Check for updates

Received: 12 May 2026 | Accepted: 31 October 2026

DOI: 10.1111/2041-210x.70453

Methods in Ecology and Evolution

PRACTICAL TOOLS

Individual natal assignment in highly migratory species: The genomic baseline and its application in loggerhead turtles

Astrid Luna-Ortiz[1] | Anna Barbanti[1] | Cinta Pegueroles[2] | Federico Alberto Abreu-Grobois[3] | Paolo Casale[4] | Daniela Freggi[5] | Silvia Giralt[6] | Elizabeth Labastida-Estrada[3] | Raúl Llera-Herrera[3] | Salima Machkour-M'Rabet[7] | Adolfo Marco[8] | Dimitris Margaritoulis[9] | Oguz Türkozan[10] | Marta Pascual[1] | Carlos Carreras[1]

[1]Departament de Genètica, Microbiologia i Estadística and IRBio, Universitat de Barcelona, Barcelona, Spain; [2]Departament de Genètica i Microbiologia i Institut de Biomedicina i de Biotecnologia (IBB), Facultat de Biociències, Universitat Autònoma de Barcelona, Barcelona, Spain; [3]Unidad Académica Mazatlán, Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Mazatlan, Sinaloa, Mexico; [4]Ethology Unit, Department of Biology, University of Pisa, Pisa, Italy; [5]Lampedusa Turtle Rescue, Lampedusa and Linosa Islands, Agrigento, Italy; [6]Fundación Para la Conservación y la Recuperación de Animales Marinos (CRAM), El Prat de Llobregat, Spain; [7]Laboratorio de Ecología Molecular y Conservación, Departamento de Conservación de la Biodiversidad, El Colegio de la Frontera Sur Unidad Chetumal, Chetumal, Mexico; [8]Estación Biológica de Doñana, CSIC, Sevilla, Spain; [9]ARCHELON, The Sea Turtle Protection Society of Greece, Athens, Greece and [10]Department of Biology, Faculty of Science, Aydın Adnan Menderes University, Aydın, Türkiye

Correspondence
Carlos Carreras
Email: carreras@ub.edu

Marta Pascual
Email: martapascual@ub.edu

Funding information
Ministerio de Ciencia, Tecnología e Innovación, Grant/Award Number: CNS2022-135205, PID2020-118550RB and PID2023-146307OB; Generalitat de Catalunya, Grant/Award Number: 2025 SGR 00608 and 2025 SGR 01194; Fundación Biodiversidad, Grant/Award Number: FB 04/2023

Handling Editor: Pablo Duchen

Abstract

1. Effective conservation of highly migratory species requires understanding the genetic structure across breeding populations and access to high-resolution markers capable of assigning individuals from mixed aggregates (e.g. bycatch or new nesting sites) to their natal origins. Genomic approaches provide unprecedented resolution but add methodological challenges; therefore, it is essential to first build a genomic baseline from known breeding areas and then evaluate strategies for assigning unknown individuals.
2. To address this, we used 2bRAD sequencing, a genomic reduction technique useful for degraded DNA, and loggerhead turtles (*Caretta caretta*) as a case study. This species shows philopatric breeding, while juveniles and adults form mixed aggregations in foraging grounds.
3. Our results highlight the importance of building baselines that include all potential source populations that contribute to mixed aggregations. We detected hierarchical genetic differentiation and high resolution and successfully assigned the natal origin of 124 unknown individuals from four Mediterranean foraging grounds. These grounds showed distinct source contributions, and comparisons with previous studies suggest possible temporal shifts in stock composition.

Marta Pascual and Carlos Carreras are senior authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

4. We provide a comprehensive genomic baseline for individual assignment of Atlanto-Mediterranean loggerhead turtles of unknown natal origin and a general framework for identifying population-specific threats in highly migratory species.

KEYWORDS
2bRAD, *Caretta caretta*, foraging grounds, migratory species, natal origin, population structure, sea turtles

1 | INTRODUCTION

The wide spatial distribution, long lifespan and use of multiple dispersed habitats for development, foraging and reproduction make highly migratory species particularly challenging to monitor and conserve. Movements between these habitats expose individuals to various anthropogenic threats throughout life stages, including fisheries interactions, pollution, habitat degradation and vessel collisions (Cooke et al., 2024). Individuals from multiple breeding populations can share foraging areas without mating (Carreras et al., 2011). Consequently, the ability to assign individuals at developmental and foraging sites to their natal origins may help identify threats affecting specific populations. Directly identifying the breeding origin of individuals sampled away from natal areas requires two complementary datasets. First, a robust genomic baseline must be established using individuals from potential source populations. Second, unknown individuals must be genotyped and compared against this baseline to infer their natal origin.

Sea turtles are emblematic highly migratory species, undertaking long-distance migrations between nesting and foraging areas that can span thousands of kilometres and more than a decade (Hays et al., 2002). Despite this, they exhibit natal philopatry, returning to their hatching beaches to reproduce, which generates a detectable genetic structure (Barbanti et al., 2026). Combining the genetic structure driven by philopatry with migratory patterns has led to the definition of Regional Management Unit (RMU) as multiple genetically distinct populations that share migration routes, foraging areas and threats (Wallace et al., 2023). Mixed stock analysis (MSA) estimates the proportion of individuals originating from different populations in a potential mixed aggregate (Grant et al., 1980). In sea turtles, MSA has primarily relied on mitochondrial DNA, revealing that Mediterranean foraging grounds (FGRs) host individuals from multiple RMUs (Mediterranean, North-West Atlantic and North-East Atlantic) (Vela-Garcia et al., 2025). However, the widespread sharing of mtDNA haplotypes among source populations results in large confidence intervals, limiting the precision of stock composition assessments (Carreras et al., 2006). Nuclear microsatellites enabled individual assignment (IA) in marine loggerhead turtles (*Caretta caretta*) (Carreras et al., 2011) but constrained IAs to the RMU level and left many individuals unassigned (Clusa et al., 2016).

Genomic advances now allow the use of thousands of markers, greatly improving resolution. In this context, 2bRAD sequencing is particularly suitable for species with large genomes and degraded DNA (Barbanti et al., 2020). A recent 2bRAD study has revealed unprecedented fine-scale population structure in Mediterranean loggerhead nesting populations (Barbanti et al., 2026), demonstrating the potential of genomics to establish robust baselines for IA. However, the absence of Atlantic samples limits its application for assigning individuals found in Mediterranean FGRs.

Our study aims to establish best practices for applying genomic baselines to assign individuals of unknown natal origin in highly migratory species. Using loggerheads as a model, we evaluate how genomic differentiation between RMUs and subregional units (SubRMUs) can improve IA. We specifically tested how alternative genotyping strategies for both baseline and unknown datasets affect assignment outcomes. By integrating genomic data across breeding and foraging habitats, this work provides a framework for quantifying population contributions to mixed aggregations and offers general guidelines for genomic IA in conservation and management.

2 | MATERIALS AND METHODS

2.1 | Sampling

We built a baseline dataset with 274 loggerhead individuals from 14 nesting areas (Table S1, Figure 1), representing three RMUs known to contribute to Mediterranean FGRs (Vela-Garcia et al., 2025; Wallace et al., 2023). The dataset included 15 individuals from the North-West Atlantic RMU (NWA; Quintana Roo [QRO, $N = 12$], and Grand Cayman [CAY, $N = 3$]), 25 from the North-East Atlantic RMU (NEA; all from Boa Vista [BOA]), and 234 previously published unrelated individuals from the Mediterranean (MED) RMU (Barbanti et al., 2026). Nesting area samples were obtained from dead hatchlings. All samples were collected under the local legislation, and appropriate CITES permits were obtained when needed. No additional permission was needed for fieldwork.

Additionally, we analysed 124 individuals of unknown natal origin collected from 4 Mediterranean FGRs (Figure 1, Table S1): Catalan coast (CAT, $N = 32$), Lampedusa (LAM, $N = 38$), West Aegean coast (WAS, $N = 23$) and East Aegean coast (EAS, $N = 31$). These samples were obtained from animals admitted in rescue centres. Skin samples were taken from dead animals, while blood samples were collected from living animals following established rescue-centre

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

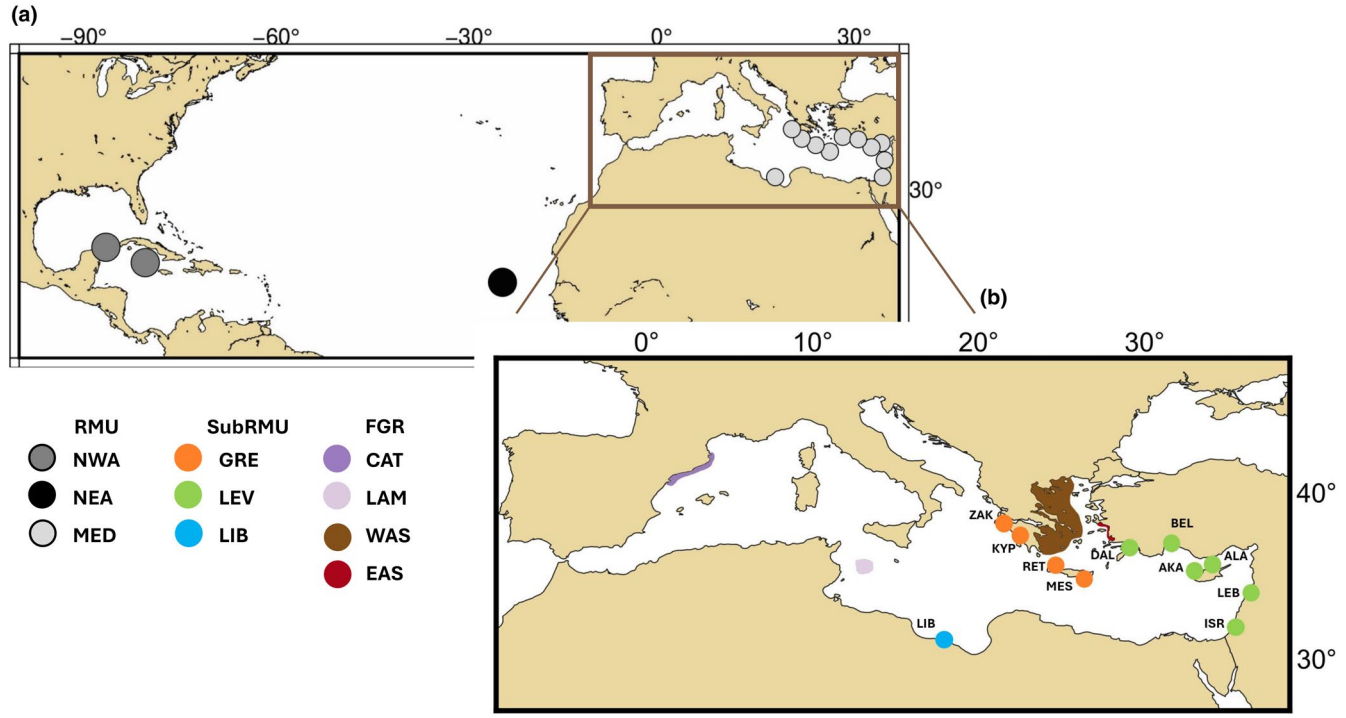

FIGURE 1 Map of nesting locations (baseline) and foraging grounds (unknown natal origin). (a) Baseline samples were from North-West Atlantic (NWA), North-East Atlantic (NEA) and Mediterranean (MED) RMUs. (b) Mediterranean locations were grouped into three SubRMUs: Libya (LIB); Levantine (LEV; Israel [ISR], Lebanon [LEB], Alagadi [ALA], Akamas [AKA], Belek [BEL] and Dalyan [DAL]); and Greece (GRE; Messara [MES], Rethymno [RET], Kyparissia [KYP] and Zakynthos [ZAK]). Coloured polygons indicate the sampled foraging areas of individuals of unknown origin. Catalan coast (CAT), Lampedusa (LAM), Western Aegean coast (WAS) and Eastern Aegean coast (EAS). Maps created with QGIS (https://qgis.org/).

protocols during health status evaluation. Thus, our study did not require ethical approval. All samples were preserved in 96% ethanol and stored at −20°C until DNA extraction.

2.2 | Laboratory and genotyping protocols

Genomic DNA was extracted using the Puregen Kit (Qiagen), following the manufacturer's instructions. Individual libraries were prepared for the 164 individuals sampled in this study (Table S2) as for the 234 previously published Mediterranean baseline individuals (Barbanti et al., 2026). Briefly, DNA was digested with AlfI and ligated to base-selective adapters (5′-WN-3′) to reduce the number of loci while preserving genetic differentiation (Galià-Camps et al., 2022). Fragments were amplified, purified using magnetic beads and quantified using the Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo Fisher Scientific). Libraries were pooled by index type and sequencing was performed at the Centre for Genomic Regulation.

Raw 2bRAD reads from new and published data (398 individuals) were processed using custom scripts (Luna-Ortiz et al., 2026). Sequences were trimmed to remove adapters, cut to the same length (34 bp) and mapped to the loggerhead reference genome (GenBank accession GCA_023653815.1; Chang et al., 2023) using Hisat2-2.2.183 (Kim et al., 2019). Two genotyping strategies were implemented (Figure S1). First, all samples were genotyped simultaneously (*Joint Genotyping*) with BCFtools (Li, 2011), and resulting genotypes saved in a VCF file. Second, the baseline and unknown samples were independently genotyped (*Non-Joint Genotyping*) and independent VCF files were generated for each group.

In both strategies, the baseline VCF file was filtered using VCFtools (Danecek et al., 2011), by removing genotypes supported by fewer than five reads, loci exceeding 50× mean depth (defined as 1.5 times the interquartile range), and SNPs with missing data in more than 30% of individuals (*--max-missing* 0.7). SNPs with a minor allele frequency >0.002 were kept, removing singletons. The SNPs retained in the baseline were extracted from the unknown dataset using the *--snps* function in VCFtools. The strategy retaining the highest number of informative SNPs in the unknown dataset was selected for subsequent analyses.

To evaluate the effect of sample size under the *Non-Joint Genotyping* approach, we performed subsampling analyses of the unknown dataset. We generated subsets of 10, 50 and 75 individuals, each with 10 random replicates sampled without replacement. Each replicate was independently genotyped, and the number of SNPs shared with the baseline was quantified. Differences between subset sizes were assessed using a Kruskal–Wallis test followed by post hoc Wilcoxon signed rank tests in R.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.3 | Baseline genomic differentiation

With the best genotyping strategy, we built the baseline including the 274 samples separated by sampling location (named *Caretta Baseline*) and characterised its population structure. Genetic differentiation among RMUs was assessed by pairwise F_{ST} values grouping sampling localities by RMU (Table S1). Statistical significance was assessed with 999 permutations using the R package 'StAMPP' (Pembleton et al., 2013). We estimated pairwise F_{ST} values and significance among the three Mediterranean SubRMUs using the *MED Baseline* dataset and the same SNPs as the *Caretta Baseline*, with localities grouped by SubRMU (Table S1). Finally, we estimated pairwise genetic differentiation between all localities with the *Caretta Baseline* dataset.

We evaluated genomic differentiation using Discriminant Analysis of Principal Components (DAPC), implemented in 'adegenet' (Jombart, 2008), following Jombart and Collins (http://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf). We extracted the density distribution of the first two discriminant axes for both baseline datasets (*Caretta* and *MED*).

2.4 | Individual assignment of samples of unknown natal origin

We used the R package 'assignPOP' version 1.1.4 (Chen et al., 2018) for IA, following Marín-Capuz et al. (2025). The 124 samples of the unknown dataset were assigned using the 'assign.X' function and the best-performing model following three strategies: (1) assignment against the *Caretta Baseline* grouped by RMU (NWA, NEA and MED); (2) assignment against the *MED Baseline* grouped by SubRMU (LIB, LEV and GRE); and (3) assignment against the *Caretta Baseline* considering five groups (NWA, NEA, LIB, LEV and GRE), referred to as 'all-at-once' approach.

3 | RESULTS

3.1 | Strategies for building a genomic baseline

Overall, 4,678,552 mean reads were obtained per individual (SD ± 1,727,584), with a mean mapping success of 93.08% (SD ± 7.35), regardless of the library preparation strategy (Table S2). After filtering the baseline datasets, 6586 SNPs were retained in both genotyping strategies (Figure S1). However, the two approaches differed in the number of SNPs recovered in the unknown dataset. The *Joint Genotyping* strategy recovered all polymorphic markers identified in the baseline, whereas the *Non-Joint Genotyping* strategy recovered only 88.2%. Subsampling analyses under the *Non-Joint Genotyping* approach (Figure S2) revealed a strong effect of sample size on SNP recovery, with a significant reduction in the number of shared loci at small sample sizes (Kruskal–Wallis $p < 0.05$, Wilcoxon signed rank test, $p < 0.05$ for all cases). Consequently, we selected the *Joint Genotyping* strategy for all subsequent analyses.

3.2 | Genomic differentiation of the baseline

All pairwise F_{ST} values between RMUs were statistically significant (Table S3a) and the DAPC analysis revealed three genetic groups corresponding to the three RMUs (Figure S3a). The first discriminant function clearly separated individuals from the three RMUs, with the greatest differentiation observed for Mediterranean individuals (Figure S3b), whereas the second discriminant function distinguished between individuals from the two Atlantic RMUs (Figure S3c). SubRMUs pairwise F_{ST} values with the *MED Baseline* were all statistically significant (Table S3b) and the DAPC analysis revealed three distinct genetic groups LIB, LEV and GRE (Figure S3d). The first discriminant axis differentiated the three SubRMUs with some overlap, whereas the second discriminant axis separated LIB from the rest (Figure S3f,e). Consequently, both baselines showed good differentiation and discrimination potential. Finally, we evaluated genetic differentiation at the locality level using the *Caretta Baseline* (Figure S4) and found significant genetic differentiation between all pairwise localities except AKA-ALA and DAL-BEL (Table S3c).

3.3 | Individual assignments in foraging areas

Natal origins of the individuals from the FGRs were inferred with 'assignPOP' using the SVM model with the highest accuracy (Figure S5). Each assignment was repeated 10 times and the mean assignment probabilities per individual were calculated. First, we assign the FGR individuals to the RMU level using the *Caretta Baseline* (Figure 2). All FGR individuals were assigned to an RMU with a probability >0.75 (Figure 2b, Table S4). We chose this threshold after plotting the assignment probabilities of the FGR individuals with the three baseline strategies (RMU, SubRMU and all-at-once) to detect break points in their distribution (Figure S6). Most of the FGR individuals (96.8%) were assigned to MED (Table S4), three were assigned to NWA (one from CAT and two from LAM) and one to NEA (from LAM). Consequently, LAM had the highest proportion of individuals from Atlantic origin (7.9%).

Second, FGR individuals were assigned using the *MED Baseline*. Assignments to a single SubRMU were considered successful with probabilities >0.75 (Table S5, Figure S6), and individuals with maximum assignment probability <0.75 are considered admixed. In total, 97 individuals were assigned to a single SubRMU (5 to LIB, 61 to LEV and 31 to GRE) and 27 individuals had an admixed origin (Figure 2c, Table S5). Of the four individuals previously assigned to an Atlantic origin, one was assigned to LEV, and the other three were identified as admixed. This highlights that excluding potential source populations from the baseline leads to wrong assignments.

Third, FGR individuals were assigned using the all-at-once strategy (Figure 2d, Figure S6). Three individuals were assigned to

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

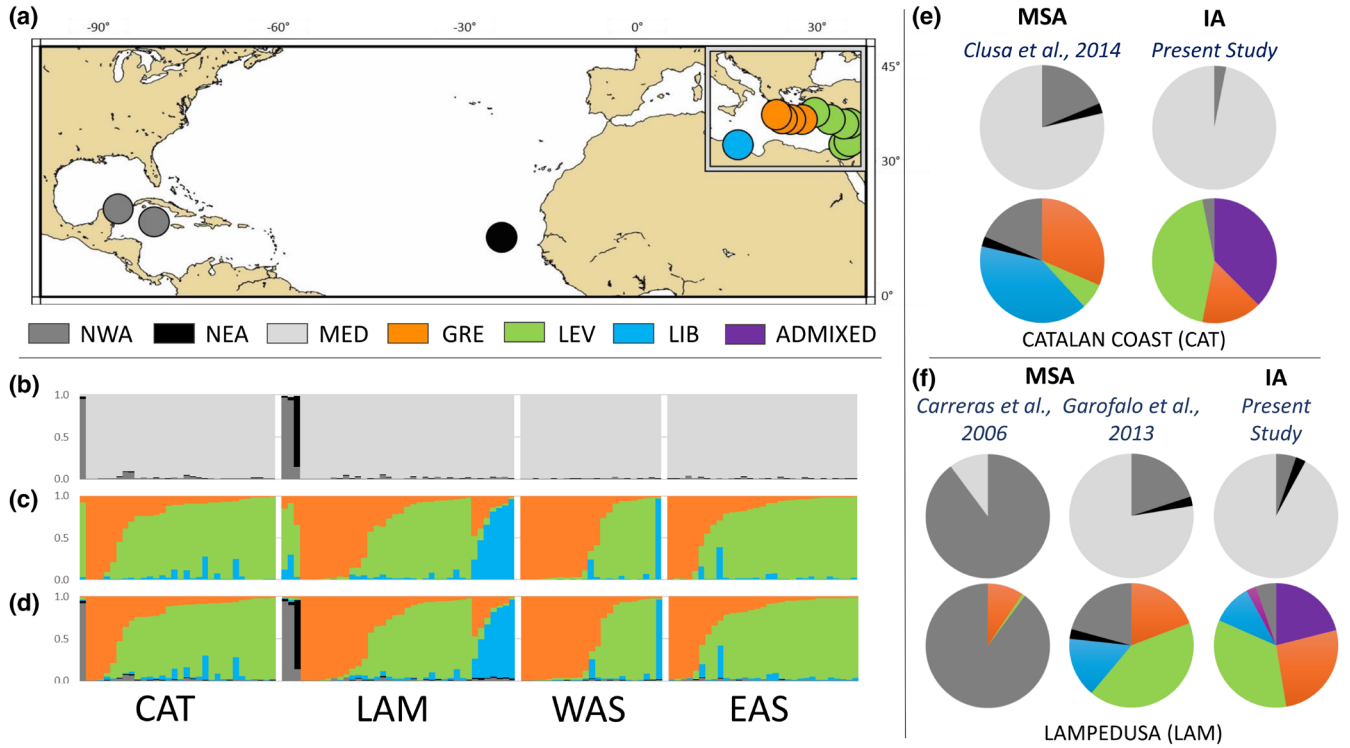

FIGURE 2 Results of the assignment strategies for the individuals sampled on the foraging grounds of CAT ($N = 32$), LAM ($N = 38$), WAS ($N = 23$) and EAS ($N = 31$). (a) Nesting locations of the genomic baseline RMU coloured (North-West Atlantic: NWA, North-East Atlantic: NEA, and Mediterranean: MED), and coloured Mediterranean SubRMU (Libya: LIB, Levantine: LEV and Greece: GRE). (b) Individual assignment (IA) probabilities using the *Caretta Baseline* grouped in RMU. (c) IA probabilities using the *MED Baseline* grouped at SubRMU. (d) All-at-once IA probabilities using the *Caretta Baseline* grouped into five genetic groups (NWA, NEA, LIB, LEV and GRE). The 124 FGRs individuals have the same order across panels. (e) Proportion of individuals from the Catalan coast (CAT) assigned to RMU and with the all-at-once approach, and its comparison with the MSA results of Clusa et al. (2014) ($N = 40$). (f) Proportion of individuals from Lampedusa (LAM) assigned to RMU and with the all-at-once approach in this study and using MSA in Carreras et al. (2006) ($N = 65$) and Garofalo et al. (2013) ($N = 123$). Individuals with probabilities <0.75 to one group are considered admixed (purple). Maps created with QGIS (https://qgis.org/).

NWA and 1 to NEA, consistent with the RMU assignment, 5 were assigned to LIB, 57 to LEV and 31 to GRE, while the remaining 27 were considered admixed. Aside from discrepancies caused by excluding Atlantic populations from the baseline, all other individuals presented similar assignment probabilities to the same SubRMU or were considered admixed, except for individuals CC2006, CC2015 and CC2019 (Tables S5 and S6). Although individuals assigned to the Levantine SubRMU were the most prevalent (57/124, 46%), origin distributions varied by FGR (Figure 2d). Lampedusa had the highest proportion assigned to Libya (10.5%), while WAS had the highest proportion assigned to Greece (47.8%). The highest frequency of admixed individuals was found on the Catalan coast (37.5%) and Lampedusa (21.1%).

4 | DISCUSSION

The present study is the first to demonstrate that RMUs (Wallace et al., 2023) are fully genomically supported, confirming a real intermediate level of genetic structuring in the species. Furthermore, our reanalysis of Mediterranean nesting areas confirmed an additional SubRMU level (Barbanti et al., 2026). Overall, we show a hierarchical genetic organisation of loggerhead turtles in at least three levels: populations, SubRMUs and RMUs. Using hierarchic genetic analyses in the assignment of unknown individuals is fundamental when baseline samples include major genetic groups but may lack information of genetic substructure within them. Consequently, future studies may consider analyses of unsampled areas, such as Florida in NWA, using the same methodology, as they may provide unused genetic groups (Silver-Gorges et al., 2026).

Our study shows that genotyping all samples of unknown natal origin simultaneously with baseline samples using the reference genome is necessary (Figure 3). To optimise computational resources and reduce the size of the output file, genotyping software only provides information of polymorphic positions in relation to the reference genome (Li, 2011). The number of recovered loci depends on the sample size and genetic diversity. As shown in our simulations, genotyping small sets of unknown samples independently can substantially reduce markers and compromise the accuracy of the assignment.

A robust baseline is also essential for accurate assignment. Across all approaches, correct assignments occurred only when the baseline captured the genetic variation of the individual's origin; otherwise, assignPop misassigned individuals with high confidence.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

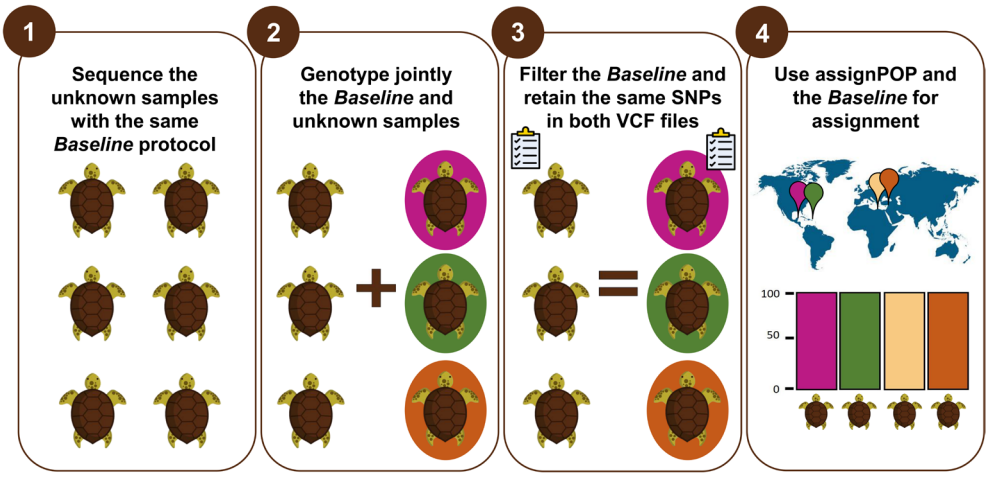

FIGURE 3 Pipeline for assigning individuals of unknown natal origin using a baseline.

This highlights the need for comprehensive baselines covering many nesting sites and hierarchical levels. Our genomic IA outperformed MSA, achieving full assignment due to more markers and a wider baseline, while MSA showed high rates of unassigned individuals (Clusa et al., 2016). Two valid strategies emerge for species with hierarchical genetic structure from our work: (1) a stepwise approach from RMU to SubRMU, removing individuals already assigned to other RMUs before finer analyses; and (2) an all-at-once approach that assigns individuals across both levels simultaneously. Both methods detected admixed individuals among SubRMUs, consistent with previous findings (Barbanti et al., 2026), suggesting that these intermediate assignments reflect true admixture rather than methodological artefacts.

The composition of origins varied markedly among Mediterranean foraging areas, consistent with earlier mitochondrial MSA studies linking this heterogeneity to prevailing currents (Vela-Garcia et al., 2025). Consistent with the hypothesis that migration from nesting to FGRs is shaped by Mediterranean current systems and geographic proximity (Casale & Mariani, 2014), the foraging areas of the Aegean Sea were composed of individuals from Greece and the Levantine region. In particular, the Aegean Sea showed subregional structure: eastern areas were dominated by Levantine individuals, while western areas had a higher contribution of Greek-origin individuals. However, we detected a substantially lower proportion of Atlantic individuals than previously reported (Figure 2). On the Catalan coast a lower percentage of individuals were assigned to the Atlantic in comparison to older samples using MSA (2002–2012: Clusa et al., 2014). Similarly, in Lampedusa a reduction in Atlantic contribution was observed across time (1991–1993: Carreras et al., 2006; 2002–2009: Garofalo et al., 2013). Although comparisons should be made cautiously due to methodological differences and large MSA confidence intervals, these discrepancies may reflect temporal shifts. A decline in Atlantic individuals in western Mediterranean FGRs may result from the recovery of Mediterranean nesting populations (Casale et al., 2018) and the stability or slight increase in Atlantic populations (Nalovic et al., 2021). In addition, the weakening of the Atlantic Meridional Overturning Circulation (Ditlevsen & Ditlevsen, 2023) may limit Atlantic turtles entering the Mediterranean in recent years. The higher prevalence of individuals from the Levantine region in CAT needs further temporal analyses using different size classes and oceanographic modelling to evaluate behavioural patterns. Future studies using the same methodology in additional foraging areas and emerging nesting beaches will be essential to provide a better picture of the migration routes of juveniles in the Mediterranean.

5 | CONCLUSIONS

In this work, we propose a methodology to assign the natal origin of unknown individuals and guide genomic baseline design for highly migratory species (Figure 3). First, use the same library building methodology employed in the baseline in the unknown samples. Minor adjustments (e.g. indexing or pooling) do not affect genotyping results, demonstrating robustness between laboratories. Second, genotype new samples jointly with the baseline. This approach is particularly important for ensuring that the same SNPs are consistently genotyped and retained in the new dataset, thereby enabling the accurate assignment of even a small number of individuals, rather than being limited to large-scale studies. Third, split the VCF into baseline and unknown files, filter the baseline and generate a list of SNPs to retain in both datasets. Fourth, use assignPOP to determine the origin of the unknown individuals. For loggerhead turtles, we provide baseline sequences and a SNP panel for future assignments of unknown samples across the Atlanto-Mediterranean region. Moreover, our methodology is transferable to other non-model species and guides the generation of key information for managing species with complex life cycles.

AUTHOR CONTRIBUTIONS

Astrid Luna-Ortiz, Marta Pascual and Carlos Carreras designed the research. Federico Alberto Abreu-Grobois, Paolo Casale, Daniela

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Freggi, Silvia Giralt, Elizabeth Labastida-Estrada, Salima Machkour-M'Rabet, Adolfo Marco, Dimitris Margaritoulis and Oguz Türkozan provided samples. Astrid Luna-Ortiz, Anna Barbanti and Raúl Llera-Herrera did the laboratory work. Astrid Luna-Ortiz did the genomic data analyses with inputs from Cinta Pegueroles, Marta Pascual and Carlos Carreras. Astrid Luna-Ortiz, Marta Pascual and Carlos Carreras provided funding acquisition. Astrid Luna-Ortiz, Marta Pascual and Carlos Carreras wrote the manuscript with inputs from all authors.

ACKNOWLEDGEMENTS

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, BlueDNA (PID2023-146307OB) funded by MICIU/AEI/10.13039/501100011033 and ERDF/EU and the contract: 'Utilización de herramientas de secuenciación masiva como nuevos marcadores genéticos en tortuga boba (*Caretta caretta*), con el fin de probar su eficacia en la asignación de zonas de origen frente a los marcadores tradicionales, FB 04/2023', in the framework of the project 'LIFE IP-PAF INTEMARES [LIFE15 IPE ES 012]' 'Gestión integrada, innovadora y participativa de la Red Natura 2000 en el medio marino español', coordinated by the Biodiversity Foundation. Marta Pascual and Carlos Carreras (2025 SGR 00608) and Cinta Pegueroles (2025 SGR 01194) are members of the research groups funded by the Generalitat de Catalunya. Dimitris Margaritoulis thanks the ARCHELON staff and volunteers for collecting the samples at the nesting beaches of Greece and at the ARCHELON Sea Turtle Rescue Centre in Glyfada. All the samples had the authorisation of the corresponding country authorities. Astrid Luna-Ortiz dedicates this paper to Juan Luis Flores Colín.

CONFLICT OF INTEREST STATEMENT

The authors declare no financial or personal relationship conflicts.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210x.70453.

DATA AVAILABILITY STATEMENT

2b-RAD raw data can be found at the European Nucleotide Archive (ENA), for the Mediterranean nesting rookeries (project PRJEB77620, from Barbanti et al., 2026), for the Atlantic nesting rookeries (project PRJEB105674) and for the individuals sampled in four Mediterranean foraging grounds (project PRJEB105699). The name of the SNP loci used in the study and obtained by mapping the raw reads to the loggerhead genome (GenBank accession GCA_023653815.1; Chang et al., 2023), the genotypes of Baseline and Unknown individuals, and custom scripts are available via https://doi.org/10.5281/zenodo.23953872 (Luna-Ortiz et al., 2026).

ORCID

Astrid Luna-Ortiz https://orcid.org/0000-0003-4736-551X
Anna Barbanti https://orcid.org/0000-0001-8349-7603
Cinta Pegueroles https://orcid.org/0000-0003-0701-9866
Federico Alberto Abreu-Grobois https://orcid.org/0000-0002-7350-757X
Paolo Casale https://orcid.org/0000-0003-2534-6158
Daniela Freggi https://orcid.org/0000-0001-5457-1050
Silvia Giralt https://orcid.org/0000-0002-7707-883X
Elizabeth Labastida-Estrada https://orcid.org/0000-0002-3237-5566
Raúl Llera-Herrera https://orcid.org/0000-0001-9873-9732
Salima Machkour-M'Rabet https://orcid.org/0000-0002-9505-3900
Adolfo Marco https://orcid.org/0000-0002-6523-9704
Dimitris Margaritoulis https://orcid.org/0009-0001-0368-3260
Oguz Türkozan https://orcid.org/0000-0001-6889-7908
Marta Pascual https://orcid.org/0000-0002-6189-0612
Carlos Carreras https://orcid.org/0000-0002-2478-6445

REFERENCES

Barbanti, A., Pegueroles, C., Margaritoulis, D., Rees, A. F., Turkozan, O., Ulger, C., Snape, R. T. E., Broderick, A. C., Demetropoulos, A., Demetropoulos, S., Hamza, A. A., Levy, Y., Pascual, M., & Carreras, C. (2026). Genomic population structure and adaptation of eastern Mediterranean loggerhead (*Caretta caretta*) nesting areas. *Scientific Reports*, *16*, 30941. https://doi.org/10.1038/s41598-026-62668-z

Barbanti, A., Torrado, H., Macpherson, E., Bargelloni, L., Franch, R., Carreras, C., & Pascual, M. (2020). Helping decision making for reliable and cost-effective 2b-RAD sequencing and genotyping analyses in non-model species. *Molecular Ecology Resources*, *20*, 795–806. https://doi.org/10.1111/1755-0998.13144

Carreras, C., Pascual, M., Cardona, L., Marco, A., Bellido, J. J., Castillo, J. J., Tomás, J., Raga, J. A., Sanfélix, M., Fernández, G., & Aguilar, A. (2011). Living together but remaining apart: Atlantic and Mediterranean loggerhead sea turtles (*Caretta caretta*) in shared feeding grounds. *Journal of Heredity*, *102*, 666–677. https://doi.org/10.1093/jhered/esr089

Carreras, C., Pont, S., Maffucci, F., Pascual, M., Barceló, A., Bentivegna, F., Cardona, L., Alegre, F., SanFélix, M., Fernández, G., & Aguilar, A. (2006). Genetic structuring of immature logger-head sea turtles (*Caretta caretta*) in the Mediterranean Sea reflects water circulation patterns. *Marine Biology*, *149*, 1269–1279. https://doi.org/10.1007/s00227-006-0282-8

Casale, P., Broderick, A. C., Camiñas, J. A., Cardona, L., Carreras, C., Demetropoulos, A., Fuller, W. J., Godley, B. J., Hochscheid, S., Kaska, Y., Lazar, B., Margaritoulis, D., Panagopoulou, A., Rees, A. F., Tomás, J., & Türkozan, O. (2018). Mediterranean sea turtles: Current knowledge and priorities for conservation and research. *Endangered Species Research*, *36*, 229–267. https://doi.org/10.3354/esr00901

Casale, P., & Mariani, P. (2014). The first 'lost year' of Mediterranean sea turtles: Dispersal patterns indicate subregional management units for conservation. *Marine Ecology Progress Series*, *498*, 263–274. https://doi.org/10.3354/meps10640

Chang, G., Jones, S., Leelakumari, S., Ashkani, J., Culibrk, L., O'Neill, K., Tse, K., Cheng, D., Chuah, E., McDonald, H., Kirk, H., Pandoh, P., Pari, S., Angelini, V., Kyle, C., Bertorelle, G., Zhao, Y., Mungall, A., Moore, R., … Jones, S. (2023). The genome sequence of the Loggerhead

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sea turtle, *Caretta caretta* Linnaeus 1758. *F1000Research*, *12*, 336. https://doi.org/10.12688/f1000research.131283.2

Chen, K. Y., Chen, K. Y., Marschall, E. A., Sovic, M. G., Fries, A. C., Gibbs, H. L., & Ludsin, S. A. (2018). *assignPOP*: An R package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution*, *9*, 439–446. https://doi.org/10.1111/2041-210X.12897

Clusa, M., Carreras, C., Pascual, M., Gaughran, S. J., Piovano, S., Avolio, D., Ollano, G., Fernández, G., Tomás, J., Raga, J. A., Aguilar, A., & Cardona, L. (2016). Potential bycatch impact on distinct sea turtle populations is dependent on fishing ground rather than gear type in the Mediterranean Sea. *Marine Biology*, *163*, 122. https://doi.org/10.1007/s00227-016-2875-1

Clusa, M., Carreras, C., Pascual, M., Gaughran, S. J., Piovano, S., Giacoma, C., Fernández, G., Levy, Y., Tomás, J., Raga, J. A., Maffucci, F., Hochscheid, S., Aguilar, A., & Cardona, L. (2014). Fine-scale distribution of juvenile Atlantic and Mediterranean loggerhead turtles (*Caretta caretta*) in the Mediterranean Sea. *Marine Biology*, *161*, 509–519. https://doi.org/10.1007/s00227-013-2353-y

Cooke, S. J., Piczak, M. L., Singh, N. J., Åkesson, S., Ford, A. T., Chowdhury, S., Mitchell, G. W., Norris, D. R., Hardesty-Moore, M., McCauley, D., Hammerschlag, N., Tucker, M. A., Horns, J. J., Reisinger, R. R., Kubelka, V., & Lennox, R. J. (2024). Animal migration in the Anthropocene: Threats and mitigation options. *Biological Reviews*, *99*, 1242–1260. https://doi.org/10.1111/brv.13066

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Ditlevsen, P., & Ditlevsen, S. (2023). Warning of a forthcoming collapse of the Atlantic meridional overturning circulation. *Nature Communications*, *14*, 1–12. https://doi.org/10.1038/s41467-023-39810-w

Galià-Camps, C., Carreras, C., Turon, X., & Pascual, M. (2022). The impact of adaptor selection on genotyping in 2b-RAD studies. *Frontiers in Marine Science*, *9*, 1079839. https://doi.org/10.3389/fmars.2022.1079839

Garofalo, L., Mastrogiacomo, A., Casale, P., Carlini, R., Eleni, C., Freggi, D., Gelli, D., Knittweis, L., Mifsud, C., Mingozzi, T., Novarini, N., Scaravelli, D., Scillitani, G., Oliverio, M., & Novelletto, A. (2013). Genetic characterization of central Mediterranean stocks of the loggerhead turtle *(Caretta caretta)* using mitochondrial and nuclear markers, and conservation implications. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *23*(6), 868–884. https://doi.org/10.1002/aqc.2338

Grant, W. S., Milner, G. B., Krasnowski, P., & Utter, F. M. (1980). Use of biochemical genetic variants for identification of sockeye salmon (*Oncorhynchus nerka*) stocks in Cook Inlet, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, *37*, 1236–1247. https://doi.org/10.1139/f80-159

Hays, G. C., Broderick, A. C., Godley, B. J., Lovell, P., Martin, C., McConnell, B. J., & Richardson, S. (2002). Biphasal long-distance migration in green turtles. *Animal Behaviour*, *64*, 895–898. https://doi.org/10.1006/anbe.2002.1975

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*, 907–915. https://doi.org/10.1038/s41587-019-0201-4

Li, H. A. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Luna-Ortiz, P. A., Pegueroles, C., Pascual, M., & Carreras, C. (2026). *Data and custom scripts from "Natal assignment in highly migratory species"*. Zenodo. https://doi.org/10.5281/zenodo.23953872

Marín-Capuz, G., Crespo-Picazo, J. L., Demetropoulos, S., Garrido, L., Hardwick, J., Jribi, I., Margaritoulis, D., Panagopoulou, A., Patrício, A. R., Robinson, N. J., Pascual, M., Pegueroles, C., & Carreras, C. (2025). Incipient range expansion of green turtles in the Mediterranean. *Molecular Ecology*, *34*, e17790. https://doi.org/10.1111/mec.17790

Nalovic, M. A., Ceriani, S. A., Fuentes, M. M. P. B., Pfaller, J. B., Wildermann, N. E., Uribe-Martínez, A., & Cuevas, E. (2021). *Sea turtles in the North Atlantic & Wider Caribbean Region* (MTSG Annual Regional Report 2021). IUCN-SSC Marine Turtle Specialist Group.

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*, 946–952. https://doi.org/10.1111/1755-0998.12129

Silver-Gorges, I., Komoroske, L. M., Adkins, J., Swenson, J. D., Addison, D. S., Burkholder, D. A., Bagley, D. A., Goodwin, G. D., Hart, K. M., Pfaller, J. B., Shamblin, B. M., & Fuentes, M. M. P. B. (2026). Population structure and genetic stock identification in southeastern United States loggerhead sea turtles (*Caretta caretta*) using genome-wide SNPs. *Conservation Genetics*, *27*, 58. https://doi.org/10.1007/s10592-026-01783-w

Vela-Garcia, H., Karaman, S., Sürücü, B., Turkozan, O., Pascual, M., Pegueroles, C., & Carreras, C. (2025). Where do sea turtles forage in the Mediterranean sea? Filling the gaps for a regional assessment of migration routes. *Frontiers in Marine Science*, *12*, 1661954. https://doi.org/10.3389/fmars.2025.1661954

Wallace, B. P., Posnik, Z. A., Hurley, B. J., DiMatteo, A. D., Bandimere, A., Rodriguez, I., Maxwell, S. M., Meyer, L., Brenner, H., Jensen, M. P., LaCasella, E., Shamblin, B. M., Abreu-Grobois, F. A., Stewart, K. R., Dutton, P. H., Barrios-Garrido, H., Dalleau, M., Dell'amico, F., Eckert, K. L., … Mas, R. B. (2023). Marine turtle regional management units 2.0: An updated framework for conservation and research of wide-ranging megafauna species. *Endangered Species Research*, *52*, 209–223. https://doi.org/10.3354/esr01243

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Flowchart of the two genotyping strategies generating a list of common SNPs for comparison between individuals of the genomic baseline and unknown origin.

Figure S2. Effect of unknown individuals sample size on SNP recovery.

Figure S3. Genetic structure of the baseline nesting locations for the loggerhead sea turtle (*Caretta caretta*) at the Regional Management Unit (RMU) and at the Sub Regional Management Unit (SubRMU) level.

Figure S4. Heatmap of the pairwise genetic differentiation (F_{ST}) between rookeries.

Figure S5. Results of the five different models implemented in 'assignPOP' package when performing self-assignments at RMU level.

Figure S6. Individual assignment probabilities of the unknown individual's dataset using different baseline strategies.

Table S1. Nesting sites and foraging ground locations of the 398 *Caretta caretta* individuals genotyped with 2b-RAD.

Table S2. Genomic data of the 398 individuals of *Caretta caretta*.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table S3. Pairwise F_{ST} values (below diagonal) and their significance (above diagonal) based on 6586 SNPs retained after filtering.

Table S4. Individual assignment of foraging ground samples to Regional Management Units (RMUs) using the 6586 SNPs from the *Caretta Baseline*.

Table S5. Individual assignment of foraging ground samples to SubRegional Management Units (SubRMUs) using the 6586 SNPs from the *MED Baseline*.

Table S6. Individual assignment of foraging ground samples with the all-at-once strategy.

How to cite this article: Luna-Ortiz, A., Barbanti, A., Pegueroles, C., Abreu-Grobois, F. A., Casale, P., Freggi, D., Giralt, S., Labastida-Estrada, E., Llera-Herrera, R., Machkour-M'Rabet, S., Marco, A., Margaritoulis, D., Türkozan, O., Pascual, M., & Carreras, C. (2026). Individual natal assignment in highly migratory species: The genomic baseline and its application in loggerhead turtles. *Methods in Ecology and Evolution*, *00*, 1–9. https://doi.org/10.1111/2041-210x.70453

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70453 by Dimitris Margaritoulis - Cochrane Greece , Wiley Online Library on [25/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License