G. Sychenko (see Lugli and Sychenko in this volume) in collaboration with Novosibirsk State Conservatory, Institute of Philology (SB RAS, Novosibirsk), Institute for Humanities and Kyzyl College of Arts (RT), Institute of Altaistics (RA).

Table 1. Estimates of genetic diversity obtained for the mitochondrial HVS-I fragment of dog populations analysed here. n: sample size, S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; and π: nucleotide diversity.

Country	Sampling date	n	S	H	*h*	*π*
Mongolia	Total	33	16	14	0.888	0.01328
	Nov. 2013	14	15	9	0.923	0.01279
	Oct. 2014	14	9	6	0.813	0.01137
Republic of Tuva	Oct. 2013	23	18	12	0.917	0.01094
Republic of Altai	Total	23	11	11	0.806	0.01118
	Oct. 2016	15	8	6	0.648	0.00832
	Jul. 2014	8	10	7	0.964	0.01252
Total		79	24	24	0.908	0.01276

Figure 2. Some of the individuals, from Mongolia, the Republic of Altai and the Republic of Tuva, whose hairs were collected to perform non-invasive DNA extractions in the present study (Photos by F. Lugli).

There was usually a preference to take the samples in traditional and isolated contexts. A few samples were also taken from a dog farm in Ulaanbaatar in Mongolia in order to analyse the Mongol Bankhar mastiffs.

The molecular analysis of a 348 base pairs-long HVS-I mitochondrial fragment was performed on 79 dogs from seven sites in Mongolia (33 individuals), and five sites from two states of the Russian Federation, being Republic of Tuva (23 individuals from two sites) and Republic of Altai (23 individuals from three sites) (Figure 1 and Table 1 for details). The individuals analysed included not only non-breed dogs, but also representatives of three canine breeds (Laika, Mongol

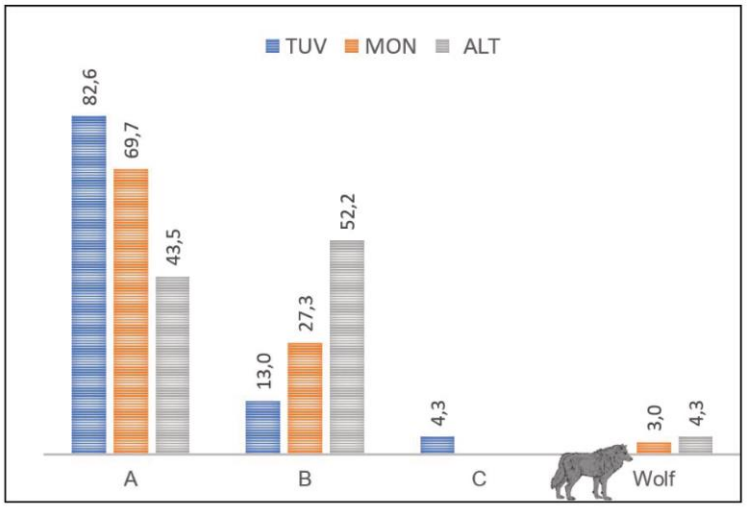


Figure 3. Distribution frequencies of the canine mitochondrial HVS-I haplogroups found for the dogs analysed here. MON: Mongolia; ALT: Republic of Altai; TUV: Republic of Tuva.

Bankhar, and Volkodav) (see Figure 2). Hair samples were collected from individuals with the help of their owners during 2013, 2014, and 2016. Genomic DNA was extracted from hairs by means of the InstaGene™ Matrix (Bio-Rad) according to the manufacturer's protocol. Sample quality and DNA concentration were determined via spectrophotometry using a ND-8000 (NanoDrop Technologies, Thermo Fisher Scientific Inc., Wilmington, DE). PCR was carried out using modified mtDNA HVS-I universal primers (Kocher *et al.* 1989) and sequencing was performed by an external service (Macrogen, The Netherlands). Newly generated sequences were aligned using the BioEdit 7.2.5. software package (Hall, 1999). The genetic variation was assessed estimating the number of polymorphic sites, the number of haplotypes (H), the haplotype diversity (h), and the nucleotide diversity (π) using the software package DnaSP 6.12 (Librado *et al.* Rozas, 2009). Genetic relationships among haplotypes were investigated by a Median-Joining network using the software package Network 10.2.0.0 (www.fluxus-engineering.com).

3 Results and discussion

High levels of genetic variability, resulting in a total of 24 haplotypes (*i.e.* nucleotide sequences corresponding to the same DNA fragment that differ among each other in some informative nucleotide sites), were found at each sampling location (see Table 1 for details). The lowest levels of genetic variation were found in the camps from the North of the Republic of Altai (Turochaksky district). In the present study, the two main worldwide distributed canine mitochondrial haplogroups (A, B) showed distributions of frequencies (Figure 3) that are

Figure 4. Images of the Laika-like individuals whose mitochondrial HVS-I sequence was also found in Siberian wolves. (a) a 3-year-old male from Mongolia; (b) a 1-year-old female from the Republic of Altai (Photos by F. Lugli).

consistent with those generally retrieved for almost all dog populations from the Old World (Savolainen *et al.* 2002 ; Ardaland *et al.* 2011; and references therein), with the haplogroup A showing the highest frequencies. Dogs from the Republic of Altai were the only exception as they showed the highest frequencies of distribution for the haplogroup B, in accordance with a similar trend already reported for Southwest Asian dogs (Ardalan *et al.* 2011). We hypothesise this latter finding may be a consequence of genetic drift mediated by human artificial selection. This evolutionary force may have acted on the Altaic population with repeated introductions of dogs carrying less common haplogroups just by chance. The haplogroup C was found only in dogs from the Republic of Tuva, and the genetic drift may be invoked in this case as well. As briefly outlined above, genetic drift is a stochastic evolutionary force whose strength depends on the population size. When a new population originates from a very small number of individuals (the so-called founders), genetic drift may trigger the loss of genetic variation changing the frequencies of haplotypes. As a consequence, the distribution of haplogroups in the new established populations may diverge from those reported for the original population.

It is noteworthy that a mitochondrial lineage, that is present in Siberian wolves, was found in two morphologically Laika-like individuals, one from Mongolia (a 3-year old male from the district of Bulgan) and one from the Republic of Altai (a 1-year-old female from the district of Kurmach Baygol) (Figure 4). Such a finding may be the result of past accidental domestications of wolf females or cubs in nomad camps. Indeed, although nomads usually prefer to select dogs directly from their canine families, they do not always follow this choice criterion and puppies can also be rescued from stray mothers or lost adults can be adopted (Lugli 2016).

However, it should be taken into account that the uncommon mitochondrial lineage found in these two individuals may also belong to the mitochondrial canine clade D, whose distribution is restricted to North Europe, Siberia, Southwest Asia and the Mediterranean Sea (Angleby and Savolainen 2005; Pang *et al.* 2009). Some sub-haplogroups of this clade are the products of a dog-wolf cross-breeding, rather than of independent domestication of wolves (Ardalan *et al.* 2011).

The network analysis (see Figure 5a and its legend for more details) evidenced that many sequences were exclusively found in single individuals, probably due to very recent multiple introductions of new dogs. The occurrence of star-like configurations in the plot suggests the lack of genetic divergence among areas, along with the occurrence of many founder effects. Indeed, here the network star-like configurations are represented by a common central haplotype, usually shared among individuals from many regions, that is surrounded by many lesser-frequent (and private to single individuals) haplotypes differing by a few mutations. The most common haplotypes in the network likely correspond to sequences belonging to the first dogs introduced in the nomads'camps, which had the opportunity to breed extensively. Such findings are consistent with the general trend of genetic homogeneity worldwide reported for dog populations (*i.a.* Pang *et al.* 2009).

A less frequent and highly divergent haplotype (Figure 5b) was found exclusively among dogs from a Mongol Bankhar breeding farm in Mongolia: we hypothesised the occurrence of a mtDNA matrilinear relationship among all individuals born in the farm, which likely descend from a group of related females.

The network analysis also evidenced that two Volkodav dogs from the Republic of Tuva show private-owned haplotypes (not shared with the other breeds). This finding suggests that the genetic divergence reported for these dogs is likely consistent with the different history and geographic origin of their breed.

4 Conclusions

The study of the genetic makeup of village dogs and central Asian local breeds represents an important step to depict the complex evolutionary history of these animals (Shannon *et al.* 2015).

In such a context, we have reported the first and preliminary molecular inference on dogs from the mountains of Mongolia and from the Russian Republics of Altai and Tuva. The results pointed out high levels of genetic divergence at each sampling site, and a lack of geographic differentiation among regions. Our findings reflect the typical marks of expanding populations with multiple origins. We hypothesise that such a scenario could be the result of repeated genetic exchanges among dogs from different nomads'camps, which were likely mediated by human activities. Indeed, Mongolian nomads usually move to villages to pick their dogs (Lugli 2016). Accordingly, the haplotype distribution frequencies and the founder effects evidenced by the network analysis, further account for the signature of artificial selection, which drastically skewed the genetic diversity within village dogs and local breeds such as the Mongol Bankhar mastiff. Within this framework, it should also be considered that the mitochondrial genetic variability reported for village dogs from nomadic camps may be sex-biased because of the maternal inheritance of the mitochondrial molecular marker here used. In the present study,

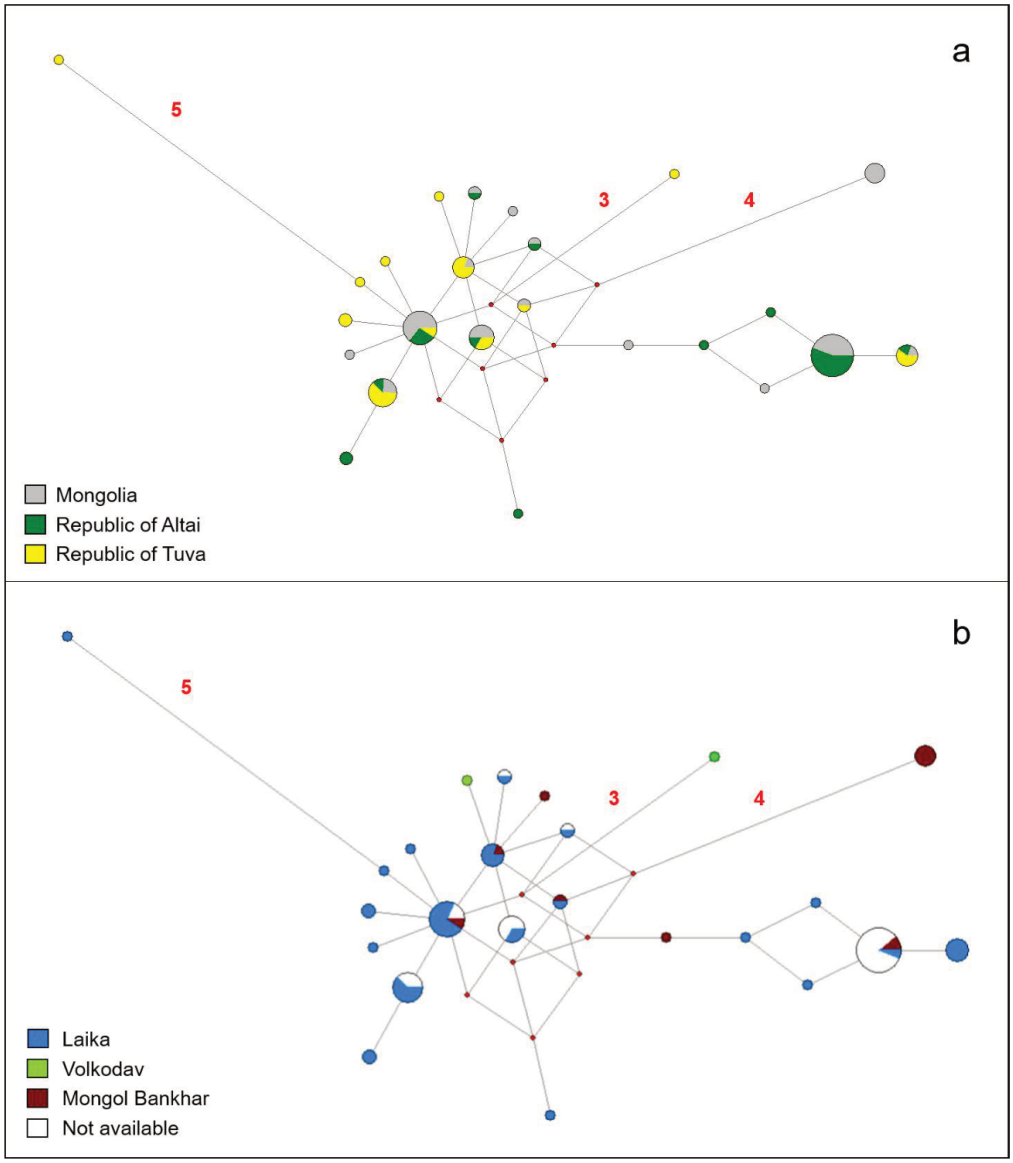


Figure 5. Median-Joining networks showing the phylogenetic relationships occurring among the mitochondrial haplotypes found in the present study. Haplotypes are represented by the circular spots on the graphic; the diameter of spots is proportional to the number of individuals that share the haplotype; the length of branches in the graphic are proportional to the number of nucleotide mutations occurring between the two haplotypes at the edges of the branch. The spots are coloured according to the characteristics of the individuals that share the haplotype. Little red spots in the graphic represent the median vectors that are crucial sequences likely existing in nature but not found among the specimens collected for the present study. In the present graphic all haplotypes diverge from each other for a single mutation except for the cases where Arabic red numbers on network branches indicate the occurrence of a higher number of point mutations. (a) the spots on the network are coloured according to the geographic origin of dogs; (b) the spots on the network are coloured according to the breed of dogs. Please note that it was not possible to identify the breed for all individuals.

the general trend of genetic homogeneity evidenced among areas may be the result of the sex-based choice criterion that nomads apply to pick their dogs. For example, 99% of Mongolian nomads' families decide to only have male dogs because they consider females too difficult to manage during their oestrus cycle (see Lugli 2016 for more details). As a consequence, the common ancestors of the dogs considered in the present study could be represented by a reduced number of females that were likely used as breeders. This choice might has decreased the effective population size (*i.e.* the number of mating individuals that contribute to the genetic pool of the next generation) and increased the genetic homogeneity evidenced by the mitochondrial DNA, which is matrilineally transmitted to the offspring.

The main mitochondrial haplogroups found in the present study among dog populations suggest a recent origin, common to other European canine populations.

In the future, the analysis of a larger number of individuals from further Mongolian, Altaic and Tuvinian sites, will shed further light on the evolutionary processes that might have shaped the genetic patterns of dog populations living in these Asiatic regions.

References

Anderson, S., A.T. Bankler, B.G. Barrel, M.H.L. de Bruijn, A.R. Cuolson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden and I.G. Young 1981. Sequence and organization of the human mitochondrial genome *Nature* 290: 457–464.

Angleby, H. and P. Savolainen 2005. Forensic informativity of domestic dog mtDNA control region sequences. *Forensic Science International* 154(2–3): 99–110.

Ardalan, A., C.F. Kluetsch, A.B. Zhang, M. Erdogan, M. Uhlén, M. Houshmand, C. Tepeli, S.Z. Miraei Ashtiani and P. Savolainen 2011. Comprehensive study of mtDNA among Southwest Asian dogs contradicts independent domestication of wolf, but implies dog-wolf hybridization. *Ecology and Evolution* 1(3): 373–385.

Boyko, A.R., R.H. Boyko, C.M. Boyko, H.G. Parker, M. Castelhano, L. Corey, J.D. Degenhardt, A. Auton, M. Hedimbi, R. Kityo, E.A. Ostrander, J. Schoenebeck, R.J. Todhunter, P. Jones, and C.D. Bustamante 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. *Proceedings of the National Academy of Sciences* 106(33): 13903–13908.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Irion, D.N., A.L. Schaffer, S. Grant, A. N. Wilton and N.C. Pedersen, 2005. Genetic variation analysis of the Bali street dog using microsatellites. *BMC Genetics* 6: 6.

Kocher, T.D., W.K. Thomas, A. Meyer, S.V. Edwards, S. Pääbo, F. Villablanca and A.C. Wilson 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the USA* 86: 6196 - 6200.

Librado, P. and J. Rozas 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.

Lugli, F. 2016. Mongolian nomads and their dogs, in S. Biagetti and F. Lugli (eds) *The Intangible Elements of Culture in Ethnoarchaeological Research:* 125–139. Springer: Cambridge.

Pang, J-F., C. Kluetsch, X-J. Zou, A-B., Zhang, L.Y. Luo, H. Angleby, A. Ardalan, A., C. Ekström, A. Sköllermo, J. Lundeberg, S. Matsumura, T. Leitner, Y.-P. Zhang and P. Savolainen 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Molecular Biology and Evolution* 26(12): 2849–2864.

Pedersen, N., H. Liu, G. Theilen and B. Sacks 2013. The effects of dog breed development on genetic diversity and the relative influences of performance and conformation breeding. *Journal of Animal Breeding and Genetics* 130(3): 236–248.

Savolainen, P., B. Rosen, A. Holmberg, T. Leitner, M. Uhlen and J. Lundeberg 1997. Sequence analysis of domestic dog mitochondrial DNA for forensic use. *Journal of Forensic Science* 42(4): 593–600.

Savolainen, P., Y.P. Zhang, J. Luo, J. Lundeberg and T. Leitner 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science* 298(5598): 1610–1613.

Shannon. L.M., R.H. Boyko, M. Castelhano, E. Corey, J.J. Hayward, C. McLean, M.E. White, M. Abi Said, B.A. Anita, N.I. Bondjengo, J. Calero, A. Galov, M. Hedimbi, B. Imam, R. Khalap, D. Lally, A. Masta, K.C. Oliveira, L. Pérez, J. Randall, N.M. Tam, F.J. Trujillo-Cornejo, C. Valeriano, N.B. Sutter, R.J. Todhunter, C.D. Bustamante and A.R. Boyko 2015. Dog genetic diversity reveals Central Asian origin. *Proceedings of the National Academy of Sciences* 112.44: 13639–13644.