



# Evidence of backcross inviability and mitochondrial DNA paternal leakage in sea turtle hybrids

Sibelle T. Vilaça<sup>1,2,3</sup>  | Francesco Maroso<sup>3</sup> | Paulo Lara<sup>4</sup> | Benoit de Thoisy<sup>5</sup> | Damien Chevallier<sup>6</sup> | Larissa Souza Arantes<sup>1,2</sup> | Fabricio R. Santos<sup>7</sup> | Giorgio Bertorelle<sup>3</sup> | Camila J. Mazzoni<sup>1,2</sup> 

<sup>1</sup>Department of Evolutionary Genetics, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

<sup>2</sup>Berlin Center for Genomics in Biodiversity Research, Berlin, Germany

<sup>3</sup>Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy

<sup>4</sup>Fundação Projeto Tamar, Mata de São João, Bahia, Brazil

<sup>5</sup>Kwata NGO, Cayenne, French Guiana

<sup>6</sup>BOREA, National Museum of Natural History (MNHN), CNRS 8067, SU, IRD 207, UCN, UA, Martinique, France

<sup>7</sup>Laboratório de Biodiversidade e Evolução Molecular, Departamento de Genética, Ecologia e Evolução, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

## Correspondence

Sibelle T. Vilaça, Dipartimento di Scienze della Vita e Biotecnologie, Università di Ferrara, Via L. Borsari 46, 44121 Ferrara, Italy.

Email: [sibelletores@gmail.com](mailto:sibelletores@gmail.com)

## Funding information

Alexander von Humboldt-Stiftung; Centre National de la Recherche Scientifique; Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Grant/Award Number: 491292795

**Handling Editor:** Katie Lotterhos

## Abstract

Hybridization is known to be part of many species' evolutionary history. Sea turtles have a fascinating hybridization system in which species separated by as much as 43 million years are still capable of hybridizing. Indeed, the largest nesting populations in Brazil of loggerheads (*Caretta caretta*) and hawksbills (*Eretmochelys imbricata*) have a high incidence of hybrids between these two species. A third species, olive ridleys (*Lepidochelys olivacea*), is also known to hybridize although at a smaller scale. Here, we used restriction site-associated DNA sequencing (RAD-Seq) markers, mitogenomes, and satellite-telemetry to investigate the patterns of hybridization and introgression in the Brazilian sea turtle population and their relationship with the migratory behaviours between feeding and nesting aggregations. We also explicitly test if the mixing of two divergent genomes in sea turtle hybrids causes mitochondrial paternal leakage. We developed a new species-specific PCR-assay capable of detecting mitochondrial DNA (mtDNA) inheritance from both parental species and performed ultra-deep sequencing to estimate the abundance of each mtDNA type. Our results show that all adult hybrids are first generation (F1) and most display a loggerhead migratory behaviour. We detected paternal leakage in F1 hybrids and different proportions of mitochondria from maternal and paternal species. Although previous studies showed no significant fitness decrease in hatchlings, our results support genetically-related hybrid breakdown possibly caused by cytonuclear incompatibility. Further research on hybrids from other populations in addition to Brazil and between different species will show if backcross inviability and mitochondrial paternal leakage is observed across sea turtle species.

## KEYWORDS

ddRAD, hybrid breakdown, hybridization, introgression, marine turtles, mitogenomes

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Interspecific hybridization is an important and fascinating evolutionary process in a species' history (Noor & Bennett, 2009). The degree of hybridization and introgression, or repeated backcrossing of hybrids, to one or both parental species in natural systems is limited by reproductive isolating barriers. Increasing evidence supports these as permeable filters to gene flow (Twyford & Ennos, 2012). Therefore, extensive gene flow may occur between species where their distributions overlap. It has been estimated that between 1% to 10% of animals hybridize with at least one species (Mallet, 2005), but low hybrid viability may limit introgression depending on the degree of genomic divergence between parental species (Edmands, 2002). Generally, recently diverged species will produce fertile hybrids more frequently than species separated by longer times (Edmands, 2002). However, exceptions exist especially in reptiles where species diverged tens of million years can still interbreed and produce fertile offspring (Karl et al., 1995; Prager & Wilson, 1975).

Sea turtles have been swimming in earth's oceans since the Cretaceous period (~100 million years ago). Two families are currently recognized: the Dermochelyidae, represented by one species, the leatherback *Dermochelys coriacea*, and the Cheloniidae, which includes the other six extant species. Cheloniidae species are known to frequently hybridize, despite their ancient initial divergence ~43 million years ago (Karl et al., 1995; Lara-Ruiz et al., 2006; Vilaça et al., 2021). In particular, four Cheloniidae species (loggerhead *Caretta caretta*, hawksbill *Eretmochelys imbricata*, olive ridley *Lepidochelys olivacea*, and green *Chelonia mydas* sea turtles) are known to hybridize and generate viable offspring in Brazil (Karl et al., 1995; Lara-Ruiz et al., 2006; Soares et al., 2017, 2018, 2021; Vilaça et al., 2012). This process is fairly common when compared to other worldwide nesting areas, especially in a small coastal region that comprises the states of Bahia and Sergipe where 32%–42% of nesting hawksbills have been identified as hybrids between hawksbills and loggerheads (Lara-Ruiz et al., 2006; Soares et al., 2017, 2018; Vilaça et al., 2012). Other less common hybrids in Bahia include those between loggerheads×olive ridleys, hawksbills×olive ridleys (Soares et al., 2017, 2018, 2021; Vilaça et al., 2012), and greens×loggerheads (Karl et al., 1995). Interestingly, most molecular studies involving morphologically-identified hawksbills found hybrids with loggerhead mitochondrial DNA (mtDNA), and very few hybrid individuals displayed hawksbill mtDNA (Bass et al., 1996; Brito et al., 2020; Lara-Ruiz et al., 2006; Proietti et al., 2014; Soares et al., 2017, 2018).

Despite the high frequency of hybrids in this Brazilian region, it is still not clear why this is a unique phenomenon there and why those hybrids are almost always the offspring of a female loggerhead and a male hawksbill. Vilaça et al. (2021) raised the possibility that this is a direct cause of population decline, combined with overlapping reproductive peaks. Hawksbills are rarer than loggerheads in the Brazilian coast; hawksbills are classified as Critically Endangered by the International Union for Conservation of Nature

(IUCN) (Mortimer & Donnelly, 2008), while loggerheads are classified as Least Concern in the southwest Atlantic Ocean and globally as Vulnerable (Casale & Tucker, 2017). In the Praia do Forte region (Bahia state) where hybrids have been found in large frequency, the nesting seasons of hawksbills and loggerheads slightly overlap (Marcovaldi & Dei Marcovaldi, 1999). The overlap of reproductive seasons, the temporal encounter between loggerhead females and hawksbill males, and the higher abundance of loggerheads in the Brazilian coastline, are all factors that may contribute to the observed gender-biased hybridization.

Although hybridization between divergent species is generally thought of as maladaptive (Delph & Demuth, 2016), reports of adaptive introgression with exchange of genes between species leading to a significant fitness benefit are rapidly increasing (Edelman & Mallet, 2021). In sea turtles, hybrids do not seem to be at great disadvantage compared to parental species. Analysis of fitness and reproductive output from Brazilian populations with largest proportions of hawksbill×loggerhead and olive ridley×loggerhead hybrids showed that fitness of hybrid hatchlings and reproductive output of hybrids (integrated over 7 years) is not different from both parental species (Soares et al., 2017, 2018, 2021). Furthermore, fitness reduction in the generations after F1, the so-called hybrid breakdown effect, was not observed among hatchlings in the Praia do Forte population (Soares et al., 2018). The lack of evident hybrid breakdown or long-term fitness decrease in loggerhead×hawksbill hatchlings poses an interesting evolutionary question, as species that diverged millions of years ago (Vilaça et al., 2021) apparently have compatible genomes and hybrids with well adapted phenotypes. The combination of two divergent genomes in hybrids can favour the colonization of unique niches (Taylor & Larson, 2019), even though hybrids can also assume intermediate niches or behavioural traits of the two parental species (Vilaça et al., 2019). An important characteristic of sea turtles is the migratory pattern between nesting and foraging sites. Both hawksbills and loggerheads travel long distances (e.g., 1400–2400 km within the Brazilian continental shelf; Marcovaldi et al., 2012), but the feeding areas of each species along the Brazilian coast are distinct. Hybrids were found to migrate to foraging areas typical of both parental species (Marcovaldi et al., 2010, 2012). However, the relationship between the hybrids' genetic composition and their migratory pattern is not clear since a combination of mtDNA sequences (Lara-Ruiz et al., 2006) and few nuclear loci identified them as first generation (F1) (Arantes, Ferreira, et al., 2020; Arantes, Vilaça, et al., 2020; Vilaça et al., 2012). The limits of these studies concern both the use of mtDNA and the number of nuclear loci.

Mitochondrial loci are by far the most common molecular markers used in sea turtle studies (Bowen & Karl, 2007). The mtDNA is a double-stranded circular genome averaging about 16,000 bp and is predominantly inherited as a single locus though the maternal lineage (Ballard & Whitlock, 2004). Previous research showed that mitochondrial heteroplasmy, or the presence of different mtDNA copies within one cell or individual, is common in green sea turtles

and plays an important role in their ability to maintain diversity during population bottlenecks and founder effects (Tikochinski et al., 2020). Although the mtDNA is regarded as having strict maternal inheritance in animals (Ballard & Whitlock, 2004), paternal leakage, the transmission of the paternal mitochondria to offspring, has been shown to occur in various species and is often associated with interspecific hybridization (Fontaine et al., 2007; Kmiec et al., 2006) and also with intraspecific crossings (Konrad et al., 2017; Kvist et al., 2003). Hybridization between divergent species was hypothesized to promote the breakdown of mechanisms that maintain strict maternal inheritance, like cytonuclear controls that limit the persistence of paternally-derived mitochondria (Breton & Stewart, 2015). Although the precise cellular mechanism responsible for enforcing mitochondrial maternal inheritance is unknown, hybridization between extremely divergent species, like sea turtles, poses an interesting model to test if paternal leakage can indeed be observed as a result of mixing two divergent genomes.

Although thousands of sea turtle individuals have been sequenced so far in different studies, the low number of molecular markers did not provide enough resolution to distinguish hybrid generations past F1 in the biggest hybrid population known so far (Soares et al., 2018; Vilaça et al., 2012). Therefore, no studies have provided enough resolution to investigate if the largest Brazilian nesting population is a hybrid swarm (when multiple hybrid generations due to a high proportion of backcross with one parental species or another hybrid), or if it is mostly formed by pure individuals and F1 hybrids. Furthermore, a better characterization of this population may enlighten behavioural patterns such as migration, and possibly improve conservation measures. Genome-wide markers, such as loci from restriction sites with associated DNA sequencing (RAD-seq, Davey et al., 2011), have been widely used in nonmodel species and could provide important genome insights by providing higher resolution in identifying later hybrid generations.

In this study, we investigated the hybridization process between hawksbill and loggerhead sea turtles from different nesting and feeding populations in the Brazilian coast using genomic markers. We were specifically interested in the introgression pattern in the Praia do Forte rookery, Bahia, the relationship between hybridization and migratory behaviour, and mtDNA paternal leakage in sea turtle hybrids. We used the double digest RADseq (ddRAD) technique to genotype individuals with a large panel of single nucleotide polymorphisms (SNPs), and used these genotypes to identify the genomic composition of hybrids and putative pure individuals. We also used telemetry data combined with hybrid generation assignment to infer if migratory patterns can be explained by introgression. To infer paternal leakage, we first sequenced mitogenomes for multiple individuals and then developed a targeted species-specific PCR test capable of detecting mtDNA inheritance from both species, combined with mtDNA ultra-deep sequencing to infer abundance of each mtDNA type.

## 2 | MATERIALS AND METHODS

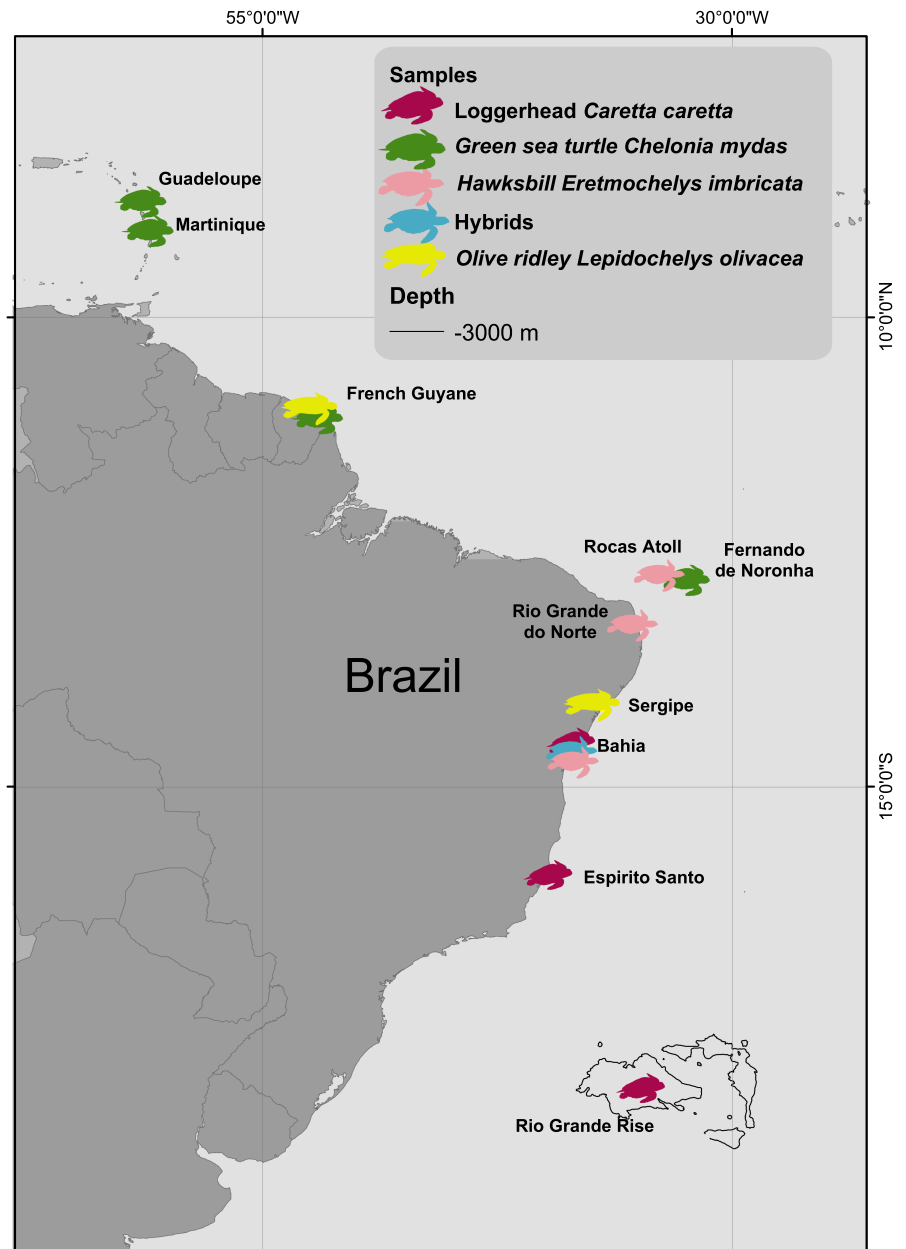
### 2.1 | Sampling

We obtained a total of 129 tissue samples from Brazilian rookery and feeding populations (Figure 1). Samples of 41 individuals (loggerhead = 24, hawksbill = 17) from three Brazilian nesting sites (states of Bahia, Espírito do Santo, and Rio Grande do Norte, Table S1) were used for this study. Samples were first sequenced for their D-loop region of the mitochondrial DNA following Vilaça et al. (2013), which is usually used as a first identification of hybrids. In addition, we analysed 46 additional samples (loggerheads = 14, hawksbills = 20, hybrid = 12) that were used in a previous study based on five nuclear markers and with a previously known hybrid/pure status (Vilaça et al., 2012). All samples were collected by trained members of Fundação Projeto TAMAR. Samples were always identified as a parental species, never as hybrids. To identify potential hybrids with other sea turtle species we also included eight green sea turtles from Brazil, French Guiana, Guadeloupe, and Martinique; and seven olive ridleys from Brazil and French Guiana.

### 2.2 | ddRAD laboratory

DNA extractions were performed using the DNeasy Blood and Tissue kit (Qiagen). Double digest RAD (ddRAD) libraries were produced using the protocol of Peterson et al. (2012) with modifications of adapter sequences according to Meyer and Kircher (2010). In brief, 1 µg of genomic DNA was digested with two enzymes (EcoRI-HF and MseI) at 37°C for at least 2 h in a reaction volume dependent of the DNA concentration per sample. The adapter ligation step was performed immediately after the digestion. The P5 adapter had inline barcodes of varying sizes (5–9 bp) to avoid low diversity problems at the restriction site positions, and P7 adapter was designed with one strand lacking the indexing primer complementary region. Ligated samples were cleaned with 1.8× CleanPCR magnetic beads (CleanNA). A 10-cycle indexing PCR was performed independently for each individual, using one of the 50 indexes in the P7 adapter (Meyer & Kircher, 2010). The PCRs were cleaned with 0.8× CleanPCR magnetic beads, and concentration was measured with Qubit 2.0 using the dsDNA HS assay (Life Technologies). The indexed libraries were equimolarly pooled before the size selection step. The size selection step was performed with the Blue Pippin (Sage Science) using a 1.5% cassette and the R2 marker, targeting fragments between 490 and 610 bp. The final libraries were characterized with a qPCR using the KAPA SYBR FAST (Kapa Biosystems) kit and also checked in the Agilent 2100 sHigh Sensitivity DNA (Agilent Technologies). The libraries were run in the in-house Illumina NextSeq 500 using the 300-cycle mid- or high-output kit.

**FIGURE 1** Map displaying sampling sites of samples used in this study



### 2.3 | ddRAD analysis

Raw reads quality was checked with *FASTQC* (Andrews, 2010) and pre-processed for quality as described in Driller et al. (2020) with modifications to accommodate the lack of overlapping reads. Adapter sequences were trimmed using *Cutadapt* (Martin, 2011). P5 inline barcodes were demultiplexed using *FLEXBAR* version 3.0.3 (Roehr et al., 2017). Paired-end reads were merged using *PEAR* version 0.9.11 (Zhang et al., 2014) setting the maximum length of the assembled sequences to 249 bp to exclude short fragments. Unassembled R1 and R2 reads were trimmed to the maximum length of 140 and 135 bp, respectively, using *Trimmomatic* version 0.3.6 (Bolger et al., 2014). The first five bases of R2 reads were also trimmed, as they had a below average Phred quality score (Q). Only sequences with overall Q > 25 were analysed. We used the *checkRestrictionSites.py* custom

script to check the presence of the *MseI* enzyme at the beginning of reads R1 and discarded all the paired reads with incorrect restriction sites. We also checked for undigested and chimeric sequences and removed all reads containing intact *MseI* and *EcoRI* restriction sites.

After quality check, the processed reads were mapped to a set of ddRAD single copy orthologue loci (R2SCO-*MseI*-*EcoRI*-384-448) for hawksbills and loggerheads from Driller et al. (2020). Reads were mapped using *BOWTIE2* version 2.3.4.1 (Langmead & Salzberg, 2012) with default parameters. The mapped reads were then processed using the reference-based pipeline in *STACKS* 2.2 (Catchen et al., 2013). We used *Stacks*' populations module to obtain loci present in at least one population and in a minimum of 70% of individuals. We also compared if SNPs and microhaplotypes had a similar resolution in defining clusters, since microhaplotypes from RAD loci might have higher power in population assignments (Baetscher et al., 2018). The

called SNPs were used to reconstruct haplotypes using the script `Create_Haplotype_Structure.py`, which was also used to filter loci based on their minimum coverage of 15 reads and the minimum number of populations covering a locus ( $\text{popCOV} = 4$ ). To maximize the number of RAD loci and variants used to describe hybrids, we chose to analyse two data sets. The first, with more SNPs, had only hawksbills, loggerheads and their hybrids. In the second data set, samples of olive ridleys, olive ridley  $\times$  loggerhead hybrids, and green turtles were also included. Both data sets were filtered for SNPs with a minimum depth of  $15\times$  and a minimum allele frequency (MAF) of 0.05.

To identify introgression in individuals and putative population structure, we ran the model-based `STRUCTURE` version 2.3.4 (Pritchard et al., 2000) without population/species a priori information, with admixed model and unlinked sites. The number of simulated clusters ( $K$ ) ranged from 1 to 10, each repeated 10 times; 50,000 iterations as burnin followed by 200,000 iterations. Results were parsed with Evanno's method using the online tool `Structure Harvester` (Earl & vonHoldt, 2012) to determine the most likely  $K$ . To evaluate the consistency of solutions to clustering (major and minor clusters) and summary results for each  $K$  across replicates, we used the online tool `CLUMPAK` (Kopelman et al., 2015). We also ran a nonmodel-based clustering method to compare assignments with `Structure` results. Discriminant analysis of principal components (DAPC) (Jombart et al., 2010) was performed in the `R` (R Core Team, 2017) package `adeigenet` (Jombart & Ahmed, 2011). The number of retained principal components (PC) was optimized with `xValDapc` method. In all cases a low number of PC (5–10) allowed for very good sample assignment and no overfitting. Two or three discriminant axis (DA) were retained depending on their eigenvalues. To estimate pairwise relatedness estimates among all individuals, we used `vcftools` version 1.9 (Danecek et al., 2011) with the command `-relatedness` (Yang et al., 2010).

We used the `NewHybrids` (Anderson & Thompson, 2002) function `GL.NHYBRIDS()` implemented in the `R` package `DARTR` (Gruber et al., 2018) to assign each individual to a "hybrid category" (pure species, F1, F2 or backcross). We extracted SNPs that were fixed in loggerheads and hawksbills ( $F_{ST} = 1$ ) and had a minimum coverage of  $15\times$ . A data set with loggerheads, hawksbills, and their hybrids was used to calculate hybrid category probability. Two hundred loci were selected by the program, either randomly or using average polymorphism information content (PIC) values based on the 200 most informative loci. To verify potential influences of setting prior populations, we ran `NewHybrids` without assigning pure individuals as priors, and also using pure individuals previously identified as nonhybrids by `Structure` and `DAPC`.

We used the  $D$ -statistics to calculate conflicting patterns between ancestral ("A" alleles) and derived ("B" alleles) to distinguish introgression from incomplete lineage sorting (ILS). The  $D$  statistic is based on a tree model of four populations, defined as  $((P1,P2),P3),O$ , where  $P1$ ,  $P2$ , and  $P3$  are ingroups and  $O$  is the outgroup. Excesses in the "ABBA" or "BABA" patterns produce deviation from  $D = 0$ , supporting introgression between  $P2$  and  $P3$  over

the null hypothesis of ILS. We used the package `D-suite` (Malinsky et al., 2021) and the command `Dtrios`. To identify potential origin of hybrids and admixture between sites, samples were divided as population of origin (parental species) based on genetic structure found in previous studies (Shamblin et al., 2014; Vilaça et al., 2013), or as hybrids. Loggerheads were divided in two nesting areas (Bahia and Espírito Santo) and one feeding (Rio Grande Elevation). Hawksbills were divided in two nesting areas (Bahia and Rio Grande do Norte) and one feeding (Rocas Atoll). Olive ridleys were considered a single population (Vilaça et al., 2022). Green turtles were set as outgroup. Significant gene-flow was considered if  $Z\text{-score} > 3$  and assessed using jackknifing. We also estimated the genome fraction involved in the admixture (Malinsky et al., 2021) by computing the  $f_4$ -ratio using the same software.

## 2.4 | Migratory behaviour of hybrids

As some of the individuals analysed were included in previous telemetry studies, we evaluated if their migratory patterns could be related to the fraction of the genome assigned to each parental species. For example, if a hybrid displays a migratory behaviour similar to a pure hawksbill, this could be caused by introgression/backcrossing with a hawksbill (F1  $\times$  hawksbill), as a larger proportion of its genome will have come from hawksbills. The same rationale is valid for a trajectory pattern similar to pure loggerheads. In the case of F1 hybrids, since they have 50% of each parental genome, we hypothesize they could either have a behaviour similar to parental species if they have similar niches/migratory trajectories as at least one of the parental species; or dissimilar to both parentals if they have a diverse or intermediate migratory behaviour. We gathered data from Marcovaldi et al. (2010, 2012) and matched hybrid generation assignment (F1, backcross) from our `Structure` and `NewHybrids` assignments or from previous studies (Soares et al., 2021; Vilaça et al., 2012). For all analysed samples, we had genetic information to confirm their hybrid/pure status, even though not all were sequenced with genome-wide markers. Therefore, we obtained ddRAD data for three hybrids and one hawksbill. For four samples (3 hawksbills, 1 hybrid) genetic information was retrieved from Vilaça et al. (2012) for multiple nuclear loci. All loggerheads ( $n = 7$ ) were previously analysed (Soares et al., 2021) and data for one mitochondrial and one nuclear locus was available. All individuals had their transmitters deployed at Praia do Forte (Bahia). Maps were done in `ArcGIS 10` (ESRI) and minimum convex polygons and core foraging areas followed the methods from Marcovaldi et al. (2010, 2012).

## 2.5 | Mitogenomes

We selected 37 samples (loggerheads = 16, hawksbills = 11, hybrids = 10) from three Brazilian rookeries and feeding sites. The only previous study with population-level mitogenomes of sea turtles showed low genetic diversity outside the  $D$ -loop region in

Caribbean green sea turtles (Shamblin et al., 2012). Given that both loggerheads and hawksbills show generally lower genetic diversity in the mtDNA compared to green sea turtles (Jensen et al., 2019; Shamblin et al., 2014; Vilaça et al., 2013), we expect that our sampling number will recover most of the polymorphisms present in the Brazilian population. Five new primer combinations were developed using two published primers (Abreu-Grobois et al., 2006; Duchene et al., 2012) and six newly developed ones (Table S2). New primers based on previously published mitogenomes (Duchene et al., 2012) were designed for overlapping fragments longer than 3000 bp using PRIMER3 (Untergasser et al., 2012) implemented in GENEIOUS 8.1.3 (Kearse et al., 2012). The amplification of four fragments consisted of a 40 µl mix containing 100–200 ng of genomic DNA, 1× Herculanase II reaction buffer, 0.4 µl of dNTPs (at 25 mM each), 1 µl of each primer at 10 mM, 1% BSA, 1% DMSO and 0.8 µl of Herculanase II Fusion DNA polymerase. The amplification reaction consisted of 95°C for 3 min, followed by 35 cycles of 94°C for 30 s, 55/57°C for 30 s, 70°C for 5 min and a final extension of 70°C for 10 min. The fragments were checked in 1% agarose gel stained with Sybr gold. If any sample presented nonspecific amplification, the fragment of expected size was excised from the agarose gel and purified with the GeneJet PCR Purification kit (Thermo Scientific). For each sample, the four fragments were purified with 1.8× CleanPCR magnetic beads, had their concentration measured in QUBIT 2.0 instrument applying the dsDNA assay (Life Technologies) and were equimolarly pooled. Each pool was sheared using a Covaris S220 ultrasonicator (Covaris) to a mean of 400 bp following the manufacturer's recommendations. For the library construction we followed the protocol from Meyer and Kircher (2010). The library was first size-selected in BluePippin using the 1.5% cassette with R2 marker (Sage Science) to a fragment size range between 490 bp to 610 bp. The final library was characterized with a qPCR using the KAPA SYBR FAST (Kapa Biosystems) kit and also checked with the Agilent Bioanalyser 2100 High Sensitivity DNA chip. The sequencing was performed in the in-house NextSeq 500 with 2× 150 bp reads.

To assemble the mitogenomes we used the software GENEIOUS 8.1.3. All sequences were imported into the software and mapped to two references (a hawksbill and a loggerhead) with haplotypes typical of the Atlantic Ocean (GenBank accession numbers JX454986 and JX454984). The mapping was done using the Geneious mapper with Custom sensitivity (minimum overlap identity: 98%, maximum gap size: 50, index word length: 12, allow gaps of maximum: 15%, minimum mapping quality: 30, maximum ambiguity: 4). Annotations were transferred from the references to the newly assembled mitogenomes. The mitogenomes were aligned with MUSCLE (Edgar, 2004) and gaps and polymorphisms were manually checked.

## 2.6 | Paternal leakage detection from PCR

We developed a species-specific PCR-based test to detect if paternal leakage is present in sea turtle hybrids. We amplified a total of 118 samples, including those sequenced for ddRAD and previously

sequenced for mtDNA and nuclear markers (Vilaça et al., 2012, 2013). We tested two olive ridley and green sea turtle samples to verify possible cross-amplification. Hybrids between other species or individuals with haplotypes characteristic of other oceans (CC-A2, CC-A33, CC-A34) present in Rio Grande Elevation feeding area were not included.

The full mitochondrial DNA of Brazilian loggerheads and hawksbills was used to find regions with enough differences to distinguish both species, and where primers could be designed. We used the same software as for the mitogenomes and searched for species-specific primers. We selected one primer pair per species with enough differences that would be species-specific (i.e., amplify one species but not in the other). The primers designed comprise a 2420 bp region between the 16S rRNA and ND2 loci. Primer sequences can be found in Table S3. Even though a smaller region could have been used, we preferred the use a longer region to avoid amplifying nuclear copies of mitochondrial DNA (NUMTs) (Richly & Leister, 2004). Sequences from noncommon Atlantic haplotypes (common haplotypes include, e.g., CC-A4 for loggerheads and Ei8 for hawksbills) or from the Indo-Pacific Ocean might have mutations in the annealing sites, and since we did not have population-level mitogenomes for those haplotypes, we could not confirm if our primers would remain species-specific in individuals with mitochondrial haplotypes in addition to the ones sequenced here. Furthermore, in our amplification and sequencing efforts for mitogenomes, we observed that the same fragments failed to amplify in individuals with other low frequency or non-Atlantic haplotypes, which did not allow us to confirm the specificity of our primers in relation to other haplotypes.

Our species-specific PCR cocktail consisted of a 40 µl mix containing at least 8 ng of genomic DNA, 1× Phusion reaction buffer, 1 µl of dNTPs (at 10 mM each), 1 µl of each primer at 10 mM, and 0.25 µl of Phusion High-Fidelity DNA Polymerase (New England Biolabs). The amplification reaction consisted of 95°C for 3 min, followed by 35 cycles of 94°C for 30 s, 68°C for 30 s, 70°C for 1 min and a final extension of 70°C for 5 min. Each PCR included a blank, and a control sample for each species (negative and positive controls). The fragments were checked in 1% agarose gel stained with Sybr gold.

To perform detection tests in our primer pairs, we quantified the DNA of samples using the QUBIT 2.0 instrument applying the dsDNA assay (Life Technologies). We chose one loggerhead and one hawksbill sample with the same amount of DNA (22 ng/µl). We tested the detection ability of each primer pair by serial diluting these samples (no dilution, 1:10, 1:100, 1:1000, 1:10,000) and amplifying them with their species-specific primer. The capacity of distinguishing both mitochondria was tested by mixing samples of both species in different concentrations (1:1, 1:10, 1:50, 1:100, 10:1, 50:1, 100:1) and amplifying them with the two species-specific primer pairs.

Our analysis consisted in observing the banding pattern for each sample. An individual was considered as having mtDNA paternal leakage only if it showed two bands of correct sizes in amplifications for both species. If a light band was detected, the PCR was repeated to confirm the presence of the second mitochondria. If it failed to show the two bands in the second PCR, the sample was considered as homoplasmic.

## 2.7 | Paternal leakage detection from ultra-deep sequencing

To further confirm the presence in the hybrids of two different mitochondrial sequences and estimate the abundance of each mitochondrial type, we sequenced 40 samples using the Illumina platform. Samples were chosen as representatives of pure and hybrid categories from Brazilian populations that had enough good quality DNA. We used the primer pair corresponding to positions 9518 and 12,629 (genes NADH dehydrogenase subunit 3 to NADH dehydrogenase subunit 5) (Table S2). This fragment was successfully amplified in all samples and did not overlap with the PCR-assay, providing further evidence that the entire mtDNA was inherited in individuals with evidence of paternal leakage. The PCR and library preparation followed the same protocol as for the mitogenome. To confirm the “leakage” pattern, 27 samples were sequenced in two different runs from independent PCRs and libraries. We subsampled 30,000 reads (expected mean coverage = 1000 $\times$ ) and mapped them to mitogenomes of Atlantic hawksbills (GenBank accession no.: JX454986) and loggerheads (JX454984) using stringent mapping parameters so only sequences of its respective species were mapped to each reference. The mapping was done using the Geneious mapper with Custom sensitivity (minimum overlap identity: 98%, maximum gap size: 50, index word length: 12, allow gaps of maximum: 15%, minimum mapping quality: 30, maximum ambiguity: 4). To be considered as an indication of paternal leakage, an individual needed to show sequences with a homogeneous coverage of the entire amplified fragment and a minimum average coverage of three for both species' mitochondrial DNA. Each individual was manually checked for mapping correctness to the right species and mapping patterns.

To infer the abundance of each parental mitochondria, we considered the ratio of sequences mapped to each species-specific mitogenome as a proxy for abundance. We divided the total number of sequences mapped to the less abundant mitochondrial reference

divided by the number of sequences mapped to the most abundant mitochondria.

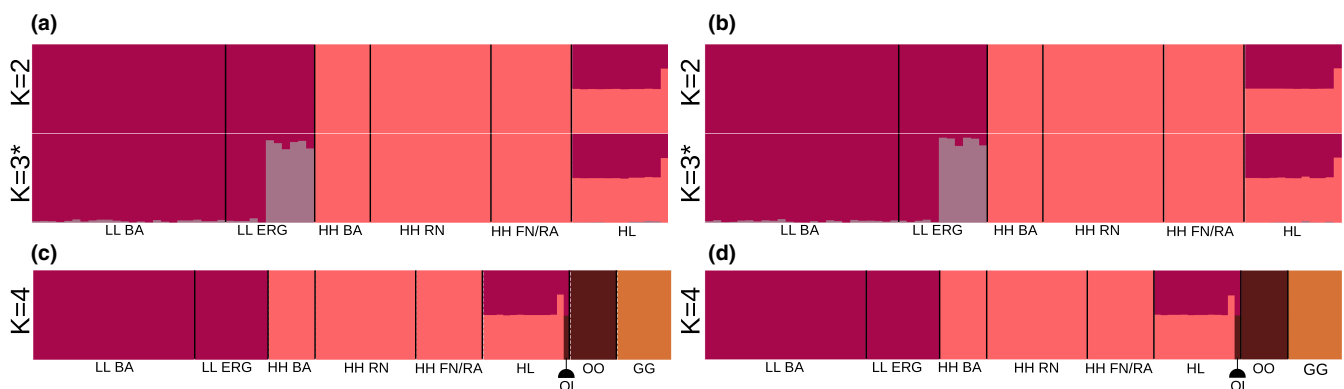
To further confirm the inheritance of mtDNA from both parents, we investigated if the second mitochondria might be derived from a NUMT. In humans, a previous study showed that biparental mtDNA transmission are cryptic mega-NUMTs that resemble paternal leakage (Wei et al., 2020). Thus, we aligned the hawksbill mtDNA to a draft assembly of the loggerhead genome (unpublished, N50 = 20.7 Mb). Alignments were done using MINIMAP2 version 2.17 (Li, 2018) with maximum 5% of sequence divergence (asm5). We also checked the haplotypes obtained for all samples, as NUMTs tend to evolve slower than the mitochondrial counterparts. This means that NUMT sequences tend to be similar to the ancestral mitochondrial sequences, have random polymorphisms with respect to codon positions, and have also stop codons that disrupt the reading frame of protein-coding genes (Baldo et al., 2011).

## 3 | RESULTS

### 3.1 | ddRAD and hybrid identification

For samples with high quality DNA ( $n = 95$ ), we obtained a total of 255,738,252 sequences with an average coverage of 38.5 $\times$  (details can be found in Tables S1 and S4). Our data set for loggerheads, hawksbills and their hybrids ( $n = 79$ ) had a total of 2138 loci and 7845 variants. When including olive ridleys and green sea turtles ( $n = 95$ ), we obtained a total of 11,740 variants in 1368 RAD loci. Details for each processing step following the protocol of Driller et al. (2020) can be found in Table S4.

The Structure results subdivided hawksbills and loggerheads when  $K$  was set to two (Figure 2). All hybrids had a 50% component of each species, indicating that they are all first generation (F1) hybrids. The only exception is a hatchling previously known to be a F1  $\times$  hawksbill backcross. At  $K = 3$ , loggerheads from feeding areas with mtDNA



**FIGURE 2** Proportional membership for hawksbills (HH), loggerheads (LL) and their hybrids (HL) using SNP (a) and haplotype (b) data; and including olive ridleys (OO), green sea turtles (GG) and their hybrids (OL), using SNP (c) and haplotype (d) information. Each individual is represented by a vertical bar, and the length of each bar indicates the probability of membership in each cluster. The asterisk denotes a minor cluster. Sampling sites are abbreviated as: BA, Bahia; ERG, Rio Grande elevation; FN, Fernando de Noronha; RA, Rocas atoll; RN, Rio Grande do Norte

haplotypes from Indo-Pacific (CC-A2, CC-A33, CC-A34) formed a different cluster, a pattern also seen in our relatedness analysis (Figure 3). When considering the four species, at  $K = 4$  they formed separate homogeneous clusters (Figure 2c,d). One individual morphologically identified as a loggerhead and with an olive ridley mtDNA (haplotype Lo67, Vilaça et al. (2022)) was shown to be an olive ridley  $\times$  loggerhead F1 hybrid (OL). We did not observe any differences in cluster definition between SNPs and microhaplotypes (Figure 2).

For the DAPC (Figures S1 and S2) and pairwise relatedness estimates (Figure 3), the results were similar to the Structure analysis, with hybrids clustering between loggerheads and hawksbills, with the exception of the backcross hatchling which is closer to hawksbills.

Individual assignments to a "hybrid category" with NewHybrids classified all hybrids as F1, with the exception of the backcrossed hatchling (Figure 4). This pattern was observed in all prior combinations.

The role of incomplete lineage sorting in generating the observed hybridization patterns was excluded using  $D$ -statistics and  $f_4$ -ratio (Table 1). First generation hawksbills  $\times$  loggerheads hybrids and the single olive ridley  $\times$  loggerhead hybrid had an estimate of admixture proportion ( $f_4$ -ratio) expected for F1 hybrids (~50%). The  $D$ -statistics,  $f_4$ -ratio, and  $Z$ -scores were higher when considering trios when both P1 and P3 were individuals from Bahia (the same region as the hybrids), indicating that the hybridization is a local phenomenon of this population. For the backcrossed hatchling, the  $f_4$ -ratio showed the expected combination of ~25% of genome derived from loggerheads and 75% from hawksbills. No combinations between parental populations had a  $Z$ -score  $>3$ .

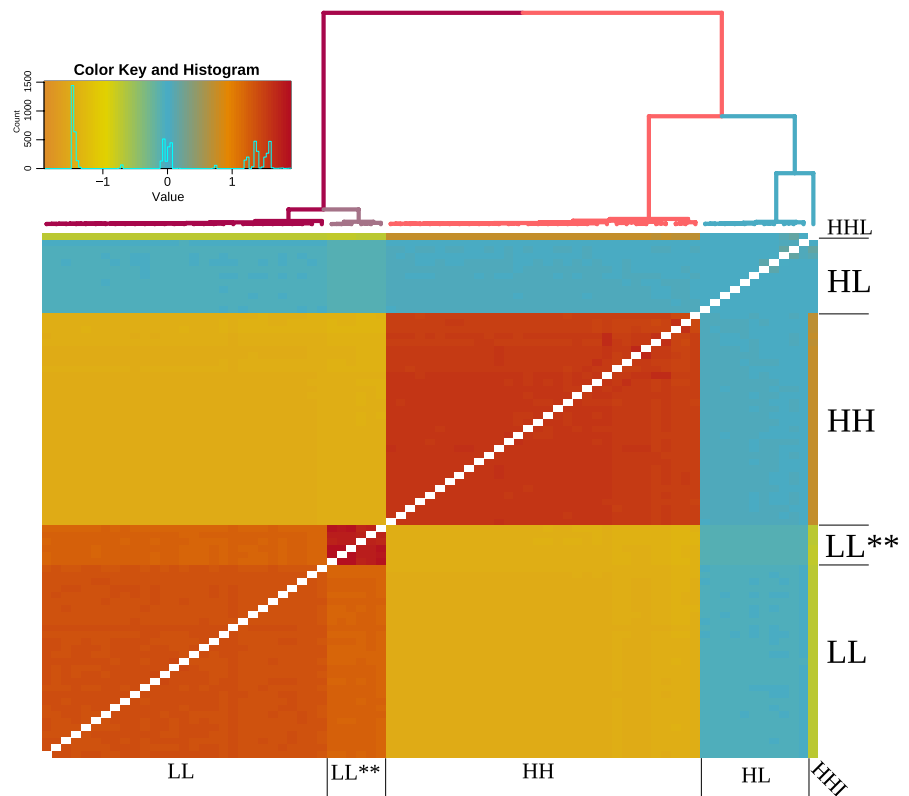
### 3.2 | Migratory patterns

We retrieved data for 15 samples with telemetry data. Our map (Figure 5) shows the post-nesting migration of loggerheads ( $n = 7$ ) are towards more northern sites along the Brazilian coast. Most loggerheads migrated to feeding aggregations near the coast of Ceará state ( $n = 5$ ), up to the Pará coast ( $n = 2$ ). Hawksbills migrated both north ( $n = 3$ ) and south ( $n = 1$ ) of the transmitter deployment site in Praia do Forte, with one individual migrating as far north as Ceará (Figure 5). Most hybrids exhibited a migratory pattern of loggerheads (migrating to northern areas of Pará coast) ( $n = 3$ ) while one hybrid migrated south of Praia do Forte.

Among the individuals with telemetry data, we obtained ddRAD data for one hawksbill and three hybrids. The remaining had only nuclear and mtDNA information. Our genomic data confirmed that three hybrids from Marcovaldi et al. (2012) were all F1 (Figure 5). The hybrid with nuclear genetic information, but no ddRAD data, exhibited a migratory pattern of loggerheads (migrating to northern areas of Pará coast). Furthermore, the individuals assigned as pure did not show indication of introgression in any molecular marker.

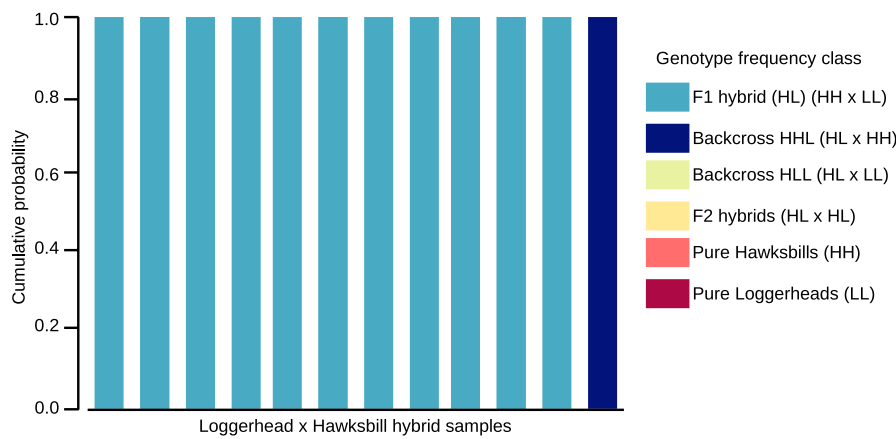
### 3.3 | Mitogenomes

Of the 37 samples from three Brazilian rookeries and feeding sites, we obtained mitogenomes with less than 5% missing data for 15 loggerheads and six hawksbills. All hybrids had a loggerhead mitogenome. The remaining genomes failed to amplify at least one fragment or had low coverage regions where variants could not be called.



**FIGURE 3** Pairwise relatedness analysis between hawksbills (HH), loggerheads (LL), F1 hybrids (HL) and backcrosses (HHL). LL\*\* denotes the loggerhead individuals from the feeding population with Indo-Pacific and South African mtDNA haplotypes





**FIGURE 4** NewHybrids assignments. Only hybrid samples are shown, the remaining samples are all classified as pure hawksbills or loggerheads. The y-axis is the cumulative posterior probability of each individual belonging to one of the six genotype frequency classes

**TABLE 1** *D*-statistics and *f*<sub>4</sub>-ratio results for hybrids and parental population combinations generated with *D*-suite. Since *D*-suite only calculates positive *D* values, only combinations which an excess of allele sharing between P2 and P3 populations are shown

P1	P2	P3	<i>D</i> -stat	Z-score	<i>p</i> -value	<i>f</i> <sub>4</sub> -ratio
Hawksbill Bahia	F1 hybrids	Loggerhead Rookery	0.980	347.9	0	0.501
Hawksbill RN	F1 hybrids	Loggerhead Rookery	0.979	360.7	0	0.501
Hawksbill Feeding	F1 hybrids	Loggerhead Rookery	0.977	309.3	0	0.500
Hawksbill Bahia	F1 hybrids	Loggerhead Feeding	0.976	279.8	0	0.497
Hawksbill RN	F1 hybrids	Loggerhead Feeding	0.976	286.8	0	0.497
Hawksbill Feeding	F1 hybrids	Loggerhead Feeding	0.973	259.7	0	0.497
Hawksbill Bahia	Backcross F1 × HH (hatchling)	Loggerhead Rookery	0.959	137.1	0	0.272
Olive ridley	Hybrid Olive ridley × Loggerhead	Loggerhead Rookery	0.977	151.8	0	0.503

Therefore, we considered for our analysis the 21 mitogenomes with fewer missing data and D-loop haplotypes typical of the Brazilian coast. We found a new D-loop haplotype for loggerheads (Cc-A4.4) which is related to frequent haplotypes present in the Brazilian population (Cc-A4). We also found one polymorphism (A/G) in the ND6 gene that was fixed between the two loggerhead haplotypes common in the Atlantic Ocean (Cc-A4.1 and Cc-A4.2).

### 3.4 | Paternal leakage

Based on our mitogenome sequences, we successfully developed two primer pairs that specifically amplify loggerhead and hawksbill mitochondrial haplotypes typical of the Brazilian coast (Figure 6). The primer detection capacity was as low as a 1:10,000 dilution (Figure S3). The mixing of two loggerhead and hawksbill samples amplified both mitochondria in concentrations of 1:99 and 99:1, respectively. This means that even if the paternal mitochondria were present in small concentrations, we were able to detect it with our PCR assay. Olive ridley samples failed to amplify, but both green turtle samples were amplified by the hawksbill primer pair.

Most loggerhead × hawksbill hybrids (10 out of 12 hybrids) amplified for both fragments, indicating that they have mitochondrial DNA from both species (Table S1). Nonhybrid (“pure”) individuals

(*n* = 98, 46 loggerheads and 52 hawksbills) only amplified for their species-specific mitochondria. One F1 hybrid amplified only for the loggerhead fragment. The remaining loggerhead × hawksbill hybrid did not amplify either fragment even with multiple PCR trials. The olive ridley × loggerhead hybrid did not amplify. Among our samples not sequenced for ddRAD (*n* = 53) because of lower DNA quality/concentration, their amplification patterns agreed with morphology or previous identification using mtDNA and nuclear markers.

### 3.5 | Ultra-deep sequencing of mtDNA paternal leakage

The relative abundance of the mapped reads mirrored the species assignment defined by ddRAD data and the PCR-assay amplifications. In hybrid individuals, the most abundant mitochondrial sequence was always the loggerhead’s (coverage = 390–1150×). This was an expected pattern since the female in hybrid crossings are loggerheads. In hybrids with heteroplasmy, the coverage for the hawksbill mitochondria varied from 3 to 490× and the abundance of the second mitochondria varied from 0.01 to 0.77 (Table S1). In nonhybrid samples, no sample reached the minimum average coverage threshold for the second (“leaked”) mitochondria (minimum

**FIGURE 5** Migratory patterns for hawksbills  $\times$  loggerheads (HL) hybrids ( $n = 4$ ), and nonhybrids hawksbill ( $n = 4$ ) and loggerhead ( $n = 7$ ) sea turtles. Specific types of genetic data (ddRAD or nuclear markers) are represented by different filling of polygons. Polygon area denotes individual interesting areas, as reported in Marcovaldi et al. (2010, 2012). Loggerheads did not have confirmed genetic data information by multiple loci, although their migration pattern was analysed and confirmed as loggerheads in Soares et al. (2021). The triangle represents Praia do forte where samples were collected, and transmitters deployed. Modified from Marcovaldi et al. (2010, 2012)



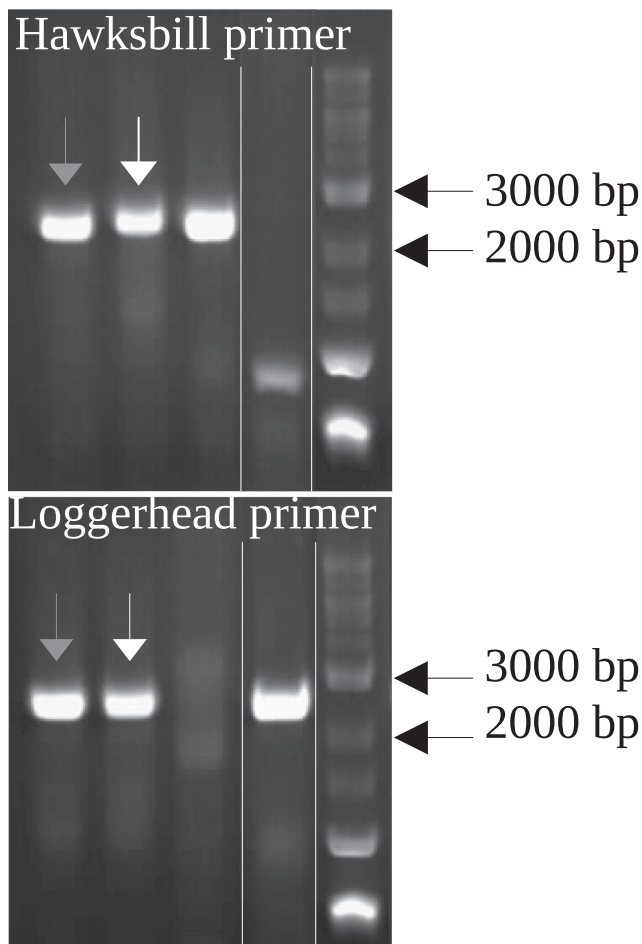
3x), and therefore they were all considered as homoplasmic. The backcross hatchling had no detectable second mtDNA, although its mother had the second mtDNA with 0.01 abundance. Regarding the presence of NUMTS, no regions similar to the hawksbill mtDNA were recovered in the loggerhead draft genome. All mitochondrial sequences had an identical sequence to previously reported mtDNA. We also did not detect frameshift mutations or early stop codons in protein-coding genes that could indicate possible NUMTs. All “leaked” hawksbill mitochondria had an identical sequence for the amplified fragment to other individuals found in the Bahia population. We also did not observe any polymorphisms in the sequenced region that would indicate heteroplasmy (i.e., presence of more than one type of species-specific mtDNA) besides the co-occurrence of two mtDNA types (i.e., paternal leakage) associated with two different species.

## 4 | DISCUSSION

Using genomic markers, here we investigate the natural hybridization between two globally distributed sea turtle species that diverged ~18 million years ago and are still capable of hybridizing in Brazil, and do so frequently. We showed that adult hybrids are F1 with migration patterns similar to loggerheads. We also showed that mitochondrial paternal leakage is present in these divergent hybrids.

### 4.1 | Hybridization

Sea turtle species are separated by several million years (Vilaça et al., 2021), but have highly syntenic genomes and conserved karyotypic organization, including all species that nest in Praia do Forte



**FIGURE 6** Agarose gel showing hybrid individuals (arrows) with amplifications in both hawksbill (upper gel) and loggerhead (bottom gel) mitochondrial DNA specific markers, and two pure samples with only one amplified band corresponding to its respective species. The grey and white arrows denote two hybrid samples. The original gel photo was modified for better visualization and includes only a relevant subset of individuals present in the original image. This figure represents a rearranged and edited version of the original gel image, and no alterations were made beyond subsetting representative individuals of each category. The white lines represent where the lanes were rearranged. All samples shown were run in the same gel and are an accurate representation of the original results, following best practice guidelines from Rossner and Yamada (2004). No brightness nor contrast adjustments were done

(Machado et al., 2020). This may explain why hybridization could occur more frequently among sea turtles, since signatures of ancient hybridization are still found in their genomes (Vilaça et al., 2021), and may be indicate why they can still hybridize (Brito et al., 2020). Using thousands of genomic markers typed in new samples from Brazilian nesting populations and in samples collected previously (Vilaça et al., 2012), we showed that hybrid individuals from Praia do Forte (Bahia) are classified as F1s, and these were all nesting females. The only exception was a single backcross hatchling (F1 × hawksbill backcross) from a known F1 mother. Other populations along the Brazilian coast and continental shelf still did not show evidence of recent hybridization

even with genome-wide markers. Furthermore, using D-statistics and  $f_4$ -ratio we show that this is a local phenomenon in Bahia state.

Previous data from the same population showed no difference in fitness or hatchling viability between hybrids and nonhybrids of both parental species (Soares et al., 2017, 2018), but we did not find later generation hybrids among adult females. This result can be simply explained assuming that the hybridization process we observe started only few decades ago. Considering that sea turtles take 20–40 years to reach sexual maturity (Vilaça et al., 2021), >F1 adult hybrids would not be expected if hybridization started in the last 40–80 years. Conversely, if this is an older phenomenon, the lack of adults >F1 could be due to hybrid breakdown in the backcrossed hatchlings because of genomic incompatibility.

Genetic studies in the Praia do Forte population have been done since the 1980s and the presence of hybrids has been reported in all studies of this rookery (Arantes, Vilaça, et al., 2020; Bass et al., 1996; Brito et al., 2020; Conceição et al., 1990; Karl et al., 1995; Lara-Ruiz et al., 2006; Proietti et al., 2014; Soares et al., 2017, 2018; Vilaça et al., 2012, 2013). However, no adult backcross has been reported yet, indicating that there is evidence of strong hybrid breakdown in the second generation. This is further supported by our RAD results, increasing marker resolution still finds all nesting females are F1s or pure species. Given the lack of >F1 adults, the time passed since the first studies detected hybrids in the 1980s, combined with similar reproductive output of F1 and backcross hybrids compared to nonhybrid individuals, we expect that F2s (F1 × F1) and backcrosses have lower or much lower fitness related to genetic incompatibilities, with no or very rare opportunities to reach sexual maturity. A similar result of lack of >F1 adult hawksbill × loggerhead hybrids was also found in a smaller nesting population in Abrolhos Archipelago (~600 km south of Praia do Forte) where it was suggested a lower reproductive success among hatchlings of F1 hybrids when compared to loggerheads (Arantes, Ferreira, et al., 2020). In general, data from whole genomes indicated that ancient hybridization occurred in sea turtles (Vilaça et al., 2021) and left a signature in their genomes. Such evidence is absent in the ddRAD data, possibly as a consequence of the much lower fraction of the genome analysed by this method. Monitoring this population for the next decades, using a reduced number of genome-wide molecular markers, is necessary to understand if the current hybridization process will produce trans-generation genomic admixture or not.

Although we did not sample male sea turtles, the use of nuclear markers can indirectly recover the paternal perspective of the hybridization. Although juvenile male loggerhead × olive ridley hybrids were found in the Brazilian coast (Medeiros et al., 2019), we have no indication that male loggerhead × hawksbill hybrids are fertile, as no F2s (F1 × F1, as detected by NewHybrids) or backcrosses with male hybrids were found. Based on mtDNA and one nuclear marker sequences for >4800 hatchlings, Soares et al. (2018) observed multiple paternity patterns that could indicate male hybrids are fertile and are capable of siring offspring. Despite the large number of hatchlings, the information from one nuclear marker is not sufficient to confidently distinguish between multiple paternity involving males

of multiple-species  $\times$  male hybrid mating. Thousands of ddRAD loci bring important information to the hybridization dynamics by classifying the individuals as hybrids (and their hybrid generation) and also provide a higher confidence in estimating hybrid generation than few nuclear markers. Future studies using genomic markers in hatchlings or adult males could elucidate if males are indeed capable of reproduce and generate viable offspring.

## 4.2 | Migratory patterns

Here, we show using migratory patterns between nesting and feeding areas that three F1 hybrids migrated to regions known to be loggerheads' feeding aggregations and one F1 hybrid migrated to a hawksbill's aggregation. Although we hypothesized that hybrids with a migratory pattern like one species might have a larger genomic component of that species (i.e., backcrosses would display a migration pattern similar to the species' having higher genomic proportion), here we show that such migratory pattern occurs instead in F1 hybrids. Sea turtles are dependent of migration between feeding and nesting aggregations, and it has been hypothesized that they will follow migratory trajectories that they were exposed as hatchlings (i.e., natal homing imprinting) (Brothers & Lohmann, 2015). Although no genetic components have yet been associated with migratory trajectories, migration directionality may either have a dominance type of inheritance, or it depends on oceanic currents combined with natal homing imprinting (Lohmann et al., 2008; Mansfield et al., 2017).

Previous studies have found hybrid juveniles in hawksbills' feeding areas in southern Brazil, and dispersal simulations showed that hybrids can potentially migrate following both species patterns (Brito et al., 2020; Proietti et al., 2014). Sea turtles are one of the classical examples of species that navigate long distances by sensing Earth's geomagnetic field (Lohmann et al., 2008, 2022). Studies have shown that specific genes are involved in animal magnetoreception and might guide long-distance migration (Wan et al., 2021), and migration routes may also be genetically determined (Gu et al., 2021). Genomic studies of hybrids coupled with satellite tracking might help elucidate if allelic differences in migratory-related genes can explain the differential migration trajectories in hybrids.

## 4.3 | Paternal leakage

Mitochondrial DNA inheritance in vertebrates is considered to be strictly maternal in most animals, although exceptions are found in various taxa (Breton & Stewart, 2015). These unusual patterns have been associated with hybridization events as a consequence of the disruption of the cytonuclear mechanisms that ensures mtDNA's strict maternal inheritance (as discussed below). Here, we report that most fertile sea turtle hybrids have mtDNA from both parental species as a result of paternal leakage. We also show that the most abundant mtDNA is from the mother, loggerheads in all our hybrids, while the hawksbill paternal mtDNA has a lower abundance. This also confirms the directionally mating crosses in the Brazilian population.

The occurrence of mitochondrial paternal leakage happens when the sperm's mitochondria are not eliminated during fertilization, and thus the zygote has both maternal and paternal mitochondria in its cells (Figure S4). Strict mitochondrial inheritance is thought to be controlled by ubiquitin-proteasome cellular machinery. During fertilization, the sperm's mitochondria are tagged with ubiquitin (Rokas et al., 2003) and degraded (autophagy) in the proteasome (Song et al., 2014). The recognition of the mitochondria is thought to be species-specific (Rokas et al., 2003). Therefore, when two divergent genomes are present in the same zygote as a consequence of hybridization, it is not surprising that the species-specific machinery that recognizes the mitochondria fails to degrade the paternal mitochondria. The two species that we investigated diverged ~18 million years ago (Vilaça et al., 2021) and although they have highly syntenic chromosomes, their mitogenome is divergent (~7% divergence) and each species has their characteristic mitochondrial lineages (Dutton et al., 2013; Jensen et al., 2019; Shamblin et al., 2014; Vargas et al., 2016; Vilaça et al., 2022). Thus, it is expected that within-species mechanisms of paternal mitochondrial recognition and elimination are effective within sea turtle species, but not when genomes of two species are present (as in F1 hybrids). Hawksbill  $\times$  loggerhead hybrids showed paternal leakage by both sequencing and the PCR-assay, demonstrating that paternal leakage was often associated with hybridization in sea turtles.

Although the majority of hybrids showed a pattern of paternal leakage, one F1 hybrid only showed loggerhead mtDNA. While our experiments based on newly developed species-specific PCR assay and ultra-deep sequencing can detect low abundance of paternal leakage, we cannot exclude that this hybrid has an abundance of the second mtDNA below our detection threshold. Another possibility is that F1 hybrids may still be viable with one species mtDNA (i.e., no paternal leakage) due to compensatory mechanisms, but backcrosses are not due to a higher imbalance in mitochondrial/nuclear proportions.

The backcrossed sample sequenced in our study (the offspring of a F1 female and hawksbill male) showed no detectable transmission of paternal mitochondria. The lack of leaked mtDNA is consistent with no transmission of the hawksbill mtDNA through the egg of the F1 female, and may indicate that paternal leakage does not happen in backcross mating. Another possibility is that the mother of this hatchling, sequenced in this study and which had the second mitochondria in low abundance (0.01), did not pass on the second mitochondria. Although we did not find evidence of leaked mtDNA in the backcross, we cannot discard the possibility that hawksbill mtDNA is present below our detection threshold. Thus, further studies are needed to understand the dynamics of mtDNA leakage in sea turtle hybrids. Nonetheless, the lack or extremely low abundance of the hawksbill mtDNA may be a contributing factor to why hybrid breakdown is observed in backcrosses. Since backcrossed hatchlings are viable (Soares et al., 2018), the fitness costs may become apparent in a different life stage and this might explain why backcrosses were not found in the adult population. Disruption in F2 hybrids has been associated with cytonuclear coadapted genes in other animal hybrids (Barreto et al., 2018; Han & Barreto, 2021). Higher mitonuclear fitness, as measured by oxidative phosphorylation (OXPHOS) ATP synthesis capacity, was observed in

hybrid crustaceans when nuclear and mitochondrial genes were from the same parental population suggesting cytonuclear adaptations (Han & Barreto, 2021). Sea turtles' mtDNA shows adaptive molecular evolution in genes involved in OXPHOS possibly as a response to active behaviour (Ramos et al., 2020), and the higher presence of loggerhead mtDNA in backcrosses with 75% of hawksbill genome may reduce ATP synthesis capacity and cause hybrid breakdown due to incompatible cytonuclear combinations. Studies using a larger sample set of F1 and backcrossed hatchlings might elucidate if paternal leakage plays a role in fitness and hybrid breakdown.

Our PCR assay is a possible cost-effective screening method for hybrids in the Bahia population. Our assay was designed specifically for Brazilian samples with mitochondrial haplotypes characteristic of the southwest Atlantic Ocean. Given that few mitogenomes are currently available for sea turtles, it is unclear if our assay will amplify or remain species-specific when used in individuals with haplotypes that are not present in the Brazilian loggerhead and hawksbill nesting population. An indication was the failed mitochondrial amplification in some individuals in both long-range mitogenome PCRs and our assay. The sequencing of mitogenomes from more individuals and populations outside the Atlantic Ocean will allow the design of species-specific primers capable of detecting paternal leakage in all sea turtle populations.

#### 4.4 | Hybrids and conservation

Sea turtles are species of conservation concern and hybridization in Brazil's biggest nesting population needs to be better understood. Our results show that although hybrids at a first glance might represent a waste of reproductive effort, since no >F1 adult females are observed, this is a natural process and might have long-term evolutionary importance. Despite the frequent hybridization, we show evidence that there is a strong hybrid breakdown in the second generation and that the Brazilian population is not a hybrid swarm. Consequently, these two species continue to maintain distinct genetic pools and separate evolutionary trajectories. Interspecies hybridization can increase genetic diversity (Ottenburghs, 2021) and exchange adaptive alleles (Edelman & Mallet, 2021). Therefore, our recommendation is to continue the recovery efforts of the Bahia population. We expect that with the growing availability of mating partners, the presence of hybrids in Bahia population should decrease in the next decades. However, climate change is predicted to cause reproductive alterations, changes in species ranges, and skewed sex ratios (Jensen et al., 2018; Lockley & Eizaguirre, 2021), which may increase opportunities of interspecific mating in other populations beyond the Brazilian. Knowing the causes and consequences of hybridization is therefore important to understand how climate change may alter mating opportunities. Nonetheless, because of the uniqueness of this phenomenon, being one of the oldest known lineages capable of natural hybridization in the animal kingdom, they might help answer many pending questions in sea turtle biology and evolution.

#### AUTHOR CONTRIBUTIONS

Experimental design was performed by Sibelle T. Vilaça and Camila J Mazzoni, Sibelle T. Vilaça performed all laboratory work, Sibelle T. Vilaça and Francesco Maroso analysed the data, Paulo Lara analysed migration data, Larissa Souza Arantes performed ddRAD data preprocessing, Benoit de Thoisy, Damien Chevallier and Fabricio R. Santos provided samples, Giorgio Bertorelle helped interpret data, Sibelle T. Vilaça wrote the manuscript with input from all authors.

#### ACKNOWLEDGEMENTS

This research was supported by an Alexander von Humboldt research fellowship awarded to STV and a CNRS grant "Incubator for interdisciplinary research projects in French Guyana" awarded to DC. Open Access funding enabled and organized by Projekt DEAL. The publication of this article was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 491292795. We thank the BeGenDiv team for help with laboratory and bioinformatic analysis, especially Susan Mbedi and Kirsten Richter for laboratory support. Samples were collected/transported under SISBIO permit 37499-2. Tissue samples were exported from Brazil under CITES permits 13BR011173/DF and 14BR015253/DF and imported into Germany under CITES permits E-013161/13 and E-03346/14, respectively. We thank Alonzo Alfaro-Núñez for sharing primer sequences; and Steven Jones and the CanSeq150 team for access to the loggerhead genome. We thank Fundação Projeto Tamar (SISBIO permit N. 42760) for logistical support in sample collection and contributions to this manuscript.

#### CONFLICT OF INTEREST

We have no conflicts of interest to disclose.

#### DATA AVAILABILITY STATEMENT

All data has been made available on NCBI Sequence Read Archive (SRA) database under BioProject PRJNA857276. All scripts used in the ddRAD analysis are available at <https://github.com/mazzoni-izw/mazzoni-begendiv/>. Other relevant data (e.g., VCF files, reference fastas, input files) are available on Zenodo DOI: [10.5281/zenodo.7038996](https://doi.org/10.5281/zenodo.7038996). The original gel image (Figure 6) is available under request from the corresponding author.

#### ORCID

Sibelle T. Vilaça  <https://orcid.org/0000-0002-6887-4703>

Camila J. Mazzoni  <https://orcid.org/0000-0001-6758-5427>

#### REFERENCES

- Abreu-Grobois, A., Horrocks, J. A., Formia, A., Dutton, P. H., LeRoux, R. a., Velez-Zuazo, X., Soares, L., & Meylan, A. B. (2006). *New mtDNA D-loop primers which work for a variety of marine turtle species may increase the resolution of mixed stock analysis*. Proceedings of the 26th Annual Symposium on Sea Turtle Biology, Book of Abstracts, 179.

- Anderson, E. C., & Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, *160*(3), 1217–1229. [https://doi.org/10.1175/1520-0442\(1992\)0052.0.CO;2](https://doi.org/10.1175/1520-0442(1992)0052.0.CO;2)
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute.
- Arantes, L. S., Ferreira, L. C. L., Driller, M., Repinaldo Filho, F. P. M., Mazzoni, C. J., & Santos, F. R. (2020). Genomic evidence of recent hybridization between sea turtles at Abrolhos archipelago and its association to low reproductive output. *Scientific Reports*, *10*(1), 12847. <https://doi.org/10.1038/s41598-020-69,613-8>
- Arantes, L. S., Vilaça, S. T., Mazzoni, C. J., & Santos, F. R. (2020). New genetic insights about hybridization and population structure of hawksbill and loggerhead turtles from Brazil. *Journal of Heredity*, *111*(5), 444–456. <https://doi.org/10.1093/jhered/esaa024>
- Baetscher, D. S., Clemente, A. J., Ng, T. C., Anderson, E. C., & Garza, J. C. (2018). Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, *18*(2), 296–305. <https://doi.org/10.1111/1755-0998.12737>
- Baldo, L., de Queiroz, A., Hedin, M., Hayashi, C. Y., & Gatesy, J. (2011). Nuclear-mitochondrial sequences as witnesses of past interbreeding and population diversity in the jumping bristletail mesomachilis. *Molecular Biology and Evolution*, *28*(1), 195–210. <https://doi.org/10.1093/molbev/msq193>
- Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology*, *13*(4), 729–744. <https://doi.org/10.1046/j.1365-294X.2003.02063.x>
- Barreto, F. S., Watson, E. T., Lima, T. G., Willett, C. S., Edmands, S., Li, W., & Burton, R. S. (2018). Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nature Ecology and Evolution*, *2*, 1250–1257. <https://doi.org/10.1038/s41559-018-0588-1>
- Bass, A. L., Good, D. A., Bjørndal, K. A., Richardson, J. I., Hillis, Z. M., Horrocks, J. A., & Bowen, B. W. (1996). Testing models of female reproductive migratory behaviour and population structure in the Caribbean hawksbill turtle, *Eretmochelys imbricata*, with mtDNA sequences. *Molecular Ecology*, *5*(3), 321–328. <https://doi.org/10.1111/j.1365-294X.1996.tb00323.x>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bowen, B. W., & Karl, S. A. (2007). Population genetics and phylogeography of sea turtles. *Molecular Ecology*, *16*(23), 4886–4907. <https://doi.org/10.1111/j.1365-294X.2007.03542.x>
- Breton, S., & Stewart, D. T. (2015). Atypical mitochondrial inheritance patterns in eukaryotes. *Genome*, *58*(10), 423–431. <https://doi.org/10.1139/gen-2015-0090>
- Brito, C., Vilaça, S. T., Lacerda, A. L., Maggioni, R., Marcovaldi, M. Á., Vélez-Rubio, G., & Proietti, M. C. (2020). Combined use of mitochondrial and nuclear genetic markers further reveal immature marine turtle hybrids along the south western Atlantic. *Genetics and Molecular Biology*, *43*(2), 1–11. <https://doi.org/10.1590/1678-4685-GMB-2019-0098>
- Brothers, J. R., & Lohmann, K. J. (2015). Evidence for geomagnetic imprinting and magnetic navigation in the Natal homing of sea turtles. *Current Biology*, *25*(3), 392–396. <https://doi.org/10.1016/j.cub.2014.12.035>
- Casale, P., & Tucker, A. D. (2017). *Caretta caretta*. (amended version published in 2015). <https://doi.org/10.2305/IUCN.UK.2017-2.RLTS.T3897A119333622.en>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140.
- Conceição, M. B., Levy, J. A., Marins, L. F., & Marcovaldi, M. A. (1990). Electrophoretic characterization of a hybrid between *Eretmochelys imbricata* and *Caretta caretta* (Cheloniidae). *Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology*, *97*(2), 275–278. [https://doi.org/10.1016/0305-0491\(90\)90280-7](https://doi.org/10.1016/0305-0491(90)90280-7)
- 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*(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499–510. <https://doi.org/10.1038/nrg3012>
- Delph, L. F., & Demuth, J. P. (2016). Haldane's rule: Genetic bases and their empirical support. *Journal of Heredity*, *107*(5), 383–391. <https://doi.org/10.1093/jhered/esw026>
- Driller, M., Vilaça, S. T., Arantes, L. S., Carrasco-Valenzuela, T., Heeger, F., Chevallier, D., De Thoisy, B., & Mazzoni, C. (2020). Optimization of ddRAD-like data leads to high quality sets of reduced representation single copy orthologs (R2SCOs) in a sea turtle multi-species analysis. *BioRxiv*, 202004.03.024331. <https://doi.org/10.1101/2020.04.03.024331>
- Duchene, S., Frey, A., Alfaro-Núñez, A., Dutton, P. H., Thomas, M., & Morin, P. A. (2012). Marine turtle mitogenome phylogenetics and evolution. *Molecular Phylogenetics and Evolution*, *65*(1), 241–250. <https://doi.org/10.1016/j.ympev.2012.06.010>
- Dutton, P. H., Roden, S. E., Stewart, K. R., LaCasella, E., Tiwari, M., Formia, A., ... Allman, P. (2013). Population stock structure of leatherback turtles (*Dermodochelys coriacea*) in the Atlantic revealed using mtDNA and microsatellite markers. *Conservation Genetics*, *14*(3), 625–636. <https://doi.org/10.1007/s10592-013-0456-0>
- Earl, D. A., & von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Edelman, N. B., & Mallet, J. (2021). Prevalence and adaptive impact of introgression. *Annual Review of Genetics*, *55*(1), 265–283. <https://doi.org/10.1146/annurev-genet-021821-020805>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Edmands, S. (2002). Does parental divergence predict reproductive compatibility? *Trends in Ecology and Evolution*, *17*(11), 520–527. [https://doi.org/10.1016/S0169-5347\(02\)02585-5](https://doi.org/10.1016/S0169-5347(02)02585-5)
- Fontaine, K. M., Cooley, J. R., & Simon, C. (2007). Evidence for paternal leakage in hybrid periodical cicadas (Hemiptera: *Magicicada* spp.). *PLoS One*, *2*(9), e892. <https://doi.org/10.1371/journal.pone.0000892>
- Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dart: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, *18*(3), 691–699. <https://doi.org/10.1111/1755-0998.12745>
- Gu, Z., Pan, S., Lin, Z., Hu, L., Dai, X., Chang, J., Xue, Y., Su, H., Long, J., Sun, M., Ganusevich, S., Sokolov, V., Sokolov, A., Pokrovsky, I., Ji, F., Bruford, M. W., Dixon, A., & Zhan, X. (2021). Climate-driven flyway changes and memory-based long-distance migration. *Nature*, *591*(7849), 259–264. <https://doi.org/10.1038/s41586-021-03265-0>
- Han, K.-L., & Barreto, F. S. (2021). Pervasive Mitonuclear coadaptation underlies fast development in Interpopulation hybrids of a marine crustacean. *Genome Biology and Evolution*, *13*(3), evab004. <https://doi.org/10.1093/gbe/evab004>
- Jensen, M. P., Allen, C. D., Eguchi, T., Bell, I. P., LaCasella, E. L., Hilton, W. A., Hof, C. A. M., & Dutton, P. H. (2018). Environmental warming and feminization of one of the largest sea turtle populations in the world. *Current Biology*, *28*(1), 154–159.e4. <https://doi.org/10.1016/j.cub.2017.11.057>
- Jensen, M. P., FitzSimmons, N. N., Bourjea, J., Hamabata, T., Reece, J., & Dutton, P. H. (2019). The evolutionary history and global

- phylogeography of the green turtle (*Chelonia mydas*). *Journal of Biogeography*, 46(5), 860–870. <https://doi.org/10.1111/jbi.13483>
- Jombart, T., & Ahmed, I. (2011). ADEGENET 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations (DAPC). *BMC Genetics*, 11, 94.
- Karl, S. A., Bowen, B. W., & Avise, J. C. (1995). Hybridization among the ancient mariners: Characterization of marine turtle hybrids with molecular genetic assays. *Journal of Heredity*, 86(4), 262–268. <https://doi.org/10.1093/oxfordjournals.jhered.a111579>
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Kmieć, B., Wołoszynska, M., & Janska, H. (2006). Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. *Current Genetics*, 50(3), 149–159. <https://doi.org/10.1007/s00294-006-0082-1>
- Konrad, A., Thompson, O., Waterston, R. H., Moerman, D. G., Keightley, P. D., Bergthorsson, U., & Katju, V. (2017). Mitochondrial mutation rate, spectrum and heteroplasmy in *Caenorhabditis elegans* spontaneous mutation accumulation lines of differing population size. *Molecular Biology and Evolution*, 34(6), 1319–1334. <https://doi.org/10.1093/molbev/msx051>
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5), 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Kvist, L., Martens, J., Nazarenko, A. A., & Orell, M. (2003). Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution*, 20(2), 243–247. <https://doi.org/10.1093/molbev/msg025>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lara-Ruiz, P., Lopez, G. G., Santos, F. R., & Soares, L. S. (2006). Extensive hybridization in hawksbill turtles (*Eretmochelys imbricata*) nesting in Brazil revealed by mtDNA analyses. *Conservation Genetics*, 7(5), 773–781. <https://doi.org/10.1007/s10592-005-9102-9>
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, 34, 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Lockley, E. C., & Eizaguirre, C. (2021). Effects of global warming on species with temperature-dependent sex determination: Bridging the gap between empirical research and management. *Evolutionary Applications*, 14, 2361–2377. <https://doi.org/10.1111/eva.13226>
- Lohmann, K. J., Goforth, K. M., Mackiewicz, A. G., Lim, D. S., & Lohmann, C. M. F. (2022). Magnetic maps in animal navigation. *Journal of Comparative Physiology A*, 208, 41–67. <https://doi.org/10.1007/s00359-021-01529-8>
- Lohmann, K. J., Putman, N. F., & Lohmann, C. M. F. (2008). Geomagnetic imprinting: A unifying hypothesis of long-distance natal homing in salmon and sea turtles. *Proceedings of the National Academy of Sciences*, 105(49), 19096–19101. <https://doi.org/10.1073/pnas.0801859105>
- Machado, C. R. D., Domit, C., Pucci, M. B., Gazolla, C. B., Glugoski, L., Nogaroto, V., & Vicari, M. R. (2020). Heterochromatin and microsatellites detection in karyotypes of four sea turtle species: Interspecific chromosomal differences. *Genetics and Molecular Biology*, 43(4), 1–9. <https://doi.org/10.1590/1678-4685-GMB-2020-0213>
- Malinsky, M., Matschiner, M., & Svoldal, H. (2021). Dsuite - fast D -statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, 21(2), 584–595. <https://doi.org/10.1111/1755-0998.13265>
- Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, 20(5), 229–237. <https://doi.org/10.1016/j.tree.2005.02.010>
- Mansfield, K. L., Mendilaharsu, M. L., Putman, N. F., Dei Marcovaldi, M. A. G., Sacco, A. E., Lopez, G., Pires, T., & Swimmer, Y. (2017). First satellite tracks of South Atlantic Sea turtle 'lost years': Seasonal variation in trans-equatorial movement. *Proceedings of the Royal Society B: Biological Sciences*, 284(1868), 20171730. <https://doi.org/10.1098/rspb.2017.1730>
- Marcovaldi, M. A., & Dei Marcovaldi, G. G. (1999). Marine turtles of Brazil: The history and structure of Projeto TAMAR-IBAMA. *Biological Conservation*, 91(1), 35–41. [https://doi.org/10.1016/S0006-3207\(99\)00043-9](https://doi.org/10.1016/S0006-3207(99)00043-9)
- Marcovaldi, M. A., Lopez, G., Soares, L., & López-Mendilaharsu, M. (2012). Satellite tracking of hawksbill turtles *Eretmochelys imbricata* nesting in northern Bahia, Brazil: Turtle movements and foraging destinations. *Endangered Species Research*, 17(2), 123–132. <https://doi.org/10.3354/esr00421>
- Marcovaldi, M. A., Lopez, G. G., Soares, L. S., Lima, E. H. S. M., Thomé, J. C. A., & Almeida, A. P. (2010). Satellite-tracking of female loggerhead turtles highlights fidelity behavior in northeastern Brazil. *Endangered Species Research*, 12(3), 263–272. <https://doi.org/10.3354/esr00308>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, 17(1), 10–12. <https://doi.org/10.14806/ej.17.1.200>
- Medeiros, L., Monteiro, D. S., Botta, S., Proietti, M. C., & Secchi, E. R. (2019). Origin and foraging ecology of male loggerhead sea turtles from southern Brazil revealed by genetic and stable isotope analysis. *Marine Biology*, 166, 1–14. <https://doi.org/10.1007/s00227-019-3524-2>
- Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 5(6), pdb-prot5448. <https://doi.org/10.1101/pdb.prot5448>
- Mortimer, J. A., & Donnelly, M. (IUCN SSC Marine Turtle Specialist Group) (2008). *Eretmochelys imbricata*. *The IUCN Red List of Threatened Species 2008*. e.T8005A12881238. <http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T8005A12881238.en>
- Noor, M. A. F., & Bennett, S. M. (2009). Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, 103(6), 439–444. <https://doi.org/10.1038/hdy.2009.151>
- Ottensburghs, J. (2021). The genic view of hybridization in the Anthropocene. *Evolutionary Applications*, 14(10), 2342–2360. <https://doi.org/10.1111/eva.13223>
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7(5), e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Prager, E. M., & Wilson, A. C. (1975). Slow evolutionary loss of the potential for interspecific hybridization in birds: A manifestation of slow regulatory evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 72(1), 200–204. <https://doi.org/10.1073/pnas.72.1.200>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- Proietti, M. C., Reisser, J., Marins, L. F., Marcovaldi, M. A., Soares, L. S., Monteiro, D. S., ... Secchi, E. R. (2014). Hawksbill × loggerhead sea turtle hybrids at Bahia, Brazil: Where do their offspring go? *PeerJ*, 2, e255. <https://doi.org/10.7717/peerj.255>
- R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.r-project.org/>

- Ramos, E. K. d. S., Freitas, L., & Nery, M. F. (2020). The role of selection in the evolution of marine turtles mitogenomes. *Scientific Reports*, 10(1), 1–13. <https://doi.org/10.1038/s41598-020-73,874-8>
- Richly, E., & Leister, D. (2004). NUMTs in sequenced eukaryotic genomes. *Molecular Biology and Evolution*, 21(6), 1081–1084. <https://doi.org/10.1093/molbev/msh110>
- Roehr, J. T., Dieterich, C., & Reinert, K. (2017). Flexbar 3.0 - SIMD and multicore parallelization. *Bioinformatics*, 33, 2941–2942. <https://doi.org/10.1093/bioinformatics/btx330>
- Rokas, A., Ladoukakis, E., & Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution*, 18(8), 411–417. [https://doi.org/10.1016/S0169-5347\(03\)00125-3](https://doi.org/10.1016/S0169-5347(03)00125-3)
- Rossner, M., & Yamada, K. M. (2004). What's in a picture? The temptation of image manipulation. *The Journal of Cell Biology*, 166(1), 11–15. <https://doi.org/10.1083/jcb.200406019>
- Shamblin, B. M., Bjørndal, K. A., Bolten, A. B., Hillis-Starr, Z. M., Lundgren, I., Naro-Maciel, E., & Nairn, C. J. (2012). Mitogenomic sequences better resolve stock structure of southern greater Caribbean green turtle rookeries. *Molecular Ecology*, 21(10), 2330–2340. <https://doi.org/10.1111/j.1365-294X.2012.05530.x>
- Shamblin, B. M., Bolten, A. B., Abreu-Grobois, F. A., Bjørndal, K. A., Cardona, L., Carreras, C., Clusa, M., Monzón-Argüello, C., Nairn, C. J., Nielsen, J. T., Nel, R., Soares, L. S., Stewart, K. R., Vilaça, S. T., Türkozan, O., Yilmaz, C., & Dutton, P. H. (2014). Geographic patterns of genetic variation in a broadly distributed marine vertebrate: New insights into loggerhead turtle stock structure from expanded mitochondrial DNA sequences. *PLoS One*, 9(1), e85956. <https://doi.org/10.1371/journal.pone.0085956>
- Soares, L. S., Bjørndal, K., Bolten, A., Wayne, M., Castilhos, J., Weber, M., ... Naro-Maciel, E. (2021). Reproductive output, foraging destinations, and isotopic niche of olive ridley and loggerhead sea turtles, and their hybrids, in Brazil. *Endangered Species Research*, 44, 237–251. <https://doi.org/10.3354/esr01095>
- Soares, L. S., Bjørndal, K. A., Bolten, A. B., de Marcovaldi, M. A. G., Luz, P. B., Machado, R., & Wayne, M. L. (2018). Effects of hybridization on sea turtle fitness. *Conservation Genetics*, 19(6), 1311–1322. <https://doi.org/10.1007/s10592-018-1101-8>
- Soares, L. S., Bolten, A. B., Wayne, M. L., Vilaça, S. T., Santos, F. R., de Marcovaldi, M. A. G., & Bjørndal, K. A. (2017). Comparison of reproductive output of hybrid sea turtles and parental species. *Marine Biology*, 164(1), 1–10. <https://doi.org/10.1007/s00227-016-3035-3>
- Song, W. H., Ballard, J. W. O., Yi, Y. J., & Sutovsky, P. (2014). Regulation of mitochondrial genome inheritance by autophagy and ubiquitin-proteasome system: Implications for health, fitness, and fertility. *BioMed Research International*, 2014(981), 867. <https://doi.org/10.1155/2014/981867>
- Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, 3(2), 170–177. <https://doi.org/10.1038/s41559-018-0777-y>
- Tikochinski, Y., Carreras, C., Tikochinski, G., & Vilaça, S. T. (2020). Population-specific signatures of intra-individual mitochondrial DNA heteroplasmy and their potential evolutionary advantages. *Scientific Reports*, 10(1), 1–14. <https://doi.org/10.1038/s41598-019-56,918-6>
- Twyford, A. D., & Ennos, R. A. (2012). Next-generation hybridization and introgression. *Heredity*, 108(3), 179–189. <https://doi.org/10.1038/hdy.2011.68>
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*, 40(15), e115. <https://doi.org/10.1093/nar/gks596>
- Vargas, S. M., Jensen, M. P., Ho, S. Y. W., Mobaraki, A., Broderick, D., Mortimer, J. A., Whiting, S. D., Miller, J., Prince, R. I. T., Bell, I. P., Hoenner, X., Limpus, C. J., Santos, F. R., & Fitzsimmons, N. N. (2016). Phylogeography, genetic diversity, and management units of hawksbill turtles in the indo-pacific. *Journal of Heredity*, 107(3), 199–213. <https://doi.org/10.1093/jhered/esv091>
- Vilaça, S. T., Hahn, A. T., Naro-Maciel, E., Abreu-Grobois, F. A., Bowen, B. W., Castilhos, J. C., Ciofi, C., FitzSimmons, N. N., Jensen, M. P., Formia, A., Limpus, C. J., Natali, C., Soares, L. S., Thoisy, B., Whiting, S. D., & Bonatto, S. L. (2022). Global phylogeography of ridley sea turtles (*Lepidochelys* spp.): Evolution, demography, connectivity, and conservation. *Conservation Genetics*, 23(6), 995–1010. <https://doi.org/10.1007/s10592-022-01465-3>
- Vilaça, S. T., Lara-Ruiz, P., Marcovaldi, M. A., Soares, L. S., & Santos, F. R. (2013). Population origin and historical demography in hawksbill (*Eretmochelys imbricata*) feeding and nesting aggregates from Brazil. *Journal of Experimental Marine Biology and Ecology*, 446, 334–344. <https://doi.org/10.1016/j.jembe.2013.06.004>
- Vilaça, S. T., Lima, C. S., Mazzoni, C. J., Santos, F. R., & de Thoisy, B. (2019). Manatee genomics supports a special conservation area along the Guianas coastline under the influence of the Amazon River plume. *Estuarine, Coastal and Shelf Science*, 226(552), 919. <https://doi.org/10.1016/j.ecss.2019.106286>
- Vilaça, S. T., Piccinno, R., Rota-Stabelli, O., Gabrielli, M., Benazzo, A., Matschiner, M., ... Bertorelle, G. (2021). Divergence and hybridization in sea turtles: Inferences from genome data show evidence of ancient gene flow between species. *Molecular Ecology*, 30(23), 6178–6192. <https://doi.org/10.1111/mec.16113>
- Vilaça, S. T., Vargas, S. M., Lara-Ruiz, P., Molfetti, É., Reis, E. C., Lobo-Hajdu, G., Soares, L. S., & Santos, F. R. (2012). Nuclear markers reveal a complex introgression pattern among marine turtle species on the Brazilian coast. *Molecular Ecology*, 21(17), 4300–4312. <https://doi.org/10.1111/j.1365-294X.2012.05685.x>
- Wan, G., Hayden, A. N., Iiams, S. E., & Merlin, C. (2021). Cryptochrome 1 mediates light-dependent inclination magnetosensing in monarch butterflies. *Nature Communications*, 12(1), 771. <https://doi.org/10.1038/s41467-021-21,002-z>
- Wei, W., Pagnamenta, A. T., Gleadall, N., Sanchis-Juan, A., Stephens, J., Broxholme, J., Tuna, S., Odhams, C. A., Genomics England Research Consortium; NIHR BioResource, Fratter, C., Turro, E., Caulfield, M. J., Taylor, J. C., Rahman, S., & Chinnery, P. F. (2020). Nuclear-mitochondrial DNA segments resemble paternally inherited mitochondrial DNA in humans. *Nature Communications*, 11(1), 1740. <https://doi.org/10.1038/s41467-020-15,336-3>
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7), 565–569.
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, 30(5), 614–620.

## SUPPORTING INFORMATION

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

**How to cite this article:** Vilaça, S. T., Maroso, F., Lara, P., de Thoisy, B., Chevallier, D., Arantes, L. S., Santos, F. R., Bertorelle, G., & Mazzoni, C. J. (2023). Evidence of backcross inviability and mitochondrial DNA paternal leakage in sea turtle hybrids. *Molecular Ecology*, 32, 628–643. <https://doi.org/10.1111/mec.16773>