



OPEN Genomic population structure and adaptation of eastern Mediterranean loggerhead (*Caretta caretta*) nesting areas

Barbanti A.^{1,14}, Pegueroles C.^{2,14}, Margaritoulis D.³, Rees A.F.³, Turkozan O.⁴, Ulger C.⁴, Snape R.T.E.^{5,6}, Broderick A.C.⁶, Demetropoulos A.⁷, Demetropoulos S.^{7,8}, Hamza A. A.^{9,10}, Levy Y.^{11,12,13}, Pascual M.^{1,15} & Carreras C.^{1,15}✉

Considering the current scenario of biodiversity crisis, it is crucial to understand how populations are structured and their adaptation drivers to aid management and conservation. In this context, sea turtles are predicted to be heavily impacted by climate change, causing feminisation of the populations and decreasing hatchling survival rates. These effects can be exacerbated in the Mediterranean, as it is predicted that global warming will exceed global trends. Here we genotyped 243 individuals from 11 nesting populations of the Mediterranean loggerhead sea turtle (*Caretta caretta*) using 2bRAD, obtaining a panel of 5,998 curated SNPs. We then combined this genomic information of with environmental, behavioural and reproductive data to investigate their population structure and understand their mechanisms of adaptation. We found substantial genetic differentiation among almost all rookeries, with a major break separating Greek locations, and a second break separating Libya from the Levantine locations. This genetic structuring, caused by male and female philopatry, was also shaped by adaptation with strong genomic signals associated with all phenotypic and environmental variables tested. A conservation prioritisation analysis, to identify the areas with higher priority for protection, concluded that protecting all rookeries should be considered a priority, as some areas are key to retain the genetic diversity and adaptive potential of the Mediterranean region, while others play a crucial role maintaining connectivity that is fundamental for the long-term survival of this vulnerable species. This case study exemplifies the potential of genomics to unveil key information for species of conservation concern.

Keywords Adaptation, Conservation genomics, Conservation prioritisation, Highly migratory species, Climate change, Environment-phenotype-genotype approach

Marine ecosystems are some of the environments most affected by anthropogenic pressures and climate change¹. Consequently, the study and monitoring of marine and coastal environments have therefore become crucial for the conservation of wildlife². Accurate knowledge of population structure is fundamental to designating management units (MUs) and to detecting genetically isolated populations, which may require independent conservation actions³. The design and conservation planning of MUs does not depend only on the target

¹Department of Genetics, Microbiology, and Statistics, and IRBio, University of Barcelona, Av.Diagonal 643, Barcelona, Spain. ²Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain. ³ARCHELON, the Sea Turtle Protection Society of Greece, Solonos 113, 10678 Athens, Greece. ⁴Faculty of Science, Department of Biology, Aydın Adnan Menderes University, 09010 Aydın, Türkiye. ⁵Society for Protection of Turtles, Kyrenia, Cyprus. ⁶Centre for Ecology and Conservation, University of Exeter, Penryn, UK. ⁷Cyprus Wildlife Society, Nicosia, Cyprus. ⁸Mediterranean Turtle Research and Conservation Society (MedTRACS), P.O.Box 61267, 8132 Paphos, Cyprus. ⁹Biology Department, Faculty of Education, University of Tripoli, Bab Tajura, Tripoli, Libya. ¹⁰Ministry of Environment, Libyan Sea Turtle Program, Janzur Road, Tripoli, Libya. ¹¹Israel Sea Turtle Rescue Center, National Nature and Parks Authority, Gan Leumi Beit Yannay, Kfar Vitkin, Israel. ¹²Department of Marine Biology, Leon H. Charney School of Marine Sciences, University of Haifa, 3498838 Haifa, Israel. ¹³Morris Kahn Marine Research Station, University of Haifa, 3498838 Haifa, Israel. ¹⁴Barbanti A. and Pegueroles C. are authors equally contributed to this work. ¹⁵Pascual M. and Carreras C. are authors jointly supervised this work. ✉ email: carreras@ub.edu

population structure but considers factors that shape and impact population structure, connectivity and self-sustainability. For instance, adaptability of a species to a constantly changing environment is critical for conservation planning since it can shape population structure. Similarly, effective population size (N_e) is a crucial factor for populations' self-sustainability, as small N_e values increase populations' vulnerability and extinction risk. Genetic assessments can provide information on these factors to advise and assist in the conservation planning of MUs. Conservation of marine species should therefore ideally include a complete genetic assessment of population structure, including the role of adaptation, and estimates of effective population sizes.

The Mediterranean Sea is a global marine biodiversity hotspot, but is also facing concerning biodiversity loss^{4,5}. In this sea, surface temperatures are increasing faster than in the rest of the world⁶ and invasive species are increasing their expansion more severely in the Mediterranean^{7,8}. Ocean acidification is also playing its part in depleting marine biodiversity⁹, as well as fisheries interaction^{10,11}. The Mediterranean loggerhead turtle populations are considered a distinct regional management unit (RMU)^{12,13}, with rookeries mostly located in the eastern basin¹⁴, where the water is warmer, and temperatures have been historically more suitable for nesting¹⁵. In recent years, the Mediterranean loggerhead turtle populations have increased, but they are still dependent on conservation actions¹⁶. The two main conservation concerns linked to Mediterranean populations are habitat degradation, including climate change, and interaction with fisheries¹⁴. Climate change is considered a global threat to marine turtles^{17,18}, and habitat alterations due to sea level rise^{19,20} causing phenological shifts as a response^{21,22}. Current Mediterranean nesting areas have been predicted to be heavily impacted^{23,24}, resulting in feminisation of populations due to its temperature-dependent sex determination and a decrease in hatchling viability and hatching success^{25,26}. Fisheries interaction is still one of the main threats to juveniles and adults in foraging areas¹⁴, which can be located thousands of kilometres away from nesting areas. This is the case in the western Mediterranean, where turtles from Mediterranean and Atlantic origins coexist^{27,28}. Consequently, assessing the area of origin (e.g., Mediterranean or Atlantic) of turtles captured at sea is crucial to identify threats to the populations of origin²⁹.

For the Mediterranean loggerhead turtle, previous studies using mitochondrial DNA and nuclear microsatellites revealed a deep global structuring favoured by female philopatry^{30,31}. However, some key questions remain unresolved due to the limitations of the markers used, such as the fine-scale population structure and connectivity, the role of male philopatry in population structure³², and the role of adaptation. High-throughput sequencing technologies can be applied to non-model species at the individual level, to answer ecological and conservation dilemmas left so far unanswered³³. A growing number of population genomic studies in the Mediterranean Sea have focused on non-model marine species, providing both an improved resolution for population structure³⁴ and genome-wide signals of adaptation^{35–37}. Understanding the genetic structure and dynamics of the nesting populations in this area is critical to inform conservation management in both nesting and foraging areas in order to improve prioritisation of conservation actions.

Here, we aim to identify the diversity and population structure of loggerhead turtle Mediterranean rookeries from a genomic perspective by combining genomic, phenotypic and environmental information. In this study, we had the following specific objectives: i) assess the diversity and effective population size of Mediterranean rookeries, ii) identify population differentiation and connectivity between rookeries, iii) understand the role of environmental factors shaping population genomic structure and local adaptation and iv) identify the importance of each population within the context of regional management, to aid prioritise conservation actions.

Results

Sampling, data collection and genomic diversity indexes

Samples were taken from Zakynthos (ZAK; $N = 16$), Kyparissia (KYP; $N = 25$), Rethymno (RET; $N = 23$), Messara (MES; $N = 18$), Dalyan (DAL; $N = 24$), Belek (BEL; $N = 24$), Akamas (AKA; $N = 25$), Alagadi (ALA; $N = 25$), Lebanon (LEB; $N = 19$), Israel (ISR; $N = 21$) and Libya (LIB; $N = 23$). Illumina sequencing resulted in an average number of 6.2 million raw reads per individual and 4.8 million trimmed reads per individual. After filtering and SNP calling, 5,998 SNPs were kept, with a mean coverage of 25.7 sequences per locus, a good sequencing depth for reliable genotyping³⁸. In terms of number of reads, mapping efficiency and missingness levels, the results were similar to previous studies using samples from living individuals³⁹. The SNP loci were distributed along the whole genome, with a highly significant correlation between chromosome size and number of SNPs (Spearman rho, $R = 0.98$, $p < 0.001$, Figure S1). The diversity indexes were relatively similar across populations (Table S1), with H_o ranging from 0.186 (KYP) to 0.200 (ZAK), H_e ranging from 0.193 (KYP) to 0.216 (LIB), and A_r ranging from 1.193 (KYP) to 1.201 (ISR and LIB). The values of inbreeding were always significantly different from zero and always positive except for Zakynthos, which was negative.

The effective population sizes (N_e) of our sampling locations ranged from 111 individuals (ISR) to 3,736.5 (KYP) as shown in Table S2. The estimated effective population size was not significantly correlated to the inferred number of adults in any of the two temporal periods considered, although the best correlation was found in the more recent years (Figure S2). Mean clutch sizes were variable among localities, with the highest values located in Greece (ZAK, KYP, RET and MES, Fig. 1b, Table S3). All temperature variables had a higher impact than salinity variables on the two first components of the environmental PCA (Fig. 1c). The first axis was associated to the minimum temperature and separated Zakynthos from the remaining populations. The second axis distributed the localities following a gradient of maximum temperature. The particle modelling data indicated that hatchlings of the Greek populations used mainly the Ionian and Adriatic seas, hatchlings from Libya used the South-Central Mediterranean, and those from the remaining populations used mainly the Levantine region and the Aegean Sea (Fig. 1d, Table S3). Finally, the isotopic signal, related to the female food intake prior to the nesting season, indicated that females from Zakynthos and Kyparissia feed on prey rich in $\delta^{15}N$ and with low values of $\delta^{13}C$, females from Türkiye (DAL and BEL) feed on prey rich in $\delta^{13}C$, and females

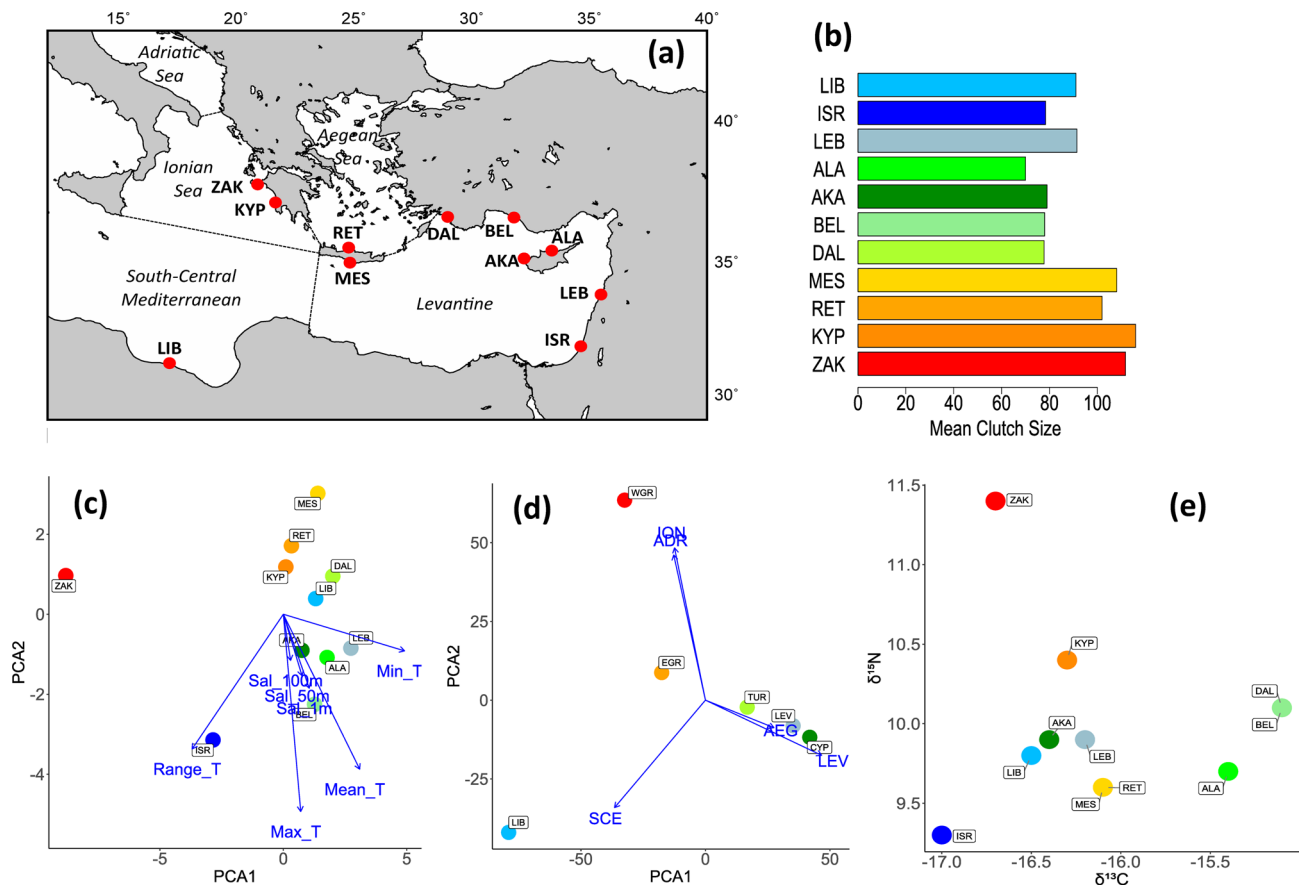


Fig. 1. Sampling and data collection. **(a).** Location of the nesting populations considered in this study: Zakynthos (ZAK; N = 16), Kyparissia (KYP; N = 25), Rethymno (RET; N = 23), Messara (MES; N = 18), Dalyan (DAL; N = 24), Belek (BEL; N = 24), Akamas (AKA; N = 25), Alagadi (ALA; N = 25), Lebanon (LEB; N = 19), Israel (ISR; N = 21) and Libya (LIB; N = 23). The dashed lines delimit the marine zones (in italics) described as potential hatchling foraging areas in the literature⁵¹ **(b).** Mean clutch size of each sampled population (mean number of eggs per nest). **(c)** Principal Component Analysis (PCA) of the environmental variables obtained per population. The contribution of each specific variable is included as a vector whose length represents its weighted contribution to the two first axis of the PCA (salinity at 1 m, 50 m and 100 m; and mean, range, maximum and minimum temperatures). **(d)** Principal Component Analysis (PCA) of the foraging areas where the hatchlings disperse during their first months as inferred using particle modelling. The contribution of each specific foraging area is included as a vector whose length represents its weighted contribution to the two first axis of the PCA (ADR: Adriatic Sea; ION: Ionian Sea; SCE: South-Central Mediterranean; AEG: Aegean Sea; LEV: Levantine of the Mediterranean). **(e)** Mean stable isotope signature ($\delta^{15}\text{N}$ and $\delta^{13}\text{C}$) of hatchlings found in the sampling locations that reflect the food intake of nesting females prior to nest. Map created using MAPTOOL (SEATURTLE.ORG Maptool. 2002. SEATURTLE.ORG, Inc. <http://www.seaturtle.org/maptool/> (accessed 12/03/2024) which uses GMT (the Generic Mapping Tool) v5.4.3¹¹⁵).

from Israel feed on prey with low values of both $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$. The females from the remaining populations stayed in intermediate positions (Fig. 1e, Table S3).

Population structure

The number of genetic clusters was three ($K=3$) for DAPC using the Bayesian Information Criterion (BIC) and the $\ln(K)$ of STRUCTURE, while delta K suggested $K=2$ (Figure S3). The plots of the averaged values of STRUCTURE at $K=2$ and $K=3$ and the self-assignment values of the DAPC (Figure S4) suggested a first partition between the Greek populations (ZAK, KYP, RET and MES) and a second differentiation of Libya from the populations in the eastern Mediterranean (DAL, BEL, AKA, ALA, LEB and ISR). Pairwise F_{ST} analyses showed that almost all populations are significantly different from each other (Fig. 2a, Table S4) with two groups of exceptions. On one hand, Zakynthos was not statistically different to Kyparissia, in Greece. On the other hand, all pairwise comparisons involving populations from Türkiye and the island of Cyprus (DAL, BEL, AKA and ALA) were not significant, except for the comparison between Belek and Akamas. The network analysis also indicated higher connectivity among the nesting areas of Türkiye and those in the Island of Cyprus and also between Zakynthos and Kyparissia (Fig. 2b). Moreover, populations of Rethymno, Messara and Akamas were revealed as stepping stones among the two different regional networks detected (ZAK-KYP and DAL-

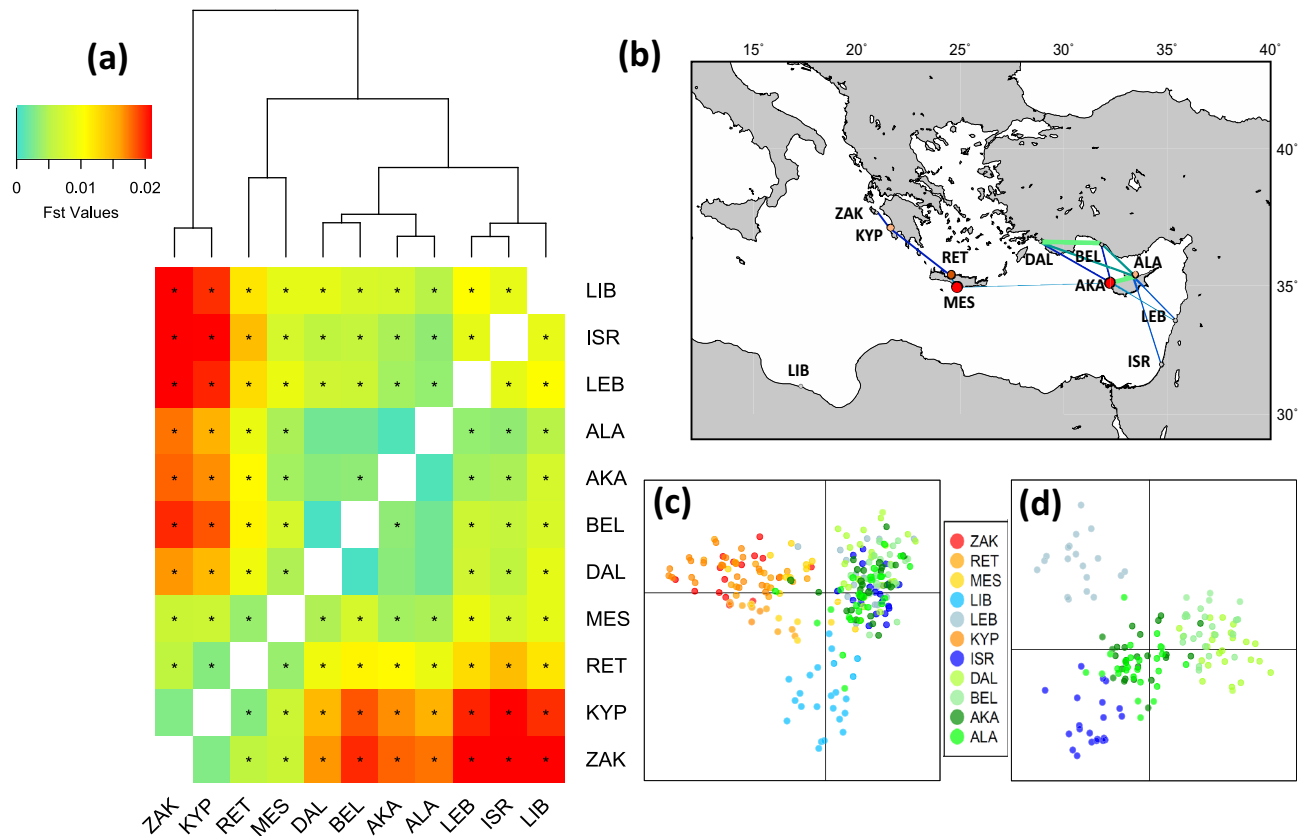


Fig. 2. Population structure of the loggerhead sea turtle (*Caretta caretta*) in the Mediterranean. **(a)** Heatmap and dendrogram based on F_{ST} values among all population pairs. The asterisks in each cell indicate the pairwise comparisons that are significantly different after FDR correction. **(b)** Connectivity network of populations overlapped on the geographic map. The size of the circles of the localities and their colours indicate their centrality of betweenness (from white for low values to red for high values). Line thickness and colour reflect the strength of the edges, and thus the connectivity among localities (from light green thick lines representing high connectivity to dark blue thin lines representing low connectivity). **(c)** Results of the DAPC with all individuals using the populations as prior groupings. **(d)** Results of the DAPC with only the individuals from the eastern Mediterranean (DAL, BEL, AKA, ALA, LEB and ISR). In c and d each individual is represented by a circle colour coded by locality. Population acronyms as in Fig. 1 and Table S1. Map created using MAPTOOL (SEATURTLE.ORG Maptool. 2002. SEATURTLE.ORG, Inc. <http://www.seaturtle.org/maptool/> (accessed 12/03/2024) which uses GMT (the Generic Mapping Tool) v5.4.3¹¹⁵).

BEL-AKA-ALA) while Libya appeared to be isolated from the remaining populations. In accordance with the network analysis, the Discriminant Analysis of Principal Components (DAPC) showed three clouds of individuals (Fig. 2c), including Greece (ZAK, KYP, RET and MES), Libya (LIB) and the eastern Mediterranean populations (DAL, BEL, AKA, ALA, LEB and ISR). A DAPC including only the individuals from the eastern Mediterranean (Fig. 2d) showed three marginal clouds: Lebanon, Israel and the Turkish populations (DAL and BEL) while the individuals from the Cyprus Island (AKA and ALA) clustered in the middle of the DAPC. No related nests were found among different locations, indicating that none of the females nor the males contributed to sampled nests from different nesting beaches. Mantel tests between genetic and geographic distances were significant when Euclidean distances were considered (p -value = 0.002) and when coastline distances (Figure S5) were used (p -value = 0.003). However, IBD plots showed a small cloud of points with high distance and relatively low differentiation (Figure S6) involving always the population of Libya, indicating that this population is less differentiated from the others than expected by distance.

Signals of adaptation

We found 97 outlier SNP loci using Pcadapt and 73 using Bayescan, with a total of 16 shared by the two programs. The structure found with the set of 16 loci and with the set of 5,844 neutral population loci was similar, but with F_{ST} values generally an order of magnitude higher in the set of outlier loci in relation to the set of neutral (Figure S7, Table S5) and total loci (Fig. 2a, Table S4). The four Redundancy Analyses provided different population distributions based on the predictors used in each analysis, with the two first RDA components explaining a high percentage of the genomic variability (Fig. 3) and globally matching the variability found in the phenotypic and environmental variables (Fig. 1b-e). A total of 237 association SNP loci (RDA_CLUTCH set of loci) were found associated to clutch size and geographic location (Fig. 3a, Table S6), with the first axis (associated to clutch size)

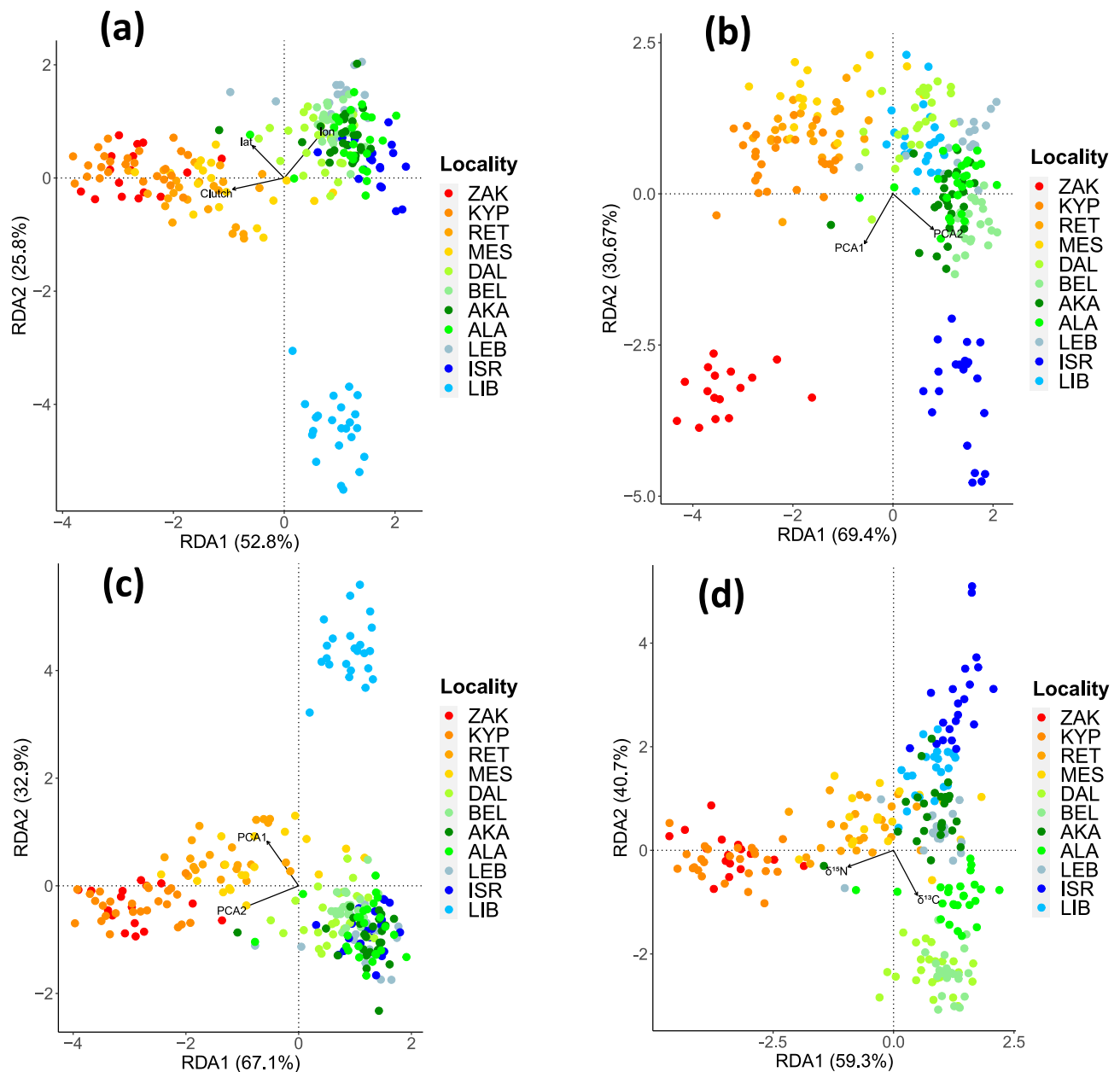


Fig. 3. Genomic association analysis to phenotypic and environmental data. Redundancy Analysis (RDA) performed with phenotypic and environmental data and SNP loci. Coloured points represent individual sea turtles coded by locality. Vectors indicate the predictors for the two first RDA components. a) RDA using clutch size and geographic location (Fig. 1b) b) RDA using the two first PCAs of the environmental data (Fig. 1c) b) RDA using the two first PCAs of the particle dispersion modelling data (Fig. 1d) and c) RDA using the stable isotopes data (Fig. 1e). Population acronyms as in Fig. 1 and Table S1.

separating the Greek populations from the remaining populations, and Libya separated in the second axis. The combination with environmental variables (Fig. 3b) yielded 230 association SNP loci (RDA_ENV set of loci, Table S6) with a separation between the Greek populations and the remaining populations on the first axis and separating Zakynthos and Israel from the remaining populations on the second axis. The RDA with the particle simulation, as an indicative of hatchling dispersal (Fig. 3c), yielded 247 association SNP loci (RDA_DISP set of loci, Table S6), with the first axis separating again the Greek populations from the others and the second axis differentiating Libya. Finally, the RDA with the stable isotope information, indicating the foraging signature of females prior to nesting (Fig. 3d), yielded 247 association SNP loci (RDA_SIA set of loci, Table S6), with the first axis (linked to $\delta^{15}\text{N}$) separating Greece and the second axis (linked to $\delta^{13}\text{C}$) separating Libya, Israel and Akamas from the populations of Türkiye (DAL and BEL) and Alagadi. No apparent clustering of outlier loci across the genome was revealed on the Manhattan plots for each one of the four sets of loci identified by RNA (Figure S8).

We evaluated the proportion of SNP loci located in genes for the total loci data set (5,998 SNPs) and for the 6 datasets of candidate loci for adaptation, including outlier loci (Pcadapt and Bayescan) and association loci (RDA_CLUTCH, RDA_ENV, RDA_DISP and RDA_SIA). The percentage of SNP loci located in genes in the total set (57.59%) was within the range of percentages found in the outlier and association SNP loci from the different datasets, ranging from 53.16% to 64.37% (Table S6). From the SNPs located in genes in all datasets, almost all (>95%) were located in protein-coding genes. From the SNPs located in protein-coding genes, a small portion were located in exonic regions, ranging from 5.56% to 8.48%. Thus, in all cases, more than 91% of the loci located in protein-coding genes were located in intronic regions (Table S6). In conclusion, all SNPs analysed in the present work, as well as candidate SNPs, are more frequently located in introns of protein-coding genes, (Fig. 4a).

When combining the association and outlier loci, we recovered 604 SNPs (Table S6) that were located in 330 protein-coding genes (Table S7). We found that 189 protein-coding candidate loci were unique from a single analysis and only 3 overlapped between the 6 different analyses (Fig. 4b). We then performed functional enrichment analyses by comparing functionalities assigned to the candidate genes uniquely associated to one of the 6 analyses, in comparison to the total set of analysed SNPs located in protein-coding genes (Figs. 4c and 4d and Table S8). Bayescan outliers were significantly enriched in anatomical structure development, Pcadapt in respiratory system process, chromatin organization and carbohydrate metabolic process, RDA_CLUTCH in cytoplasmic translation, RDA_SIA in renal system process, regulation of DNA transcription, chromosome segregation and carbohydrate derivative metabolic process, and RDA_DISP in carbohydrate metabolic process and sulfur compound metabolic process. The 3 overlapped genes shared among all analyses are involved in important cell functions. The gene FAM19A3 (XP_048682780.1) regulates microglia activation, the microglia being a specialised population of macrophages, NKAIN3 (XP_048696741.1) regulates sodium ion transport and ATP8A2 (XP_048702040.1) in the generation and maintenance of phospholipid asymmetry in photoreceptor disk membranes and neuronal axon membranes. These genes might be involved in vesicle trafficking in neuronal cells and are required for normal visual and auditory function; they are also involved in photoreceptor and inner ear spiral ganglion cell survival. All the candidate SNPs associated with these three protein-coding genes

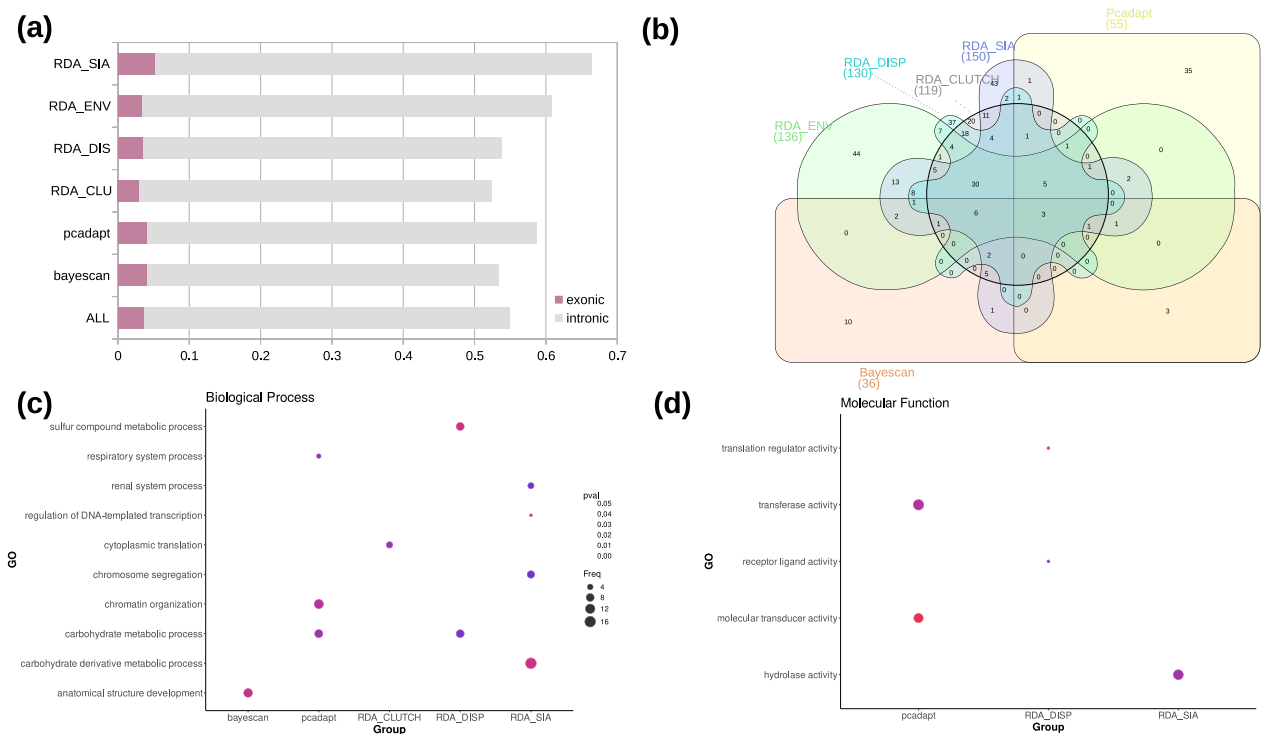


Fig. 4. Functional analysis of candidate protein-coding genes. **(a)** Location in the genome of the different sets of candidate SNPs. The length of the bar shows the percentage of the SNPs found in protein-coding genes in the 4 association analyses performed (RDA_CLUTCH: clutch size and geographic location, RDA_DISP: short-term dispersal data, RDA_SIA: stable isotopes and RDA_ENV: temperature and salinity) and the 2 outlier analyses (Bayescan, Pcadapt) and all the SNPs of this study (ALL). Purple designates percentage of SNPs in exonic regions, grey designates percentage of SNPs in intronic regions and the remaining percentage to reach the 100% correspond to SNPs found outside protein coding genes. **(b)** Venn diagram showing the number of candidate protein-coding genes from the 4 association analyses and 2 outlier analyses. Functional enrichment analysis of the candidate protein-coding genes uniquely found on a single association or outlier set of candidate genes, for biological processes **(c)** and molecular functions **(d)**. Circles designate the significant cases, where the size of the circle is proportional to the number of genes found in each category (Freq), and their colour to the adjusted p-value (pval).

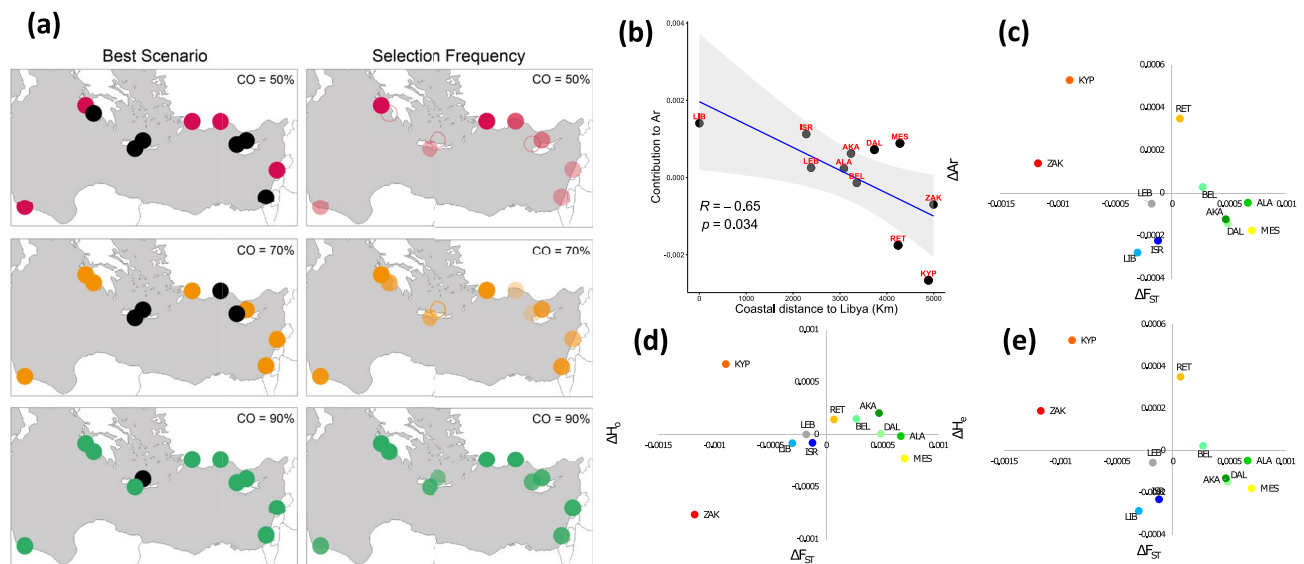


Fig. 5. Conservation prioritization analyses based on genomic information. **(a)** Panels on the left represent the best scenario chosen by MARXAN to achieve different conservation objectives (CO) to preserve 50% of the genomic diversity (pink), 70% (yellow) and 90% (green). Coloured circles indicate prioritized populations, whereas black circles show populations not considered a conservation priority to reach the CO. The graphs on the right display the frequency of selection for each population to be preserved for each scenario. Populations with more intense colours were selected at higher frequencies. **(b)** Correlation between the contribution of each population to the overall allelic richness (Ar) and its geographic distance (Km) from Libya. **(c-e)** Impact of the removal of each population to the overall structuring and connectivity. Increment of F_{ST} on the x axis for all plots (ΔF_{ST}) and genetic diversity indicated on the y axis: **(c)** allelic richness (ΔAr), **(d)** observed Heterozygosity (ΔHo), **(e)** expected Heterozygosity (ΔHe). Localities with positive values of ΔF_{ST} indicate that their loss will reduce connectivity. Localities with negative values of ΔAr , ΔHo or ΔHe indicate that their loss will reduce the corresponding genetic diversity index.

are located in intronic regions. The 24 protein-coding genes having candidate SNPs in exons, found in one or several analyses to unveil adaptation, have a myriad of domains and functions, including receptors, zinc fingers and kinases among others (Table S7).

Conservation prioritization

When assuming an equal cost for conserving genetic diversity among populations, MARXAN suggested different best solutions according to the three degrees of conservation of genetic potential provided (Fig. 5a). The populations of Zakynthos, Dalyan, Lebanon and Libya were considered as conservation priority for genetic diversity in all three scenarios, while the simulation also considered that almost all populations, should be preserved to retain 90% of their genomic potential. However, the frequency of selection of each population across the 100 runs was uneven, suggesting that different combinations of prioritization could yield similar results in terms of genetic diversity preservation. The analysis of allelic richness contribution indicated that not all populations contributed equally to the overall genetic diversity, with a significant gradient from Libya as the population with the highest contribution, following the coastline to Kyparissia with the lowest contribution. Finally, leave-one-out analyses (Fig. 5c-e) revealed the impact of removing specific populations in terms of genetic diversity (Ho , He and Ar) and structuring (F_{ST}). In terms of genetic diversity, the Allelic richness and the expected Heterozygosity would be negatively more impacted by the loss of LIB and ISR while the observed Heterozygosity would be negatively impacted, particularly by the loss of ZAK. In terms of genetic structuring, the populations of ZAK and KYP have a higher contribution to the overall structuring, while the populations of MES, RET, ALA, AKA, DAL and BEL contribute to the overall connectivity, as found with the network (Fig. 2b). Consequently, all populations had a significant role in the global genomic architecture, either by containing high diversity or to contributing to the needed connectivity among genetic clusters.

Discussion

Here we present the first genomic regional assessment of the loggerhead sea turtle populations breeding in the Mediterranean by combining genomic, phenotypic and environmental information to provide detailed insights into their genomic complexity. Our results not only improve the previous knowledge on their population structure, but also include the role of adaptation that has to be considered when conserving biodiversity⁴⁰. Finally, we provide informed conservation strategies to preserve the regional genetic variation genome-wide⁴¹.

Understanding the effective population size of wild populations is critical for assessing their status and planning conservation measures. This parameter is directly related to the impact of genetic drift and is then associated with the risk of losing genetic variation when populations have small effective populations sizes, that

may lead to deleterious effects. However, effective population size is rarely used for conservation purposes, even if it has been recently proposed as a key indicator for adaptive potential under the new Convention on Biological Diversity's Kunming-Montreal Global Biodiversity framework^{42,43}. Instead, changes in adult census size are still commonly used for risk assessments⁴⁴. Although these parameters are necessarily linked, understanding the variations in the relation between effective and census population sizes is key to assess the potential impact of genetic drift on the genomic diversity of the populations. Our study demonstrates that, in the case of the Mediterranean loggerhead turtles, effective population sizes are not correlated to adult census sizes and thus need to be considered independently. Previous studies suggest that an effective population of 500 is necessary to maintain equilibrium between the loss of adaptive genetic variation due to genetic drift and its replacement by mutation⁴⁵. Consequently, it is considered that the target N_e for conservation programs of endangered species should range from 500 to 1000⁴⁶. Although these estimates likely depend on the target species, previous studies on sea turtles considered populations with N_e of over a thousand breeders as healthy⁴⁷ and N_e ranging between 90–220 individuals as a sign of vulnerability⁴⁸. Among the Mediterranean region, our study finds that the Turkish populations of Belek and Dalyan, the Greek populations of Kyparissia and Zakynthos and the populations of Akamas and Alagadi all have estimates of effective population sizes above the suggested threshold of 500. Thus, these populations are less vulnerable to detrimental genetic effects related to small population sizes, such as the loss of alleles due to genetic drift. Furthermore, all these populations exhibit an effective population size similar to or even much larger than the adult census size. This can be the result of current connectivity or the reflection of genetic variability inherited from a past abundance whose loss had no time to be detected due to the long generation time of the species. For instance, recent studies find out that transient nesting individuals from other places, such as Israel, Syria and Türkiye, are not rare in Alagadi^{49,50}, thus suggesting that connectivity here can be higher than in other places, as also suggested by our connectivity analysis. Regarding historical processes, individuals from the populations of the eastern Mediterranean, thought to mainly forage in the area of the Mediterranean Levantine (Fig. 1a)^{28,51}, suffered an intense harvest of sea turtles for edible consumption in the XXth century until the 1970s^{52,53}. Therefore, the populations of Dalyan, Belek, Akamas and Alagadi were likely much larger than the present populations, suffered a decline at the end of the XXth century, and are now recovering after intense conservation efforts¹⁴. In this scenario of a recent bottleneck, failure to recover the past numbers of individuals with strong conservation measures will potentially erode on the long term the genomic diversity of these populations. Consequently, the application of conservation management to recover populations with contrasting differences between effective and adult population sizes, like Alagadi, is crucial to prevent this long-term loss of biodiversity.

The populations with low census and effective population sizes should also be of high concern. As such, the populations of Israel, Lebanon and Rethymno showed the smallest effective population sizes and are therefore more vulnerable to genetic drift. The case of Rethymno is curious, since this site hosts one of the largest populations in terms of the estimated number of adults coupled with one of the lowest effective population sizes. According to mtDNA haplotype composition, the populations of Crete and Cyprus have been proposed as being colonised more recently in comparison to the other populations of Greece, Libya and Türkiye³¹. As a consequence, the lower effective population sizes of the insular populations of Crete can be the result of a more recent origin, while in the case of Cyprus, current connectivity and past higher abundance may explain much higher effective population size. Finally, the populations of Israel and Lebanon presented low effective and census population sizes suggesting that these nesting populations were not much larger than now prior to the documented harvesting in the area during the XX century^{52,53}. In this scenario, the situation of these populations could be critical, as sustained reduced population sizes can be even more devastating for population genomic diversity than high levels of harvesting in long-living species⁵⁴. Considering our results, we recommend for conservation purposes the use of effective population size based on a genome-wide panel of markers, as a complementary measure of population census size. Based on these different classes of extinction risk, our results support the management of Mediterranean rookeries as different units, given the range of N_e presented by the studied populations.

Our genomic data also indicate a generalised significant level of inbreeding in Mediterranean loggerhead populations. The Mediterranean region was colonised by this species in the Pleistocene (ca. 65,000 years ago) and has suffered a series of contractions and expansions associated with the glacial periods during the Holocene³¹ that could explain these levels of inbreeding. Furthermore, the mean number of nests in the Mediterranean¹⁴ is generally lower than in other regions of the world⁵⁵. Depending on the patterns of connectivity and structure, partial gene flow across populations may alleviate this inbreeding to avoid potential associated negative effects.

We detected a complex population genomic structure in the Mediterranean with different levels of differentiation and connectivity. The loggerhead turtle populations first clustered into three groups, including those of Greece, Libya and the populations of the eastern Mediterranean. A previous study using 15 microsatellites³² found two clusters that are congruent with the first axis of our DAPC (Fig. 2c) or the STRUCTURE plot considering two groups (Figure S4a). A study using particle modelling to infer hatchling trajectories⁵¹ detected that some nesting populations share trajectories and foraging areas. Based on these modelling results, the authors proposed the definition of three subregional management units for conservation of marine turtles in the Mediterranean identified by dispersal patterns including 1) Greece nesting populations, 2) Libya nesting population and 3) the eastern Mediterranean nesting populations. Our genomic results, aligning with these subdivisions, demonstrate that these subregional management units correspond to real intermediate levels of genomic differentiation above the population level, and that hatchling dispersal patterns may have a role on building genetic differentiation.

Once the populations of the eastern Mediterranean are analysed alone, three additional groups emerged, including Türkiye (with Belek and Dalyan), Lebanon and Israel, all three groups connected by the populations of the Island of Cyprus. Our connectivity analysis network indicated that the populations of the Islands of Crete and Cyprus may play an important role in connecting different genetic clusters. Previous studies have already

suggested the role of these two islands as stepping stones among Mediterranean differentiated populations³⁰. Although genetic structure favours overall genetic diversity, connectivity networks are fundamental for species resilience since they could alleviate the negative effects of small population sizes on the local genetic diversity and reduce inbreeding in otherwise isolated populations⁵⁴. Consequently, the combination of deep differentiation with selected connectivity may potentially enhance the overall and local genetic diversity to allow an evolutionary rescue. Below the subregional level of structuring, almost all location pairwise comparisons were significant, meaning that almost all populations are genetically different and can be considered independent Management Units. This result contradicts some of the findings from previous studies^{30,32}. In particular, these studies found no differences between Rethymno and Zakynthos neither between Akamas, Lebanon and Libya. The lack of differentiation found in previous studies between Libya and other populations³² probably compromised the detection of Libya as an independent population. The authors suggested the existence of a complex pattern of breeding behaviour in common foraging areas to explain the lack of differentiation between these three populations. Considering that our study used exactly the same samples that this prior study for all the populations with discordant results (Libya, Akamas and Lebanon), the only possible explanation of such discrepancies necessarily relies on the different methodologies and the different number of markers used. In our study we used 5,998 SNPs while only 15 microsatellites were used in this prior study, meaning that the old set of markers could lack power to detect genetic differentiation and something similar may be happening with the comparison between Rethymno and Zakynthos. This agrees with the observation that increasing the number of loci improves the ability of identifying genetic differentiation³² and that the power of the genetic markers used is crucial to correctly infer population genetic structuring⁵⁶. Finding these discordant results highlights the importance of using markers with enhanced resolution for management and conservation, and specifically the potential of genomics studies for delineating conservation units. Consequently, we can conclude that the previously reported lack of differentiation between Libya, Akamas and Lebanon, and between Zakynthos and Rethymno was due to a lack of resolution rather than male mediated gene flow.

The deep genetic structuring found in the Mediterranean populations has also implications for the study of foraging grounds. Despite being philopatric, turtles from different populations undertake long migrations to foraging areas and thus may be affected by threats far away from their nesting beaches⁵⁷. The assignment of individuals at sea to the populations of origin has been neglected in this species due to a lack of resolution, and thus could only be analysed at the regional level^{29,58} or considering each mixed stock as a group without assigning at the individual level²⁸. The use of genomics is thus very promising for individual assignments to assess the origin of individuals incidentally captured by fisheries²⁹, stranded turtles⁵⁹ or individuals colonising new habitats^{39,60}. The clustering analysis using DAPC indicated much better clustering at individual level than previous studies, but we also detected some admixed individuals potentially resulting from discrete migration events among populations and subregional units. Consequently, our results will provide not only the baseline for identifying the origin of individual turtles of this highly migratory species along its complex life cycle to adapt conservation strategies, but also the detection of potential admixed individuals.

Beyond increased resolution, perhaps the most game changing potential of genomics is the power to detect candidate regions for adaptation⁴⁰. We found a strong signal of candidate regions for adaptation, either using outlier loci (potential local adaptation) or association loci with the phenotypic and environmental variables tested. Although some loci were shared across the multiple analyses performed, a large portion of the loci were unique to a specific test. This highlights the need to assess the role of adaptation on the genomic composition of populations using different approaches to have a wide comprehension of the drivers of adaptation³⁶. Additionally, the large number of loci found in each one of the analyses, and the high percentage of the genomic variability explained, indicate that the selected phenotypic and environmental variables have a substantial impact on the loggerhead sea turtle genomes. The fact that some loci are shared among different analyses is not surprising, considering that prior research has shown that some of the potential factors for adaptation may interact. For instance, the combination of stable isotopes with reproductive information⁶¹ indicated that the isotopic signature is correlated at the individual level with the clutch size, as females feeding in more productive areas lay a larger number of eggs. Other studies combining telemetry and stable isotope analysis showed that the foraging areas used by the individuals are associated with their isotopic signal, to the point that the isotopic signal can be used to infer the foraging areas used by the individuals⁶². Furthermore, the combination of particle modelling and mixed stock analysis on foraging grounds suggested that marine turtle migratory pathways as adults could be influenced by the dispersal patterns learned as hatchlings⁶³. Consequently, the candidate loci shared by different RDA analyses may reflect these interactions between phenotypic and environmental variables, while the unique loci may explain features unique to each set of variables. All SNP loci analysed in the present work, including candidate loci, were significantly enriched in intronic regions. This enrichment is already reported to be due to the restriction enzyme used in the reduced sequencing technique rather than selection preferentially targeting intronic regions^{64,65}. However, most loci are found in genic regions which provide a detailed set of candidate genes to explore the genomic footprint of the selected phenotypic and environmental variables impacting the loggerhead genomes. Future work on genome diversity using whole-genome sequencing across different populations would help to understand the adaptation processes driving intraspecific genome diversity⁶⁶.

Genomic and genetic data can not only provide information on the status, structure, and diversity of populations, but can also serve as an important scientifically based methodology to make crucial decisions for conservation. Considering the important role of conservation in loggerhead turtle survival in the Mediterranean region, it is imperative that decision makers have the full picture when it comes to prioritising conservation objectives. For this reason, conservation recommendations using genetic data should not be limited by a single analysis or parameter tested, but be complete enough to cover all relevant conservation targets, from genetic diversity and structuring to connectivity and adaptation. The prioritisation analysis software (MARXAN) allowed us to identify what populations of the region should be prioritised in conservation action to maximise

the retention of the genomic potential of the region⁴¹. The populations of Libya, Lebanon, Belek, Dalyan and Zakynthos were considered as conservation priorities in the three scenarios. However, some authors suggested that even a 90% threshold can lead to significant losses of genetic diversity on the long term⁶⁷, implying that almost all, if not all, populations should be prioritized to guarantee the long term persistence of genetic diversity. Furthermore, additional analysis on connectivity and diversity provided new insights to better understand the role of each population and its contribution to the overall population genomic potential, that should not be neglected. For instance, the population contribution to overall allelic richness decreased from Libya following the coastline, as suggested in previous studies³² and was consistent with the colonisation history of the Mediterranean³¹. Moreover, the role of populations in structuring and connectivity also needs to be considered. On one side, the populations of Zakynthos, Kyprisia or Libya provide overall structuring, thus enhancing the potential for genomic differentiation within the Mediterranean. On the other side, the populations of the islands of Crete (Rethymno and Messara) and Cyprus (Akamas and Alagadi) are essential hubs connecting genetic clusters, thus potentially alleviating the negative effects of inbreeding and genetic drift. As a result of the complex population structure of the Mediterranean loggerhead nesting populations, conservation prioritization should consider the different roles that the populations have on the overall structure, rather than focusing only on a particular aspect of the genomic diversity or differentiation. In conclusion, these results illustrate the potential of genomics for conservation decision making and implementation, showing the added value of incorporating genetic studies in conservation planning. Now more than ever, science and conservation must work together to speed up action and revert the trend of biodiversity decline.

In summary, our findings allowed for a better understanding of the status and population dynamics of loggerhead turtles breeding in the Mediterranean, revealed significant genetic differentiation among nesting populations and uncovered the role of adaptation. Our results set a baseline for evaluating future conservation measures, since the pressure that loggerhead turtles face in the Mediterranean goes beyond the nesting beaches. Consequently, conservation efforts at sea and on nesting beaches should complement to be effective. For this reason, future population genomic studies should focus on other high-risk areas such as feeding grounds, to understand the origin of turtles congregating there and set appropriate conservation measures. This study shows that genomics represents a step forward in the field of population genetics, and the results can be highly beneficial to improve conservation management of threatened species. These methods not only apply to loggerhead turtles but can be used with all sea turtle species and, more broadly, with any complex long lived highly migratory species. We therefore suggest the application of genomic analysis for future studies focusing on marine organisms' population and conservation genetics, and, in specific cases, we suggest reanalysing data published using traditional genetic markers since the power of genomics could unravel past misleading results.

Methods

Sampling and data collection

We collected tissue samples from 243 dead loggerhead hatchlings from 11 nesting populations across the Eastern Mediterranean basin (Fig. 1, Table S1). We reanalysed samples from different previous studies including a) Sirte (Libya), El Mansouri (Lebanon) and Akamas (Cyprus)^{31,32}, b) Belek and Dalyan (Türkiye)⁶⁸ and c) Kyprisia (Greece)⁶⁹. Samples from Akamas came mainly from two different nesting areas, the west coast of the peninsula of Akamas and Chrysochou Bay, east of the peninsula. However, these samples were pooled together as in previous studies using other genetic approaches^{31,32} since no genetic differentiation was found between the two sites also with genomic data ($F_{ST}=0.0005$, non-significant test, see procedures below). We sampled the five remaining sites, Zakynthos, Rethymno and Messara (Greece), Alagadi (Cyprus) and samples scattered along the Israeli coastline (Israel) between 2016 and 2019 (Table S1). To avoid pseudoreplication (e.g., sampling hatchlings from nests of the same female), dead hatchlings were collected either from nests of known females (through tagging) or from nests laid within a 14-day window, or in two consecutive nesting seasons, as in previous studies³⁰. One hatchling per nest was sampled, and the tissue was stored in 96% ethanol. All sampling was performed according to the regulations of each country and the appropriate CITES permits were obtained when needed.

For all populations, we collected environmental and phenotypic data on processes representing the different life stages of loggerhead turtles (Fig. 1, Table S3). Mean population output, recorded as mean clutch size, was retrieved from the literature for Zakynthos, Rethymno, Akamas, Lebanon, Israel and Libya⁶¹, Kyprisia, Messara and Alagadi¹⁴, Dalyan⁷⁰ and Belek⁷¹. Air temperature was taken, as it is known to correlate with sand temperature⁷² which directly affects the development of sea turtle embryos and determines their sex during incubation⁷³. We obtained data for air temperature from 1948 to present ($0.5^\circ \times 0.5^\circ$ grid) from the NOAA Earth System Research Laboratory (ESRL) database (<https://www.esrl.noaa.gov/>). Considering that sand temperature is not routinely recorded on nesting beaches at a constant depth, this approach allowed us to obtain a standardised parameter of temperature across all nesting sites during long data series. Data for water salinity was retrieved from the Copernicus Marine Service database⁷⁴, from 1989 to present ($0.042^\circ \times 0.042^\circ$ grid). For both parameters, we considered data from such a long period of time to be able to detect an impact of the environmental variable on sea turtle genomes, considering the long lifespan of marine turtles. We extracted minimum, maximum and mean temperatures, and salinity means for 1-, 50- and 100-m depth for each one of our sampling sites, considering all data within a 10 km distance radius for each site using R packages 'RANN'⁷⁵ and 'ncdf4'⁷⁶. Potential dispersal routes and foraging grounds used by newborn hatchlings during their first months were inferred from previous studies combining particle modelling with marine turtle specific nesting locations and biological characteristics⁵¹. We gathered the percentage of drifting particles reaching each one of the different foraging marine zones defined in this previous study (Adriatic Sea, Ionian Sea, South-Central Mediterranean Sea, Aegean Sea or Levantine area, Fig. 1a) after three months of drifting (measured at 1st December) that were released in the different nesting sites during the hatch period (August–September). We grouped the populations in the following nesting areas to match the approach used in this published study⁵¹:

western Greece (Zakynthos and Kyparissia), eastern Greece (Rethymno and Messara), Türkiye (Dalyan and Belek), Cyprus (Akamas and Alagadi) and Libya (Sirte). This previous study did not include the populations of Israel and Lebanon for the loggerhead turtle, and we used for these two sites the results for the green turtle (*Chelonia mydas*) in the nearby region of Syria. We used the information of the green turtle for these two sites, as the parameters used for both species were identical. Despite green turtle hatchlings have a slightly higher swimming speed (1.57 km/h) in relation to the loggerhead hatchlings (1.25 km/h), the offshore movement was set for 30 km for both species⁵¹. Additionally, the results that we used are considered from the perspective of the nesting areas (percentage of particles from a nesting area reaching each marine zone) and not from the perspective of the marine zones (percentage of particles from each marine zone coming from each nesting area) and thus nesting abundance of each species is irrelevant. Finally, mean stable isotope ratios $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ were obtained from the literature for hatchlings and nesting females, as both are indicative of the food intake of the females prior to the nesting seasons. Thus, the $\delta^{15}\text{N}$ is indicative of the trophic level, while the $\delta^{13}\text{C}$ indicates the source of carbon from primary producers from pelagic to neritic. Data from hatchlings was available from Zakynthos, Rethymno, Akamas, Lebanon, Israel and Libya⁶¹. This previous study also included data from Fethiye beach in western Türkiye, a nesting beach located between Belek and Dalyan, so we assumed the same values for these two populations. Since stable isotope data for Messaras was unavailable, we used as proxy the values of the closest populations of Rethymno⁶¹. Information from nesting females was available for Alagadi⁶² and Kyparissia⁷⁷ and was transformed into hatchling data following the formulas found in the literature⁷⁸: $\delta^{13}\text{C}_{\text{female}} = 0.51 \times \delta^{13}\text{C}_{\text{hatchling}} - 7.38$ and $\delta^{15}\text{N}_{\text{female}} = 1.02 \times \delta^{15}\text{N}_{\text{hatchling}} - 1.02$. Environmental variables were summarised using a Principal Components Analysis (PCA). Hatchling dispersal data was also summarised using a PCA. Environmental and phenotypic data was plotted using the 'ggplots' R package⁷⁹. In the set of variables reduced by PCA, the corresponding individual variables were included in the scatter plots as vectors using the envfit function of the 'vegan' R Package⁸⁰, whose direction indicates the main gradient of change in the variable and whose length is scaled to reflect the correlation between the spatial configuration of the populations in the PCA and the particular variable (Fig. 1).

DNA extraction and library building

Genomic DNA was extracted using Qiagen® Gentra Puregene blood cell kit, adjusting the manufacturer's protocol to our samples. DNA concentration was measured with Nanodrop® or PicoGreen®, and DNA degradation was assessed on 1% agarose gels. 2b-RAD libraries were prepared by adjusting the original protocol⁸¹ as optimised for marine turtles³⁸. We performed DNA digestion using AflI enzyme and customised adaptors were then attached to the digested sequences with sticky ends 5'-WN-3' (W = A and T)³⁹. Selective-adaptors (5'-WN-3') have been shown to reduce the number of analysed markers without compromising genetic differentiation⁸². Barcodes and Illumina primers were attached to the adaptors, sequences were amplified by PCR and then purified using magnetic beads to remove primers and sequences longer and shorter than 165 bp. The DNA concentration of individual purified libraries was quantified using PicoGreen®. To ensure a good depth of coverage for our data, we pooled individual libraries to sequence no more than 48 samples in a single lane. We performed single read 50bp sequencing per lane with a HiSeq 2500 Illumina at the Center for Genomic Regulations (CRG) of Barcelona.

Genotyping and filtering

We processed the 2bRAD sequences using customized scripts³⁸, trimming the raw sequences to remove ligation adaptors and cutting all the fragments to the same length (34 bp). We mapped the trimmed sequences to the published reference genome of the loggerhead turtle (GenBank accession GCA_023653815.1)⁸³ using Hisat2-2.2.1⁸⁴, to identify polymorphic nucleotides (SNPs) with BCFtools⁸⁵. Individual genotypes were outputted as SNPs in a VCF file. Filtering was performed with VCFtools v1.12⁸⁶. Loci were filtered for a minimum depth of 5 reads and should be present in at least 70% of individuals. Loci with outlier values of mean depth across all individuals (above the upper whisker of the mean depth value boxplot, corresponding to the 95% CI) were removed as they could be potential paralog DNA regions. Only polymorphic loci were retained by applying a -maf of 0.003 with VCFtools, which guarantees at least one alternative allele among the 243 individuals. This value was set to guarantee a Minimum Allele Count (MAC) of two alternative alleles thus removing singletons, an approach recommended for population structure analysis with large sample sizes⁸⁷.

Diversity indexes and effective population size

We computed basic genetic statistics for each population, such as observed and expected heterozygosities (H_o and H_e) and inbreeding coefficient (F_{IS}) using the function basic.stats of 'hierfstat'⁸⁸ and we used boot.ppfis of the same package to check for significant deviations of the inbreeding coefficient. We also calculated the allelic richness (A_r) using the function allelic.richness in 'hierfstat'. We estimated the effective population size (N_e) of our populations using the function ldne from the R package 'strataG'⁸⁹. The R function ldne estimates N_e from linkage disequilibrium (LD) with a Pearson correlation approximation. We only used loci present in all individuals, as the function used to calculate N_e requires no missing data. We also estimated the number of adults (N_a) in the population using information on nests counts and population trends from the literature¹⁴, that provided nest counts from two different time periods: prior to 1999 (corresponding to a published review⁹⁰) and after 2000 to reflect more recent abundance¹⁴. Nest counts from the two different periods (prior to 1999 and posterior to 2000) were then transformed to the number of adults following the formula of $Ad = nr/pd$ ⁹¹, where n is the number of nests, r is the remigration interval (years between consecutive nesting seasons for a given female) p is the proportion of females in the population and d is the number of nests per female and season. Values for these parameters were taken from the literature ($r = 2.3$; $p = 0.4$ and $d = 1.9$)⁹¹.

Population genetic differentiation

We assessed the overall population structure using two approaches. We first performed a Discriminant Analysis of Principal Components (DAPC) using the function `dapc` from the R package ‘*adegenet*’⁹² using localities as prior groupings. We used the function ‘`xvalDapc`’ to determine the number of PCs to retain, and we used the Bayesian Information Criterion (BIC) to determine the optimum number of groups in our samples. We also performed a STRUCTURE analysis, which implements a Bayesian clustering method to identify the most likely number of genetic groups (K) even in the presence of linked loci⁹³. Following established strategies⁹⁴, we carried out 10 runs per each value of K ranging from 1 to 11. We used the model of correlated allele frequencies and a burn-in of 50,000 steps, followed by 500,000 Markov Chain Monte Carlo steps. In addition to the mean $\ln(K)$, we also estimated the statistic Delta K to infer the most likely number of groups using STRUCTURE HARVESTER⁹⁵. The 10 runs of STRUCTURE for the most probable K, considering all available evidences, were averaged using CLUMPP version 1.1.2⁹⁶.

Pairwise F_{ST} values between populations were computed using the pairwise.neifst function from the R package ‘*hierfstat*’⁸⁸. We performed 999 permutations to obtain the respective p-values, and significance thresholds were set after B-Y FDR correction^{97,98} at an overall corrected 0.05 alpha-level. The results were plotted with heatmaps and dendrograms using the ‘*gplots*’ R package⁷⁹. A network analysis was done using the F_{ST} matrix with the program EDENetworks⁹⁹. We used the F_{ST} values among localities as a dissimilarity matrix and then transformed them into a network. We chose the automatic thresholding option, that sets the maximal distance values required to keep an edge just above the percolation threshold (at which the network breaks down into its main components). Geographic coordinates of sampled populations were introduced to adjust the network to the actual geographic location. Additionally, we tested if the breeders of the sampled nests contributed to nests from more than one location by performing relatedness tests among samples from different nesting beaches using the (`-relatedness2`) function of VCFTOOLS, based on the method described in the literature¹⁰⁰. Finally, we carried out an analysis of isolation by distance (IBD) to detect the correlation between geographic and genetic distance. Euclidean geographic distance between sites was computed using the function ‘`distm`’ from the R package ‘*geosphere*’¹⁰¹. Following the strategy of a prior study³², we also ran the analysis using the coastline distance matrix, which was measured using Google maps (<https://www.google.com/maps>) (Figure S5). We used the `mantel.randtest` function from R package ‘*adegenet*’⁹², using the linearized pairwise F_{ST} distances among populations ($F_{ST}/(1-F_{ST})$) as genetic distance.

Signals of adaptation

Genomic signals of adaptation were assessed using two different approaches: by looking for loci showing outlier signals of population differentiation (outlier loci) and by looking for loci associated with environmental and phenotypic variables using Redundancy Analyses (association loci). Candidate outlier loci were identified using two different programs: Pcadapt¹⁰² and BayeScan¹⁰³. For the first analysis, we used the Pcadapt function from the R package ‘*pcadapt*’¹⁰², with Mahalanobis distance. This package uses Bayesian factor models to determine both population structure and outlier loci. The Pcadapt function assumes that markers excessively related to population structure are candidates for local adaptation. The number of principal components (PC) was selected by plotting the proportion of explained variance in relation to the number of PCs. We calculated q-values using the `qvalue` function in ‘*qvalue*’ R package¹⁰⁴ which performs a false discovery rate (FDR) estimation from a collection of p-values. We considered outliers all loci having a q-value lower than an alpha-level of 0.05. BayeScan search was implemented with a total of 150,000 iterations, a burn-in period of 100,000, and 40 pilot runs of 10,000 iterations each. Outlier loci in BayeScan were identified using a `qual` of 0.05. In order to avoid false positives, only loci detected with both Pcadapt and Bayescan were kept as candidate outliers, and we used this set of loci to define an “outlier” dataset. Additionally, we defined a “neutral” dataset in a conservative way by removing from the general dataset all the loci found in any of the two programs. In this way, we made sure that no potentially selected locus was included in the “neutral” dataset. The analyses of population structure carried out with the complete loci dataset (pairwise F_{ST} and heatmap with dendrogram) were repeated using only the “outlier loci” and the “neutral loci” datasets.

We used a Redundancy Analysis (RDA)¹⁰⁵ to identify candidate loci significantly associated with the environmental and phenotypic data collected for the populations (Fig. 1, Table S2). We performed four separate RDAs, using the RDA function of the package ‘*vegan*’⁸⁰, each one representing a different life process by using different sets of variables: a) clutch size and geographic location b) environmental variables (the two first PCAs, Fig. 1) c) short-term dispersal data (the two first PCAs, Fig. 1) and d) stable isotopes ($\delta^{15}N$ and $\delta^{13}C$). Given the restriction of the analysis that does not allow missing data, we imputed missing genotypes by replacing them with the most common genotype across all individuals following the standard procedure¹⁰⁵. The results were plotted using the ‘*ggplot2*’ package¹⁰⁶ in R and associated loci were identified. The distribution of the candidate loci identified for each RDA analysis was visualised using a Manhattan plot using the ‘*qqman*’ package¹⁰⁷.

We performed a functional analysis for all sets of candidate loci for adaptation, including outlier and association loci (Figure S9). Using the `intersect` tool from `bedtools2` v2.30.0¹⁰⁸, we identified for all the sets of loci its location in the genome in genic and exonic regions. Furthermore, the SNPs classified as genic but not located in exonic regions were identified as being in intronic regions. We also distinguished those SNPs located in protein coding genes from those located in non-protein coding genes. We generated Venn diagrams for protein coding genes considering the different sets of loci using the InteractiVenn online tool¹⁰⁹. To decipher the functionality of the candidate proteins, we identified the longest transcript of each protein-coding gene and attached its functional information using eggNOG-mapper¹¹⁰ v5 (as implemented in <http://eggno-mapper.embl.de/>, last accessed November 2023) with default parameters. We parsed the eggNOG-mapper output file to subtract the Gene Ontology (GO) terms information for each protein, and generate the association file necessary to run a functional enrichment analysis (FEA). Using GOATOOLS¹¹¹, we did two types of FEA. On

one hand, we ran a FEA for each outlier and association set of proteins compared with the total set using the goslim_generic.obo database obtained from www.geneontology.org (release 2023–07–27). On the other hand, we did the same analysis but provided as an input file the list of protein coding genes that were unique for each candidate set. The GO terms of each candidate set from the Molecular Function and Biological Processes categories were visualized using ggplot2 for R.

Conservation prioritisation analysis

Conservation prioritization helps decision-making in the planning of protected areas. We used different molecular metrics to identify which of the studied populations had a higher potential for conservation planning, with the aid of the decision-support tool Marxan¹¹². Marxan uses a simulated annealing algorithm to identify sets of sites that meet given conservation targets at minimum costs. Each population was considered as a planning unit, to which we attributed the cost of 1, since actual data on the cost of conserving for each of the populations was not readily available. We considered as conservation features allelic richness (Ar), observed heterozygosity (Ho), population size (Na) (based on the number of nests since 2000) and nucleotide diversity (π) for each population. Ho , Ar and Ne had already been computed in previous analyses, while π was calculated using vcftools. Since Ho , Ar and π had low variability, each one of these metrics was normalised into three equal-interval categories: low, medium and high. Each metric was a single feature with three different levels pertaining to the low, medium and high classes derived from the range of values specific to that metric for that species¹¹³. Ne was not divided into categories, as its variability among all populations was very high. We considered 3 scenarios, setting our conservation objectives at 50, 70 and 90% for each feature, to understand which population would retain the highest amount of genetic variability under the different requirements. Marxan was run for 100 repetitions with 100,000 iterations and the temperature set at 10,000.

To further understand the genetic importance of each population in the broader Mediterranean context, we also calculated their contribution in terms of genetic diversity, measured as the allelic richness of each population to the total, or, in other words, the loss that would result from the extinction of any of these populations. We calculated the total allelic richness of the region and then computed the allelic richness of the total, removing one population at a time¹¹⁴. We then used the formula $(Ar_{(i)} - Ar_{(-i)}) / (Ar_{(i)} - 1)$, where $Ar_{(i)}$ is total allelic richness and $Ar_{(-i)}$ is allelic richness over all populations, excluding the one in question, to calculate the contribution of each population to the whole. Finally, we performed a leave-one-out analysis to evaluate the impact on the overall region of the loss of a specific population. To do so, we removed one population from the overall dataset, and we assessed the change in the overall measures of genetic diversity (Ar , Ho and He) and structure (F_{ST}) in relation to the complete dataset. For each measure, we calculated the rate of change (e.g., for ZAK, $\Delta Ar = \bar{Ar}_{all} - \bar{Ar}_{(all-ZAK)}$ where $\bar{Ar}_{all}$ is the mean allelic richness including all populations and $\bar{Ar}_{(all-ZAK)}$ is the mean allelic richness including all populations except ZAK). We plotted each change in diversity against the change in structure for visualisation.

Data availability

2bRAD raw data can be found at the European Nucleotide Archive (ENA) project PRJEB77620. Data on the phenotypic and environmental variables can be found in the Supplementary material.

Received: 20 November 2025; Accepted: 14 July 2026

Published online: 22 July 2026

References:

1. Dietz, S. & Adger, W. N. Economic growth, biodiversity loss and conservation effort. *J. Environ. Manage.* **68**, 23–35 (2003).
2. Harris, L. R., Holness, S., Finke, G., Kirkman, S. & Sink, K. Systematic conservation planning as a tool to advance ecologically or biologically significant area and marine spatial planning processes. In *Maritime Spatial Planning: Past, Present, Future* 71–96 (Springer International Publishing, 2019).
3. Moritz, C. Defining evolutionarily significant units for conservation. *Trends Ecol. Evol.* **9**, 373–375 (1994).
4. Coll, M. et al. The Mediterranean Sea under siege: Spatial overlap between marine biodiversity, cumulative threats and marine reserves. *Glob. Ecol. Biogeogr.* **21**, 465–480 (2012).
5. Giorgi, F. Climate change hot-spots. *Geophys. Res. Lett.* <https://doi.org/10.1029/2006GL025734> (2006).
6. Pastor, F., Valiente, J. A. & Khodayar, S. A warming Mediterranean: 38 years of increasing sea surface temperature. *Remote Sens.* <https://doi.org/10.3390/rs12172687> (2020).
7. Giangrande, A. et al. The Mediterranean in check: Biological invasions in a changing sea. *Marine Ecol.* **41**, Preprint at <https://doi.org/10.1111/maec.12583> (2020).
8. Zenetos, A. & Galanidi, M. Mediterranean non indigenous species at the start of the 2020s: Recent changes. *Marine Biodiv. Rec.* **13**, Preprint at (2020).
9. Hassoun, A. E. R. et al. Ocean acidification research in the Mediterranean Sea: Status, trends and next steps. *Front. Marine Sci.* **9**, Preprint at <https://doi.org/10.3389/fmars.2022.892670> (2022).
10. Casale, P. Sea turtle by-catch in the Mediterranean. *Fish Fish.* **12**, 299–316 (2011).
11. FAO. *The State of Mediterranean and Black Sea Fisheries 2023* (FAO, 2023). <https://doi.org/10.4060/cc8888en>.
12. Wallace, B. P. et al. Regional management units for marine turtles: A novel framework for prioritizing conservation and research across multiple scales. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0015465> (2010).
13. Wallace, B. et al. Marine turtle regional management units 2.0: An updated framework for conservation and research of wide-ranging megafauna species. *Endanger. Species Res.* **52**, 209–223 (2023).
14. Casale, P. et al. Mediterranean sea turtles: Current knowledge and priorities for conservation and research. *Endanger. Species Res.* **36**, 229–267. <https://doi.org/10.3354/esr00901> (2018).
15. Pike, D. A. Climate influences the global distribution of sea turtle nesting. *Glob. Ecol. Biogeogr.* **22**, 555–566 (2013).
16. Casale, P. *Caretta caretta* (Mediterranean subpopulation). *The IUCN Red List Threat. Species* **2014**, (2015).
17. Hawkes, L. A., Broderick, A. C., Godfrey, M. H. & Godley, B. J. Climate change and marine turtles. *Endanger. Species Res.* **7**, 137–154 Preprint at (2009).

18. Turkozan, O., Almpanidou, V., Yilmaz, C. & Mazaris, A. D. Extreme thermal conditions in sea turtle nests jeopardize reproductive output. *Clim. Change* **167**, 1–16 (2021).
19. Sönmez, B., Karaman, S. & Turkozan, O. Effect of predicted sea level rise scenarios on green turtle (*Chelonia mydas*) nesting. *J. Exp. Mar. Biol. Ecol.* **541**, 151572 (2021).
20. Varela, M. R. et al. Assessing climate change associated sea-level rise impacts on sea turtle nesting beaches using drones, photogrammetry and a novel GPS system. *Glob. Change Biol.* **25**, 753–762 (2019).
21. Mazaris, A. D., Kallimanis, A. S., Pantis, J. D. & Hays, G. C. Phenological response of sea turtles to environmental variation across a species' northern range. *Proc. R. Soc. B Biol. Sci.* <https://doi.org/10.1098/rspb.2012.2397> (2013).
22. Almpanidou, V., Katragkou, E. & Mazaris, A. D. The efficiency of phenological shifts as an adaptive response against climate change: A case study of loggerhead sea turtles (*Caretta caretta*) in the Mediterranean. *Mitig. Adapt. Strateg. Glob. Change* **23**, 1143–1158 (2018).
23. Hawkes, L. A., Broderick, A. C., Godfrey, M. H. & Godley, B. J. Investigating the potential impacts of climate change on a marine turtle population. *Glob. Chang. Biol.* **13**, 923–932 (2007).
24. Witt, M. J., Hawkes, L. A., Godfrey, M. H., Godley, B. J. & Broderick, A. C. Predicting the impacts of climate change on a globally distributed species: The case of the loggerhead turtle. *J. Exp. Biol.* **213**, 901–911 (2010).
25. Pike, D. A. Forecasting the viability of sea turtle eggs in a warming world. *Glob. Chang. Biol.* **20**, 7–15 (2014).
26. Turkozan, O., Yilmaz, C., Almpanidou, V., Godfrey, M. H. & Mazaris, A. D. Thermal conditions of green turtle (*Chelonia mydas*) nests in the largest rookery in the eastern Mediterranean. *Endanger. Species Res.* **50**, 63–73 (2023).
27. Carreras, C. et al. Genetic structuring of immature loggerhead sea turtles (*Caretta caretta*) in the Mediterranean Sea reflects water circulation patterns. *Mar. Biol.* **149**, 1269–1279 (2006).
28. Clusa, M. et al. Fine-scale distribution of juvenile Atlantic and Mediterranean loggerhead turtles (*Caretta caretta*) in the Mediterranean Sea. *Mar. Biol.* **161**, 509–519 (2014).
29. Clusa, M. et al. Potential bycatch impact on distinct sea turtle populations is dependent on fishing ground rather than gear type in the Mediterranean Sea. *Mar. Biol.* **163**, 122 (2016).
30. Carreras, C. et al. The genetic structure of the loggerhead sea turtle (*Caretta caretta*) in the Mediterranean as revealed by nuclear and mitochondrial DNA and its conservation implications. *Conserv. Genet.* **8**, 761–775 (2007).
31. Clusa, M. et al. Mitochondrial DNA reveals pleistocene colonisation of the Mediterranean by loggerhead turtles (*Caretta caretta*). *J. Exp. Mar. Biol. Ecol.* **439**, 15–24 (2013).
32. Clusa, M. et al. Philopatry in loggerhead turtles *Caretta caretta*: Beyond the gender paradigm. *Mar. Ecol. Prog. Ser.* **588**, 201–213 (2018).
33. Hohenlohe, P. A., Funk, W. C. & Rajora, O. P. Population genomics for wildlife conservation and management. *Mol. Ecol.* **30**, 62–82 (2021).
34. Casso, M., Turon, X. & Pascual, M. Single zooids, multiple loci: Independent colonisations revealed by population genomics of a global invader. *Biol. Invasions* **21**, 3575–3592 (2019).
35. Torrado, H., Carreras, C., Raventos, N., Macpherson, E. & Pascual, M. Individual-based population genomics reveal different drivers of adaptation in sympatric fish. *Sci. Rep.* **10**, 12683 (2020).
36. Carreras, C. et al. East is east and west is west: Population genomics and hierarchical analyses reveal genetic structure and adaptation footprints in the keystone species *Paracentrotus lividus* (Echinoidea). *Divers. Distrib.* **26**, 382–398 (2020).
37. Carreras, C. et al. The two sides of the Mediterranean: Population genomics of the black sea urchin *Arbacia lixula* (Linnaeus, 1758) in a warming sea. *Front. Mar. Sci.* **8**, 739008 (2021).
38. Barbanti, A. et al. Helping decision making for reliable and cost-effective 2b-RAD sequencing and genotyping analyses in non-model species. *Mol. Ecol. Resour.* **20**, 795–806 (2020).
39. Luna-Ortiz, A. et al. New colonisers drive the increase of the emerging loggerhead turtle nesting in Western Mediterranean. *Sci. Rep.* <https://doi.org/10.1038/s41598-024-51664-w> (2024).
40. Funk, W. C., McKay, J. K., Hohenlohe, P. A. & Allendorf, F. W. Harnessing genomics for delineating conservation units. *Trend. Ecol. Evol.* **27**, 489–496 (2012).
41. Kardos, M. et al. The crucial role of genome-wide genetic variation in conservation. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2104642118 (2021).
42. Hoban, S. et al. How can biodiversity strategy and action plans incorporate genetic diversity and align with global commitments?. *Bioscience* **75**, 47–60 (2025).
43. Da Silva, J. M. et al. Conserving genetic and genomic diversity in accordance with the global biodiversity framework. **30**, 2025 (2025).
44. IUCN. *IUCN Red List Categories and Criteria, Version 3.1*. vols iv, 32p. : ill. (IUCN Species Survival Commission (SSC), Gland and Cambridge, 2012).
45. Franklin, I. R. & Frankham, R. How large must populations be to retain evolutionary potential?. *Anim. Conserv.* **1**, 69–70 (1998).
46. Lynch, M. & Lande, R. The critical effective size for a genetically secure population. *Anim. Conserv.* **1**, 70–72 (1998).
47. Theissinger, K., FitzSimmons, N. N., Limpus, C. J., Parmenter, C. J. & Phillott, A. D. Mating system, multiple paternity and effective population size in the endemic flatback turtle (*Natator depressus*) in Australia. *Conserv. Genet.* **10**, 329–346 (2009).
48. Rivalan, P., Dutton, P. H., Baudry, E., Roden, S. E. & Giron-dot, M. Demographic scenario inferred from genetic data in leatherback turtles nesting in French Guiana and Suriname. *Biol. Conserv.* **130**, 1–9 (2006).
49. Snape, R. T. E. et al. Shelf life: Neritic habitat use of a turtle population highly threatened by fisheries. *Divers. Distrib.* **22**, 797–807 (2016).
50. Snape, R. T. E. et al. Off-the-shelf GPS technology to inform marine protected areas for marine turtles. *Biol. Conserv.* **227**, 301–309 (2018).
51. Casale, P. & Mariani, P. The first 'lost year' of Mediterranean sea turtles: Dispersal patterns indicate subregional management units for conservation. *Mar. Ecol. Prog. Ser.* **498**, 263–274 (2014).
52. Sella, I. Sea turtles in the eastern Mediterranean and northern Red Sea. In *Biology and Conservation of Sea Turtles* (ed. Bjørndal, A.) 417–423 (Smithsonian Institution Press, 1982).
53. Hornell, J. *Report on Fisheries of Palestine* (Published on behalf of the Govt. of Palestine by the Crown Agents for the Colonies, 1935).
54. Nigenda-Morales, S. F. et al. The genomic footprint of whaling and isolation in fin whale populations. *Nature Communications* **2023** *14*:1 14, 1–18 (2023).
55. Bolten, A. B. & Witherington, B. *Loggerhead Sea Turtles* (Smithsonian, 2003).
56. Bradshaw, P. J. et al. Defining conservation units with enhanced molecular tools to reveal fine scale structuring among Mediterranean green turtle rookeries. *Biol. Conserv.* <https://doi.org/10.1016/j.biocon.2017.12.014> (2018).
57. Vela-Garcia, H. et al. Where do sea turtles forage in the Mediterranean sea? Filling the gaps for a regional assessment of migration routes. *Front. Mar. Sci.* <https://doi.org/10.3389/fmars.2025.1661954> (2025).
58. Carreras, C. et al. Living together but remaining apart: Atlantic and Mediterranean loggerhead sea turtles (*Caretta caretta*) in shared feeding grounds. *J. Hered.* **102**, 666–677 (2011).
59. Turkozan, O. et al. Local differentiation in the origin of stranded loggerhead turtles, *Caretta caretta*, within an eastern Turkey foraging area. *Ocean Coast. Manag.* **153**, 70–75 (2018).

60. Carreras, C. et al. Sporadic nesting reveals long distance colonisation in the philopatric loggerhead sea turtle (*Caretta caretta*). *Sci. Rep.* **8**, 1435 (2018).
61. Cardona, L. et al. Distribution patterns and foraging ground productivity determine clutch size in Mediterranean loggerhead turtles. *Mar. Ecol. Prog. Ser.* **497**, 229–241 (2014).
62. Haywood, J. C. et al. Spatial ecology of loggerhead turtles: Insights from stable isotope markers and satellite telemetry. *Divers. Distrib.* **26**, 368–381 (2020).
63. Hays, G. C., Fossette, S., Katselidis, K. A., Mariani, P. & Schofield, G. Ontogenetic development of migration: Lagrangian drift trajectories suggest a new paradigm for sea turtles. *J. R. Soc. Interface* **7**, 1319–1327 (2010).
64. López, A., Carreras, C., Pascual, M. & Pegueroles, C. Evaluating restriction enzyme selection for reduced representation sequencing in conservation genomics. *Mol. Ecol. Resour.* <https://doi.org/10.1111/1755-0998.13865> (2023).
65. Galià-Camps, C., Pegueroles, C., Turon, X., Carreras, C. & Pascual, M. Genome composition and GC content influence loci distribution in reduced representation genomic studies. *BMC Genomics* **25**, 410 (2024).
66. Pegueroles, C., Pascual, M. & Carreras, C. Going beyond a reference genome in conservation genomics. *Trend. Ecol. Evol.* **39**, 13–15 (2024).
67. Frankham, R. Evaluation of proposed genetic goals and targets for the convention on biological diversity. *Conserv. Genet.* **23**, 865–870 (2022).
68. Yilmaz, C., Turkoz, O. & Bardakci, F. Genetic structure of loggerhead turtle (*Caretta caretta*) populations in Turkey. *Biochem. Syst. Ecol.* **39**, 266–276 (2011).
69. Carreras, C., Rees, A. F., Broderick, A. C., Godley, B. J. & Margaritoulis, D. Mitochondrial DNA markers of loggerhead marine turtles (*Caretta caretta*) [Testudines: Cheloniidae] nesting at Kyparissia Bay, Greece, confirm the western Greece unit and regional structuring. *Sci. Mar.* **78**, 115–124 (2014).
70. Turkoz, O. & Yilmaz, C. Loggerhead Turtles, *Caretta caretta*, at Dalyan Beach, Turkey: Nesting Activity (2004–2005) and 19-year Abundance Trend (1987–2005). <https://doi.org/10.2744/CCB-0719.1> **7**, 178–187 (2008).
71. Sak, S. & Baran, İ. Research on the sea turtle population of Belek beach. *Turk. J. Zool.* **25**, 361–367 (2001).
72. Cardona, L. et al. Global warming facilitates the nesting of the loggerhead turtle in the Mediterranean coast of Spain. *Anim. Conserv.* (2022).
73. Mrosovsky, N., Kamel, S., Rees, A. F. & Margaritoulis, D. Pivotal temperature for loggerhead turtles (*Caretta caretta*) from Kyparissia Bay, Greece. *Can. J. Zool.* **80**, 2118–2124 (2002).
74. Escudier, R. et al. A High Resolution Reanalysis for the Mediterranean Sea. *Front. Earth Sci. (Lausanne)* **9**, (2021).
75. Arya, S., Mount, D., Kemp, S. E. & Jefferis, G. RANN: Fast Nearest Neighbour Search (Wraps ANN Library) Using L2 Metric. Preprint at <https://cran.r-project.org/package=RANN> (2019).
76. Pierce, D. ncd4: interface to Unidata netCDF (version 4 or earlier) format data files.--R package ver. 1.16. Preprint at (2017).
77. Haywood, J. C., Margaritoulis, D., Theodorou, P. & Rees, A. F. Identifying critical marine habitats of the largest nesting population of loggerhead sea turtles in the Mediterranean: insights from stable isotope markers and satellite telemetry. *Testudo* **9**, 80–88 (2021).
78. Frankel, N. S., Zanden, H. B. Vander, Reich, K. J., Williams, K. L. & Bjørndal, K. A. Mother–offspring stable isotope discrimination in loggerhead sea turtles *Caretta caretta*. *Endanger. Species Res.* **17**, 133–138 (2012).
79. Warnes, G. R. et al. gplots: various R programming tools for plotting data. <http://CRAN.R-project.org/package=gplots>. (2016).
80. Oksanen, J. et al. Package ‘vegan’ Community Ecology Package. (2019).
81. Wang, S., Meyer, E., McKay, J. K. & Matz, M. V. 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nat. Methods* **9**, 808–810 (2012).
82. Galià-Camps, C., Carreras, C., Turon, X. & Pascual, M. The impact of adaptor selection on genotyping in 2b-RAD studies. *Front. Mar. Sci.* **9**, 1079839 (2022).
83. Chang, G. et al. The genome sequence of the Loggerhead sea turtle, *Caretta caretta* Linnaeus 1758. *F1000Res.* **12**, 336. (2023).
84. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
85. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
86. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
87. Hemstrom, W., Grummer, J. A., Luikart, G. & Christie, M. R. Next-generation data filtering in the genomics era. *Nat. Rev. Genet.* **25**, 750–767 (2024).
88. Goudet, J. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **5**, 184–186 (2005).
89. Archer, F. I., Adams, P. E. & Schneiders, B. B. stratag: An r package for manipulating, summarizing and analysing population genetic data. *Mol. Ecol. Resour.* **17**, 5–11 (2017).
90. Margaritoulis, D. et al. Loggerhead turtles in the Mediterranean Sea: Present knowledge and conservation perspectives. In *Loggerhead Sea Turtles* (eds Bolten, A. & Witherington, B. E.) 175–198 (Smithsonian Books, 2003).
91. Casale, P. & Heppell, S. S. How much sea turtle bycatch is too much? A stationary age distribution model for simulating population abundance and potential biological removal in the Mediterranean. *Endanger. Species Res.* **29**, 239–254 (2016).
92. Jombart, T. ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
93. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
94. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
95. Earl, D. A. & von Holdt, B. M. Structure harvester: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
96. Jakobsson, M. & Rosenberg, N. A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806 (2007).
97. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188 (2001).
98. White, T., van der Ende, J. & Nichols, T. E. Beyond Bonferroni revisited: Concerns over inflated false positive research findings in the fields of conservation genetics, biology, and medicine. *Conserv. Genet.* **20**, 927–937 (2019).
99. Kivela, M., Arnaud-Haond, S. & Saramäki, J. EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Mol. Ecol. Resour.* **15**, 117–122 (2015).
100. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
101. Hijmans, R. geosphere: spherical trigonometry. *R package version 1.5–19* (2023).
102. Luu, K., Bazin, E. & Blum, M. G. B. Pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* **17**, 67–77 (2017).
103. Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **180**, 977–993 (2008).
104. Storey, J., Bass, A., Dabney, A. & Robinson, D. qvalue: Q-value estimation for false discovery rate control. *R package version 2.36.0* (2024).

105. Forester, B. R., Lasky, J. R., Wagner, H. H. & Urban, D. L. Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Mol. Ecol.* **27**, 2215–2233 (2018).
106. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
107. Turner, S. D. Qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* **3**, 731 (2018).
108. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
109. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **16**, 169 (2015).
110. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucl. Acid. Res.* **47**, D309–D314 (2019).
111. Klopfenstein, D. V. *et al.* GOATOOLS: A Python library for Gene Ontology analyses. *Sci. Rep.* **8**, 10872 (2018).
112. Ball, I., Possingham, H. & Watts, M. Marxan and relatives: Software for spatial conservation prioritization. In *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools* (ed. Moilanen, A.) 185–210 (Oxford University Press, 2009).
113. Nielsen, E. S., Beger, M., Henriques, R. & von der Heyden, S. A comparison of genetic and genomic approaches to represent evolutionary potential in conservation planning. *Biol. Conserv.* **251**, 108770 (2020).
114. von Takach, B. *et al.* Population genomics and conservation management of a declining tropical rodent. *Heredity (Edinb)* **126**, 763–775 (2021).
115. Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. & Wobbe, F. Generic mapping tools: Improved version released. *EOS Trans. Am. Geophys. Union* **64**, 409–420 (2013).

Acknowledgements

Our special gratitude goes to Andreas Demetropoulos (1938–2022), the Mediterranean region's pioneer marine turtle conservationist and a co-author of this paper. Andreas was a key figure in the Mediterranean region through his active, long-term involvement in turtle conservation, including as a long-standing member of the IUCN's Marine Turtle Specialist Group. He started to work on the marine turtles of Cyprus in the late 70s which helped provide valuable information for the conservation of the species. He was an inspirational mentor to many marine turtle conservationists in the Mediterranean and beyond who have followed in his footsteps. He is remembered for his charismatic personality, his unbending enthusiasm and dedication towards conservation and for his willingness to share his immense knowledge. We thank Dr. Mona Khalil (Medasset Lebanon), who provided the samples from Lebanon for past published studies that were reused in the present work. We thank all of the supporters, funders, staff and volunteers of the Marine Turtle Conservation Project in Cyprus, a collaboration between the Society for the Protection of Turtles and the University of Exeter. We thank all field assistants and volunteers of ARCHELON that collected samples in Zakynthos, Kyparissia Bay, Rethymno and Messaras Bay. We wish to express our gratitude to the anonymous reviewers for their critical peer review, which improved the clarity and quality of this manuscript.

Author contributions

CC and MP conceived and designed the study. DM, AR, OT, CU, RTES, ACB, AD, SD, AAH and YL obtained the biological samples, applied for permits and organised sample or DNA transportation with CC. OT and CU performed the DNA extractions of the samples from Belek and Dalyan. AB undertook all the remaining laboratory analysis with the aid of MP and CC. AB, CC and CP conducted the data analyses with inputs from MP. AB, CC, CP and MP wrote the manuscript with inputs from all the authors.

Funding

This work was supported by the projects MarGech (PID2020-118550RB) funded by MICIU/AEI/10.13039/501100011033, GenoMarTur (CNS2022-135205) funded by MICIU/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR), and BlueDNA (PID2023-146307OB) funded by MICIU/AEI/10.13039/501100011033 and ERDF/EU. AB, CP, MP and CC are members of the research group SGR2021-01271 funded by the Generalitat de Catalunya.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-62668-z>.

Correspondence and requests for materials should be addressed to C.C.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026