UNISS
UNIVERSITÀ
DEGLI STUDI
DI SASSARI

University of Sassari

Department of Biomedical Sciences

Ph.D. School in Life Sciences and Biotechnologies

XXXVIII cycle

Population genetics across aquatic species: evolutionary history, phylogeography, and taxonomic complexity

PhD candidate

Chiara Locci

Supervisors

Prof. Daria Sanna

Department of Biomedical Sciences

Prof. Marco Casu

Department of Veterinary Medicine

Academic Year
2024/2025

"There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved."

Charles Darwin, *The Origin of Species*, 1872

Index

Preamble

This PhD thesis was carried out at the Department of Biomedical Sciences and the Department of Veterinary Medicine of the University of Sassari.

My PhD fellowship was funded under the DM 351/2022 by the National Recovery and Resilience Plan (PNRR).

Abstract

This doctoral thesis investigates genetic diversity, population structure, and evolutionary dynamics of aquatic species across marine and freshwater systems through population genetic approaches. By integrating phylogenetics, phylogeography, and molecular taxonomy, four case studies were conducted on native and invasive species under distinct ecological contexts. The results provide insights into the complex interplay among evolutionary history, invasion dynamics, taxonomic complexity, and population connectivity and dispersal, highlighting the role of historical and contemporary factors in shaping genetic differentiation. Particular attention is given to processes including demographic decline, habitat fragmentation, and biological invasions, which represent major challenges for aquatic biodiversity. This research contributes to the understanding of evolutionary resilience, diversification, and adaptation in aquatic environments and highlights the role of population genetics as a unifying framework for interpreting evolutionary processes. Additionally, the findings offer relevant information to support biodiversity monitoring, management and conservation strategies.

Riassunto

Questa tesi di dottorato ha come obiettivo principale lo studio della variabilità genetica, della struttura di popolazione e delle dinamiche evolutive di specie acquatiche, sia marine che di acqua dolce, mediante l'utilizzo di diversi approcci di genetica di popolazione. Integrando la filogenetica, filogeografia e tassonomia molecolare, sono stati condotti quattro casi di studio su specie native e invasive in differenti contesti ecologici. I risultati forniscono approfondimenti sull'articolata interazione tra storia evolutiva, dinamiche di invasione, complessità tassonomica, connettività e dispersione, evidenziando il ruolo di fattori storici e contemporanei nel modellare il differenziamento genetico. Particolare attenzione è stata dedicata a processi quali il declino demografico, la frammentazione degli habitat e le invasioni biologiche, che rappresentano sfide significative per la biodiversità acquatica. Questa ricerca contribuisce alla comprensione delle dinamiche di resilienza, della diversificazione e dell'adattamento negli ecosistemi acquatici, sottolineando il ruolo della genetica di popolazione come filo conduttore nell'interpretazione dei processi evolutivi. Inoltre, i risultati forniscono informazioni rilevanti a supporto delle strategie di monitoraggio, gestione e conservazione della biodiversità.

General introduction

1. Population genetics

Population genetics is a central discipline of evolutionary biology that investigates the genetic composition of populations and how it changes under the influence of evolutionary processes (Okasha, 2024). At its foundation lies the connection between evolution and natural selection: evolution can be defined as a change in the genetic composition of populations across generations, while natural selection represents one of the several mechanisms capable of driving such change. Alongside selection, processes including mutation, migration, genetic drift, and non-random mating can shape the genetic structure of populations (Okasha, 2024; Charlesworth & Charlesworth, 2017).

Genetic diversity represents the basis upon which evolutionary change occurs, and it enables species to evolve and adapt in response to environmental and climatic shifts, emerging diseases, competitors, and predators (Frankham et al., 1996, 1997; Montgomery, 2000; Frankham et al., 2004, 2010). A reduction in genetic diversity decreases a population's adaptive capacity and is often associated with reduced reproductive success and survival, referred to as reduced fitness (Frankham et al., 2004, 2010). Given its fundamental role in evolutionary resilience, genetic diversity is a key parameter evaluated in population genetics. It refers to the variation in DNA sequences within and among populations and species. This diversity can influence functional, morphological, or behavioural differences among individuals (Frankham et al., 2004, 2010).

The discipline emerged in the early 20th century through the foundational contributions of Fisher (1930), Wright (1931), and Haldane (1930), who independently developed mathematical models describing the behaviour of genes in populations governed by Mendelian inheritance and subject to Darwinian natural selection (Okasha, 2024). Their work demonstrated that Mendelian genetics and natural selection are not only compatible but essential for understanding evolutionary change (Okasha, 2024).

A primary aim of population genetics is to understand the origin, maintenance, and distribution of genetic variation within populations, and to explain how this variation gives rise to genetic differentiation among populations over space and time. This objective relies on the integration of theoretical models with empirical data obtained from the analyses of evolutionary processes in both natural and artificial populations (Charlesworth & Smith, 1983). Because evolutionary change depends on the existence of genetic variation within populations, quantifying such variation is crucial in population genetic research (Lewontin, 1974). While early studies relied largely on observable phenotypic traits, advances in molecular genetics made it possible to directly characterise variation at the level of genes and DNA sequences (Charlesworth & Smith, 1983).

A major shift in the discipline occurred from the 1960s onwards, when molecular approaches enabled the direct measurement of genetic variation in natural populations. In this context, the application of protein electrophoresis revealed levels of genetic polymorphism that were unexpected at that time, prompting renewed debate about the roles of random genetic drift and natural selection in shaping genetic variation (Kimura & Crow, 1964). These empirical findings coincided with fundamental theoretical developments, such as the formulation of the neutral theory of molecular evolution (Kimura, 1983; Charlesworth & Smith, 1983; Charlesworth & Charlesworth, 2017).

In subsequent decades, the introduction of increasingly informative molecular markers and DNA sequencing technologies further expanded the scope of population genetics. These advances strengthened the interaction between theory and empirical data, consolidating the discipline as a powerful tool for reconstructing demographic histories, interpreting evolutionary patterns, analysing population structure, and estimating gene flow (Moritz, 1995; Charlesworth & Charlesworth, 2017).

1.1 Conservation genetics

Conservation genetics builds upon the theoretical foundations of evolutionary and population genetics established by Wright, Fisher, Haldane, and further developed by Dobzhansky, Crow, and Kimura (e.g., Fisher, 1930; Haldane, 1930; Wright, 1931;

Dobzhansky, 1951; Crow, 1961; Kimura, 1969). It integrates molecular genetic concepts with ecological principles to understand and mitigate threats to the species persistence (O'Brien, 1994).

Conservation genetics is an applied subdiscipline that uses genetic theory and methodologies to support the preservation of biodiversity. In this context, biodiversity encompasses the variety of ecosystems, species, populations, as well as the genetic diversity both among and within these populations (Frankham et al., 2004, 2010).

The conservation of genetic diversity is recognised as one of the three fundamental levels of biodiversity and as a global conservation priority by the World Conservation Union (IUCN), the leading international authority on biodiversity conservation. The primary goals of conservation genetics include the genetic management of small or declining populations to maintain their genetic diversity and limit its loss, the clarification of taxonomic uncertainties, the identification of biologically relevant management units, and the application of molecular tools to support conservation strategies aimed at preserving genetic variation (Hedrick & Miller, 1992; Frankham et al., 2004, 2010).

The theoretical background developed in the following sections provides the basis for the empirical chapters of this doctoral thesis. The research primarily focuses on the application of population genetics to aquatic organisms, reflecting the broader scope of population genetics as a discipline able to investigate a wide range of biological contexts, from species on the brink of extinction to populations affected by climate change, habitat loss, fragmentation, biological invasions, and emerging diseases. Across these different case studies, population genetics is applied as a unifying conceptual and methodological framework that integrates approaches, such as phylogenetics, phylogeography and molecular taxonomy to explore the mechanisms shaping genetic variation, population connectivity, and long-term evolutionary viability, with the ultimate aim of informing monitoring, management, and conservation strategies.

2. Fundamentals of population genetics

2.1 Genetic diversity: measures and quantification

Quantifying genetic diversity is an essential stage in population genetics because it provides direct information on the evolutionary potential and health status of populations. Genetic diversity can be described using various metrics, each capturing different aspects of the underlying variation (Frankham et al., 2004, 2010).

➢ Metrics of genetic diversity

Heterozygosity (H_o, H_e): reflects the level of genetic diversity at a single locus. The observed heterozygosity (H_o) is the proportion of the individuals that are heterozygous at a given locus, whereas the expected heterozygosity (H_e) represents the probability that two alleles randomly chosen from the population are different (Frankham et al., 2004, 2010).

Polymorphism: is the proportion of loci that are variable within a population (Frankham et al., 2010).

Allelic diversity (A): is the number of alleles per locus, averaged across loci (Frankham et al., 2004, 2010).

Haplotype diversity (h): is the probability that two randomly sampled haplotypes are different (Frankham et al., 2004, 2010).

Nucleotide diversity (π): is the average number of nucleotide differences per site between two randomly selected sequences (Frankham et al., 2004, 2010).

Fixation index (F): the Wright's fixation index or inbreeding coefficient represents the probability that two alleles at a locus in an individual are identical by descent; it quantifies heterozygosity reduction due to non-random mating (Wright, 1969; Frankham et al., 2004, 2010).

➢ Molecular markers for measuring genetic diversity

The accurate evaluation of genetic diversity relies on distinguishing genotypes that differ at loci relevant to evolutionary and conservation processes. Because such differences are not always detectable at the phenotypic level, molecular markers are employed as indirect indicators of allelic variation (Schlotterer, 2004). These tools are widely used across several biological disciplines, including conservation genetics, phylogenetics, gene mapping, forensics, and parentage analysis (Schlotterer, 2004). Advances in molecular biology have progressively improved the resolution of genetic assays, giving rise to performative marker systems capable of detecting fine-scale genetic variation (Schlotterer, 2004; Le Roux & Wieczorek, 2009). Patterns of genetic diversity reflect underlying demographic, migration, and evolutionary processes. Because different classes of genetic markers vary in their mutation rates, modes of inheritance and levels of polymorphism, marker selection is a crucial step in any population and conservation genetic studies (Baverstock & Moritz, 1996; Cavalli-Sforza, 1998; Sunnucks, 2000; Le Roux & Wieczorek, 2009).

The first molecular markers employed to quantify genetic variation were allozymes, introduced in the 1970s. Allozymes consist of enzyme variants separated through gel electrophoresis, which enable the detection of allelic differences based on shifts in protein mobility (Frankham et al., 2004; Schlotterer, 2004; Le Roux & Wieczorek, 2009; Frankham et al., 2010). Despite their foundational role, allozymes identify only protein-level differences representing a limited portion of underlying DNA variation, often underestimating true levels of genetic diversity (Frankham et al., 2004, 2010). Their application is also constrained by the requirement for fresh or fresh-frozen samples, which can necessitate invasive sampling procedures. Nevertheless, allozyme analysis remains a cost-effective option when financial resources are limited and fine-scale genetic resolution is not essential (Schlotterer, 2004; Frankham et al., 2004; Le Roux 2009; Frankham et al., 2010).

The development of DNA-based molecular markers represented a major methodological advance, as they characterise genetic variation at the DNA sequence level rather than at proteins level (Schlotterer, 2004). This shift was accelerated by the introduction of polymerase chain reaction (PCR), developed by Kary Mullis in 1983, for which he was awarded the Nobel Prize in Chemistry (Bartlett & Stirling, 2003). PCR allows the amplification of specific genomic regions from small amounts of DNA, enabling non-invasive sampling and fine-scale genetic analyses by using samples of different nature (Schlotterer, 2004; Frankham et al., 2004, 2010).

Among DNA markers, mitochondrial DNA (mtDNA) has become widely used in population and conservation genetic studies. Generally, in animal taxa, mtDNA is maternally inherited, occurs in multiple copies per cells, and does not undergo recombination (Cavalli-Sforza, 1998). Moreover, the mitochondrial genome typically has a high mutation rate, generating substantial sequence variation even within species (Frankham et al., 2004, 2010). These biological features facilitate amplification and sequencing from degraded or low-quality samples and make mtDNA a valuable marker for inferring evolutionary relationships, as well as dispersal and phylogeographic patterns (Frankham et al., 2004, 2010). However, since it represents a single locus that reflects the maternal lineage, mtDNA alone may provide an incomplete picture of population-level processes (Frankham et al., 2004, 2010), therefore, it is often complemented with nuclear markers (Schlotterer, 2004). In some taxa, particularly in large mammals, mitochondrial DNA analyses are complemented by Y chromosome markers, representing the paternal counterpart of mtDNA (Cavalli-Sforza, 1998; Greminger et al., 2010; Kharkov, 2021), useful for tracing male-mediated dispersal patterns and demographic histories (Underhill et al., 2001; Cavalli-Sforza & Feldman, 2003; Greminger et al., 2010; Panigrahi et al., 2025).

Microsatellites, also known as short tandem repeats (STRs) or simple sequence repeats (SSRs), represent one of the most widely used classes of nuclear DNA markers in conservation and population genetic studies. Composed of short tandem repeats (2-6 bp), they exhibit significant levels of polymorphism due to elevated mutation rates and codominant inheritance, allowing robust estimates of allelic diversity, heterozygosity, and population structure (Frankham et al., 2004; Schlotterer 2004; Le Roux & Wieczorek, 2009; Frankham et al., 2010). Their PCR-based amplification enables the use of minimal amounts of DNA, allowing their application for a wide range of sample types, including those obtained through non-invasive sampling (Schlotterer, 2004; Frankham et al., 2004, 2010). Although species-specific primer development may be time-consuming and PCR artefacts, including stutter bands, may challenge allele scoring and interpretation, well-validated microsatellite panels remain highly informative across ecological, evolutionary and conservation applications (Schlotterer, 2004; Frankham et al., 2004, Le Roux & Wieczorek, 2009; Frankham et al., 2010).

Minisatellites, also known as variable number tandem repeats (VNTRs) or DNA fingerprints, were among the earliest DNA markers used in forensic applications and individual identification. They consist of longer repeat units (10-100 bp), often producing highly distinctive multilocus patterns in which individual has a unique "barcode". However, they require high-quality and high-quantity DNA, which makes them unsuitable for degraded or non-invasively collected samples. Consequently, they have been largely replaced by PCR-based markers such as microsatellites (Schlotterer, 2004; Frankham et al., 2004, Le Roux & Wieczorek, 2009; Frankham et al., 2010).

Several additional markers exist but are currently less commonly used in population and conservation genetics. Restriction fragment length polymorphisms (RFLPs) detect DNA polymorphisms based on differences in restriction enzyme recognition sites. However, they are labour-intensive and not cost-effective, with low reproducibility (Le Roux & Wieczorek, 2009). Random amplified polymorphic DNA (RAPDs) markers use short oligonucleotide sequences as primers to amplify uncharacterised genomic regions through PCR. Their dominant inheritance and low reproducibility significantly restrict their application (Schlotterer, 2004; Le Roux & Wieczorek, 2009). Amplified fragment length polymorphisms (AFLPs) combine restriction digestion and selective PCR amplification, enabling the detection of a high number of polymorphisms, without prior genomic information. However, the dominant inheritance reduces their informativeness (Schlotterer, 2004; Le Roux & Wieczorek, 2009). In contrast, single nucleotide polymorphisms (SNPs) represent single-nucleotide variation and occur at high density across the genome. Despite their effectiveness, SNPs discovery and genotyping typically require next-generation sequencing (NGS) approaches (e.g., whole genome sequencing), making their development time and cost intensive (Schlotterer, 2004).

The rise of NGS technologies is shifting population genetics towards genome-wide analyses. Traditional markers, particularly mtDNA and microsatellites, have been highly efficient, cost-effective, and pivotal in this field for decades, reducing the incentive to adopt genomic approaches in routine studies (Kohn et al., 2006). Nevertheless, modern genomic tools allow the simultaneous genotyping of thousands of loci, potentially offering finer reconstruction of demographic histories and assessment of population structure and genetic variability (Ryder, 2005; Kohn et al., 2006; Allendorf et al., 2010). Genome-wide approaches have also facilitated the investigation of functional variation. For instance, microarray technology has enabled large-scale analyses of gene-expression within and among species, providing insights into evolutionary processes at both structural and regulatory level (Frankham et al., 2004; Ranz & Machado, 2006; Whitehead & Crawford, 2006; Frankham et al., 2010). Even though the application of genomic approaches has so far been restricted to a limited number of species with available genomic resources, the rapid expansion of sequencing technologies indicates that genomic data will become increasingly accessible and broadly integrated into future population and conservation research (Allendorf et al., 2010).

2.2 Origin and persistence of genetic diversity

Genetic variation is the fundamental substrate upon which evolutionary processes act, and its origin and persistence determine the adaptive potential of populations (Frankham, 1996; Bohonak, 1999). Understanding how variation arises, is maintained, lost, and regenerated through time is central to evolutionary biology, population, and conservation genetics. In natural populations, levels of genetic diversity reflect the balance among processes that generate new alleles, primarily mutation, those that modify allele frequencies, including migration (gene flow), natural selection, and genetic drift, and those that alter genotype frequencies, including non-random mating (Frankham et al., 2004, 2010). These evolutionary forces shape the distribution of genetic variation among and within populations and influence their evolutionary paths (Frankham et al., 2004, 2010).

In nature, genetic variation can be classified into three categories, each with distinct implications for conservation (Kohn et al., 2006):

- Neutral genetic variation encompasses differences that do not directly influence fitness and are shaped primarily by evolutionary forces such as mutation, gene flow, and genetic drift (Kohn et al., 2006).

- Detrimental genetic variation includes genetic changes that negatively affect individual fitness. Such variation can accumulate through mutations, limited gene flow, or genetic drift, particularly in small populations. Its effects often emerge as inbreeding depression, driven by increased homozygosity for recessive deleterious alleles (Kohn et al., 2006).
- Adaptive genetic variation refers to variation in genes that allow populations to respond to environmental pressures, such as climate change, habitat alteration, or emerging pathogens (Kohn et al., 2006).

Assessing how evolutionary forces shape genetic variation requires a theoretical baseline describing how allele and genotype frequencies behave in the complete absence of evolutionary change. This foundation is provided by the Hardy-Weinberg equilibrium, formulated independently by the British mathematician Godfrey H. Hardy and the German physician Wilhelm Weinberg in 1908 (Hardy, 1908; Weinberg, 1908). The Hardy-Weinberg principle states that in a large, randomly mating diploid population with non-overlapping generations and no mutation, migration or selection, allele and genotype frequencies reach an equilibrium after a single generation and remain constant over time (Frankham et al., 2004; Frankham et al., 2010; Chen, 2010; Charlesworth & Charlesworth, 2010; Okasha, 2024). Under these conditions, the expected genotype frequencies are given by p^2, $2pq$ and q^2, where p and q represent the allele frequencies (Hardy, 1908; Weinberg, 1908; Hartl & Clark, 1997; Frankham et al., 2004; Frankham et al., 2010; Chen, 2010; Charlesworth & Charlesworth, 2010; Okasha, 2024).

➢ Mutation

Spontaneous mutation is recognised as the primary source of genetic variation (Dobzhansky, 1982; Charlesworth & Smith, 1983; Frankham et al., 2004, 2010; Charlesworth & Charlesworth, 2010). A mutation can be defined as a heritable change in an allele or chromosome, and the term refers both to molecular events that generate new variants, such as errors arising during DNA replication, and to the resulting altered genetic products (Frankham et al., 2004, 2010). Since every new allele originates from mutational events, mutation represents the starting point for all evolutionary processes (Frankham et al., 2004, 2010).

The most frequent forms of mutations are single-nucleotide substitutions, whereas small insertions and deletions caused by replication slippage are also common (Charlesworth & Smith, 1983). Other classes of mutations include transposable element insertions, large-scale deletions and duplications, changes in chromosome number, and structural rearrangements such as inversions and translocations. These alterations can have significant consequences for genome organisation and evolution (Charlesworth & Smith, 1983). Notably, mutations may vary widely in their effects, for instance, those occurring at loci influencing fitness, particularly deleterious or lethal mutations, are evolutionarily important. In contrast, many mutations in non-coding regions or those not altering amino acid sequences, have little or no effect on fitness. Despite their selective neutrality, these variants are extremely informative for assessing genetic differences among individuals, populations, and species (Frankham et al., 2004, 2010).

Overall, mutation rates tend to be very low in most organisms (Charlesworth & Smith, 1983; Lande, 1995; Frankham et al., 2004, 2010), largely because of the efficiency of DNA repair and error-correction mechanisms (Charlesworth & Smith, 1983; Charlesworth & Charlesworth, 2010). Genomes that lack such mechanisms, including those of RNA viruses and mitochondrial DNA, exhibit higher mutation rates (Charlesworth & Smith, 1983; Charlesworth & Charlesworth, 2010). Although mutation is essential as the source of new genetic variation, its low rate makes it a weak evolutionary force at individual loci. For a new mutation to contribute to evolutionary change, it must increase to a detectable frequency within the population (Charlesworth and Charlesworth, 2010). Consequently, the regeneration of genetic diversity lost through evolutionary processes is slow, often requiring thousands to millions of generations for variation at a single locus to be restored (Frankham et al., 2004, 2010).

➢ Migration

Migration, or gene flow, refers to the movement of reproductively successful individuals among populations (Allendorf, 1983; Hartl & Clark, 1997). When immigrants enter a population, they introduce alleles from genetically distinct sources, mixing gene pools and potentially restoring genetic diversity that has been lost. The degree of this contribution depends on both the proportion of immigrants, and the difference in allele frequencies between source and recipient populations (Frankham et al., 2004; 2010). Because even a small number of migrants can alter allele frequencies, gene flow is far more efficient than mutation at restoring genetic variation (Frankham et al., 2004, 2010).

In isolated or captive populations of species that remain abundant elsewhere, the introduction of individuals from large, genetically distinct source populations provides a powerful means of reintroducing genetic diversity. While comparable to mutation in its effect, gene flow typically occurs at higher rates (Lacy, 1987). Additionally, the genetic impact of migration on allele frequencies depends on how much the recipient population has diverged from the source. For this reason, immigration produces proportionally stronger genetic effects in small populations than in larger ones (Lacy, 1987). Finally, gene flow may have detrimental consequences for endangered species when it results in introgression from closely related, non-endangered taxa (Frankham et al., 2004, 2010).

➢ Natural selection

Natural selection is defined as the differential success of genotypes in contributing offspring to the next generation (Allendorf, 1983; Hartl & Clark, 1997). The concept was introduced by Charles Darwin in 1859 in "On the Origin of Species", drawing on his studies of artificial selection in domesticated species as well as observations from natural populations (Evans, 1984). For species to persist over long timescales, they must respond to changing physical and biotic conditions, such as shifts in climate, the emergence of new diseases, or the arrival of novel competitors and predators. In this context, adaptive evolution plays a leading role (Frankham et al., 2004, 2010). Depending on the selective mode, natural selection may operate in three forms: directional, which favours a single phenotype, shifting the mean trait to one direction; stabilizing, which favours intermediate phenotypes and reduces genetic variation; and disruptive, which favours extreme phenotypes over intermediates and increases genetic variation. Through these processes, the frequency of alleles that enhance reproduction and survival increases, thereby enabling populations to adapt to environmental change (Frankham et al., 2004, 2010).

Selection decreases the frequency of deleterious alleles while increasing the prevalence of advantageous ones, gradually shaping the genetic composition of populations toward higher average fitness (Frankham et al., 2004, 2010).

Notably, a mutation-selection balance is established when the continual introduction of deleterious alleles through mutation is mitigated by their removal through natural selection (Hartl & Clark, 1997; Frankham et al., 2004, 2010). As a result, low frequencies of harmful alleles persist in all naturally outbreeding populations, a phenomenon known as the mutation load, originally conceptualised by Haldane (1937) and Muller (1950) (Charlesworth & Smith, 1983; Hartl & Clark, 1997; Frankham et al., 2004, 2010).

Selective sweep

Natural selection is commonly described as a gradual change in allele frequencies driven by differences in fitness among genotypes. However, selection may also generate rapid and strong effects on patterns of genetic variation when an advantageous mutation arises. In such cases, positive selection may drive a beneficial allele to increase rapidly in frequency, often leading to its fixation within the population (Stephan, 2019). These processes are described by the concept of "selective sweep" (Berry et al., 1991; Stephan, 2019), historically known as the genetic hitch-hiking effect, first described by Maynard Smith and Haigh (1974). During a selective sweep, the rapid spread of a favourable allele leads to a parallel increase in the frequency of linked neutral variants, resulting in a reduction of genetic variation in genomic regions surrounding the selected site (Frankham et al., 2004, 2010; Charlesworth & Charlesworth, 2017; Stephan, 2019).

➢ Genetic drift

Genetic drift refers to random fluctuations in allele frequencies that arise from the stochastic sampling of gametes from one generation to the next (Allendorf, 1983; Hartl & Clark, 1997). The power of genetic drift intensifies as population effective size (N_e, the size of an idealised population that would undergo genetic drift or inbreeding at a rate equivalent to that observed in the population under study) declines, causing its effects to be most evident in small populations. In this context, it may cause the loss of genetic diversity and the fixation of alleles, with the increase of homozygosity and the reduction of evolutionary potential

(Charlesworth & Smith, 1983; Frankham et al., 2004, 2010). It can also drive divergence among populations that originated from the same ancestral source, such as in the case of habitat fragmentation (Charlesworth & Smith, 1983; Frankham et al., 2004, 2010). When such fragmented populations are analysed as a unique, large genetic group, their differentiation can lead to an overall reduction in heterozygosity, a phenomenon known as the Wahlund's effect, which arises from differences in allele frequencies among subpopulations (Wahlund, 1928; Charlesworth & Smith, 1983; Hartl & Clark, 1997; Garnier-Géré & Chikhi, 2013).

Finally, genetic drift may even overwhelm the effects of natural selection, allowing alleles to change in frequency regardless of their fitness consequences (Frankham et al., 2004, 2010). Since the power of drift increases as population size decreases, its effects are particularly severe during episodes of sharp population decline or colonisation events. These processes, known as bottlenecks and founder effects represent key scenarios where genetic drift can reshape genetic variation (Frankham et al., 2004, 2010). Additionally, these demographic models lay the foundation for the neutral theory of molecular evolution, which proposes that most of the genetic variation observed among and within species is governed by genetic drift acting on neutral mutations (Kimura, 1983; Charlesworth & Smith, 1983; Kreitman & Akashi, 1995).

Bottleneck

A bottleneck occurs when a population experiences a sudden and severe reduction in size. This often leads to the loss of alleles, particularly those occurring at low frequencies, to a decline in overall genetic diversity, and to random changes in allele frequencies due to genetic drift (O'Brien, 1994; Frankham et al., 2004, 2010). Consequently, such bottlenecks reduce the evolutionary potential of the affected population (Frankham et al., 2004, 2010).

Founder effect

The founder effect arises when a new population is formed by a reduced number of individuals originating from a larger source population. This typically leads to a reduction in genetic diversity, pronounced genetic drift and an increased likelihood of inbreeding (Frankham et al., 2004, 2010).

➢ Non-random mating

In most studies of natural, outbreeding populations, mating is assumed to occur at random; similarly, in populations subdivided in small subpopulations, individuals are generally considered to mate randomly. Non-random mating therefore refers to any pattern of mate choice that deviates from the expectation of random mating under Hardy-Weinberg equilibrium (Caballero & Hill, 1991; Keller & Waller, 2002). It can take different forms, each with distinct consequences for offspring, including variation in individual mating success, assortative mating, and frequency-dependent mating success. In particular, assortative mating can occur in two forms: positive assortative mating, where individuals preferentially mate with those with similar phenotypes, and negative assortative mating, where individuals preferentially mate with those with dissimilar phenotypes (Partridge, 1983). Additionally, inbreeding represents an extreme form of positive assortative mating and has the greatest impact on genetic variation and offspring fitness (Partridge, 1983; Pierce, 2016).

Inbreeding

Inbreeding is defined as the mating between related individuals. In small or closed populations, it is unavoidable because of the limited number of founders and the reduced population size (Hartl & Clark, 1997; Keller & Waller, 2002; Frankham et al., 2004, 2010). While inbreeding can arise also in larger populations, it accumulates more slowly due to the greater availability of unrelated mating partners. Its evolutionary and conservation relevance is considerable, indeed, as related individuals mate, heterozygosity declines and the risk of expressing deleterious recessive alleles rises, ultimately leading to reduced reproductive fitness and survival, a condition known as inbreeding depression (Reed et al., 2002; Frankham et al., 2004, 2010).

2.3 Neutral theory of molecular evolution

The neutral theory of molecular evolution, first proposed by Kimura (1983), suggests that a large portion of genetic variation is generated by neutral mutations whose evolutionary fate is shaped by random genetic drift rather than natural selection (Kimura, 1983; Charlesworth

& Smith, 1983; Frankham et al., 2004, 2010; Kern & Hahn, 2018). Even though most neutral mutations are rapidly lost, new mutations are continually introduced into populations, and a small fraction increase in frequency purely by chance, with some eventually reaching fixation (Frankham et al., 2004, 2010).

A key prediction of the neutral theory is that the rate of molecular evolution remains constant over long timescales and is largely independent of population size, since it is determined primarily by the mutation rate (Kimura, 1983; Charlesworth & Smith, 1983; Frankham et al., 2004, 2010). The neutral theory also predicts higher levels of polymorphism in genomic regions experiencing weak selective constraints, such as non-functional sequences or synonymous sequences, where neutral mutations are more likely to accumulate (Kimura, 1983; Charlesworth & Smith, 1983; Frankham et al., 2004, 2010).

➢ The coalescent theory

Within the framework of the neutral theory of molecular evolution, the coalescent theory provides a retrospective approach for describing genetic processes by tracing the ancestry of sampled alleles backward in time (Charlesworth & Smith, 1983; Hartl & Clark, 1997; Rosenberg & Nordborg, 2002; Frankham et al., 2004, 2010; Charlesworth & Charlesworth, 2017). Originally formalised by Kingman (1982) and subsequently developed by Hudson (1983) and Tajima (1983) (Charlesworth & Charlesworth, 2017), this concept describes how gene lineages merge, or coalesce, toward a single ancestral sequence as a consequence of genetic drift acting on a finite population (Charlesworth & Smith, 1983; Frankham et al., 2004, 2010; Charlesworth & Charlesworth, 2017). In this context, coalescent theory provides a means of modelling the persistence and transmission of alleles over time, with genealogical relationships among sequences commonly represented as branching trees (Charlesworth & Smith, 1983, Rosenberg, 2002; Frankham et al., 2004, 2010; Charlesworth & Charlesworth, 2017). Moreover, it establishes the conceptual basis for interpreting genetic data in terms of population divergence, historical relationships, and evolutionary history (Frankham et al., 2004, 2010).

2.4 Population genetic structure

Investigating population genetic structure is fundamental for understanding species evolution, as it reveals whether populations are diverging along independent evolutionary paths or remain connected through ongoing gene flow. Weak genetic structure typically reflects high connectivity, whereas strong differentiation may indicate initial stages of speciation. Patterns of genetic structure are also shaped by life-history traits, geographic distribution, and other ecological characteristics that directly or indirectly influence gene dispersal (Duminil et al., 2007).

Most natural populations exhibit some degree of fragmentation rather than behaving as a single, panmictic unit. The genetic consequences of fragmentation depend strongly on the extent of gene flow among habitat fragments (Charlesworth & Smith, 1983; Frankham et al., 2004, 2010). When fragmentation occurs without sufficient gene exchange, subpopulations may experience elevated levels of inbreeding, accelerated loss of genetic diversity and an increased risk of extinction (Frankham et al., 2004, 2010).

In this framework, several theoretical models describe how populations may be genetically structured. At one extreme, there are completely isolated fragments, where no gene flow occurs. At the opposite end, there are systems in which gene flow is sufficiently frequent that fragments function as a single, large population (Frankham et al., 2004, 2010). Intermediate scenarios include the island model, in which equally sized subpopulations exchange migrants at uniform rates; the linear stepping-stone model, where gene flow occurs only between neighbouring populations (e.g., along a river or coastline); and the two-dimensional stepping-stone model, where migration is limited to adjacent populations (e.g., along a landscape with units in multiple directions) (Charlesworth & Smith, 1983; Frankham et al., 2004, 2010). Additional models include mainland-island (source-sink) structures, where gene flow moves predominantly from a large central population into smaller peripheral ones, and metapopulations, characterised by recurrent local extinction events and subsequent recolonisation (Frankham et al., 2004, 2010).

The extent of genetic differentiation among population fragments can be quantified using Wright's F-statistics, which partition overall inbreeding into components that reflect variation within and among sub-populations. Wright (1969) divided the total inbreeding coefficient, F_{IT}, into F_{IS}, which measures the reduction in heterozygosity of individuals relative to their own subpopulation (averaged across all fragments), and F_{ST}, which quantifies the proportion of genetic variance attributable to population subdivision, reflecting genetic differences among subpopulations in relation to the total population. For multi-allelic loci, this component is commonly denoted G_{ST} (Frankham et al., 2004, 2010).

➢ Genetic distance

Genetic distance was originally developed to quantify the degree of genetic differentiation among populations. It indicates any numerical measure of genetic divergence derived from molecular data, whether based on differences in DNA sequences or on variation in allele frequencies, and it can be calculated between individuals, populations, or species. Conceptually, genetic distances can be represented as distances among points in a multi-dimensional space. Many distance measures based on population genetic models includes distances derived from F_{ST} estimates, as well as Nei's *D* (Nei, 1972; Beaumont et al., 1998). Among the various metrics available, Nei's genetic distance remains one of the most widely used measures of genetic differentiation across populations and species. In general, genetic distance increases with the degree of reproductive isolation (Frankham et al., 2004, 2010).

➢ Idealised population and effective population size

Population and conservation genetic theory often relies on "idealised" population models, which assume a closed population with random mating, constant size, non-overlapping generations, equal sex ratios, equal reproductive success among individuals, and the absence of evolutionary forces. Because natural populations rarely satisfy these conditions, idealised models provide a theoretical baseline against which real populations can be evaluated, and deviations from these ideal conditions can be interpreted (Frankham et al., 2004, 2010).

Effective population size (N_e)

The effective population size (N_e) describes the size of an idealised population that would undergo genetic drift or inbreeding at a rate equivalent to that observed in the population under study (Hartl & Clark, 1997; Frankham et al., 2004, 2010). In a scenario where a natural population fully conforms to the assumptions of an ideal population, N_e would be equal to the census size (N). However, real populations almost always depart from idealised conditions, causing N_e to differ from, and typically be lower than, N. Factors such as variance in individual reproductive success, unequal sex-ratios, generational overlap, and temporal fluctuations in population size contribute to reducing N_e relative to the census size (Frankham et al., 2004, 2010).

3. Fundamentals of evolutionary biology applied to population genetics

3.1 The biological species concept and speciation

Defining species boundaries is a central issue in evolutionary biology and has direct implications for conservation genetics, as taxonomic decisions directly influence conservation priorities, protection initiatives, and management strategies. Within this framework, the principal aim is the identification of taxa with unresolved or unknown taxonomic status, to prevent their inadvertent extinction. These include the so-called "cryptic" or "sibling" species, representing individuals which are morphologically indistinguishable, but genetically different and reproductively isolated (Frankham et al., 2004, 2010).

Because no single, universally accepted definition of species exists, numerous species concepts have been proposed over time. Early approaches were dominated by the morphological species concept until the 1930s (Coyne, 1994), followed by alternative frameworks such as the internodal species concept introduced by Kornet in 1993 (Kornet, 1993; Coyne, 1994). Of particular relevance to modern genetic studies is the phylogenetic species concept (PSC), which defines species as the smallest monophyletic groups sharing a common evolutionary history (Cracraft, 1989; de Queiroz & Donoghue, 1990). Despite the diversity of perspectives, the biological species concept (BSC) has emerged as the most influential evolutionary definition of species and remains central to population, evolutionary, and conservation genetics (Coyne, 1994, Frankham et al., 2004, 2010).

Originally formulated by Dobzhansky in 1935 and subsequently refined by Mayr in 1942, the biological species concept defines a species as a group of actually or potentially interbreeding natural populations that can exchange genetic material among themselves but are reproductively isolated from other such groups (Dobzhansky, 1935; Mayr, 1942; Coyne 1994). Reproductive isolation is therefore a defining feature of species boundaries under the BSC, as it prevents gene flow between distinct species and maintains their evolutionary independence (Frankham et al., 2004, 2010). This concept has remained the dominant evolutionary framework largely because it integrates biological and genetic perspectives, defines species as distinct evolutionary entities, and provides a coherent basis for interpreting biological diversity (Coyne, 1994). Nevertheless, the concept is considered controversial, as it is difficult to apply to asexual and inbreeding organisms, to taxa that hybridise, and fossil species. Despite these limitations, the BSC represents a practical and widely used approach for genetically delimiting species (Coyne, 1994; Frankham et al., 2004, 2010).

In this context, subspecies are typically defined as geographically or ecologically distinct groups of populations within a species that exhibit partial genetic differentiation from other conspecific populations. Unlike species under the biological species concept, subspecies rarely display complete reproductive isolation. However, they often represent populations undergoing evolutionary divergence, potentially constituting early stages of speciation (Frankham et al., 2004, 2010).

➢ Speciation

Speciation refers to the evolutionary process through which new species originate and can be defined as the splitting of a single ancestral lineage into two or more independently evolving daughter species (Mayr, 1963). In most cases, speciation proceeds through the evolution of reproductive isolation, which limits gene flow and leads genetic differences to accumulate over time (Mayr, 1963; Templeton, 1981; Coyne, 1994; Frankham et al., 2004, 2010). This process is driven by evolutionary forces such as mutation, genetic drift, and natural selection, and is often reinforced by geographic separation, ecological divergence, or behavioural barriers (Mayr, 1963; Coyne, 1994; Frankham et al., 2004, 2010).

Based on the mechanisms and spatial context through which reproductive isolation evolves, speciation can be classified into several major modes. These comprise instantaneous speciation, as well as forms of gradual speciation, including allopatric speciation and sympatric speciation (Mayr, 1963; Coyne, 1994; Frankham et al., 2004, 2010).

Instantaneous speciation

Instantaneous speciation denotes the emergence of a new species from a single individual or from the offspring of a single reproductive event that is both reproductively isolated from its parental species and capable of persisting as an independent population. In such cases, reproductive isolation arises in a single generation (Mayr, 1963). This mode of speciation is rare in most sexually reproducing animals but may occur in parthenogenetic organisms or self-fertilizing hermaphrodites, where reproduction does not require outcrossing. In plants, however, instantaneous speciation is relatively common and is most often associated with polyploidy (Mayr, 1963; Frankham et al., 2004, 2010).

Allopatric speciation

Allopatric or geographic speciation is generally considered the most common mode of speciation in animals and is also likely to be prevalent in plants (Mayr, 1963; Frankham et al., 2004, 2010). It occurs when populations of a species become physically separated by geographic barriers or through colonisation of new areas, resulting in the interruption of gene flow (Frankham et al., 2004, 2010). Over time, they evolve independently, leading to the development of reproductive isolating mechanisms that prevent interbreeding if populations come into secondary contact. These barriers allow the formerly isolated populations to potentially coexist in sympatry as distinct species (Mayr, 1963; Coyne, 1994). Divergence is often reinforced by adaptation to different environments, ecological specialisation, and sexual selection, all of which contribute to the establishment and maintenance of reproductive isolation (Templeton, 1982; Frankham et al., 2004, 2010).

Sympatric speciation

Sympatric speciation occurs when evolutionary divergence takes place within the geographic range of the ancestral species, in the absence of physical barriers to gene flow.

Therefore, sympatric populations occupy the same or overlapping areas and have the potential to exchange genes (Mayr, 1963; Frankham et al., 2004, 2010).

This mode of speciation is based on two key conditions. First, subsets of a population become established in different ecological niches within the normal dispersal range of the ancestral population. Second, reproductive isolation must evolve between these emerging groups and the parental population, thereby restricting gene flow (Mayr, 1963). Strong ecological differentiation or behavioural isolation can rapidly reinforce divergence, for this reason most models of sympatric speciation imply a relatively rapid process of species formation (Mayr, 1963).

Although sympatric speciation is generally considered less frequent than allopatric speciation, the number of documented cases has increased in recent decades, particularly where ecological specialisation plays a central role (Coyne, 1994; Frankham et al., 2004, 2010).

4. Applications of population genetics

Molecular population genetics studies the role of evolutionary forces in shaping the molecular sequence variation observed within populations (Frankham et al., 2004, 2010). This is achieved through the integration of mathematical models, statistical estimation of population genetic parameters, and empirical observations (Hartl & Clark, 1997). Today, the discipline addresses a broad range of biological questions, spanning phylogenetic relationships, phylogeographic patterns, taxonomic uncertainties, and phylodynamic processes.

4.1 Phylogenetics

Phylogenetics is the study of evolutionary relationships among genes, populations, and species, with the aim of reconstructing their ancestral history and patterns of divergence through time. Within evolutionary biology, it provides a framework for interpreting genetic similarities and differences as a result of historical evolutionary processes (Nei, 1996; Hartl & Clark, 1997).

The introduction of molecular techniques for determining nucleotide and amino acid sequences represented a major turning point in phylogenetic research. In classical phylogenetic approaches, phenotypic traits and morphological observations were the primary basis for inferring relationships among organisms and reconstructing evolutionary trees (Nei, 1996). However, morphological characters may not always be heritable and can be influenced by environmental factors. In contrast, molecular characters provide heritable information that accumulates over generations, making them particularly informative for evolutionary inference. This advance led to major shifts in the reconstruction of evolutionary relationships across almost all levels of biological classification, from kingdoms to species, and even among intraspecific populations, and revealed deep evolutionary relationships previously unrecognised (Nei, 1996; Okazaki, 2021).

Over time, several researchers have made relevant contributions to the development of phylogenetics by introducing statistical approaches and analytical tool for tree reconstruction and interpretation. These contributions include early work by Fisher (1922), followed by foundational developments by Zuckerkandl and Pauling (1965), Cavalli Sforza and Edwards (1967), Felsenstein (1973, 1978), and Adachi and Hasegawa (1992, 1996), establishing the modern phylogenetic basis (Okazaki, 2021).

Phylogeny refers to the evolutionary history of a group of taxa or species descending from common ancestors, including the order of branching events and, in some cases, estimates of divergence times (Patwardhan et al., 2014). These relationships are typically represented as phylogenetic trees composed of branches and nodes, where branches represent the persistence of genetic lineages through time and nodes correspond to lineage-splitting events (Yang & Rannala, 2012). Different methods are available for reconstructing phylogenies, and their application depends on the scope and objectives of the analysis. These include distance-based methods and character-based methods, the latter encompassing maximum parsimony, maximum likelihood, and Bayesian inference (Yang & Rannala, 2012; Okazaki, 2021).

Today, phylogenetic analyses are applied across a wide range of biological areas, including the reconstruction of hierarchical patterns of species evolution, the investigation of evolutionary history in gene families, and the study of adaptive evolution at the molecular level (O'Brien, 1994; Nei, 1996; Yang & Rannala, 2012). A range of molecular markers is commonly used in phylogenetic studies, including ribosomal RNA genes (e.g. 18S, 16S, 28S,) and mitochondrial DNA genes (e.g., Cytochrome c Oxidase subunit I - COI, cytochrome b, and Control Region), which have proven to be particularly effective for resolving phylogenetic relationships at both species and population levels (Patwardhan et al., 2014).

Overall, phylogenetics represents a powerful approach for addressing fundamental questions in evolutionary biology, improving the understanding of how genetic diversity is structured across lineages and through time, and providing crucial insights into the evolutionary processes and relationships shaping species and populations (Yang & Rannala, 2012).

4.2 Phylogeography

Phylogeography is a subfield of population genetics that focuses on the geographic distribution of genetic lineages, particularly within and among closely related species (Avise et al., 1987; Avise, 1995; Emerson & Hewitt, 2005; Avise, 2008). The term was first introduced by Avise and colleagues in 1987 (Avise et al., 1987; Avise, 2008).

Early phylogeographic studies relied primarily on mitochondrial DNA (mtDNA) to explore the genealogical connections between conspecific individuals through shared ancestors (Torroni et al., 1996; Avise, 2008). Due to its maternal inheritance, lack of recombination, and relatively high mutation rate, mtDNA has proven to be suitable for detecting population structure and historical migration patterns, particularly at intraspecific levels (O'Brien, 1994). The genetic variants of mtDNA are referred to as haplotypes, and it is typical to find multiple haplotypes coexisting within a species (Torroni et al., 1996; Avise, 2008).

Nuclear DNA can also be employed for phylogeographic studies, but its generally slower evolutionary rate at many loci and the challenges in isolating nuclear haplotypes limited its widespread application in phylogeographic research (Avise, 2008).

Phylogeographic studies typically investigate multiple populations, distributed across a geographic region, which may have experienced past or contemporary barriers to dispersal. These barriers could be physical, such as rivers or mountains, or behavioural, limiting gene flow and leading to varying degrees of population isolation. Populations may therefore be partially isolated by distance, with occasional gene exchange, or completely isolated for extended periods. Notably, species with high dispersal capability, such as many marine organisms, have shown evidence of historical barriers to gene flow (Avise, 2008).

The primary goal of phylogeographic research is to use gene trees or genealogical networks to reconstruct the historical and ongoing processes that shape the genetic structure of populations and species (Posada & Crandall, 2001; Emerson & Hewitt, 2005). When these networks are integrated with information about the geographic distribution of populations, phylogeographic analyses can provide insights into demographic history, including past population bottlenecks, range expansions, size of ancestral populations, refugial areas, and migration paths (Hemerson & Hewitt, 2005). These studies can also highlight the genetic consequences of habitat fragmentation, which may vary depending on both the persistence of geographic barriers and the historical demographic trajectories of isolated populations (Hemerson & Hewitt, 2005; Avise, 2008).

Although comprehensive phylogeographic studies may integrate data from both nuclear and mitochondrial genomes, analyses based only on mtDNA have often been sufficient to reveal phylogeographic patterns (Avise, 1991, 2008). Moreover, advances in DNA sequencing technologies, including whole-genome approaches, have greatly expanded the scope of phylogeographic studies, enabling more detailed investigations into demographic histories of populations (Emerson & Hewitt, 2005).

4.3 Phylodynamics

Phylodynamics is an interdisciplinary field within population genetics that focuses on inferring changes in population size and transmission dynamics over time using genetic data sampled from pathogen populations. The term was introduced by Grenfell and colleagues in 2004 to describe an integrative framework combining evolutionary biology, epidemiology, and immunodynamics (Grenfell et al., 2004; Lan et al., 2015; Baele et al., 2017, 2018). Phylodynamic approaches connect patterns of genetic variation to demographic and epidemiological processes and are highly informative for investigating evolutionary dynamics over short timescales. The central aim of phylodynamics is to reconstruct epidemic processes through phylogenetic inference, by examining how these dynamics are reflected in the accumulation of mutations in pathogen genomes during outbreaks (Lan et al., 2015; Baele et al., 2017). In this context, the integration of statistical modelling and computational inference is essential (Lan et al., 2015).

Phylodynamic analyses often incorporate information on evolutionary ancestry with estimates of changes in population size, genealogical branching rates, and, in many cases, spatial dynamics (Lan et al., 2015). Phylodynamic models, when combined with phylogeographic approaches, can reveal patterns of connectivity underlying transmission, estimate movement rates among regions, and trace the spatial dissemination of viral lineages (Baele et al., 2018; Attwood et al., 2022). This integrative approach makes phylodynamics particularly valuable for investigating the spread of rapidly evolving infectious agents, especially RNA viruses, such as influenza (Lan et al., 2015; Scarpa et al., 2024).

From a methodological point of view, phylodynamics relies on molecular clock models, which describe the rate at which nucleotide substitutions accumulate over time, and on population dynamic models that reconstruct changes in population size from genetic sequence data (Guinat et al., 2021; Attwood et al., 2022). Bayesian statistical tools are widely used, because they provide a coherent and flexible approach for integrating genetic data with temporal information (Volz et al., 2013; Baele et al., 2017). Many phylodynamic methods are based on coalescent theory and are usually implemented through skyline-based plots to estimate historical changes in the effective population size (Attwood et al., 2022).

Besides human health, phylodynamic approaches are increasingly applied to investigate disease emergence and transmission across the wildlife-livestock-human system, providing insights into zoonotic spillover and cross-species transmission dynamics (Guinat et al., 2021).

Developments in whole-genome sequencing technologies have further expanded the application of phylodynamics beyond rapidly evolving pathogens to include more slowly evolving organisms, broadening its relevance within evolutionary biology and population genetics (Guinat et al., 2021).

4.4 Molecular taxonomy and species delimitation

Taxonomy is a discipline focused on the identification, classification, and naming of organisms, providing a framework for recognising biological diversity and assigning identities to newly discovered taxa (Wheeler et al., 2004; Kim & Byrne, 2006; Schlick-Steiner et al., 2010). By defining species and higher categories, taxonomy plays a fundamental role in biological research and provide the basis for comparative, evolutionary and conservation studies (O'Brien, 1994; Tautz, 2002; Kim & Byrne, 2006).

Taxonomic research is particularly important in the context of biodiversity conservation. A substantial proportion of Earth's biodiversity remains undescribed, and many species have likely gone extinct before being formally recognised (Frankham et al., 2004, 2010; Rannala & Yang, 2020).

In this regard, taxonomy covers several key roles, including the assignment of individuals to known species or higher taxonomic units, the delimitation and description of new species, and the identification of cryptic diversity within morphologically similar taxa (Rannala & Yang, 2020). Advances in taxonomic research have led to an increasing number of discoveries of cryptic species, particularly in understudied groups such as marine organisms and tropical terrestrial taxa, while also demonstrating that some nominal species previously considered distinct belong to the same species (Schlick-Steiner et al., 2010).

Early taxonomic investigations relied primarily on morphological traits, which remain an important source of information but may be insufficient for reliable species delimitation. Moreover, morphological approaches may be time-consuming, require considerable taxonomic expertise, and may fail to detect species boundaries. For these reasons, traditional taxonomy can be complemented by the integration of complementary approaches, particularly molecular data, to support species identification and delimitation (Schlick-Steiner et al., 2010; Rannala & Yang, 2020). Since the early 2000s, DNA barcoding has been widely used as a standard molecular tool for species identification, with the mitochondrial Cytochrome c Oxidase Subunit I (COI) gene proposed as a universal diagnostic marker for many animal taxa (Hebert et al., 2003; Frankham et al., 2004, 2010). Hence, the DNA barcoding facilitates the recognition of genetically distinct groups and can highlight levels of sequence divergence consistent with a differentiation at the species level (Hebert & Gregory, 2005; Rannala & Yang, 2020).

Additionally, a range of statistical species delimitation methods has been developed to support and refine taxonomic inference. These models are particularly useful for resolving taxonomic uncertainties and they can be broadly classified into two main categories. The first comprises methods based on the Phylogenetic Species Concept (PSC), which delimits species considering the phylogenetic relationships within the dataset, including the Poisson Tree Processes (PTP) and its Bayesian implementation (bPTP) (Zhang et al., 2013), as well as the Generalised Mixed Yule Coalescent (GMYC; Pons et al., 2006). The second category comprises models based on genetic distances, including the Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2012), the Nucleotide Divergence Threshold (NDT; Scarpa et al., 2019), and the Assemble Species by Automatic Partitioning (ASAP; Puillandre et al., 2020). Additionally, the K/θ method represents a hybrid approach that integrates both the PSC and genetic distances (Birky et al., 2010).

5. Contemporary challenges of population genetics

5.1 Aquatic organisms as model systems

Aquatic organisms, including marine, freshwater, and estuarine species, represent valuable model systems for population genetic studies due to their diverse life-history strategies, dispersal capacities, adaptive processes, and patterns of population connectivity and isolation (Plough, 2016; Manel et al., 2020). Aquatic ecosystems have unique physical and ecological dynamics, including thermal and oxygen gradients, seasonal fluctuations in currents, flow rates, and water depth, as well as natural environmental barriers, all of which influence gene flow and promote genetic divergence among populations (Grummer et al., 2019). These characteristics make aquatic systems particularly suitable for investigating how ecological and environmental processes, along with evolutionary forces interact to shape genetic structure and species diversity. In this context, population genetic approaches provide powerful tools for elucidating the mechanisms driving genetic structure and diversification across spatial and temporal scales (Manel et al., 2020; Yi et al., 2023). Additionally, several aquatic organisms (e.g., marine and freshwater fish and invertebrates) have biological traits that make them accessible for genetic investigations, including ease of sampling and laboratory handling (Ribas & Piferrer, 2014).

➢ Threats to aquatic populations

The continuous growth of global human population poses increasing challenges to the long-term persistence of animal populations (Reed, 2004; Frankham et al., 2004, 2010; Hogue & Breon, 2022). Aquatic biodiversity is currently exposed to a wide range of anthropogenic and environmental threats, including habitat loss and fragmentation, environmental degradation, biological invasions, emerging diseases, water pollution, overexploitation, and climate change (Moyle & Leidy, 1992; Frankham et al., 2004; Reusch et al., 2005; Frankham et al., 2010; Arthington et al., 2016; Hogue & Breon, 2022; Nikolaou & Katsanevakis, 2023; Zhou, 2024; Mayer & Pŝenîcka, 2024). In addition, population viability is influenced by different stochastic factors, comprising demographic, environmental, including catastrophic events, and genetic processes (Reed, 2004; Frankham et al., 2004, 2010).

Collectively, these pressures affect population connectivity, reduce genetic diversity, and potentially increase the extinction risk (Frankham et al., 2004; 2010), particularly in small, fragmented, and island populations (Frankham, 1997; Reed 2004). Consequently, effective monitoring and management of aquatic ecosystems are essential to preserve the long-term sustainability of aquatic resources (Zhou, 2024).

Fragmented populations and habitat loss

Habitat alteration and destruction resulting from direct human impacts represent major threats to aquatic species, shaping population spatial structure and connectivity (Hogue & Breon, 2022). Small populations are particularly vulnerable under these conditions, as habitat fragmentation and environmental change can reduce population viability and promote genetic erosion (Willi et al., 2006). In both marine and freshwater systems, habitat and population fragmentation frequently arise from the construction of artificial barriers, such as dams and aquaculture facilities, which can alter natural dispersal routes and interfere with ecological processes, affecting population persistence through time (Arthington et al., 2016).

Invasive species

Biological invasions represent one of the main drivers of ecological change in aquatic ecosystems, contributing to habitat degradation, competition with native biodiversity, introduction of novel pathogens, and shifts in the evolutionary pathways of recipient populations (Le Roux & Wieczorek, 2009). The study of invasive species is essential for resolving taxonomic uncertainties, identifying cryptic diversity, reconstructing invasion origins and dynamics, assessing dispersal patterns, and determining potential hybridization and introgression events with native populations. Notably, high levels of genetic variability within invasive populations may further influence their ability to spread and adapt to new environments (Sakai et al., 2001; Mooney & Cleland, 2001; Lee, 2002; Le Roux & Wieczorek, 2009). In this context, population genetics provide a crucial basis for inferring genetic structure and colonisation dynamics of invasive species, supporting effective monitoring and management strategies to limit their impacts (Sakai et al., 2001).

6. Aims of the doctoral project and species involved

Population genetics represented the unifying thread linking the different studies I undertook during my doctoral programme, allowing the investigation of multiple genetic and evolutionary aspects of aquatic biodiversity across different biological systems. The main aim of this doctoral project was to investigate genetic diversity, population structure, and evolutionary dynamics under distinct ecological contexts through the application of population genetic approaches, including phylogenetics, phylogeography, and molecular taxonomy.

To this end, different mitochondrial molecular markers have been employed to explore genetic processes shaping aquatic biodiversity, including the effects of demographic decline and habitat fragmentation, population resilience to anthropogenic pressures and environmental changes, patterns of connectivity and isolation, and the genetic dynamics associated with biological invasions. The ultimate goal was to provide relevant information and support biodiversity monitoring and conservation management strategies in both marine and freshwater environments.

The aims of this thesis are addressed through four case studies focusing on different aquatic organisms, spanning both marine and freshwater habitats. Each chapter, dedicated to a specific species, begins with a general introduction to the species and a presentation of the study objectives, followed by the corresponding scientific article.

From a marine perspective, Chapter 1 focuses on the Mediterranean *Pinna nobilis*, a marine bivalve species endemic to the Mediterranean Sea (Butler et al., 1993), currently on the brink of extinction due to the combined effects of multiple anthropogenic pressures and a multifactorial disease that began affecting its populations in 2016 (Scarpa et al., 2020). Chapter 2 investigates the invasive blue crab *Callinectes sapidus*, a crustacean species native to the western Atlantic that has recently undergone a rapid expansion into the Mediterranean Sea, where it represents a significant threat to native biodiversity and human activities (Nehring, 2011).

From a freshwater perspective, Chapter 3 examines *Salmo letnica*, an endemic trout species of Lake Ohrid, one of the oldest lakes in Europe, located on the border between Albania and North Macedonia (Stanković, 1960). The presence of four distinct morphotypes within the lake makes the taxonomic status of this species particularly puzzling (Sušnik et al., 2007). Chapter 4 focuses on *Salariopsis fluviatilis*, a benthic fish species inhabiting freshwater environments across several Mediterranean countries, whose populations are affected by environmental and anthropogenic stressors leading to population fragmentation (Laporte et al., 2014).

Finally, Appendix A includes a list of additional studies in which I took part during the doctoral programme, as well as a study applying population genetic and phylodynamic approaches to Orf virus (ORFV; Family: Poxviridae), a potentially zoonotic virus mainly affecting small ruminants (Bergqvist et al., 2017). While not focused on aquatic organisms, this study illustrates the broader applicability of population genetics beyond aquatic systems.

7. References

Adachi, J. (1992). MOLPHY: programs for molecular phylogenetics I. PROTML. maximam likelihood inference of protein phylogeny. *Computer science monographs*, *27*, 1–14. https://doi.org/10.1266/jjg.67.187

Adachi, J., & Hasegawa, M. (1996). *MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood*. Institute of Statistical Mathematics.

Allendorf, F. W. (1983). Isolation, gene flow, and genetic differentiation among populations. *Genetics and conservation*, *18*(3), 51-65.

Allendorf, F. W. (1983). Isolation, gene flow, and genetic differentiation among populations. *Genetics and conservation*, *18*(3), 51-65.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature reviews genetics*, *11*(10), 697-709. https://doi.org/10.1038/nrg2844

Arthington, A. H., Dulvy, N. K., Gladstone, W., & Winfield, I. J. (2016). Fish conservation in freshwater and marine realms: status, threats and management. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *26*(5), 838-857. https://doi.org/10.1002/aqc.2712

Attwood, S. W., Hill, S. C., Aanensen, D. M., Connor, T. R., & Pybus, O. G. (2022). Phylogenetic and phylodynamic approaches to understanding and combating the early SARS-CoV-2 pandemic. *Nature Reviews Genetics*, *23*(9), 547-562. https://doi.org/10.1038/s41576-022-00483-8

Avise, J. C. (1991). Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annual review of genetics*, *25*(1), 45-69.

Avise, J. C. (2008). Phylogeography: retrospect and prospect. *Journal of biogeography*, *36*(1), 3-15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual review of ecology and systematics*, 489-522.

Avise, J.C. (1995) Mitochondrial DNA polymorphism and a connection between genetics and demography of relevance to conservation. *Conservation Biology*, *9*, 686–690.

Baele, G., Dellicour, S., Suchard, M. A., Lemey, P., & Vrancken, B. (2018). Recent advances in computational phylodynamics. *Current opinion in virology*, *31*, 24-32. https://doi.org/10.1016/j.coviro.2018.08.009

Baele, G., Suchard, M. A., Rambaut, A., & Lemey, P. (2017). Emerging concepts of data integration in pathogen phylodynamics. *Systematic biology*, *66*(1), e47-e65. https://doi.org/10.1093/sysbio/syw054

Bartlett, J. M., & Stirling, D. (2003). A short history of the polymerase chain reaction. In J.M.S. Bartlett, D. Stirling (Eds.), *PCR protocols* (pp. 3-6). Humana Press.

Baverstock P.R. & Moritz C. (1996) Project design. In D.M. Hillis, C. Moritz and B.K. Mable. (edz), *Molecular Systematics* (2nd ed., pp. 17–28). Sinauer Associates.

Beaumont, M.A., Ibrahim, K.M., Boursot, P., Bruford, M.W. (1998). Measuring Genetic Distance. In A. Karp, P.G. Isaac, D.S. Ingram (Eds.) Molecular Tools for Screening Biodiversity. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-0019-6_58

Bergqvist, C., Kurban, M., & Abbas, O. (2017). Orf virus infection. *Reviews in medical virology*, *27*(4), e1932. https://doi.org/10.1002/rmv.1932

Berry, A. J., Ajioka, J. W., & Kreitman, M. (1991). Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. *Genetics*, *129*(4), 1111-1117. https://doi.org/10.1093/genetics/129.4.1111

Berry, A. J., Ajioka, J. W., & Kreitman, M. (1991). Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. *Genetics*, *129*(4), 1111-1117. https://doi.org/10.1093/genetics/129.4.1111

Birky Jr, C. W., & Maughan, H. (2020). Evolutionary genetic species detected in prokaryotes by applying the K/θ ratio to DNA sequences. *bioRxiv*, 2020-04. https://doi.org/10.1101/2020.04.27.062828

Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *The Quarterly review of biology*, *74*(1), 21-45.

Butler, A., Vicente, N., & de Gaulejac, B. (1993). Ecology of the pterioid bivalves Pinna bicolor Gmelin. *Marine Life*, *3*(1-2), 37-45.

Caballero, A., & Hill, W. G. (1991). Effective size of nonrandom mating populations. *Genetics*, *130*(4), 909-916.

Cavalli-Sforza, L. L. (1998). The DNA revolution in population genetics. *Trends in Genetics*, *14*(2), 60-65. https://doi.org/10.1016/S0168-9525(97)01327-9

Cavalli-Sforza, L. L., & Edwards, A. W. (1967). Phylogenetic analysis. Models and estimation procedures. *American journal of human genetics*, *19*(3 Pt 1), 233.

Cavalli-Sforza, L. L., & Feldman, M. W. (2003). The application of molecular genetic approaches to the study of human evolution. *Nature genetics*, *33* (Suppl 3), 266-275. https://doi.org/10.1038/ng1113

Charlesworth, B. (2017). Haldane and modern evolutionary genetics. *Journal of Genetics*, *96*(5), 773-782. https://doi.org/10.1007/s12041-017-0833-4

Charlesworth, B., & Charlesworth, D. (2010). *Elements of Evolutionary Genetics*. Roberts and Company.

Charlesworth, B., & Charlesworth, D. (2017). Population genetics from 1966 to 2016. *Heredity*, *118*(1), 2-9. https://doi.org/10.1038/hdy.2016.55

Charlesworth, B., & Smith, J. M. (1983). Population genetics. *Nature*, *303*(5920), 748-748.

Chen, J. J. (2010). The Hardy-Weinberg principle and its applications in modern population genetics. *Frontiers in Biology*, *5*(4), 348-353. https://doi.org/10.1007/s11515-010-0580-x

Coyne, J. A. (1994). Ernst Mayr and the origin of species. *Evolution*, 19-30. https://doi.org/10.2307/2409999

Cracraft, J. (1989). Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. *Speciation and its Consequences*.

Crow, J. F. (1961). Population genetics. *American journal of human genetics*, *13*(1 Pt 2), 137.

Darwin, C. (1859). *On the origin of species*. John Murray.

de Queiroz, K., & M. J. Donoghue (1990). Phylo-genetic systematics or Nelson's version of cladistics? *Cladistics*, *6*, 61-75.

Dobzhansky, T. (1935). A critique of the species concept in biology. *Philosophy of Science*, *2*(3), 344-355.

Dobzhansky, T. (1951). *Genetics and the Origin of Species*. Columbia University Press.

Dobzhansky, T. (1982). *Genetics and the Origin of Species*. Columbia university press.

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits?. *The American Naturalist*, *169*(5), 662-672. https://doi.org/10.1086/513490

Emerson, B. C., & Hewitt, G. M. (2005). Phylogeography. *Current biology*, *15*(10), R367-R371. https://doi.org/10.1016/j.cub.2005.05.016

Evans, L. T. (1984). Darwin's use of the analogy between artificial and natural selection. *Journal of the History of Biology*, *17*(1), 113-140. https://doi.org/10.1007/BF00397504

Felsenstein, J. (1973a). Maximum-likelihood estimation of evolutionary trees from continuous characters. *American journal of human genetics*, *25*(5), 471.

Felsenstein, J. (1973b). Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Biology*, *22*(3), 240-249. https://doi.org/10.1093/sysbio/22.3.240

Felsenstein, J. (1978). The number of evolutionary trees. *Systematic zoology*, *27*(1), 27-33.

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. https://doi.org/10.1017/S0080456800012163

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, *222*(594-604), 309-368.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. The Clarendon Press.

Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation biology*, *10*(6), 1500-1508. https://doi.org/10.1046/j.1523-1739.1996.10061500.x)

Frankham, R. (1997). Do island populations have less genetic variation than mainland populations?. *Heredity*, *78*(3), 311-327. https://doi.org/10.1038/hdy.1997.46

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2004). *A primer of conservation genetics*. Cambridge University Press.

Frankham, R., Briscoe, D. A., & Ballou, J. D. (2010). *Introduction to conservation genetics*. Cambridge University Press.

Garnier-Géré, P., & Chikhi, L. (2013). Population subdivision, Hardy–Weinberg equilibrium and the Wahlund effect. Els. https://doi.org/10.1002/9780470015902.a0005446.pub3

Greminger, M. P., Krützen, M., Schelling, C., Pienkowska–Schelling, A., & Wandeler, P. (2010). The quest for Y-chromosomal markers–methodological strategies for mammalian non-model organisms. *Molecular ecology resources*, *10*(3), 409-420. https://doi.org/10.1111/j.1755-0998.2009.02798.x

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, *303*(5656), 327-332. https://doi.org/10.1126/science.1090727

Grummer, J. A., Beheregaray, L. B., Bernatchez, L., Hand, B. K., Luikart, G., Narum, S. R., & Taylor, E. B. (2019). Aquatic landscape genomics and environmental effects on genetic variation. *Trends in ecology & evolution*, *34*(7), 641-654. https://doi.org/10.1016/j.tree.2019.02.013

Guinat, C., Vergne, T., Kocher, A., Chakraborty, D., Paul, M. C., Ducatez, M., & Stadler, T. (2021). What can phylodynamics bring to animal health research?. *Trends in Ecology & Evolution*, *36*(9), 837-847. https://doi.org/10.1016/j.tree.2021.04.013

Haldane J. B. S. (1937) The effect of variation on fitness. *American Naturalist*, *71*, 337–349.

Haldane, J. B. (1930). A mathematical theory of natural and artificial selection. (Part VI, Isolation.). In *Mathematical proceedings of the Cambridge philosophical society* (Vol. 26, No. 2, pp. 220-230). Cambridge University Press. https://doi.org/10.1017/S0305004100015450

Hardy, G. H. (1908) Mendelian proportions in a mixed population. *Science*, *28*, 49–50. https://doi.org/10.1126/science.28.706.49

Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics*. Sinauer associates.

Hebert, P. D., & Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Systematic biology*, *54*(5), 852-859. https://doi.org/10.1080/10635150500354886

Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313-321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313-321. https://doi.org/10.1098/rspb.2002.2218

Hedrick, P. W., & Miller, P. S. (1992). Conservation Genetics: Techniques and Fundamentals. *Ecological Applications*, *2*(1), 30–46. https://doi.org/10.2307/1941887

Hogue, A. S., & Breon, K. (2022). The greatest threats to species. *Conservation Science and Practice*, *4*(5), e12670. https://doi.org/10.1111/csp2.12670

Hudson, R. R. (1983). Properties of a neutral allele model with intragenic recombination. *Theoretical population biology*, *23*(2), 183-201. https://doi.org/10.1016/0040-5809(83)90013-8

Hudson, R. R. (1983). Testing the constant-rate neutral allele model with protein sequence data. *Evolution*, 203-217. https://doi.org/10.2307/2408186

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in ecology & evolution*, *17*(5), 230-241. https://doi.org/10.1016/S0169-5347(02)02489-8

Kern, A. D., & Hahn, M. W. (2018). The neutral theory in light of natural selection. *Molecular biology and evolution*, *35*(6), 1366-1371. https://doi.org/10.1093/molbev/msy092

Kharkov, V. N. (2021). Y-Chromosome Markers in Population Genetics: Fundamental and Applied Results of Ethnogenomic Research. *Russian Journal of Genetics*, *57*(9), 989-1001. https://doi.org/10.1134/S1022795421090040

Kim, K. C., & Byrne, L. B. (2006). Biodiversity loss and the taxonomic bottleneck: emerging biodiversity science. *Ecological research*, *21*(6), 794-810. https://doi.org/10.1007/s11284-006-0035-7

Kimura, M. (1969). The rate of molecular evolution considered from the standpoint of population genetics. *Proceedings of the National Academy of Sciences*, *63*(4), 1181-1188. https://doi.org/10.1073/pnas.63.4.1181

Kimura, M. (1983). Rare variant alleles in the light of the neutral theory. *Molecular Biology and Evolution*, *1*(1), 84-93. https://doi.org/10.1093/oxfordjournals.molbev.a040305

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Kimura, M., & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics*, *49*(4), 725. https://doi.org/10.1093/genetics/49.4.725

Kingman, J. F. C. (1982). The coalescent. *Stochastic processes and their applications*, *13*(3), 235-248. https://doi.org/10.1016/0304-4149(82)90011-4

Kohn, M. H., Murphy, W. J., Ostrander, E. A., & Wayne, R. K. (2006). Genomics and conservation genetics. *Trends in ecology & evolution*, *21*(11), 629-637. https://doi.org/10.1016/j.tree.2006.08.001

Kornet, D. J. (1993). Permanent splits as speciation events: a formal reconstruction of the internodal species concept. *Journal of Theoretical Biology*, *164*(4), 407-435. https://doi.org/10.1006/jtbi.1993.1164

Kreitman, M., & Akashi, H. (1995). Molecular evidence for natural selection. *Annual Review of Ecology and Systematics*, 403-422.

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation biology*, *1*(2), 143-158. https://doi.org/10.1111/j.1523-1739.1987.tb00023.x

Lan, S., Palacios, J. A., Karcher, M., Minin, V. N., & Shahbaba, B. (2015). An efficient Bayesian inference framework for coalescent-based nonparametric phylodynamics. *Bioinformatics*, *31*(20), 3282-3289. https://doi.org/10.1093/bioinformatics/btv378

Lande, R. (1995). Mutation and conservation. *Conservation biology*, *9*(4), 782-791. https://doi.org/10.1046/j.1523-1739.1995.09040782.x

Laporte, M., Perrier, C., Magnan, P., & Berrebi, P. (2016). Genetic evidence of recent migration among isolated-by-sea populations of the freshwater blenny (Salaria fluviatilis). *Conservation Genetics*, *17*(2), 389-399. https://doi.org/10.1007/s10592-015-0791-4

Le Roux, J., & Wieczorek, A. M. (2009). Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology*, *154*(1), 1-17. https://doi.org/10.1111/j.1744-7348.2008.00280.x

Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in ecology & evolution*, *17*(8), 386-391. https://doi.org/10.1016/S0169-5347(02)02554-5

Lewontin, R. C. (1974). *The genetic basis of evolutionary change*. Columbia University Press.

Manel, S., Guerin, P. E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature communications*, *11*(1), 692. https://doi.org/10.1038/s41467-020-14409-7

Mayer, I., & Pšenička, M. (2024). Conservation of teleost fishes: application of reproductive technologies. *Theriogenology Wild*, *4*, 100078. https://doi.org/10.1016/j.therwi.2024.100078

Mayr, E. (1942). *Systematics and the Origin of Species*. The Columbia Classics in Evolution Series.

Mayr, E. (1963). *Animal species and evolution*. Harvard University Press.

Meynell, L., & Lopez, A. (2025). Gaia: the Earth is an organism (not a Darwinian individual). *Biology & Philosophy*, *40*(6), 30. https://doi.org/10.1007/s10539-025-10002-8

Montgomery, M. E., Woodworth, L. M., Nurthen, R. K., Gilligan, D. M., Briscoe, D. A., & Frankham, R. (2000). Relationships between population size and loss of genetic diversity: comparisons of experimental results with theoretical predictions. *Conservation Genetics*, *1*(1), 33-43. https://doi.org/10.1023/A:1010173401557

Mooney, H. A., & Cleland, E. E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences*, *98*(10), 5446-5451. https://doi.org/10.1073/pnas.091093398

Moritz, C. (1995). Uses of molecular phylogenies for conservation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *349*(1327), 113-118. https://doi.org/10.1098/rstb.1995.0097

Moyle, P. B., & Leidy, R. A. (1992). Loss of biodiversity in aquatic ecosystems: evidence from fish faunas. In P.L. Fiedler, S.K. Jain (Eds.), *Conservation biology.* Springer US. https://doi.org/10.1007/978-1-4684-6426-9_6

Muller H. J. (1950) Our load of mutations. *American Journal of Human Genetics*, *2*, 111–176.

Nehring, S. (2011). Invasion history and success of the American blue crab *Callinectes sapidus* in European and adjacent waters. In B. Galil, P. Clark, J. Carlton (Eds.), *In the wrong place-alien marine crustaceans: distribution, biology and impacts*. Springer Netherlands. https://doi.org/10.1007/978-94-007-0591-3_21

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283-292. https://doi.org/10.1086/282771

Nei, M. (1996). Phylogenetic analysis in molecular evolutionary genetics. *Annual review of genetics*, *30*(1), 371-403. https://doi.org/10.1146/annurev.genet.30.1.371

Nikolaou, A., & Katsanevakis, S. (2023). Marine extinctions and their drivers. *Regional environmental change*, *23*(3), 88. https://doi.org/10.1007/s10113-023-02081-8

O'Brien, S. J. (1994). Perspective on conservation genetics. In *Molecular Ecology and Evolution: Approaches and Applications* (pp. 275-280). Basel: Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-7527-1_16

Okasha, S. (2024). Population genetics. *The Stanford Encyclopedia of Philosophy*.

Okazaki, A., Yamazaki, S., Inoue, I., & Ott, J. (2021). Population genetics: past, present, and future. *Human Genetics*, *140*(2), 231-240. https://doi.org/10.1007/s00439-020-02208-5

Panigrahi, M., Rajawat, D., Nayak, S. S., Verma, S. K., Bose, A., Bharia, N., & Dutt, T. (2025). Genetic diversity of the Y-chromosome in domestic animals: from evolutionary insights to functional implications. *Molecular Biology Reports*, *52*(1), 1-18. https://doi.org/10.1007/s11033-025-11049-2

Partridge, L. (1983). Non-random mating and offspring fitness. *Mate choice*, 227-255.

Patwardhan, A., Ray, S., & Roy, A. (2014). Molecular markers in phylogenetic studies-a review. *Journal of Phylogenetics & Evolutionary Biology*, *2*(2), 131.

Pierce, B. A. (2016). *Genetics: A conceptual approach* (6th ed.). W. H. Freeman and Company.

Plough, L. V. (2016). Genetic load in marine animals: a review. *Current Zoology*, *62*(6), 567-579. https://doi.org/10.1093/cz/zow096

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. *Systematic biology*, *55*(4), 595-609. https://doi.org/10.1080/10635150600852011

Posada, D., & Crandall, K. A. (2001). Intraspecific gene genealogies: trees grafting into networks. *Trends in ecology & evolution*, *16*(1), 37-45. https://doi.org/10.1016/S0169-5347(00)02026-7

Puillandre, N., Brouillet, S., & Achaz, G. (2021). ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources*, *21*(2), 609-620. https://doi.org/10.1111/1755-0998.13281

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. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Rannala, B., & Yang, Z. (2020). Species delimitation. In C. Scornavacca, F. Delsuc, & N. Galtier (Eds.), *Phylogenetics in the Genomic Era*.

Ranz, J. M., & Machado, C. A. (2006). Uncovering evolutionary patterns of gene expression using microarrays. *Trends in Ecology & Evolution*, *21*(1), 29-37. https://doi.org/10.1016/j.tree.2005.09.002

Reed, D. H. (2004). Extinction risk in fragmented habitats. *Animal conservation*, *7*(2), 181-191. https://doi.org/10.1017/S1367943004001313

Reed, D. H., Briscoe, D. A., & Frankham, R. (2002). Inbreeding and extinction: the effect of environmental stress and lineage. *Conservation Genetics*, *3*(3), 301-307. https://doi.org/10.1023/A:1019948130263

Reusch, T. B., Ehlers, A., Hämmerli, A., & Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences*, *102*(8), 2826-2831. https://doi.org/10.1073/pnas.0500008102

Ribas, L., & Piferrer, F. (2014). The zebrafish (*Danio rerio*) as a model organism, with emphasis on applications for finfish aquaculture research. *Reviews in Aquaculture*, *6*(4), 209-240. https://doi.org/10.1111/raq.12041

Rosenberg, N. A., & Nordborg, M. (2002). Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics*, *3*(5), 380-390. https://doi.org/10.1038/nrg795

Ryder, O. A. (2005). Conservation genomics: applying whole genome studies to species conservation efforts. *Cytogenetic and genome research*, *108*(1-3), 6-15.https://doi.org/10.1159/000080796

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N., & Weller, S. G. (2001). The population biology of invasive species. *Annual review of ecology and systematics*, *32*(1), 305-332. https://doi.org/10.1146/annurev.ecolsys.32.081501.114037

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*(5), 2075-2087. https://doi.org/10.1007/s12526-017-0785-1

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*(5), 2075-2087. https://doi.org/10.1007/s12526-017-0785-1

Scarpa, F., Cossu, P., Sanna, D., Lai, T., Norenburg, J. L., Curini-Galletti, M., & Casu, M. (2015). An 18S and 28S-based clock calibration for marine Proseriata (Platyhelminthes). *Journal of experimental marine biology and ecology*, *463*, 22-31. https://doi.org/10.1016/j.jembe.2014.10.020

Scarpa, F., Sanna, D., Azzena, I., Mugetti, D., Cerruti, F., Hosseini, S., Cossu, P., Pinna, S., Grech, D., Cabana, D., Pasquini, V., Esposito, G., Cadoni, N., Atzori, F., Antuofermo, E., Addis, P., Sechi, L. A., Prearo, M., Peletto, S., ... Casu, M. (2020). Multiple non-species-specific pathogens possibly triggered the mass mortality in Pinna nobilis. *Life*, *10*(10), 238. https://doi.org/10.3390/life10100238

Scarpa, F., Sernicola, L., Farcomeni, S., Ciccozzi, A., Sanna, D., Casu, M., Vitale, M., Cicenia, A., Giovanetti, M., Romano, C., Branda, F., Ciccozzi, M., & Borsetti, A. (2024). Phylodynamic and evolution of the hemagglutinin (HA) and neuraminidase (NA) genes of influenza A (H1N1) pdm09 viruses Circulating in the 2009 and 2023 seasons in Italy. *Pathogens*, *13*(4), 334. https://doi.org/10.3390/pathogens13040334

Scarpa, F., Sernicola, L., Farcomeni, S., Ciccozzi, A., Sanna, D., Casu, M., Vitale, M., Cicenia, A., Giovanetti, M., Romano, C., Branda, F., Ciccozzi, M., & Borsetti, A. (2024). Phylodynamic and evolution of the hemagglutinin (HA) and neuraminidase (NA) genes of influenza A (H1N1) pdm09 viruses Circulating in the 2009 and 2023 seasons in Italy. *Pathogens*, *13*(4), 334. https://doi.org/10.3390/pathogens13040334

Schlick-Steiner, B. C., Steiner, F. M., Seifert, B., Stauffer, C., Christian, E., & Crozier, R. H. (2010). Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual review of entomology*, *55*(1), 421-438. https://doi.org/10.1146/annurev-ento-112408-085432

Schlötterer, C. (2004). The evolution of molecular markers—just a matter of fashion?. *Nature reviews genetics*, *5*(1), 63-69. https://doi.org/10.1038/nrg1249

Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, *23*(1), 23-35. https://doi.org/10.1017/S0016672300014634

Stanković, S. (1960). *The Balkan Lake Ohrid and its living world* (Monographiae Biologicae, Vol. 9). Dr. W. Junk.

Stephan, W. (2019). Selective sweeps. *Genetics*, *211*(1), 5-13. https://doi.org/10.1534/genetics.118.301319

Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in ecology & evolution*, *15*(5), 199-203. https://doi.org/10.1016/S0169-5347(00)01825-5

Sušnik, S., Snoj, A., Wilson, I. F., Mrdak, D., & Weiss, S. (2007). Historical demography of brown trout (Salmo trutta) in the Adriatic drainage including the putative S. letnica endemic to Lake Ohrid. *Molecular Phylogenetics and Evolution*, *44*(1), 63-76. https://doi.org/10.1016/j.ympev.2006.08.021

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, *105*(2), 437-460. https://doi.org/10.1093/genetics/105.2.437

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., & Vogler, A. P. (2002). DNA points the way ahead in taxonomy. *Nature*, *418*(6897), 479-479. https://doi.org/10.1038/418479a

Templeton, A. R. (1981). Mechanisms of speciation--a population genetic approach. *Annual review of Ecology and Systematics*, 23-48.

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savontaus, M. L., & Wallace, D. C. (1996). Classification of European mtDNAs from an analysis of three European populations. *Genetics*, *144*(4), 1835. https://doi.org/10.1093/genetics/144.4.1835

Underhill, P. A., Passarino, G., Lin, A. A., Shen, P., Lahr, M. M., Foley, R. A., Oefner, P. J., & Cavalli-Sforza, L. L. (2001). The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Annals of human genetics*, *65*(1), 43-62. https://doi.org/10.1046/j.1469-1809.2001.6510043.x

Volz, E. M., Koelle, K., & Bedford, T. (2013). Viral phylodynamics. *PLoS computational biology*, *9*(3), e1002947. https://doi.org/10.1371/journal.pcbi.1002947

Wahlund, S. (1928). Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, *11*(1), 65-106. https://doi.org/10.1111/j.1601-5223.1928.tb02483.x

Weinberg, W. (1908) Uber den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*, *64*, 369–382.

Wheeler, Q. D., Raven, P. H., & Wilson, E. O. (2004). Taxonomy: impediment or expedient?. *Science*, *303*(5656), 285-285. https://doi.org/10.1126/science.303.5656.285

Whitehead, A., & Crawford, D. L. (2006). Variation within and among species in gene expression: raw material for evolution. *Molecular ecology*, *15*(5), 1197-1211.

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annu. Rev. Ecol. Evol. Syst.*, *37*(1), 433-458. https://doi.org/10.1146/annurev.ecolsys.37.091305.110145

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*(2), 97. https://doi.org/10.1093/genetics/16.2.97

Wright, S. (1969). Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press.

Yang, Z., & Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nature reviews genetics*, *13*(5), 303-314.

Yi, S., Zeng, C., Li, Y., & Muangmai, N. (2023). Population genetics and conservation of aquatic species. *Frontiers in Genetics*, *13*, 1052740. https://doi.org/10.3389/fgene.2022.1052740

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, *29*(22), 2869-2876. https://doi.org/10.1093/bioinformatics/btt499

Zhou, Y. (2024). Adaptation or extinction: survival strategies of marine organisms under the pressure of climate change. *International Journal of Marine Science*, *14*, 57-65. https://doi.org/10.5376/ijms.2024.14.0008

Zuckerkandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In B. Vernon, J. V Henry (Eds.), *Evolving genes and proteins* (pp. 97-166). Academic press.

Chapter 1

Pinna nobilis

Pinna nobilis Linnaeus 1758, commonly known as the noble pen shell or fan mussel, is a large marine bivalve species endemic to the Mediterranean Sea (Butler et al., 1993). It belongs to the family Pinnidae (Mollusca: Bivalvia) and is a sessile filter-feeder typically associated with soft-bottom substrates within *Posidonia oceanica* meadows (see Figure 1), where individuals are partially buried in the sediment and anchored to the substratum by byssus threads (Butler et al., 1993; García-March, 2005). Due to its large size and longevity, *P. nobilis* plays an important ecological role in Mediterranean coastal ecosystems, acting both as a flagship and a keystone species through its contribution to ecosystem functioning (Butler et al., 1993; Scarpa et al., 2021; Nebot-Colomer et al., 2022).

Figure 1. Individuals of *Pinna nobilis* in *Posidonia oceanica* meadows. Photo credits: Fabio Scarpa

Throughout its history, *P. nobilis* has been exposed to multiple anthropogenic threats. These include direct exploitation for the harvesting of byssus threads used in the production of precious textiles (sea silk), collection for human consumption (particularly the adductor muscle), and use as fishing bait (Greenwald, 1996; Rabaoui et al., 2010; Basso et al., 2015; Scarpa et al., 2021). In parallel, indirect human activities such as anchoring, pollution, and habitat degradation have further impacted its populations (Vázquez-Luis et al., 2015; Öndes et al., 2020). Consequently, these pressures led to a rapid population decline across the Mediterranean Sea, which peaked during the 1980s (Öndes et al., 2020), resulting in the species being placed under European protection in the early 1990s.

However, since 2016 the species has experienced severe mass mortality events (MMEs) across the Mediterranean basin (e.g., Vázquez-Luis et al., 2017; Catanese et al., 2018; Katsanevakis et al., 2019). These events have been associated with a multifactorial disease caused by different etiological agents (protozoan, bacteria and viruses) acting in combination with environmental stressors, whereas the exact mechanisms are still under investigation (Carella et al., 2019; Scarpa et al., 2020, Prado et al., 2020; Carella et al., 2023, 2024). As a consequence, the conservation status of the species was reassessed and upgraded from "Endangered" to "Critically Endangered" (Kersting et al., 2019). In response to this dramatic decline, several Mediterranean institutions have implemented ex situ conservation programmes, focusing on captive breeding and reintroduction actions, alongside targeted efforts to protect the few remaining wild populations (Kersting et al., 2019; Haberle et al., 2020).

Within this critical conservation context, the research on *Pinna nobilis* presented in this thesis is organised into two main studies aimed at investigating the genetic and evolutionary dynamics of the species in relation to its recent demographic collapse, both based on analyses of the mitochondrial Cytochrome c Oxidase subunit I (COI) gene.

The first study aimed to shed light on the evolutionary history of the species before MMEs and to better understand the evolutionary processes underlying its adaptation to the Mediterranean Sea. Specifically, the objectives were: (i) to reconstruct the phylogeographic patterns and assess levels of genetic variation in *P. nobilis* prior to MMEs; (ii) to test the hypothesis of a genetic boundary between western and eastern Mediterranean populations, proposed to occur eastward of the Sicilian Strait (Sanna et al., 2013); and (iii) to investigate the temporal and geographical origin of *P. nobilis* by integrating phylogeographic and phylogenetic approaches. This latter objective also aimed to provide insights into the evolutionary relationships between *P. nobilis* and other members of the family Pinnidae, particularly its congeneric species *Pinna rudis*.

The second study aimed to further investigate the evolutionary history of *Pinna nobilis* through the analysis of biological samples from ancient byssus glands. These samples, dating back up to approximately three hundred years, offered a unique opportunity to explore temporal changes in the genetic composition of the species and its potential to respond to human-mediated stressors. By comparing ancient specimens with populations sampled both before the MMEs (modern populations) and after the MMEs (surviving populations), this study sought to: (i) examine population dynamics during the early Pleistocene and (ii) assess the effects of long-term environmental and anthropogenic pressures on genetic diversity across different time periods.

References

Basso, L., Vázquez-Luis, M., García-March, J. R., Deudero, S., Alvarez, E., Vicente, N., Duarte, C. M., & 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. https://doi.org/10.1016/bs.amb.2015.06.002

Butler, A., Vicente, N., & de Gaulejac, B. (1993). Ecology of the pterioid bivalves *Pinna bicolor* Gmelin. *Marine Life, 3*, 37–45.

Carella, F., Aceto, S., Pollaro, F., Miccio, A., Iaria, C., Carrasco, N., Prado, P., & 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. https://doi.org/10.1038/s41598-018-37217-y

Carella, F., Palić, D., Šarić, T., Župan, I., Gorgoglione, B., Prado, P., Andree, K. B., Giantsis, I. A., Michaelidis, B., Lattos, A., Theodorou, J. A., Barja Perez, J. L., Rodriguez, S., Scarpa, F., Casu, M., Antuofermo, E., Sanna, D., Otranto, D., Panarese, R., … & 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. https://doi.org/10.1177/03009858231186737

Carella, F., Prado, P., García-March, J. R., Tena-Medialdea, J., Melendreras, E. C., Porcellini, A., & Feola, A. (2024). Measuring immunocompetence in the natural population and captive individuals of noble pen shell Pinna nobilis affected by Pinna nobilis Picornavirus (PnPV). *Fish & Shellfish Immunology*, *151*, 109664. https://doi.org/10.1016/j.fsi.2024.109664

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. (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. https://doi.org/10.1016/j.jip.2018.07.006

Greenwald, M. (1996). *The cruising chef cookbook*. Paradise Cay Publications.

Haberle, I., Marn, N., Geček, S., & Klanjšek, T. (2020). Dynamic energy budget of endemic and critically endangered bivalve *Pinna nobilis*: A mechanistic model for informed conservation. *Ecological Modelling, 434*, 109207. https://doi.org/10.1016/j.ecolmodel.2020.109207

Katsanevakis, S., Tsirintanis, K., Tsaparis, D., Doukas, D., Sini, M., Athanassopoulou, F., Kolygas, M. N., Tontis, D., Koutsoubas, D., & Bakopoulos, V. (2019). The cryptogenic parasite *Haplosporidium pinnae* invades the Aegean Sea and causes the collapse of *Pinna nobilis* populations. *Aquatic Invasions, 14*(2), 150–164. https://doi.org/10.3391/ai.2019.14.2.01

Kersting, D., Benabdi, M., Cizmek, H., Grau, A., Jimenez, C., Katsanevakis, S., Oztürk, B., Tuncer, S., Tunesi, L., & Vázquez-Luis, M. (2019). *Pinna nobilis*. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org

March, J. R. G. (2005). *Aportaciones al conocimiento de la biología de Pinna nobilis Linneo, 1758 (Mollusca: Bivalvia) en el litoral Mediterráneo ibérico* (Doctoral dissertation). Universitat de València, Valencia, Spain.

Nebot-Colomer, E., Vázquez-Luis, M., Boissin, E., Peyran, C., Deudero, S., & Planes, S. (2022). 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. https://doi.org/10.3354/esr01207

Öndes, F., Kaiser, M. J., & Güçlüsoy, H. (2020). Human impacts on the endangered fan mussel *Pinna nobilis*. *Aquatic Conservation: Marine and Freshwater Ecosystems, 30*(1), 31–41. https://doi.org/10.1002/aqc.3237

Prado, P., Carrasco, N., Catanese, G., Grau, A., Cabanes, P., Carella, F., García-March, J. R., Tena, J., Roque, A., Bertomeu, E., Gras, N., Caiola, N., Furones, M. D., & Andree, K. B. (2020). Presence of Vibrio mediterranei associated to major mortality in stabled individuals of Pinna nobilis L. *Aquaculture*, *519*, 734899. https://doi.org/10.1016/j.aquaculture.2019.734899

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. https://doi.org/10.1093/mollus/eyq023

Sanna, D., Cossu, P., Dedola, G. L., Scarpa, F., Maltagliati, F., Castelli, A., Franzoi, P., Lai, T., Cristo, B., Curini-Galletti, M., Francalacci, P., & Casu, M. (2013). Mitochondrial DNA reveals genetic structuring of Pinna nobilis across the Mediterranean Sea. *PLoS One*, *8*(6), e67372. https://doi.org/10.1371/journal.pone.0067372

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? *Veterinaria, 70*(1), 1–14.

Scarpa, F., Sanna, D., Azzena, I., Mugetti, D., Cerruti, F., Hosseini, S., Cossu, P., Pinna, S., Grech, D., Cabana, D., Pasquini, V., Esposito, G., Cadoni, N., Atzori, F., Antuofermo, E., Addis, P., Sechi, L. A., Prearo, M., Peletto, S., ... Casu, M. (2020). Multiple non-species-specific pathogens possibly triggered the mass mortality in Pinna nobilis. *Life*, *10*(10), 238. https://doi.org/10.3390/life10100238

Vázquez-Luis, M., Álvarez, E., Barrajón, A., García-March, J. R., Grau, A., Hendriks, I. E., Jimenéz, S., Kersting, D., Moreno, D., Pérez, M., Ruiz, J. M., Sánchez, J., Villalba, A., & Deudero, S. (2017). SOS Pinna nobilis: a mass mortality event in western Mediterranean Sea. *Frontiers in Marine Science*, *4*, 220. https://doi.org/10.3389/fmars.2017.00220

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. https://doi.org/10.1071/MF14285

Research articles

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	Sanna et al. [14]
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	
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	Sanna et al. [14]
MON	2011	Mondello	11	JX854948–JX854958	
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.

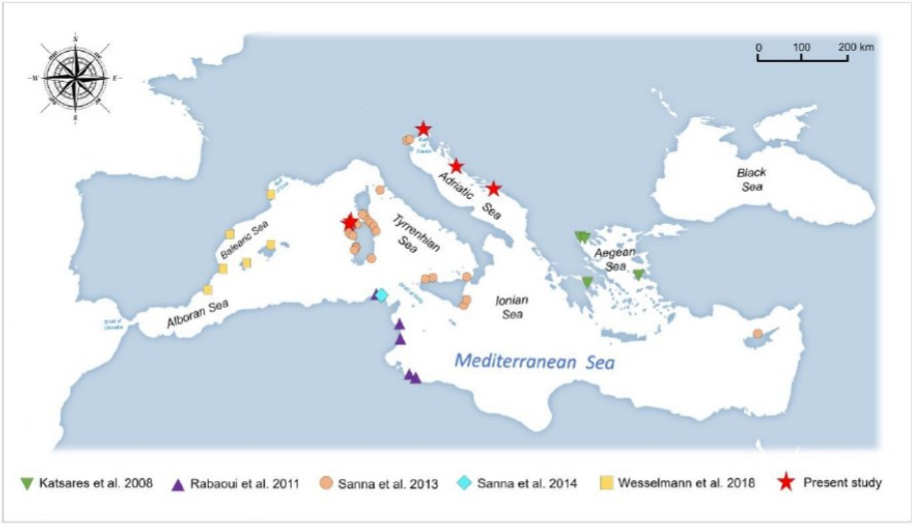


Figure 1. Map of the sampling sites. The map shows the geographical locations for the sequences obtained in the present study along with those from previous research [14,22,25,57,58].

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, Mljlet, 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.

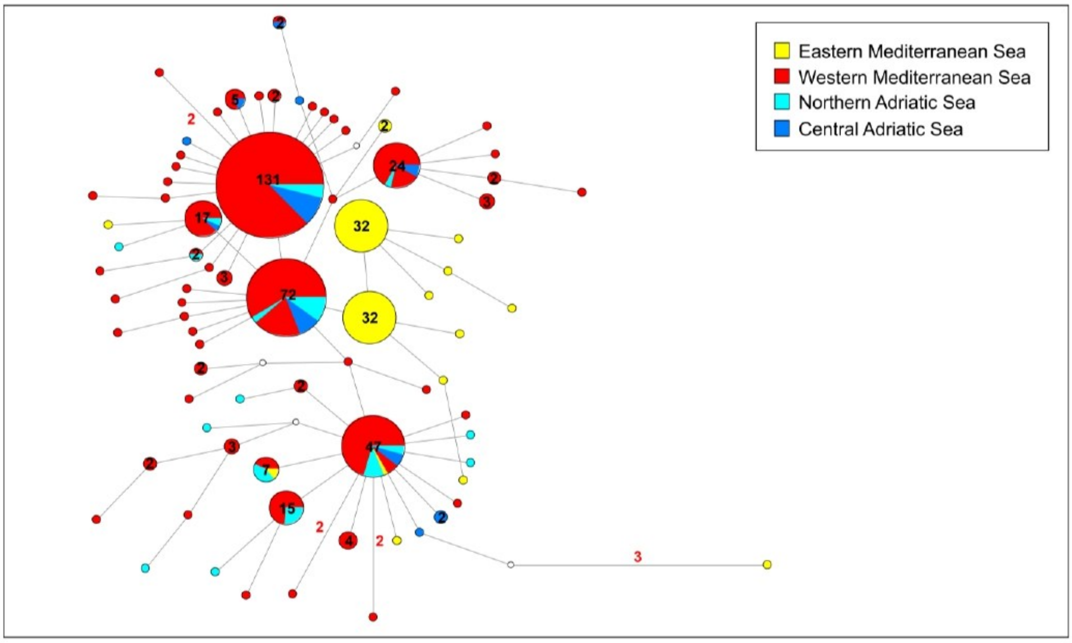


Figure 2. Median-joining network analysis performed on the COI gene fragment. The small white spots on the nodes show median vectors representing the hypothetical sequences that were calculated using the maximum parsimony method. The number of mutations between haplotypes that are greater than n = 1 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).

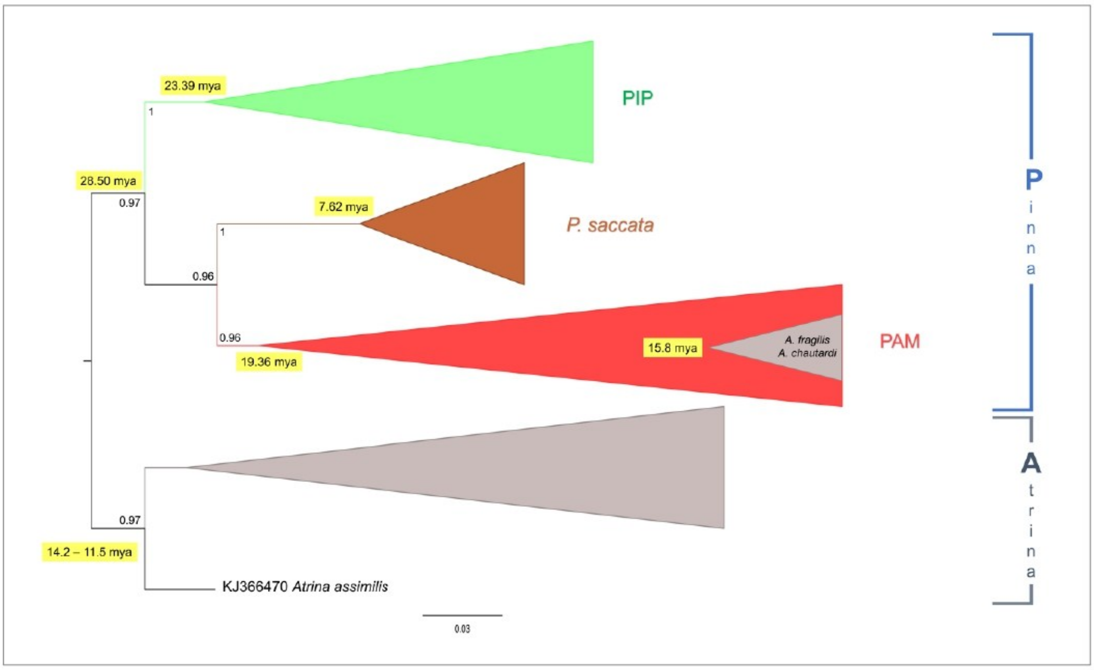


Figure 3. Schematic representation of the Bayesian phylogenetic tree based on the COI gene fragment, which is provided as Figure S2. The values of the node supports are expressed as posterior probabilities.

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.

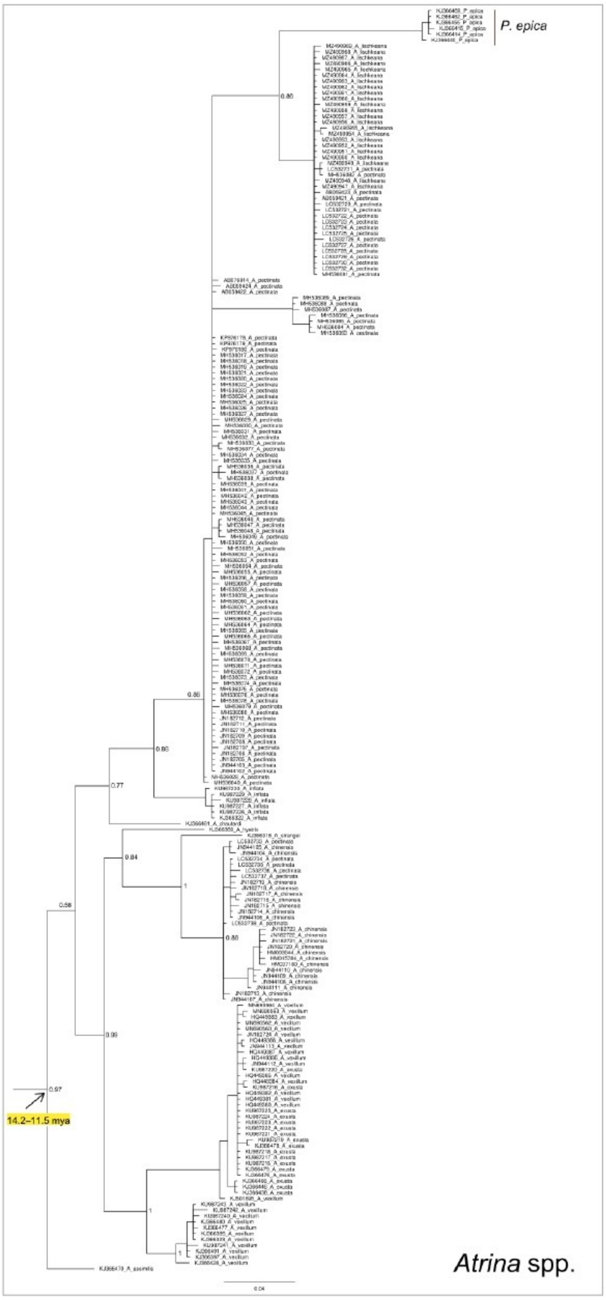


Figure 4. Enlarged detail of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the clade A. The values of the node supports are expressed as posterior probabilities.

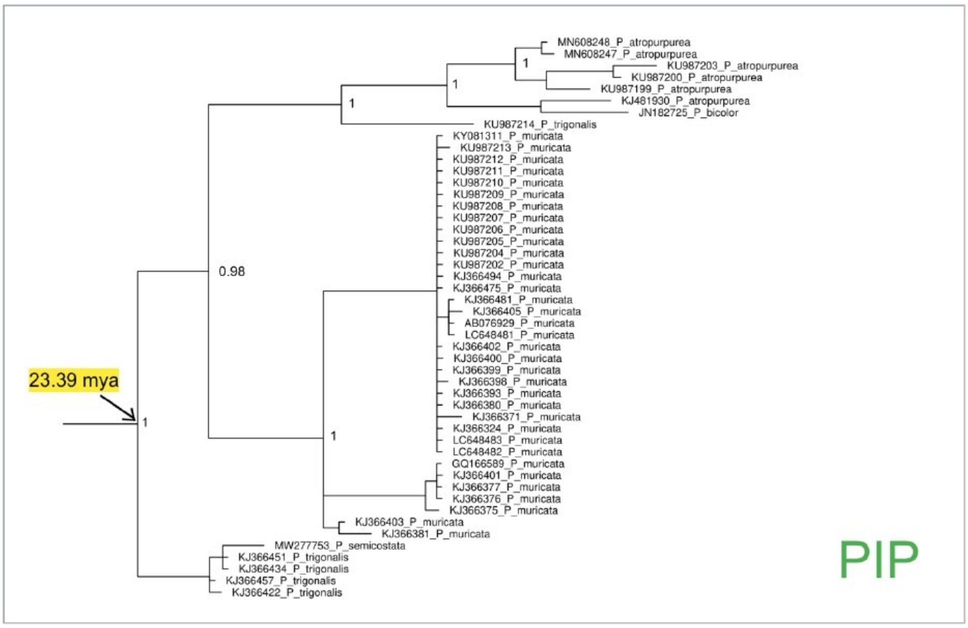


Figure 5. Enlarged detail of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the cluster PIP. The values of the node supports are expressed as posterior probabilities.

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*.

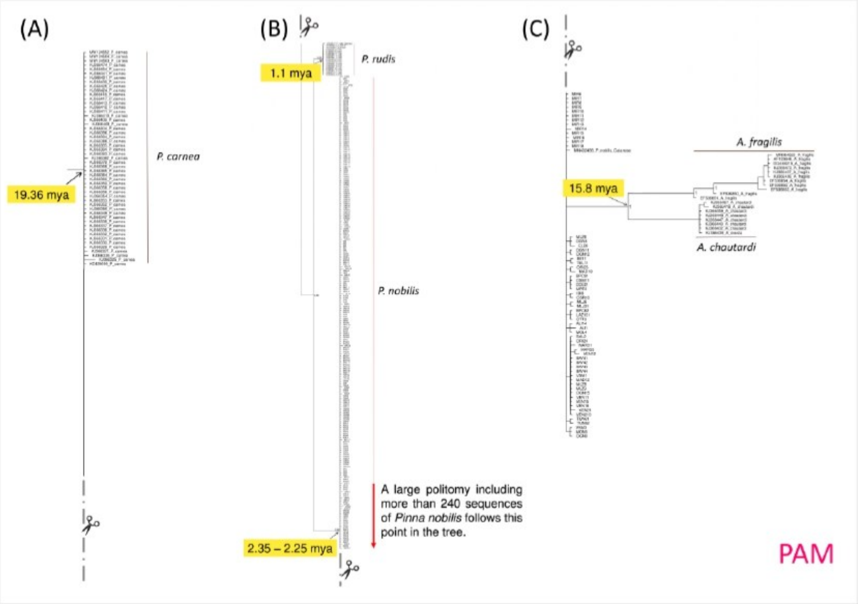


Figure 6. Enlarged details of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the cluster PAM. Due to the large size of the PAM cluster, the corresponding figure was here segmented into three parts (**A**–**C**). The section (**A**) represents all sequences 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.

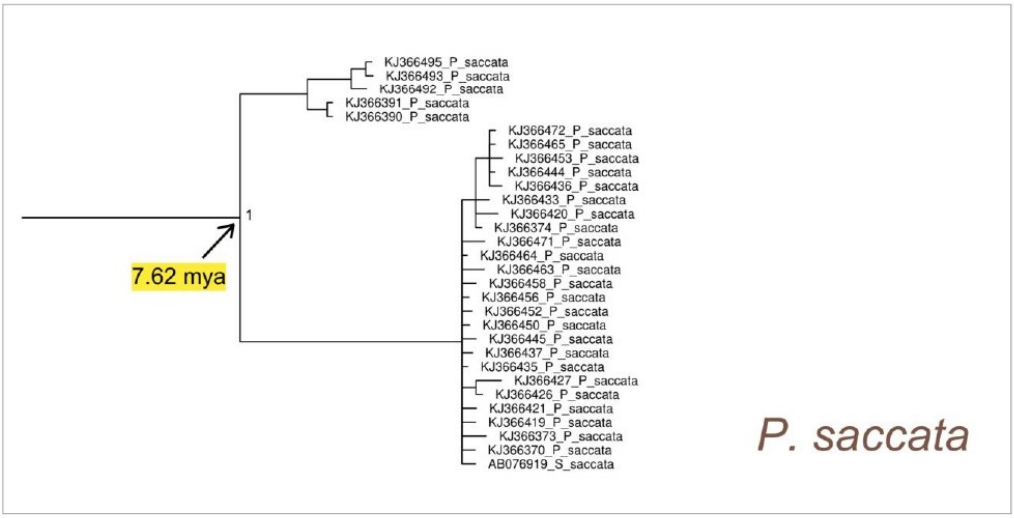


Figure 7. Enlarged detail of the Bayesian phylogenetic tree, which is provided as Supplementary Figure S2, corresponding to the *Pinna saccata* cluster. The values of the node supports are expressed as posterior probabilities.

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.

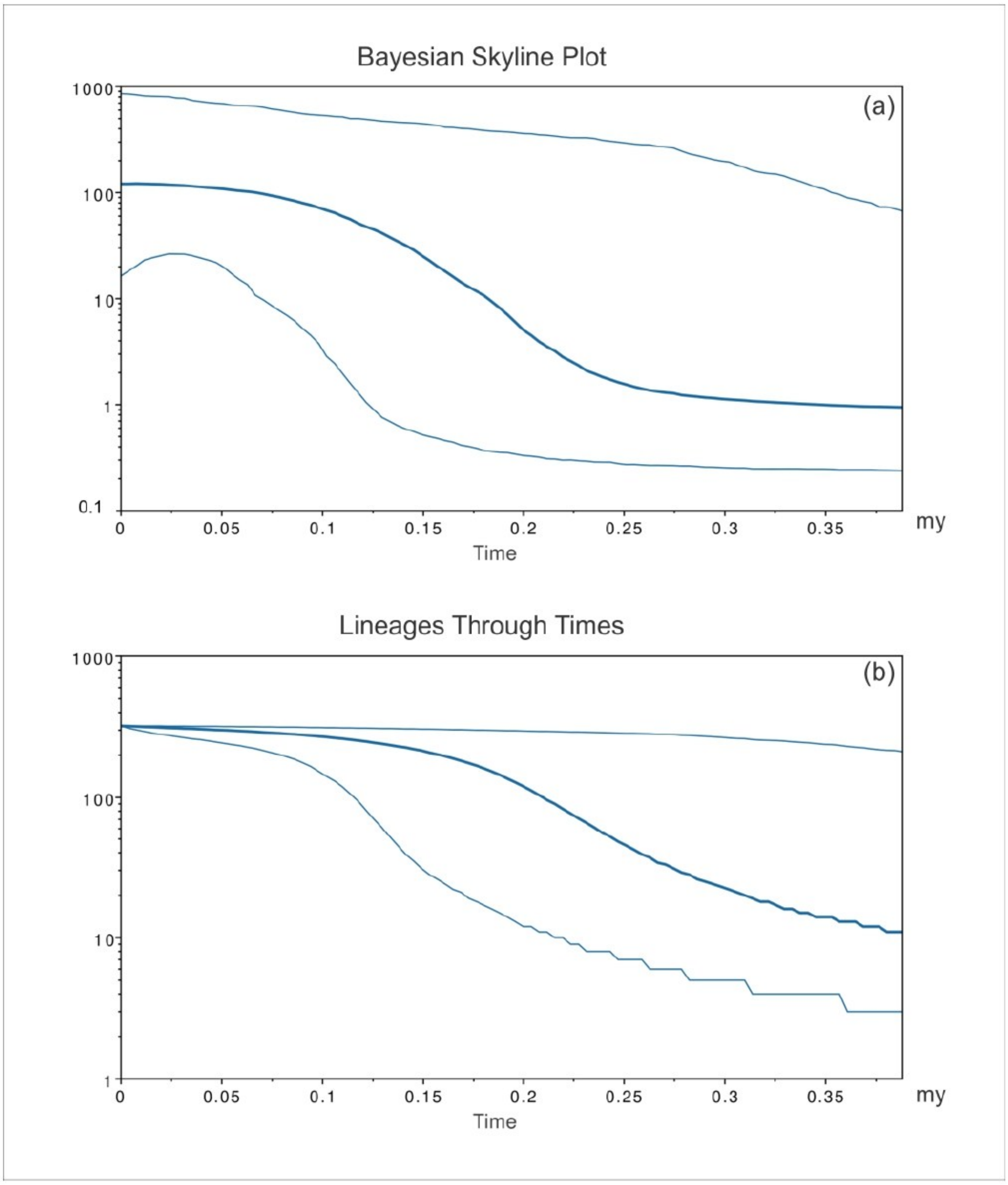


Figure 8. Bayesian skyline plot (**a**) and lineages through time (**b**) for the western Mediterranean populations of *Pinna nobilis*. The effective population size and the number of lineages in the y-axis are shown as a function of million years (my). The thicker central line represents the median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

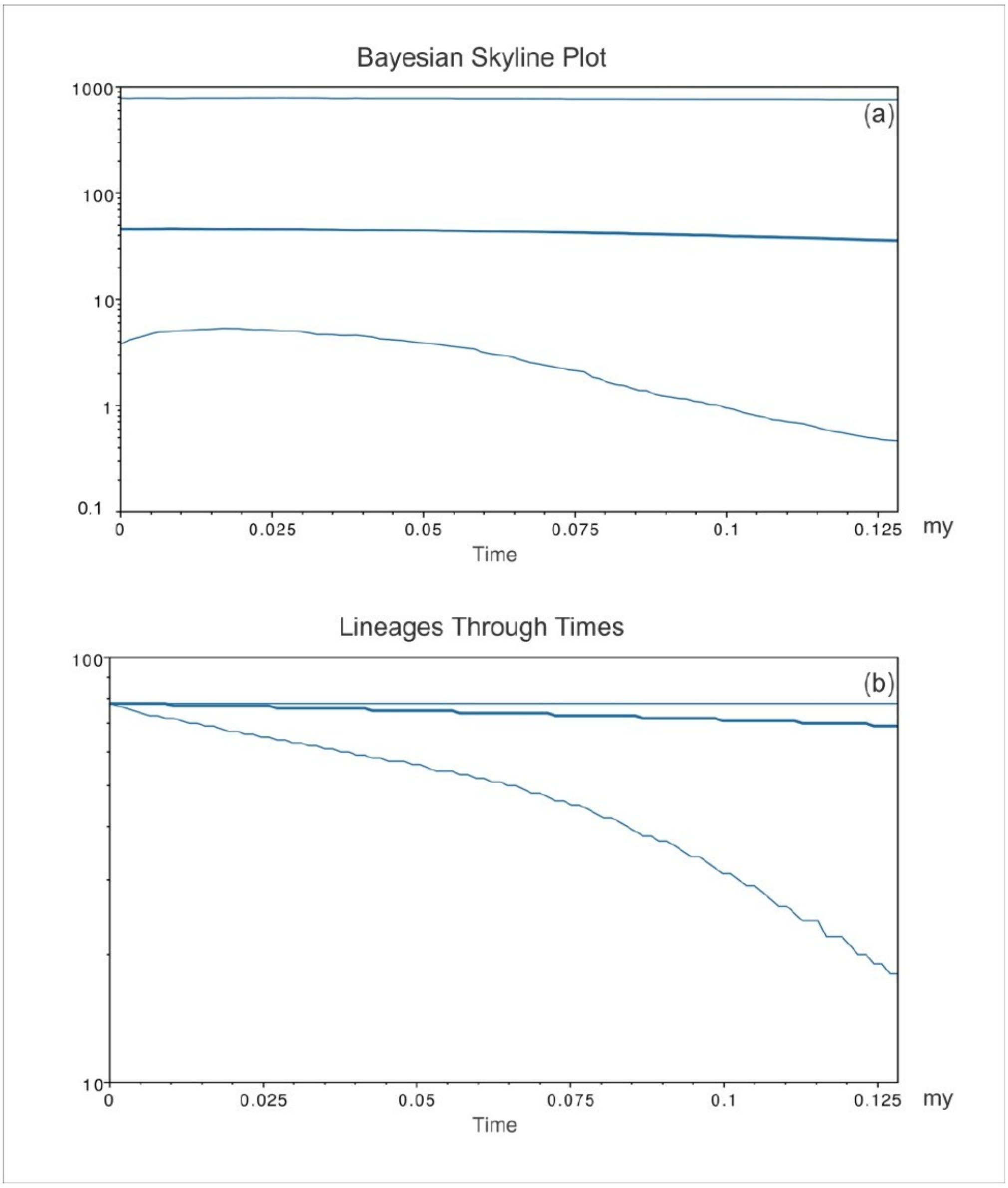


Figure 9. Bayesian skyline plot (**a**) and lineages through time (**b**) for the eastern Mediterranean populations of *Pinna nobilis*. The effective population size and the number of lineages in the y-axis are shown as a function of million years (my). The thicker central line represents the median 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.

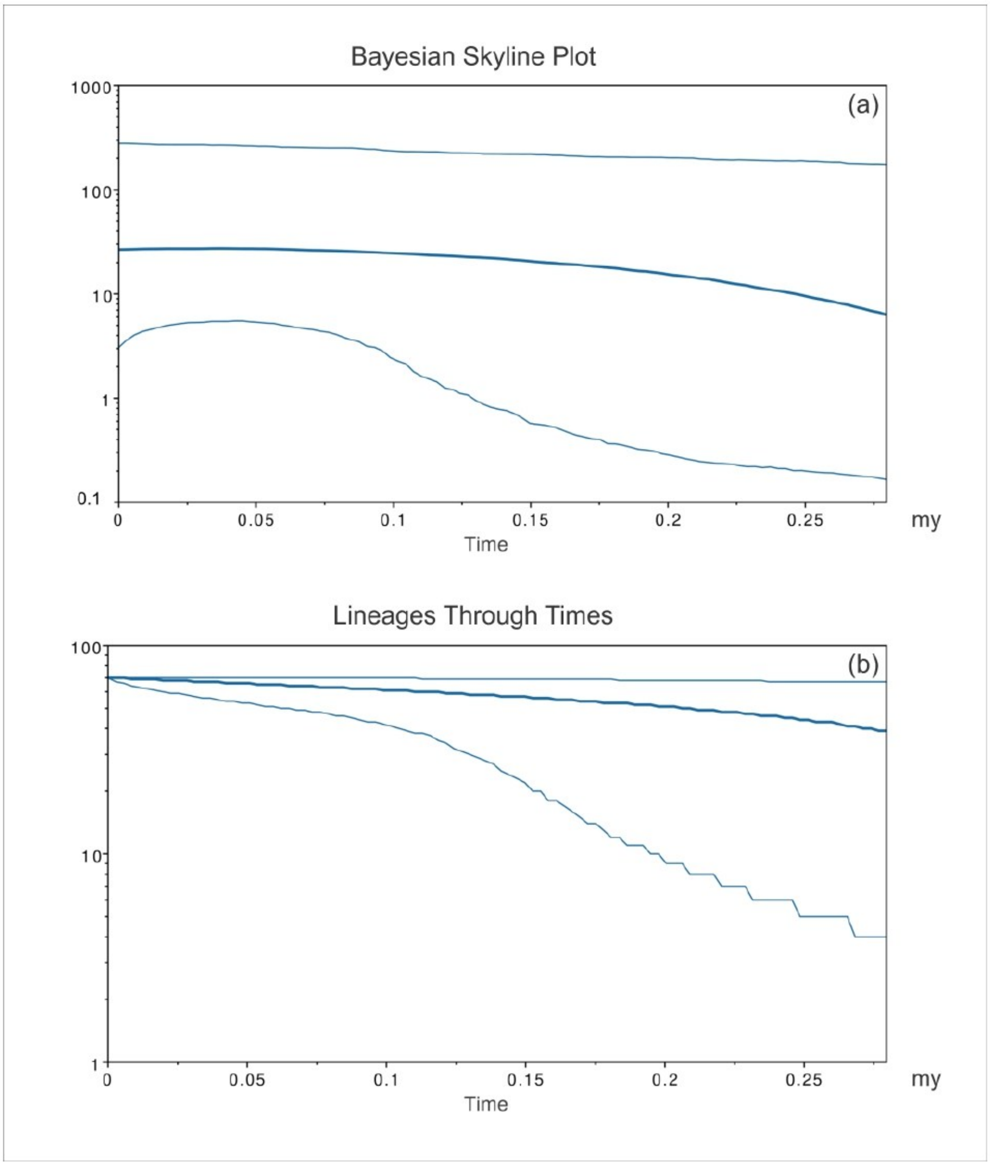


Figure 10. Bayesian skyline plot (**a**) and lineages through time (**b**) for the Adriatic Sea populations of *Pinna nobilis*. The effective population size and the number of lineages in the y-axis are shown as a function of million years (my). The thicker central line represents the median value, while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

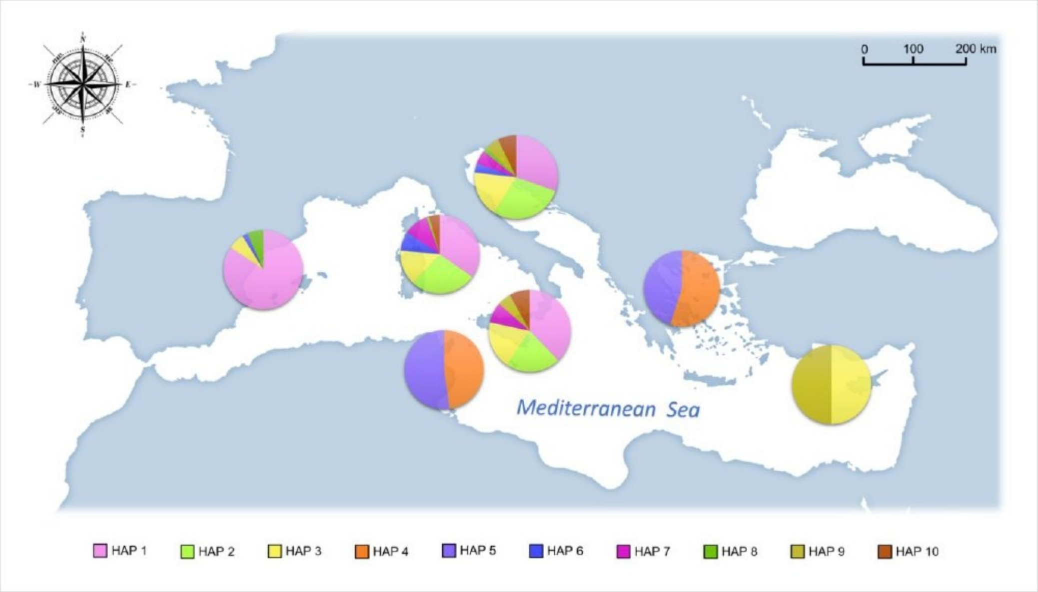


Figure A1. Map showing the distribution of the most common COI haplotypes within the whole Mediterranean basin. In the figure, only haplotypes shared by 5 or more individuals are represented. The number of each haplotype, which is reported in the bottom of the figure, do not refer to a specific nomenclature.

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]

www.nature.com/scientificreports

scientific reports

OPEN

Conservation hints for *Pinna nobilis* from a century-old genetic time capsule

Ilenia Azzena[1,15], Chiara Locci[1,2,15], Noemi Pascale[1,2,3], Ilaria Deplano[2], Riccardo Senigaglia[1], Edoardo Batistini[4], Daniela Caracciolo[5], Mariachiara Chiantore[6,7], Saul Ciriaco[4,8], Maria Paola Ferranti[6,7], Daniele Grech[9], Arianna Liconti[10], Monica Montefalcone[6,7], Alice Oprandi[6], Valentina Pitacco[11], Marco Segarich[4], Rym Zakhama-Sraieb[12], Ahmed Ben Hmida[13], Salma Zribi[14], Fabio Scarpa[2], Marco Casu[1,16] & Daria Sanna[2,16]✉

The noble pen shell, *Pinna nobilis*, is an iconic marine bivalve endemic to the Mediterranean Sea, playing a key role as an ecosystem engineer. Over the past century, it has faced severe threats from overharvesting, pollution, and catastrophic mass mortality events. This study analysed 119 mitochondrial COI gene sequences from historical (1700s, 1920s, 1970s, 1990s) and modern (2000s) samples, including survivors of recent mass mortality crises. We standardised a protocol to extract DNA from ancient byssus samples over a century old and dated the emergence of the mitochondrial lineages of *Pinna nobilis*, uncovering its evolutionary history in unprecedented detail. Our findings suggest two main temporal origins for the species' genetic variation: (i) a group of modern lineages directly descended from *Pinna nobilis* early ancestors originating 2.5 mya, and (ii) a large group derived from the first Pleistocene radiation of the species, approximately 1.5 mya. Importantly, our research depicts the evolutionary response of *Pinna nobilis* to three major challenges in the last century: human overexploitation, pollution, and environmental changes. Our results highlight the species' remarkable resilience, likely mediated by Pleistocene genetic traits, whose persistence over time mainly depends on the maintaining of a high effective population size to ensure successful recruitment.

Pinna nobilis (Linnaeus, 1758), known as the noble pen shell, is a key species endemic to the Mediterranean Sea. It belongs to the family Pinnidae (Mollusca: Bivalvia)[1] and is commonly found in coastal waters at depths between 0.5 and 60 m, favouring soft-bottom habitats, especially those with mixed seagrass meadows. It is primarily associated with *Posidonia oceanica* beds, where its filter-feeding activity contributes significantly to ecosystem functioning[1].

Pinna nobilis supports marine ecosystems by providing surfaces for organisms to attach to, filtering up to 2500 Liters of water daily, and sustaining symbiotic crustaceans and local food webs[1–3]. Historically, *Pinna nobilis* was renowned for its byssus, used to produce the sea silk (a luxurious textile), as well as its muscle (a delicacy) and decorative shell. Overharvesting for these purposes led to a decline in *Pinna nobilis* populations during the twentieth century[1,4,5].

To protect the species, a complete ban on harvesting was introduced in 1992 across European countries (Annex IV of European Council Directive 92/43/EEC). This strict protection regime led to a gradual increase in

[1]Department of Veterinary Medicine, University of Sassari, Via Vienna 2, 07100 Sassari, Italy. [2]Department of Biomedical Sciences, University of Sassari, Viale San Pietro 43B, 07100 Sassari, Italy. [3]Department of Chemical, Physical, Mathematical, and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy. [4]Shoreline Soc. Coop., AREA Science Park, Località Padriciano 99, 34149 Trieste, Italy. [5]ARPAL, Regional Agency for the Environmental Protection Liguria, Genova, Italy. [6]DiSTAV (Department of Earth, Environment and Life Sciences), University of Genoa, Corso Europa, 26, 16132 Genova, Italy. [7]National Biodiversity Future Center (NBFC), Palermo, Italy. [8]WWF Fondazione, Area Marina Protetta Di Miramare, Via Beirut 2/4, 34151 Trieste, Italy. [9]IMC - International Marine Centre, Loc. Sa Mardini, 09170 Torregrande, Oristano, Italy. [10]OutBe, Via Nino Bixio 19, 16043 Chiavari, Genova, Italy. [11]Marine Biology Station Piran, National Institute of Biology, Fornače 41, 6330 Piran, Slovenia. [12]Institut Supérieur de Biotechnologie de Sidi Thabet, University of Manouba, Sidi Thabet, Tunisia. [13]Coastal Protection and Development Agency (APAL), 5000 Monastir, Tunisia. [14]Faculty of Sciences of Tunis, Laboratory of Diversity, Management and Conservation of Biological System, University of Tunis El Manar, LR18ES06, Tunis, Tunisia. [15]Ilenia Azzena and Chiara Locci contributed equally to this work as first authors. [16]Marco Casu and Daria Sanna contributed equally to this work as senior authors. ✉email: darsanna@uniss.it

population size in the Mediterranean. As a result of this recovery, researchers were able to investigate the genetic variability of *Pinna nobilis* throughout the basin[6–14]. Initial studies on *Pinna nobilis* focused on mitochondrial markers[6,7,11] revealing low genetic differentiation among populations in the Eastern Mediterranean, suggesting a lack of strong genetic structuring in that area[6]. Subsequent broader assessment across the Mediterranean[11] identified patterns of genetic variability that appeared to follow three main marine ecoregions: (1) the western Mediterranean and Ionian Sea, (2) the Adriatic Sea, and (3) the Aegean Sea along with the Tunisian coasts. However, a more recent study by Sanna et al.[14], which included a significantly larger number of samples from the previously underrepresented Adriatic Sea, revealed genetic homogeneity between Adriatic and Western Mediterranean populations, indicating a higher degree of interregional connectivity than previously thought.

Later genetic research incorporated microsatellite markers[8–10,12,13], sometimes in combination with mitochondrial data[8], and confirmed both a general high genetic variability for the species and low inter-population differentiation. These findings support the existence of a single genetic lineage that likely experienced a recent population expansion[10]. This hypothesis is further supported by the most recent work by Sanna et al.[14], which suggests that *Pinna nobilis* diverged from an Atlantic ancestor approximately 2.5 million years ago, likely colonizing the Western Mediterranean following the Zanclean flood. The species then expanded into the Adriatic and Eastern Mediterranean, shaping its current genetic distribution and structure.

Despite the recovery experienced in the early 2000's, *Pinna nobilis* populations faced a new threat in 2016 in the form of Mass Mortality Events (MMEs), which severely reduced their numbers. These events began along the Spanish coasts and spread across the Mediterranean[15]. They were initially attributed to the parasite *Haplosporidium pinnae*[16] however, further research revealed that the die-offs were likely driven by a complex interplay of pathogens, including bacteria and viruses, compounded by environmental stressors[17–21]. As populations declined, the surviving *Pinna nobilis* individuals were confined to isolated refuges in coastal lagoons and estuaries across the coastlines of Spain, France, Italy, Greece, and Turkey[22–31]. However, the reasons why certain populations have managed to survive, while others have not, remain unclear, warranting further investigation.

In response to this crisis, large-scale conservation projects were launched, with the European LIFE PINNA project standing out. This effort aims to protect the remaining populations and restore affected areas through measures such as breeding *Pinna nobilis* in captivity and transplanting individuals to safer areas. "Citizen Science" initiatives have also played a key role, with divers and local communities helping locate surviving populations, including a resilient group found near Sant'Antioco Island in Sardinia, Italy (Western Mediterranean).

To gain deeper insights into the species evolutionary history, we analysed ancient byssus samples, obtained through the collaboration of two Sardinian weavers, Assuntina Pes and Giuseppina Pes, who continue to work with the byssus threads from the collection of Italo Diana, one of the most important Sardinian byssus weavers—a legacy of the atelier he founded in the 1920s in Sant'Antioco—and of the so-called "sea silk master" weavers Chiara Vigo. These samples, up to 300 years old, provided a unique opportunity to study how the genetic make-up of *Pinna nobilis* has evolved over time. By comparing ancient samples with those from modern (before MMEs[6,7,10,11,14,24,32,33]) and surviving populations (after MMEs[24,33]), valuable insights into genetic changes and adaptations were gained.

In this context, the focus of this study is twofold: first, it aims to understand the population dynamics of *Pinna nobilis* during the early Pleistocene; second, it examines the impact of human and environmental stressors in the species' recent history. To enable a comparison among ancient, modern, and surviving populations, these authors relied on the Cytochrome c Oxidase Subunit I (COI) gene as the main tool for this research, as it is the only molecular marker that allows the broadest comparisons with sequences[6,7,10,11,14,24,32,33] from different areas and time periods.

Through a combination of genetic knowledge and dedicated conservation efforts, there is hope that this iconic species will continue to thrive in the Mediterranean for generations to come.

Results

Phylogenetics and phylogeographic analyses

The whole dataset, consisting of 667 sequences, was divided into three main cohorts: ancient populations (1920s, 1970s, 1990s, and one individual from 300 years ago), modern populations (early 2000s, pre-MMEs), and survivor populations (post-MMEs). The 119 newly generated sequences were obtained from three areas: Sant'Antioco Island (Sardinia, Italy), which included ancient, modern and survivor populations; Tunisia and North Adriatic Sea, which included survivor populations.

Genetic variation analysis across the whole Mediterranean dataset revealed 59 polymorphic sites, resulting in 100 haplotypes. Among the three temporal cohorts, the ancient one, exhibited the lowest rates of genetic variation in terms of haplotype (h) and nucleotide diversity (π) (see Table 1). However, it should be taken into consideration that such a reduced diversity in the ancient group may be explained by the fact that modern and surviving cohorts are represented by a larger number of individuals from both eastern and western Mediterranean populations, while the ancient cohort is exclusively represented by samples from one western population.

In contrast, between the early 2000s (moderns) and 2024 (survivors) (see Table S1 for details), the species has maintained high genetic variation, with a slight decrease in the survivor populations (see Table 1). However, the survivors showed reduced polymorphic sites (S), and haplotypes (H) compared to modern populations, potentially influenced by the larger effective population size of modern individuals.

The ancient cohort from Sant'Antioco Island was categorized into three groups based on the sampling period: 1920s, 1970s, and 1990s. The individuals sampled in 1920s showed the lowest level of genetic variability (see Table 1). In contrast, individuals from the 1970s and 1990s exhibited higher levels of genetic variability, despite fewer available sequences from the 1990s. Notably, individuals from the 1970s displayed the highest genetic variability within the ancient cohort, with six haplotypes identified out of a total of 7 (see Table 1).

Samples	**N**	**S**	**H**	**h**	**π**
Ancients—Sant'Antioco Island (Western Mediterranean)					
1920'	17	2	3	0.412	0.00131
1970'	20	6	6	0.721	0.00366
1990'	6	2	3	0.733	0.00276
Total ancients	44*	7	7	0.613	0.00263
Moderns					
Total moderns	552	55	92	0.892	0.00675
Survivors					
Total survivors	71	13	16	0.837	0.00525

Table 1. Sample sizes and genetic diversity estimates for the mitochondrial COI gene fragment analysed in *Pinna nobilis* individuals from various Mediterranean regions across distinct time periods: N: sample size; S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; π: nucleotide diversity. *Ancient dataset from Sant'Antioco also included 1 sample of 300 years ago, which could not be incorporated into the other groups or treated separately as it corresponds to only one individual.

A subset of 100 sequences, corresponding to all *Pinna nobilis* COI haplotypes identified in this study and shared among ancient, modern, and survivor cohorts, was used to perform a time-calibrated tree analysis (see Fig. 1 and Figure S1 in supplementary materials for details on molecular dating).

The time calibrated tree (see Fig. 1 and Figure S1), generated with BEAST, indicates that the ancestor of *Pinna nobilis* originated in the Mediterranean ~ 2.5 mya.

The common ancestor gave rise to two clades: Clade A and Clade B. Clade A, which diverged ~ 860 kya, comprises 9 of the 100 haplotypes (9%) directly descending from the early ancestor of *Pinna nobilis*. This clade (A) was once widespread across the whole Mediterranean (Western Mediterranean, Adriatic Sea, and Eastern Mediterranean), but most haplotypes seem to have gone extinct after MMEs, or at least they are extremely rare and not detected in the sampled area of this study, except H6 (Fig. 1). H6, the sixth most common haplotype, persists in survivor populations from the Western Mediterranean (Mar Menor, Spain) and North Adriatic Sea (Trieste and Miramare, Italy). Notably, 55.5% of Clade A haplotypes diverged during the Donau-Günz interglacial (949–681 kya; see the red-underlined haplotype in Clade A, Fig. 1). Additionally, H6, along with the modern haplotype from the Western Mediterranean (Isola Piana, Sardinia, Italy) HAP73, are among the most recently originated haplotypes. HAP73, which now could be extinct or extremely rare and not detected in this area of the Mediterranean, and H6, diverged during the post-Würm period ~ 6 kya (see Clade A in Fig. 1).

Clade B includes 91% of all haplotypes and exhibits two main subclades, B1 and B2 (see Fig. 1), that share a common ancestor which originated 1.55 mya. Subclade B1, which differentiated 1.20 mya, comprises 62% of the whole lineages. Within this subclade, two main, almost contemporary subgroups are included. Among them, the subgroup that differentiated 0.93 mya includes the greatest number of haplotypes, but proportionally, it also groups the highest number lineages that could be now extinct or extremely rare.

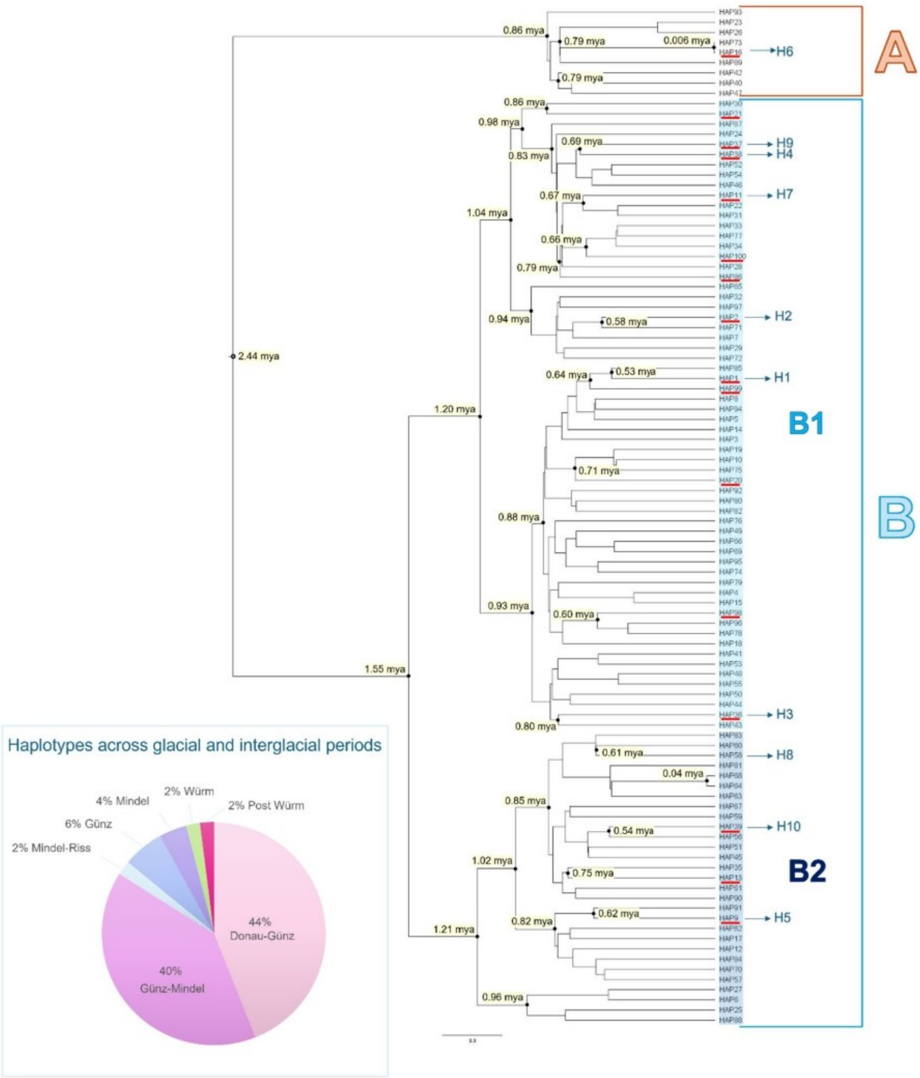
Subclade B2, which originated 1.21 mya, accounts for 29% of haplotypes. Subclade B1 includes 6 of the 10 most common haplotypes (H1, H2, H3, H4, H7, and H9 in Fig. 1). Among these, haplotypes H3, H4, and H9 diverged during the Donau-Günz interglacial (~ 800 kya for H3, ~ 690 kya for H4 and H9), preceding the first major Quaternary glaciations (Günz, Mindel, Riss, and Würm). Haplotype H7 diverged during the Günz glacial period (~ 670 kya), while the most widespread H1 and H2 arose during the Günz-Mindel interglacial (~ 530 kya and 580 kya respectively). Haplotype H1 was also retrieved in a 300-year-old sample from Sant'Antioco Island (Western Mediterranean), suggesting that it was among the most common haplotypes historically. Moreover, within subclade B1, 12 haplotypes persist in surviving Mediterranean populations (see red-underlined haplotypes in subclade B1, Fig. 1), representing 75% of total surviving haplotypes. Of these 12 haplotypes, 6 are among the most common (H1-H4, H7 and H9 in subclade B1, Fig. 1), with the divergence times previously reported. In particular, haplotypes H2 and H7 are mainly found in the Western Mediterranean and the Adriatic Sea, while H3, H4, and H9 are prevalent in the Eastern Mediterranean and the Ionian Sea. Notably, haplotype H1 is primarily distributed in the Western Mediterranean and Adriatic Sea, but it has also been detected in two survivor individuals from the Tunisian Kerkennah Islands in the Eastern Mediterranean.

Whereas the remaining 6 haplotypes diverged during the following periods: 3 during the Donau-Günz (~ 949–681 kya), 1 during the Günz-Mindel (~ 619–456 kya), and 2 during the Günz (~ 680–620 kya).

Subclade B2 includes the three remaining prevalent haplotypes (H5, H8, and H10, Fig. 1), originating during the Günz-Mindel interglacial period and broadly distributed across the Mediterranean Sea. Among these, haplotypes H5 and H10 persist in surviving populations. Specifically, H5 is found in the Western Mediterranean (Sant'Antioco Island) and the Northern Adriatic Sea, while H10 is distributed in the Ionian Sea (Amvravikikos Gulf, Greece). In contrast, haplotype H8 appears extinct or extremely rare and not detected.

Similar to Clade A, Clade B includes two recently diverged haplotypes, HAP68 and HAP64 (Fig. 1), which diverged during the Würm glacial period (~ 41 kya). These haplotypes, found in modern populations from Sicily and Elba Islands (Western Mediterranean), now appear extinct or extremely rare and not detected.

In summary, most haplotypes identified in *Pinna nobilis* populations across ancient, modern, and surviving individuals diverged during the Donau-Günz (44%) and Günz-Mindel (40%) interglacial periods (see inbox in Fig. 1).



Posterior probability values for all nodes in the tree are > 0.9

Fig. 1. Time-calibrated tree based on the mitochondrial COI gene fragment. All nodes are highly supported with value exceeding 0.9. Node values indicate divergence times, expressed in millions of years ago (mya). The 100 haplotypes are labelled with the code "HAP". Haplotypes underlined in red represent those still present in the surviving Mediterranean populations. The 10 most common haplotypes are indicated with the code "H". The inbox in the bottom left shows the percentage of haplotypes that evolved during different time periods.

The network analysis (see Fig. 2) described the phylogeographic patterns of the 100 haplotypes identified within the whole dataset.

Ten of these haplotypes are the most prevalent (H1–H10, Fig. 2), forming at least three main star-like patterns. These ten haplotypes are defined by a total of eight mutations. Of these, seven mutations are silent, resulting in no change to the amino acid sequence. The remaining mutation, located at position 62 of the analysed COI fragment and found exclusively in haplotype H8, is a missense mutation. This substitution leads to a change in the amino acid sequence, replacing glycine with glutamic acid.

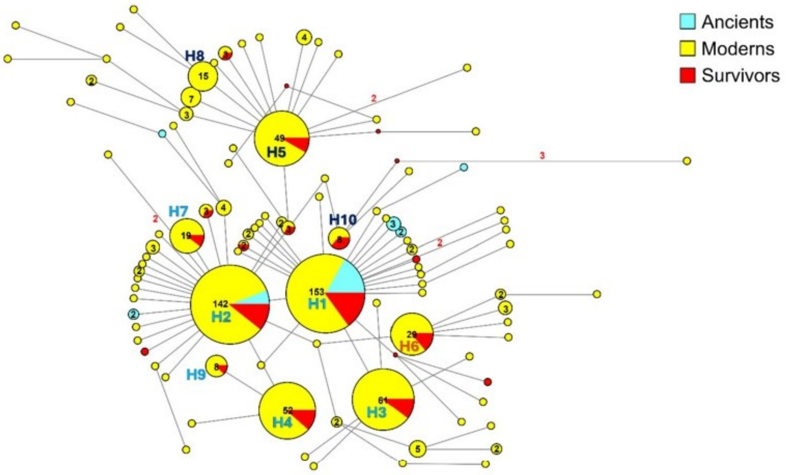


Fig. 2. Median-joining network analysis performed on the mitochondrial COI gene fragment for *Pinna nobilis* populations over time, with the resulting haplotypes color-coded based on their sampling period. The small red spots on the nodes show median vectors representing the hypothetical sequences that 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. Haplotype codes (H1–10) are reported on the network with the colours used in the phylogenetic tree (Fig. 1) to indicate Clade A and Subclades B1 and B2.

The two most common haplotypes, H1 and H2, are shared across all sampled periods—ancient, modern, and survivor—comprising 22.9% and 21.3% of the whole dataset, respectively. In particular, they were present in 79.5% of the ancient cohort. The remaining 20.5% of ancient samples show haplotypes exclusive to this cohort, which probably went extinct in modern and survivor populations. Overall, 71.4% of ancient haplotypes appear to have no remaining descendants in surviving populations.

The two most common haplotypes among the Eastern and Ionian samples (H3 and H4 in Fig. 2) are directly derived from the two most common haplotypes in the dataset (H1 and H2 in Fig. 2).

Haplotypes H3 and H4 differ from H1 and H2 due to two main single-nucleotide polymorphisms (SNPs) within the last 25 nucleotides of the analysed COI fragment. Specifically, one polymorphism is prevalent in nearly all eastern individuals, resulting in a silent mutation at the third base of a codon encoding glycine, involving a transition between the purine bases A and G. This mutation does not cause any amino acid changes in the protein.

The Mantel test, used to test the hypothesis of the isolation by distance (IBD) among Mediterranean sites, yielded a correlation coefficient of $r = 0.15$, with a very high level of significance ($p = 9.999 \times 10^{-5}$, based on 10,000 permutations). The absolute value of the correlation coefficient ($r = 0.15$) indicates that the strength of the association is relatively weak (Fig. 3a). This observed value was significantly higher than the quantiles of the null distribution, generated through random permutations of the matrices (i.e. 90% = 0.0289; 95% = 0.0376; 97.5% = 0.0448; 99% = 0.0540). Such a result confirms that the observed correlation between genetic distance and geographic distance is not compatible with a random model that would have supported the presence of IBD among the *Pinna nobilis* sampling sites in the present study: indeed, a non-random spatial structure in the distribution of genetic variability was found for the dataset which was analysed.

Finding obtained by the Mantel test highlight the existence of a weak IBD among Mediterranean sites for *Pinna nobilis* and suggest that factors other than the geographic distance, (e.g., historical events, geographical barriers, demographic or selective effects) account for the observed genetic structure in accordance with a complex biological system subjected to multiple influences.

In this context, Fig. 3b shows how the first two principal axes (85.71% of the total variance explained), obtained for the spatial PCA (sPCA), explain the limited portion of the IBD retrieved by Mantel test (Fig. 3a). Axis 1 (63.5%) reveals a spatial structure with an east–west gradient, suggesting major spatial genetic differentiation. Indeed, this axis highlights a spatial structuring, with a concentration of positive values (in red) in the eastern part of the Mediterranean basin, suggesting significant geographical differentiation among sampling sites, likely due to ecological barriers or environmental factors. In contrast, axis 2 (22.3%) captures a weaker, less structured component of variability, potentially reflecting local processes or stochastic patterns not related to geographic distances. This axis shows weaker and more diffuse variation, indicating potential local or secondary differences not associated with a clear geographical gradient.

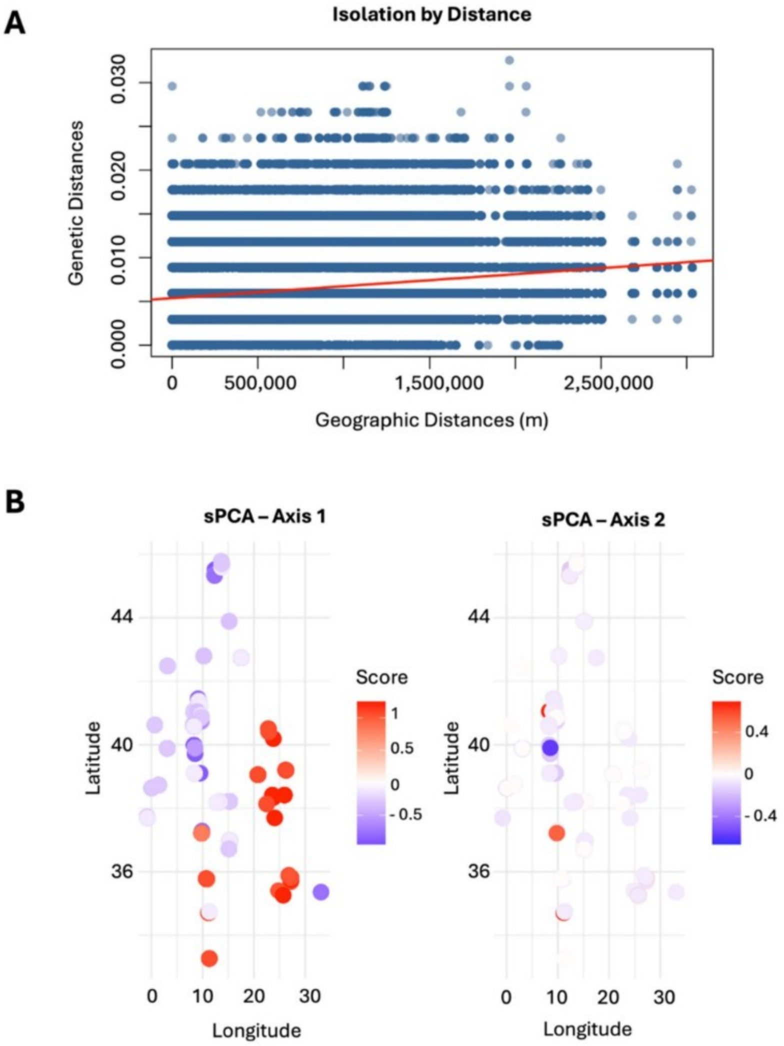


Fig. 3. Analysis of spatial genetic structure. (**A**) Plot of the results obtained for the isolation by distance that show the relationship between genetic distance (Y-axis) and geographic distance (X-axis, in meters): the linear regression (red line) indicates a weak, although significant positive correlation. (**B**) Spatial Principal Component Analysis (sPCA). Scores represent combinations of allele frequencies that maximize both genetic variance and spatial autocorrelation.

Discussion

This study provides a comprehensive overview of the evolutionary history of the critically endangered Mediterranean bivalve *Pinna nobilis*. By comparing genetic data from ancient, modern, and surviving populations, we examined how genetic diversity, and evolutionary trends have shifted over the past century. Our findings shed new light on how *Pinna nobilis* has managed to survive major environmental changes, human overexploitation, and the most recent MMEs that have dramatically decreased its populations.

In this context, we found that the mitochondrial genetic structure of the species has remained stable, with key lineages (haplotypes) persisting across ancient, modern, and surviving populations, likely because they provide survival advantages.

Furthermore, the analyses conducted in the present study did not detect signals of isolation by distance, corroborating the results obtained in previous studies[10,11,32] and suggesting effective larval dispersal for *Pinna* *nobilis*. However, rather than being correlated with geographic distances, the observed genetic structure appears to reflect the evolutionary dynamics resulting from the early Pleistocene radiation of the species, shaped by environmental conditions and the ecological characteristics of fan mussels.

Previous studies suggested that *Pinna nobilis* originated in the Western Mediterranean around 2.5 mya[14] and began a rapid spread across the basin about 1.5 mya. However, it should be considered that this finding is based on a fragment of the mitochondrial COI gene, and therefore, the evolutionary trend could be modified or implemented by analysing a larger fragment. Nonetheless, the present study corroborates the previous finding[14] that the Mediterranean expansion process, which preceded the Pleistocene radiation, took about one million years, allowing *Pinna nobilis* ancestor to achieve a stable and well-adapted genetic make-up that persists to the present day.

Examining the temporal origin of the mitochondrial haplotypes, we identified two major genetic groups within *Pinna nobilis*: Clade A and Clade B. These groups trace back to different points in the species' evolutionary history. Clade A is representative of the species' early ancestral genetic pool, while Clade B appears to be derived from the first Pleistocene radiation event, reflecting changes that helped the species adapt to new Mediterranean environments, after its origin 2.5 mya. Interestingly, nearly all mitochondrial lineages which characterised the species in the last century diverged between 900 and 500 kya, suggesting that the genetic diversity observed in the 1900s originated in ancient times. The persistence of certain haplotypes over long periods may be linked to their role in enhancing the species' resilience.

The more common mitochondrial haplotypes found in this study may represent ancient lineages associated with mitochondrial or nuclear genes that likely conferred adaptive advantages during the Pleistocene radiation, such as greater resilience to fluctuations in temperature, salinity, and food availability. These mitochondrial haplotypes may harbour advantageous genetic variants positively selected during glacial and interglacial cycles, enabling *Pinna nobilis* to endure severe bottlenecks caused by pollution, pathogens, environmental changes, and human overexploitation in the twentieth and twenty-first centuries.

For these reasons, the long-term stability of *Pinna nobilis* haplotypes, as highlighted in this study, is a key factor in the species' resilience and survival. Remarkably, *Pinna nobilis* has exhibited minimal genetic changes over time, indicating that its haplotypes are highly adapted to prevailing environmental conditions, resulting in a state of genetic equilibrium. This stability has likely played a crucial role in the species' persistence, enabling it to adapt effectively to environmental changes and withstand human-mediated evolutionary pressures.

However, this genetic stability has not made the species invulnerable to different stressors. Over the last century, *Pinna nobilis* populations have faced multiple threats that significantly impacted their genetic variability and survival. One of the earliest pressures was human overexploitation, driven by the demand for sea silk (byssus) production and the harvesting of the shells for ornaments and collectibles[35]. Moreover, its abductor muscle was also consumed as a delicacy[4]. This overexploitation caused a reduction in genetic variability, as random harvesting indiscriminately removed individuals regardless of their haplotype, thereby limiting the haplotype richness of the genetic pool. Later, pollution produced by anthropogenic activities, such as mining and intensive agriculture[36] introduced additional stressors that likely affected population health and fitness (reproductive success) across the whole Mediterranean basin. This could have affected the size of populations where only individuals with resilient haplotypes survived. Moreover, from 2016 onwards, MMEs linked to a multifactorial disease had a devastating impact on *Pinna nobilis* populations across the Mediterranean, experiencing a sharp decline in population size. This led to the persistence of very few individuals with the most common mitochondrial haplotypes which likely are involved in resilience and capability to survive.

Notably, genetic divergence has remained high, enabling natural selection to act on a broad pool of allelic variants and favour those most beneficial for long-term survival. In this context, while natural bottlenecks (such as those produced by MMEs) reduce population size, they may facilitate the selection and fixation of beneficial traits in remaining populations. In contrast, human-driven bottlenecks (such as overexploitation for sea silk) reduce genetic variability randomly, increasing the risk of losing advantageous genetic traits (and preserving maladaptive traits just by chance) and thus posing a greater threat to long-term adaptability. Furthermore, the beneficial variants that support the species' persistence are likely not strongly linked to maladaptive ones. This potentially helps preserve their positive effects, allowing the overall advantage to outweigh any potential disadvantages.

Importantly, the study offers a surprising and hopeful discovery: present-day surviving populations exhibit higher genetic variability than those recorded in the early 1900s. This increase in genetic variability may be attributed to the persistence of haplotypes that trace back to ancient lineages. It also evidences that even small "founder populations" (those remained after significant population declines) can retain the genetic potential required for recovery. This finding reinforces the importance of protecting remnant populations (even isolated individuals), as they may harbour the genetic diversity essential for the species' future survival.

The genetic legacy of *Pinna nobilis* reflects its stability, resilience, and ability to adapt. Despite facing major environmental changes and human impact over thousands of years, its genetic structure has stayed largely unchanged.

Moreover, the dispersal potential of *Pinna nobilis* larvae, aided by surface currents and hydrodynamic factors, promotes connectivity and reinforces genetic flow among populations. In particular, temperature appears to be an important abiotic factor shaping genetic differentiation between eastern and western populations[37]. Variations in temperature can regulate spawning periods and affect larval development and settlement. Warmer waters, typical of the eastern Mediterranean, may lead to earlier spawning and shortened larval durations, promoting settlement closer to parental habitats. In contrast, cooler waters in the western Mediterranean could prolong the planktonic larval phase, allowing for wider dispersal and greater connectivity even among distant areas.

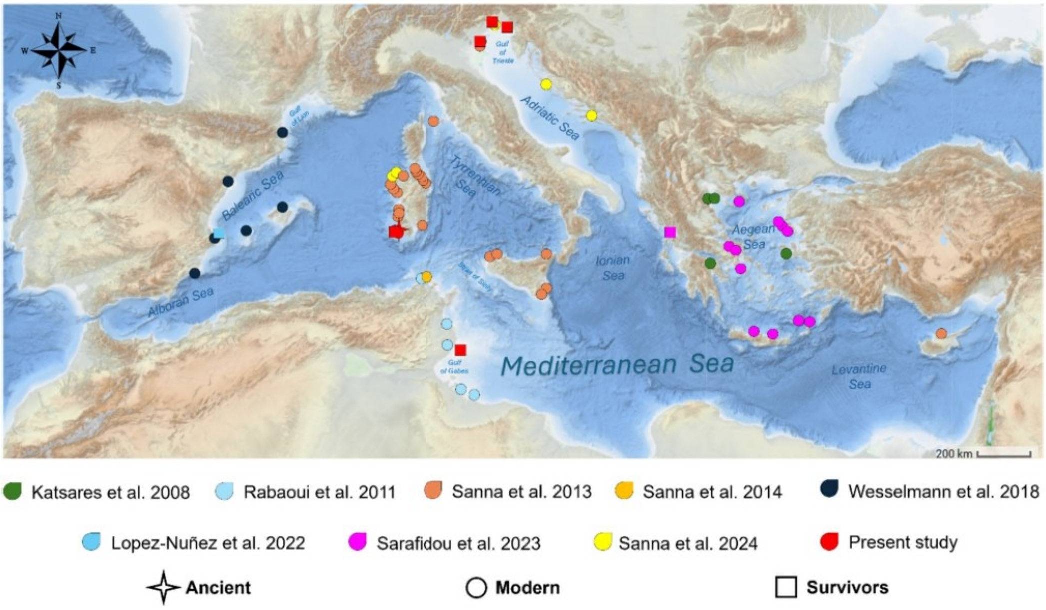


Fig. 4. Map of the sampling sites. The map shows the geographical locations for the sequences obtained in the present study along with those from previous research[6,7,10,11,14,24,32,33]. The different studies are indicated using distinct colours, while unique symbols represent key time periods: ancient populations (300 years ago, 1920s, 1970s, 1990s), modern populations (early 2000s, pre-MMEs), and survivor populations (post-MMEs). Detailed symbol and colour designations are provided in the map legend. The map was generated using the online website ScribbleMaps, freely available at the site https://www.scribblemaps.com/. Minor edits, such as adding labels, symbols and North arrow, were performed using Microsoft PowerPoint (Office 16). Credits: Ilenia Azzena.

These temperature-driven dynamics, along with ocean currents and other hydrodynamic forces, may therefore help to explain the general genetic homogeneity of *Pinna nobilis* populations in Western Mediterranean and the slight level of isolation by distance highlighted by the Mantel test in the Eastern Mediterranean.

For this reason, conservation efforts should focus on protecting populations by avoiding the harvesting of individuals across the Mediterranean to preserve this diversity.

The persistence of key mitochondrial haplotypes overtime and the retention of genetic diversity in surviving populations offer hope for the recovery of this iconic Mediterranean species.

Materials and methods

Sample collection

The dataset analysed in this study comprises all sequences of *Pinna nobilis* collected before, during, and after the MMEs, either sourced from GenBank[6,7,10,11,14,24,32,33] (n = 548, last updated on 8th December 2024) or generated in this study (n = 119) from both fresh tissues and museum specimens (1700s -2024) (see sampling details in Fig. 4 and Supplementary Table S1).

For living individuals from Western Mediterranean and North Adriatic Sea, a non-lethal sampling was conducted following the methodology of Sanna et al.[11]. This method was officially approved by the Italian "Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA)" and the "Ministero dell'Ambiente e della Tutela del Territorio e del Mare", as reported in the Ethics & Inclusion Statement.

A non-invasive sampling method[24] was also employed, using a cotton swab to scrape tissue and mucus from the interior surfaces of the species. However, this technique yielded ~ 10 times less DNA than tissue sampling.

Ancient byssus samples and standardization of extraction protocol

We extracted total genomic DNA from biological tissue fragments belonging to the byssus gland that were still adherent to the basal portion of the keratinous filaments. These old byssus samples sourced from various collections, including samples from Sant'Antioco. These latter included samples donated by the sea silk master Chiara Vigo, the Pes sisters (via Italo Diana's family), as well as samples from Efisia Murroni's collection, held in the Museo Etnografico of Sant'Antioco. Additionally, we included byssus samples donated by local fishermen to the same museum.

Since previous inclusive phylogeographic studies on *Pinna nobilis* highlighted the species' effective dispersal capability across the Mediterranean[11] and the genetic homogeneity between the Western and Adriatic basins[14], ancient samples from Sant'Antioco Island can be considered as part of a representative population of Western Mediterranean. In this way, we assume to have access not only to ancient *Pinna nobilis* individuals from Sant'Antioco Island, but to a substantial representation of entire Western Mediterranean populations from key time periods: 1920s, 1970s, and 1990s.

A standardized protocol was developed using the Macherey–Nagel Nucleo Spin Tissue Kit (MACHEREY–NAGEL GmbH & Co. KG; Neumann Neander Str. 6–8 D-52355 Düren, Germany). Since the lack of prior studies on DNA extraction from byssus, we conducted tests using two (SAV1 and SAV2) samples (see Table S2 for the full list of hydrating and incubation time conditions) to optimize the procedure.

These tests were conducted using various extraction combinations on both dry and rehydrated tissues samples (~ 25 mg), with rehydration performed using Milli-Q water. Additionally, we tested different incubation times for the action of the proteolytic enzyme, proteinase K (see Table S2 for the full list of hydrating and incubation time conditions).

These trials were performed with the aim of identifying the most efficient method for extracting DNA from old biological tissue fragments belonging to the byssus gland still adherent to the basal portion of filaments. After completing the initial phase of experimentation, we moved forward with the protocol as recommended by the manufacturers, making only a single variation: rather than eluting the final extract in 100 µl of elution buffer, we opted for 50 µl to obtain a more concentrated DNA solution.

The DNA concentration obtained for the samples were quantified using the Nanodrop™ Lite Spectrophotometer (Thermo Scientific, Waltham, MA, USA) (see Table S3 for the protocol details and DNA yield).

All the combinations were tested through standard PCR with species-specific primers[11] for Cytochrome c Oxidase subunit I of *Pinna nobilis*. The test (number 5 in Table S2), which contemplate 5 h of hydration + 2 h of incubation at 56 °C, produced the most reliable results, yielding average concentrations of 2.5 ng/µl (see Table S3).

Although some other combinations provided higher DNA concentrations, they simultaneously exhibited lower absorbance values (see Table S3 for details), and the electrophoresis did not display bands or, if yes, the bands did not correspond to the expected base pair (bp) size (see Figure S2).

It is worth noting that for the dry or highly hydrated attempts (e.g., tests 1, 2, 3, 6, and 7 in Table S2), the PCR products which were twice the expected size (as shown in Figure S2). Subsequent sequencing reactions and BLAST analyses confirmed the presence of fungal amplification, specifically belonging to the *Aspergillus* genus. These findings further emphasize the effectiveness of test number 5 (5 h of hydration + 2 h of incubation at 56 °C, see Table S2), in obtaining reliable DNA extracts from old biological samples of byssus gland.

Molecular analyses

For living individuals total genomic DNA was extracted from both mantle tissue (sampling method from Sanna et al.[11]) and cotton swab (sampling method from Sarafidou et al.[24]) using the Macherey–Nagel NucleoSpin Tissue Kit (MACHEREY–NAGEL GmbH & Co. KG, Düren, Germany) following the manufacturer's instructions. DNA was quantified with a Nanodrop™ Lite Spectrophotometer (Thermo Scientific, Waltham, MA, USA), yielding average concentrations of: (i) 79 ng/µl for the tissue sampling method and (ii) 7.5 ng/µl for the cotton swab sampling method. The mitochondrial Cytochrome c Oxidase subunit I gene (COI) was partially amplified using standard PCR with species-specific primers (L: 5′-GGTTGAACTATHTATCCNCC-3′ and H: 5′-GAAATCATYCCAAAAGC-3′) following the protocol used in Sanna et al.[11], producing a 338 base pair fragment. Electrophoresis was performed on 2% agarose gels in 1 × TAE buffer (Tris–acetate-EDTA, pH 8.3) stained with GelRed Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). PCR products were purified using ExoSAP-IT (USB Corporation, Cleveland, OH, USA) and sequenced in both forward and reverse directions using the same PCR primers through an external Sanger sequencing service (Macrogen Europe, Amsterdam, The Netherlands, and Macrogen Europe, Milano, Italy).

Phylogenetics and phylogeographic analyses

The 119 newly generated sequences (GenBank PQ728133-PQ728251, see Table S1 for details) were aligned with all previously deposited sequences in GenBank[6,7,10,11,14,24,32,33] utilizing Clustal Omega[38].

Genetic variation within datasets was assessed using DnaSP version 6.12.03[39].

The best probabilistic model of sequence evolution was determined using jModeltest 2.1.3[40], identifying GTR + G[41] as the best-fit model.

Phylogeographic and phylogenetic analyses included a comprehensive dataset spanning all *Pinna nobilis* sequences from the 1700s to 2024 (see Table S1) to evaluate potential diachronic variations in genetic variability in response to human and environmental stressors.

Genetic relationships between haplotypes and their distribution frequencies were assessed through a median-joining network[42] using Network 10.2.0.0 (www.fluxus-engineering.com).

To assess the presence of isolation by distance (IBD), a Mantel test was performed using Pearson's correlation to evaluate the relationship between genetic and geographic distances. Genetic distances were computed from the DNA sequence alignment using the APE package v.5.8-1[43] in R v.4.3.1[44], applying the "raw" model (p-distance). Geographic distances were calculated as Euclidean distances (in meters) between sampling site coordinates[45] using the GEOSPHERE package v.1.5-20[46]. The Mantel test, performed with the VEGAN package v.2.6-10[47], using 10,000 permutations to assess statistical significance. The strength and significance of the association were evaluated using the Mantel statistic (r) and its associated p-value. Spatial Principal Component Analysis (sPCA) was performed by using ADEGENET package v.2.1.22[48] in order to detect spatial patterns of genetic variation, including both global structures (e.g., gradients) and local structures (e.g., genetic discontinuities), by incorporating spatial information into multivariate analysis. Positive values of Moran's I (Moran 1950) [49] indicate positive spatial autocorrelation, where nearby individuals tend to share similar genetic profiles. In contrast, negative values reflect negative spatial autocorrelation, suggesting sharp local genetic discontinuities between geographically close individuals, potentially due to barriers to gene flow or local isolation. Values near zero indicate no spatial structure, corresponding to a random spatial distribution of genetic variation.

Molecular dating and estimation of the divergence time analyses

An additional dataset was created by selecting the 100 haplotypes identified within the whole dataset of 667 sequences by means of DnaSP 6.12.03[39]. Subsequently, this subset was used to estimate the divergence time of each lineage found.

Phylogenetic relationships among haplotypes were inferred through Bayesian Inference (BI) using MrBayes 3.2.6[49] following Scarpa et al.[50].

Divergence times of the clades, identified in the Cytochrome c Oxidase subunit I haplotypes phylogenetic tree, were estimated with BEAST 1.10.4[51] applying evolutionary rates for marine bivalves with pelagic larval dispersal as suggested by Luttikhuizen et al.[52] and recently verified by Sanna et al.[14]. In BEAUti (part of the BEAST package), mutation rates were specified using a normal distribution ranging from 0.14 to 0.52% divergence per site per million years. Substitution models were defined based on the results of jModeltest, setting the model to GTR with estimated base frequencies, a Gamma distribution plus invariant sites for heterogeneity, and four gamma rate categories. A lognormal uncorrelated relaxed clock model was chosen to allow for rate variation among branches, assuming evolutionary rates vary independently across the tree. The Speciation Yule Process[53,54] was used as the tree prior to model a constant speciation rate. Model priors and parameters were tailored to calibrate the time-tree using the mutation rates per million years, with divergence estimates derived from a normal distribution reflecting minimum, mean, and maximum mutation rate values. Operator settings followed the BEAST user manual guidelines. The lognormal uncorrelated relaxed clock model also provided insight into how clock-like the data were, using the ucld.stdev parameter: values close to zero indicated a clock-like pattern, while values well above one suggested considerable rate heterogeneity across lineages. To ensure sufficient sampling of all model parameters (ESS > 200), the analysis was run for 400 million generations, with tree samples taken every 40,000 generations, following the protocol of Scarpa et al.[55]. Tracer v1.7.1[56] was employed to inspect the log files, verify convergence, check ESS values, and estimate node ages. Final visualization and editing of the time-calibrated phylogenetic tree were carried out using TreeAnnotator v1.10.4 and FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Data availability

The datasets used and analysed during the current study are available from the corresponding author (Daria Sanna darsanna@uniss.it) on reasonable request. The new sequences generated and analysed during the current study are available in the GenBank repository under the accession numbers PQ728133-PQ728251.These sequences were deposited on December 10, 2024 (SUB14920067) and are set for release on March 18, 2026, or upon manuscript publication. GenBank accession numbers for all data analysed during this study, except for newly generated sequences, are included in this published article [and its supplementary information files, Table S1].

Received: 8 January 2025; Accepted: 22 September 2025

References

1. Butler, A., Vicente, N. & de Gaulejac, B. Ecology of the pterioid bivalves *Pinna bicolor* Gmelin. *Mar. Life* **3**(1–2), 37–45 (1993).
2. Rabaoui, L., Tlig-Zouari, S., Cosentino, A. & Hassine, O. K. B. Associated fauna of the fan shell *Pinna nobilis* (Mollusca: Bivalvia) in the northern and eastern Tunisian coasts. *Sci. Mar.* **73**(1), 129–141 (2009).
3. Hernandis Caballero, S. Avances en la ecología y biología para la conservación de la especie críticamente amenazada *Pinna nobilis* (Linnaeus, 1758), endemismo mediterráneo (Doctoral dissertation, 2021).
4. Greenwald, M. *The cruising chef cookbook* 300 (Paradise Cay Publication, 1996).
5. Katsanevakis, S. et al. The fan mussel *Pinna nobilis* on the brink of extinction in the Mediterranean. In *Imperiled: The Encyclopedia of Conservation* (eds DellaSala, D. A. & Goldstein, M. I.) 700–709 (Elsevier, 2022).
6. Katsares, V., Tsiora, A., Galinou-Mitsoudi, S. & Imsiridou, A. Genetic structure of the endangered species *Pinna nobilis* (Mollusca: Bivalvia) inferred from mtDNA sequences. *Biologia* **63**, 412–417 (2008).
7. Rabaoui, L. et al. Genetic variation among populations of the endangered fan mussel Pinna nobilis (Mollusca: Bivalvia) along the Tunisian coastline. *Hydrobiologia* **678**, 99–111 (2011).
8. Gonzalez-Wanguemert, M. et al. Highly polymorphic microsatellite markers for the Mediterranean endemic fan mussel *Pinna nobilis*. *Mediterr. Mar. Sci.* **16**, 31–35 (2015).
9. González-Wangüemert, M. et al. Gene pool and connectivity patterns of *Pinna nobilis* in the Balearic Islands (Spain, Western Mediterranean Sea): Implications for its conservation through restocking. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **29**, 175–188 (2019).
10. Wesselmann, M. et al. Genetic and oceanographic tools reveal high population connectivity and diversity in the endangered pen shell *Pinna nobilis*. *Sci. Rep.* **8**(1), 4770 (2018).
11. Sanna, D. et al. Mitochondrial DNA reveals genetic structuring of *Pinna nobilis* across the Mediterranean Sea. *PLoS ONE* **8**(6), e67372 (2013).
12. Peyran, C. et al. Genetic homogeneity of the critically endangered fan mussel, *Pinna nobilis*, throughout lagoons of the Gulf of Lion (North-Western Mediterranean Sea). *Sci. Rep.* **11**, 7805 (2021).
13. Peyran, C. et al. Investigating population dynamics from parentage analysis in the highly endangered fan mussel *Pinna nobilis*. *Ecol. Evol.* **12**, e8482 (2022).
14. Sanna, D. et al. Reconstructing the evolutionary history of *Pinna nobilis*: New genetic signals from the past of a species on the brink of extinction. *Animals* **14**(1), 114 (2023).
15. Vázquez-Luis, M. et al. SOS *Pinna nobilis*: A mass mortality event in western Mediterranean Sea. *Front. Mar. Sci.* **4**, 220 (2017).

16. Catanese, G. et al. *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.* **157**, 9–24 (2018).
17. Carella, F. et al. A mycobacterial disease is associated with the silent mass mortality of the pen shell *Pinna nobilis* along the Tyrrhenian coastline of Italy. *Sci. rep.* **9**(1), 2725 (2019).
18. Prado, P. et al. Presence of *Vibrio mediterranei* associated to major mortality in stabled individuals of *Pinna nobilis* L. *Aquaculture* **519**, 734899 (2020).
19. Scarpa, F. et al. Multiple non-species-specific pathogens possibly triggered the mass mortality in *Pinna nobilis*. *Life* **10**(10), 238 (2020).
20. Carella, F. et al. Multipathogen infections and multifactorial pathogenesis involved in noble pen shell (*Pinna nobilis*) mass mortality events: Background and current pathologic approaches. *Vet. Pathol.* **60**(5), 560–577 (2023).
21. Carella, F. et al. Measuring immunocompetence in the natural population and captive individuals of noble pen shell *Pinna nobilis* affected by *Pinna nobilis* Picornavirus (PnPV). *Fish Shellfish Immunol.* 109664 (2024).
22. García-March, J. R. et al. Can we save a marine species affected by a highly infective, highly lethal, waterborne disease from extinction?. *Biol. Conserv.* **243**, 108498 (2020).
23. Silvestri, S., Capra, V., Cucchiaro, S., Pivato, M. & Tarolli, P. Tides, topography, and seagrass cover controls on the spatial distribution of *Pinna nobilis* on a coastal lagoon tidal flat. *J. Geophys. Res. Biogeosci.* **127**(3), e2021JG006667 (2022).
24. Sarafidou, G. et al. Insights on *Pinna nobilis* population genetic structure in the Aegean and Ionian Sea. *PeerJ* **11**, e16491 (2023).
25. Foulquié, M., Coupé, S., Vicente, N. & Bunet, R. Population status of the Critically Endangered fan mussel *Pinna nobilis* in Thau Lagoon (France). *Endanger. Species Res.* **55**, 21–36 (2024).
26. Lunetta, A., Spinelli, A., Donato, G., Gatì, I. A. & Giacobbe, S. Lake faro (Central Mediterranean): A potential short-term reservoir for *Pinna nobilis*. *J. Nat. Conserv.* **81**, 126690 (2024).
27. Martínez-Martínez, P. et al. Assessing habitat microscale conditions underlying with mortality and survival of *Pinna nobilis* in the Mar Menor coastal lagoon (SE Spain). *J. Sea Res.* **200**, 102518 (2024).
28. Cinar, M. E., Bilecenoglu, M., Yokeş, M. B. & Güçlüsoy, H. A. The last fortress fell: mass mortality of *Pinna nobilis* in the Sea of Marmara. *Mediterr. Mar. Sci.* **22**(3), 669–676 (2021).
29. Donato, G., Vázquez-Luis, M., Nebot-Colomer, E., Lunetta, A. & Giacobbe, S. Noble fan-shell, *Pinna nobilis*, in Lake Faro (Sicily, Italy): Ineluctable decline or extreme opportunity?. *Estuar. Coast Shelf Sci.* **261**, 107536 (2021).
30. Acarlı, S., Acarlı, D. & Kale, S. Distribution of the critically endangered fan mussel *Pinna nobilis* population in the Çanakkale Strait and Marmara Sea. *Ege J. FAS* **41**(1), 46–53. https://doi.org/10.12714/egejfas.41.1.07 (2024).
31. Karadurmuş, U., Benli, T. & Sarı, M. Discovering new living *Pinna nobilis* populations in the Sea of Marmara. *Mar. Biol.* **171**(4), 90 (2024).
32. Sanna, D. et al. New mitochondrial and nuclear primers for the Mediterranean marine bivalve *Pinna nobilis*. *Mediterr. Mar. Sci.* **15**(2), 416–422 (2014).
33. Lopez-Nuñez, R. et al. Detection of *Haplosporidium pinnae* from *Pinna nobilis* Faeces. *J. Mar. Sci. Eng.* **10**(2), 276 (2022).
34. Moran, P. A. P. Notes on continuous stochastic phenomena. *Biometrika* **37**(1/2), 17–23 (1950).
35. Basso, L. et al. The pen shell, *Pinna nobilis*: A review of population status and recommended research priorities in the Mediterranean Sea. *Adv. Mar. Biol.* **71**, 109–160 (2015).
36. Béjaoui, B. et al. Synthesis review of the Gulf of Gabes (eastern Mediterranean Sea, Tunisia): Morphological, climatic, physical oceanographic, biogeochemical and fisheries features. *Estuar. Coast. Shelf Sci.* **219**, 395–408 (2019).
37. Kersting, D. K. & García-March, J. R. Long-term assessment of recruitment, early stages and population dynamics of the endangered Mediterranean fan mussel *Pinna nobilis* in the Columbretes Islands (NW Mediterranean). *Mar. Environ. Res.* **130**, 282–292 (2017).
38. Sievers, F. & Higgins, D. G. Clustal omega, accurate alignment of very large numbers of sequences. In *Multiple Sequence Alignment Methods* Vol. 1079 (ed. Russell, D.) 105–116 (Humana Press, 2014).
39. Librado, P. & Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
40. Posada, D. jModelTest: Phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256 (2008).
41. Tavaré, S. Some probabilistic and statistical problems in the analysis of DNA sequences. In *Some Mathematical Questions in Biology—DNA Sequence Analysis* (ed. Miura, R. M.) 57–86 (American Mathematical Society, 1986).
42. Bandelt, H. J., Forster, P. & Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
43. Paradis, E. & Schliep, K. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
44. Team, R.D.C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2024).
45. Karney, C. F. F. Algorithms for geodesics. *J. Geod.* **87**, 43–55 (2013).
46. Hijmans, R. J., Williams, E., Vennes, C. & Hijmans, M. R. J. Package 'geosphere'. *S. Trig.* **1**(7), 1–45 (2017).
47. Oksanen, J. et al. Vegan: Community Ecology Package. R package version 2.5-6 (2019).
48. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129 (2008).
49. Ronquist, F. et al. MrBayes 32: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
50. Scarpa, F. et al. New insights on the genus *Otoplana* Du Plessis, 1889 (Platyhelminthes: Proseriata), with description of two new species from the Canary Islands. *Mar. Biodiv.* **49**, 2075–2087. https://doi.org/10.1007/s12526-017-0785-1 (2019).
51. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, e214 (2007).
52. 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.* **12**, 2215–2229 (2003).
53. Gernhard, T. J. The conditioned reconstructed process. *J. Theor. Biol.* **253**, 769–778 (2008).
54. 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.* **213**, 21–87 (1925).
55. Scarpa, F. et al. Meiofaunal cryptic species challenge species delimitation: the case of the *Monocelis lineata* (Platyhelminthes: Proseriata) species complex. *Contrib. Zool.* **85**, 123–145 (2016).
56. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).

Acknowledgements

The authors are grateful to the citizen science group Subacquei per la Scienza (Underwater Divers for Science) for their support and involvement in the documentation and search for live specimens. The authors are also indebted to the personnel of OutBe for their assistance during the detection of individuals in Sant'Antioco, as well as to the citizens who reported the presence of *Pinna nobilis* in Sant'Antioco as part of the citizen science project.

The Tunisian samples from Kerkennah Island were collected thanks to the support of the Specially Protected Areas Regional Activity Centre (SPA/RAC) of the United Nations Environment Programme / Mediterranean Action (UNEP/MAP) Barcelona Convention, who implemented a project funded by the UNEP Regional Seas Programme—2021 Swedish International Development Cooperation Agency (SIDA) with the aim of contributing to the restoration of Pinna nobilis through two actions: The drafting of a restoration plan and training activities at the Kerkennah Islands, Tunisia, 23-25 June 2022. We are also grateful to the project funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4—Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union–NextGenerationEU; Award Number: Project code CN_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP D33C22000960007, Project title "National Biodiversity Future Center". The dedication and passion that animated Daria Sanna during this genetic study are a tribute to the memory of Giuseppino Sanna, an extraordinary man and her wonderful father.

Author contributions

I.A. and D.S. conceived the study; I.A., C.L, M.C., and D.S. wrote the original draft of the manuscript and prepared figures and tables; I.A., C.L, N.P., and I.D. performed molecular analyses; I.A., C.L, and F.S. performed bioinformatic analyses; R.S., E.B., D.C., M.C., S.C., M.P.F., D.G., A.L., M.M., A.O., V.P., M.S., R.Z., A.B.H., collected samples, provided resources and managed the data curation; M.C. and D.S. supervised and administrated the study and acquired the funding to develop the whole bulk of activities involved in the study. All authors reviewed and edited the manuscript and visualized the results. All authors have read and agreed to the published version of the manuscript.

Funding

This whole 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. In Slovenia the LIFE PINNA project is co-financed by the Ministry of Environment and Spatial Planning of Slovenia, with additional financial support from the Slovenian Research Agency (research core funding no. P1-0237).

Declarations

Competing interests

The authors declare no competing interests.

Ethics & inclusion statement

For the collection of tissues, we used a specific non-lethal sampling method that did not cause significant damage to the shells and soft tissues of *Pinna 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". Small mantle tissue fragments were collected via scuba diving with minimal harm to the shells or soft tissues of *Pinna nobilis* individuals, under permits from the respective authorities for each protected area. These include the derogation authorization under DPR 357/97, for research activities on *Pinna nobilis* as part of the European LIFE20 NAT/IT/001122—LIFE PINNA project (Prot. MASE 0049391 of 30/03/2023), issued by the Ministry of the Environment and Energy Security—Directorate General for Biodiversity and the Sea, and the Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA). Additionally, a permit was released on February 22, 2022, by the Ministry of Natural Resources and Spatial Planning of the Republic of Slovenia (former Ministry for the Environment and Spatial Planning) (n. 35606-9/2022-2550-5).

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-21574-6.

Correspondence and requests for materials should be addressed to D.S.

Reprints and permissions information is available at www.nature.com/reprints.

Chapter 2

Callinectes sapidus

The blue crab, *Callinectes sapidus* Rathbun 1896 (Arthropoda: Malacostraca: Portunidae), is a crustacean species native to the western Atlantic Ocean and considered an invasive alien species in the Mediterranean Sea (Zenetos et al., 2017; Figure 2). It is an opportunistic predator and a scavenger that primarily inhabits estuaries and lagoons and is able to tolerate a wide range of environmental conditions, including broad temperature fluctuations (7°-32°C) (Marchessaux et al., 2022, 2023a).

Figure 2. *Callinectes sapidus* specimens collected from Sardinian lagoons (Italy): (a) ovigerous female and (b) male specimen from the Avalè–Su Pedrosu lagoon; (c) individuals of *C. sapidus* from the Tortolì lagoon. Photo credits: Noemi Pascale.

Over the last decades, *Callinectes sapidus* has rapidly expanded across the Mediterranean basin, mainly as a result of intentional or accidental introductions, particularly via ballast water transport (Nehring, 2011). Its successful colonisation of new habitats is further facilitated by a combination of biological and behavioural traits, including strong swimming abilities, high dispersal capacity, long-distance movements of adult females, and larval dispersal (Hill et al., 1989; Nehring, 2011; Marchessaux et al., 2023a, 2023b).

Callinectes sapidus is widely recognised for its negative impact on biodiversity, fisheries, and aquaculture activities. Due to its aggressiveness, the species is known to damage fishing nets, cause fish mutilations, and compete with autochthonous species, particularly for food resources (Streftaris & Zenetos, 2006; Marchessaux et al., 2023a, 2023b). Additionally, the establishment of *C. sapidus* populations raises concerns regarding the potential transmission of pathogens and viral diseases to native crustaceans. Notable examples include the dinoflagellate *Hematodinium* sp. (e.g., Small et al., 2007; Lattos et al., 2024) and *Callinectes sapidus* reovirus 1 (CsRV1) (e.g., Flowers et al., 2018; Lively et al., 2023).

The earliest record of *C. sapidus* in European waters dates back to 1901, along the French coast (Bouvier, 1901), while historical reports suggest its presence in the Aegean Sea in 1935. However, the first confirmed observation in the Mediterranean region was reported in 1949 in Italy, in the northern Adriatic Sea (Mizzan, 1993; Nehring, 2011).

In light of this background, the study presented in this thesis focused on investigating genetic variation and evolutionary history of *Callinectes sapidus* across both its native and introduced ranges of distribution, based on analyses of the mitochondrial Cytochrome c Oxidase subunit I (COI) gene. Specifically, the main aims were to: (i) assess the levels of genetic variability among and within native and introduced populations; (ii) reconstruct phylogeographic patterns across a large part of the species distribution range; (iii) infer on the evolutionary history of *C. sapidus* populations; and (iv) evaluate the taxonomic status of the species across its geographical distribution range.

References

Bouvier, E. L. (1901). Sur un *Callinectes sapidus* M. Rathbun trouvé à Rochefort. *Bulletin du Muséum national d'Histoire naturelle, 7*, 16–17.

Flowers, E. M., Johnson, A. F., Aguilar, R., & Schott, E. J. (2018). Prevalence of the pathogenic crustacean virus *Callinectes sapidus* reovirus 1 near flow-through blue crab aquaculture in

Chesapeake Bay, USA. *Diseases of Aquatic Organisms, 129*, 135–144. https://doi.org/10.3354/dao03232

Hill, J., Fowler, D. L., & Van Avyle, M. J. (1989). *Species profiles: Life histories and environmental requirements of coastal fishes and invertebrates (Mid-Atlantic)—Blue crab*. U.S. Army Corps of Engineers. https://apps.dtic.mil/sti/pdfs/ADA210181.pdf

Lattos, A., Papadopoulos, D. K., Giantsis, I. A., Stamelos, A., & Karagiannis, D. (2024). Histopathology and phylogeny of the dinoflagellate *Hematodinium perezi* and the epibiotic peritrich ciliate *Epistylis* sp. infecting the blue crab *Callinectes sapidus* in the Eastern Mediterranean. *Microorganisms, 12*(3), 456. https://doi.org/10.3390/microorganisms12030456

Lively, J. A., Spitznagel, M. I., Schott, E. J., & Small, H. J. (2023). Investigating conspecific CsRV1 transmission in *Callinectes sapidus*. *Journal of Invertebrate Pathology, 201*, 107987. https://doi.org/10.1016/j.jip.2023.107987

Marchessaux, G., Bosch-Belmar, M., Cilenti, L., Lago, N., Mangano, M. C., Marsiglia, N., & Sarà, G. (2022). The invasive blue crab *Callinectes sapidus* thermal response: Predicting metabolic suitability maps under future warming Mediterranean scenarios. *Frontiers in Marine Science, 9*, 1055404. https://doi.org/10.3389/fmars.2022.1055404

Marchessaux, G., Chevalier, C., Mangano, M. C., & Sarà, G. (2023b). Larval connectivity of the invasive blue crabs *Callinectes sapidus* and *Portunus segnis* in the Mediterranean Sea: A step toward improved cross-border management. *Marine Pollution Bulletin, 194*, 115272. https://doi.org/10.1016/j.marpolbul.2023.115272

Marchessaux, G., Gjoni, V., & Sarà, G. (2023a). Environmental drivers of size-based population structure, sexual maturity and fecundity: A study of the invasive blue crab *Callinectes sapidus* (Rathbun, 1896) in the Mediterranean Sea. *PLoS ONE, 18*(8), e0289611. https://doi.org/10.1371/journal.pone.0289611

Mizzan, L. (1993). Presence of swimming crabs of the genus *Callinectes* (Stimpson) (Decapoda, Portunidae) in the Venice Lagoon (North Adriatic Sea, Italy): First record of *Callinectes danae* Smith in European waters. *Bollettino del Museo Civico di Storia Naturale di Venezia, 42*, 31–43.

Small, H. J., Shields, J. D., Hudson, K. L., & Reece, K. S. (2007). Molecular detection of *Hematodinium* sp. infecting the blue crab, *Callinectes sapidus*. *Journal of Shellfish Research, 26*(1), 131–139. https://doi.org/10.2983/0730-8000(2007)26[131:MDOHSI]2.0.CO;2

Streftaris, N., & Zenetos, A. (2006). Alien marine species in the Mediterranean—The 100 "worst invasives" and their impact. *Mediterranean Marine Science, 7*(1), 87–118. https://doi.org/10.12681/mms.180

Zenetos, A., Çinar, M. E., Crocetta, F., Golani, D., Rosso, A., Servello, G., Shenkar, N., Turon, X., & Verlaque, M. (2017). Uncertainties and validation of alien species catalogues: The Mediterranean as an example. *Estuarine, Coastal and Shelf Science, 191*, 171–187. https://doi.org/10.1016/j.ecss.2017.03.031

Research article

life

MDPI

Article

A Sister Species for the Blue Crab, *Callinectes sapidus*? A Tale Revealed by Mitochondrial DNA

Chiara Locci [1,2], Ilenia Azzena [1], Noemi Pascale [1,3], Alessandra Ciccozzi [2], Ilaria Deplano [2], Ioannis A. Giantsis [4], Dimitrios K. Papadopoulos [4], Athanasios Lattos [4], Flavio Orrù [5], Cesare M. Puzzi [5], Fabio Scarpa [2], Marco Casu [1,†] and Daria Sanna [2,*,†]

[1] Department of Veterinary Medicine, University of Sassari, Via Vienna 2, 07100 Sassari, Italy; c.locci3@phd.uniss.it (C.L.); iazzena@uniss.it (I.A.); npascale@uniss.it (N.P.); marcasu@uniss.it (M.C.)
[2] Department of Biomedical Sciences, University of Sassari, Viale San Pietro 43b, 07100 Sassari, Italy; aciccozzi@uniss.it (A.C.); ilariadeplano1@gmail.com (I.D.); fscarpa@uniss.it (F.S.)
[3] Department of Chemical, Physical, Mathematical, and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy
[4] Laboratory of Ichthyology and Fisheries, Faculty of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece; igiants@auth.gr (I.A.G.); dkpapado@bio.auth.gr (D.K.P.); lattosad@bio.auth.gr (A.L.)
[5] GRAIA Gestione Ricerca Ambientale Ittica Acque, Via Repubblica 1, 21020 Varano Borghi, Italy; flaorru@tiscali.it (F.O.); cesare.puzzi@graia.eu (C.M.P.)
* Correspondence: darsanna@uniss.it
† These authors contributed equally to this work.

Abstract: The Atlantic blue crab, *Callinectes sapidus*, is acknowledged as one of the worst invasive alien species in the Mediterranean, impacting biodiversity and human activities. Native to the western Atlantic, it has expanded to European coastal waters since the early 1900s. Despite its ecological and commercial importance, genetic research on this species is limited. Here we show a comprehensive investigation of the genetic variation and evolutionary history in *Callinectes sapidus* using 667 mitochondrial COI gene sequences. Our dataset encompasses 36 newly generated sequences from previously understudied Mediterranean sites and 631 from worldwide locations obtained from the GenBank public database. Our findings reveal two distinct, but closely related, genetic groups within the species' distribution range, suggesting the occurrence of a potential species complex. Furthermore, in the Mediterranean, low levels of genetic variability were observed except for substantial haplotypic differentiation in Turkish samples. This study depicts the global genetic diversity and evolutionary patterns of *Callinectes sapidus*, offering new insights into the taxonomic status of the species.

Keywords: alien species; MtDNA; Mediterranean invasion; molecular dating; cryptic speciation; sibling species; evolutionary history

check for updates

Citation: Locci, C.; Azzena, I.; Pascale, N.; Ciccozzi, A.; Deplano, I.; Giantsis, I.A.; Papadopoulos, D.K.; Lattos, A.; Orrù, F.; Puzzi, C.M.; et al. A Sister Species for the Blue Crab, *Callinectes sapidus*? A Tale Revealed by Mitochondrial DNA. *Life* **2024**, *14*, 1116. https://doi.org/10.3390/life14091116

Academic Editor: Koichiro Tamura

Received: 30 July 2024
Revised: 30 August 2024
Accepted: 3 September 2024
Published: 5 September 2024

1. Introduction

Non-indigenous species, known as invasive alien species (IAS), pose a significant threat to biodiversity, ecosystem stability, socioeconomic activities, and human health [1–4]. In this regard, among the widely recognized dangers to biodiversity, the introduction of exotic species stands out as one of the most severe, second only to habitat loss and fragmentation [5]. Over the past decade, the rapid establishment of invasive alien species in marine and coastal ecosystems has emerged as a significant environmental concern, particularly in the Mediterranean region. This latter area is currently acknowledged as one of the most drastically impacted globally, considering both the number of detected alien species and the rate of introduction [6].

In 2006, Streftaris et al. undertook a selection process to identify the one hundred most problematic invasive species in the Mediterranean [2]. In accordance with this list, the

Atlantic blue crab, *Callinectes sapidus* Rathbun 1896 (Crustacea, Decapoda, Portunidae), has been designated as one of the most invasive crustacean species in this region [7,8]. Indeed, its presence is noted for its significant impact on biodiversity, fisheries, and aquaculture activities [2], with negative economic and social consequences [9,10]. The blue crab is notorious for causing important damages to fishing nets, primarily through entanglement and cutting [11–14] and fish mutilations [15]. Due to its aggressiveness [14], this species is known to induce competition with autochthonous species, potentially leading to local species extinctions [15,16].

The blue crab is an opportunistic predator and a scavenger inhabiting the continental shelf in estuaries, lagoons (where it can tolerate freshwater conditions), and near-shore waters. It is found at depths ranging from 0 to 90 m on sandy and muddy substrates, regardless of the presence of vegetation [14,17–19]. Additionally, the species is known to thrive in a broad temperature range, from 7 °C to 32 °C, with an optimal temperature of approximately 24 °C in introduced areas [20]. In its native range, however, it can tolerate a wider thermal range, with seawater temperatures reaching up to 40 °C [21] and thermal preferences between 26 °C and 33 °C [22].

Callinectes sapidus is a species native to the western Atlantic [23] and holds a dominant position along the eastern coasts of North and South America, including the Gulf of Mexico. In its native range, it represents one of the most important fishery resources for its commercial value and catch volume [24], and it is largely consumed [25,26].

Furthermore, *Callinectes sapidus* is well-suited for colonizing new habitats due to its biology, particularly for its long larval phase and robust swimming capability [14–16,27–33]. Factors such as adult migration patterns, local current and wind patterns influencing larval dispersion, and larval behavior can shape the genetic population structure of the blue crab [15,16].

Since the year 1900, the blue crab has expanded its geographical range to encompass numerous European coastal waters [23]. The presence of *Callinectes sapidus* in Europe was initially reported along the French coasts by Bouvier in 1901 [23,34], followed by reports along the Baltic Sea and the Atlantic coasts of Spain and Portugal [23,35]. The species has also established colonies in Japan since 1975 and Hawaii since 1985 [36], with subsequent expansion to the Black Sea [37] and the whole Mediterranean basin [10,38–40]. Historical records suggest its existence in the Aegean Sea as far back as 1935 [23,41,42], but the first confirmed observation in the Mediterranean region was reported in 1949 in Marina di Grado, Italy, in the northern Adriatic Sea. Initial misidentification of specimens as *Neptunus pelagicus* [43] was later corrected to *Callinectes sapidus*, marking the species' introduction to the region [44]. Subsequent records in Israel, Lebanon, and Egypt further highlighted the spread of the blue crab in the Mediterranean [45–47]. In Italy, since the late 1950s, numerous specimens have been collected along the Ligurian [48,49], Tyrrhenian [50,51], Ionian [6,29,52,53] and Adriatic coasts [44,54–61]. Furthermore, in the last decade, the species has exhibited a dynamic expansion in the western parts of the Mediterranean (e.g., [62,63]). Indeed, some specimens were recently recorded along the Sardinian coasts [64–66]. Noteworthy, *Callinectes sapidus* has been present for nearly a century in Turkish waters [67–69], where its abundance led to the development of a relevant commercial fishery. The significant rise in Turkish fishing activity since the 1980s highlights the potential economic importance of blue crabs for the entire Mediterranean basin [70,71]. However, while blue crab landings are primarily associated with Turkey, they have experienced a significant decline in recent years [28]. Furthermore, although blue crabs have been present in the North Aegean Sea for a similar period [23,41,42], their commercial value in Greece remains lower, as they are considered a secondary fishery product [13,28,72]. Additionally, it is important to note that exploitation of blue crabs also occurs in other Mediterranean regions, including Egypt [47], Tunisia [63], and Italy [57].

Notably, ballast water has been identified as the primary vector for its expansion [23,28,41–43,52]. Indeed, initial findings in introduced areas occurred in or near ports where ballast waters were released, suggesting that the species' introduction is primarily through transport in ballast tanks [23,45,73]. Apart from intentional or accidental introductions via ballast water, potential releases from holding tanks used for live crabs imported for food or the aquarium trade have also been considered [23,74].

However, the rapid and widespread dispersal of *Callinectes sapidus* from introduced areas may also be influenced by long-distance migrations from established populations [23,27]. Adult blue crabs, known for their swimming abilities, can cover substantial distances, with females capable of moving several hundred kilometers [27].

Remarkably, the potential for dispersal of this species not only facilitates the spread of individuals, but also raises concerns about the transmission of pathogens and viral diseases. For instance, the dinoflagellate *Hematodinium* sp. commonly associated with blue crabs can lead to significant mortality in marine crustaceans. The transmission of this parasite from introduced blue crabs to native species can have detrimental effects on indigenous populations [6,72,75–78]. Another pathogen of concern is the *Callinectes sapidus* reovirus 1 (CsRV1), which was first isolated in blue crabs in the 1970s. While the mechanisms of its transmission and evolutionary dynamics remain largely unknown, CsRV1 has been associated with mortality events in blue crab populations [79–81].

Despite the ecological and economic importance of blue crab, genetic research on this species is so far limited. Over the years, several studies have contributed to our understanding concerning the genetic structure and diversity of blue crab populations in both native and introduced ranges. McMillen-Jackson et al. (1994) conducted a pioneering genetic survey of *Callinectes sapidus* across its native range in the United States of America [82]. Place et al. (2005) determined the complete sequence of the blue crab's mitochondrial genome [83], while Yednock and Neigel (2014) assessed its genetic structure within the Gulf of Mexico, focusing on nuclear protein-coding genes [84]. More recently, Rodrigues et al. (2017) explored the genetic structure of *Callinectes sapidus* populations in Brazil and North America, revealing two distinct lineages in the southern and northern parts of its distribution [85]. Windsor et al. (2019) reported the first analysis of genetic variation across the entire natural range, identifying two lineages within the United States of America to Venezuela area and a third in the region spanning from Brazil to Argentina [86].

Regarding the Mediterranean invasion, Gonzalez-Ortegon et al. (2019) analyzed Atlantic blue crab populations along the Mediterranean coast, discovering two mitochondrial haplotypes (CSWM1 and CSWM2), with varying distribution [68]. Schubart et al. (2023) delved into genetic diversity and gene flow in Mediterranean regions inhabited by *Callinectes sapidus*, confirming three distinct lineages (North American, Caribbean, and South American) in its native range [69]. The same authors also reported that the Mediterranean invasion exhibited low genetic diversity, indicating a recent genetic bottleneck associated with the species' introduction. The presence of only two frequent haplotypes (CSWM1 and CSWM2) in the Mediterranean suggested a founder effect, emphasizing the potential role of introduction events and environmental factors [69]. The Turkish population, with high genetic diversity, stood out as an exception, likely influenced by a longer divergence time and intentional introductions for enhancement of fisheries [69]. Accordingly, Vecchioni et al. (2022) conducted a focused study on the findings of *Callinectes sapidus* individuals in two rivers of southern Sicily [87]. Subsequently, in 2023, Vella et al. aimed to provide updated records and distribution information regarding *Callinectes sapidus* in Malta and along the eastern coast of Sicily [88].

In such a context, this study aimed to provide first insights into the invasion dynamics of *Callinectes sapidus* across both its native and introduced ranges, with a focus on the Mediterranean areas. To achieve this goal, we assessed the levels of genetic variability within and among *Callinectes sapidus* populations from the Atlantic and Mediterranean coastal regions. To place our Mediterranean samples within a broader geographic context, we analyzed their mitochondrial sequences alongside those reported in the literature to shed new light on phylogeographic patterns. Additionally, species delimitation methods, supported by phylogenetic analysis tools, were employed to enhance our understanding of the evolutionary history of blue crab populations and to evaluate the taxonomy of this species across a large part of its geographical distribution range. Our genetic analyses were based on the analysis of the mitochondrial Cytochrome c Oxidase gene's subunit I (COI), as, for *Callinectes sapidus*, it is the molecular marker that offers the most extensive sequence data available for comparison in the literature. Our study incorporates over six hundred blue crab sequences representative of various global locations, along with newly obtained sequences from individuals collected in previously underexplored areas of the Mediterranean basin.

This extensive study elucidates the worldwide genetic diversity and evolutionary patterns of *Callinectes sapidus*, providing insights on species' expansion and adaptation to new environments. Future genetic and morphological surveys, along with pathogen monitoring, are essential for a fully comprehension of speciation events and genetic structuring within the blue crab species. These findings may support effective management strategies in areas where the species is not native and provide valuable hints on the taxonomic status of *Callinectes sapidus*.

2. Materials and Methods

2.1. Sample Collection

A total of 36 individuals belonging to the *Callinectes sapidus* species was initially identified through morphological observation [44] and collected from various sites along the coasts of the Island of Sardinia and the continental coastal side of Greece in the Aegean Sea (Figure 1 and Supplementary Table S1). Between July 2022 and July 2023, 21 specimens were obtained from four different areas in the northern, central, and southern areas of Sardinia (refer to Figure 1). Among these, 14 individuals were captured in St. Teodoro Lagoon (SS), where the presence of blue crabs is reported for the first time in the present study. Additionally, 2 individuals were collected from Calich lagoon, 1 from the mouth of the Tirso river, and 4 from Santa Gilla lagoon (details provided in Supplementary Table S1). Furthermore, during the month of November 2020, 15 samples were gathered from two locations in the North area of continental Greece: 7 from the Thermaikos Gulf, Chalastra, and 8 from the Rodopi region in Vistonikos Bay. The samples were obtained by collecting a piece of tissue from the walking legs of blue crabs caught by local professional fishermen using trammel nets at the specified locations before being distributed in the local market. No animals were sacrificed in the sampling campaign process, and the sampling method employed was non-invasive. Tissues were stored in absolute ethanol and used for DNA extraction.

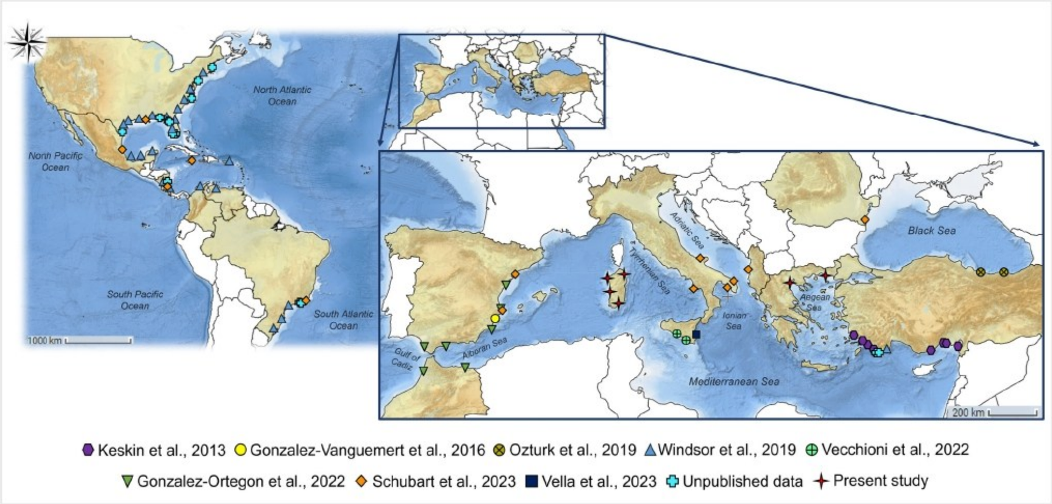


Figure 1. *Callinectes sapidus* sampling plan. The map illustrates the geographical locations where the COI sequences analyzed in the current study were obtained, as well as the locations of sequences from previous research [67–69,86–90].

This sampling approach received approval from the Ethics Committee ("Organismo Preposto al Benessere e alla Sperimentazione Animale—OPBSA") of the University of Sassari, under the Protocol number 82235, dated 19 July 2024.

2.2. Diagnostic Molecular Analysis

Genomic DNA extraction was performed from a muscle tissue segment using the Macherey-Nagel Nucleo Spin Tissue Kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany) in accordance with the manufacturer's guidelines. Quantification of DNA solutions was conducted using the Nanodrop™ Lite Spectrophotometer (Thermo Scientific; Waltham, MA, USA), revealing an average yield of approximately 30 ng/µL. A fragment of the mitochondrial Cytochrome c Oxidase gene's subunit I (COI) was amplified through standard PCR using universal primers [91]. Reactions were executed in a total volume of 25 µL, with an average combination of 10 ng of total genomic DNA, 0.6 µM of each primer, and one PuReTaq Ready-To-Go PCR bead (GE Healthcare, Wauwatosa, WI, USA), encompassing stabilizers, bovine serum albumin (BSA), deoxynucleotide triphosphates, 2.5 units of PuReTaq DNA polymerase, and reaction buffer. Upon reconstitution of a bead to a final volume of 25 µL, the concentrations of each dNTP and MgCl2 were set at 200 µM and 1.5 mM, respectively. PCR cycles were carried out in a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems, Waltham, MA, USA) following the specified program: 1 cycle of 4 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at 51 °C, and 30 s at 72 °C. Post-treatment involved 10 min at 72 °C and final cooling at 4 °C. Positive (high-quality DNA samples) and negative controls were employed to validate PCR protocol effectiveness and check for possible contaminations. Electrophoresis was conducted on 2% agarose gels, created with 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 with ExoSAP-IT (USB Corporation, Cleveland, OH, USA). Sequencing was performed for both the forward and reverse strands using the same primers employed in the PCR process, and this was carried out by an external sequencing core service (Macrogen Europe, Milan, Italy).

2.3. Phylogeographic and Phylogenetic Analysis

Through the Basic Local Alignment Search Tool (BLAST) analysis on the GenBank nucleotide database (www.ncbi.nlm.nih.gov, accessed on 6 May 2024), all newly acquired sequences were confirmed to belong to the species *Callinectes sapidus*, exhibiting a 100% identity match.

Sequence alignment was performed using the Clustal Omega 1.2.4 package [92] (accessible at https://www.ebi.ac.uk/Tools/msa/clustalo/, accessed on 6 May 2024). Manual inspection and editing were carried out using Unipro UGENE v.35 (developed by the Unipro Center for Information Technologies, Novosibirsk, Russia) [93].

A total of 36 blue crab's sequences spanning the mitochondrial COI region (633 bp) were obtained in the present study (deposited in GenBank under the accession numbers PQ067263–PQ067298). For conducting phylogenetic and phylogeographic analyses within a broader geographic context, these sequences were aligned with 631 COI sequences, reported in literature in previous studies and available on GenBank [67–69,86,87,89,90], belonging to blue crabs from the North Atlantic (United States of America—USA and United Mexican States—Mexico)—including the Caribbean Sea (Puerto Rico, Jamaica, Costa Rica, Nicaragua, and Colombia), South Atlantic (Venezuela and Brazil), Western Mediterranean (Spain and Italy), Eastern Mediterranean (Greece and Turkey) and Black Sea (Turkey). Moreover, a subset of 86 sequences was created by combining all Sardinian and Greek sequences acquired in the current study with only those from Turkey, Spain, peninsular Italy, and Sicily (refer to Supplementary Table S2 and Figure S1 for details). This merging aimed to create a wide scenario to depict the phylogeographic and phylogenetic patterns of *Callinectes sapidus* across the Mediterranean and Black Sea regions.

Genetic variation among sequences was evaluated by estimating key parameters, including the number of polymorphic sites (S), number of haplotypes (H), nucleotide diversity (π), and haplotype diversity (h). This analysis was conducted using the software package DnaSP 6.12.03 (developed by Universitat de Barcelona, Barcelona, Spain) [94].

The software JmodelTest 2.1.7 [95] was employed to identify the probabilistic model of sequence evolution that best fit the sequence data, with a maximum likelihood optimized search by the Akaike (AIC) and Bayesian Information Criterion (BIC). The GTR + I + G was suggested by both criteria as the best-fitting model for the whole dataset.

A median-joining network [96] was generated using the software package Network 10.2.0.0 (available at www.fluxus-engineering.com, accessed on 9 May 2024) (Colchester, UK) to deduce genetic relationships among haplotypes and identify potential discrete genetic clusters. Both transitions and transversions were given equal weighting. Given the uncertainty regarding the occurrence of retromutation events, each observed polymorphism was assigned the same weight (10).

To assess the reliability of the entire dataset for taxonomic and phylogenetic investigations, the phylogenetic signal was evaluated using the software TREEPUZZLE 5.3 (Wien, Austria) [97]. A likelihood-mapping analysis of 10,000 random quartets was conducted following Scarpa et al. [98,99]. The Likelihood Map test partitioned the dataset into quartets, representing the smallest set of taxa with multiple unrooted tree topologies [100]. Quartet puzzling operates on groups of four sequences to create a map assessing data reliability for phylogenetic and taxonomic inferences. The key information in the map is the percentage of star-like trees, where a value exceeding 30% indicates that the dataset may not be suitable for analysis due to factors such as noisy data, alignment errors, recombination events, a lack of informative sites, or inadequate taxonomic coverage [98,99].

Phylogenetic relationships were determined through the construction of a Bayesian phylogenetic species tree, generated with the software MrBayes 3.2.7 [101]. The model parameters were set as follows: NST = 6, rates = invgamma, ngammacat = 4. Two independent runs were conducted, each comprising four Metropolis-Coupled MCMC chains (one cold and three heated chains), for a total of 5,000,000 generations. Trees were sampled every 1000 generations, and the initial 25% of the 10,000 sampled trees were discarded as burn-in. Convergence of chains was assessed by ensuring that the Average Standard Deviation of Split Frequencies (ASDSF) approached 0 [101] and the Potential Scale Reduction Factor (PSRF) was around 1 [102], following the methodology outlined by Scarpa et al. [103]. The resulting phylogenetic tree was visualized and edited using FigTree 1.4.0 (accessible at http://tree.bio.ed.ac.uk/software/figtree/, accessed on 15 June 2024).

To validate the taxonomic assessment of the specimens, three species delimitation methods were employed. Firstly, the Automatic Barcode Gap Discovery (ABGD) method [104] was applied. It relies on K2P genetic distances [105] and does not consider the phylogenetic relationships within the dataset. It operates on sequences and identifies the barcode gap as the initial significant gap beyond a specified limit and utilizes it to partition the data [104]. Species evaluation was conducted using the ABGD online tool, accessible at http://www.abi.snv.jussieu.fr/public/abgd/abgdweb.html (accessed on 15 June 2024), with default settings. The appropriate species estimate was chosen following the guidelines of Puillandre et al. [104], employing gene-specific priors for the maximum divergence of intraspecific diversity, corresponding to $p = 0.001$ (refer to Scarpa et al. [106]). The second method used was the nucleotide divergence threshold (NDT), which is based on genetic distances as well. It works on sequences to categorize taxa into taxonomic groups by applying a fixed threshold of 2% established by Hebert et al. (2003) for DNA barcodes [107]. This method employs the pairwise Kimura [105] two-parameter model (K2P) to calculate the genetic distance matrix. The analysis was conducted using a script [98] written in the R statistical environment (available at https://cran.r-project.org, accessed on 15 June 2024). The last method employed for species delimitation was the Poisson Tree Processes (PTP) model and its Bayesian implementation, known as bPTP [108], implemented on the PTP web server (available at http://species.h-its.org/ptp/, accessed on 15 June 2024). This method identifies species boundaries through the phylogenetic species concept (PSC), distinguishing between population and speciation patterns on a given species tree [109].

To conduct this test, the Bayesian species tree constructed using the software MrBayes 3.2.7 [101] was utilized.

2.4. Estimation of Divergence Time

The software package Beast 1.10.4 [110] was utilized to estimate the divergence time for the clades identified in the phylogenetic tree. The evolutionary rate of the COI locus in arthropods, which ranges between 0.7% and 2.0% per Myr, with a mean rate of 1.35 per Myr [111,112], was used for calibrating the time tree. The site parameters were configured as follows: Substitution Model = GTR, Bases Frequencies = Estimated, Site Heterogeneity Model = Gamma + Invariant Sites, Number of Gamma Categories = 4, aligning with the evolutionary model chosen by jModeltest. The lognormal uncorrelated relaxed clock model, which assumes independent rates on different branches, was chosen for the molecular clock rate variation, and a "uniform distribution" was selected. The Yule prior process was applied to the tree prior, reflecting the speciation model. Priors for model parameters and statistics were determined to calibrate the time tree based on mutation rates per million years. Operator parameters were set in accordance with the instructions provided in the user manual. Additionally, the utilization of the lognormal uncorrelated relaxed clock model offers insights into the clock-like nature of the data, measured by the ucld.stdev parameter. If the ucld.stdev parameter estimate is close to 0, it indicates that the dataset is relatively clock-like; if the estimated value is significantly greater than 1, it suggests substantial rate heterogeneity among lineages. To achieve an effective sample size (ESS) exceeding 200 for all statistical parameters, a run was conducted with 200,000,000 generations, with a tree sampled at intervals of 20,000 generations, as per the methodology described by Scarpa et al. [103]. The software Tracer 1.6 (©2022 BEAST Developers. All rights reserved.) was employed to examine the resulting log file, aiming to confirm the convergence of parameter values, verify ESS values exceeding 200, and estimate node ages [113]. Tree Annotator (Beast package) and FigTree 1.4.0 were then employed for the drawing and visualization of the time-calibrated tree, following the procedures detailed by Scarpa et al. [103].

3. Results

3.1. Phylogeographic and Phylogenetic Relationships among *Callinectes sapidus* Populations from the Whole Distribution Range

A total of 124 polymorphic sites were identified out of 667 sequences analyzed, resulting in 196 distinct COI haplotypes (Table 1). Across the dataset, the highest levels of genetic variation were observed in North American populations, particularly in the Atlantic coast of the United States of America (USA) and in the Mexican coast of the Gulf of Mexico. A general lower rate of genetic diversity was found for Central (Caribbean Sea coasts) and South (Atlantic coasts) America. Within the Mediterranean populations, genetic variation was notably low, except for the Levantine Sea Turkish samples, which displayed a level of variation similar to that of North America.

The network analysis conducted on the entire dataset (Figure 2) revealed a genetic structuring between two primary haplogroups (i.e., groups of haplotypes sharing a common ancestor), designated as groups A and B. Group A included two sub-haplogroups, named A-1 and A-2. Sub-haplogroup A-1 encompassed individuals from South America (Colombia, Venezuela), Central America (Nicaragua and Costa Rica), and North America (USA, Mexico, Puerto Rico, and Jamaica) with a predominant star-like shape. Notably, significant genetic variation was observed in these regions due to the presence of several distinct haplotypes, originating from and surrounding the most frequent haplotype, which was shared by 51.30% of South American (Colombia and Venezuela) individuals and 8.20% of North American (USA, Mexico, Puerto Rico, and Jamaica) individuals. In contrast, sub-haplogroup A-2 was private to 14 Brazilian individuals and diverged for 7-point mutations from A-1. Specifically, it was characterized by one common haplotype and three derived lineages. Group B differed from Group A by 14-point mutations. Notably, 10 of these mutations were silent, meaning they did not result in any amino acid change within COI gene haplotypes. One mutation resulted in missense substitutions, replacing the codon GGC encoding for glycine with two different codons: GTG for valine and AGT for serine. These substitutions can have varying effects on the chemical and physical structure of the protein. The remaining 3 mutations were characterized by degenerate bases, making their precise identification and effect on protein synthesis unattainable. Group B included two sub-haplogroups, B-1 and B-2. Sub-haplogroup B-1 was constituted by samples from North America (USA and Mexico) (likely biased by the higher number of sequences from these areas in the dataset). It also encompassed samples from the Levantine Sea (Turkey), the Aegean Sea (Greece), the Western Mediterranean (Spain, peninsular Italy, Sardinia, and Sicily), the Black Sea (Turkish samples from the Black Sea region), and, to a lesser extent, South America (Venezuela and Brazil). Specifically, sub-haplogroup B-1 included 5 widely distributed haplotypes (named from H1 to H5, in Figure 2), with the most common one (H1) centrally positioned and shared by individuals from North America (USA and Mexico) and a few from South America (Brazil). Surrounding this most common lineage, several derived haplotypes differed by 1- to 3-point mutations, according to a typical star-like shape. One of the five common haplotypes (H2) corresponded to the CSWM1 haplotype, previously reported by Gonzalez-Ortegon et al. (2022) and Schubart et al. (2023) [68,69]. In the current study, CSWM1 was shared by specimens from the Western Mediterranean, Aegean, and Levantine Sea regions, along with a few individuals from the Black Sea coastline. This haplotype was notably prevalent among Mediterranean Turkish specimens (13.9%), 3 individuals from Spain, 52.40% of Sardinian specimens, all Italian peninsular individuals, 2 Sicilian individuals, and all continental Greek samples. Conversely, the CSWM2 haplotype (H3) [68,69] was shared by 2 individuals from Spain, 2 from Sicily, 10 from Sardinia, and 2 from North America (USA). Furthermore, several samples from the southwestern Mediterranean coastline of Turkey exhibited divergent haplotypes that formed the sub-haplogroup B-2. This sub-haplogroup showed the highest genetic affinity with those haplotypes belonging to B-1, which were found in North America (USA). B-2 diverged from these North American haplotypes by 5-point mutations.

Table 1. Sample sizes and genetic diversity estimates obtained for the mitochondrial COI region analyzed for *Callinectes sapidus* individuals. N: sample size; S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; π: nucleotide diversity. Sites with gaps were not considered.

Geographic Region	N	S	H	h	π
United States of America	397	100	146	0.922	0.00766
Puerto Rico	22	34	10	0.749	0.00600
Jamaica	1	0	0	0.000	0.00000
United Mexican States	43	38	22	0.844	0.01777
North America	463	108	162	0.931	0.01249
Nicaragua	4	2	3	0.833	0.00184
Costa Rica	1	0	0	0.000	0.00000
Central America	5	4	4	0.900	0.00284
Venezuela	90	45	30	0.614	0.00216
Colombia	4	2	3	0.833	0.00158
Brazil	19	22	7	0.667	0.01168
South America	113	53	39	0.729	0.00712
Turkey—Black Sea	3	5	3	1.000	0.00527
Greece	15	0	1	0.000	0.00000
Turkey—Levantine Sea	36	46	11	0.744	0.01317
Eastern Mediterranean	51	45	11	0.733	0.01289
Spain	5	3	2	0.600	0.00292
Peninsular Italy	2	0	1	0.000	0.00000
Sardinia	21	3	2	0.524	0.00248
Sicily	4	3	2	0.500	0.00237
Western Mediterranean	32	3	2	0.508	0.00247
Whole Mediterranean	83	45	12	0.725	0.01103
Whole dataset	667	124	196	0.937	0.01793

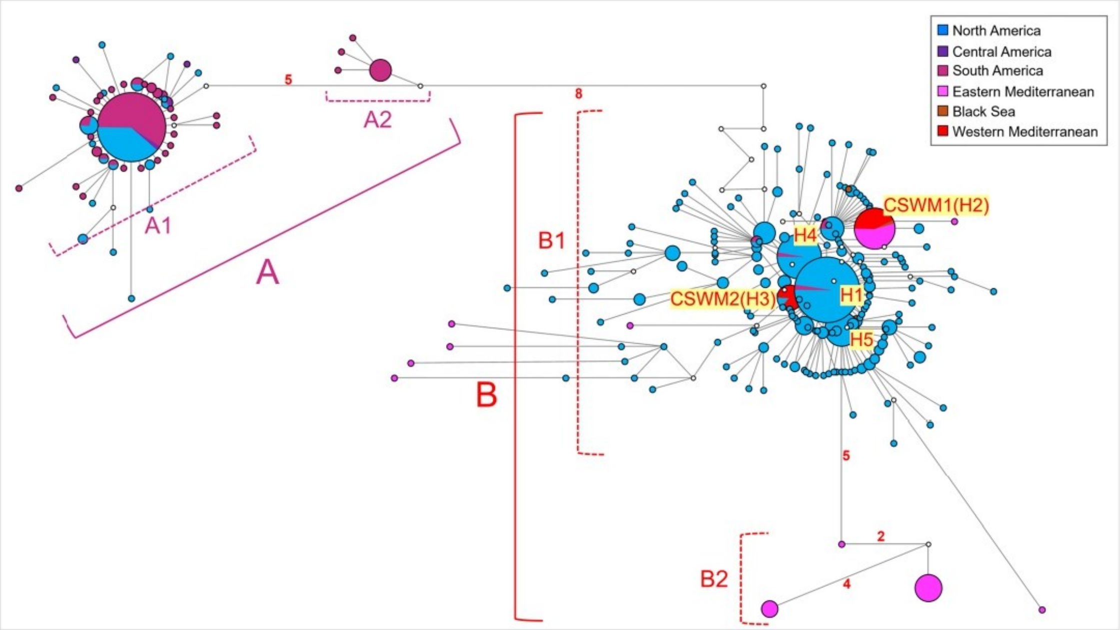


Figure 2. Network (median-joining) analysis performed on the whole dataset of *Callinectes sapidus* mitochondrial COI sequences. Haplogroups A and B and sub-haplogroups A1–A2 and B1–B2 are described in the text. Each circle represents a unique haplotype, with the diameter proportional to its frequency. The number of mutations higher than 1 is reported along the branch length. The small white circles show median vectors, representing intermediate missing or unsampled haplotypes. Network colors indicate the localities of populations investigated, according to the legend.

Likelihood-mapping analysis, conducted on the entire dataset, reveals a strong phylogenetic signal with a value of white noise of 25.6%, which is below the rejection threshold of 33% [114] (see Supplementary Figure S1).

Both Bayesian phylogenetic tree and time tree analyses conducted on the complete dataset yielded congruent topologies at the major nodes, fully supported by values of posterior probabilities (pp) of 1. Consequently, to illustrate these findings, a schematic representation (see Figure 3) of the phylogenetic tree generated through MrBayes was crafted, incorporating divergence time information derived from the Beast tree (see Supplementary Figures S2 and S3). Notably, within this representation, emphasis was placed on two distinct monophyletic clades (groups A and B), thereby reinforcing the outcomes of the above-reported genetic clustering analyses (see Figure 2). Groups A and B separated from a common ancestor, which dated back to about 710,000 ya (95% HPD C.I. = 252,000–918,000). Group A, originating approximately 500,000 ya (95% HPD C.I. = 146,000–642,000), included two internal sister clusters (A-1 and A-2). The first, A-1, was predominantly composed of samples from South America (Colombia and Venezuela), with additional representation from Central America (Nicaragua and Costa Rica) and North America (USA, Mexico, Puerto Rico, and Jamaica). The second internal sister cluster, A-2, included most of Brazilian samples. On the other hand, Group B included individuals from North America (USA, Mexico, and Puerto Rico), the Mediterranean region (south-western Turkey, Greece, Spain, peninsular Italy, Sardinia, and Sicily), and the Black Sea region (northern Turkey), along with some remaining samples from South America (Venezuela and Brazil). Group B originated contemporaneously with Group A, dating back to approximately 510,000 ya (95% HPD C.I. = 199,000–696,000). Within Group B, Mediterranean and Black Sea samples were grouped together with most North American and a few South American samples in a large polytomic clade, which included some internal subclusters. Interestingly, a well-supported monophyletic subcluster was identified within this polytomic clade, corresponding to the sub-group B-2 previously identified in the network analysis (Figure 2). This subcluster encompassed the majority of Mediterranean sequences from south-western Turkey. Notably, these Turkish samples displayed an incipient divergence process from the main Mediterranean polytomic clade within Group B. The genetic structuring highlighted in the phylogenetic analyses is consistent with the former clustering analyses, thereby corroborating the outcomes already described for this dataset.

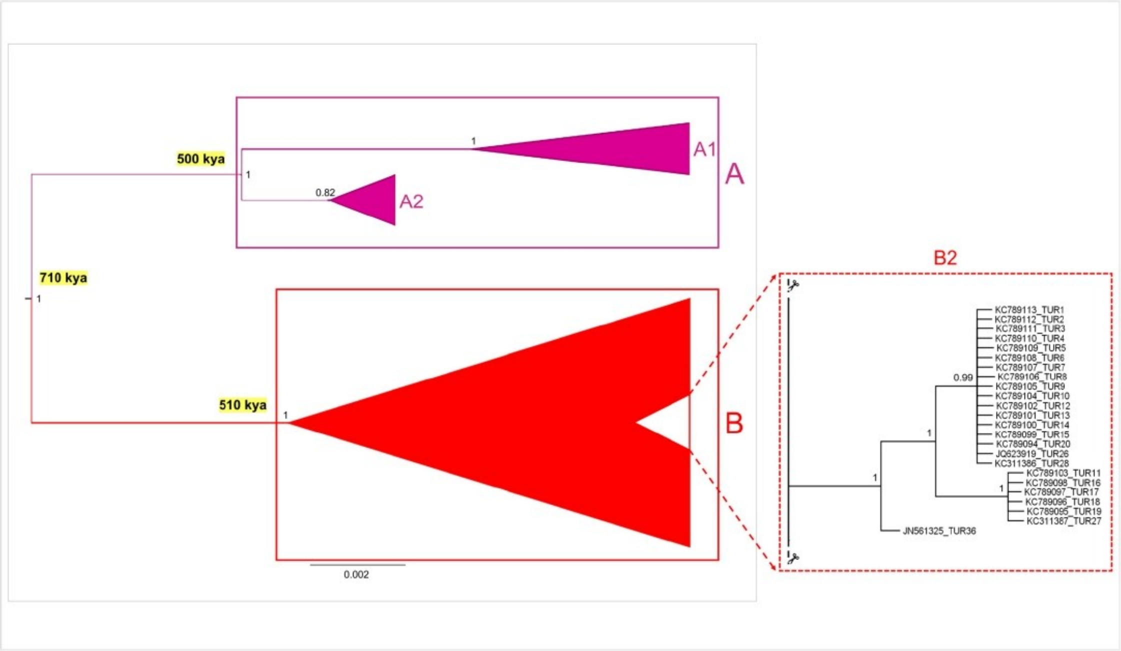


Figure 3. Schematic representation of the Bayesian phylogenetic tree performed on the whole dataset of *Callinectes sapidus* mitochondrial COI sequences (provided at Supplementary Figure S2), integrated with the divergence time estimations at the major nodes (provided at Supplementary Figure S3). Groups A and B and internal clusters A1–A2 and B2 are described in the text. The values at the nodes are represented as posterior probabilities.

All the species delimitation methods (see Supplementary Table S3) performed on the dataset confirmed the genetic structuring observed between the two main groups of sequences mentioned above. They identified two distinct taxonomic entities, precisely encompassing the sequences structured within the two groups, A and B, already detected in the above reported analyses.

To assess the genetic variation levels within groups A and B, parameters of genetic divergence, such as the number of polymorphic sites (S), number of haplotypes (H), nucleotide diversity (π), and haplotype diversity (h), were calculated also for these two putative taxonomic entities (see Table 2). Group A, consisting of 176 sequences, displayed 50 haplotypes and a lower genetic variation in respect to Group B, which displayed 156 haplotypes out of 491 sequences.

Table 2. Sample sizes and genetic diversity estimated for the whole dataset of *Callinectes sapidus* mitochondrial COI sequences categorized into the two groups A and B retrieved by genetic clustering and phylogenetic analyses. N: sample size; S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; π: nucleotide diversity. Sites with gaps were not considered.

Group	N	S	H	h	π
A	176	60	50	0.685	0.00349
B	491	110	156	0.933	0.00699

3.2. Evolutionary Relationships among Callinectes sapidus *Populations, along with Species of the Genus* Callinectes

A Bayesian analysis was performed to explore the evolutionary relationships among *Callinectes sapidus* and the other species within the genus *Callinectes*. The analysis generated a phylogenetic tree based on a dataset that included sequences of *Callinectes sapidus* categorized in the two groups (A and B) identified in this study, along with all available sequences belonging to *Callinectes* species from GenBank. Additionally, an outgroup sequence of *Carcinus maenas* was included in the dataset (Supplementary Figure S4).

The results revealed two sequences of *Callinectes bellicosus* that separated from a monophyletic clade containing all the other sequences of the dataset, divided into two main clusters. This monophyletic clade was fully supported at the major nodes by a value of posterior probabilities (pp) of 1 and at the internal nodes by values of posterior probabilities ranging between 0.70 and 1. The main internal cluster of sequences encompassed *Callinectes sapidus* and other *Callinectes* species and included two internal sister subclusters. The first corresponds to a monophyletic clade encompassing sequences of *Callinectes danae*, *Callinectes similis*, *Callinectes arcuatus*, *Callinectes ornatus*, *Callinectes exasperatus*, and *Callinectes pallidus*. The second corresponded to a large monophyletic clade grouping sequences of *Callinectes sapidus* along with sequences belonging to other species of the genus *Callinectes*. Within this clade, *Callinectes larvatus* and *Callinectes marginatus* branched off separately from *Callinectes sapidus* species, which clustered together with *Callinectes bocourti*, *Callinectes rathbunae*, *Callinectes amnicola*, and *Callinectes toxotes*.

3.3. Genetic Variation of Callinectes sapidus *within Its Mediterranean and Black Sea Ranges of Distribution*

The network analysis conducted on the subset of samples from the Mediterranean and Black Sea regions revealed significant genetic structuring between two distinct groups, Group 1 and Group 2, both exhibiting notable founder effects as suggested by the star-like shapes represented in the graphic (Figure 4). Group 1 consisted of 24 Turkish individuals, mainly from the south-western Mediterranean area of the country, showing a divergence of 6-point mutations from Group 2. In contrast, Group 2 comprised two main common haplotypes. The first haplotype, corresponding to the already reported CSWM1 [68,69], was shared by 39 individuals, mainly from Sardinia and continental Greece, with some representation from Spain, peninsular Italy, Sicily, and the Levantine Mediterranean coast of Turkey. The second haplotype, corresponding to the already reported CSWM2 [68,69], was found in 14 individuals, mainly from Sardinia and a few from Spain and Sicily as well. Additionally, within Group 2, several private haplotypes, corresponding to the Levantine and Black seas Turkish sequences, diverged by 2- to 10-point mutations from the main Mediterranean haplotypes.

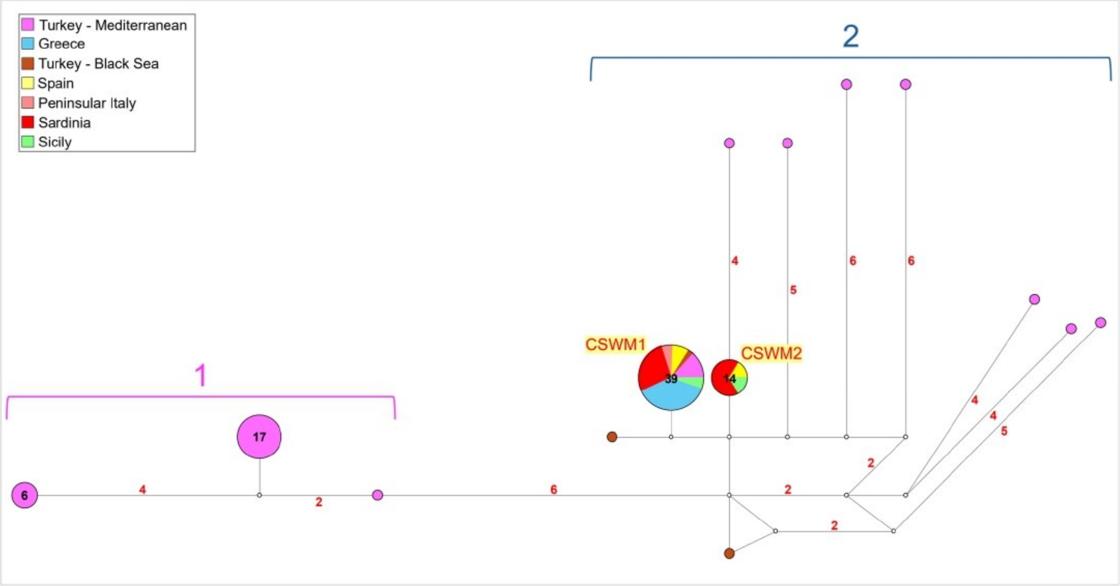


Figure 4. Network (median-joining) analysis performed on the Mediterranean and Black Sea subsets of mitochondrial COI sequences analyzed for *Callinectes sapidus*. Groups 1 and 2 are described in the text. Each circle represents a unique haplotype, with the diameter proportional to its frequency. The number of mutations higher than 1 is reported along the branch length. The small white circles show median vectors, representing intermediate missing or unsampled haplotypes. Network colors indicate the localities of populations investigated, according to the legend. The number of individuals sharing the same haplotype, with a frequency greater than 1, is displayed within the circle.

4. Discussion

The current study represents an extensive molecular investigation conducted on the blue crab (*Callinectes sapidus*) populations, encompassing a dataset sourced from the Atlantic coasts of North, Central and South America, and various Mediterranean and Black Sea regions hitherto underexplored. Notably, this study examined the distribution of genetic variability within the Mediterranean Sea, acquiring sequences from continental Greece and Sardinia in greater abundance compared to extant sequences from the Mediterranean already reported in literature [67–69,86,87,89,90].

Combined phylogenetic and species delimitation analyses led to the most striking result of the ongoing research, which is the appraisal of two closely related, but sharply distinct, genetic groups within a large part of the distribution range of *Callinectes sapidus*. Such a finding suggests the existence of two possible sister species, indicating the occurrence of a species complex for *Callinectes sapidus*.

Accordingly, clustering analyses have highlighted the presence of two strongly differentiated genetic groups, separating most of the South Atlantic American blue crab populations from those in the North Atlantic, Mediterranean, and Black Sea. Phylogenetic analyses have further corroborated this finding by identifying two distinct *Callinectes sapidus* mitochondrial lineages. Notably, Schubart et al. (2023) have already recognized genetic disparities between the North and South Atlantic blue crab populations through mitochondrial COI gene analyses [69]. This previous study reported restricted gene flow within the two regions, likely due to repeated isolation in water bodies of different sizes. Similarly, Rodrigues et al. (2017) reported two distinct lineages in the northern and southern parts of blue crab American distribution [85], while Windsor et al. (2019) identified two lineages within the United States of America to Venezuela area and a third lineage from Brazil to Argentina [86], both using the mitochondrial COI gene as molecular marker. It is worth mentioning that the study by Rodrigues et al. [85] employed a portion of the COI gene that did not align with our dataset, which is why their sequences were not included in our current analysis.

Interestingly, phylogenetic analysis gave us the possibility to evidence the contemporary divergence of the two *Callinectes sapidus* putative sister species on the North and South Atlantic American coasts, which originated around 500,000 years ago. We hypothesized that these two sister taxa were raised from a common ancestor that likely spread along the entire American Atlantic coastline after its early origin about 700,000 years ago. Interestingly, this evolutionary period coincided with the Early-Middle Pleistocene Transition (EMPT), a significant climatic shift occurring around 1.2 million years ago [115]. The EMPT, previously known as the Mid-Pleistocene Transition, was characterized by intensified climate oscillations with 100,000-year periodicity, notable asymmetry in global ice volume cycles, and a reduction in the duration of interglacial periods starting from 1.4 million years ago [115]. This transitional phase persisted until about 420,000 years ago, after which high-amplitude interglacial periods occurred. Marine records from this period revealed high-frequency climate variability influenced by factors such as precessional forcing, changes in ice sheet dynamics, temperature variations, and alterations in oceanic currents [115]. Notably, during the interval from 712,000 to 676,000 years ago, an interglacial period occurred, with an extended deglaciation and sustained interglacial conditions spanning across two summer insolation peaks [115].

Furthermore, the patterns of genetic diversity between the two taxonomic entities we observed in *Callinectes sapidus* can be explained by the environmental conditions during the Marine Isotope Stage 13 (MIS 13) period, approximately from 533,000 to 478,000 years ago. During MIS 13, the Northern Hemisphere experienced a robust summer monsoon and warmer interglacial conditions, marked by rising temperatures and declining ice volumes [115,116]. These environmental changes likely had a significant impact on the migration patterns and adaptation of various tropical marine organisms, including invertebrates, temperate fish, and marine mammals [117].

For what the blue crab is concerned, in the southern regions of American coastline, the observed lower genetic diversity compared to the Northern Atlantic area can be attributed to a positive reinforcement of the two distinct evolutionary forces: genetic drift, as the founder effect, and natural selection, as the selective sweep. Indeed, the warmer and more favorable environmental conditions during MIS 13 may have led to the selection and the spreading of specific, adaptive haplotypes in the blue crab southern populations. This rapid adaptation and massive expansion of the adaptive haplotypes resulted in a reduction of the overall genetic diversity in the southern region, as the population became dominated by positive selected lineages. As the adaptive founder haplotypes became established in the southern populations, the evolutionary rate likely gradually decreased over time. In this context, since the blue crab populations had already acclimated to the prevailing environmental conditions, further adaptation was likely unnecessary. Additionally, a continuous and effective gene flow due to larval dispersal across the western South Atlantic, as reported by Lacerda et al. (2016) in modern populations, could have homogenized genetic variation within past populations, further reducing diversity over time [118].

In such a context, our study has evidenced an incipient, yet statistically relevant, mitochondrial genetic structuring between the two putative sister species of the blue crab. Notably, despite these differences, the two species maintain a considerable level of genetic similarity among each other. Furthermore, our results highlighted a relevant genetic divergence between the two putative blue crab sister species that are representative of the *Callinectes sapidus* species complex and the other species belonging to the genus *Callinectes*. It is important to take into consideration that the glacial stage preceding MIS 13 may have affected connectivity between different marine organisms in the North and South

Atlantic. The lowering of sea levels during this period could have created, or increased, the geographic barriers, potentially prompting species diversification [115,117].

Despite the emerging genetic structuring between the North-Central and South American populations of *Callinectes sapidus*, significant morphological distinctions have not been reported to date. This suggests that the mitochondrial incipient speciation events may be related to environmental conditions, but nucleotide diversification has possibly not yet progressed sufficiently to yield detectable morphological differences between the two putative sister species. To fully understand the extent of genetic divergence between North-Central and South American populations, as well as the potential presence of underlying morphological variations that may have been previously undetected, further comprehensive genetic and morphological investigations are essential. Future studies should focus on analyzing both nuclear and mitochondrial markers in *Callinectes sapidus* individuals collected from the entire Atlantic coastline of the Americas.

The abovementioned scenario, which depicts the presence of two genetically divergent putative sister species for *Callinectes sapidus* inhabiting North-Central and South America coastlines, is strongly and independently further corroborated by a study conducted by our research group (Scarpa et al., submitted), which aims to perform a phylodynamic genome-based reconstruction incorporating all *Callinectes sapidus* Reovirus 1 (CsRV1) genomes and segments available in the NCBI virus database. Indeed, it should be taken into consideration that pathogen dispersal is typically influenced by host movement, providing crucial insights into host migration and population connectivity [81,119]. In this context, preliminary results of our phylodynamic analysis on the CsRV1 highlighted a genetic structuring between the viral strains from North Atlantic and South Atlantic American coastlines, which is consistent with the significant genetic divergence evidenced in the present study among the blue crab populations from these two oceanic waterbodies.

Moving our discussion to the Mediterranean and Black Seas, the haplotypes reported for this area in the present study belong to the *Callinectes sapidus* taxonomical entity that we found to be exclusive of North America Atlantic coastline. This observation suggests that the demographic expansion of *Callinectes sapidus* in the Mediterranean might have been facilitated by the introduction of blue crabs into the basin mainly through ballast waters [23,28,41–43,52].

In this context, the predominant Mediterranean and Black Sea haplotypes correspond to, or closely resemble, those distributed in North America. Consistently with Schubart et al. (2023), the genetic landscape evidenced for this area in the present study is characterized by the presence of the two most frequent reported mitochondrial lineages (i.e., the so-called CSWM1 and CSWM2) (see [68,69] for details), along with several closely derived haplotypes suggesting that the occurrence of a marked founder effect may have affected the spreading of blue crabs in the Mediterranean and Black Sea areas [69].

The present study has revealed that the CSWM1 haplotype is predominant in the Mediterranean basin, corroborating the findings of previous research. Specifically, González-Ortegón et al. (2022) reported this haplotype to be prevalent in the Alboran Sea and the Gulf of Cadiz [68], and Schubart et al. (2023) found it to be also widespread in the southern Adriatic and Black Sea [69]. Furthermore, the CSWM1 haplotype was also detected in Sicily [87]. Importantly, the current study has provided evidence of the presence of CSWM1 also in geographic areas that were previously poorly investigated, such as continental Greece and the island of Sardinia. Additionally, in the present study, the CSWM2 haplotype has been found to be predominant along the western Mediterranean coasts of Spain, as previously reported by González-Ortegón et al. (2022) and Schubart et al. (2023) [68,69]. We also detected for the first time this haplotype on the island of Sardinia, and, as it was previously reported [87], on the island of Sicily. These findings contribute to a more thorough understanding of the genetic diversity and distribution of *Callinectes sapidus* within the Mediterranean and Black Sea areas. To further expand this knowledge, future research efforts should focus on collecting data from additional locations in these areas, which would deepen our comprehension of the distribution patterns of the CSWM1, CSWM2, and derivate haplotypes across this large geographic region.

Interestingly, a significant genetic differentiation has been evidenced in the present study between individuals originating from Turkey and those from other Mediterranean areas, despite the relatively close geographic proximity to some other Mediterranean regions, such as the North Aegean (see Figure 4). Particularly, Turkish individuals from the Levantine Mediterranean south-western coastline of the country displayed the highest levels of mitochondrial divergence from the other Mediterranean populations. The high level of genetic divergence in Turkish blue crabs, with the occurrence of at least two large groups of mitochondrial lineages, likely stems from their early arrival in the region around a century ago [67–69]. This predates the much more recent, massive appearance of the species in the western Mediterranean. Although blue crab may have faced difficulties establishing in other parts of the Mediterranean after its initial introductions, the species could have experienced a rapid adaptation in Turkey, where local environmental conditions might have been conducive to its survival and genetic differentiation since its first arrival.

In particular, first sightings of the blue crab in Thermaikos Gulf date back to 1935 [23,41,42], suggesting that vigorous commercial maritime interactions between the Aegean coast of Greece and Mediterranean Turkey might have accelerated the spread of *Callinectes sapidus* in this latter area earlier than in other Mediterranean regions [120].

Consequently, the high level of internal genetic variation observed in the present study among Turkish blue crabs may be the consequence of multiple introductions over the past century that produced an overlapping of different haplogroups (distinct groups of similar haplotypes with different common ancestors) in the country. However, an alternative scenario should also be considered to explain the occurrence of multiple, and divergent, mitochondrial lineages in Turkey. Considering that commercial shipping trade is less intense in southern Turkey compared to other areas of the country and also considering that fisheries and blue crab harvesting represent an important economic national resource, the genetic pattern reported in this study for the Turkish Levantine Mediterranean region may be the result of intensive human-mediated activities that acted as artificial evolutionary forces driving divergence from other Mediterranean areas.

Conversely, the limited genetic variability and the marked signs of founder effects detected for the other Mediterranean regions here analyzed align with a species newly established in the area, predominantly comprising descendants of recently introduced founder individuals. In this context, based on our data, the early reports of blue crabs in the western Mediterranean, dating back to the mid-20th century (i.a., [6,29,48–53,62–66]), might have been followed by a rapid disappearance of the species due to incompatibility with the environmental conditions that, at that time, were not suitable with the expansion of the species in those geographical areas.

On the other hand, the present general mitochondrial homogeneity observed in the modern Mediterranean blue crab populations is likely attributable to the species' high larval dispersal capability [15,69]. This has resulted in the predominance of common haplotypes, likely adapted to environmental conditions, across the region. Moreover, marine currents may have further facilitated the gene flow among blue crab populations, thus contributing to the reduced genetic variability observed among distant Mediterranean sites for this species [121,122]. Additionally, the genetic similarities observed between populations from the Black Sea and the Mediterranean suggest a possible pathway for the introduction of *Callinectes sapidus* into the Black Sea from the northern Aegean Sea, likely via the Gulf of Saros, where sightings have been recorded since 1935 [23,67,69]. Therefore, our results confirm what was proposed by Ozturk et al. in 2020, which hypothesized that individuals of *Callinectes sapidus* migrated from the Saros Bay to establish a thriving population in the Black Sea. Again, in this case, the high dispersal capability of larvae and notable adults' swimming capacity may also account for these results [67,69].

Interestingly, the exclusive haplotypes that are private to the Mediterranean Levantine Sea Turkish population were not found in the modern North American population, which is the likely source of blue crab introductions to Turkey a century ago. This finding could be attributed to the effect of evolutionary forces, such as genetic drift or selective sweep, acting on the first founding population, which was introduced to Turkey in the early 1900s.

This phenomenon is consistent with the fundamentals of adaptive evolution, according to which the molecular lineages of introduced populations can be rare, or became absent, in the source population but prevalent among individuals in the newly colonized region [123]. This is likely due to founder events or natural selection favoring the spread of rare, potentially adaptative alleles, thereby enhancing the species' adaptability to the new environment. This scenario is supported by similar findings in other invasive Mediterranean and European species, such as the teleost *Fistularia commersonii* [124] and the crustacean *Procambarus virginalis* [125], where recently introduced populations lacked mitochondrial lineages present in the source populations. However, another plausible explanation for the existence of private haplotypes in the Turkish population, which are nowadays absent in North America, is the potential extinction of these lineages in their native range during the last century. This extinction could have occurred over time, possibly due to natural cycles of demographic fluctuations in the North American blue crab populations. Consequently, these haplotypes have become endemic and private to the Turkish population only.

5. Conclusions

Our study presents a wide geographical examination of the genetic diversity of the blue crab (*Callinectes sapidus*) populations, encompassing a large mitochondrial dataset from the Americas and previously under investigated European and Mediterranean regions. This research reveals significant insights into the species' evolutionary history and its phylogeography, including the identification of two distinct genetic groups within the blue crab populations, suggesting the potential existence of a never-reported sister species predominantly spanning in southern America. This study also sheds light on the introduction pathways and genetic adaptation of *Callinectes sapidus* in new environments of the Mediterranean and Black Sea. Furthermore, our findings highlight the role of natural selection and genetic drift in shaping the genetic diversity of introduced populations.

In conclusion, this research provides valuable insights into the genetic diversity, evolutionary history, and phylogeographic patterns of *Callinectes sapidus* with important implications for the understanding of the species' expansion across new environments. Future genetic and morphological investigations, along with studies on reproductive biological traits, are essential to deepen our comprehension of the evolutionary processes and taxonomical status of *Callinectes sapidus*. These studies are fundamental for corroborating the occurrence of speciation events within the species complex of *Callinectes sapidus* and for determining whether the two putative sister species could be considered as cryptic or sibling species. Additionally, monitoring pathogens, particularly CsRV1, would be crucial for better depicting the genetic structuring among blue crab populations.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/life14091116/s1. Figure S1: Phylogenetic signal performed on the whole dataset of *Callinectes sapidus* mitochondrial COI sequences. The likelihood-mapping method partitions the area of the equilateral triangle into seven regions. The three trapezoids at the corners represent the areas supporting strictly bifurcating trees, that is, the presence 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 sets of points P (posterior probabilities of the unrooted trees) where all three trees are equally supported (Scarpa et al., 2019) [98]. Figure S2: Bayesian phylogenetic tree obtained by the software MrBayes 3.2.7 and performed on the whole dataset of *Callinectes sapidus* mitochondrial COI sequences. The values at the nodes are represented as posterior probabilities. The sequences obtained in the present study are reported in red (Sardinia) and light blue (Greece). Figure S3: Ultrametric tree obtained by the software Beast 1.10.4 and performed on the whole dataset of *Callinectes sapidus* mitochondrial COI sequences. Divergence times among taxa are shown. The sequences obtained in the present study are reported in red (Sardinia) and light blue (Greece). Figure S4: Bayesian phylogenetic tree obtained by the software MrBayes 3.2.7 and performed, including the mitochondrial COI gene sequences of *Callinectes sapidus* along with sequences from other *Callinectes* species sourced from GenBank. The values at the nodes are represented as posterior probabilities. Table S1: *Callinectes sapidus* specimens collected in the present study. The table reports the samples' code, GenBank accession numbers, sampling locations, sampling date, and species. Table S2: COI dataset. The table shows *Callinectes sapidus* COI sequences used in the present study, taken from the GenBank database. Sample code, GenBank accession numbers, collection locations and period, species, and references are reported. Table S3: Species delimitation results. The table shows the outcomes from the species delimitation methods applied to the whole *Callinectes sapidus* mitochondrial COI dataset. All three methods employed (ABGD, NDT, and PTP) produced identical results. Table S4: *Callinectes* genus COI dataset. The table shows the sample code and the GenBank accession numbers of species belonging to the *Callinectes* genus, taken from the GenBank database.

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

Funding: This research was funded by: 1. "Regione Autonoma della Sardegna" with the project "Attività di ricerca e monitoraggio finalizzate alla cattura e alla quantificazione della specie aliena invasiva granchio blu (*Callinectes sapidus*)" entrusted to the University of Sassari with the Regional law no. 17 of 19 December 2023, article 8, paragraph 1 (CUP E78H23000990002); 2. PNRR, Missione 4, Componente 2 "Dalla ricerca all'impresa", Investimento 1.4 Project CN00000033, which supports the National Biodiversity Future Center of the University of Sassari.

Institutional Review Board Statement: The samples used for this study were obtained by collecting a piece of tissue from the walking legs of blue crabs caught by local professional fishermen using trammel nets at the specified locations before being distributed in the local market. No animals were sacrificed in the sampling campaign process, and the sampling method employed was non-invasive. Tissues were stored in absolute ethanol and used for DNA extraction. This sampling method received approval from the Ethics Committee ("Organismo Preposto al Benessere e alla Sperimentazione Animale—OPBSA") of the University of Sassari, under the Protocol number 82235, dated 19 July 2024.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequences obtained in the present study for the mitochondrial COI gene were deposited in the GenBank database under the accession numbers PQ067263–PQ067298.

Acknowledgments: Daria Sanna would like to express her sincere gratitude to Surveyor Giuseppino Sanna for his never-ending support and guidance throughout their whole journey together. Furthermore, we would like to thank Maria Perra for her support during manuscript preparation and Elisabetta Soi for her support during the sampling campaign.

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

References

1. IUCN; Species Survival Commission. IUCN guidelines for the Prevention of Biodiversity Loss Caused by Alien Invasive Species. *Aliens* **2000**, *11*. Available online: https://portals.iucn.org/library/efiles/documents/Rep-2000-052.pdf (accessed on 19 March 2024).
2. Streftaris, N.; Zenetos, A. Alien marine species in the Mediterranean-the 100 'Worst Invasives' and their impact. *Mediterr. Mar. Sci.* **2006**, *7*, 87–118. [CrossRef]
3. Lucy, F.E.; Caffrey, J.; Dick, J.T.; Davis, E.; Coughlan, N.E. Prevention, Control and Eradication of Invasive Alien Species (No. 368). EPA Research Report. 2021. Available online: https://www.epa.ie/publications/research/biodiversity/Research_Report_368 (accessed on 19 March 2024).
4. European Commission; Directorate-General for Environment; Sundseth, K. Invasive Alien Species—A European Union Response. Publications Office, 2016. Available online: https://data.europa.eu/doi/10.2779/374800 (accessed on 19 March 2024).

5. Genovesi, P. Guidelines for Eradication of Terrestrial Vertebrates: A European Contribution to the Invasive Alien Species Issue. Other Publications in Wildlife Management. 2000. Available online: https://digitalcommons.unl.edu/icwdmother/24 (accessed on 19 March 2024).
6. Mancinelli, G.; Carrozzo, L.; Costantini, M.L.; Rossi, L.; Marini, G.; Pinna, M. Occurrence of the Atlantic blue crab *Callinectes sapidus* Rathbun, 1896 in two Mediterranean coastal habitats: Temporary visitor or permanent resident? *Estuar. Coast. Shelf Sci.* **2013**, *135*, 46–56. [CrossRef]
7. Zenetos, A.; Çinar, M.E.; Crocetta, F.; Golani, D.; Rosso, A.; Servello, G.; Shenkar, N.; Turon, X.; Verlaque, M. Uncertainties and validation of alien species catalogues: The Mediterranean as an example. *Estuar. Coast. Shelf Sci.* **2017**, *191*, 171–187. [CrossRef]
8. Galil, B.S.; Marchini, A.; Occhipinti-Ambrogi, A. East is east and West is west? Management of marine bioinvasions in the Mediterranean Sea. *Estuar. Coast. Shelf Sci.* **2018**, *201*, 7–16. [CrossRef]
9. Vilà, M.; Hulme, P.E. *Impact of Biological Invasions on Ecosystem Services*; Springer: Dordrecht, The Netherlands, 2017. [CrossRef]
10. Mancinelli, G.; Bardelli, R.; Zenetos, A. A global occurrence database of the Atlantic blue crab *Callinectes sapidus*. *Sci. Data* **2021**, *8*, 111. [CrossRef] [PubMed]
11. Hines, A.H.; Haddon, A.M.; Wiechert, L.A. Guild structure and foraging impact of blue crabs and epibenthic fish in a subestuary of Chesapeake Bay. *Mar. Ecol. Prog. Ser.* **1990**, *67*, 105–126. [CrossRef]
12. Laughlin, R.A. Feeding habits of the blue crab, *Callinectes sapidus* Rathbun, in the Apalachicola estuary, Florida. *Bull. Mar. Sci.* **1982**, *32*, 807–822.
13. Kampouris, T.E.; Porter, J.S.; Sanderson, W.G. *Callinectes sapidus* Rathbun, 1896 (Brachyura: Portunidae): An assessment on its diet and foraging behaviour, Thermaikos Gulf, NW Aegean Sea, Greece: Evidence for ecological and economic impacts. *Crustac. Res.* **2019**, *48*, 23–37. [CrossRef]
14. Millikin, M.R.; Williams, A.B. Synopsis of biological data on blue crab, *Callinectes sapidus* Rathbun. In *FAO Fisheries Synopsis*; NMFS (National Marine Fisheries Service): Silver Spring, MD, USA, 1984; Volume 138, pp. 32–39.
15. Marchessaux, G.; Chevalier, C.; Mangano, M.C.; Sarà, G. Larval connectivity of the invasive blue crabs *Callinectes sapidus* and *Portunus segnis* in the Mediterranean Sea: A step toward improved cross border management. *Mar. Pollut. Bull.* **2023**, *194*, 115272. [CrossRef]
16. Marchessaux, G.; Gjoni, V.; Sarà, G. Environmental drivers of size-based population structure, sexual maturity and fecundity: A study of the invasive blue crab *Callinectes sapidus* (Rathbun, 1896) in the Mediterranean Sea. *PLoS ONE* **2023**, *18*, e0289611. [CrossRef] [PubMed]
17. Mizerek, T.; Regan, H.M.; Hovel, K.A. Seagrass habitat loss and fragmentation influence management strategies for a blue crab *Callinectes sapidus* fishery. *Mar. Ecol. Prog. Ser.* **2011**, *427*, 247–257. [CrossRef]
18. Hines, A.H. Ecology of juvenile and adult blue crabs: Summary of discussion of research themes and directions. *Bull. Mar. Sci.* **2003**, *72*, 423–433.
19. Walters, E.A.; Bojko, J.; Crowley, C.E.; Gandy, R.L.; Martin, C.W.; Shea, C.P.; Bateman, K.S.; Stentiford, G.D.; Behringer, D.C. Salinity and temperature affect the symbiont profile and host condition of Florida USA blue crabs *Callinectes sapidus*. *J. Invertebr. Pathol.* **2023**, *198*, 107930. [CrossRef] [PubMed]
20. Marchessaux, G.; Bosch-Belmar, M.; Cilenti, L.; Lago, N.; Mangano, M.C.; Marsiglia, N.; Sarà, G. The invasive blue crab *Callinectes sapidus* thermal response: Predicting metabolic suitability maps under future warming Mediterranean scenarios. *Front. Mar. Sci.* **2022**, *9*, 1055404. [CrossRef]
21. Lewis, D.H.; Roer, R.D. Thermal preference in distribution of blue crabs, *Callinectes sapidus*, in a power plant cooling pond. *J. Crustac. Biol.* **1988**, *8*, 283–289. [CrossRef]
22. García-Rueda, A.; Tremblay, N.; Mascaró, M.; Díaz, F.; Paschke, K.; Caamal-Monsreal, C.; Rosas, C. The thermal tolerance of a tropical population of blue crab (*Callinectes sapidus*) modulates aerobic metabolism during hypoxia. *J. Therm. Biol.* **2021**, *102*, 103078. [CrossRef] [PubMed]
23. Nehring, S. Invasion history and success of the American blue crab *Callinectes sapidus* in European and adjacent waters. In *The Wrong Place—Alien Marine Crustaceans: Distribution, Biology and Impacts*; Galil, B., Clark, P., Carlton, J., Eds.; Springer: Dordrecht, The Netherlands, 2011; pp. 607–624. [CrossRef]
24. Guillory, V.; Perry, H.; Steele, P.; Wagner, T.; Keithly, W.; Pellegrin, B.; Petterson, J.; Floyd, T.; Buckson, B.; Hartman, L.; et al. *The Blue Crab Fishery of the Gulf of Mexico, United States: A Regional Management Plan*; Guillory, V., Perry, H.M., VanderKooy, S., Eds.; Gulf States Marine Fisheries Commission: Ocean Springs, MS, USA, 2001; Publication 96. Available online: https://www.impactassessment.com/reports/Petterson_2001_Sociocultural_Characterization_Gulf_Mexico_Blue_Crab_Fishery.pdf (accessed on 19 March 2024).
25. Churchill, E.P. Life history of the blue crab. *Bull. Bur. Fish* **1919**, *36*, 95–128.
26. Van Engel, W.A. The blue crab and its fishery in Chesapeake Bay. Part 1. Reproduction, early development, growth and migration. *Comm. Fish. Rev.* **1958**, *20*, 6.
27. Hill, J.; Fowler, D.L.; Avyle, M.V. Species Profiles: Life Histories and Environmental Requirements of Coastal Fishes and Invertebrates (Mid-Atlantic)—Blue Crab. U.S. Army Corps of Engineers, Vicksburg, 1989. Available online: https://apps.dtic.mil/sti/pdfs/ADA210181.pdf (accessed on 19 March 2024).

28. Kevrekidis, K.; Kevrekidis, T.; Mogias, A.; Boubonari, T.; Kantaridou, F.; Kaisari, N.; Malea, P.; Dounas, C.; Thessalou-Legaki, M. Fisheries Biology and Basic Life-Cycle Characteristics of the Invasive Blue Crab *Callinectes sapidus* Rathbun in the Estuarine Area of the Evros River (Northeast Aegean Sea, Eastern Mediterranean). *J. Mar. Sci. Eng.* **2023**, *11*, 462. [CrossRef]
29. Gennaio, R.; Scordella, G.; Pastore, M. Occurrence of blue crab *Callinectes sapidus* (Rathbun, 1986, Crustacea, Brachyura), in the Ugento ponds area (Lecce, Italy). *Thalass. Salentina* **2006**, *29*, 29–39. [CrossRef]
30. Lipcius, R.N.; Eggleston, D.B.; Heck Jr, K.L.; Seitz, R.D.; Van Montfrans, J. Post-settlement abundance, survival, and growth of postlarvae and young juvenile blue crabs in nursery habitats. In *The Blue Crab: Callinectes sapidus*; Kennedy, V.S., Cronin, L.E., Eds.; Maryland Sea Grant College: College Park, MD, USA, 2007; Volume 13, pp. 535–565. Available online: https://www.researchgate.net/publication/284418968 (accessed on 20 March 2024).
31. Hines, A.H. Ecology of Juvenile and Adult Blue Crabs. In *The Blue Crab: Callinectes sapidus*; Kennedy, V.S., Cronin, L.E., Eds.; Maryland Sea Grant College: College Park, MD, USA, 2007; pp. 565–654. Available online: https://repository.si.edu/bitstream/handle/10088/18447/serc_Hines_2007_Chapter_14_565_654.pdf (accessed on 20 March 2024).
32. Epifanio, C.E. Early life history of the blue crab *Callinectes sapidus*: A review. *J. Shellfish Res.* **2019**, *38*, 1–22. [CrossRef]
33. Taylor, D.L.; Fehon, M.M. Blue crab (*Callinectes sapidus*) population structure in southern New England tidal rivers: Patterns of shallow water, unvegetated habitat use and quality. *Estuar. Coast.* **2021**, *44*, 1320–1343. [CrossRef] [PubMed]
34. Bouvier, E.L. Sur un *Callinectes sapidus* M. Rathbun trouvé à Rochefort. *Bull. Mus. Hist. Nat. Paris* **1901**, *7*, 16–17.
35. Ribeiro, F.; Veríssimo, A. A new record of *Callinectes sapidus* in a western European estuary (Portuguese coast). *Mar. Biodivers. Rec.* **2014**, *7*, e36. [CrossRef]
36. Eldredge, L.G. First record of the blue crab (*Callinectes sapidus*) in Hawaii (Decapoda: Brachyura). *Bishop Mus. Occas. Pap.* **1995**, *42*, 55–58.
37. Monin, V. New find of the blue crab *Callinectes sapidus* (Decapoda, Brachyura) in the Black Sea. *Zool. Zh.* **1984**, *63*, 1100–1101.
38. Labrune, C.; Amilhat, E.; Amouroux, J.; Jabouin, C.; Gigou, A.; Noël, P. The arrival of the American blue crab, *Callinectes sapidus* Rathbun, 1896 (Decapoda: Brachyura: Portunidae), in the Gulf of Lions (Mediterranean Sea). *BioInvasions Rec.* **2019**, *8*, 876–881. [CrossRef]
39. Mancinelli, G.; Chainho, P.; Cilenti, L.; Falco, S.; Kapiris, K.; Katselis, G.; Ribeiro, F. The Atlantic blue crab *Callinectes sapidus* in southern European coastal waters: Distribution, impact and prospective invasion management strategies. *Mar. Pollut. Bull.* **2017**, *119*, 5–11. [CrossRef]
40. Cerri, J.; Chiesa, S.; Bolognini, L.; Mancinelli, G.; Grati, F.; Dragičević, B.; Dulčic, J.; Azzurro, E. Using online questionnaires to assess marine bio-invasions: A demonstration with recreational fishers and the Atlantic blue crab *Callinectes sapidus* (Rathbun, 1986) along three Mediterranean countries. *Mar. Pollut. Bull.* **2020**, *156*, 111209. [CrossRef]
41. Serbetis, C. Un nouveau Crustacé comestible en Mer Egée *Callinectes sapidus* Rathb. (Décapode brachyoure). *Proc. Gen. Fish. Comm. Mediterr.* **1959**, *5*, 505–507.
42. Georgiadis, C.; Georgiadis, G. Zur Kenntnis der Crustacea Decapoda des Golfes von Thessaloniki. *Crustaceana* **1974**, *26*, 239–248. [CrossRef]
43. Giordani Soika, A. Il *Neptunus pelagicus* (L.) nell'Alto Adriatico. *Natura* **1951**, *42*, 18–20.
44. Mizzan, L. Presence of swimming crabs of the genus *Callinectes* (Stimpson) (Decapoda, Portunidae) in the Venice Lagoon (North Adriatic Sea-Italy): First record of *Callinectes danae* Smith in European waters. *Boll. Mus. Civ. Stor. Nat. Venezia* **1993**, *42*, 31–43.
45. Holthuis, L.B.; Gottlieb, L. The occurrence of the American blue crab, *Callinectes sapidus* Rathbun, in Israel waters. *Bull. Res. Counc. Isr. B* **1955**, *5*, 154–156.
46. Shiber, J.G. Brachyurans from Lebanese waters. *Bull. Mar. Sci.* **1981**, *31*, 864–875.
47. Abdel-Razek, F.A. Crab fishery of the Egyptian waters with notes on the bionomics of *Portunus pelagicus* (L.). *Acta Adriat.* **1987**, *28*, 143–154.
48. Tortonese, E. La comparsa di *Callinectes sapidus* Rath. (Decapoda, Brachyura) nel Mar Ligure. *Ann. Mus. Civ. Stor. Nat. Genova* **1965**, *4*, 1–3.
49. Suaria, G.; Pierucci, A.; Zanello, P.; Fanelli, E.; Chiesa, S.; Azzurro, E. *Percnon gibbesi* (H. Milne Edwards, 1853) and *Callinectes sapidus* (Rathbun, 1896) in the Ligurian Sea: Two additional invasive species detections made in collaboration with local fishermen. *BioInvasions Rec.* **2017**, *6*, 147–151. [CrossRef]
50. Cavaliere, A.; Berdar, A. Presenza di *Callinectes sapidus* Rathbun (Decapoda, Brachyura) nello Stretto di Messina. *Boll. Pesca Piscicolt. Idrobiol.* **1975**, *30*, 315–322.
51. Bisconti, M.; Silvi, E. Prima segnalazione di *Callinectes sapidus* Rathbun, 1896 (Crustacea, Decapoda, Brachyura) nella provincia di Livorno. *Quad. Mus. Stor. Nat. Livorno* **2005**, *18*, 1–6.
52. Stasolla, G.; Innocenti, G. New records of the invasive crabs *Callinectes sapidus* Rathbun, 1896 and *Percnon gibbesi* (H. Milne Edwards, 1853) along the Italian coasts. *BioInvasions Rec.* **2014**, *3*, 39–43. [CrossRef]
53. Carrozzo, L.; Potenza, L.; Carlino, P.; Costantini, M.L.; Rossi, L.; Mancinelli, G. Seasonal abundance and trophic position of the Atlantic blue crab *Callinectes sapidus* Rathbun 1896 in a Mediterranean coastal habitat. *Rend. Lincei* **2014**, *25*, 201–208. [CrossRef]
54. Florio, M.; Breber, P.; Scirocco, T.; Specchiulli, A.; Cilenti, L.; Lumare, L. Exotic species in Lesina and Varano lakes new guest in Lesina and Varano lakes: Gargano National Park (Italy). *Transit. Waters Bull.* **2008**, *2*, 69–79. [CrossRef]

55. Mancinelli, G.; Glamuzina, B.; Petrić, M.; Carrozzo, L.; Glamuzina, L.; Zotti, M.; Raho, D.; Vizzini, S. The trophic position of the Atlantic blue crab *Callinectes sapidus* Rathbun 1896 in the food web of Parila lagoon (South Eastern Adriatic, Croatia): A first assessment using stable isotopes. *Mediterr. Mar. Sci.* **2016**, *17*, 634–643. [CrossRef]
56. Cilenti, L.; Pazienza, G.; Scirocco, T.; Fabbrocini, A.; D'Adamo, R. First record of ovigerous *Callinectes sapidus* (Rathbun, 1896) in the Gargano lagoons (south-west Adriatic Sea). *BioInvasions Rec.* **2015**, *4*, 281–287. [CrossRef]
57. Manfrin, C.; Turolla, E.; Chung, J.S.; Giulianini, P.G. First occurrence of *Callinectes sapidus* (Rathbun, 1896) within the Sacca di Goro (Italy) and surroundings. *Check List* **2015**, *11*, 1640. [CrossRef]
58. Manfrin, C.; Comisso, G.; dall'Asta, A.; Bettoso, N.; Chung, J.S. The return of the Blue Crab, *Callinectes sapidus* Rathbun, 1896, after 70 years from its first appearance in the Gulf of Trieste, northern Adriatic Sea, Italy (Decapoda: Portunidae). *Check List* **2016**, *12*, 2006. [CrossRef]
59. Mordenti, O.; Scaravelli, D. Segnalazioni Faunistiche n° 83–87 (*Callinectes sapidus*). *Quad. Studi Notiz. Stor. Nat. Romagna* **2007**, *24*, 155–156.
60. Giansante, C. Segnalazioni faunistiche n. 119–124. *Quad. Studi Notiz. Stor. Nat. Romagna* **2012**, *36*, 207–208.
61. Castriota, L.; Andaloro, F.; Costantini, R.; De Ascentiis, A. First record of the Atlantic crab *Callinectes sapidus* Rathbun, 1896 (Crustacea: Brachyura: Portunidae) in Abruzzi waters, central Adriatic Sea. *Acta Adriat.* **2012**, *53*, 467–470.
62. Castejón, D.; Guerao, G. A new record of the American blue crab, *Callinectes sapidus* Rathbun, 1896 (Decapoda: Brachyura: Portunidae), from the Mediterranean coast of the Iberian Peninsula. *BioInvasions Rec.* **2013**, *2*, 141–143. [CrossRef]
63. El Zrelli, R.; Rabaoui, L.; Ragkousis, M.; Abdelali, N.; Azzurro, E.; Badreddine, A.; Bariche, M.; Bitar, G.; Crocetta, F.; Denitto, F.; et al. First record of the Atlantic blue crab *Callinectes sapidus* Rathbun, 1896 in the Gulf of Gabès (south-eastern Tunisia). New Alien Mediterranean Biodiversity Records (October 2020). *Medit. Mar. Sci.* **2020**, *21*, 631–652. [CrossRef]
64. Culurgioni, J.; Diciotti, R.; Satta, C.T.; Camedda, A.; de Lucia, G.A.; Pulina, S.; Lugliè, A.; Brundu, R.; Fois, N. First data on the appearance of *Callinectes sapidus* (Rathbun, 1896) in Sardinia (western Mediterranean). In Proceedings of the Atti del XXVIII Congresso Società Italiana di Ecologia, Cagliari, Italy, 12–14 September 2018; p. 112.
65. Piras, P.; Esposito, G.; Meloni, D. On the occurrence of the blue crab *Callinectes sapidus* (Rathbun, 1896) in Sardinian coastal habitats (Italy): A present threat or a future resource for the regional fishery sector? *BioInvasions Rec.* **2019**, *8*, 134–141. [CrossRef]
66. Culurgioni, J.; Diciotti, R.; Satta, C.T.; Camedda, A.; de Lucia, G.A.; Pulina, S.; Lugliè, A.; Brundu, R.; Fois, N. Distribution of the alien species *Callinectes sapidus* (Rathbun, 1896) in Sardinian waters (western Mediterranean). *BioInvasions Rec.* **2020**, *9*, 65–73. [CrossRef]
67. Öztürk, R.Ç.; Terzi, Y.; Feyzioğlu, A.M.; Şahin, A.; Aydın, M. Genetic characterization of the invasive Blue crab, *Callinectes sapidus* (Rathbun, 1896), in the Black Sea. *Reg. Stud. Mar. Sci.* **2020**, *39*, 101412. [CrossRef]
68. González-Ortegón, E.; Berger, S.; Encarnação, J.; Chairi, H.; Morais, P.; Teodosio, M.A.; Oliva-Paterna, F.J.; Schubart, C.D.; Cuesta, J.A. Free pass through the pillars of Hercules? Genetic and historical insights into the recent expansion of the Atlantic blue crab *Callinectes sapidus* to the West and the East of the Strait of Gibraltar. *Front. Mar. Sci.* **2022**, *9*, 918026. [CrossRef]
69. Schubart, C.D.; Deli, T.; Mancinelli, G.; Cilenti, L.; Gil Fernández, A.; Falco, S.; Berger, S. Phylogeography of the Atlantic blue crab *Callinectes sapidus* (Brachyura: Portunidae) in the Americas versus the Mediterranean Sea: Determining origins and genetic connectivity of a large-scale invasion. *Biology* **2022**, *12*, 35. [CrossRef]
70. Bilen, C.T.; Yesilyurt, I.N. Growth of blue crab, *Callinectes sapidus*, in the Yumurtalik Cove, Turkey: A molt process approach. *Cent. Eur. J. Biol.* **2014**, *9*, 49–57. [CrossRef]
71. Öndes, F.; Gökçe, G. Distribution and Fishery of the Blue Crab (*Callinectes sapidus* Rathbun, 1896) in Turkey Based on Local Ecological Knowledge of Fishers. *J. Anatol. Environ. Animal Sci.* **2021**, *6*, 325–332. [CrossRef]
72. Lattos, A.; Papadopoulos, D.K.; Giantsis, I.A.; Stamelos, A.; Karagiannis, D. Histopathology and Phylogeny of the Dinoflagellate *Hematodinium perezi* and the Epibiotic Peritrich Ciliate *Epistylis* sp. Infecting the Blue Crab *Callinectes sapidus* in the Eastern Mediterranean. *Microorganisms* **2024**, *12*, 456. [CrossRef]
73. Wolff, T. Occurrence of Two East American Species of Crabs in European Waters. *Nature* **1954**, *174*, 188–189. [CrossRef]
74. ICES WGITMO. Report of the Working Group on Introductions and Transfers of Marine Organisms. In Proceedings of the ICES CM 2006/ACME:05, Oostende, Belgium, 16–17 March 2006; p. 334.
75. Small, H.J.; Shields, J.D.; Hudson, K.L.; Reece, K.S. Molecular detection of *Hematodinium* sp. infecting the blue crab, *Callinectes sapidus*. *J. Shellfish Res.* **2007**, *26*, 131–139. [CrossRef]
76. Lohan, K.M.P. *Molecular Genetic Studies of Hematodinium sp., a Parasitic Dinoflagellate of the Blue Crab, Callinectes sapidus*; The College of William & Mary, Virginia Institute of Marine Science: Gloucester Point, VA, USA, 2011. [CrossRef]
77. Pagliara, P.; Mancinelli, G. Parasites affect hemocyte functionality in the hemolymph of the invasive Atlantic blue crab *Callinectes sapidus* from a coastal habitat of the Salento Peninsula (SE Italy). *Mediterr. Mar. Sci.* **2018**, *19*, 193–200. [CrossRef]
78. Small, H.J.; Huchin-Mian, J.P.; Reece, K.S.; Pagenkopp Lohan, K.M.; Butler, M.J., IV; Shields, J.D. Parasitic dinoflagellate Hematodinium perezi prevalence in larval and juvenile blue crabs *Callinectes sapidus* from coastal bays of Virginia. *Dis. Aquat. Org.* **2019**, *134*, 215–222. [CrossRef] [PubMed]
79. Flowers, E.M.; Johnson, A.F.; Aguilar, R.; Schott, E.J. Prevalence of the pathogenic crustacean virus *Callinectes sapidus* reovirus 1 near flow-through blue crab aquaculture in Chesapeake Bay, USA. *Dis. Aquat. Org.* **2018**, *129*, 135–144. [CrossRef]
80. Lively, J.A.; Spitznagel, M.I.; Schott, E.J.; Small, H.J. Investigating conspecific CsRV1 transmission in *Callinectes sapidus*. *J. Invertebr. Pathol.* **2023**, *201*, 107987. [CrossRef]

81. Zhao, M.; Plough, L.V.; Behringer, D.C.; Bojko, J.; Kough, A.S.; Alper, N.W.; Xu, L.; Schott, E.J. Cross-hemispheric genetic diversity and spatial genetic structure of *Callinectes sapidus* reovirus 1 (CsRV1). *Viruses* **2023**, *15*, 563. [CrossRef]
82. McMillen-Jackson, A.L.; Bert, T.M.; Steele, P. Population genetics of the blue crab *Callinectes sapidus*: Modest population structuring in a background of high gene flow. *Mar. Biol.* **1994**, *118*, 53–65. [CrossRef]
83. Place, A.R.; Feng, X.; Steven, C.R.; Fourcade, H.M.; Boore, J.L. Genetic markers in blue crabs (*Callinectes sapidus*): II. Complete mitochondrial genome sequence and characterization of genetic variation. *J. Exp. Mar. Biol. Ecol.* **2005**, *319*, 15–27. [CrossRef]
84. Yednock, B.K.; Neigel, J.E. An investigation of genetic population structure in blue crabs, *Callinectes sapidus*, using nuclear gene sequences. *Mar. Biol.* **2014**, *161*, 871–886. [CrossRef]
85. Rodrigues, M.A.; Dumont, L.F.C.; dos Santos, C.R.M.; D'Incao, F.; Weiss, S.; Froufe, E. Two distinct mtDNA lineages of the blue crab reveal large-scale population structure in its native Atlantic distribution. Estuar. *Coast. Shelf Sci.* **2017**, *197*, 45–53. [CrossRef]
86. Windsor, A.M.; Moore, M.K.; Warner, K.A.; Stadig, S.R.; Deeds, J.R. Evaluation of variation within the barcode region of Cytochrome c Oxidase I (COI) for the detection of commercial *Callinectes sapidus* Rathbun, 1896 (blue crab) products of non-US origin. *PeerJ* **2019**, *7*, e7827. [CrossRef]
87. Vecchioni, L.; Russotto, S.; Arculeo, M.; Marrone, F. On the occurrence of the invasive Atlantic blue crab *Callinectes sapidus* Rathbun 1896 (Decapoda: Brachyura: Portunidae) in Sicilian inland waters. *Nat. Hist. Sci.* **2022**, *9*, 43–46. [CrossRef]
88. Vella, A.; Giarrusso, E.; Monaco, C.; Misfud, C.M.; Darmanin, S.A.; Raffa, A.; Tumino, C.; Peri, I.; Vella, N. New Records of *Callinectes sapidus* (Crustacea, Portunidae) from Malta and the San Leonardo River Estuary in Sicily (Central Mediterranean). *Diversity* **2023**, *15*, 679. [CrossRef]
89. Keskin, E.; Atar, H.H. DNA barcoding commercially important aquatic invertebrates of Turkey. *Mitochondrial DNA* **2013**, *24*, 440–450. [CrossRef]
90. Gonzalez-Wanguemert, M.; Pujol, J.A. First record of the atlantic blue crab *Callinectes sapidus* (Crustacea: Brachyura: Portunidae) in the Segura river mouth (Spain, southwestern Mediterranean Sea). *Turk. J. Zool.* **2016**, *40*, 615–619. [CrossRef]
91. 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.
92. 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]
93. Okonechnikov, K.; Golosova, O.; Fursov, M.; Ugene Team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [CrossRef] [PubMed]
94. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
95. 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]
96. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef] [PubMed]
97. 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]
98. 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]
99. 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]
100. Strimmer, K.; von Haeseler, A. Quartet Puzzling: A quartet Maximum-Likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **1996**, *13*, 964–969. [CrossRef]
101. 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] [PubMed]
102. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
103. 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]
104. Puillandre, N.; Brouillet, S.; Achaz, G. ASAP: Assemble species by automatic partitioning. *Mol. Ecol. Resour.* **2020**, *21*, 609–620. [CrossRef]
105. 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]
106. 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]
107. 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] [PubMed]

108. Scarpa, F.; Cossu, P.; Delogu, V.; Lai, T.; Sanna, D.; Leasi, F.; Norenuburg, 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]
109. Zhang, J.; Kapli, P.; Pavlidis, P.; Stamatakis, A. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **2013**, *29*, 2869–2876. [CrossRef] [PubMed]
110. Drummond, A.J.; Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **2007**, *7*, e214. [CrossRef]
111. Schubart, C.D.; Diesel, R.; Hedges, S.B. Rapid evolution to terrestrial life in Jamaican crabs. *Nature* **1998**, *393*, 363–365. [CrossRef]
112. Daniels, S.R.; Phiri, E.E.; Klaus, S.; Albrecht, C.; Cumberlidge, N. Multilocus phylogeny of the Afrotropical freshwater crab fauna reveals historical drainage connectivity and transoceanic dispersal since the Eocene. *Syst. Biol.* **2015**, *64*, 549–567. [CrossRef]
113. Rambaut, A.; Suchard, M.A.; Xie, D. Drummond. Tracer v1.6. A.J. 2014. Available online: http://tree.bio.ed.ac.uk/software/tracer/ (accessed on 15 June 2024).
114. Strimmer, K.; von Haeseler, A. Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6815–6819. [CrossRef]
115. Head, M.J.; Gibbard, P.L. Early–Middle Pleistocene transitions: Linking terrestrial and marine realms. *Quat. Int.* **2015**, *389*, 7–46. [CrossRef]
116. Oliveira, D.; Desprat, S.; Yin, Q.; Rodrigues, T.; Naughton, F.; Trigo, R.; Su, Q.; Grimalt, J.O.; Alonso-Garcia, M.; Voelker, A.H.L.; et al. Enhanced humidity in SW Iberia driven by the combination of insolation and ice-sheet forcing during MIS 13 interglacial. In Proceedings of the EGU General Assembly Conference Abstracts, EGU General Assembly 2021, Online, 19–30 April 2021; p. EGU21-11408. [CrossRef]
117. Ludt, W.B.; Rocha, L.A. Shifting seas: The impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. *J. Biogeogr.* **2015**, *42*, 25–38. [CrossRef]
118. Lacerda, A.L.F.; Kersanach, R.; Cortinhas, M.C.S.; Prata, P.F.S.; Dumont, L.F.C.; Proietti, M.C.; Maggioni, R.; D'Incao, F. High connectivity among blue crab (*Callinectes sapidus*) populations in the Western South Atlantic. *PLoS ONE* **2016**, *11*, e0153124. [CrossRef] [PubMed]
119. Olival, K.J.; Latinne, A.; Islam, A.; Epstein, J.H.; Hersch, R.; Engstrand, R.C.; Gurley, E.S.; Amato, G.; Luby, S.P.; Daszak, P. Population genetics of fruit bat reservoir informs the dynamics, distribution and diversity of Nipah virus. *Mol. Ecol.* **2020**, *29*, 970–985. [CrossRef] [PubMed]
120. Akdaş, O.; Ayaz, İ.S.; Çaylan, D.Ö. Analysing the trade between Turkey and the Aegean Islands: The hidden opportunities. *Dokuz Eylül Üniv. Denizcilik Fak. Derg.* **2018**, *10*, 109–126. [CrossRef]
121. Cowen, R.K.; Sponaugle, S. Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.* **2009**, *1*, 443–466. [CrossRef] [PubMed]
122. 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]
123. Willis, J.H.; Orr, H.A. Increased heritable variation following population bottlenecks: The role of dominance. *Evolution* **1993**, *47*, 949–957. [CrossRef]
124. 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]
125. 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]

Chapter 3

Salmo letnica

Salmo letnica Karaman 1924 (Chordata: Actinopterygii: Salmonidae), also known as the Ohrid trout, is an endemic salmonid species of Lake Ohrid, one of the oldest lakes in Europe, located between Albania and North Macedonia (Karaman, 1924; Stanković, 1960; Figure 3). Within the lake, four distinct morphotypes have been described so far (*S. l. typicus*, *S. l. aestivalis*, *S. l. balcanicus*, and *S. l. lumi*), which differ in morphological traits and reproductive behaviours (Rakaj and Flloko, 1995).

Figure 3. (a) Lake Ohrid. Photo credits: D. Kelly, 2014 (https://www.lakescientist.com/across-the-pond-lake-ohrid/). (b) *Salmo letnica* specimen in Lake Ohrid. Photo credits: S. Zogaris (https://www.marinelifeid.com/identification/ohrid-trout-salmo-letnica/).

Lake Ohrid hosts high levels of endemicity and likely functioned as a glacial refugium in the past (Albrecht & Wilke, 2008). In this context, *Salmo letnica* coexists in Lake Ohrid with another endemic salmonid species, *Salmo ohridanus* (Sušnik et al., 2006; Albrecht & Wilke, 2008). In addition, historical records indicate that *Salmo farioides* was also observed within the lake, despite currently restricted to its smaller tributaries (Drin River) (Kottelat, 1997).

Over the decades, several studies have attempted to characterise *Salmo letnica* from both morphological (e.g., Karaman, 1924; Stefanoviç, 1948; Rakaj & Flloko, 1995) and genetic perspectives (e.g., Sell & Spirkovski, 2004; Sušnik et al., 2007; Snoj et al., 2009). Nevertheless, the taxonomic status of the species, along with its evolutionary origin, remain subject of ongoing uncertainties.

Beyond its evolutionary interest, the species plays an important ecological role in Lake Ohrid ecosystem and represents an important local food resource (Rakaj & Flloko, 1995). However, population declines have been reported in recent decades as a consequence of environmental and anthropogenic pressures, highlighting the need for targeted conservation actions (Prifti & Cake, 2017; Spirkovski et al., 2019).

On this basis, the research presented in this thesis chapter aimed to investigate the taxonomic status and adaptive processes of *Salmo letnica* within Lake Ohrid, using the mitochondrial Control Region (D-loop) as molecular marker. Precisely, the objectives were threefold: (i) to assess the genetic variability and potential structuring among the putative species morphotypes; (ii) to investigate the genetic relationships between Ohrid trout specimens and other *Salmo* species occurring in the lake; and (iii) to shed light on the evolutionary history and origin of *S. letnica* within Lake Ohrid.

References

Albrecht, C., & Wilke, T. (2008). Ancient Lake Ohrid: Biodiversity and evolution. *Hydrobiologia, 615*, 103–140. https://doi.org/10.1007/s10750-008-9558-y

Karaman, S. (1924). *Pisces Macedoniae* (pp. 1–90). Hrvatska Štamparija.

Kelly, D. (2014, February 13). *Across the pond: Lake Ohrid is Europe's oldest*. Lake Scientist. https://www.lakescientist.com/across-the-pond-lake-ohrid/

Kottelat, M. (1997). *European freshwater fishes: An heuristic checklist of the freshwater fishes of Europe (exclusive of former USSR), with an introduction for non-systematists and comments on nomenclature and conservation*. *Biologia, 52*, 1–271.

Prifti, V., & Cake, A. (2017). The impact of environmental conditions on growth and development of *Salmo letnica* smolt. *Scientific Papers: Animal Science and Biotechnologies, 50*, 131.

Rakaj, N., & Flloko, A. (1995). Conservation status of freshwater fish of Albania. *Biological Conservation, 72*, 195–199. https://doi.org/10.1016/0006-3207(94)00082-2

Sell, J., & Spirkovski, Z. (2004). Mitochondrial DNA differentiation between two forms of trout *Salmo letnica*, endemic to Balkan Lake Ohrid, reflects their reproductive isolation. *Molecular Ecology, 13*, 3633–3644. https://doi.org/10.1111/j.1365-294X.2004.02362.x

Snoj, A., Marić, S., Berrebi, P., Crivelli, A. J., Shumka, S., & Sušnik, S. (2009). Genetic architecture of trout from Albania as revealed by mtDNA control region variation. *Genetics Selection Evolution, 41*, 22. https://doi.org/10.1186/1297-9686-41-22

Spirkovski, Z., Ilik-Boeva, D., Ritterbusch, D., Peveling, R., & Pietrock, M. (2019). Ghost net removal in ancient Lake Ohrid: A pilot study. *Fisheries Research, 211*, 46–50. https://doi.org/10.1016/j.fishres.2018.10.023

Stanković, S. (1960). The Balkan Lake Ohrid and its living world. In *Monographiae Biologicae* (Vol. 9). Uitgeverij Dr. W. Junk.

Stefanović, D. (1948). *Recherches systématiques et écologiques sur les Salmonidés d'Ohrid*. Académie Serbe des Sciences (Éditions spéciales), 139, 1–207.

Sušnik, S., Knizhin, I., Snoj, A., & Weiss, S. (2006). Genetic and morphological characterization of a Lake Ohrid endemic, *Salmo (Acantholingua) ohridanus*, with a comparison to sympatric *Salmo trutta*. *Journal of Fish Biology, 68*, 2–23. https://doi.org/10.1111/j.0022-1112.2006.00902.x

Sušnik, S., Snoj, A., Wilson, I. F., Mrdak, D., & Weiss, S. (2007). Historical demography of brown trout (*Salmo trutta*) in the Adriatic drainage including the putative *S. letnica* endemic to Lake Ohrid. *Molecular Phylogenetics and Evolution, 44*, 63–76. https://doi.org/10.1016/j.ympev.2006.08.021

Zogaris, S. (n.d.). *Ohrid trout (Salmo letnica)*. Marine Life ID. https://www.marinelifeid.com/identification/ohrid-trout-salmo-letnica/

life

MDPI

Article

Exploring Mitochondrial Evolutionary Pathways: Insights into the Origin of the Endemic Ohrid Trout

Anila Hoda [1,†], Chiara Locci [2,3,†], Ilenia Azzena [3], Noemi Pascale [2,3,4], Ilaria Deplano [2], Roland Kristo [5], Arian Demiri [6], Fabio Scarpa [2], Marco Casu [3,‡] and Daria Sanna [2,*,‡]

[1] Department of Animal Sciences, Agricultural University of Tirana, Kodër Kamëz, 1029 Tirana, Albania; ahoda@ubt.edu.al
[2] Department of Biomedical Sciences, University of Sassari, Viale San Pietro 43b, 07100 Sassari, Italy; c.locci3@phd.uniss.it (C.L.); npascale@uniss.it (N.P.); i.deplano@phd.uniss.it (I.D.); fscarpa@uniss.it (F.S.)
[3] Department of Veterinary Medicine, University of Sassari, Via Vienna 2, 07100 Sassari, Italy; iazzena@uniss.it (I.A.); marcasu@uniss.it (M.C.)
[4] Department of Chemical, Physical, Mathematical, and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy
[5] Ministry of Agriculture and Rural Development, Blvd Deshmoret e Kombit, Nr. 2, 1001 Tirana, Albania; roland.kristo@bujqesia.gov.al
[6] Directory of Fishery and Aquaculture Services, Blvd Deshmoret e Kombit, Nr. 2, 1001 Tirana, Albania; arjan.demiri@dshpa.gov.al
[*] Correspondence: darsanna@uniss.it
[†] These authors contributed equally to this work as co-first author.
[‡] These authors contributed equally to this work as co-senior author.

Abstract: The Ohrid trout, *Salmo letnica*, is an endemic species of Lake Ohrid, one of Europe's oldest lakes, located on the Albania-North Macedonia border. This species exhibits distinct morphotypes—*Salmo letnica typicus*, *Salmo letnica aestivalis*, *Salmo letnica balcanicus*, and *Salmo letnica lumi*—that differ in morphology and spawning behaviour. However, the extent of their genetic differentiation remains unclear. This study aimed to investigate the genetic variability and population structure of *Salmo letnica* morphotypes using the mitochondrial Control Region as molecular marker. We obtained 127 sequences from *Salmo letnica* morphotypes and compared them with sequences from other species within the genus *Salmo*. Phylogenetic and clustering analyses revealed no significant genetic structuring among the four morphotypes, suggesting an ecological differentiation not (yet) fixed at mitochondrial level. Additionally, our findings suggest that the modern *Salmo letnica* population likely originated in Lake Ohrid from *Salmo farioides* founders through evolutionary differentiation, potentially driven by environmental changes. Future studies incorporating a larger number of samples from both *Salmo letnica* and *Salmo farioides* are essential to fully understand the evolutionary and ecological dynamics of *Salmo letnica* morphotypes.

Keywords: Lake Ohrid; Korani trout; mitochondrial DNA; Control region; *Salmo l. typicus*; *Salmo l. aestivalis*; *Salmo l. lumi*; *Salmo l. balcanicus*

check for updates

Academic Editor: Koichiro Tamura

Received: 12 November 2024
Revised: 11 December 2024
Accepted: 25 December 2024
Published: 3 January 2025

Citation: Hoda, A.; Locci, C.; Azzena, I.; Pascale, N.; Deplano, I.; Kristo, R.; Demiri, A.; Scarpa, F.; Casu, M.; Sanna, D. Exploring Mitochondrial Evolutionary Pathways: Insights into the Origin of the Endemic Ohrid Trout. *Life* **2025**, *15*, 52. https://doi.org/10.3390/life15010052

1. Introduction

Lake Ohrid, nestled among the Dassaretes Lakes in the Balkans (central Europe), holds a distinguished status as one of Europe's oldest bodies of water, with estimates of its age ranging between two and four million years old [1]. The Dassaretes Lakes consist of a group of lakes that includes Lake Ohrid (Albania, North Macedonia), Lake Prespa (Albania, Greece, North Macedonia), Lake Mikri Prespa (Albania, Greece,), and Lake Maliq

(Albania) [2]. Lake Ohrid, positioned at an altitude of 693 m [3], is the most renowned among these lakes, with a watershed extending over 1000 km² [4]. Moreover, it ultimately drains into the Adriatic Sea via the Drim River [4]. Remarkably, Lake Ohrid's maximum water depth of 288 m and an average depth of 155 m [5] have shielded it from glaciation, setting it apart as a geological marvel [1]. The lake benefits from a steady influx of water from karstic surface and sub lacustrine springs, ensuring a relatively stable water level and fostering a habitat conducive to a high degree of endemism [1].

Predominantly fed by spring water, including contributions from neighboring Lake Prespa [6], Lake Ohrid maintains oligotrophic conditions characterized by high water clarity [3,5]. Since the early 20th century, the lake has been the subject of scientific scrutiny, with studies particularly focused on its endemic and relict species (e.g., [1–3,6–8]). The lake's distinctive characteristics render it an invaluable natural laboratory for comprehending long-term environmental changes in the central northern Mediterranean region [3,9].

The enduring lineage of Lake Ohrid's flora and fauna, preserved over millions of years, is a testament to its unique characteristics [1]. Approximately 40% of its fish species are endemic [2], underscoring its significance as a biodiversity hotspot. Among these are several endemic trout species, including the Ohrid trout, *Salmo letnica* Karaman, 1924, locally known as "Korani trout" [10], which stands out as a key subject for further investigation due to its distinct evolutionary adaptation (i.a., [11]). Additionally, another endemic trout species, *Salmo ohridanus* (Steindachner, 1892) [12,13], previously classified as *Acantholingua ohridana* (Hadžišče, 1961), has been identified in this lake [2].

Notably, the taxonomic classification and evolutionary history of the endemic trout species *Salmo letnica* remain subjects of ongoing debate. Initially described by Karaman (1924) [6,14], subsequent research by Stefanoviç (1948) delineated three distinct populations within the *letnica* species: *Salmo l. typicus*, *Salmo l. aestivalis*, and *Salmo l. balcanicus* [7]. These populations exhibit significant differences in various morphological traits and behaviors, including spawning preferences [7]. *Salmo l. typicus* population spawns in winter, specifically in January and February, in littoral and sublittoral areas, whereas *Salmo l. aestivalis* population prefers deeper rocky habitats during the summer months. In contrast, *Salmo l. balcanicus* population favors spawning near the lake outlet between December and January [7]. Additionally, two other trout populations have been identified in the Lake Ohrid area: *Salmo l. lumi* Poljakov, Filipi & Basho, 1958 and *Salmo farioides* Karaman, 1938 [15]. *Salmo l. lumi*, a river spawner, was found along the northern and western shores of the lake. Meanwhile, *Salmo farioides* has been detected in the smaller tributaries along the eastern lakeshore [16,17].

In more recent times, Rakaj and Flloko (1995) provided a detailed morphological description of the *Salmo letnica* (species) group, highlighting distinct intraspecies characteristics, especially between males and females [10]. Additionally, these authors detected a more pronounced differentiation of *Salmo l. lumi* in respect to the other three morphotypes of Korani trout, being distinguished by its longer, more robust body with distinct black and red spots [10].

Salmo letnica plays a significant role in Lake Ohrid's ecosystem, constituting a substantial portion of salmonid catches, with efforts in artificial reproduction dating back to the 1950s [10]. However, over the past few decades, environmental changes and human pressures have significantly affected its populations. Currently, only two forms, *Salmo l. typicus* and *Salmo l. aestivalis*, are definitively observed. However, *Salmo l. lumi* and *Salmo l. balcanicus* still exist but are possibly dwindling in population [17]. This decline evidences the urgent need for targeted conservation efforts.

In response to these challenges, Kostoski et al. (2010) emphasized the importance of immediate conservation efforts to protect freshwater resources globally, particularly ancient lakes like Lake Ohrid [9]. This lake is increasingly threatened by anthropogenic pressures, non-native species, and climate change. To mitigate these threats, effective conservation strategies should primarily focus on watershed management and establishing conservation and coastal zone management areas to safeguard the lake's unique ecosystem [9].

It is important to note that this area requires sustainable conservation and management strategies, as highlighted by Smederevac-Lalić et al. (2023) [18]. Indeed, in the Western Balkans, balancing agricultural modernization with environmental sustainability remains challenging, as economic priorities frequently overshadow ecological concerns. The fisheries sector, essential for both food security and income, suffers from weak organization and insufficient enforcement, leading to unsustainable practices [18].

In this context, the Lake Ohrid Conservation Project (LOCP), launched in the late 1990s, represents a collaborative effort between Albania and Macedonia, aimed at jointly managing the lake's watershed to preserve its biodiversity [19].

Recent studies have further highlighted the decline in the trout population in Lake Ohrid, reinforcing the need for coordinated management efforts to ensure the fishery's sustainability [20,21]. Additionally, understanding the environmental factors influencing the growth and development of *Salmo letnica* is essential for its long-term conservation [22].

From a conservation perspective, studies specifically investigating the genetic structure and evolutionary history of Ohrid trout are limited. Most research has focused primarily on mitochondrial DNA analyses [13,17,23–28]. These studies have consistently placed *Salmo letnica* within the Adriatic phylogeographical lineage of brown trout (*Salmo trutta*) [13,17,27]. Efforts have also been made to explore the historical demography of Ohrid trout, with a particular focus on the genetic differentiation between its winter (*Salmo l. typicus*) and summer (*Salmo l. aestivalis*) spawning forms. This distinction is crucial due to the unique genetic makeup and ecological significance of *Salmo letnica* [23,25]. For instance, Sušnik et al. (2007) conducted a study using both mitochondrial DNA and microsatellites as molecular markers to investigate the genetic structuring between these two forms [26]. Their results did not reveal significant genetic differentiation within the species [26]. These results, however, contrasted with the findings of Sell and Spirkovski (2004), who reported a high degree of genetic differentiation between the winter and summer spawning forms of Ohrid trout [17]. The discrepancy was later explained by Sušnik et al. (2006, 2007), who attributed the variation observed in that earlier study to instability at the 3′ end of the mitochondrial Control Region. This instability, caused by intramolecular mechanisms, was deemed unreliable for phylogenetic investigations [13,26].

The limited availability of genetic data and the possible ambiguity in previous findings underscore the need for more detailed genetic research to clarify the extent of differentiation within the species and to better understand its genetic variability and evolutionary origin.

In light of these concerns, this study aimed to investigate the genetic variability and potential genetic structuring of the four distinct morphotypes of the Ohrid trout (*Salmo letnica*) population through mitochondrial genotyping. To achieve this, a portion of the mitochondrial Control Region was employed as molecular marker. This non-coding region was chosen due to its general high variability and neutrality to selective pressures, making it effective in detecting genetic divergence due to distinct phylogenetic origins. This region is recognized for its utility in assessing genetic structure within populations and identifying interspecific differences [29,30]. Additionally, the availability of numerous sequences from this mitochondrial region in the GenBank database, associated with selected *Salmo* species, enabled an examination of the genetic relationships and a comparison between the Lake Ohrid trout obtained in the present study and other trout species inhabiting or observed in the lake. This analysis sought to shed light on the evolutionary history of *Salmo letnica* and provide actionable data essential for effective conservation strategies aimed at preserving the biodiversity of Lake Ohrid.

2. Materials and Methods

2.1. Sample Collection

Ohrid Trout (*Salmo letnica*) is a commercial species fished in both Albania and North Macedonia. In Albania, fishing is conducted by the Fishermen's Association through a co-management process, while in North Macedonia, it is managed via concession agreements. Listed as "Endangered" on the IUCN (International Union for Conservation of Nature) Red List of Threatened Species [31], Ohrid trout is subject to strict fishing regulations: minimum landing sizes (32 cm in Albania and 35 cm in North Macedonia) and designated closed fishing seasons (December 1 to February 28 in Albania, and December 1 to March 20 in North Macedonia) (Albanian legislation: https://qbz.gov.al/eli/urdher/2022/03/31/149/cc5c689b-5245-417a-978d-52346bbb2543, accessed on 1 August 2022; North Macedonian legislation: https://mzsv.gov.mk, accessed on 1 August 2022).

In the present study, samples of Ohrid trout were collected from multiple locations within the Albanian side of Lake Ohrid between September 2022 and July 2023, excluding the closed fishing season previously mentioned (Figure 1 and Supplementary Table S1), representing four distinct morphological forms: *typicus*, *balcanicus*, *lumi* and *aestivalis*. A total of 127 individuals were included in the analyses, comprising 35 specimens of *Salmo l. typicus*, 36 specimens of *Salmo l. balcanicus*, 22 specimens of *Salmo l. lumi*, and 34 specimens of *Salmo l. aestivalis*.

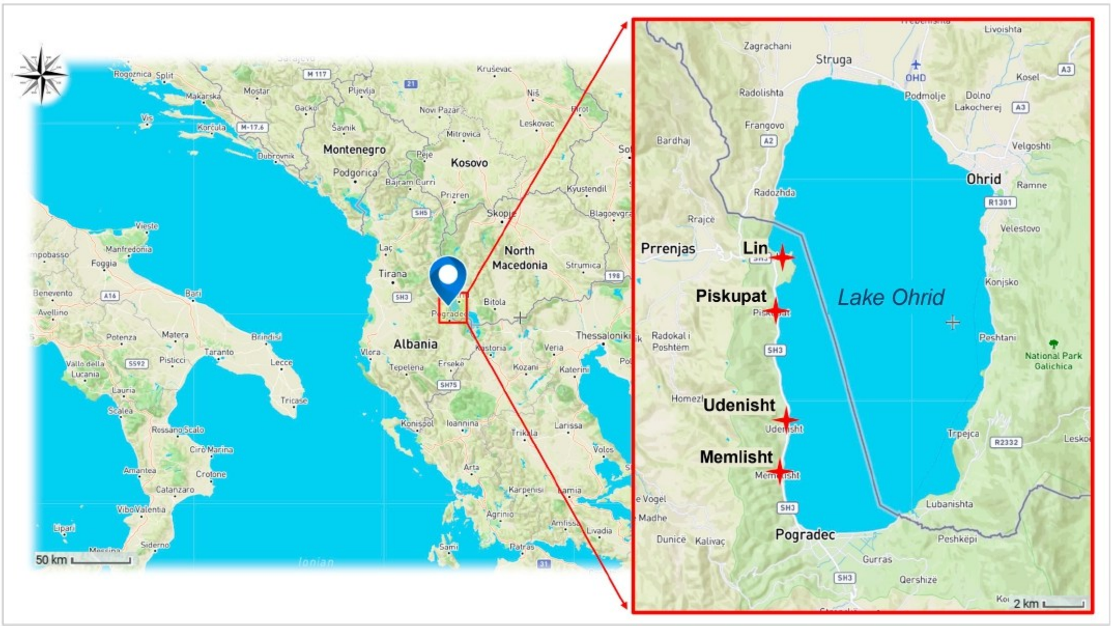


Figure 1. Sampling locations of Ohrid brown trout (*Salmo letnica*). The map highlights the four sampling sites (marked in red) where the Control Region sequences analyzed in the current study were obtained.

The sampling process was particularly time-intensive due to the rarity of certain forms, notably *Salmo l. lumi*. While *Salmo l. lumi* is not extinct, it is exceedingly rare, which made it challenging to reach the desirable target of 30 specimens for this form. During some expeditions, only 1–2 specimens were found, while on other occasions, none were located. This rarity significantly extended the sampling period, requiring nearly a year to complete.

Despite these challenges, we aimed to achieve the required sample size of 30 individuals per morphological form of *Salmo letnica*, though this was difficult for the rarer forms.

All samples were obtained exclusively from deceased animals destined for the fish market upon the return of fishermen from fishing expeditions. Local members of the Fishermen's Association facilitated access to licensed fishermen and boats specializing in Ohrid trout fishing, both for family consumption and commercial purposes. These professionals, who have studied *Aquaculture and Fishery*, are also involved in operations at the Koran Restocking Centre in Lin.

Their logistical support was essential. Indeed, at designated shore centers (Lin, Piskupat, Udenisht, and Memlisht), they assisted in distinguishing the four forms based on their morphological characteristics and supported the separation of samples for subsequent molecular analyses. Additionally, specimens were collected by taking fin clips from each individual.

The fin clips were preserved in absolute ethanol and subsequently utilized for DNA extraction.

2.2. Diagnostic Molecular Analyses

Total genomic DNA was extracted from a portion of fin tissue using the Mackerey-Nagel NucleoSpin Tissue Kit (MACHEREY-NAGEL GmbH and Co. KG, Duren, Germany), following the manufacturer's protocol. DNA quantification was carried out with the Nanodrop™ Lite Spectrophotometer (Thermo Scientific; Waltham, MA, USA), yielding an average concentration of 80 ng/μL. A portion of the mitochondrial Control Region was amplified by standard PCR using the following primers: L19CR (forward) (5′-CCACTAGCTCCCAAAGCTA-3′) [23] and K-Rev CR (reverse) ((5′-CAGGACCAAGCTTTT-GTGCTTACG-3′) [32]. The PCR reactions were prepared in a total volume of 25 μL, containing approximately 10 ng of genomic DNA, 0.6 μM of each primer, and a PuReTaq Ready-To-Go PCR bead (GE Healthcare, Wauwatosa, WI, USA), which included stabilizers, bovine serum albumin (BSA), deoxynucleotide triphosphates, 2.5 units PuReTaq DNA polymerase, and buffer. Upon reconstitution to a final volume of 25 μL, the dNTP and MgCl2 concentrations were set to 200 μM and 1.5 mM, respectively. PCR cycles were performed in a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems, Waltham, MA, USA), setting the thermal cycling conditions as follows: an initial denaturation at 94 °C for 4 min, followed by 35 cycles of 30 s at 94 °C, 30 s at 51 °C, and 30 s at 72 °C, with a final extension step of 10 min at 72 °C and cooling at 4 °C. To ensure the reliability of the PCR and to verify the absence of contamination, positive controls (high-quality DNA) and negative controls were also included in the analysis. The PCR products were analyzed through electrophoresis on 2% agarose gels prepared in 1x TAE buffer (Tris-Acetate-EDTA, pH 8.3) and stained with Gel Red Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). Purification of the PCR products was performed using ExoSAP-IT (USB Corporation, Cleveland, OH, USA). Both forward and reverse strands were sequenced using the same primers as in the PCR, with sequencing performed by an external core service (Macrogen Europe, Milan, Italy).

2.3. Phylogeographic and Phylogenetic Analysis

A total of 127 Control Region sequences (553 bp) from Ohrid trout individuals were obtained in the present study (Supplementary Table S1) and have been deposited in the GenBank online database (Accession numbers PQ583195–PQ583321).

All newly obtained sequences were verified through BLAST analysis on the GenBank nucleotide database (www.ncbi.nlm.nih.gov, accessed on 6 June 2024) and confirmed to belong to the species *Salmo letnica*, with a 100% identity match.

Sequences were aligned using the Clustal Omega algorithm, implemented in the MPI Bioinformatics Toolkit (available at https://toolkit.tuebingen.mpg.de/tools/hhpred, accessed on 6 June 2024). The alignment was manually checked and verified for the occurrence of polymorphisms with Unipro UGENE v.35 [33].

The new sequences obtained in the present study were merged with all available sequences from the GenBank database for *Salmo* species (*Salmo letnica*, *Salmo l. lumi*, *Salmo ohridanus*) inhabiting Lake Ohrid. This integration was based on prior studies documenting these species in the area [13,26,28,34,35]. Additionally, sequences of *Salmo farioides* from the White Drin River in Albania [36] were included (see Supplementary Table S2 for details). Hence, this comprehensive dataset, encompassing both newly generated and publicly available sequences, consisted of a total of 164 sequences.

To enhance our understanding of the relationship between our samples and the major brown trout (*Salmo trutta*) phylogeographical lineages (AT—Atlantic, AD—Adriatic, ME—Mediterranean, MA—Marmoratus, DA—Danubian) [37–41], five representative sequences were incorporated into the whole dataset (see Supplementary Table S2 for details). Furthermore, a sequence belonging to *Salmo salar* species was used as outgroup.

The genetic diversity among sequences was evaluated by calculating several key parameters, including the number of polymorphic sites (S), number of haplotypes (H), nucleotide diversity (π), and haplotype diversity (h). This analysis was carried out using the DnaSP 6.12.03 software package (developed by Universitat de Barcelona, Barcelona, Spain) [42].

The best probabilistic model for sequence evolution was identified using jModeltest 2.1.1 [43] via a maximum likelihood optimized search, evaluated by the Akaike (AIC) and Bayesian Information Criterion (BIC). Both criteria identified the GTR + I + G model as the best fit for the whole dataset.

Evolutionary relationships among haplotypes and taxonomic entities were examined by using MrBayes 3.2.7 [44]. The parameters for the analysis were set as follows: NST = 6, rates = invgamma, ngammacat = 4. The procedure included two independent runs of four Metropolis-coupled Markov-chain Monte Carlo (MCMCMC) simulations (one cold and three heated chains) for 5,000,000 generations, sampling trees every 1000 generations. The first 25% of the 10,000 sampled trees were discarded as burn-in. Nodes with a posterior probability above 0.9 were considered statistically well supported. Convergence of the chains was tested by verifying that the Average Standard Deviation of Split Frequencies (ASDSF) approached zero, following Ronquist et al. [44], and that the Potential Scale Reduction Factor (PSRF) was approximately equal to 1, as suggested by Gelman and Rubin [45]. The resulting phylogenetic tree was edited and visualized using FigTree 1.4.0 (available online: http://tree.bio.ed.ac.uk/software/figtree/, accessed on 15 July 2024).

To identify potential subgroups within genetic clusters and assess genetic variability among sequences, we conducted a principal coordinate analysis (PCoA) using GenAlEx 6.5 [46]. This PCoA reconstruction was based on a pairwise p-distance matrix of genetic data, aiming to examine the dissimilarity represented by the genetic variability among the analyzed sequences. This analysis was performed on both the entire dataset and a subset, based on the results obtained from the initial run.

A median-joining network [47] was constructed using the Network 10.2.0.0 software package (available at www.fluxus-engineering.com, accessed on 6 June 2024) from Colchester, UK. This network was generated to analyze the genetic relationships among haplotypes and to identify potential discrete genetic clusters. In this analysis, both transitions and transversions were weighted equally. Due to uncertainties surrounding the occurrence of retromutation events, each observed polymorphism was assigned an equal weight of 10.

3. Results

A total of 6 polymorphic sites were detected among the 127 analyzed sequences, resulting in 6 distinct Control Region haplotypes (see Table 1). Moreover, reduced and comparable genetic diversity estimates were observed across the four forms, except for *Salmo letnica lumi*, which showed relatively lower nucleotide divergence compared to the other morphotypes, specifically with a nucleotide diversity an order of magnitude lower.

Table 1. Sample sizes and genetic diversity estimates for the mitochondrial Control Region sequences obtained in this study, representing *Salmo letnica* individuals classified into four morphotypes. N: sample size; S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; π: nucleotide diversity. Sites with gaps were not considered.

Sample	N	S	H	h	π
Salmo l. typicus	35	3	4	0.264	0.00050
Salmo l. balcanicus	36	5	4	0.348	0.00150
Salmo l. lumi	22	1	2	0.247	0.00045
Salmo l. aestivalis	34	4	4	0.401	0.00134
Whole dataset *	**127**	**6**	**6**	**0.319**	**0.00101**

*** Total values for sample size and genetic diversity for the whole dataset, which includes all the *Salmo letnica* sequences obtained in the present study.**

The Bayesian evolutionary tree (Figure 2), derived from a dataset comprising both newly sequenced samples and all relevant sequences of specimens belonging to the genus *Salmo*, displayed well-supported branches with a posterior probability of 1 at the major node. Overall, the tree revealed a statistically significant lack of genetic structure based on taxonomic entities. With midpoint rooting, it correctly placed *Salmo salar* as the outgroup, positioned externally to the main clade. The main clade was further structured into several sub-clades, forming two primary groups. The first group consisted of *Salmo ohridanus*, which stood out as an exception to the general lack of taxonomic structuring; this species clustered outside the most inclusive clade and was characterized by 11 haplotypes. The second group, moving inward through the tree, represented the most inclusive clade. This large polytomy encompassed *Salmo letnica*, *Salmo l. typicus*, *Salmo l. balcanicus*, *Salmo l. lumi*, *Salmo l. aestivalis*, and *Salmo farioides*. This clade constituted an heterogenous non-monophyletic group, that, based on the results, was at least polyphyletic. Notably, the phylogenetic tree did not reveal a division consistent with the taxonomic framework to create discrete groups corresponding to the number of species included. All the sub-clades within this clade comprised multiple *Salmo* species. Interestingly, a fully supported sub-clade included *Salmo letnica* (AY926572), *Salmo l. balcanicus* (b_5, b_13, b_36), and *Salmo l. aestivalis* (SA_4, SA_32).

The absence of genetic structuring based on taxonomic criteria was also confirmed by the Principal Coordinates Analysis (PCoA, Supplementary Figure S1a), which accounted for 93.71% of the genetic variability in the dataset. The graph showed that *Salmo ohridanus* was separated from the other terminals, appearing as the most heterogenous and variable among those included in the analyses. Moreover, the PCoA revealed the occurrence of two main genetic groups: Group 1, comprising *Salmo farioides* (MW251458, MT505423), *Salmo letnica* (AY926572), *Salmo l. balcanicus* (b_5, b_13, b_36), and *Salmo l. aestivalis* (SA_4, SA_32), and Group 2, a more inclusive group, containing the remaining terminals belonging to the species *Salmo letnica*, *Salmo farioides*, *Salmo l. typicus*, *Salmo l. balcanicus*, *Salmo l. lumi*, and *Salmo l. aestivalis*. Notably, Group 1 was partially consistent with the fully supported sub-clade observed in the phylogenetic tree. However, while the sub-clade did not include *Salmo farioides*, the PCoA grouped a few *Salmo farioides* sequences with *Salmo letnica*, *Salmo l. balcanicus*, and *Salmo l. aestivalis*. This suggests a slight discrepancy between the results of the phylogenetic analysis and the PCoA, highlighting differences in how these analyses represent the genetic relationships of *Salmo farioides*, along with *Salmo letnica* species. This difference could be the consequence of methodological variations between the two approaches (phylogenetic tree vs. PCoA), which base their results on evolutionary divergences, along with genetic differences, versus general similarities, respectively.

In particular, the phylogenetic tree is based on evolutionary models and evidences divergences driven by specific mutations accumulated over time. This branching structure reflects the evolutionary history of the species and highlights ancestral relationships. In contrast, PCoA is based on a dissimilarity matrix in which samples are distributed across a multidimensional space according to genetic similarities among sequences, without considering ancestral relationships. For this reason, in PCoA, genetically similar samples may appear close in space even if they do not share a recent common ancestor and are not phylogenetically related.

Additionally, phylogenetic trees are sensitive to mutations that are significant for lineage distinctions, especially recent ones, whereas PCoA considers only overall genetic distance, ignoring the temporal sequence of mutations.

Group 2 was further investigated with an additional PCoA (Supplementary Figure S1b), which explains 75.65% of the genetic variability within this subset. The graph revealed several outliers: a few *Salmo farioides*, one *Salmo letnica typicus* (T_35), and one *Salmo l. lumi* (MW251454). The analysis also identified two genetic groups labelled as A and B. Group A included *Salmo l. lumi* (MW251455, L_7, L_8, L_15), *Salmo letnica* (AY926571, AY926573), *Salmo l. typicus* (T_5, T_7, T_34), *Salmo l. balcanicus* (b_17, b_28), and *Salmo l. aestivalis* (SA_8, SA_14, SA_20, SA_26, SA_28, SA_35). Group B encompassed all remaining terminals of *Salmo letnica*, *Salmo l. lumi*, *Salmo l. aestivalis*, *Salmo l. balcanicus*, and *Salmo l. typicus*. In general, the PCoA did not reveal any genetic structuring corresponding to the subdivision into morphotypes.

The network analysis of the whole dataset (Figure 3) was represented according to two categories: *Salmo* species (a) and major brown trout (*Salmo trutta*) phylogeographical lineages (b). In Figure 3a, three principal haplotypes were identified, predominantly including *Salmo letnica* species. The first and most common haplotype encompassed 106 individuals, comprising approximately 85% of *Salmo l. typicus*, 80% of *Salmo l. balcanicus*, 81% of *Salmo l. lumi*, 73% of *Salmo l. aestivalis*, and 1 *Salmo l.* sequence representing haplotype 12 (AY926570). This haplotype showed an incipient star-like shape, indicating a marked founder effect. It was surrounded by several derived haplotypes, differing by a few mutations, which mostly represented *Salmo farioides* sequences. Furthermore, 3 and 6 to 7-point mutations away from this haplotype were 3 private haplotypes belonging to *Salmo trutta*. The second most frequent haplotype, separated from the first by a single point mutation, contained 9% of *Salmo l. typicus*, 6% of *Salmo l. balcanicus*, 14% of *Salmo l.a lumi*, 15% of *Salmo l. aestivalis*, and 1 *Salmo letnica* sequence corresponding to haplotype 13 (AY926573). Moreover, a single *Salmo trutta* sequence diverged from this haplotype 6-point mutations. The third haplotype, which was 3-point mutation away from the second, comprised 8% of *Salmo l. balcanicus*, 6% of *Salmo l. aestivalis*, and a sequence of *Salmo l.* representing haplotype 15 (AY926572). Additionally, all sequences belonging to *Salmo ohridanus* diverged by 7-point mutations from the most common lineage, forming a separate cluster characterized by one common haplotype with six derived ones. This was consistent with findings from the phylogenetic tree and the PCoA analyses. In Figure 3b, the majority of *Salmo farioides* sequences, along with all *Salmo letnica* sequences, clustered within the Adriatic lineage (AD). This large haplogroup was surrounded by the other selected brown trout lineages (Atlantic, Mediterranean, Danubian, Marmoratus). Notably, as previously reported, *Salmo ohridanus* formed a separate group, distinct from all recognized lineages in the analysis.

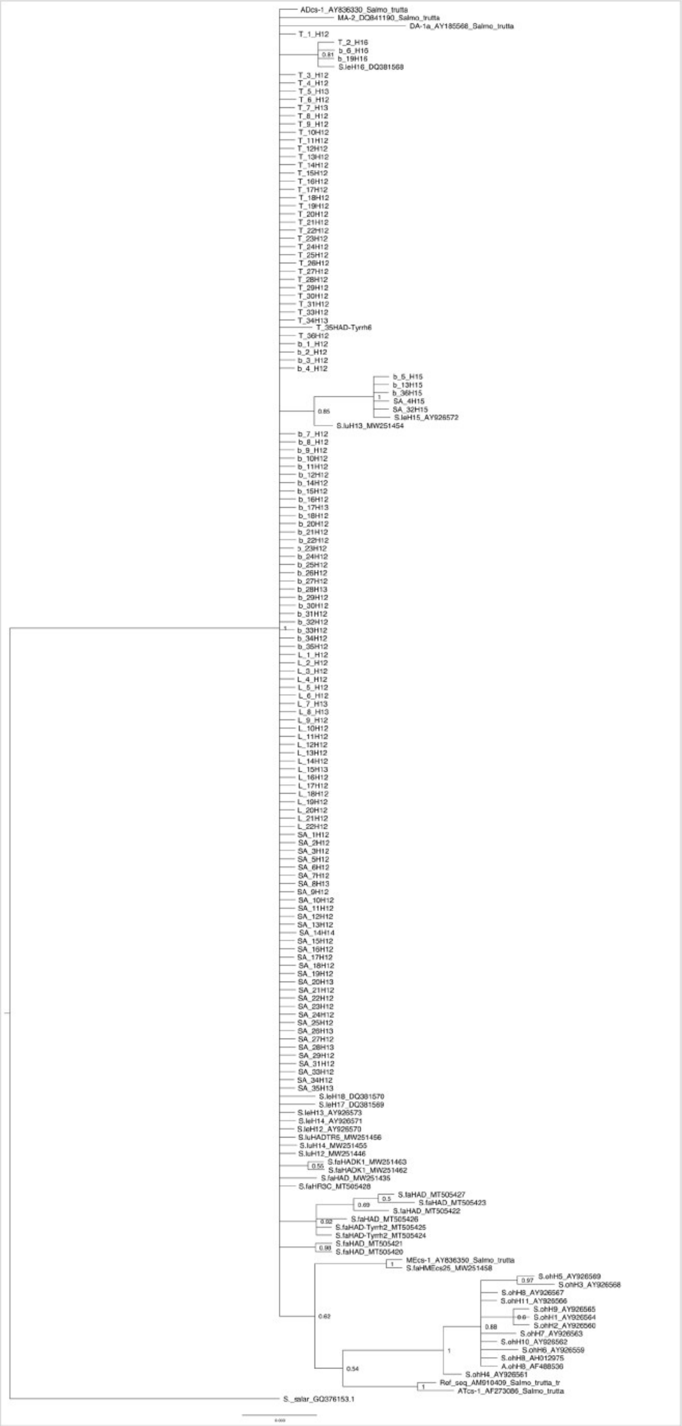


Figure 2. Bayesian phylogenetic tree based on the complete dataset, including *Salmo letnica* sequences and selected *Salmo* species. Specimens from the present study are labelled according to their species

abbreviations: "T" for *Salmo l. typicus*, "b" for *Salmo l. balcanicus*, "L" for *Salmo l. lumi*, and "SA" for *Salmo l. aestivalis*, followed by their corresponding assigned haplotype ("H"). GenBank sequences are labelled using species abbreviations: "S.le" for *Salmo letnica*, "S.lu" for *Salmo l. lumi*, "S.fa" for *Salmo farioides*, and "S.oh" for *Salmo ohridanus*, followed by their respective assigned haplotype and GenBank accession numbers. *Salmo trutta* sequences are labelled with the assigned haplotype, GenBank accession number, and species name. Node values represent posterior probabilities. In general, support values exceeding 0.9 were considered as well supported.

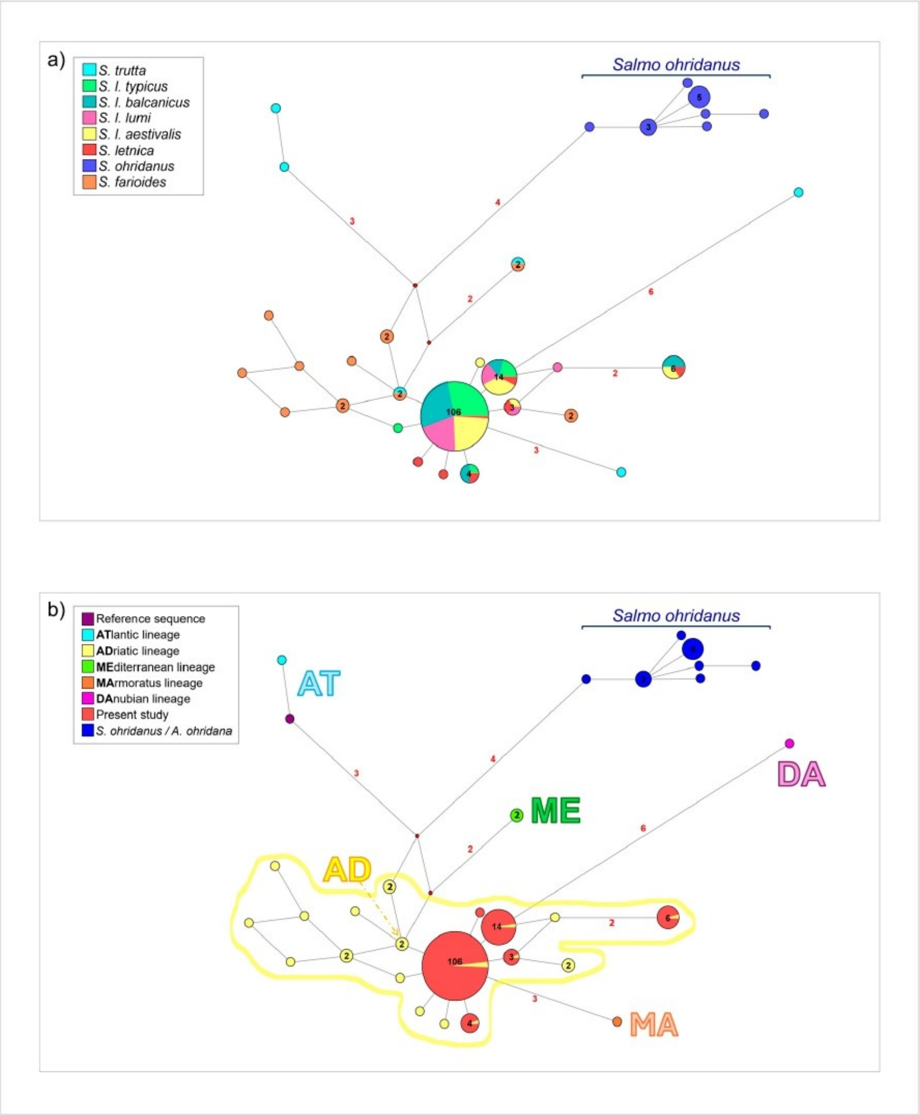


Figure 3. Network (Median-Joining) analysis of the whole dataset, including *Salmo letnica* sequences and selected *Salmo* species. The network is represented based on two categories: (**a**) *Salmo* species and (**b**) major phylogeographical lineages of brown trout (*Salmo trutta*). Each circle represents a unique haplotype, with its diameter proportional to the haplotype frequency. The number of mutations greater than 1 is indicated along the branches. Small red circles represent median vectors, which

correspond to intermediate missing or unsampled haplotypes. Colors in the network indicate the *Salmo* species (**a**) and the main phylogeographical lineages of brown trout (*Salmo trutta*) (**b**), as shown in the legend. The number of individuals sharing the same haplotype, when the frequency exceeds 1, is displayed inside the circle.

4. Discussion

The present study represents a preliminary investigation that provides insights into the genetic diversity of the endemic *Salmo letnica* in Lake Ohrid, using the mitochondrial Control Region as molecular marker. This region, widely used in genetic diversity studies, was chosen because it was the only molecular marker that permitted an extensive comparison with all the molecular data already available in literature. Moreover, it is particularly valuable for identifying incipient genetic structuring within populations and for detecting interspecific divergence [30,48].

Indeed, in this study, mitochondrial DNA has been effective in detecting genetic divergence between *Salmo letnica* and its sympatric congener, *Salmo ohridanus*. The latter species consistently exhibited a relevant level of genetic differentiation, forming a distinct cluster in all the analyses performed. This outcome aligns with previous studies, further supporting the genetic divergence of *Salmo ohridanus* from its related species [13,26,27]. It is likely that these two species have coexisted in Lake Ohrid over time, evolving independently due to differences in their ecological niche and reproductive behaviors [13,33]. The exact timing of *Salmo ohridanus* colonization of the lake, as well as its morphological divergence from *Salmo letnica* (Ohrid trout), remain unclear [13]. Interestingly, the distinct taxonomic position of *Salmo ohridanus* is attributed to its closer genetic relationship with *Salmo obtusirostris* [27], a species found in the Dalmatian River systems, which has been identified as its sister taxon [13,26,34]. The Dalmatian River systems are located in Croatia along the Adriatic coast and include the Neretva, Cetina, Krka, and Zrmanja rivers. These rivers are situated quite a distance from Lake Ohrid, which lies within a separate watershed in the western Balkans. Of note, Lake Ohrid's basin is distinct from the Adriatic drainage basin that encompasses the Dalmatian rivers.

Interestingly, despite the use of mitochondrial DNA as marker, our findings underscored a general genetic homogeneity within the different *Salmo letnica* populations analyzed, with no distinctive genetic traits among the *Salmo letnica* morphotypes (*Salmo l. typicus*, *Salmo l. balcanicus*, *Salmo l. lumi*, *Salmo l. aestivalis*), as revealed by both clustering and phylogenetic analyses. Indeed, the four morphotypes, which are in general associated with ecological adaptation phenomena [7], did not show relevant nucleotide differences. This lack of genetic divergence is consistent with results reported by Sušnik et al. (2007), who did not report the presence of population sub-structuring within the lake using both nuclear and mitochondrial markers. This occurrence suggests a potential common origin of Ohrid trout mitochondrial lineages, likely due to a founder effect that characterized the ancestral population in Lake Ohrid [26].

During the first half of the twentieth century, two main hypotheses have been proposed to explain the origin of Ohrid trout morphological forms. Stefanoviç (1948) suggested that these morphotypes evolved through intralacustrine speciation [7], while Karaman (1957) proposed independent invasions by distinct species [49].

Our results support the first hypothesis. While we acknowledge that the use of a single DNA fragment in our analyses might have influenced the results, and that informative genetic variation might have exclusively involved coding regions of the genome, a plausible scenario is that the adaptive processes leading to the emergence of the four *Salmo letnica* morphotypes may have been driven by differential gene expression in mitochondrial or nuclear DNA, possibly boosted by environmental pressures. This hypothesis warrants further investigation to validate its potential role in shaping the observed patterns. However, it should also be considered that differential gene expression, eventually prompted by epigenetic mechanisms, could result in morphological and ecological differences, producing phenotypic variation without altering DNA sequences.

Our findings may challenge the current taxonomic status of *Salmo letnica*, indicating that the adaptive processes shaping the ecological differentiation of the four Ohrid trout morphological forms may not be mirrored in their mitochondrial DNA. Similarly, two endemic species of the genus *Champsochromis* (Vertebrata, Perciformes, Cichlidae) inhabiting Lake Malawi exhibit differences in the morphology but show no mitochondrial genetic diversity [48]. This phenomenon has been associated with the absence of strong reproductive barriers and recent speciation, with morphological changes reflecting evolutionary processes while genetic differentiation remains limited. This highlights that environmental factors, such as habitat specializations or diet, can drive morphological diversification without corresponding changes in mitochondrial genetic structure [48].

Notably, ancient lakes may allow existing species to evolve through intralacustrine speciation, while also functioning as evolutionary reservoirs [2,50,51]. This variation suggests that intralacustrine speciation plays a significant role in diversity processes [2]. Furthermore, the processes of speciation could be influenced by varying degrees of isolation within the Ohrid watershed, potentially caused by intralacustrine barriers existing in both vertical and horizontal dimensions [2,52]. For instance, several morphological forms of the endemic *Proasellus* species (Crustacea, Isopoda, Asellidae) have been observed in Lake Ohrid, each associated with different bathymetric zones of the lake. Accordingly, the development of the four morphotypes of Ohrid trout may be a response to the different environmental features of the lake [2]. A similar case of intralacustrine diversification is seen in the Caucasian Lake Sevan, where the Sevan trout, *Salmo ischchan* Kessler, 1877, has differentiated into four sympatric ecomorphs or species [11]. These morphological forms are each associated with distinct spawning resources, showing how resource partitioning within a single lake system may drive speciation and contribute to ecological diversity [11].

Our hypothesis is further supported by the case of lake char, *Salvelinus namaycush* (Walbaum, 1792), in Lake Superior (North America) [53]. In this species, four morphologically distinct ecotypes showed no significant genetic structuring. This suggests that the species may be in an early stage of speciation, where phenotypic differences are associated with ecological characteristics, such as water depth, and potentially driven by diversifying selection acting on adaptive genes [53]. Likewise, *Salmo letnica* in Lake Ohrid may exhibit similar adaptive processes, with phenotypic variation likely emerging due to ecological factors despite limited genetic differentiation. Future studies are needed, incorporating additional molecular markers, particularly nuclear DNA, alongside mitochondrial DNA to validate this hypothesis. Such studies could also further explore the ecological divergence and niche specialization of Ohrid trout in relation to environmental gradients within the lake. This could offer valuable insights into the mechanisms underlying speciation in this ecosystem.

Another important outcome is that the network analysis clustered *Salmo letnica* and *Salmo farioides* sequences within the brown trout's Adriatic lineage, consistent with the findings of Snoj et al. (2009) [27]. According to Bianco (2014), *Salmo farioides* would represent the native population of the Adriatic trout lineage [54]. Commonly known as the brook-dwelling trout, *Salmo farioides* inhabits streams ranging from the Krka to the Drim–Skadar River systems [55]. The Drim River has two main branches: the Black Drim (Crni Drim), which originates from Lake Ohrid and flows through North Macedonia into Albania, and the White Drim (Beli Drim), which begins in Kosovo. These branches merge in northern Albania, forming the Drim River, which ultimately flows into the Adriatic Sea [55]. Historically, prior to the construction of hydropower dams in 1960s [27], Lake

Ohrid was naturally connected to the Drim River, creating a continuous aquatic network that spatially connected the two species. However, in 1969, the Spilje Hydropower Plant was activated [56], likely disrupting the direct connection between Lake Ohrid and the Drim River. This interruption may have played a role in isolating *Salmo letnica* and *Salmo farioides*, contributing to their independent evolutionary paths.

Additionally, the absence of haplotype sharing between *Salmo letnica* and *Salmo farioides* suggests that the present-day *Salmo letnica* populations likely evolved from a small group of *Salmo farioides* founders that became landlocked in Lake Ohrid as a consequence of the dam construction. Over time, evolutionary forces, such as genetic drift first and selective sweep second, may have led to the specialization and adaptation of this founding population to the environmental conditions of Lake Ohrid, resulting into the modern *Salmo letnica* population, while *Salmo farioides* continued evolving in its riverine habitat. This phenomenon aligns with adaptive mechanisms of those allochthonous species in which haplotypes that are common in the derivative population are rare in the source population. Such a pattern was also observed for the alien invasive species *Fistularia commersonii*, *Procambarus virginalis* and *Callinectes sapidus* [57–59].

This scenario can be further explained by the punctuated equilibrium model, which proposes that species experience extended periods of relative evolutionary stability, followed by brief intervals of rapid evolutionary change and speciation, often prompted by radical environmental shifts [60,61]. The isolation of *Salmo farioides* in Lake Ohrid, following the damming of the Drim river, may have led the species to undergo a rapid evolutionary adaptation, accounting for the differentiation into the contemporary *Salmo letnica*. This scenario was potentially followed by the ecological adaptation of Ohrid trout to the lake's environmental conditions, which may have contributed to the emergence of the four *Salmo letnica* morphotypes. These morphological forms may currently be in the early stages of an incipient speciation process, which might not yet have produced detectable genetic differentiation (at least at the mitochondrial Control Region) within the species.

5. Conclusions

This pilot study provides valuable hints for future inferences on the complex dynamics between genetic variation and ecological differentiation in the endemic Ohrid trout, *Salmo letnica*. Despite the presence of distinct *Salmo letnica* morphotypes in Lake Ohrid, our analysis of the mitochondrial Control Region revealed no evidence of genetic structuring, suggesting a possible overall low genetic variability for the mitochondrial DNA. In this context, morphological variability may have been acquired as a consequence of rapid adaptation phenomena mediated by a reduced level of polymorphisms in coding regions. For this reason, the present morphotypes may represent ecologically adapted variants of *Salmo letnica* rather than genetically distinct subspecies. This interpretation suggests the potential adaptive ability of the Ohrid trout to the lake's diverse environmental conditions.

Additionally, interesting results suggest that the modern *Salmo letnica* population of Lake Ohrid likely evolved from a few *Salmo farioides* founders, adapting to the lake's environmental conditions over time after the interruption of the connection between Lake Ohrid and Drim River. This scenario emphasizes the role of adaptive evolution in shaping the biodiversity of Lake Ohrid. In contrast, the clear genetic divergence of *Salmo ohridanus* reinforces the distinct evolutionary trajectory of *Salmo letnica* within the lake.

Further investigations, integrating additional molecular markers, such as nuclear DNA, and, if possible, employing a genome-wide approach, are crucial to validate and support the results reported in this study. Detailed ecological studies are also essential to fully understand the drivers of phenotypic variation and speciation in the Ohrid Lake's ancient ecosystem.

In particular, research involving a larger number of *Salmo farioides* and *Salmo letnica* individuals will be necessary to clarify their evolutionary relationship, contributing to a better comprehension of the origin of Ohrid trout and its morphotypes.

Finally, our findings may have significant implications for the management and conservation of Lake Ohrid's ecological diversity, particularly in supporting efforts to protect its endemic trout species.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/life15010052/s1, Figure S1: Principal Coordinates Analysis (PCoA) performed on the dataset including *Salmo letnica* sequences and selected sequences belonging to *Salmo* species reported from Lake Ohrid. Graphs "a" and "b" are described in the text. The bi-dimensional plot illustrates genetic differentiation among samples based on the nucleotide substitutions per site within the dataset. Table S1: *Salmo letnica* samples collected in the present study. The collection period ranged from September 2022 to November 2022 and from March 2023 to July 2023. The table includes the sample code, GenBank accession number, sampling area and site, and species for each specimen; Table S2: Control Region Dataset. The table lists Control Region sequences belonging to *Salmo letnica* and other *Salmo* species which were used in the present study, sourced from the GenBank online database.

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

Funding: This research was funded under: (1) the project titled "Study of the genetic diversity of 4 forms of the Koran (*Salmo letnica*), as an endemic species of Lake Ohrid" supported by the Global Environment Facility/Small Grants Programme ("GEF/SGP"), grant number ALB/SGP/OP7/Y2/CORE/BD/2022/03; (2) the project titled "Protection of Ohrid Lake trout through capture and release of breeders involved in egg collection for restocking program" supported by United Nations Office for Project Services (UNOPS) ALB/SGP/OP6/Y5/CORE/BD/2020/09), Programi Mbeshtetjes se Shoqatave Rurale (Rural Association Support Programme), Tirana, Albania.

Institutional Review Board Statement: Samples used for this study were obtained following the return of licensed fishermen from their fishing expeditions. Local fishermen contributed to the sampling process by identifying the four forms based on morphological characteristics. Specimens were collected using a standardized non-lethal sampling method, in which fin clips were taken from each individual. Fin clips were preserved in absolute ethanol and subsequently utilized for DNA extraction. The fish caught by fisherman are dead and are used for commercial reasons (catering and restaurants) and/or family commercial reasons (restaurants) and/or family consume.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequences obtained in the present study for the mitochondrial Control Region were deposited in the GenBank database under the accession numbers PQ583195–PQ583321.

Acknowledgments: Daria Sanna would like to express her deepest gratitude to Giuseppino Sanna whose never-ending support and guidance has been an invaluable gift throughout their journey together.

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

References

1. Stanković, S. The Balkan Lake Ohrid and its living world. In *Monographiae Biologicae*; Uitgeverij Dr. W. Junk: Den Haag, The Netherlands, 1960; Volume 9.
2. Albrecht, C.; Wilke, T. Ancient Lake Ohrid: Biodiversity and evolution. *Hydrobiologia* **2008**, *615*, 103–140. [CrossRef]
3. Wagner, B.; Lotter, A.F.; Nowaczyk, N.; Reed, J.M.; Schwalb, A.; Sulpizio, R.; Valsecchi, V.; Wessels, M.; Zanchetta, G. 40,000-year record of environmental change from ancient Lake Ohrid (Albania and Macedonia). *J. Paleolimnol.* **2009**, *41*, 407–430. [CrossRef]
4. Popovska, C.; Bonacci, O. Basic data on the hydrology of Lakes Ohrid and Prespa. *Hydrol. Process.* **2007**, *21*, 658–664. [CrossRef]
5. Matzinger, A.; Spirkovski, Z.; Patceva, S.; Wüest, A. Sensitivity of ancient Lake Ohrid to local anthropogenic impacts and global warming. *J. Great Lakes Res.* **2006**, *32*, 158–179. [CrossRef]
6. Karaman, S. *Pisces Macedoniae*; Hrvatska Stamparija: Split, Croatia, 1924; pp. 1–90.
7. Stefanoviç, D. Recherches systematiques et écologiques sur les Salmonides d'Ohrid. *Acad. Serbe Sci. Éd. Spéc.* **1948**, *139*, 1–207.
8. Wagner, B.; Wilke, T.; Francke, A.; Albrecht, C.; Baumgarten, H.; Bertini, A.; Combourieu-Nebout, N.; Cvetkoska, A.; D'Addabbo, M.; Donders, T.H.; et al. The environmental and evolutionary history of Lake Ohrid (FYROM/Albania): Interim results from the SCOPSCO deep drilling project. *Biogeosciences* **2017**, *14*, 2033–2054. [CrossRef]
9. Kostoski, G.; Albrecht, C.; Trajanovski, S.; Wilke, T. A freshwater biodiversity hotspot under pressure—Assessing threats and identifying conservation needs for ancient Lake Ohrid. *Biogeosciences* **2010**, *7*, 3999–4015. [CrossRef]
10. Rakaj, N.; Flloko, A. Conservation status of freshwater fish of Albania. *Biol. Conserv.* **1995**, *72*, 195–199. [CrossRef]
11. Levin, B.; Simonov, E.; Gabrielyan, B.K.; Mayden, R.L.; Rastorguev, S.M.; Roubenyan, H.R.; Sharko, F.S.; Nedoluzhko, A.V. Caucasian treasure: Genomics sheds light on the evolution of half-extinct Sevan trout, *Salmo ischchan*, species flock. *Mol. Phylogenet. Evol.* **2022**, *167*, 107346. [CrossRef]
12. Phillips, R.B.; Matsuoka, M.P.; Konon, I.; Reed, K.M. Phylogenetic Analysis of Mitochondrial and Nuclear Sequences Supports Inclusion of *Acantholingua ohridana* in the Genus *Salmo*. *Copeia* **2000**, *2000*, 546–550. [CrossRef]
13. Sušnik, S.; Knizhin, I.; Snoj, A.; Weiss, S. Genetic and morphological characterization of a Lake Ohrid endemic, *Salmo* (*Acantholingua*) *ohridanus* with a comparison to sympatric *Salmo trutta*. *J. Fish Biol.* **2006**, *68*, 2–23. [CrossRef]
14. Karaman, S. Salmonidi Balkana. *Glas. Skopskog Naučn. Druš.* **1927**, *2*, 253–268.
15. Karaman, S. Beitrag zur kenntnis der su¨sswasserfische Jugoslawiens. *Bull. Soc. Sci. Skoplje* **1938**, *18*, 131–139.
16. Kottelat, M. European freshwater fishes. An heuristic checklist of the freshwater fishes of Europe (exclusive of former USSR), with an introduction for non-systematists and comments on nomenclature and conservation. *Biologia* **1997**, *52*, 1–271.
17. Sell, J.; Spirkovski, Z. Mitochondrial DNA differentiation between two forms of trout *Salmo letnica*, endemic to the Balkan Lake Ohrid, reflects their reproductive isolation. *Mol. Ecol.* **2004**, *13*, 3633–3644. [CrossRef] [PubMed]
18. Smederevac-Lalić, M.; Špelić, I.; Đug, S.; Pengal, P.; Çobani, M.; Mrdak, D.; Piria, M. Political and Socio-Economic Aspects of Fisheries in Inland and Coastal Waters of the Western Balkan. In *Ecological Sustainability of Fish Resources of Inland Waters of the Western Balkans: Freshwater Fish Stocks, Sustainable Use and Conservation*; Simić, V., Simić, S., Pešić, V., Eds.; Springer International Publishing: Cham, Switzerland, 2023; Volume 43, pp. 655–688. [CrossRef]
19. Prifti, V. Artificial fertilization of the Ohrid trout and the presence of its summer form in the lake. *Sci. Pap. Ser. D. Anim. S.* **2016**, *59*, 342–345.
20. Spirkovski, Z. Changes in the spawning ecology of the Lake Ohrid trout, *Salmo letnica* (Karaman). In Proceedings of the 2nd Congress of Ecologists of the Republic of Macedonia with International Participation. Special issues of Macedonian Ecological Society, Ohrid, Macedonia, 25–29 October 2003.
21. Spirkovski, Z.; Ilik-Boeva, D.; Ritterbusch, D.; Peveling, R.; Pietrock, M. Ghost net removal in ancient Lake Ohrid: A pilot study. *Fish. Res.* **2019**, *211*, 46–50. [CrossRef]
22. Prifti, V.; Cake, A. The Impact of Environmental Conditions on Growth and Development of *Salmo letnica* Smolt. *Sci. Pap. Anim. Sci. Biotechnol.* **2017**, *50*, 131.
23. Bernatchez, L.; Guyomard, R.; Bonhomme, F. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. *Mol. Ecol.* **1992**, *1*, 161–173. [CrossRef]
24. Apostolidis, A.P.; Karakousis, Y.; Triantaphyllidis, C. Genetic differentiation and phylogenetic relationships among Greek *Salmo trutta* L. (brown trout) populations as revealed by RFLP analysis of PCR amplified mitochondrial DNA segments. *Heredity* **1996**, *77*, 608–618. [CrossRef]
25. Apostolidis, A.P.; Triantaphyllidis, C.; Kouvatsi, A.; Economidis, P.S. Mitochondrial DNA sequence variation and phylogeography among *Salmo trutta* L. (Greek brown trout) populations. *Mol. Ecol.* **1997**, *6*, 531–542. [CrossRef]
26. Sušnik, S.; Snoj, A.; Wilson, I.F.; Mrdak, D.; Weiss, S. Historical demography of brown trout (*Salmo trutta*) in the Adriatic drainage including the putative *S. letnica* endemic to Lake Ohrid. *Mol. Phylogenet. Evol.* **2007**, *44*, 63–76. [CrossRef] [PubMed]
27. Snoj, A.; Marić, S.; Berrebi, P.; Crivelli, A.J.; Shumka, S.; Sušnik, S. Genetic architecture of trout from Albania as revealed by mtDNA control region variation. *Genet. Sel. Evol.* **2009**, *41*, 22. [CrossRef]

28. 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]
29. Wan, Q.H.; Wu, H.; Fujihara, T.; Fang, S.G. Which genetic marker for which conservation genetics issue? *Electrophoresis* **2004**, *25*, 2165–2176. [CrossRef]
30. 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]
31. Ford, M.; Salmo letnica. The IUCN Red List of Threatened Species 2024, e.T19858A137328386. Available online: https://www.iucnredlist.org/species/19858/137328386 (accessed on 9 December 2024).
32. Milana, V.; Ciampoli, M.; Sola, L. mtDNA sequences of *Sphyraena viridensis* (Perciformes: Sphyraenidae) from Italy: Insights into historical events and the phylogeny of the genus. *Biol. J. Linn. Soc.* **2014**, *113*, 635–641. [CrossRef]
33. Okonechnikov, K.; Golosova, O.; Fursov, M.; The UGENE Team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [CrossRef]
34. Snoj, A.; Melkič, E.; Sušnik, S.; Muhamedagić, S.; Dovč, P. DNA phylogeny supports revised classification of *Salmothymus obtusirostris*. *Biol. J. Linn. Soc.* **2002**, *77*, 399–411. [CrossRef]
35. Sušnik, S.; Schöffmann, J.; Snoj, A. Phylogenetic position of Salmo (Platysalmo) platycephalus Behnke 1968 from south-central Turkey, evidenced by genetic data. *J. Fish Biol.* **2004**, *64*, 947–960. [CrossRef]
36. Grapci-Kotori, L.; Vavalidis, T.; Zogaris, D.; Šanda, R.; Vukić, J.; Geci, D.; Ibrahimi, H.; Bilalli, A.; Zogaris, S. Fish distribution patterns in the White Drin (Drini i Bardhë) river, Kosovo. *Knowl. Manag. Aquat. Ecosyst.* **2020**, *421*, 29. [CrossRef]
37. Duftner, N.; Weiss, S.; Medgyesy, N.; Sturmbauer, C. Enhanced phylogeographic information about Austrian brown trout populations derived from complete mitochondrial control region sequences. *J. Fish Biol.* **2003**, *62*, 427–435. [CrossRef]
38. Cortey, M.; García-Marín, J.L. Evidence for phylogeographically informative sequence variation in the mitochondrial control region of Atlantic brown trout. *J. Fish Biol.* **2002**, *60*, 1058–1063. [CrossRef]
39. Cortey, M.; Pla, C.; García-Marín, J.L. Historical biogeography of Mediterranean trout. *Mol. Phylogenet. Evol.* **2004**, *33*, 831–844. [CrossRef]
40. Meraner, A.; Baric, S.; Pelster, B.; Dalla Via, J. Trout (*Salmo trutta*) mitochondrial DNA polymorphism in the centre of the marble trout distribution area. *Hydrobiologia* **2007**, *579*, 337–349. [CrossRef]
41. Splendiani, A.; Fioravanti, T.; Giovannotti, M.; Olivieri, L.; Ruggeri, P.; Nisi Cerioni, P.; Vanni, S.; Enrichetti, F.; Caputo Barucchi, V. Museum samples could help to reconstruct the original distribution of Salmo trutta complex in Italy. *J. Fish Biol.* **2017**, *90*, 2443–2451. [CrossRef]
42. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
43. 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]
44. 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]
45. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
46. 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] [PubMed]
47. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef]
48. Hashem, S.; Kawai, K.; Fatsi, P.S.K.; Kodama, A.; Saito, H. Genetic relationships of cichlid fishes from Lake Malawi based on mitochondrial DNA sequences. *Limnology* **2020**, *21*, 151–163. [CrossRef]
49. Karaman, S. The Radika river trouts. *Folia Balc.* **1957**, *1*, 57–70.
50. Wilson, A.B.; Glaubrecht, M.; Meyer, A. Ancient lakes as evolutionary reservoirs: Evidence from the thalassoid gastropods of Lake Tanganyika. *Proc. R. Soc. Lond. B.* **2004**, *271*, 529–536. [CrossRef] [PubMed]
51. Albrecht, C.; Trajanovski, S.; Kuhn, K.; Streit, B.; Wilke, T. Rapid evolution of an ancient lake species flock: Freshwater limpets (Gastropoda: Ancylidae) in the Balkan lake Ohrid. *Org. Divers. Evol.* **2006**, *6*, 294–307. [CrossRef]
52. Radoman, P. Hydrobioidea a superfamily Prosobranchia (Gastropoda), II. Origin, Zoogeography, Evolution in the Balkans and Asia Minor. In *Monographs Institute of Zoology*; Institute of Zoology: Belgrade, Serbia, 1985.
53. Baillie, S.M.; Muir, A.M.; Hansen, M.J.; Krueger, C.C.; Bentzen, P. Genetic and phenotypic variation along an ecological gradient in lake trout *Salvelinus namaycush*. *BMC Evol. Biol.* **2016**, *16*, 1–16. [CrossRef]
54. Bianco, P.G. An update on the status of native and exotic freshwater fishes of Italy. *J. Appl. Ichthyol.* **2014**, *30*, 62–77. [CrossRef]

55. Marić, S.; Sušnik Bajec, S.; Schöffmann, J.; Kostov, V.; Snoj, A. Phylogeography of stream-dwelling trout in the Republic of Macedonia and a molecular genetic basis for revision of the taxonomy proposed by *S. karaman*. *Hydrobiologia* **2017**, *785*, 249–260. [CrossRef]
56. Ivanova-Davidovic, J.; Cingoski, V.; Pavleski, V.; Saveski, V. Upgrading of the Spilje HPP. In Proceedings of the 13th Seminar on Hydropower Plants, Vienna, Austria, 22–24 November 2004.
57. 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]
58. 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]
59. Locci, C.; Azzena, I.; Pascale, N.; Ciccozzi, A.; Deplano, I.; Giantsis, I.A.; Papadopoulos, D.K.; Lattos, A.; Orrù, F.; Puzzi, C.M.; et al. A sister species for the blue crab, *Callinectes sapidus*? A tale revealed by mitochondrial DNA. *Life* **2024**, *14*, 1116. [CrossRef] [PubMed]
60. Pennell, M.W.; Harmon, L.J.; Uyeda, J.C. Is there room for punctuated equilibrium in macroevolution? *Trends Ecol. Evol.* **2014**, *29*, 23–32. [CrossRef] [PubMed]
61. Eldredge, N.; Gould, S. Punctuated equilibria: An alternative to phyletic gradualism. In *Models in Paleobiology*; Schopf Tomas, J.M., Ed.; Freeman, Cooper & Company: San Francisco, CA, USA, 1972; pp. 82–115.

Chapter 4

Salariopsis fluviatilis

Salariopsis fluviatilis Asso 1801 (Chordata: Actinopterygii: Blenniidae), commonly known as the freshwater blenny, is a benthic fish species inhabiting river basins and lakes of the Mediterranean region (Vecchioni et al., 2022; Figure 4). The species is typically characterised by small, spatially fragmented, and often isolated populations, which have increased its vulnerability to anthropogenic pressures and ongoing environmental and climate changes (Freeman et al., 1990; Côté et al., 1999; Alp & Kara, 2007; Méndez et al., 2019).

Figure 4. *Salariopsis fluviatilis* specimen from Strona di Omegna, Lake Maggiore. Photo credits: Stefano Bovero.

Across its distribution range, populations of *Salariopsis fluviatilis* exhibit morphological and behavioural variability in response to local environmental conditions and habitat characteristics (Perdices et al., 2000; Neat et al., 2003). In this context, various studies have proposed that *S. fluviatilis* represents a species complex, rather than a single taxon (Belaiba et al., 2019; Wagner et al., 2021). Consistent with this hypothesis, two additional freshwater blenny species have been formally described to date: *S. economidisi*, endemic to Lake Trichonis in Greece (Kottelat, 2004), and *S. atlantica*, endemic to the Seboui river basin in Morocco (Doadrio et al., 2011).

This taxonomic framework supports the presence of distinct evolutionary lineages within the *Salariopsis fluviatilis* species complex. In particular, Wagner et al. (2021) identified at least six divergent genetic lineages across the species' distribution range, Middle East, Guadiana, Algeria-Verde, Occidental basin, North Oriental basin, and the Island of Crete, thus supporting previous findings by Belaiba et al. (2019), who suggested that Middle Eastern Mediterranean populations of *S. fluviatilis* could represent a differentiated taxon (*S.* cf *fluviatilis*).

However, several gaps remain in the understanding of the species genetic structure and taxonomic status across its distribution range. These gaps are particularly evident in Italy, where *S. fluviatilis* is distributed in a discontinuous way (Zava & Violani, 1991; Gallo et al., 2012) and where populations have remained largely under-investigated. In this context, the island of Sardinia deserves a particular attention, as *S. fluviatilis* represents one of the few putative native freshwater fish species known from the island (Orrù et al., 2010).

Considering this background, the study presented in this chapter primarily focused on *Salariopsis fluviatilis* populations from under-explored Italian regions, with genetic analyses conducted using the mitochondrial Control Region and the 16S ribosomal gene as molecular markers. In particular, the main objectives of the study were to: (i) draw inferences on the phylogeographic relationships among *S. fluviatilis* populations coming from Sardinia and north-western Italian mainland; (ii) reconstruct phylogenetic patterns among the species populations from the whole Mediterranean region; and (iii) clarify the taxonomic status of the species across its distribution range.

References

Alp, A., & Kara, C. (2007). Distribution pattern and morphological differences between the sexes of river blenny, *Salaria fluviatilis* (Asso, 1801), in the Ceyhan River basin, Turkey. *Turkish Journal of Zoology, 31*, 113–120.

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. https://doi.org/10.1016/j.ympev.2019.02.026

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*, 359–367. https://doi.org/10.1016/S0006-3207(98)00072-X

Doadrio, I., Perea, S., & Yahyaoui, A. (2011). Morphological and molecular analyses of freshwater blennids: A new species of the genus *Salaria* Forsskål, 1775 (Actinopterygii, Blenniidae) from Morocco. *Graellsia, 67*, 151–173. https://doi.org/10.3989/graellsia.2011.v67.042

Freeman, M. C., Viñolas, D., Grossman, G. D., & De Sostoa, A. (1990). Microhabitat use by *Blennius fluviatilis* in the Río Matarraña, Spain. *Freshwater Biology, 24*, 335–345. https://doi.org/10.1111/j.1365-2427.1990.tb00714.x

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*, 246–251. https://doi.org/10.1080/11250003.2011.631943

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*, 121–137. https://doi.org/10.5962/bhl.part.80231

Méndez, L., Perdices, A., & Machordom, A. (2019). Genetic structure and diversity of Iberian populations of the freshwater blenny *Salaria fluviatilis* (Asso, 1801) and its conservation implications. *Conservation Genetics, 20*, 1223–1236. https://doi.org/10.1007/s10592-019-01205-0

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*, 374–387. https://doi.org/10.1046/j.1095-8649.2003.00159.x

Orrù, F., Deiana, A. M., & Cau, A. (2010). Introduction and distribution of alien freshwater fishes on the island of Sardinia (Italy): An assessment based on existing data sources. *Journal of Applied Ichthyology, 26*, 46–52. https://doi.org/10.1111/j.1439-0426.2010.01501.x

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*, 723–731. https://doi.org/10.1643/0045-8511(2000)000[0723:GDAOOM]2.0.CO;2

Vecchioni, L., Ching, A. C., Marrone, F., Arculeo, M., Hundt, P. J., & Simons, A. M. (2022). Multi-locus phylogenetic analyses of the *Almadablennius* clade reveal inconsistencies with the present taxonomy of blenniid fishes. *Diversity, 14*, 53. https://doi.org/10.3390/d14010053

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*, 1832–1847. https://doi.org/10.1111/ddi.13372

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 di 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*

check for updates

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.

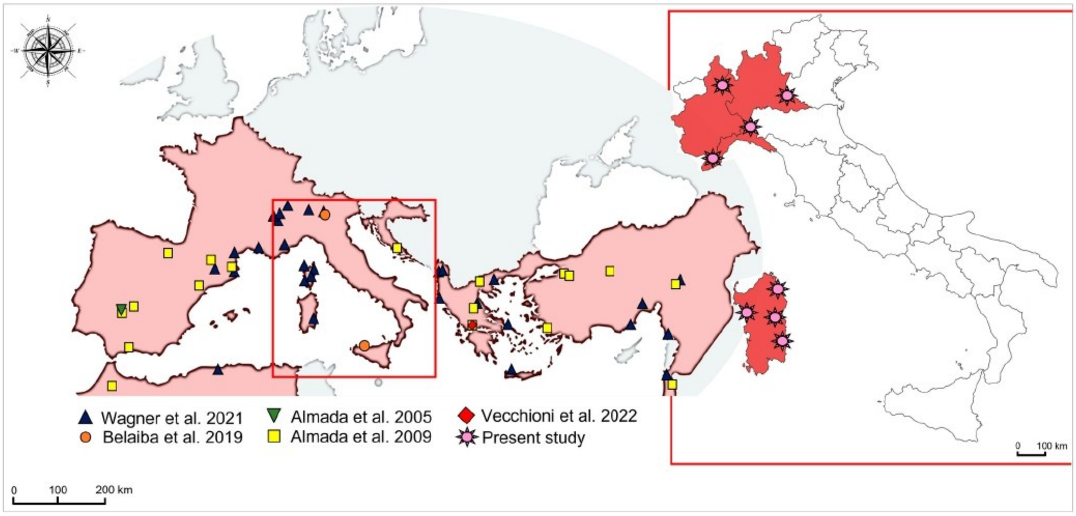


Figure 1. Map of samples' collection sites. The map shows the geographical origin of the sequences isolated in the present study along with those from previous research [1,8,15,16,24].

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.

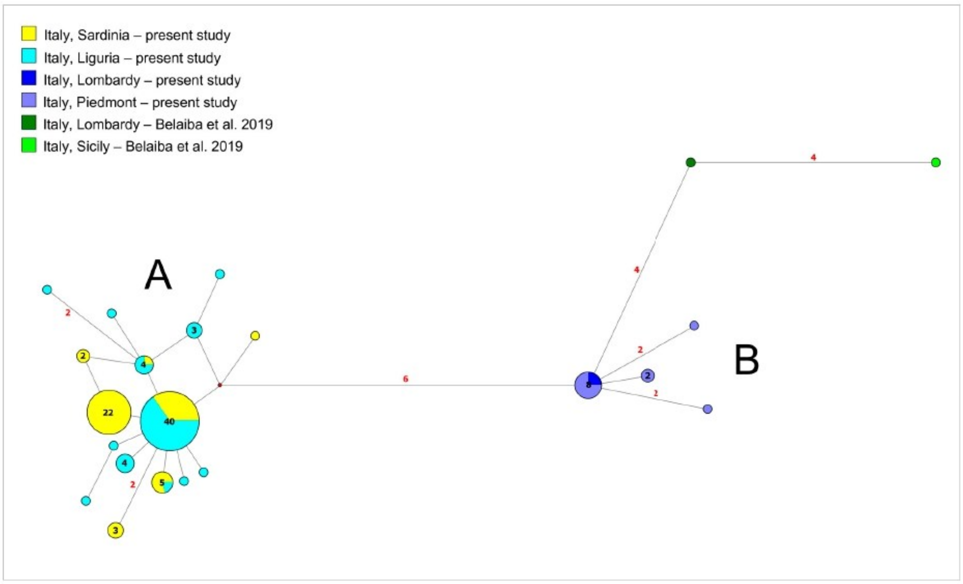


Figure 2. Median-joining network analysis. The network includes all the Italian sequences (from the present study and also from a previous study [15]) for the Control Region and 16s (merged in a concatenated dataset). Clusters A and B are described in the text. The number of mutations between sequences 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%.

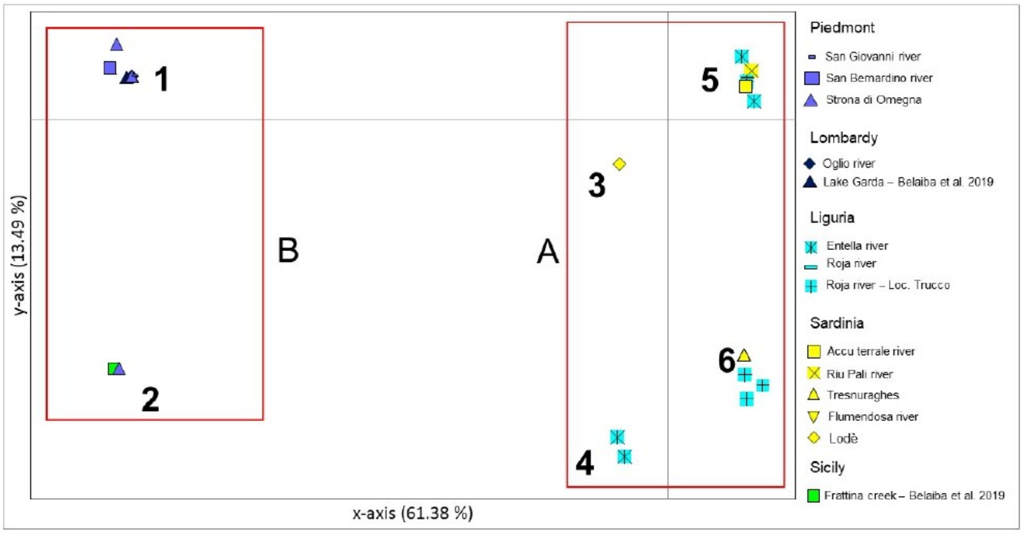


Figure 3. Principal coordinates analysis performed on Italian sequences from the present study and also from a previous study [15]. Bidimensional plot shows the genetic differentiation among specimens due to the nucleotide substitutions per site found in the dataset. Belaiba et al., 2019 in the legend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

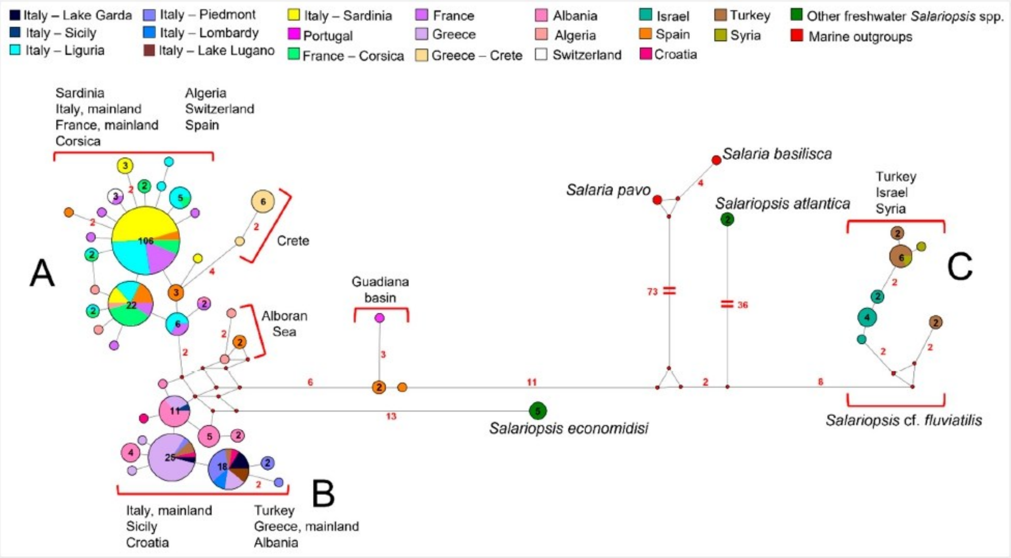


Figure 4. Median-joining network analysis performed on the whole Control Region dataset. Clusters A, B and C are described in the text. The number of mutations between sequences that are greater than $n = 1$ are reported on network branches. Additionally, the number of individuals showing the same haplotype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

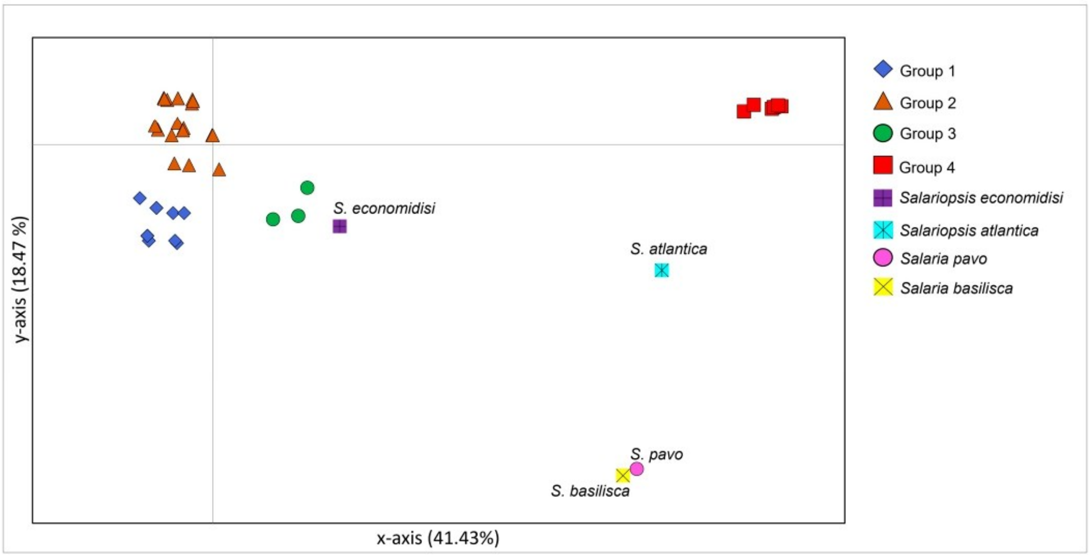


Figure 5. Principal coordinates analysis performed on the whole Control Region dataset. Bidimensional plot shows the genetic differentiation among specimens due to the nucleotide substitutions per site found in the dataset.

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]

General conclusions

This doctoral thesis comprises four case studies conducted during my doctoral programme, focusing on different aspects of aquatic biodiversity. The thesis includes two research articles centred on *Pinna nobilis*, and one research article for each of the other species investigated, *Callinectes sapidus*, *Salmo letnica* and *Salariopsis fluviatilis*. During my research path, I explored a wide range of genetic and evolutionary processes across various biological systems, encompassing marine and freshwater environments, as well as native, endemic, and invasive taxa. By adopting multiple population genetic approaches, including phylogenetics, phylogeography, and molecular taxonomy, this thesis provides an integrative perspective on the evolutionary history, genetic structure, dispersal dynamics, and taxonomic complexity of aquatic organisms in response to historical and contemporary challenges.

The two studies on *Pinna nobilis* that I reported in Chapter 1 explored the evolutionary history and the genetic make-up of this species within the Mediterranean region over time. The first study revealed that *P. nobilis* diverged from its Atlantic ancestor approximately 2.5 million years ago, following the entrance of Atlantic waters into the Mediterranean Sea after the Zanclean flood, which occurred around 5.3 million years ago. The results further indicated that the central western Mediterranean represented the first area where the early population of the species originated, and from this region, it later expanded towards the Adriatic Sea and the eastern Mediterranean. These findings allowed the reconstruction of the historical expansion pathways of *P. nobilis*, providing insights into how the species adapted to the Mediterranean environment through time. Based on this evolutionary framework, the second study showed that the mitochondrial genetic structure of *P. nobilis* has remained stable over time. Pleistocene mitochondrial haplotypes persisted across ancient, modern (pre-MMEs), and surviving populations (post-MMEs), indicating the long-term persistence of the species' genetic diversity, despite major environmental changes and human pressures. These results suggest that these haplotypes are associated with adaptative or protective allelic variants, which contributed to the resilience of the species and its ability to persist under changing conditions, even in the face of the dramatic reduction in population numbers. Overall, these outcomes offer important information for the development of effective conservation strategies and yield encouraging perspectives for the future survival of the species, provided that Pleistocene allelic variants can be preserved or maintained among surviving individuals through time.

Chapter 2, which focuses on *Callinectes sapidus*, showed how invasion dynamics can shape the genetic composition of invasive species, as well as their expansion and adaptation to new environments. In this context, the main results revealed the existence of two distinct genetic groups (taxonomic units) across the distribution range of *C. sapidus*. This pattern suggests the potential presence of two putative sister species, one mainly diffused in Central and South America and the other primarily distributed in North America and the Mediterranean Sea. These findings revealed that *C. sapidus* may represent a species complex, rather than a single taxon across its range of distribution. This outcome is of particular importance, as it may have significant implications for the monitoring, management and control of the species in both its native and introduced areas.

The study on *Salmo letnica* presented in Chapter 3 addressed the controversial taxonomic status of this species endemic to Lake Ohrid. The study evidenced the absence of significant genetic structure among the four putative morphological forms described to date, suggesting either an incipient divergent process not yet resulted in mitochondrial genetic differentiation or the existence of ecologically adapted variants whose morphological traits have been shaped by environmental pressures. This result reflects the complexity of adaptive and evolutionary mechanisms as a consequence of habitat changes and anthropogenic stressors. Hence, distinct morphological traits may not always be associated with underlying genetic differentiation, but may rather be the result of adaptive processes, highlighting the complex dynamics between genetic composition and ecological differentiation.

In Chapter 4, the phylogeographic patterns of *Salariopsis fluviatilis* populations and the taxonomic uncertainties of the species were investigated across the whole Mediterranean basin. The main findings highlighted the presence of genetic structuring among Italian *S.*

fluviatilis populations, offering insights into the dispersal trajectories of the species across the Italian mainland and the island of Sardinia. Additionally, three taxonomic units were identified within the distribution range of the species, supporting the existence of a species complex: *Salariopsis fluviatilis sensu stricto*, with two sub-groups (Occidental and North Oriental Mediterranean), a taxonomic entity from the Middle East, and another from the Iberian Guadiana River basin. These results evidenced how limited population size and high degree of fragmentation can influence genetic differentiation and evolutionary processes within freshwater systems.

Final considerations

The results obtained in the four case studies presented in this thesis highlight how the choice of molecular markers represents a crucial methodological component for interpreting evolutionary and demographic processes across different categories of species, including marine and freshwater taxa, as well as native, endemic, and invasive species.

Throughout this thesis, mitochondrial molecular markers were primarily employed. Mitochondrial DNA (mtDNA) is characterised by maternal inheritance, a compact genome, a high copy number per cell, and the absence of recombination. In general, the mitochondrial genome comprises 37 genes (protein-coding genes, tRNAs, and rRNAs) and a single major non-coding region, known as the Control Region or D-Loop, although minor variations in gene content have been reported in some taxa. Moreover, mtDNA has a relatively high mutation rate compared to nuclear DNA.

In the case of *Pinna nobilis*, the protein-coding gene Cytochrome c Oxidase subunit I (COI) proved to be highly informative for phylogenetic and phylogeographic reconstruction among and within populations, enabling the identification of historical divergence patterns and the detection of genetic structures consistent with palaeogeographical events and colonisation dynamics. For *Callinectes sapidus*, the use of COI gene was essential for placing the study within a global geographic framework, allowing the assessment of invasion dynamics and the detection of a potential species complex characterised by the presence of two divergent taxonomic units distributed in the geographic range of the species.

In the case of *Salmo letnica*, the use of the mitochondrial Control Region, more variable than the coding COI gene, allowed the exploration of more recent genetic differentiation patterns and contributed to deepen on taxonomic uncertainties. Similarly, its application in *Salariopsis fluviatilis* enabled a comprehensive investigation of the phylogenetic patterns across the whole Mediterranean region and supported the evaluation of the species' taxonomic status. Moreover, the combined use of the Control Region and the 16S ribosomal gene strengthened the phylogeographic reconstruction of the species across the Italian mainland and the Island of Sardinia.

Overall, the methodological framework adopted in this thesis underscores a key point: mitochondrial molecular markers, despite certain intrinsic limitations, demonstrated to be powerful tools for reconstructing historical genetic patterns, identifying evolutionary lineages, and conducting preliminary molecular taxonomic assessments. Nevertheless, future research would benefit from the analysis of complete mitochondrial genomes integrated with nuclear markers (e.g., microsatellites, SNPs). Such an approach would provide a more comprehensive understanding of invasion and adaptive processes, fine-scale genetic structuring, and taxonomic delimitation. In this perspective, follow-up investigations are currently underway to expand and refine the findings presented in this thesis. In particular, a whole-genome sequencing study on *Pinna nobilis* is currently ongoing, aiming to provide a higher-resolution assessment of genomic diversity and resilience patterns. Moreover, a new research project on *Salmo letnica* is in progress, incorporating microsatellite markers and the whole mitochondrial genome sequencing to further clarify its unresolved taxonomic status and to achieve a more detailed characterisation of its genetic structure and evolutionary history.

Acknowledgements

My deepest gratitude goes to my supervisors, Prof. Daria Sanna and Prof. Marco Casu, who have been exemplary mentors and have supported me at every stage of my doctoral path.

I would also like to sincerely thank all the other members of my research group. In particular, Prof. Fabio Scarpa, for his invaluable suggestions and guidance, as well as my colleagues and friends Dr. Ilenia Azzena and Dr. Noemi Pascale, for their support, companionship, laughter, and indispensable presence throughout every stage of this journey.

I am especially grateful to Prof. Anila Hoda and Prof. Ioannis Giantsis for their collaboration and invaluable support throughout my three years of PhD.

I also wish to express my gratitude to Dr. Alessio Niffoi for the insightful discussions regarding *Salariopsis fluviatilis* and *Salmo letnica*, as well as Mr. Cristian Sechi and Mr. Cristian Meloni for their assistance with sampling activities and their scientific support concerning trout populations.

I am deeply grateful to my family, who have consistently believed in and supported me, and to all my closest friends, near and far, for their constant presence.

My final thanks go to all the people I have met over these years who have contributed to achieving this long-awaited goal.

Appendix A

Supplementary academic outcomes

A1. Publications in international peer-reviewed journals (2022 – 2026), including 5 as first author:

1. Lostia, G.*, **Locci**, C.*, Rocchigiani, A. M., Cacciotto, C., Fiori, M. S., Azzena, I., Rodriguez-Valera, Y., Pascali, A. M. D., Brandolini, M., Pintus, D., Ligios, C., Scagliarini, A., Casu, M., Coradduzza, E., Scarpa, F., Sanna, D., & Puggioni, G. (2026). Unravelling the Evolutionary Complexity of Orf Virus: A Global and Multi-Host Perspective. *Viruses*, *18*(2).

   *These authors contributed equally to this work.
2. Azzena, I.*, **Locci**, C.*, Pascale, N., Deplano, I., Senigaglia, R., Batistini, E., Caracciolo, D., Chiantore, M., Ciriaco, S., Ferranti, M. P., Grech, D., Liconti, A., Montefalcone, M., Oprandi, A., Pitacco, V., Segarich, M., Zakhama, R., Ben Hmida, A., Zribi, S., Scarpa, F., Casu, M., & Sanna, D. (2025). A century-old genetic time capsule for *Pinna nobilis*: New hints for conservation from ancient, modern, and surviving populations. *Scientific Reports, 15*, 37660.

   *These authors contributed equally to this work.
3. Pitacco, V., Trkov, D., Caracciolo, D., Ciriaco, S., Segarich, M., Batistini, E., Orlando-Bonaca, M., Lipej, L., Mavrič, B., Rogelja, M., Azzena, I., **Locci**, C., Scarpa, F., Sanna, D., Casu, M., Manfrin, C., Pallavicini, A., & Fortič, A. (2025). Recruitment and controlled growth of juveniles of the critically endangered fan mussel *Pinna nobilis* in the Northern Adriatic. *Diversity, 17*, 666.
4. Azzena, I.*, **Locci**, C.*, Pascale, N., Deplano, I., Senigaglia, R., Scarpa, F., Casu, M., & Sanna, D. (2025). *Pinna nobilis*, the vanishing giant: A comprehensive review on the decline of a Mediterranean icon. *Animals, 15*, 2044.

   *These authors contributed equally to this work.

5. Branda, F., Romano, C., Pavia, G., Bilotta, V., **Locci**, C., Azzena, I., Deplano, I., Pascale, N., Perra, M., Giovanetti, M., Ciccozzi, A., De Vito, A., Quirino, A., Marascio, N., Matera, G., Madeddu, G., Casu, M., Sanna, D., Ceccarelli, G., ... Scarpa, F. (2025). Human T-lymphotropic virus (HTLV): Epidemiology, genetics, pathogenesis, and future challenges. *Viruses, 17*, 664.
6. Scarpa, F., Branda, F., Ceccarelli, G., Romano, C., **Locci**, C., Pascale, N., Azzena, I., Fiori, P. L., Casu, M., Pascarella, S., Quaranta, M., Benvenuto, D., Cauda, R., Ciccozzi, M., & Sanna, D. (2025). SARS-CoV-2 XEC: A genome-based survey. *Microorganisms, 13*(2), 253.
7. Hoda, A.*, **Locci**, C.*, Azzena, I., Pascale, N., Deplano, I., Kristo, R., Demiri, A., Scarpa, F., Casu, M., & Sanna, D. (2025). Exploring mitochondrial evolutionary pathways: Insights into the origin of the endemic Ohrid trout. *Life, 15*(1), 52.
   *These authors contributed equally to this work.
8. Pavia, G., Branda, F., Ciccozzi, A., Romano, C., **Locci**, C., Azzena, I., Pascale, N., Marascio, N., Quirino, A., Gigliotti, S., Divenuto, F., Matera, G., Giovanetti, M., Casu, M., Sanna, D., Ceccarelli, G., Ciccozzi, M., & Scarpa, F. (2024). The issue of climate change and the spread of tropical diseases in Europe and Italy: Vector biology, disease transmission, genome-based monitoring, and public health implications. *Infectious Diseases*, 1–16.
9. Pascale, N., Azzena, I., **Locci**, C., Ciccozzi, A., Orrù, F., Are, F., Scarpa, F., Sanna, D., & Casu, M. (2024). New insights into the crayfish *Procambarus clarkii* (Girard, 1852) (Crustacea, Cambaridae): A morphometric approach. *Animals, 14*(24), 3558.
10. Ferranti, M. P., Azzena, I., Batistini, E., Caracciolo, D., Casu, M., Chiantore, M., Ciriaco, S., Firpo, V., Intini, L., **Locci**, C., Montefalcone, M., Oprandi, A., Sanna, D., Scarpa, F., & Segarich, M. (2024). Handling of the bivalve *Pinna nobilis*, an endangered and pathogen-affected species: Precautions taken. *Ecology and Evolution, 14*, e70565.
11. Branda, F., Pavia, G., Ciccozzi, A., Quirino, A., Marascio, N., Gigliotti, S., Matera, G., Romano, C., **Locci**, C., Azzena, I., Pascale, N., Sanna, D., Casu, M., Ceccarelli, G., Ciccozzi, M., & Scarpa, F. (2024). Human papillomavirus (HPV) vaccination:

Progress, challenges, and future directions in global immunization strategies. *Vaccines, 12*, 1293.

12. Scarpa, F., Azzena, I., Ciccozzi, A., Branda, F., **Locci**, C., Perra, M., Pascale, N., Romano, C., Ceccarelli, G., Terrazzano, G., Fiori, P. L., Ciccozzi, M., Casu, M., & Sanna, D. (2024). Update of the genetic variability of monkeypox virus clade IIb lineage B.1. *Microorganisms, 12*, 1874.
13. **Locci**, C., Azzena, I., Pascale, N., Ciccozzi, A., Deplano, I., Giantsis, I. A., Papadopoulos, D. K., Lattos, A., Orrù, F., Puzzi, C. M., Scarpa, F., Casu, M., & Sanna, D. (2024). A sister species for the blue crab *Callinectes sapidus*? A tale revealed by mitochondrial DNA. *Life, 14*, 1116.
14. Quirino, A., Marascio, N., Branda, F., Ciccozzi, A., Romano, C., **Locci**, C., Azzena, I., Pascale, N., Pavia, G., Matera, G., Casu, M., Sanna, D., Giovanetti, M., Ceccarelli, G., Alaimo di Loro, P., Ciccozzi, M., Scarpa, F., & Maruotti, A. (2024). Viral hepatitis: Host immune interaction, pathogenesis, and new therapeutic strategies. *Pathogens, 13*, 766.
15. Pavia, G., Branda, F., Ciccozzi, A., Romano, C., **Locci**, C., Azzena, I., Pascale, N., Marascio, N., Quirino, A., Matera, G., Giovanetti, M., Casu, M., Sanna, D., Ceccarelli, G., Ciccozzi, M., & Scarpa, F. (2024). Integrating digital health solutions with immunization strategies: Improving immunization coverage and monitoring in the post-COVID-19 era. *Vaccines, 12*, 847.
16. Coradduzza, E., Scarpa, F., Rocchigiani, A. M., Cacciotto, C., Lostia, G., Fiori, M. S., Rodríguez-Valera, Y., De Pascali, A. M., Brandolini, M., Azzena, I., **Locci**, C., Casu, M., Bechere, R., Pintus, D., Ligios, C., Scagliarini, A., Sanna, D., & Puggioni, G. (2024). The global evolutionary history of Orf virus in sheep and goats revealed by whole-genome data. *Viruses, 16*, 158.
17. Sanna, D., Azzena, I., **Locci**, C., Ankon, P., Kružić, P., Manfrin, C., Pallavicini, A., Ciriaco, S., Segarich, M., Batistini, E., Scarpa, F., & Casu, M. (2024). Reconstructing the evolutionary history of *Pinna nobilis*: New genetic signals from the past of a species on the brink of extinction. *Animals, 14*, 114.

18. Scarpa, F., Pascarella, S., Ciccozzi, A., Giovanetti, M., Azzena, I., **Locci**, C., Casu, M., Fiori, P. L., Quaranta, M., Cella, E., Sanna, D., & Ciccozzi, M. (2023). Genetic and structural analyses reveal the low potential of the SARS-CoV-2 EG.5 variant. *Journal of Medical Virology, 95*, e29075.
19. Scarpa, F., **Locci**, C., Azzena, I., Casu, M., Fiori, P. L., Ciccozzi, A., Giovanetti, M., Quaranta, M., Ceccarelli, G., Pascarella, S., Ciccozzi, M., & Sanna, D. (2023). SARS-CoV-2 recombinants: Genomic comparison between XBF and its parental lineages. *Microorganisms, 11*, 1824.
20. Scarpa, F., Giovanetti, M., Azzena, I., **Locci**, C., Casu, M., Fiori, P. L., Ciccozzi, A., Imperia, E., Bazzani, L., Borsetti, A., Maruotti, A., Pascarella, S., Sanna, D., & Ciccozzi, M. (2023). Genome-based survey of the SARS-CoV-2 BF.7 variant from Asia. *Journal of Medical Virology, 95*(4), e28714.
21. Scarpa, F., Azzena, I., **Locci**, C., Casu, M., Fiori, P. L., Ciccozzi, A., Angeletti, S., Imperia, E., Giovanetti, M., Maruotti, A., Borsetti, A., Cauda, R., Cassone, A., Via, A., Pascarella, S., Sanna, D., & Ciccozzi, M. (2023). Molecular in-depth on the epidemiological expansion of SARS-CoV-2 XBB.1.5. *Microorganisms, 11*(4), 912.
22. Scarpa, F., Imperia, E., Azzena, I., Giovanetti, M., Benvenuto, D., **Locci**, C., Casu, M., Fiori, P. L., Maruotti, A., Ceccarelli, G., Borsetti, A., Caruso, A., Cauda, R., Cassone, A., Via, A., Pascarella, S., Sanna, D., & Ciccozzi, M. (2023). Genetic and structural genome-based survey reveals the low potential for epidemiological expansion of the SARS-CoV-2 XBB.1.5 sublineage. *Journal of Infection, 86*(6), 596–598.
23. Scarpa, F., Azzena, I., Ciccozzi, A., Giovanetti, M., **Locci**, C., Casu, M., Fiori, P. L., Borsetti, A., Cella, E., Quaranta, M., Pascarella, S., Sanna, D., & Ciccozzi, M. (2023). Integrative genome-based survey of the SARS-CoV-2 Omicron XBB.1.16 variant. *International Journal of Molecular Sciences, 24*, 13573.
24. Casu, M., Azzena, I., Scarpa, F., **Locci**, C., Niffoi, A., Delmastro, G. B., Lo Conte, P., Varcasia, A., Bovero, S., Candiotto, A., Sanna, D., & Cossu, P. (2023). Appraising the genetic makeup of an allochthonous southern pike population: An opportunity to

predict the evolution of introgressive hybridization in isolated populations? *Animals, 13*, 380.

25. Azzena, I., Scarpa, F., **Locci**, C., Cossu, P., Niffoi, A., Orrù, F., Bovero, S., Sotgiu, G., Sanna, D., & Casu, M. (2022). 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, 12*, 3403.
26. Scarpa, F., Sanna, D., Azzena, I., Cossu, P., **Locci**, C., Angeletti, S., Maruotti, A., Ceccarelli, G., Casu, M., Fiori, P. L., Petrosillo, N., & Ciccozzi, M. (2022). Genetic variability of the monkeypox virus clade IIb B.1. *Journal of Clinical Medicine, 11*(21), 6388.

A2. Orf virus (ORFV)

Orf, also known as contagious ecthyma, is a highly contagious viral skin disease with a global distribution (Nandi et al., 2011), caused by Orf virus (ORFV), a member of the genus *Parapoxvirus* within the family Poxviridae (Lawan et al., 2021). ORFV is most frequently associated with skin disease in sheep and goats (Spyrou & Valiakos, 2015), however, infections have also been documented in other ruminant and non-ruminant species (Kumar et al., 2015), including humans (Bala et al., 2018). From evolutionary and epidemiological perspectives, ORFV has been widely investigated, with the viral interferon resistance (VIR) gene emerging as one of the most used molecular markers for this virus (e.g., Peralta et al., 2018; Coradduzza et al., 2021, 2024; Li et al., 2023).

In light of this, population genetic approaches were applied in this study to investigate viral genetic diversity, population structure, host-specific transmission potential, and phylodynamic patterns, using an expanded dataset based on VIR gene, including natural, occasional, and human hosts analysed at a global scale.

The main results indicate that ORFV comprises three major genetic lineages: two host-specific groups (one primarily in goats and one in sheep), and a third generalist group capable of infecting several ruminant species. Moreover, the virus can occasionally infect humans, but only under specific circumstances, such as rare viral variants, recombination events, human susceptibility, and intense exposure to infected animals. Thus, human infection appears to be a sporadic spillover event rather than a zoonosis as a consequence of viral adaptation to human hosts.

References

Bala, J. A., Balakrishnan, K. N., Abdullah, A. A., Mohamed, R., Haron, A. W., Jesse, F. F. A., Noordin, M. M., & Mohd-Azmi, M. L. (2018). The re-emergence of orf virus infection: A call for surveillance, vaccination, and effective control measures. *Microbial Pathogenesis, 120*, 55–63. https://doi.org/10.1016/j.micpath.2018.04.057

Coradduzza, E., Sanna, D., Rocchigiani, A. M., Pintus, D., Scarpa, F., Scivoli, R., Bechere, R., Dettori, M.A., Montesu, M.A., Marras, V., Lobrano, R., Ligios, C., & Puggioni, G. (2021). Molecular insights into the genetic variability of ORF Virus in a Mediterranean region (Sardinia, Italy). *Life*, *11*(5), 416. https://doi.org/10.3390/life11050416

Coradduzza, E., Scarpa, F., Rocchigiani, A. M., Cacciotto, C., Lostia, G., Fiori, M. S., Rodriguez Valera, Y., De Pascali, A. M., Brandolini, M., Azzena, I., Locci, C., Casu, M., Bechere, R., Pintus, D., Ligios, C., Scagliarini, A., Sanna, D., & Puggioni, G. (2024). The global evolutionary history of Orf virus in sheep and goats revealed by whole genomes data. *Viruses*, *16*(1), 158. https://doi.org/10.3390/v16010158

Kumar, R., Trivedi, R. N., Bhatt, P., Khan, S. H., Khurana, S. K., Tiwari, R., Karthik, K., Malik, Y. S., Dhama, K., & Chandra, R. (2015). Contagious pustular dermatitis (orf disease): Epidemiology, diagnosis, control, and public health concerns. *Advances in Animal and Veterinary Sciences, 3*, 649–676. http://dx.doi.org/10.14737/journal.aavs/2015/3.12.649.676

Lawan, Z., Bala, J. A., Bukar, A. M., Balakrishnan, K. N., Mangga, H. K., Abdullah, F. F. J., Noordin, M. M., & Mohd-Azmi, M. L. (2021). Contagious ecthyma: How serious is the disease worldwide? *Animal Health Research Reviews, 22,* 40–55. https://doi.org/10.1017/S1466252320000018

Li, S., Jing, T., Zhu, F., Chen, Y., Yao, X., Tang, X., Zuo, C., Liu, M., Xie, Y., Jiang, Y., Wang, Y., Li, D., Li, L., Gao., S., Chen, D., Zhao, H., & Ma, W. (2023). Genetic analysis of Orf virus (ORFV) strains isolated from goats in China: insights into epidemiological characteristics and evolutionary patterns. *Virus Research*, *334*, 199160. https://doi.org/10.1016/j.virusres.2023.199160

Nandi, S., De, U. K., & Chowdhury, S. (2011). Current status of contagious ecthyma or orf disease in goats and sheep: A global perspective. *Small Ruminant Research, 96*, 73–82. https://doi.org/10.1016/j.smallrumres.2010.11.018

Peralta, A., Robles, C. A., Micheluod, J. F., Rossanigo, C. E., Martinez, A., Carosio, A., & König, G. A. (2018). Phylogenetic analysis of Orf viruses from five contagious ecthyma outbreaks in Argentinian goats. *Frontiers in Veterinary Science, 5*, 134. https://doi.org/10.3389/fvets.2018.00134

Spyrou, V., & Valiakos, G. (2015). Orf virus infection in sheep or goats. *Veterinary Microbiology, 181*, 178–182. https://doi.org/10.1016/j.vetmic.2015.08.010

Research article

viruses

MDPI

Article

Unravelling the Evolutionary Complexity of Orf Virus: A Global and Multi-Host Perspective

Giada Lostia [1,2,†], Chiara Locci [2,3,†], Angela Maria Rocchigiani [1], Carla Cacciotto [2,4], Mariangela Stefania Fiori [1], Ilenia Azzena [2], Yoel Rodriguez-Valera [5], Alessandra Mistral De Pascali [6], Martina Brandolini [6], Davide Pintus [1], Ciriaco Ligios [1], Alessandra Scagliarini [6], Marco Casu [2], Elisabetta Coradduzza [1,*], Fabio Scarpa [3], Daria Sanna [3,*,‡] and Giantonella Puggioni [1,‡]

[1] Istituto Zooprofilattico Sperimentale della Sardegna, 07100 Sassari, Italy; giada.lostia@izs-sardegna.it (G.L.); angelamaria.rocchigiani@izs-sardegna.it (A.M.R.); mariangela.fiori@izs-sardegna.it (M.S.F.); davide.pintus@izs-sardegna.it (D.P.); ciriaco.ligios@izs-sardegna.it (C.L.); giantonella.puggioni@izs-sardegna.it (G.P.)
[2] Dipartimento di Medicina Veterinaria, University of Sassari, 07100 Sassari, Italy; c.locci3@phd.uniss.it (C.L.); ccacciotto@uniss.it (C.C.); iazzena@uniss.it (I.A.); marcasu@uniss.it (M.C.)
[3] Dipartimento di Scienze Biomediche, University of Sassari, 07100 Sassari, Italy; fscarpa@uniss.it
[4] Mediterranean Center for Disease Control, University of Sassari, 07100 Sassari, Italy
[5] Faculty of Agricultural Sciences, University of Granma, Bayamo 95100, Cuba; yrodriguezvalera@udg.co.cu
[6] Dipartimento di Scienze Mediche e Chirurgiche, University of Bologna, 40138 Bologna, Italy; alessandra.depascal3@unibo.it (A.M.D.P.); martina.brandolini3@unibo.it (M.B.); alessand.scagliarini@unibo.it (A.S.)
* Correspondence: elisabetta.coradduzza@izs-sardegna.it (E.C.); darsanna@uniss.it (D.S.)
† These authors contributed equally to this work as first authors.
‡ These authors contributed equally to this work as last authors.

Abstract

Orf virus (ORFV), a member of the *Parapoxvirus* genus, is commonly associated with a highly infectious skin disease primarily affecting sheep and goats, with a reported zoonotic potential. Initially identified in the 18th century, ORFV has been sporadically reported in other species, including humans. The present study analyzed the genetic variability and phylodynamic patterns of ORFV using the highly variable VIR gene, focusing on global strains from multiple hosts, including various species of ruminants and humans. A dataset of 267 ORFV strains from around the world, including sequences from the understudied island of Cuba, was used for the analyses. Results revealed greater genetic variability for ORFV than previously reported. While the virus may be defined as a generalist pathogen, capable of infecting various ruminant species and less frequently humans, signs of host-specific specialization are emerging exclusively for sheep and goats. Other ruminant species and humans may be categorized as occasional hosts, with infections likely linked to habitat overlap with sheep and goats and sporadic transmission that appears influenced by specific risk factors. In conclusion, these findings contribute to a better understanding of the transmission risks posed by ORFV, highlighting the need for further investigations into its potential to infect a broader range of hosts, particularly humans.

Keywords: ORFV; phylodynamics; VIR gene; accidental hosts; *Parapoxvirus*

Academic Editor: Yoon-Seok Chung

Received: 7 January 2026
Revised: 1 February 2026
Accepted: 6 February 2026
Published: 10 February 2026

1. Introduction

Orf, also referred to as contagious pustular dermatitis (CPD) or contagious ecthyma (CE), is a highly infectious viral skin disease [1] with a worldwide distribution [2]. Initially identified in sheep by Steeb in 1787 and later observed in goats and humans by Hansen in

Viruses **2026**, *18*, 222 https://doi.org/10.3390/v18020222

1879 [3], the disease is caused by the Orf virus (ORFV), a member of the *Parapoxvirus* genus, within the Chordopoxviridae subfamily of the Poxviridae family [4].

In this context, contagious ecthyma has been reported worldwide, with documented cases across Europe (Greece, Norway, Italy, England, Romania, Bulgaria, Hungary), North America (USA), South America (Brazil, Argentina), Asia (India, Taiwan, China, Israel, Japan, Iran, Iraq, Malaysia, Korea, Turkey, Lao, Bangladesh, Mongolia, Saudi Arabia, Pakistan), Africa (Nigeria, Kenya, Sudan, Zambia, Gabon, South Africa, Cameroon, Egypt, Ethiopia, Morocco), and Australia [4–6].

Outbreaks of CPD most frequently occur between autumn and spring months [7], with the highest severity observed during the autumn and winter months [8–10], particularly in December and January in the Northern Hemisphere, as reported by Robinson et al. (1983) [7]. However, sporadic cases have also been documented during the summer months [11–13].

Although ORFV is commonly associated with a skin disease affecting sheep and goats [14], it has also been reported sporadically in other ruminants and non-ruminant species [15]. With regard to other ruminants' species, several studies have reported ORFV in members of the Bovidae family, including domestic cattle [16], free-ranging and captive musk ox [17–20], captive Sichuan takin [19], and wild chamois [21]. Cases have also been documented in the Antilocapridae family, such as in wild Japanese serow [22], and wild ibex [23], as well as in the Cervidae family, specifically in reindeer [24–26].

With regard to non-ruminant species, ORFV has been reported in member of the Camelidae family, including domestic dromedaries [27–30], cats [31], and humans [5,32,33]. Finally, Cargnelutti et al. [34] reported that mice and rabbits became susceptible to ORFV when inoculated with the IA-82 ORFV strain via skin scarification at various anatomical sites under experimental conditions. Notably, the infected rabbits developed skin lesions that closely resembled those observed in both naturally and experimentally infected sheep [34].

It is important to highlight that ORFV is considered to have a zoonotic potential in humans [35]. The transmission usually occurs through direct contact (through abrasion or breaks in the skin) with infected animals, or through contaminated fomites [7,36,37]. Individuals who work closely with animals (e.g., shepherds, farmers, and veterinarians) are considered to be at higher risk of infection [7]. Additionally, a case report has described ORFV transmission to a 53-year-old woman with psoriasis, living in a farming community, after she was scratched by a stray cat [38]. This extensive host range contributes to the virus's endemicity in areas where animals are abundant, suggesting a persistent opportunity for dynamic interaction between the virus and its hosts [39].

The virus causes persistent and localized skin lesions, that typically last between three and four weeks [40]. It is also capable of reinfection, due to its ability to evade the host's immune response [41]. In animals, lesions commonly develop on the lips, muzzle, ears, eyelids, and the nasal mucosa, and, less frequently, on the mammary glands, genitals, and feet [15,42]. In humans, lesions are often located on exposed areas of the skin, such as the hands, fingers, and forearms, which are most likely to come into contact with animals [43,44].

The virus primarily affects young animals, including lambs and kids, although adult animals can also be infected [41]. The disease has been reported to occur irrespective of age, sex, or breed [15]. During ORFV outbreaks, the morbidity rate is generally higher than the mortality rate [35]. However, young animals (e.g., lambs), which are more susceptible to the disease [15], experience higher mortality rate compared to adults [41]. The disease usually results in death due to the presence of painful lesions around the mouth that prevent

https://doi.org/10.3390/v18020222

suckling, maternal rejection because of severe udder lesions, or complications arising from secondary bacterial or fungal infections and maggot infestation [41].

Research efforts have focused on exploring the genetic diversity and evolutionary history of ORF virus and providing insights into the potential geographic clustering of circulating strains [45].

Traditionally, the largest number of studies have focused on variation at single-gene loci (e.g., [46–48]), with particular attention on the gene VIR (encoding for viral interferon resistant). This gene has widely been investigated to infer the phylogeography and evolutionary dynamics of ORFV (e.g., [48–57]). VIR is considered a major virulence factor of ORFV and represents a variable genomic region within the *Parapoxvirus* genus [51,56]. Although other ORFV genes (e.g., orf109, orf110, orf132) show higher levels of nucleotide variability [46,47,58], they are comparatively less represented in the literature, limiting the use of their sequences in large-scale phylogenetic and comparative analyses. Within this framework, phylogeographic studies based on the analysis of VIR variation reported that it is highly informative for genetic characterization within *Parapoxvirus*. Notably, VIR has proven to be an efficient indicator of genetic variability and population structure in ORFV [48,50,52,55–57]. Several studies across this locus have indicated the potential existence of host-associated specific strains for sheep and goats [50,57,59]. However, it should be taken into consideration that this gene was reported to exhibit conservation in amino acid residues, and this pattern is not associated with phylogenetic groupings based on host species, geographical origin, or time of isolation [52].

In addition to VIR, other ORFV genes, such as B2L (gene encoding for the immunogenic major envelope protein) and O45 (gene encoding for the late transcription factor VTLF-1), have been employed in genetic analyses aimed at clarifying the evolutionary history of this virus. Although results obtained from these loci are generally consistent with those based on VIR, their corresponding indices of genetic variation and informativeness are generally lower (e.g., [50,55–57,59]).

More recently, research has moved towards genomic-scale analysis. In particular, based on the limited number of whole genomes available to date [46,49,50,60–64], Coradduzza et al. [50] reported that modern sheep and goat ORFV lineages diverged from one another in the second half of the 20th century, originating from two geographically distinct clusters. Their study also hypothesized the possible existence of additional viral strains capable of infecting either goats or sheep.

In the present study, the ORFV molecular marker for which the largest number of sequences is currently available in the literature was analyzed. For this reason, the highly variable VIR gene was employed to analyze the genetic diversity of ORFV on a global scale, across natural, occasional or accidental, and human hosts, using the most comprehensive approach applied to date. Specifically, VIR gene has been used to investigate in detail the phylodynamic patterns of the virus and to explore early signs of host-specific adaptation in sheep and goats, based on an expanded dataset that includes understudied geographical areas and strains isolated from various host species. Furthermore, 31 sheep and goat samples collected from an under-investigated area of the island of Cuba were combined with a dataset comprising VIR sequences available in GenBank, contributing to a more comprehensive assessment of the global diversity and evolutionary dynamics of ORFV.

2. Materials and Methods

2.1. Sampling

The Cuban new pathological samples in this study were collected between 2005 and 2009 (as summarized in Table 1) from sheep and goats naturally infected by the disease during various outbreaks in the municipalities of Holguin, Granma, Santiago De Cuba, and Guantanamo of the Eastern Provinces of Cuba (Supplementary Figure S1).

Table 1. Table providing the information on the 31 Cuban strains analyzed and their related GenBank accession numbers (GB#).

Sample ID	GB#	Collection Site	Collection Date	Host
S1	OR338563	Santiago de Cuba	21 March 2006	sheep
S2	OR338564	Santiago de Cuba	22 March 2006	sheep
S3	OR338565	Guantanamo	7 March 2006	sheep
S8	OR338566	Guantanamo	22 June 2006	sheep
S46	OR338567	Guantanamo	17 April 2008	goat
S48	OR338568	Guantanamo	29 December 2008	sheep
S18	OR338569	Granma	15 June 2006	goat
S21	OR338570	Guantanamo	18 July 2006	goat
S23	OR338571	Guantanamo	25 September 2006	sheep
S37	OR338572	Guantanamo	17 April 2008	goat
S40	OR338573	Guantanamo	25 April 2008	goat
S42	OR338574	Guantanamo	28 July 2008	sheep
S835-3	OR338575	Granma	3 June 2005	goat
S28	OR338576	Holguin	14 September 2005	sheep
S30	OR338577	Granma	4 June 2006	goat
S5	OR338578	Guantanamo	5 March 2006	sheep
S44	OR338579	Granma	25 December 2008	sheep
S51	OR338580	Holguin	5 June 2009	sheep
S10	OR338581	Guantanamo	23 June 2006	goat
S15	OR338582	Granma	4 June 2006	sheep
S31	OR338583	Guantanamo	9 June 2007	sheep
S33	OR338584	Guantanamo	10 November 2007	goat
S34	OR338585	Guantanamo	1 February 2008	goat
S36	OR338586	Guantanamo	14 January 2008	sheep
S38	OR338587	Guantanamo	13 November 2008	goat
S39	OR338588	Guantanamo	10 October 2008	sheep
S41	OR338589	Guantanamo	29 December 2008	goat
S43	OR338590	Guantanamo	17 January 2009	sheep
S45	OR338591	Guantanamo	10 October 2008	sheep
S50	OR338592	Holguin	5 June 2009	sheep
S49	OR338593	Santiago de Cuba	10 April 2009	sheep

To obtain a comprehensive dataset representing all viral sequences available worldwide from several host species, the ORF sequences obtained from these Cuban samples (refer to Table 1 for accession number) were merged with available VIR gene sequences deposited in GenBank (see Supplementary Table S1 for details). The resulting dataset included a total of 267 strains from 21 countries and 5 continents, belonging to 6 different hosts, with the exception of 20 strains for which the host species of origin was not indicated (14) or specified accurately (6 referred as "small ruminants").

2.2. Viral DNA Extraction

To confirm the presence of ORFV, DNA was extracted from 25 mg of each lesion tissue sample using the DNeasy Blood and Tissue Kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions. Then, VIR was subjected to amplification using the VIR1 and VIR2 primers [19]. This process provided the generation of fragments measuring 617 bp. To amplify VIR from samples negative to the first PCR, an additional primer set (VIR 3 and VIR 4) was designed, which amplifies a fragment of approximately 817 bp. The PCR was performed as previously described by Kottaridi et al. [52] with slight changes in annealing temperatures [52]. Positive (previously characterized high quality ORFV DNA samples from reference collection) and negative controls were used to evaluate the efficacy of PCR protocols and to ascertain the absence of any potential inhibitors or contaminants. Electrophoreses were carried out with Invitrogen E-Gel EX 2% agarose kit (Carlsbad, CA, USA) (gel precast). Specific DNA PCR bands were excised from the gel and purified using the QIAquick Gel Extraction Kit (Qiagen). When non-specific bands were also present during electrophoresis, PCR products were purified using the CleanSweep PCR Purification kit (Thermo Fisher Scientific, Waltham, MA, USA).

2.3. Sequencing

All reliable PCR products were then subjected to sequencing for both the forward and reverse strands, using the same primers utilized for the PCR. In detail, the viral DNA was quantified using an Epoch microplate spectrophotometer (BioTek, Winooski, VT, USA) and a Qubit 2.0 Fluorometer (Thermo Fisher Scientific) according to the manufacturers' instructions. DNA viral libraries were then prepared using the Nextera DNA Flex Library Prep kit (Illumina Inc., San Diego, CA, USA). The following whole-genome sequencing was performed by BMR Genomics (an external core service; BMR Genomics, Padua, Italy) using the whole-genome shotgun technique (WGS) on the MiSeq platform (Illumina) to generate paired-end reads of 2 × 300 bp.

2.4. Phylogenetic Analysis

A total of 31 VIR gene sequences were obtained in the present study.

Noteworthy, to maximize the dataset, we utilized a 382 bp VIR gene fragment, enabling the inclusion of all available GenBank sequences regardless for their size. Retaining a comprehensive library representative of all host species was critical for ensuring the reliability of results.

After checking sequences using Unipro UGENE 35 [65], they were aligned with the Clustal Omega 1.2.4 software package [66] available at https://www.ebi.ac.uk/Tools/msa/clustalo/ [67] (accessed on 24 August 2025). The sequences have been submitted to the GenBank online database, and their accession numbers are provided in Supplementary Table S1.

To place our data within a broader geographic framework, we built a dataset that included the 31 VIR gene sequences obtained from Cuban samples along with the publicly available ORF VIR gene sequences deposited in GenBank. This extensive dataset encompassed sequences from across all continents, comprising Asia (145), Europe (57), South America (21), North America (41), Oceania (2), and Africa (1). Only sequences that perfectly overlapped with the VIR gene fragment analyzed in this study were included, resulting in a final dataset of 267 sequences used for further analyses. Additionally, a Pseudocowpox virus sequence (GQ329670) was used as outgroup.

Notably, a total of 252 ORF VIR sequences out of the 267 available in the dataset could be assigned to specific host species. In particular, 247 have been isolated from ruminants (136 sequences belonged to goats, 93 to sheep, 10 to Japanese goat-antelopes, 6 to small ruminants, 1 to Sichuan takin, 1 to musk ox), and 5 from humans.

To evaluate genetic variation across the whole dataset and within and among host-associated populations, the number of polymorphic sites (S), number of haplotypes (H), nucleotide diversity (π), and haplotype diversity (h), have been estimated by means of the software DnaSP 6.12.03 [68].

The best probabilistic model of sequence evolution was identified using jModeltest 2.1.1 [69], applying a maximum likelihood optimized search by the Akaike (AIC) and

Bayesian Information Criterion (BIC). Both criteria consistently identified GTR + I + G as the best-fitting model.

To assess the suitability of the VIR gene dataset for phylogenetic and phylodynamic analyses, the phylogenetic signal [70] was evaluated through a likelihood mapping of 10,000 randomly selected quartets, conducted using TreePuzzle 5.3 [71].

A Bayesian phylogenetic tree was constructed using MrBayes 3.2.7 [72], applying the following model parameters: setting nst = 3, rates = gamma, and ngammacat = 4. Two separate runs were performed simultaneously, each involving 4 Metropolis-coupled Markov chain Monte Carlo (MCMC) chains (1 cold and 3 heated). The MCMC process was run for 5 million generations, with trees sampled every1000 generations. The initial 25% of sampled trees were excluded as burn-in. Analyses were carried out through the CIPRES Science Gateway platform [73]. To confirm convergence of chains, the average standard deviation of split frequencies (expected to approach) [66] and the potential scale reduction factor (expected to be approximately close to 1) [74], have been assessed following the guidelines described by Scarpa et al. [75]. The resulting tree was then visualized and edited using FigTree 1.4.0 [76] (available at http://tree.bio.ed.ac.uk/software/figtree/, accessed on 22 December 2025).

Molecular dating was conducted using a Bayesian approach under the MCMC algorithm implemented in Beast 1.10.4 software [77], incorporating the sampling dates of the sequences. To determine the best molecular clock model, both strict and uncorrelated log-normal relaxed clocks were evaluated through preliminary runs of 100 million generations. Model selection was guided by the comparison of Bayes factor values using Tracer 1.7 [78]. All available demographic models, both parametric and non-parametric, were also evaluated to determine the best fit for the sequence data. Following model selection, time-calibrated phylogenetic trees and evolutionary rates were co-estimated for the dataset by running the analysis for 300 million generations, sampling every 30,000 generations. The output log files were assessed in Tracer 1.7 software [79], retaining only those with ESS (effective sample size) values above 200. The maximum clade credibility tree was then generated using TreeAnnotator (Beast package) and subsequently visualized and annotated in FigTree 1.4.4.

Finally, Principal Coordinate Analysis (PCoA) was performed in GenAlEx 6.5 [80] to explore genetic variation and identify potential subgrouping within clusters. The analysis relied on a pairwise p-distance matrix to represent genetic dissimilarity among sequences. PCoA was first conducted on the full dataset and then repeated on subsets identified from patterns observed in the initial run.

3. Results

Phylogenetic and Phylodynamic Analyses

Genetic variation analysis conducted on the whole dataset of 267 sequences identified 41 polymorphic sites, resulting in 44 VIR haplotypes (h = 0.924; π = 0.0423). Overall, the dataset showed high levels of genetic diversity. A subset of 252 sequences, including only strains to which it has been possible to assign a specific host species, was used to evaluate genetic diversity within and among host-associated strains (see Supplementary Table S2: 136 goats, 93 sheep, 10 Japanese goat-antelopes, 6 small ruminants, 1 Sichuan takin, 1 musk ox, and 5 humans).

Overall, for this subset of 252 sequences, genetic diversity remained high (h = 0.932; π = 0.0418) with 43 haplotypes, thus reflecting the high genetic variability present among the host-associated lineages.

Within this subset, the 247 ruminant-associated sequences (excluding the 5 human isolates) also exhibited high overall genetic diversity, with 43 haplotypes detected (h = 0.931;

π = 0.0419). In particular, goats and sheep showed the highest levels of genetic diversity. Goats displayed 70 haplotypes among 136 sequences (h = 0.978; π = 0.0361), while sheep showed 54 haplotypes among 93 sequences (h = 0.971; π = 0.0366), both indicating high within-host genetic variation. On the other hand, the occasional hosts, which included Japanese-goat antelope, small ruminants, musk ox, Sichuan takin, and human ORFV strains, exhibited lower levels of genetic diversity (see Supplementary Table S2 for details on host species groups), with 10 haplotypes among 23 sequences (h: 0.783; π = 0.0266). Within the occasional hosts group, the 5 human-associated sequences were particularly differentiated among each other: each of the 5 human isolates presented distinct private haplotypes/strains (h = 1), with nucleotide diversity values (π = 0.0377) comparable to that observed for sheep and goat lineages. Notably, polymorphic sites characterizing human lineages included both neutral and missense mutations, and all the resulting amino acid changes were also found for ruminant-associated polymorphic strains.

The likelihood-mapping analysis performed on the complete dataset of 267 sequences (Supplementary Figure S2) revealed a robust phylogenetic signal, with 2.8% of unresolved quartets, significantly under the 30% threshold commonly indicative of poor phylogenetic resolution [69].

The Bayesian phylogenetic tree, also conducted on the whole dataset of 267 sequences, was consistent with the topology of the ultrametric tree. Based on this concordance, a graphical representation of the Bayesian tree, incorporating divergence times at major nodes, was generated (see in Appendix A, Figure A1). Phylogenetic reconstruction revealed the presence of two principal genetic clusters, designated as A and B, strongly supported at their basal node (*pp* = 1.0). These clusters diverged from a common ancestor which dated back to approximately 1943 (81 years before 2024, the year in which the most recent strain was collected) and appeared to have emerged contemporaneously, around 1948 (76 years before 2024). Cluster A included the majority of sequences in the dataset, comprising viral strains isolated from sheep, goats, humans, and small ruminants, collected worldwide between 1995 and 2024. Several sub-clusters were detected; however, most of these subdivisions were weakly supported (*pp* < 0.90), suggesting limited phylogenetic resolution. A few sub-groups exceeded the support threshold (*pp* > 0.90), but did not show structuring by geography, collection date, or host species. Cluster B contained fewer sequences, primarily from goats and sheep, along with a small number from humans, Japanese goat-antelopes, a musk ox, and a Sichuan takin. These isolates were collected between 1970 and 2023, mostly from North America (including most of the Cuban samples obtained in the present study), South America, Asia, and Europe. As in Cluster A, sub-structuring was observed but generally poorly supported (*pp* < 0.90).

A second Bayesian phylogenetic tree with divergence times estimates was constructed, this time including only sheep and goat isolates (n = 229) to investigate potential occurrence of genetic structuring between these two host-associated viral populations (Figure 1). The most recent common ancestor, fully supported at the main node (*pp* = 1.0), was estimated to date back to approximately 1949 (75 years before 2024). Notably, this estimation is consistent with the root divergence time of the first tree (Figure A1) including natural and accidental hosts.

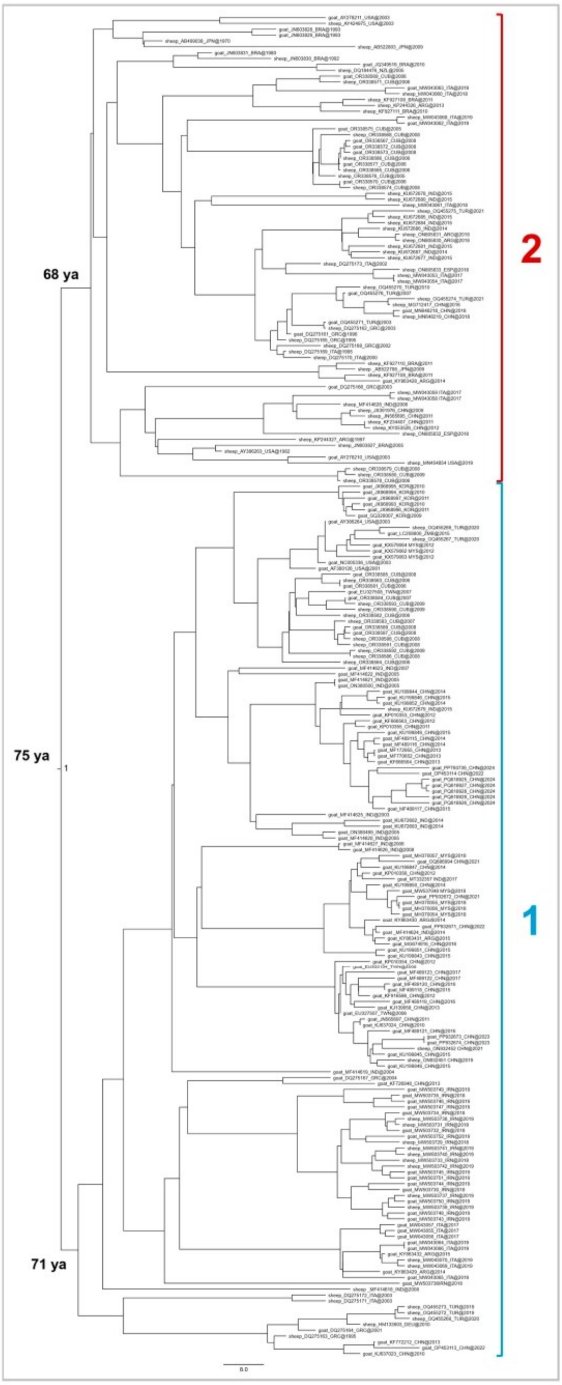


Figure 1. Bayesian phylogenetic tree of the sheep and goat subset with divergence time estimates at major nodes obtained by the software Beast 1.10.4. Group subdivisions are described in the text. Posterior probability values are indicated at the basal node (*pp* = 1.0).

https://doi.org/10.3390/v18020222

From the most recent common ancestor departed two main clusters, designated as 1 and 2, which emerged almost contemporaneously. Cluster 1 diverged around 1953 (71 years before 2024) and included approximately 85% of all goat-derived sequences, mainly from Asia, with a few isolates from Europe, North and South America, and Africa, collected between 2001 and 2024. The few sheep-derived strains included within the cluster 1 accounted for 37% of all sheep-derived sequences, originating predominantly from Asia, North America and Europe, and collected between 1995 and 2021. A sub-structuring was observed but generally not strongly supported at the main nodes ($pp < 0.90$), except for one sub-group ($pp = 0.97$) containing primarily Asian goat isolates.

Cluster 2 diverged around 1956 (68 years before 2024) and comprised about 63% of all sheep-derived strains, mostly from North America and Asia, with additional sequences from Europe, South America and Oceania, collected between 1970 and 2021. The few goat-derived isolates included within the cluster 2 represented 15% of all goat-derived sequences, originating from North America and Europe, with a few from South America and Asia, and isolated between 1980 and 2019. Similar to cluster 1, sub-clusters were present but generally lacked full statistical support ($pp < 0.90$).

Principal Coordinates Analysis (PCoA) performed on the whole dataset of 267 sequences (thus excluding the outgroup) explained a cumulative variability of 75.40 (Axis 1: 43.23%, Axis 2: 20.87%, Axis 3: 11.30%, see Supplementary Figure S3 for details). In accordance with first phylogenetic tree (Figure A1), this analysis, which included various host groups, revealed low genetic structuring among most of the sequences. A central cluster showed substantial overlap, primarily comprising sequences from sheep, goats, and to a lesser extent, other ruminant species, and humans. At the margins of this main cluster, an incipient divergence was observed in a few sequences, mostly belonging to sheep and goats. One divergent strain from sheep originated from Spain (ON805832) and was isolated in 2018, while the divergent goat sequences were from China (KF916586, KU199850, MF489117, MF489118, MF489120, MF489122) and were isolated between 2012 and 2017.

Based on these findings, a second Principal Coordinates Analysis (PCoA) was conducted on a subset, including only sequences from sheep and goats, totaling 229 sequences (Figure 2). This analysis aimed to investigate the genetic similarity between ORFV strains from these two host species. The PCoA explained a total genetic variability of 78.02% (Axis 1: 43.92%, Axis 2: 21.89%, Axis 3: 12.21%). Similar to the first PCoA (see Supplementary Figure S3), a homogeneous central cluster of overlapping for goats and sheep strains was observed, indicating low genetic differentiation among lineages. However, a slightly more pronounced divergence was evident between sequences located at the opposite x-axis extremes of the central overlapping cluster. Accordingly, a few Chinese goat sequences showed strong divergence from sheep sequences found at the opposing margins of the central overlapping area. These outlying sheep sequences originated from Argentina, India, Cuba, Turkey, Spain, and Italy.

An additional Principal Coordinates Analysis (PCoA) was performed based on sequences from the occasional hosts, including a total of 23 isolates (Figure 3). This analysis aimed to investigate the genetic diversity and relationships among occasional hosts, comprising the Japanese goat-antelope, Sichuan-takin, musk ox, small ruminants, and human. This analysis explained a comprehensive genetic variation of 96.33% (Axis 1: 47.09%, Axis 2: 34.89%, Axis 3: 14.35%). Three main clusters, designated as C1, C2 and C3, were identified. Cluster C1 comprised the majority of sequences, including all Japanese goat-antelope isolates from Japan collected between 1985 and 2009, as well as those from Sichuan takin and musk ox, both originating from the United States and collected in 2003. This cluster also included a human-isolated strain from Hungary (OR372162—2023). Cluster C2 comprised three human-derived isolates from Germany (KF837136—1996), Italy (MW043067

–2019), and Argentina (MH161456—2015). Cluster C3 included small ruminant sequences from Turkey collected in 2020. However, one Turkish sequence (MW492052) and a human sequence from Chile (MH161457—2017) behaved as outliers, with respect to Clusters C1, C2 and C3.

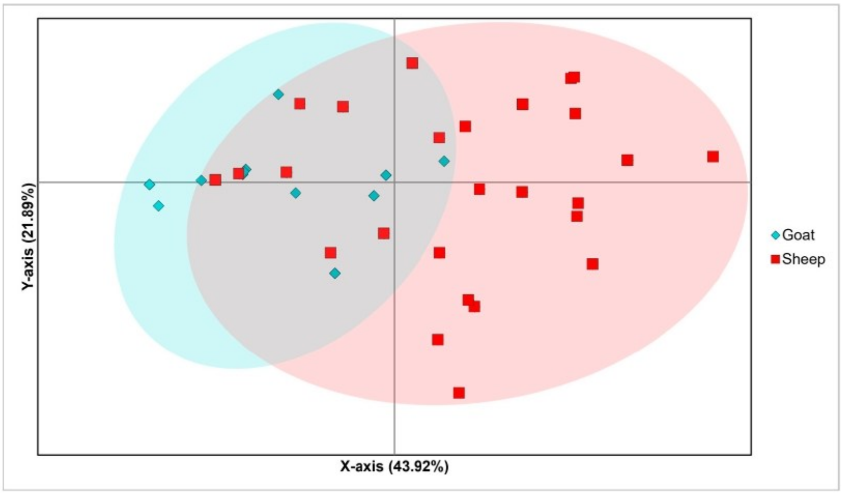


Figure 2. Principal Coordinates Analysis (PCoA) conducted on the subset including sequences from sheep and goats. The bi-dimensional plot displays the genetic differentiation among samples, by representing the number of nucleotide substitutions per site.

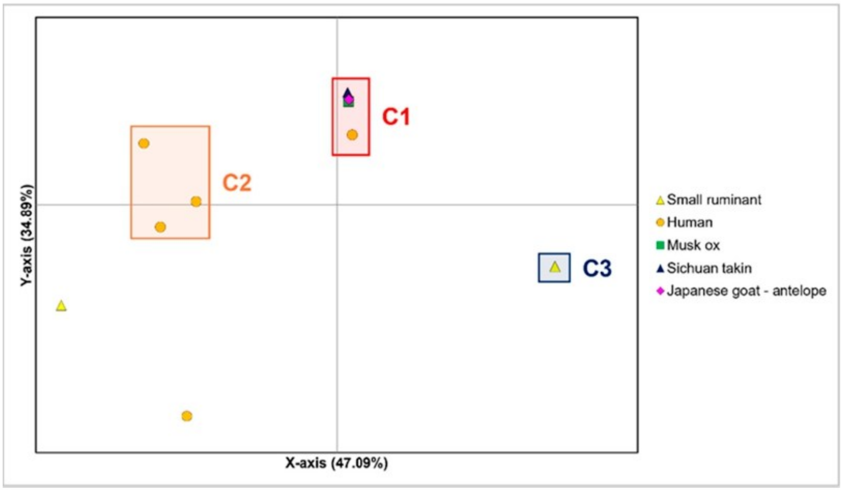


Figure 3. Principal Coordinates Analysis (PCoA) conducted on the subset including sequences from occasional hosts. Group subdivisions are described in the main text. The bi-dimensional plot displays the genetic differentiation among samples, by representing the number of nucleotide substitutions per site.

4. Discussion

The present study explores the genetic variability and phylodynamic patterns of globally distributed Orf virus strains using a comprehensive dataset based on VIR gene variation across natural and occasional hosts. The inclusion of isolates from an under-sampled region of the island of Cuba allowed a broader evaluation of ORFV diversity at the global scale. Moreover, the VIR gene, a variable region of the ORFV genome [48,56], has proven to be highly informative for reconstructing the evolutionary history of the virus and assessing its genetic structuring.

The relevant number of sequences from a wide range of hosts used for our analyses suggests that Orf virus is a pathogen more generalist than previously thought, capable of infecting different ruminant species as well as humans (e.g., [49,50,57,81,82]). In the present study, although the virus does not exhibit a strict host specialization, as it has been suggested in previous studies (e.g., [5,25,31,41,83–85]), early signs of specific genetic differentiation can be observed for strains isolated from sheep and goats.

In particular, the expanded sampling of goat and sheep strains has allowed to deepen on the phylogenetic relationships among ORFV lineages, thus evidencing emerging signs of host specialization. The genetic differentiation between sheep- and goat-associated lineages has already been reported by Coradduzza et al. [49,50,57], whose divergence time estimates for the sheep and goat clusters are largely consistent with those reported in the present study. Indeed, based on complete genome data [50], the divergence of sheep and goat lineages was dated to approximately 1957 and 1947, respectively. In the present study, based on VIR gene analysis, the sheep cluster was dated to 1956 and the goat cluster to 1953. Our findings align with previous whole-genome data, further suggesting that goat strains may have diverged slightly earlier. The minor discrepancies between present and previous estimates may be attributed to the larger number of sequences analyzed in this study and to the smaller genomic region employed (382 bp for the VIR gene versus 131,449 bp for the whole genome).

Notably, consistently with Coradduzza et al. [57], that indicated VIR gene as an informative marker from a phylogenetic perspective, the results obtained here confirm that this gene broadly reflects the evolutionary dynamics of the ORF virus whole genome.

The divergence times for sheep and goat ORFV lineages differentiation correspond to the post-World War II period, a time characterized by major socio-economic and demographic changes. Indeed, historical data indicate that between 1950 and 1960 the global population of ovine species increased significantly [86]. This growth was driven by the rapid post-war agricultural recovery, marked by the development of intensive livestock breeding and a rising movement of both animals and people [87,88]. This phenomenon was accompanied by a substantial increase in the global human population beginning in the 1950s [89], which may have facilitated the spread of pathogens, including Orf virus, through the associated rise in meat consumption and the expansion of sheep and goats farming. In this context, the expansion of ovine and caprine host populations created new opportunities for viral lineages diversification and increasing of population size. The Cuban isolates fell within the major global ORFV clusters, showing genetic similarities to European and Asian strains. This suggests that historical livestock movements and human-mediated introductions shaped the viral population on the island, supporting earlier studies involving full-genome analysis [50].

Phylogenetic findings are partly supported by the PCoAs performed on both the whole dataset and the sheep-goat subset, which showed a central cluster of overlapping strains with divergence at the margins, likely reflecting the typical Wahlund's effect [90]. This principle in population genetics describes a reduction in overall heterozygosity when a population is subdivided into smaller internal subpopulations. If a large, randomly mating population is separated into geographically isolated populations, they may develop different allele frequencies due to random genetic drift [91]. This phenomenon becomes evident when analyzing a broadly distributed population, where viral strains at the extreme edges of the host's range begin to differentiate from each other, despite belonging to a unique, larger genetic group. In this study, the genetic variability explained by the first two axes of the PCoAs, including sheep and goat sequences, indicated that while some strains still co-infect both goats and sheep, others are undergoing divergence, suggesting ongoing host specialization.

This phenomenon may be furthered by the specific livestock management strategies adopted in different regions. Indeed, some areas are characterized by mixed sheep and goat farming, where different species share the same pastures; conversely, other contexts favor specialized production systems, which minimize the opportunities for interspecies contact. For this reason, in isolated and non-promiscuous contexts, the virus evolves by adapting to the specific local host population (either sheep or goats). Notably, the virus appears more advanced in its adaptation to goats, with only early signs of specialization in sheep.

Regarding the additional hosts included in the present study, other ruminants (different from sheep and goats) and humans appear to act as occasional hosts for ORFV, likely due to habitat sharing with sheep and goats, without establishing effective transmission among co-generics or related species. In accordance, in the phylogenetic tree most of these incidental infections are represented by isolated lineages with no descendants.

Noteworthy, in the PCoA performed on the occasional hosts group, human isolates did not show genetic affinities with the other species, including the Japanese goat-antelope, Sichuan takin and musk ox. Indeed, with the only exception of a single human-derived strain from Hungary, which clustered with the occasional hosts group, all human isolates appeared as private variants unique to individual cases, potentially reflecting the sporadic and isolated nature of these zoonotic infections. Such cases may result from rare and independent viral mutations. Human infection could therefore depend on specific risk factors that manifest only in certain individuals and under specific environmental conditions. These factors may include randomly mutated viral allelic variants, predisposing host genetic traits, immunosuppression, or environmental exposures such as direct contact with infected animals. Consequently, Orf virus infections in humans might not represent a consistent zoonotic pattern but rather sporadic, multifactorial zoonotic spillover events. A similar scenario is observed with the cowpox virus, another member of the Poxviridae family, which has expanding animal hosts, but it causes rare human infections. The severity of cowpox virus in humans depends on the host's immune status, skin integrity, and environmental exposure [92,93]. Future research should aim to better disentangle the mechanisms that drive human infection and focus on genome-wide association analyses integrating viral genomic data with human genetic, immunological, and environmental conditions. The availability of additional complete human Orf virus genomes would also be essential to identify potential viral allelic variants associated with specific environmental or host factors.

In this context, the broader dataset used in the present study has also revealed a complex network of strains, including potential recombinant or mutated unique lineages that fail to persist over time and spread among a large number of hosts. This perspective challenges the strong host-specific structure reported in earlier studies, which were based on a smaller number of sequences from sheep, goats, and a few human isolates [52,53]. Those findings may have been biased by reduced sampling, overlooking the virus's ability to infect occasional hosts without sustaining transmission.

Two possible scenarios may explain the presence of ORFV strains in occasional hosts. First, occasional hosts might become infected with viral variants that arise from random nucleotide mutations, enabling the infection of new species. These events are likely sporadic and stochastic, not driven by selection, whose effects are probably destined to disappear rapidly. Second, recombination events may occur within co-infected hosts, generating unique variants that disappear with the host's death and are not transmitted further.

Studying these strains, isolated from occasional hosts, is crucial for understanding the genome evolutionary dynamics of ORFV and determining whether such lineages are exclusive to single individuals or restricted to specific ruminant populations which have been in contact with primary hosts.

Notably, both scenarios may plausibly explain the origin of the human isolated variants private to single individuals. Interestingly, for these human-derived variants, mutations facilitating spillover may have occurred in genomic regions other than the VIR gene. Indeed, the polymorphisms observed for this gene in humans do not result in human-specific amino acid changes but rather correspond to those found in sheep and goat lineages that likely spread among ovine and caprine that live in the same areas.

Further investigations into the genetic variability of the Orf virus, using either complete genomes or polymorphic genes such as VIR, are crucial to better understanding its capacity to infect a broader range of hosts beyond sheep and goats, with a particular focus on human infections.

Such research would be essential for assessing the likelihood of future cross-species transmission and spillover events. The findings may contribute to a more accurate evaluation of the zoonotic risk posed by Orf virus, particularly regarding the potential that its occasional transmission to humans and other mammal species could increase as a result of a possible enhanced infective potential in non-ovine and non-caprine hosts.

5. Conclusions

In conclusion, our results evidence that the VIR gene of Orf virus exhibits higher genetic variability than previously reported [50,57]. The findings highlight the virus ability to infect several ruminant species and humans as occasional hosts, without enabling transmission within those populations. Based on these outputs, we hypothesize the existence of three main genetic groups of lineages for this virus: two host-specific groups (one for goats and one for sheep), showing early genetic differentiation among each other, and a third generalist group, still capable of infecting elective hosts (sheep and goats), but also diffused in other ruminant host species and humans.

Interestingly, the results obtained in the present study suggest that the dynamics of ORFV transmission to humans may be more sporadic than traditionally assumed.

Although ORFV is commonly considered as a zoonotic pathogen, this classification generally assumes that the animal host acts as a reservoir and that the pathogen can efficiently cross the species barrier under natural conditions, even if the frequency of transmission may vary. In contrast, the present findings indicate that ORFV transmission from sheep and goats may require a combination of uncommon and highly specific circumstances (e.g., a particular viral mutation, a rare human genetic susceptibility, and intense exposure). Furthermore, transmission may occasionally arise from co-infection of the same individual with different viral strains, enabling recombination events. Under these conditions, this virus may not behave as a typical, efficiently transmitted zoonosis and human infection appears to be a dead-end spillover event rather than the result of viral adaptation to the human host. However, it should be taken into consideration that human infections in immunocompetent individuals are rarely reported in endemic areas, as the disease is generally considered self-limiting and requires no medical intervention. Consequently, this leads to a limited number of isolated viral strains from human hosts, which may slightly bias the present results.

In conclusion, these new findings suggest that ORFV may have a limited capacity for effective spillover and replication in human cells. The requirement for specific stochastic viral mutation or recombination events indicates that the general viral genome structure is not naturally adapted to human hosts. Furthermore, the species barrier between sheep and goats and humans appears to remain intact, with transmission occurring only when several unlikely conditions coincide. From this perspective, human infection may arise not from adaptation driven by natural selection, but from stochastic events facilitated by sporadic nucleotide mutations that provide no advantage to either the virus or the host and thus do not persist over time. In this context, human infections may be driven more by habitat overlap with the elective hosts species and specific risk factors, rather than a high intrinsic zoonotic potential. This perspective challenges the long-held assumptions regarding ORFV's zoonotic threat, suggesting that transmission to humans could be a sporadic event rather than an evolutionary trait.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/v18020222/s1, Figure S1: Map showing the geographical distribution of the Cuban samples analyzed in the present study; Figure S2: Phylogenetic signal analysis of the complete dataset. Likelihood mapping divides the equilateral triangle into seven distinct regions. The three corner trapezoids correspond to quartets that strongly support fully bifurcating, tree-like phylogenies. The side rectangles denote areas where the data do not clearly distinguish between two alternative tree topologies. The central region represents points P (posterior probabilities of the unrooted trees), where all three trees receive equal support; Figure S3. Principal Coordinates Analysis (PCoA) conducted on the whole dataset. The bi-dimensional plot displays the genetic differentiation among samples, by representing the number of nucleotide substitutions per site. Table S1: ORFV dataset reporting VIR gene sequences used in the present study. The table includes sample codes, GenBank accession numbers, and, where available, sequence origins, host species, collection dates, and corresponding references; Table S2: Sample sizes and genetic diversity estimates obtained for the VIR gene analyzed for host-associated populations. N: sample size; S: number of polymorphic sites; H: number of haplotypes; h: haplotype diversity; π: nucleotide diversity. Sites with gaps were not considered. Group subdivisions are explained within the text. Sequences from Sichuan takin and musk ox were represented by single individuals and were therefore not used for host-specific diversity estimates, although they were retained in the other analyses that included occasional hosts and all host populations.

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

Funding: This research was funded by the Italian Ministry of Health, grant number RC IZS SA 03/20 and Fondi Ordinari/Ordinary funds.

Institutional Review Board Statement: Ethical review and approval were waived for this study, as this study did not involve any animal experiments. Samples were collected from sheep and goats using standard procedures as suggested by the Sardinian Veterinary Services A.T.S. and submitted to the Experimental Zooprophylactic Institute of Sardinia for ORFV testing. Special authorization for sampling activities was not necessary; this action is regulated by the Italian Ministry of Health and performed in the case of infectious diseases.

Data Availability Statement: Sequences obtained in the present study for the ORFV VIR gene were deposited in the GenBank database under the accession numbers OR338563–OR338593.

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

Appendix A

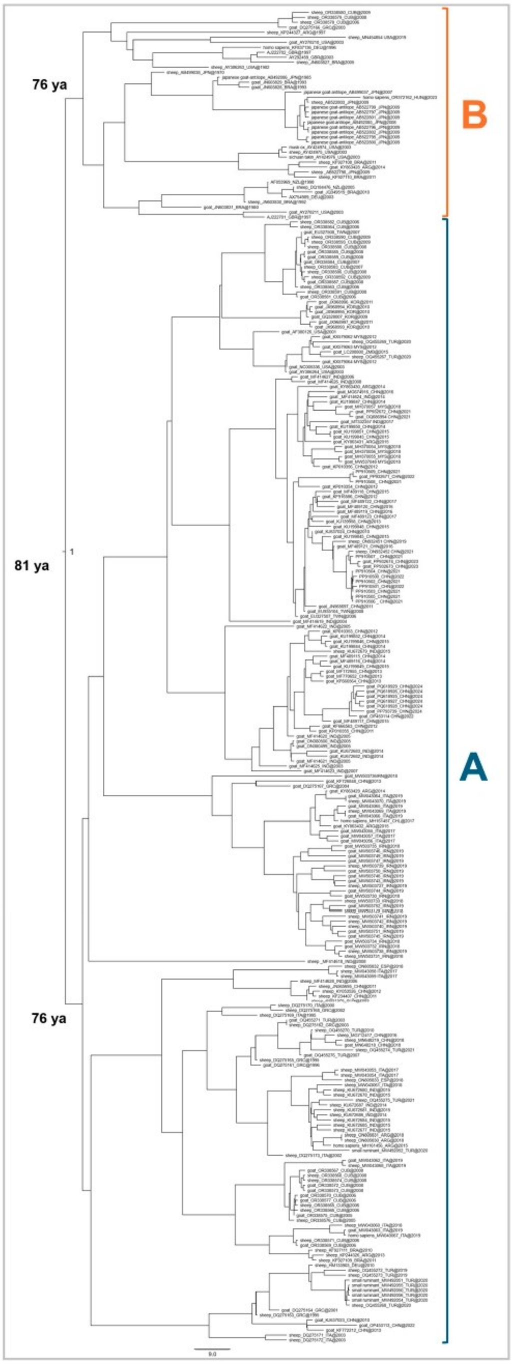


Figure A1. Bayesian phylogenetic tree of the whole dataset with divergence time estimates at major nodes obtained by the software Beast 1.10.4. Group subdivisions are described in the text. Posterior probability values are indicated at the basal node (*pp* = 1.0).

References

1. Alam, M.S.; Akther, S.; Rahman, M.H.; Ali, M.Z.; Ahmed, S. Epidemiological Diversity, Isolation, and Molecular Characterization of Circulating Contagious Ecthyma Virus (Orf Virus) in Goats of Bangladesh. *Egypt. J. Vet. Sci.* **2024**, *56*, 2979–2987. [CrossRef]
2. Nandi, S.; De, U.K.; Sumit, C. Current status of contagious ecthyma or orf disease in goat and sheep—A global perspective. *Small Rumin. Res.* **2011**, *96*, 73–82. [CrossRef]
3. Mathiesen, S.D.; Jørgensen, T.; Traavik, T.; Blix, A.S. On contagious ecthyma and its treatment in muskoxen (*Ovibos moschatus*). *Acta Vet. Scand.* **1985**, *26*, 120–126. [CrossRef] [PubMed]
4. Lawan, Z.; Bala, J.A.; Bukar, A.M.; Balakrishnan, K.N.; Mangga, H.K.; Abdullah, F.F.J.; Noordin, M.M.; Mohd-Azmi, M.L. Contagious ecthyma: How serious is the disease worldwide? *Anim. Health Res. Rev.* **2021**, *22*, 40–55. [CrossRef] [PubMed]
5. Németh, C.; Boros, Á.; Mészáros, E.; Gyömörei, C.; Albert, E.; Pankovics, P.; Reuter, G. Human orf virus (family *Poxviridae*) infection following a lamb bite in Hungary. *Arch. Virol.* **2024**, *169*, 59. [CrossRef] [PubMed]
6. Bamouh, Z.; Tifrouin, I.; Elkarhat, Z.; Abid, L.; Fellahi, S.; Elharrak, M. Pathogenicity and phylogenetic analysis of ovine contagious ecthyma virus isolated during a sheeppox outbreak in Morocco. *Microb. Pathog.* **2024**, *197*, 107023. [CrossRef]
7. Robinson, A.J.; Petersen, G.V. Orf virus infection of workers in the meat industry. *N. Z. Med. J.* **1983**, *96*, 81–85.
8. Zanin, E.; Vicenzoni, G.; Toller, G.; Allegretti, M. Contagious ecthyma in chamois. *Obiet. Doc. Vet.* **1987**, *8*, 61–64.
9. Sarr, J.; Diop, M.; Cissokho, S. Two foci of contagious ecthyma in sheep and goats in Senegal. *Rev. Élev. Méd. Vét. Pays Trop.* **1988**, *41*, 337–338. [CrossRef]
10. Gerstl, F. Diagnosis and morphology of the agent of contagious ecthyma of chamois. *Zentralbl. Bakteriol.* **1964**, *195*, 182–189.
11. Pekelder, J.J.; Talmon, F.P.; Boer, M.J.D. Contagious ecthyma a little known disease of sheep in the Netherlands. *Tijds. Voor Dierg.* **1980**, *105*, 232–239.
12. Tontis, A.; König, H.; Kaderli, R.; Walter, L. Outbreaks of ecthyma contagiosum in 2 Swiss sheep herds and 1 goat flock. *Schweiz. Arch. Tierheilkd.* **1981**, *123*, 19–28.
13. Hussain, K.A.; Yousuf, Y.A.; Qubit, T.S. Contagious ecthyma in sheep in Mosul area. *Iraqi J. Vet. Sci.* **1992**, *5*, 77–84.
14. Spyrou, V.; Valiakos, G. Orf virus infection in sheep or goats. *Vet. Microbiol.* **2015**, *181*, 178–182. [CrossRef] [PubMed]
15. Kumar, R.; Trivedi, R.N.; Bhatt, P.; Khan, S.H.; Khurana, S.K.; Tiwari, R.; Karthik, K.; Malik, Y.S.; Dhama, K.; Chandra, R. Contagious Pustular Dermatitis (Orf Disease)—Epidemiology, Diagnosis, Control and Public Health Concerns. *Adv. Anim. Vet.* **2015**, *3*, 649–676. [CrossRef]
16. Şevik, M. Orf virus circulation in cattle in Turkey. *Comp. Immunol. Microbiol. Infect. Dis.* **2019**, *65*, 1–6. [CrossRef]
17. Vikøren, T.; Lillehaug, A.; Akerstedt, J.; Bretten, T.; Haugum, M.; Tryland, M. A severe outbreak of contagious ecthyma (orf) in a free-ranging musk ox (*Ovibos moschatus*) population in Norway. *Vet. Microbiol.* **2008**, *127*, 10–20. [CrossRef] [PubMed]
18. Dalton, C.S.; Workentine, M.L.; Leclerc, L.M.; Kutz, S.; van der Meer, F. Next-generation sequencing approach to investigate genome variability of Parapoxvirus in Canadian muskoxen (*Ovibos moschatus*). *Infect. Genet. Evol.* **2023**, *109*, 105414. [CrossRef]
19. Guo, J.; Rasmussen, J.; Wünschmann, A.; de La Concha-Bermejillo, A. Genetic characterization of orf viruses isolated from various ruminant species of a zoo. *Vet. Microbiol.* **2004**, *99*, 81–92. [CrossRef]
20. Dieterich, R.A.; Spencer, G.R.; Burger, D.; Gallina, A.M.; Vander Schalie, J. Contagious ecthyma in Alaskan musk-oxen and Dall sheep. *J. Am. Vet. Med. Assoc.* **1981**, *179*, 1140–1143. [CrossRef]
21. Huemer, H.P.; Zobl, A.; Windisch, A.; Glawischnig, W.; Büttner, M.; Kitchen, M.; Trevisiol, K. Serological evidence for Parapoxvirus infection in chamois from the Tyrol regions of Austria and Italy. *Vet. Ital.* **2014**, *50*, 233–236.
22. Inoshima, Y.; Takasu, M.; Ishiguro, N. Establishment of an on-site diagnostic procedure for detection of orf virus from oral lesions of Japanese serows (*Capricornis crispus*) by loop-mediated isothermal amplification. *J. Vet. Med. Sci.* **2017**, *78*, 1841–1845. [CrossRef]
23. Scagliarini, A.; Vaccari, F.; Turrini, F.; Bianchi, A.; Cordioli, P.; Lavazza, A. Parapoxvirus Infections of Red Deer, Italy. *Emerg. Infect. Dis.* **2011**, *17*, 684–687. [CrossRef]
24. Palatsi, R.; Oksanen, A.; Sormunen, R.; Kallioinen, M.; Karvonen, J. The first Orf virus epidemic diagnosed in man and reindeer in 1992-1993 in Finland. *Duodecim* **1993**, *109*, 1945–1950. [PubMed]
25. Klein, J.; Tryland, M. Characterisation of parapoxviruses isolated from Norwegian semi-domesticated reindeer (*Rangifer tarandus tarandus*). *Virol. J.* **2005**, *2*, 79. [CrossRef] [PubMed]
26. Tryland, M.; Nymo, I.H.; Sánchez Romano, J.; Mørk, T.; Klein, J.; Rockström, U. Infectious Disease Outbreak Associated with Supplementary Feeding of Semi-domesticated Reindeer. *Front. Vet. Sci.* **2019**, *6*, 126. [CrossRef] [PubMed]
27. Al-Bayati, H.A.M. Assessment of Contagious Ecthyma Virus in Camels of Wasit Province, Iraq. *Arch. Razi Inst.* **2022**, *77*, 565–571. [PubMed]
28. Abu Ghazaleh, R.; Al-Sawalhe, M.; Abu Odeh, I.; El Ibrahim, J.; Al-Turman, B.; Makhamreh, J. Host range, severity and trans boundary transmission of Orf virus (ORFV). *Infect. Genet. Evol.* **2023**, *112*, 105448. [CrossRef]
29. Khalafalla, A.I.; Elhag, A.E.; Ishag, H.Z.A. Field investigation and phylogenetic characterization of orf virus (ORFV) circulating in small ruminants and Pseudocowpoxvirus (PCPV) in dromedary camels of eastern Sudan. *Heliyon* **2020**, *6*, e03595. [CrossRef]

30. Oryan, A.; Mosadeghhesari, M.; Zibaee, S.; Mohammadi, A. Identification and phylogenetic analysis of contagious ecthyma virus from camels (*Camelus dromedarius*) in Iran. *Onderstepoort J. Vet. Res.* **2017**, *84*, e1–e5. [CrossRef]
31. Fairley, R.A.; Whelan, E.M.; Pesavento, P.A.; Mercer, A.A. Recurrent localised cutaneous parapoxvirus infection in three cats. *N. Z. Vet. J.* **2008**, *56*, 196–201. [CrossRef]
32. Thompson, H.J.; Harview, C.L.; Swick, B.; Powers, J.G. Orf Virus in Humans: Case Series and Clinical Review. *Cutis* **2022**, *110*, 48–52. [CrossRef]
33. Alajlan, A.M.; Alsubeeh, N.A. Orf (Ecthyma Contagiosum) Transmitted from a Camel to a Human: A Case Report. *Am. J. Case. Rep.* **2020**, *21*, e927579. [CrossRef]
34. Cargnelutti, J.F.; Masuda, E.K.; Martins, M.; Diel, D.G.; Rock, D.L.; Weiblen, R.; Flores, E.F. Virological and clinico-pathological features of orf virus infection in experimentally infected rabbits and mice. *Microb. Pathog.* **2011**, *50*, 56–62. [CrossRef]
35. Bala, J.A.; Balakrishnan, K.N.; Abdullah, A.A.; Mohamed, R.; Haron, A.W.; Jesse, F.F.A.; Noordin, M.M.; Mohd-Azmi, M.L. The re-emerging of orf virus infection: A call for surveillance, vaccination and effective control measures. *Microb. Pathog.* **2018**, *120*, 55–63. [CrossRef]
36. Bayindir, Y.; Bayraktar, M.; Karadag, N.; Ozcan, H.; Kayabas, U.; Otlu, B.; Durmaz, R.; Doganay, M. Investigation and analysis of a human orf outbreak among people living on the same farm. *New Microbiol.* **2011**, *34*, 37–43. [PubMed]
37. Key, S.J.; Catania, J.; Mustafa, S.F.; Logan, R.; Kalavala, M.; Hodder, S.C.; Patton, D.W. Unusual Presentation of Human Giant Orf (Ecthyma Contagiosum). *J. Craniofacial Surg.* **2007**, *18*, 1076–1078. [CrossRef]
38. Frandsen, J.; Enslow, M.; Bowen, A.R. Orf parapoxvirus infection from a cat scratch. *Dermatol. Online J.* **2011**, *17*, 9. [CrossRef] [PubMed]
39. Kassa, T. A Review on Human Orf: A Neglected Viral Zoonosis. *Res. Rep. Trop. Med.* **2021**, *12*, 153–172. [CrossRef] [PubMed]
40. Lojkic, I.; Cac, Z.; Beck, A.; Bedekovic, T.; Cvetnic, Z.; Sostaric, B. Phylogenetic analysis of Croatian orf viruses isolated from sheep and goats. *Virol. J.* **2010**, *7*, 314. [CrossRef]
41. Bergqvist, C.; Kurban, M.; Abbas, O. Orf virus infection. *Rev. Med. Virol* **2017**, *27*, e1932. [CrossRef]
42. Concha-Bermejillo, A.; Guo, J.; Zhang, Z.; Waldron, D. Severe Persistent Orf in Young Goats. *J. Vet. Diagn. Investig.* **2003**, *15*, 423–431. [CrossRef]
43. Efridi, W.; Jain, H.; Sathe, N.C.; Ahmed, K. Orf Viral Infection. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2023.
44. Kromer, C.; Bierschenk, C.; Czerny, C.P.; Schön, M.P.; Buhl, T. Orf (ecthyma contagiosum) in a sheep and a shepherd. *Lancet Infect. Dis.* **2018**, *18*, 122. [CrossRef] [PubMed]
45. Fleming, S.B.; Wise, L.M.; Mercer, A.A. Molecular genetic analysis of orf virus: A poxvirus that has adapted to skin. *Viruses* **2015**, *7*, 1505–1539. [CrossRef] [PubMed]
46. Delhon, G.; Tulman, E.R.; Afonso, C.L.; Lu, Z.; De la Concha-Bermejillo, A.; Lehmkuhl, H.D.; Piccone, M.E.; Kutish, G.F.; Rock, D.L. Genomes of the parapoxviruses ORF virus and bovine papular stomatitis virus. *J. Virol.* **2004**, *78*, 168–177. [CrossRef]
47. Hautaniemi, M.; Ueda, N.; Tuimala, J.; Mercer, A.A.; Lahdenperä, J.; McInnes, C.J. The genome of pseudocowpoxvirus: Comparison of a reindeer isolate and a reference strain. *J. Gen. Virol.* **2010**, *91*, 1560–1576. [CrossRef] [PubMed]
48. Peralta, A.; Robles, C.A.; Micheluod, J.F.; Rossanigo, C.E.; Martinez, A.; Carosio, A.; König, G.A. Phylogenetic Analysis of ORFViruses From Five Contagious Ecthyma Outbreaks in Argentinian Goats. *Front. Vet. Sci.* **2018**, *5*, 134. [CrossRef]
49. Coradduzza, E.; Sanna, D.; Scarpa, F.; Azzena, I.F.; Fiori, M.S.; Scivoli, R.; Rocchigiani, A.M.; Bechere, R.; Dettori, M.A.; Pintus, D.; et al. Deeper Insight into Evolutionary Patterns and Phylogenetic History of ORF Virus through the Whole Genome Sequencing of the First Italian Strains. *Viruses* **2022**, *14*, 1473. [CrossRef]
50. Coradduzza, E.; Scarpa, F.; Rocchigiani, A.M.; Cacciotto, C.; Lostia, G.; Fiori, M.S.; Rodriguez Valera, Y.; De Pascali, A.M.; Brandolini, M.; Azzena, I.; et al. The Global Evolutionary History of Orf Virus in Sheep and Goats Revealed by Whole Genomes Data. *Viruses* **2024**, *16*, 158. [CrossRef]
51. Peralta, A.; Robles, C.; Martínez, A.; Alvarez, L.; Valera, A.; Calamante, G.; König, G.A. Identification and molecular characterization of Orf virus in Argentina. *Virus Genes* **2015**, *50*, 381–388. [CrossRef]
52. Kottaridi, C.; Nomikou, K.; Lelli, R.; Markoulatos, P.; Mangana, O. Laboratory diagnosis of contagious ecthyma: Comparison of different PCR protocols with virus isolation in cell culture. *J. Virol. Methods* **2006**, *134*, 119–124. [CrossRef]
53. Oem, J.K.; Chung, J.Y.; Kim, Y.J.; Lee, K.K.; Kim, S.H.; Jung, B.Y.; Hyun, B.H. Isolation and characterization of orf viruses from Korean black goats. *J. Vet. Sci.* **2013**, *14*, 227. [CrossRef]
54. Oem, J.K.; Roh, I.S.; Lee, K.H.; Lee, K.K.; Kim, H.R.; Jean, Y.H.; Lee, O.S. Phylogenetic analysis and characterization of Korean orf virus from dairy goats: Case report. *Virol. J.* **2009**, *6*, 167. [CrossRef]
55. Li, S.; Jing, T.; Zhu, F.; Chen, Y.; Yao, X.; Tang, X.; Chenxiang, Z.; Mingjie, L.; Yanfei, X.; Yuecai, J.; et al. Genetic analysis of Orf virus (ORFV) strains isolated from goats in China: Insights into epidemiological characteristics and evolutionary patterns. *Virus Res.* **2023**, *334*, 199160. [CrossRef]
56. Kaplan, M.; Koç, B.; Pekmez, K.; Cagirgan, A.; Arslan, F. Updated Molecular Characterization of Orf Virus in Türkiye. *Kocatepe Vet. J.* **2023**, *16*, 375–382.

https://doi.org/10.3390/v18020222

57. Coradduzza, E.; Sanna, D.; Rocchigiani, A.M.; Pintus, D.; Scarpa, F.; Scivoli, R.; Bechere, R.; Dettori, M.A.; Montesu, M.A.; Marras, V.; et al. Molecular insights into the genetic variability of ORF virus in a Mediterranean region (Sardinia, Italy). *Life* **2021**, *11*, 416. [CrossRef]
58. Martins, M.; Cargnelutti, J.F.; Weiblen, R.; Flores, E.F. Pathogenesis in lambs and sequence analysis of putative virulence genes of Brazilian orf virus isolates. *Vet. Microbiol.* **2014**, *174*, 69–77. [CrossRef]
59. Ahanger, S.A.; Parveen, R.; Nazki, S.; Dar, Z.; Dar, T.; Dar, K.H.; Dar, A.; Rai, N.; Dar, P. Detection and phylogenetic analysis of Orf virus in Kashmir Himalayas. *Virus Dis.* **2018**, *29*, 405–410. [CrossRef] [PubMed]
60. Li, W.; Hao, W.; Peng, Y.; Duan, C.; Tong, C.; Song, D.; Gao, F.; Li, M.; Rock, D.L.; Luo, S. Comparative genomic sequence analysis of Chinese orf virus strain NA1/11 with other parapoxviruses. *Arch. Virol.* **2015**, *160*, 253–266. [CrossRef] [PubMed]
61. Chi, X.; Zeng, X.; Li, W.; Hao, W.; Li, M.; Huang, X.; Rock, D.L.; Luo, S.; Wang, S. Genome analysis of orf virus isolates from goatsin the Fujian Province of southern China. *Front. Microbiol.* **2015**, *6*, 1135. [CrossRef] [PubMed]
62. Du, G.; Wu, J.; Zhang, C.; Cao, X.; Li, L.; He, J.; Zhang, Y.; Shang, Y. Correction to: The whole genomic analysis of the Orf virus strains ORFV-SC and ORFV-SC1 from the Sichuan province and their weak pathological response in rabbits. *Funct. Integr. Genom.* **2023**, *23*, 176. [CrossRef]
63. Andreani, J.; Fongue, J.; Khalil, J.Y.B.; David, L.; Mougari, S.; Le Bideau, M.; Abrahão, J.; Berbis, P.; La Scola, B. Human infection with orf virus and description of its whole genome. *France. Emerg. Infect. Dis.* **2017**, *25*, 2197. [CrossRef]
64. Mercer, A.A.; Ueda, N.; Friederichs, S.M.; Hofmann, K.; Fraser, K.M.; Bateman, T.; Fleming, S.B. Comparative analysis of genomesequences of three isolates of Orf virus reveals unexpected sequence variation. *Virus Res.* **2006**, *116*, 146–158. [CrossRef] [PubMed]
65. Okonechnikov, K.; Golosova, O.; Fursov, M.; UGENE team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [CrossRef]
66. 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.
67. Available online: https://www.ebi.ac.uk/Tools/msa/clustalo/ (accessed on 24 August 2025).
68. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
69. 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]
70. 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, 2012; pp. 181–209.
71. 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]
72. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574. [CrossRef]
73. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November 2010; pp. 1–8.
74. Gelman, A.; Rubin, D.B. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]
75. 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.* **2018**, *49*, 2057–2073. [CrossRef]
76. Available online: http://tree.bio.ed.ac.uk/software/figtree (accessed on 24 August 2025).
77. Drummond, A.J.; Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **2007**, *7*, 214. [CrossRef]
78. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef] [PubMed]
79. Drummond, A.J.; Ho, S.Y.; Phillips, M.J.; Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **2006**, *4*, e88. [CrossRef] [PubMed]
80. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research—An update. *Oxf. J.* **2012**, *28*, 2537–2539. [CrossRef]
81. Haig, D.M.; Mcinnes, C.J. Immunity and counter-immunity during infection with the parapoxvirusorf virus. *Virus Res.* **2002**, *88*, 3–16. [CrossRef]
82. Haig, D.M. Orf virus infection and host immunity. *Curr. Opin. Infect. Dis.* **2006**, *19*, 127–131. [CrossRef] [PubMed]
83. Hosamani, M.; Scagliarini, A.; Bhanuprakash, V.; McInnes, C.J.; Singh, R.K. Orf: An update on current research and future perspectives. *Expert Rev. Anti. Infect. Ther.* **2009**, *7*, 879–893. [CrossRef]
84. Inoshima, Y.; Ito, M.; Ishiguro, N. Spatial and temporal genetic homogeneity of orf viruses infecting Japanese serows (*Capricornis crispus*). *J. Vet. Med. Sci.* **2010**, *72*, 701–707. [CrossRef]

https://doi.org/10.3390/v18020222

85. Guo, J.; Zhang, Z.; Edwards, J.F.; Ermel, R.W.; Taylor, C., Jr.; de la Concha-Bermejillo, A. Characterization of a North American orf virus isolated from a goat with persistent, proliferative dermatitis. *Virus Res.* **2003**, *93*, 169–179. [CrossRef]
86. Food and Agriculture Organization of the United Nations (FAO). *FAOSTAT Crops and Livestock Products Database* (Livestock Counts: Asses); Accessed via Our World in Data. Available online: https://ourworldindata.org/grapher/livestock-counts (accessed on 22 December 2025).
87. Martín-Retortillo, M.; Pinilla, V. Patterns and causes of the growth of European agricultural production, 1950 to 2005. *Agric. Hist. Rev.* **2015**, *63*, 132–159.
88. Hartung, J. A short history of livestock production. In *Livestock Housing*; Wageningen Academic: Wageningen, The Netherlands, 2013; pp. 21–34.
89. Roser, M.; Ritchie, H. World population growth. *Our World Data* **2013**, *412*, 543.
90. Wahlund, S. Zusammensetzung von population und korrelationserscheinungvom standpunkt der vererbungslehre aus betrachtet. *Hereditas* **1928**, *11*, 65–106. [CrossRef]
91. Garnier-Géré, P.; Chikhi, L. Population Subdivision, Hardy–Weinberg Equilibrium and the Wahlund Effect. In *eLS*; John Wiley & Sons, Ltd.: Chichester, UK, 2013.
92. Vorou, R.M.; Papavassiliou, V.G.; Pierroutsakos, I.N. Cowpox virus infection: An emerging health threat. *Curr. Opin. Infect. Dis.* **2008**, *21*, 153–156. [CrossRef]
93. Pelkonen, P.M.; Tarvainen, K.; Hynninen, A.; Kallio, E.R.; Henttonen, K.; Palva, A.; Vaheri, A.; Vapalahti, O. Cowpox with severe generalized eruption, Finland. *Emerg. Infect. Dis.* **2003**, *9*, 1458. [CrossRef] [PubMed]

https://doi.org/10.3390/v18020222