**MEMORIA CIENTÍFICO-TÉCNICA PROYECTOS INDIVIDUALES**
**Convocatoria 2023 - «Proyectos de Generación de Conocimiento» y actuaciones para la formación de personal investigador predoctoral asociadas a dichos proyectos.**

## 1. *PROPOSAL DATA*

**IP 1** (*Name and surname*)**:** Annie Machordom

***TITLE OF THE PROJECT*:** Genomic exploration of the Atlanto-Mediterranean transition: unveiling genetic patterns and environmental drivers through a seascape genomic approach.

***ACRONYM:*** GAMSeaGA

## 2. *JUSTIFICATION AND NOVELTY OF THE PROPOSAL*

### *2.1. Adequacy of characteristics and the purpose of selected modality.*

Biodiversity is not uniformly distributed or random; instead, biogeographical or phylogeographical patterns are shaped by the biology of species, the characteristics of their living environment and historical pathway. The intricate interplay between the evolutionary history of species and the environmental factors in which they reside establishes distinct biogeographic regions, each with its unique assembly of species. These patterns are a reflection of evolutionary processes, ecological interactions, and historical events that have influenced the distribution of life on our planet. Understanding these biogeographic patterns allows to identify key areas to maintaining connectivity, implement targeted conservation efforts, and comprehend the dynamic relationship between organisms and their environments, especially in the face of contemporary challenges such as climate change. Comprehensive knowledge of the distributions of genetic diversity across multiple taxa in marine biodiversity has been emphasized (Bowen et al., 2016), underscoring the pivotal roles that geography, oceanography, and climate play in shaping marine ecosystems.

Naturally, biodiversity evolves and changes continuously in a natural way at different temporal scales, from geological to seasonal. However, in recent times, a series of changes have been occurring at an unprecedented rate compared to past epochs, as a direct consequence of human activity impacts or, indirectly, derived from the so-called "global or climate change" (Luypaert et al., 2020, for the marine environment). The speed of these changes exacerbates the possibility of adaptation for many species, leading to a significant loss of biodiversity at all levels, and there is an overwhelming amount of information and evidence on this matter (e.g., Karl & Trenberth, 2003; Boisvenue & Running, 2006).

Regardless of the direct impacts of human activities, especially those related to the destruction and fragmentation of natural habitats, climate change influences species by reducing or eliminating local populations, altering migration patterns and connectivity, and causing changes in geographic distributions. All of these factors directly affect the genetic structure of populations (Lo Brutto et al., 2011). To face these rapid changes, and undertake appropriate measures, we have to approach the study of both, species and factors associated to their connectivity, settlement, recruitment and survival in the best conditions. However, while genetic diversity is the foundation of evolutionary changes and significantly influences a species' ability to adapt to environmental changes (Hughes et al., 2008) and resist potential emerging infectious diseases, the methods used to assess the effects of climate change on biodiversity are largely ecological. They are primarily based on demography and distribution patterns of species in relation to environmental parameters and tend to ignore the potential

intrinsic factors of species that determine their adaptability. A low genetic diversity will decrease the ability to adapt and increase the risk of extinction (García-Dorado & Caballero, 2021). Likewise, there is growing evidence of the interrelation between the biological characteristics of species, their dispersal capacity, and gene flow with the genetic diversity of populations. The greater or lesser ability to acclimate or adapt to changes resulting from climate change will determine the "winners" and "losers" in the near future, sensu Somero (2010).

The distribution of neutral genetic diversity in marine environments, as in others, is predominantly shaped by barriers that restrict the dispersal of species. However, unlike the more conspicuous barriers found on land, identifying environmental discontinuities in the ocean poses a greater challenge (Galarza et al., 2019). Oceanic fronts, characterized by abrupt changes in physical and biochemical variables, are generated through various physical processes and are widespread across all oceans. Furthermore, their presence may hinder population connectivity, potentially impacting the survival or differentiation of more or less isolated populations. Some of these barriers are already more or less well-known (see for instance, for the Mediterranean Sea, Galarza et al., 2019 and Repullés et al., 2022). The ability to pinpoint such barriers is crucial for determining the extent of the potential genetic exchange among marine populations. This knowledge, in turn, plays a fundamental role in enhancing our understanding of the dynamic genetic structure of populations and is essential for effective marine species management and conservation of marine species.

Undoubtedly, time plays a pivotal role in shaping the evolutionary history of populations and, thus, species. Historical environmental factors, such as habitat alteration, sea level changes, ocean currents, and past glaciations, along with biological species traits, crucially shape both inter- and intraspecific differentiation patterns (Patarnello et al., 2007). The intricate interplay between these historical interactions and contemporary environmental conditions is considered a focal point in marine phylogeography. We intend here to deal with different taxa, because species exhibiting consistent geographical patterns of intraspecific structure not only provide valuable evidence but also document instances of phylogeographical breaks (Patarnello et al., 2007). These disruptions, both ancient and modern, manifest in the biogeographical realm, highlighting the complexity of marine ecosystems (Longhurst, 1998).

The Atlanto-Mediterranean transition provides an excellent scenario to inquire the influence of different environmental factors on species adaptation. The narrow connection between the Atlantic Ocean and the Mediterranean Sea through the Strait of Gibraltar has been proposed to represent an important phylogeographic break in several marine species, based on oceanographic models and phylogeographic studies. Additionally, the distinct features of the Atlantic and Mediterranean, including complex water circulation, temperature, and salinity, among other factors, provide an ideal foundation for comparing genetic patterns and assessing the influence of environmental factors.

These factors are compounded by the intricate paleogeography of the region, historically characterized by extensive changes in configuration and climate. The biodiversity of the area is significantly shaped by its prolonged evolutionary history throughout the Cenozoic era, involving the closure of the Tethys (Torfstein & Steinberg, 2020), the Messinian Salinity Crisis (approximately 5.3 million years ago, Krijgsman et al., 1999), and well-documented post-Pliocene sea level regressions and Pleistocene glaciations. Consequently, the Mediterranean experienced multiple isolation events from the Atlantic, followed by repopulation by species of Atlantic origin (Taviani, 2002).

More recently, the marine environment is undergoing significant impacts due to rising temperatures, coupled with associated alterations in precipitation, ocean biogeochemistry, sea levels, and extreme storms and heat waves. In the examined region, projections suggest that the absolute sea-level rise will be more pronounced in the Atlantic than in the Mediterranean (Losada et al., 2014). Concurrently, temperatures are anticipated to rise more significantly in Mediterranean areas (Volosciuk et al., 2016). These climatic shifts may give rise to various changes and threats, including species displacement, alterations in food webs and life cycles,

mass mortalities, and the spread of diseases, as highlighted by Templado (2014). Invasive species introductions represent yet another menace, as non-native species can outcompete and displace native aquatic organisms, creating ecological imbalances. It is imperative that we address these multifaceted challenges (Rilov & Crooks, 2009).

The general situation of environmental changes and loss of diversity is even more pronounced in coastal areas (Halpern et al., 2015). The cumulative impacts concentrated on the coast, with accelerated destruction and fragmentation of habitats, are compounded by the consequences of rapid environmental change, where alterations in parameters such as temperature, salinity, and sea level rise will be combined. The loss or alteration of coastal habitats due to changes in the coastline can lead to the inevitable disappearance of multiple populations of species linked to coastal environments (McCauley et al., 2015). In fact, climate change is already causing shifts in the geographic distribution of many species, changes that can be predicted with various models (Parmesan & Yohe, 2003).

Thus, coastal conservation not only helps prevent erosion, flooding, and damage to infrastructure but is also crucial for preserving the biodiversity it harbours. This includes species directly settled in these areas (such as sessile species) and those that use them as breeding grounds.

In addition to their biodiversity and ecological significance, it is crucial to highlight the far-reaching economic and social repercussions posed by the challenges confronting coastal marine ecosystems, with a particular emphasis on the gendered impact. The degradation of these environments can profoundly affect local economies, especially those dependent on activities like fishing and water-related activities, disproportionately impacting women who are traditionally engaged in activities such as shellfish harvesting, a practice particularly prevalent in certain regions of the Iberian Peninsula. Disruptions in these sectors not only jeopardize livelihoods but can also lead to food security concerns within communities, with woman bearing a significant burden. Recognizing the interconnectedness of ecological, economic, and social factors is essential for a comprehensive approach to preserving these critical coastal habitats (de Jonge et al., 2012). Ultimately, the conservation and restoration of aquatic ecosystems are not only crucial for the survival of countless species but also for the well-being of our planet and future generations.

In the face of global change and widespread habitat degradation, understanding the underlying processes and consequences of dispersal and connectivity are of critical importance for conservation management (Nanninga & Berumen, 2014). Hence, it is imperative to take action to safeguard the marine biodiversity and the ecological services they provide, quantifying changes, identifying patterns and mechanisms and testing solutions. Further, there must be a match between the priorities of academic researchers and the needs of stakeholders. Therefore, one part of the solution is to identify the research needs of practitioners (Sutherland et al., 2009).

After the comprehensive review by Patarnello et al. (2007) on the Atlanto-Mediterranean transition as a phylogeographical decisive break, some new approaches have continued to be published on the subject (e.g., Alberto et al., 2008; Galarza et al., 2009a,b; Pellerito et al., 2009; Fernández et al., 2010; Maltagliati et al., 2010; Fruciano et al., 2011; García-Merchán et al., 2012; Deli et al., 2016; Pascual et al., 2017; Heras et al., 2019; González-Castellano et al., 2020; Ojeda et al., 2022). However, most studies largely consists of "single marker - single species" and the phylogeographical insights do not reach the same level of resolution. It is necessary moving towards establishing a genome-based reconstruction of bio- and phylogeographic patterns. This involves enhancing our understanding of genetic and phylogeographic aspects by utilizing advanced genomic approaches. Our future endeavours should focus on acquiring in-depth species knowledge, coupled with refined phylogeographical calibration from a seascape perspective. This comprehensive approach is essential for fully grasping the influence of species-specific biological traits on the observed patterns.

### 2.2. *Justification and expected contribution of the project to the generation of knowledge on the theme of the proposal. Starting hypothesis*

In the presented background we would like to frame this project. We plan to select several non-model marine species of different phyla (Cnidaria, Crustacea, Mollusca and Echinodermata), which represent a wide range of biological traits, types of larval development and ecological requirements (see below under “species selection” in point 3.2. "Description of the methodology: Species selection").

This approach will provide crucial knowledge of genetic variation along two biogeographic areas and their transition zone. Moreover, it aims to offer essential data that will serve as the basis for the effective management of a diverse range of aquatic organisms, encompassing species inhabiting coastal marine environments. The genomic information will provide fundamental genetic parameters of diversity and connectivity, which we will link with abiotic factors, to understand where the actions are more urgent and to establish priority areas of conservation, considering the possible corridors. Thus, we will emphasize two critical pieces of information to characterize the state of these species: their genetic diversity and the connectivity between their populations. Based on the results obtained, we can conduct a comprehensive final analysis, using systematic conservation planning techniques (Kukkala & Moilanen, 2013) to propose scenarios that not only maximize the preservation of different species but also address something as fundamental as their genetic variability. This involves maximizing the representativeness, connectivity, and efficiency of the network of areas to protect all levels of biodiversity.

In this dual pursuit of acquiring basic knowledge and facilitating potential conservation applications, we present this proposal as "non-oriented", although some of its consequences could be associated with the thematic priorities of "Food, bioeconomy, natural resources, and the environment". This topic is also fully aligned with the Recovery, Transformation, and Resilience Plan approved by the European Commission on June 16, 2021, specifically regarding its first transversal axis, that of ecological transition. The objectives of this axis include "mitigating climate change, adapting to climate change, sustainable use and protection of water and marine resources, circular economy, prevention and control of pollution, and protection and recovery of biodiversity and ecosystems", as reflected in the regulation approved by the European Parliament and the Council on June 18, 2020, known as the "Taxonomy Regulation". In addition, this proposal aligns perfectly with the United Nations' Sustainable Development Goals (SDGs, ODS in Spanish). By advancing our knowledge of the aquatic ecosystems functioning and contributing to their conservation, the project addresses several key SDGs, including Life Below Water (SDG 14) and Life on Land (SDG 15). Furthermore, the project's commitment to environmental stewardship and the sustainable management of aquatic resources resonates with broader SDGs, such as Zero Hunger (SDG 2) and Climate Action (SDG 13). Thus, this initiative not only contributes to scientific knowledge but also promotes global efforts to achieve a sustainable future.

Therefore, we could cite the following **hypotheses** to be tested:

1. *Species complex*: some of the analysed species may constitute species complexes, with cryptic or pseudocryptic species differentiated in the studied area.

2. *Genetic diversity and connectivity*: the genetic diversity and connectivity between populations will vary based on the different biological traits of the species, thereby influencing the overall genetic structure.

3. *Molecular markers*: the use of massive sequencing will provide a greater differentiation power and result in better recognition of population differentiation.

4. *Biogeographic barriers*: different biogeographic or oceanographic barriers, such as the Gibraltar Strait, the Almeria/Oran front or the Ibiza Channel, will explain population differentiation.

5. *Human-mediated dispersal*: human activities, including shipping and aquaculture, may contribute to the dispersal of species, thereby impacting their genetic structure and potentially altering the influence of natural barriers.

6. *Environmental parameters*: certain environmental parameters (e.g., temperature, water pH, salinity, or currents) may exhibit correlations with genetic variation among populations.

7. *Mediterranean vs. Atlantic impact*: due to its greater environmental impact, Mediterranean populations are expected to exhibit smaller population sizes and lower genetic diversity compared to Atlantic populations.

8. *IPCC projections*: anticipated alterations in different environmental factors are expected to negatively impact the adaptation and survival of coastal species, potentially resulting in a considerable loss of biological diversity.

## 3. *OBJETIVES, METHODOLOGY AND WORK PLAN*

### 3.1. *General and specific objectives*.

The main goal of our project is to leverage state-of-the-art tools offered by cutting-edge sequencing technologies and data analytics to unveil differentiation patterns in Atlanto-Mediterranean marine coastal species. Subsequently, we aim to apply the acquired data on variability and factors correlated with population genetic structure to inform conservation strategies, all while considering the projections of climate change.

As a result, the specific objectives are as follow:

1. Verify the primary differentiation of species in their distribution range, to ensure the analysis of unique species.

2. Acquire fundamental parameters of population genetic differentiation, structure, and connectivity.

3. Analyse, within a seascape genetic framework, the correlation among collected environmental factors and observed genetic structure, assessing the impact of predicted environmental changes on genetic variability and population connectivity.

4. Provide essential data identifying source and sink populations in the studied area and potential corridors to protect current diversity.

5. Recognize the magnitude of larval exchange among (sub)populations (coined "population connectivity" in contemporary literature).

6. Differentiate the levels of genetic connectivity: low levels ("adaptive connectivity" sufficient to spread advantageous alleles), moderate levels (sufficient to avoid inbreeding), and high levels ("demographic connectivity").

7. Investigate the impact of human activities, including shipping, aquaculture, and coastal modifications, on the dispersal dynamics of marine coastal species. Evaluate how these activities influence genetic patterns and population connectivity, contributing insights into effective conservation strategies.

8. Align acquired genetic data with existing conservation policies and marine protected areas design, aiming to provide actionable insights for developing or adjusting conservation strategies in response to climate change.

To achieve this overarching goal, we plan to employ cutting-edge methodologies, leveraging molecular insights and utilizing advanced bioinformatic tools. The results obtained will enable a comprehensive final analysis, incorporating systematic conservation planning techniques. This analysis aims to propose scenarios that maximize the conservation of different species, considering the fundamental aspect of their genome variability. The approach involves optimizing the representativeness, connectivity, and efficiency of the network of areas to protect all levels of biodiversity.

### 3.2. *Description of the methodology*.

The methodology will proceed through these outlined steps, with the flexibility to adapt to evolving methodological advancements. We aspire to integrate innovative methods as they emerge, ensuring not only cutting-edge approaches but also prioritizing those that provide advantageous cost-effectiveness and deliver more robust results.

**Species selection.** We have chosen a set of non-model marine coastal species based on their representation of distinct biological features, such as the type of larval development (directly linked to their dispersal capacity and, consequently, connectivity), habitat and biogeographical affinities.

These species include six gastropods (the limpets *Patella rustica* and *Patella ulyssiponensis*, the false limpet *Siphonaria pectinata*, the keyhole limpet *Fissurella nubecula*, the abalone *Haliotis tuberculata*, and *Bittium reticulatum*), two decapods (the crabs *Pisa tetraodon* and *Porcelana platycheles*), the scleractinian coral *Leptopsammia pruvoti*, and the cushion seastar *Asterina gibbosa*.

All these species are distributed throughout the Atlantic-Mediterranean region. Among them, two species exhibit Mauritanian affinities (*S. pectinata* and *F. nubecula*), two Lusitanian affinities (*P. platycheles* and *A. gibbosa*), two Mediterranean affinities (*Bittium reticulatum* and *Leptopsammia pruvoti*), while the rest have a wide Atlanto-Mediterranean distribution with no defined affinity.

Four of the species are intertidal (*Patella rustica, P. ulyssiponensis, Siphonaria pectinate* and *Fissurella nubecula),* three live in shallow rocky bottoms (*Haliotis tuberculata*, *Asterina gibbosa* and *Porcelana platycheles*), two are common inhabitants in sublittoral algal communities, while *Leptopsammia pruvoti* lives on shady rock walls at varying depths.

Regarding the larval dispersal, two of the species have no or low dispersal (0-5 days): *Asterina gibbosa* (without free larval phase) and *Leptopsammia pruvoti* (with a short demersal planulae); four species have medium dispersal larval capacity (between 5 and 20 days): *Patella rustica, P. ulyssiponensis*, *Fissurella nubecula* and *Haliotis tuberculata*); and four species with high dispersal capacity (> 20 days): *Siphonaria pectinata*, *Bittium reticulatum, Pisa tetraodon* and *Porcelana platycheles*. It must be taken into account that larvae of species with medium dispersal capacity may be competent for settlement in a few days, but can considerably delay it if a suitable substrate is not found. That is why the duration of its larval phase can vary. Likewise, in species with a high dispersal capacity, their larval phase also varies depending on the species and can last up to about two months in *P. platycheles.*

Data references: Fretter & Graham (1976, 1981), González-Gordillo et al. (1996), Rodríguez (1997), Gofredo et al. (2006), Sá-Pinto et al. (2012), López-Márquez et al. (2018).

**Sampling Design.** We will identify suitable locations along both the Atlantic and Mediterranean coasts, covering both sides of the main barriers. As a preliminary guide, these locations may encompass one population in the Catalan coast, another in the Balearic Islands, one from Murcia, others from Almería, Cádiz, Canary Islands, southern of Portugal, western Portugal, Galicia, and Cantabria. Additionally, when feasible, we will include a reference sample from the northern area (France or the United Kingdom) or from the North African coast (Algeria, facilitated by one of our collaborators' sampling), and from the Tyrrhenian Sea.

Sampling will involve a combination of scuba diving, snorkelling, and manual collection from the coastline. Non-invasive techniques will be prioritized to minimize any potential impact on the studied ecosystems. Regrettably, this minimal tissue recovery will hinder gender recognition, as the sex of most selected species can only be identified through a lethal biopsy or the presence of eggs (in which case, it will be duly noted, in case any observed parameters could potentially be correlated with gender).

Furthermore, we will take advantage of the extensive resources available in the MNCN collections, which provide valuable specimens for comparative analyses. The use of existing collections enhances our ability to study historical samples and contributes to a more robust understanding of the species' genetic dynamics over time.

Depending on the natural distribution of the species, we plan to sample approximately 12-15 localities (that will be georeferenced). We aim to obtain a minimum of 10 populations for each species, with 20-30 individuals per location (averaging around 250 samples per species).

**DNA Extraction.** DNA will be extracted from the sampled specimens preserved in suitable conditions (e.g., absolute ethanol or RNAlater) using Qiagen kits (such as DNeasy 96 Blood & Tissue Kit). The total genomic DNA will be quantified using a Qubit with the High Sensitivity dsDNA Assay kit (ThermoFisher Scientific).

**Species molecular delimitation.** A specific fragment, most likely the cytochrome oxidase subunit I (COI), will be amplified and sequenced to detect and exclude cryptic or pseudocryptic species from population analyses. Forward and reverse primers previously used in our lab for closely related species will be employed to amplify this region (e.g., LCOI 1490, Folmer et al., 1994 and COH, Machordom et al., 2003) when specific primers are not available. Amplification cycles will be adjusted for different species, following the protocols routinely used in our lab (e.g., Rodríguez-Flores et al., 2019).

A novel approach will involve sequencing the obtained amplicons using an Oxford Nanopore Technologies (ONT) MinION MK1C sequencer, developed in our lab. To achieve this, primers will incorporate tails during the PCR process, and individual barcodes will be added for multiple sample inclusion in a single sequencing run. Amplicons will be cleaned with Ampure XP magnetic beads (Beckman Coulter) and ligated using ligation kit v14 (SQK-LSK114) to generate libraries for subsequent sequencing. DNA library quality will be assessed on a TapeStation 2200 system (Agilent). Libraries will be loaded into an ONT R10.4 flow cell for MinION, and samples will run for five hours.

The CSIC's DRAGO cluster will be utilized for high-precision basecalling of the resulting fast5 files using Guppy. Since the sequenced fragment usually does not show gaps among specimens of the same species, the alignment will only be performed to remove artifact-generated gaps (MAFFT, Katoh & Toh, 2008). Sequence sets will be curated in Sequencher (Gene Code Corporation) or Geneious (Kearse et al., 2012). The consensus sequence of each specimen will be integrated into a nexus matrix for subsequent data treatment, including Bayesian inference (in MrBayes, Ronquist et al., 2012) and, if necessary, species delimitation analysis (e.g., bGMYC, Reid & Carstens, 2012; bPTP, Zhang et al., 2013; ABGD, Puillandre et al., 2012). If differentiation were found, we will also apply divergence dating treatment to relate such differentiation to the possible epoch of isolation between the areas involved (BEAST, Bouckaert et al., 2019).

**MobiSeq analysis:**

Library preparation and whole genome sequencing. One representative DNA sample per species will undergo low-coverage whole genome sequencing to identify the most suitable transposable element (TE) for the MobiSeq experiment, following the methodology outlined by Rey-Iglesia et al. (2019). Library construction will employ Illumina DNA Prep, and sequencing will be performed on a fraction of a NovaSeq PE150 flow cell, targeting a total output of 90 gigabases. Subsequently, raw fastq files will undergo quality assessment using FastQC v0.11.5 (Andrews, 2010), and removal of duplicate reads will be carried out with the

clumplify.sh script from the BBmap package 38.90 (Bushnell, 2014). Adapter removal and quality filtering will be performed using Trimmomatic 0.39 (Bolger et al., 2014).

MobiSeq primer design custom pipeline. Utilizing the low-coverage whole genome sequencing data generated, an in-silico selection of the most suitable transposable element (TE) or variable repetitive region, termed the candidate region, for the MobiSeq experiment (Rey-Iglesia et al., 2019) will be conducted. This selection will follow a custom pipeline designed for this specific purpose. The processed reads will undergo normalization to achieve a 1× flat coverage distribution using the bbnorm.sh script from the BBmap package 38.90 (Bushnell, 2014). Subsequently, the kmercountexact.sh script from the same package will be employed to extract and count the occurrences of kmers in the filtered reads. Kmers with an occurrence lower than 250 counts (mincount = 250) will be discarded. Additionally, kmers with more than three repeats of the same nucleotide and those ending with three consecutive GC bases will be filtered out.

Concerning the flanking regions, any homogeneous or microsatellite-rich amplified flanking regions will be excluded using a bash script and the vsearch software (Rognes et al., 2016) with the "–cluster_fast" option set. Flanking regions of the remaining candidate kmers will be extracted from the pre-processed reads using Cutadapt 3.2 (Martin, 2011), allowing for one mismatch. The estimated number of recoverable loci will be calculated using the ustacks module of Stacks (Rochette et al., 2019), achieved by clustering the kmer flanking regions. These flanking regions will also undergo visual inspection for the presence of highly repetitive reads using Geneious, ensuring a meticulous selection of candidate regions for subsequent MobiSeq experiments.

Characterization and primer design. Four distinct primers for each species, associated with varying numbers of estimated recoverable loci, will be designed. A pilot project will be initiated to assess primer performance, guiding the selection of the most efficient primer for each species.

High-Throughput Sequencing. The equimolar mixture will be sequenced on the NovaSeq PE150 sequencer or equivalent. The expected maximum output is approximately ~0.5 gigabases of raw data per sample, including possible controls that need to be added to the sequencing mix (PhiX phage DNA or unmethylated Lambda phage DNA). These values represent the theoretical limits specified by the manufacturer for this sequencing platform, and there may be discrepancies with the final results obtained due to intrinsic variability in each sequencing mix.

SNP calling and filtering. The SNP calling process, basic filtering, and preliminary population analysis will be conducted using the dDocent pipeline (Puritz et al., 2014). Basic filtering parameters using VCFtools (Danecek et al., 2011) include retaining variants genotyped in at least 50% of individuals, a minimum quality score of 30, a minor allele count of three, a minimum depth for a genotype call of five reads, and loci thinning to avoid linkage disequilibrium, ensuring that no two sites are within 600 base pairs from each other. A second round of filtering will be applied to refine the SNP dataset for population studies. This includes filtering based on minor allele frequencies (MAF), Hardy-Weinberg Equilibrium (HWE), missing SNPs, and Linkage Disequilibrium (LD). These filters aim to enhance the accuracy and meaningful exploration of genetic diversity and relationships between populations (Linck & Battey, 2019; Falush et al., 2003; Danecek et al., 2011; Malomane et al., 2018).

**Environmental parameters.** Accurate foundational data on potential factors correlated with genomic diversity and differentiation is paramount. Among them, we can highlight temperature (range, maximum, and minimum in sea surface and air), salinity, pH, chlorophyl, topography, winds, and currents. To achieve this, we will procure data from, for instance, Bio-ORACLE (extended to marine data, Assis et al., 2018) or generate them through satellite data. Leveraging Google Earth Engine, we will operate at an optimal resolution aligning with the scale of the obtained genomic data. This meticulous approach ensures a comprehensive understanding of the environmental landscape surrounding our genetic investigations.

**Genomic data analysis.** The results will be analysed, as previously stated, with the support of the CSIC's DRAGO cluster. Initially, outlier loci detection will be conducted through both Bayesian and principal component analyses. Bayesian analysis will be executed using BayeScan (Foll & Gaggiotti, 2008), while principal component analysis will be conducted with PCAdapt (Privé et al., 2020). The identified potential outlier loci will be isolated from the data matrix of each species, resulting in the creation of three distinct matrices per species for subsequent data treatment: one inclusive of all loci (global), one excluding the outliers, and a third comprising only the outliers. These matrices will be utilized to test for the role or correlation of specific loci with adaptation or selection, aiming to uncover their significance in the studied populations.

Fundamental parameters, including allele and genotype frequencies, observed and expected heterozygosities, inbreeding coefficient (Fis), and tests for adherence to Hardy-Weinberg equilibrium, will be obtained using software tools such as PLINK (Purcell et al., 2007).

Population differentiation will be assessed by computing metrics such as Fst, calculated using VCFTools, to estimate gene flow and connectivity. Migration patterns will be evaluated using either the EEMS software (Petkova et al., 2015) or MIGRATE-n (Beerli & Felsenstein, 2001). Visual representation of differentiation will be conducted through PCA analyses and DAPC, leveraging the R package adegenet (Jombart & Ahmed, 2011). ADMIXTURE (Alexander et al., 2009) will be employed to unravel intricate patterns of genetic diversity within and between populations. To analyse molecular variance in spatial terms, incorporating geographic information, SAMOVA (Dupanloup et al., 2002) will be utilized, considering K values ranging from one to a number surpassing the total populations analysed.

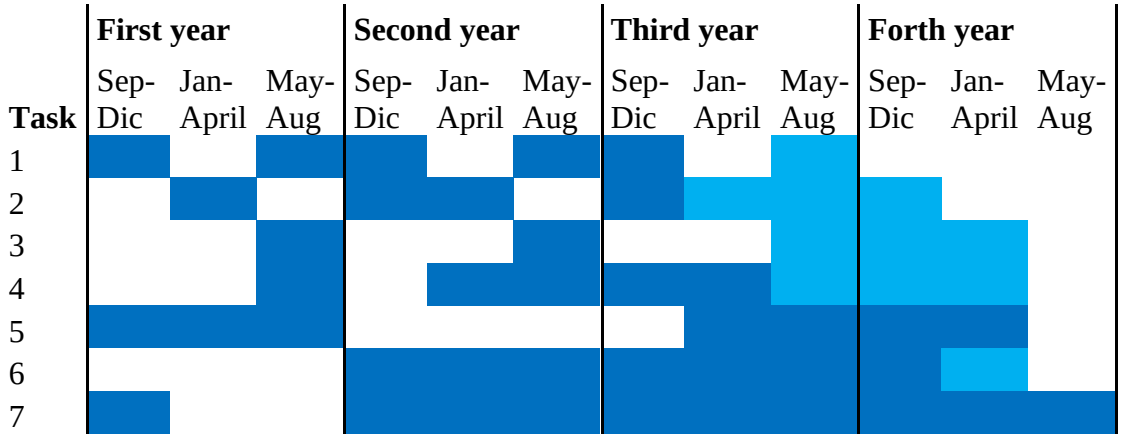
Demographic parameters and history will be comprehensively analysed using multiple tools. NeEstimator (Do et al., 2014) will be employed to estimate effective population sizes, LDNe (Waples & Do, 2010) will be used to assess linkage disequilibrium-based effective population size, and BEAST (Drummond et al., 2012) will be utilized to delve into the demographic history within a temporal framework, providing nuanced insights beyond punctual values.

Estimates of isolation by distance (IBD), environment (IBE), and resistance (IBR) will be derived using a diverse set of principles, software, and packages. Specifically, tools such as marmap (Pante & Simon-Bouhet, 2013), Mantel tests, GGally (Schloerke et al., 2018), vegan (Oksanen et al., 2019), and Circuitscape (McRae, 2006) will be harnessed for these analyses. Furthermore, individual-based population and genetic modelling will be conducted with CDfish (Landguth et al., 2012) and CDPOP (Landguth & Cushman, 2010). These methods not only provide insights into individual-level population dynamics but also play a pivotal role in predicting essential factors such as population sizes, genetic diversity, gene flow, and connectivity across populations. It is noteworthy that our collaboration with Samuel Cushman, a leading expert in the development of this approach, will significantly enhance the robustness of our analyses.

**Communication, dissemination.** Effective communication of our research findings is integral to the success of our project. We plan to disseminate our results through various channels, ensuring that our scientific contributions reach a diverse audience. In the upcoming sections, we will provide comprehensive insights into our outreach efforts. Our communication strategy will target the scientific community, early-career researchers, university students, biodiversity managers, and the general public. We will leverage our affiliation with a centre that includes exhibitions and public engagement opportunities, in addition to utilizing the communication channels available at the MNCN and the CSIC. We will provide more information in sections 4.3 and 4.4.

### 3.3. *Work plan and schedule*.

Proposed schedule: in dark blue are the most probable periods for performing the tasks, and in light blue those periods considered as contingency.




Task numbering:

1. Sampling. We will begin by verifying the samples harboured by the collections from the MNCN and collaborators, selecting those useful for the project. Most of the upcoming samplings are scheduled during the most favourable weather periods. Responsible: JT. Participants: IA, PRF, VLM, MO, MK, HP, PS, AM.

2. DNA extraction. As samples are collected, their DNA will be extracted. Responsible: IA. Participants: PS, HP, AM.

3. Species delimitation: Concurrently with sample collection and DNA extraction, we will amplify the selected marker to identify possible cryptic or pseudocryptic species. Responsible: AM. Participants: IA, HP, PS.

4. MobiSeq analyses. The first step involves performing whole genome sequencing to select informative regions. Most of these analyses will be carried out by an external service. Responsible: AM. Participants: HP, PS.

5. Obtaining environmental parameters. Different sources will be consulted to prepare the environmental dataset. Responsible: JT. Participants: PRF, VLM, SC.

6. Data analyses. Data will be processed for various purposes, including species delimitation, SNP characterization, genomic parameters obtention, population structure, and correlation with environmental features. Responsible: AM. Participants: PRF, VLM, SC, MO, HP, PS.

7. Communication, presentation and dissemination of results. Responsible: JT and AM. Participants: PRF, VLM, SC, MO, MK, HP, PS.

Abbreviations used: JT, José Templado; IA, Iván Acevedo; PRF, Paula Rodríguez Flores; VLM, Violeta López Márquez; SC, Samuel Cushman; MO, Marco Oliverio; MK, Mohammed Kallouche; HP, hired person; PS, predoctoral student; AM, Annie Machordom.

As the CVs of the "working team" cannot be included in the online submission, here we provide a brief description of the main scientific contributions of the collaborators outside of Spain:

Paula Rodríguez Flores is a postdoc at the Smithsonian National Museum of Natural History. She has 8 years of research experience and has mentored several students and co-supervised 2 BSc theses, 1 master's student, and 1 PhD student. She has published 42 peer-reviewed papers, and 3 book chapters. Her work experience includes research at the Smithsonian National Museum of Natural History, Washington DC; Museum of Comparative Zoology, Harvard University, Cambridge; Benthic Invertebrate Collection at SCRIPPS Institution of Oceanography, San Diego; Muséum national d'Histoire naturelle, Paris; Centre d'Estudis Avançats de Blanes; and Museo Nacional de Ciencias Naturales, Madrid. In national and international collaborations, she has successfully obtained three international research awards and three research grants from Harvard University. ORCID: 0000-0003-1555-9598.

Violeta López Márquez is a postdoctoral researcher at the University of Aruba and KU Leuven University completed her PhD in 2021 with nine years of research experience. She has mentored several students and co-supervised 2 BSc theses and 1 master's student. She has published 12 peer-reviewed papers, and 1 book chapter. Her work experience includes field expeditions, research and teaching at the University of Aruba, Dutch Caribbean; KU Leuven University, Belgium; US Forest Service Rocky Mountain Research Station, Arizona and Museo Nacional de Ciencias Naturales, Madrid. In national and international collaborations, she has participated successfully in 9 projects. ORCID: 0000-0002-2486-6031.

Samuel Cushman, currently a Senior Research Fellow at the University of Oxford and former Senior Scientist at the United States Forest Service, boasts over 20 years of expertise in landscape ecology, ecological informatics, and conservation biology. His impressive scholarly record includes 340 publications, an h-index of 78, and 28,000+ citations. Cushman specializes in integrating spatial population modelling, landscape genetics, and machine learning to advance global biodiversity research. He has developed influential analytical tools, including FRAGSTATS, RMLANDS, CDPOP, UNICOR, sGD, Geomorphometry and Gradient Metrics Toolbox, and Pathwalker, shaping cutting-edge ecological research and conservation strategies. As an educator, he has conducted international workshops and graduate courses in 20+ countries, mentoring over 80 students and co-authoring papers with 50+. With a decade as the founding director of the Centre for Landscape Science, he received top research awards. His wealth of experience greatly enriches our project, benefiting from his knowledge, tools, and commitment to advancing ecological research and conservation.

Marco Oliverio: Professor of Zoology at Sapienza University of Rome (and Director of the Dept. BBCD). His publications total 141 (WoS), citations 2221, and his H index is 27. He has supervised 8 PhDs and over 90 bachelor's theses for the degree in Biology in the last 20 years. He has participated in national and international research projects for the last 30 years (1989-2023), including two LIFE projects totalling over €1 million in funding. He has led field expeditions and oceanographic cruises in the Mediterranean Sea, the Atlantic, Pacific and Indian oceans, and Antarctica. He is a member of the Scientific Italian Fauna Committee. Researcher. ID: F-2229-2010; ORCID: 0000-0002-0316-4364.

Mohammed Kallouche: Served as an assistant at Oran 1 University from 2011 to 2018, progressing to the position of lecturer 'B' from 2018-2021, and subsequently achieving the position of lecturer 'A' qualified to conduct research in the same university since 2021. His responsibilities included supervising 10 master's theses and overseeing 2 PhD students. With a focus on North African islands, he conducted field expeditions in Essaouira (Morocco), Habibas, Plane (Algeria), and Kuriat (Tunisia). Engaging in various national and international projects, including a collaboration with Museo Nacional de Ciencias Naturales, Madrid since 2013, he has contributed to the publication of 5 peer-reviewed papers and one book. ORCID: 0000-0003-1510-7583.

### 3.4. *Identification of critical points and contingency plan*.

1. Working with animal wildlife and in natural environments always represents a challenge. Thus, the first critical point could be sampling species.

Contingency Plan: We have selected more species than those initially planned for analysis. The sampling campaigns will serve as a guide to definitively decide on the four or five species for detailed analysis.

Given the uncertainties related to unforeseen events such as the COVID-19 pandemic and climate abnormalities, we have opted for a four-year project duration to ensure sufficient time to complete the sampling process, and posterior analyses.

At the MNCN collections there are specimens belonging to some of the species we would like to analyse, mainly from the Mediterranean, some of them sampled by members of our team. We could start with these samples.

2. Unexpected changes in environmental conditions during sampling campaigns.

Contingency Plan: Regularly monitor weather forecasts and adapt the sampling schedule accordingly. Maintain flexibility in the timeline to accommodate unforeseen circumstances.

3. Sample Preservation. Challenges in preserving samples adequately during fieldwork.

Contingency Plan: Implement robust sample preservation protocols, including backup preservation methods. Conduct periodic checks on sample quality during the sampling process.

4. Delays or complications in obtaining necessary permits for sampling in certain regions.

Contingency Plan: Initiate the permitting process well in advance, maintain open communication with relevant authorities, and have alternative sampling sites in case of unexpected permitting issues.

5. We plan to apply new techniques for a sort barcoding of the species, avoiding treating different cryptic or pseudocryptic species as only one. This methodology has been previously employed in the Molecular Systematics Laboratory of the MNCN, but not with invertebrates, which may pose specific challenges.

Contingency Plan: The barcoding process could alternatively be conducted using the traditional Sanger methodology, which has been extensively utilized in our laboratory and may provide a reliable alternative for species identification.

5. Technical difficulties or failures during the genomic sequencing process.

Contingency plan: Collaborate with a reliable sequencing service provider as a backup option. Additionally, having experience with different companies and establishing multiple contacts will be advantageous.

### *3.5. Previous results of the team in the theme of the proposal.*

The research team spearheading this proposal boasts an extensive track record in the realm of marine biodiversity, contributing significantly from diverse perspectives. The team's collective expertise spans biological feature description, species delimitation, phylogenetic inferences, evolutionary histories, to domains of populations and seascape genetics. Each member brings a wealth of experience, and their individual and collaborative contributions have been extensively documented in publications.

Over the years, the team has actively engaged in numerous research projects, fostering a collaborative spirit that has led to numerous joint publications.

The senior members of the research team have supervised nearly 20 PhD theses (with the majority being led by female researchers), all connected to the overarching themes of marine biodiversity, molecular tools, and evolutionary patterns. This rich supervisory experience underscores the team's commitment to nurturing the next generation of scholars and researchers.

Beyond research and supervision, the team actively engages in educational activities, delivering lectures and courses at various universities across different countries. The scope of teaching encompasses diverse facets of marine biodiversity and the molecular tools employed to unravel its variability and evolutionary intricacies.

Much of these results are consequence of our participation (and leadership) in various national and international projects (MINECO, MICINN, AEI, Fundación Biodiversidad-MITECO, BBVA, LIFE, European Commission, etc.), centred on marine diversity and conservation. These experiences further underscore our dedication to advancing scientific understanding.

### **3.6. *Human, material and equipment resources available for the execution of the Project*.**

At the Museo Nacional de Ciencias Naturales (MNCN), the team has access to much of the material needed for samplings and tissue sample collection. As previously noted, some of the

samples required for this project have already been collected and are currently stored in several museum collections, including the Malacology or Invertebrates collections (http://www.mncn.csic.es/es/colecciones).

The facilities at the MNCN also include the Molecular Systematics and Population Genetics Laboratory, led by the PI of this proposal, which is perfectly equipped for DNA/RNA extraction, amplification, sample preparation for sequencing (massive or Sanger), libraries preparation, and genotyping (http://www.mncn.csic.es/es/investigación/servicios-cientifico-tecnicos/laboratorio-de-sistematica-molecular). The lab is managed by an individual (included in the research team of this project) and five technicians with a long trajectory and knowledge of current cutting-edge molecular techniques.

The metadata projections and the coastal and other oceanographic profiles will be conducted at the MNCN BI lab (Laboratorio de Biogeografía Informática) (https://www.mncn.csic.es/es/investigaci%C3%B3n/servicios-cientifico-tecnicos/biogeografia-informatica).

For data processing, we also have access to the mentioned DRAGO facility. The SGAI (CSIC Deputy General Secretariat for Information Technology) offers facilities for high-performance computing (HPC) to its different centres, with the Drago supercomputer based on batch job management, in this case, relying on SLURM.

In our institution there is also a vice director in charge of overseeing the formation of new students. The incorporation of new members into our team would be supported by this vice director, who has already prepared a welcome protocol to receive them. Moreover, we have an equality committee to ensure that no discrimination exists.

Other MNCN or CSIC facilities will assist us in dissemination activities and media contact, including our Communication Service (Vice Directorate of Communication and Scientific Culture), as well as our exhibitions, where we can access the public.

## 4. *EXPECTED IMPACT OF THE RESULTS.*

### 4.1. *Expected impact on the generation of scientific-technical knowledge in the thematic area of the proposal.*

While genetic patterns in the Atlanto-Mediterranean transition have been assessed in previous studies, these endeavours typically focused on individual species and employed conventional or unique molecular markers. In this proposal, we aim to transcend these limitations by expanding our investigation to comprise on diverse non-model species, employing state-of-the-art methodologies that enable comprehensive genomic-level approach. The efficacy of our procedure extends not only to advanced molecular techniques but also encompasses cutting-edge data treatment methods. By contrasting genomic data with environmental parameters within a seascape framework, our results, as well as the developed protocols, are anticipated to contribute significant advancements to the thematic field in which this proposal is framed.

### 4.2. *Social and economic impact of the expected results.*

Biodiversity richness holds intrinsic value, but it also serves as a livelihood for numerous communities engaged in its more or less direct exploitation, providing sustenance for the general population. In our region, over 100,000 tons of marine products are consumed annually, translating to nearly €9 billion in revenue for the year 2019 alone (https://www.mapa.gob.es/es/alimentacion/temas/consumo-tendencias/informe2019_v2_tcm30-540250.pdf). Additionally, more than 900,000 tons are exported to other countries. In addition to the provisioning services offered by marine biodiversity (such as food and pharmacological resources), we must not overlook other non-material but equally crucial benefits tied to psychological well-being. These include aesthetic, inspirational, recreational, educational, cultural, and heritage values. The well-being of our coastlines also significantly influences one of our country's vital resources: tourism.

Socially, it is crucial to consider the circumstances of individuals whose livelihoods are intertwined with coastal resource exploitation, such as shellfish harvesting, a sector where women make noteworthy contributions.

The United Nations has declared the period 2021-2030 as "the United Nations Decade of Ocean Sciences for Sustainable Development". The objective is to motivate the international community to make efforts aimed at improving the health of the oceans by providing the science necessary for their sustainable development. In short, the goal is to guide us towards "the ocean we need for the future we want", as stated in the declaration. Achieving these objectives requires an urgent, coordinated effort from the entire society, including scientists, politicians, managers, sectors involved, and citizens as a whole.

Importantly, increasing knowledge about biodiversity patterns and factors influencing shaping processes can provide tools for better management, thereby positively impacting aspects like those mentioned above. Thus, contact with responsible of administrators, is one of our primary goals.

### 4.3. *Plan for scientific communication and internationalization of the results (indicate the forecast of open access publications).*

We intend to disseminate our findings within the scientific community through conventional channels, including publication in peer-reviewed, discipline-specific journals such as Journal of Biogeography, BMC Evolutionary Biology, Molecular Ecology, Scientific Reports, Evolution, Trends in Ecology and Evolution, Diversity and Distribution, PLoS ONE, Molecular Systems Biology, Frontiers in Ecology and Evolution, or Advances in Marine Biology. Additionally, we aim for publication in top-ranking generalist journals whenever possible. Anticipating approximately six publications, each studied species will merit at least one dedicated paper presenting its results. We plan to include a technical article introducing new methodological approaches and a comprehensive review discussing the knowledge and principal barriers in the Atlanto-Mediterranean transition. Aligned with current recommendations from the European Commission, as well as national and CSIC practices, we will prioritize submitting our scientific manuscripts to open-access journals.

Our findings will also be presented at four to six national and international congresses and workshops, covering diverse themes such as marine biology (e.g., International Conference on Oceanography and Marine Biology, EMBS), evolution (e.g., SESBE, ESEB) or those specifically focused on malacological (World Congress of Malacology, SEM Congress), crustacean (e.g., International Crustacean Congress), or coral (e.g., International Coral Reefs Congress) studies.

### 4.4. *Plan for dissemination of the results to the most relevant groups for the theme of the project and to society in general.*

We intend to provide environmental managers with reports that outline our main findings, particularly as they relate to the possible impacts derived from changes in species distribution and/or diversity, and possible strategies for their management and conservation. The members of this team are highly committed to species conservation, as evidenced by our previous collaborations with various sections of the Ministry of the Environment (under its different names).

As mentioned above, transmitting knowledge to society is a major goal of this team. We will achieve this through two main routes: direct contact with students (particularly at universities) and with the general public, through exhibitions (e.g., *El Museo Investiga*), outreach programmes and conferences (e.g., *La Mujer y la Niña en la Ciencia, Semana de la Ciencia, La Noche de los Investigadores, Científic@s en Prácticas*), and through the media, both written (e.g., peer-reviewed journals, popular science and outreach magazines and press releases to local, national and international media) and digital (e.g., project webpage, individual researchers' webpages and social networking sites). All of these actions represent an added

value to our research: we will be inviting society to participate in and shape our scientific activities.

### 4.5. *Summary´s management plan of the planned data*

We will deposit the raw sequences and databases (alignments and environmental data collection) into GenBank and other repositories (e.g., Animal-SNPATLAS, Zenodo), that can be mined by researchers of all disciplines (e.g., molecular biology, systematics, biogeography, ecology, developmental biology, physiology).

### 4.6. *Effects of gender inclusion in the content of the proposal*

Gender does not play a definitive role in our proposal, as determining the sex of the specimens sampled might not be feasible for most of the selected species. The sampling approach aims to be non-lethal, obtaining only the smallest possible piece of tissue. Identifying sex without additional data is quite challenging. Nonetheless, when available, the sex of samples will be annotated to confirm the presence or absence of data bias.

Regarding the team composition, gender balance is relatively equal (three women and five men), with the proposed Principal Investigator (PI) being a woman. Furthermore, we are seeking to incorporate two more individuals, one for hire and another as a predoctoral student. In the selection process for both positions, we will ensure equality and actively work to avoid any form of discrimination.

## *5. JUSTIFICATION OF THE REQUESTED BUDGET*

Three expenses determine the budget allocation: personnel hiring, SNP analyses, and sampling.

We propose hiring a bioinformatics expert to manage the substantial volume of data generated. This individual should possess expertise in molecular techniques, contribute to laboratory work, and be proficient in utilizing informatics tools for processing massive sequencing data, forming the foundation for population genomic analyses. With a four-year project timeline, we anticipate a thirty-eight months contract for this role, reserving the initial 10 months primarily for sampling, DNA extraction, and species delimitation, tasks well-suited for the existing team.

While cost-effective techniques will be implemented for barcoding, a significant portion of the budget will be allocated to obtaining raw SNP data. To optimize costs, library preparation will be conducted in-house, but sequencing expenses are unavoidable.

In addition to laboratory fees and essential materials (e.g., plastics, plates, enzymes, polymerase, primers), sampling constitutes a substantial budgetary component. Our extensive sampling plan covers diverse locations across the Iberian Peninsula, Balearic and Canary Islands, as well as Algeria and the Tyrrhenian Sea. Leveraging our research team's presence in these areas will contribute to cost savings by eliminating the need for external travel.

Other planned expenditures include manuscript publications and minor non-inventory laboratory supplies.

## 6. *TRAINING CAPACITY*

### 6.1. *Training program planned in the context of the requested project*

The research team possesses extensive experience in training researchers across various educational levels, having supervised or co-supervised the research projects of bachelor's, master's, and doctoral candidates.

The selection of predoctoral students will be based on their previous accomplishments, ensuring a non-discriminatory approach with regard to religion, gender, age, etc. Scholarships associated with this project will be open to both national and international applicants.

The specifics of a candidate's doctoral program will align with the guidelines of the registered university. We propose that the predoctoral student joins the Biology program at the Universidad Autónoma of Madrid, given its excellent reputation and our satisfactory experience with previous students there.

The predoctoral student will attend specific courses, including those provided at the MNCN (e.g., Advanced Statistical Analyses in R, *Curso Avanzado de SIG en ecología, conservación de la biodiversidad y del paisaje*), and the CSIC (e.g., *La carrera investigadora: de los inicios a la consolidación, Técnicas aplicadas de laboratorio, Introducción al procesado de datos de metabarcoding*).

Additional training, as needed, will be offered at other institutions through short courses and workshops. Due to the international nature of the team, exchanges or short stays between institutions and relevant centres will be encouraged to enhance various aspects of the candidate's training, including knowledge, experience, and networking. Two stages of around three months each are planned, allowing the candidates to spend time in centres of foreign collaborators or other institutions, such as the Natural History Museum of London supervised by Suzanne Williams.

Predoctoral students will gain teaching experience by participating in academic courses offered by team members, including undergraduate courses, seminar series, and labs. They will be encouraged to engage with the public through outreach programs and activities at the Museo Nacional de Ciencias Naturales. This involvement will enhance their public speaking skills, especially when addressing a lay audience, and underscore the importance of effective science communication. Under our guidance, they will present their research at national and international congresses and publish their findings in peer-reviewed journals. These experiences will equip them with the skills to present their results clearly and efficiently, maximizing their impact.

Moreover, the vice direction of the MNCN has prepared a welcome plan and a specific PhD training program for MNCN students. As part of this plan, all predoctoral students will spend dedicated time in various MNCN facilities, including collections, laboratories, exhibitions, and communication, to broaden their education.

### 6.2. *Theses completed or in progress within the scope of the research team (last 10 years).*

Systematics and phylogeography of the deep-sea coral *Desmophyllum dianthus* (Anthozoa: Hexacorallia): morphological and molecular evidences. Anna M. Addamo. Universidad Autónoma de Madrid. 28/11/2014. Graded with distinction *cum laude*. Co-supervised by Annie Machordom and Marco Taviani.

Caracterización molecular, genética poblacional y relación biológica del nemertino *Malacobdella arrokeana* endocomensal de la almeja gigante *Panopea abbreviata*. José Elías Fernández Alfaya. Facultad de Ciencias Naturales. Universidad Nacional de la Patagonia San Juan Bosco. 27/3/2015. Graded with distinction. Co-supervised by Gregorio Bigatti and Annie Machordom.

Avances en el catálogo mitogenómico de Mollusca. Estructura génica y relaciones filogenéticas de Caenogastropoda. David Osca Ferriol. Universidad Autónoma de Madrid. 2015. Graded with distinction *cum laude*. Co-supervised by Rafael Zardoya and José Templado.

Mitogenómica y filogenia de linajes de gasterópodos altamente diversificados (Vetigastropoda, Neritimorpha y Conoidea). Juan Esteban Uribe Arboleda. Universidad Autónoma de Madrid. 2015. Graded with distinction *cum laude*. Co-supervised by Rafael Zardoya and José Templado.

Biología de la Conservación de las comunidades de gorgonias tropicales en el Pacífico oriental (Ecuador). Mª Mar Soler Hurtado. Universidad Menéndez Pelayo. 7/10/2016. Graded with distinction *cum laude*. Co-supervised by Pablo José López González and Annie Machordom.

Conectividade genética de *Panulirus echinatus* Smith, 1869 (Decapoda: Palinuridae) nas ilhas oceânicas do Atlântico. Juliana de Carvalho Gaeta. Universidad Federal do Ceará. 9/3/2018. Graded (there is no distinction in Brazil). Co-supervised by Raúl Cruz Izquierdo, Rodrigo Maggioni and Annie Machordom.

Biodiversidad, biogeografía y patrones evolutivos en crustáceos (Anomura, Galatheoidea) de zonas tropicales y templadas. Paula Carolina Rodríguez Flores. Universidad de Barcelona. 5/2/2021. Graded with distinction *cum laude*. Co-supervised by Enrique Macpherson and Annie Machordom.

Genetic structure and connectivity in coastal marine invertebrate species. Violeta López Márquez. Universidad Autónoma de Madrid. 16/4/2021. Graded with distinction *cum laude*. Co-supervised by Annie Machordom and José Templado.

Historia evolutiva y demográfica de las especies del género *Asterina* Nardo, 1834 (Echinodermada, Asteroidea, Asterinidae) en el Atlántico nororiental y Mediterráneo. Iván Acevedo García. Universidad Autónoma de Madrid. 12/4/2023. Graded with distinction *cum laude*. Co-supervised by Annie Machordom and José Templado.

Two more PhD theses are being supervised by the research team.

### 6.3. *Scientific or professional development of graduate doctors*.

Anna M. Addamo. Post-doctoral researcher at the Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway.

José Elías Fernández Alfaya. Researcher at the IBIOMAR, CONICET, Argentina.

David Osca Ferriol. Post-doctoral researcher, University of Las Palmas, Gran Canaria.

Juan Esteban Uribe Arboleda. Post-doctoral researcher at the MNCN-CSIC.

Mª Mar Soler Hurtado. Currently undergoing competitive exams for the position of Museum Curator.

Juliana de Carvalho Gaeta. Researcher at LABOMAR, Universidade Federal do Ceará, Brazil.

Paula Rodríguez Flores. Post-doctoral researcher at the Smithsonian National Museum of Natural History, Washington DC, USA.

Violeta López Márquez. Post-doctoral researcher at the University of Aruba, Aruba and KU Leuven University, Belgium.

Iván Acevedo García. Lab manager at the MNCN-CSIC.

**References**:


Alberto, F., Massa, S. I., Manent, P., Diaz-Almela, E., Arnaud-Haond, S., Duarte, C. M. & Serrão, E. A. (2008). Genetic differentiation and secondary contact zone in the seagrass *Cymodocea nodosa* across the Mediterranean–Atlantic transition region. Journal of Biogeography, 35(7), 1279-1294.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A. & De Clerck, O. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. Global Ecology and Biogeography, 27(3), 277-284.

Beerli, P. & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proceedings of the National Academy of Sciences, 98(8), 4563-4568.

Boisvenue, C. & Running, S. W. (2006). Impacts of climate change on natural forest productivity—evidence since the middle of the 20th century. Global Change Biology, 12(5), 862-882.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... & Drummond, A. J. (2019). BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Computational Biology, 15(4), e1006650.

Bowen, B. W., Rocha, L. A., Toonen, R. J. & Karl, S. A. (2016). The origins of tropical marine biodiversity. Trends in Ecology & Evolution, 31(11), 820-832.

Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics, 27(15), 2156-2158.
de Jonge, V. N., Pinto R., Turner R. K. (2012). Integrating ecological, economic and social aspects to generate useful management information under the EU Directives' "ecosystem approach". Ocean & Coastal Management, 68, 169-188.
Deli, T., Fratini, S., Ragionieri, L., Said, K., Chatti, N. & Schubart C. D. (2016). Phylogeography of the marbled crab *Pachygrapsus marmoratus* (Decapoda, Grapsidae) along part of the African Mediterranean coast reveals genetic homogeneity across the Siculo–Tunisian Strait versus heterogeneity across the Gibraltar Strait. Marine Biological Research, 12, 471-487.
Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution, 29(8), 1969-1973.
Falush, D., Stephens, M. & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 164(4), 1567-1587.
Fernández, M., Heras, S., Maltagliati, F., Turco, A. & Roldan, M. (2011). Genetic structure in the blue and red shrimp *Aristeus antennatus* and the role played by hydrographical and oceanographical barriers. Marine Ecology Progress Series, 421, 163-171.
Foll, M. & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics, 180(2), 977-993.
Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology, 3(5), 294-299.
Fretter, V. & Graham, A. (1976). The Prosobranch Molluscs of Britain and Denmark. Part 1 - Pleurotomariacea, Fissurelacea and Patellacea. Journal of Molluscan Studies, Suppl. 1, 1-37.
Fretter, V. & Graham, A. (1981). The Prosobranch Molluscs of Britain and Denmark. Part 6 - Cerithiacea, Strombacea, Hipponicacea, Calyptraeacea, Lamellariacea, Cypraeacea, Naticacea, Tonnacea and Heteropoda. Journal of Molluscan Studies, Suppl. 9, 285-363.
Fruciano, C., Hanel, R., Debes, P.V., Tigano, C. & Ferrito, V. (2011). Atlantic-Mediterranean and within Mediterranean molecular variation in *Coris julis* (L. 1758) (Teleostei, Labridae). Marine Biology, 158, 1271-1286.
Galarza, J. A., Carreras-Carbonell, J., Macpherson, E., Pascual, M., Roques, S., Turner, G. F. & Rico, C. (2009). The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. Proceedings of the National Academy of Sciences, 106(5), 1473-1478.
García-Dorado, A. & Caballero, A. (2021). Neutral genetic diversity as a useful tool for conservation biology. Conservation Genetics, 22, 541-545.
García-Merchán, V. H., Robainas-Barcia, A., Abelló, P., Macpherson, E, Palero, F., García-Rodríguez, M., Gil de Sola, L. & Pascual, M. (2012). Phylogeographic patterns of decapod crustaceans at the Atlantic-Mediterranean transition. Molecular Phylogenetics and Evolution, 62, 664-672.
Goffredo, S., Airi, V., Radeti, J. & Zaccanti, F. (2006). Sexual reproduction of the solitary sunset cupcoral *Leptopsammia pruvoti* (Scleractinia: Dendrophylliidae) in the Mediterranean. 2. Quantitative aspects of the annual reproductive cycle. Marine Biology 148,: 923-931.
González-Castellano, I., González-López, J., González-Tizón, A.M. & Martínez-Lage, A. (2020). Genetic diversity and population structure of the rockpool shrimp *Palaemon elegans* based on microsatellites: evidence for a cryptic species and differentiation across the Atlantic-Mediterranean transition. Scientific Reports, 10, 10784.
González-Gordillo, J. I., Cuesta, J. A. & Rodríguez, A. (1996). Studies on the larval development of northeastern Atlantic and Mediterranean Porcellanidae (Decapoda, Anomura). I - Redescription of the larval stages of *Porcellana platycheles* (Pennant, 1777) reared under laboratory conditions. Helgoländer Meeresuntersuchungen, 50(4), 517–531.
Halpern, B. S., Frazier, M., Potapenko, J. et al. (2015). Spatial and temporal changes in cumulative human impacts on the world's ocean. Nature Communications, 6, 7615.
Heras, S., Planella, L., García-Marín, J. L., Vera, M. & Roldán, M. I. (2019). Genetic structure and population connectivity of the blue and red shrimp *Aristeus antennatus*. Scientific Reports, 9, 13531.
Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N. & Vellend, M. (2008). Ecological consequences of genetic diversity. Ecology Letters, 11(6), 609-623.
Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics, 27(21), 3070-3071.

Karl, T. R. & Trenberth, K. E. (2003). Modern global climate change. Science, 302(5651), 1719-1723.
Katoh, K. & Toh, H. (2008). Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. BMC bioinformatics, 9(1), 1-12.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.
Krijgsman, W., Hilgen, F.J., Raffi, I., Sierros, F.J. & Wilson, D.S. (1999). Chronology causes and progression of the Messinian Salinity Crisis. Nature, 400, 652–655.
Kukkala, A. S. & Moilanen, A. (2013). Core concepts of spatial prioritization in systematic conservation planning. Biological Reviews, 88(2), 443-464.
Landguth, E. L., Cushman, S. A., Schwartz, M. K., McKelvey, K. S., Murphy, M. & Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. Molecular Ecology, 19(19), 4179-4191.
Landguth, E. L., Fedy, B. C., Oyler-McCance, S. J., Garey, A. L., Emel, S. L., Mumma, M., ... & Cushman, S. A. (2012). Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Molecular Ecology Resources, 12(2), 276-284.
Linck, E. & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. Molecular Ecology Resources, 19(3), 639-647.
López-Márquez, V., Acevedo, I., Manjón-Cabeza, M. E., García-Jiménez, R., Templado, J. & Machordom, A. (2018). Looking for morphological evidence of cryptic species in *Asterina* Nardo, 1834 (Echinodermata: Asteroidea). The redescription of *Asterina pancerii* (Gasco, 1870) and the description of two new species. Invertebrate Systematics, 32, 505-523.
Losada, I. J., Camus, P., Mösso, C., Izaguirre, C. & Cáceres, I. (2014). Sea-level projections for the Mediterranean coast of the Iberian Peninsula. Climatic Change, 122(1-2), 157-171.
Luypaert, T., Hagan, J. G., McCarthy, M. L. & Poti, M. (2020). Status of Marine Biodiversity in the Anthropocene. In: Jungblut, S., Liebich, V., Bode-Dalby, M. (Eds.) YOUMARES 9 - The Oceans: Our Research, Our Future. Springer, Cham. Pp. 57–82
Machordom, A., Araujo, R., Erpenbeck, D., Ramos, M. A. & Thorpe, J. P. (2003). Phylogeography and conservation genetics of endangered European Margaritiferidae (Bivalvia: Unionoidea). Biological Journal of the Linnean Society, 78(2), 235-252.
Malomane, D. K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A. R., Simianer, H. & Weigend, S. (2018). Impact of missing marker data on the accuracy of genomic selection. Journal of Animal Breeding and Genetics, 135(4), 309-320.
Maltagliati, F., Di Giuseppe, G., Barbieri, M., Castelli, A. & Dini, F. (2010). Phylogeography and genetic structure of the edible sea urchin *Paracentrotus lividus* (Echinodermata: Echinoidea) inferred from the mitochondrial cytochrome b gene. Biological Journal of the Linnean Society, 100, 910-923.
McCauley, D. J., Pinsky, M. L., Palumbi, S., Estes, J. A., Joyce, F.H. & Warner, R. R. (2015). Marine defaunation: animal loss in the global ocean. Science, 347 (6219), 1255641.
McRae, B. H. (2006). Isolation by resistance. Evolution, 60(8), 1551-1561.
Nanninga, G. B. & Berumen, M. L. (2014). The role of individual variation in marine larval dispersal. Frontiers of Marine Science, 1(71), 1-17.
Ojeda, V., Serra, B., Lagares, C., Rojo-Francàs, E., Sellés, M., Marco-Herrero, E., García, E., Farré, M., Arenas, C., Abelló, P. & Mestres, F. (2022). Interannual fluctuations in connectivity among crab populations (*Liocarcinus depurator*) along the Atlantic-Mediterranean transition. Scientific Reports, 12, 9797.
Parmesan, C. & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. Nature, 421(6918), 37-42.
Pascual, M., Rives, B., Schunter, C. & Macpherson, E. (2017). Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. PLoS ONE 12(5), e0176419.
Patarnello, T., Volckaert, F. A. M. J. & Castilho, R. (2007). Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? Molecular Ecology, 16(21), 4426-4444.
Pellerito, R., Arculeo, M. & Bonhomme, F. (2009). Recent expansion of Northeast Atlantic and Mediterranean populations of *Melicertus (Penaeus) kerathurus* (Crustacea: Decapoda). Fisheries Science, 75, 1089-1095.
Petkova, D., Novembre, J. & Stephens, M. (2015). Visualizing spatial population structure with estimated effective migration surfaces. Nature genetics, 48(1), 94-100.
Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155(2), 945-959.

Privé, F., Luu, K., Vilhjálmsson, B. J. & Blum, M. G. (2020). Making the most of Clines: inference of local adaptation from joint SNP and allele frequency data. Genetics, 215(2), 611-631.
Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology, 21(8), 1864-1877.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics, 81(3), 559-575.
Puritz, J. B., Hollenbeck, C. M. & Gold, J. R. (2014). dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ, 2, e431.
Reid, N. M. & Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evolutionary Biology, 12(1), 1-14.
Repullés, M., López-Márquez, V., Templado, J., Taviani, M. & Machordom, A. 2022. Genetic structure of the endangered coral *Cladocora caespitosa* matches the main bioregions of the Mediterranean Sea. Frontiers in Genetics, 13, 889672.
Rey-Iglesia, A., Gopalakrishan, S., Caroe, C., Alquezar-Planas, D. E., et al. (2019). MobiSeq: de novo SNP discovery in model and non-model species through sequencing the flanking region of transposable elements. Molecular Ecology Resources, 19, 512-525.
Rilov, G. & Crooks, J. A. (2009). Biological invasions in marine ecosystems: ecological, management, and geographic perspectives. Springer Science & Business Media.
Rochette, N. C., Rivera-Colón, A. G. & Catchen, J. M. (2019). Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular Ecology, 28(21), 4737-4754.
Rodríquez, A. (1997). Larval and postlarval development of *Pisa tetraodon* (Pennant, 1777) (Decapoda: Majidae) reared in the laboratory. Journal of Plankton Research, 19(1), 29-41.
Rodríguez-Flores, P. C., Macpherson, E., Buckley, D. & Machordom, A. (2019). High morphological similarity coupled with high genetic differentiation in new sympatric species of coral-reef squat lobsters (Crustacea: Decapoda: Galatheidae). Zoological Journal of the Linnean Society 185, 984-1017.
Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ, 4, e2584.
Ronquist, F., Teslenko, M., van der Mark, P., et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space**.** Systematic Biology, 61, 539-542.
Sá-Pinto, A., Branco, M. S., Alexandrino, P. B., Fontaine, M. C. & Baird, S. J. E. (2012). Barriers to gene flow in the marine environment: insights from two common intertidal limpet species of the Atlantic and Mediterranean. PLoS ONE, 7, 50330.
Schloerke, B., Crowley, J., Cook, D., Briatte, F., Marbach, M., Thoen, E., ... & Larmarange, J. (2018). GGally: Extension to 'ggplot2'. R package version 1.4. 0.
Sutherland W. J., Adams W. M., Aronson R. B. et al. (2009). One Hundred Questions of Importance to the Conservation of Global Biological Diversity. Conservation Biology, 23(3), 557-567.
Taviani, M. (2002). The Mediterranean benthos from late Miocene up to present: ten million years of dramatic climatic and geologic vicissitudes. Biologia Marina Mediterranea, 9, 445-463.
Templado, J. (2014). Future trends of Mediterranean biodiversity. In: The Mediterranean Sea. Its history and present challenges (Goffredo, S. & Dubinsky, Z. Eds.). Springer, pp. 479-498.
Torfstein, A. & Steinberg, J. 2020. The Oligo–Miocene closure of the Tethys Ocean and evolution of the proto-Mediterranean Sea. Scientific Reports, 10, 13817.
Volosciuk, L. C., Uehara-Prado, M., Diniz-Filho, J. A. F. & Terribile, L. C. (2016). Climate change and the eco-evolutionary responses of aquatic vertebrates in Cerrado streams: an uncertain future. PloS ONE, 11(1), e0145050.
Waples, R. S. & Do, C. (2010). Linkage disequilibrium estimates of contemporary N e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications, 3(3), 244-262.
Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), 2869-2876.