



Characterization of 25 new microsatellite markers for the green turtle (*Chelonia mydas*) and cross-species amplification in other marine turtle species

Violaine Dolfo¹ · Emilie Boissin^{1,2} · Miri Tatarata³ · Serge Planes^{1,2}

Received: 9 November 2022 / Accepted: 15 February 2023 / Published online: 6 March 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

Background The green sea turtle, *Chelonia mydas*, is a migratory species with a strong natal homing behavior leading to a complex population structure worldwide. The species has suffered severe declines in local populations; it is therefore crucial to understand its population dynamics and genetic structure to adopt appropriate management policies. Here, we describe the development of 25 new microsatellite markers specific to *C. mydas* and suitable for these analyses.

Methods and results They were tested on 107 specimens from French Polynesia. An average allelic diversity of 8 alleles per locus was reported and observed heterozygosity ranged from 0.187 to 0.860. Ten loci were significantly deviant from the Hardy-Weinberg equilibrium, and 16 loci showed a moderate to high level of linkage disequilibrium (4–22%). The overall F_{is} was positive (0.034, p -value < 0.001), and sibship analysis revealed 12 half- or full-sibling dyads, suggesting possible inbreeding in this population. Cross-amplification tests were performed on two other marine turtle species, *Caretta caretta* and *Eretmochelys imbricata*. All loci successfully amplified on these two species, though 1 to 5 loci were monomorphic.

Conclusion These new markers will not only be relevant for further analyses on the population structure of the green turtle and the two other species, but they will also be invaluable for parentage studies, for which a high number of polymorphic loci are necessary. This can provide important insight into male reproductive behavior and migration, an aspect of sea turtle biology that is of critical importance for the conservation of the species.

Keywords Microsatellite · Green turtle · Cross-species transfer · Parentage analysis · French polynesia · *Chelonia mydas*

Introduction

The green turtle, *Chelonia mydas*, is a marine reptile present in all tropical and subtropical waters across the globe. The species exhibits strong philopatry and natal homing: individuals return to the same reproductive ground over years and females tend to nest on the beach from which they

originate [1]. This leads to a complex network of migrations between feeding, reproductive, and nesting grounds, and to a complex genetic structure worldwide with populations breeding separately but occurring at the same feeding grounds [2].

Due to unregulated harvest over the past centuries, green turtle populations are now threatened with extinction according to the IUCN Red List assessment (International Union for Conservation of Nature), with the exception of the Hawaiian and South Atlantic populations which have benefited from conservation actions [3–5]. Despite international and local regulations, the species is still exposed to numerous threats including unsustainable harvest, poaching, coastal degradations, and bycatch [3].

It is therefore important to understand and fully characterize the population structure as well as the reproduction dynamics of the species, to facilitate adequate management

✉ Violaine Dolfo
violaine.dolfo@univ-perp.fr

¹ PSL Research University, EPHE - UPVD - CNRS, UAR 3278 CRIOBE, 66860 Perpignan, France

² Laboratoire d'Excellence « CORAIL », 66860 Perpignan, France

³ Direction de l'Environnement de la Polynésie Française, Papeete, Tahiti, French Polynesia

policies. Mitochondrial DNA has been used to reveal a female-based genetic structure, from oceanic to regional scales [6, 7]. However, this female-inherited marker is unable to provide any insights concerning the gene flow and connectivity driven by the male component of populations or about contemporaneous genetic exchanges. As such, nuclear markers, such as microsatellites, have been used to understand the population structure as a whole. Microsatellites are short repeated DNA motifs of typically 1 to 6 nucleotides that are widely distributed throughout the genome of eukaryotes, whose number of repeats can vary from one individual to another [8]. They are highly polymorphic and exhibit a high mutation rate (10^{-4} to 10^{-3} mutations per locus per generation). They are co-dominant markers, making them suitable for a wide number of genetic analyses [9].

Discrepancies between the degrees of connectivity inferred from mtDNA and microsatellite markers in green sea turtles suggested male mediated connectivity [10–12], or male philopatry [13–15]. In addition to determining population structure, microsatellite markers are useful for parentage analysis [9]. In the case of sea turtles, while individual behaviors are still largely unknown, parentage analyses conducted on samples collected from females and hatchlings can reveal information concerning male reproductive behavior and migration. This provides access to important conservation metrics such as the operational sex ratio (OSR). However, more than 15 loci are needed for accurate parentage inference or population assignments [16, 17]. It is therefore important to have a variety of markers available, with different sizes, repeat motives, and annealing temperatures to facilitate PCR multiplexing and allele scoring in order to conduct such analyses. While the development of a species-specific bank of microsatellites is essential for parentage analysis, creating such a bank is extremely costly and time-consuming [9]. This is why testing the transferability of microsatellites to related species is vitally important. The cross-species transferability of microsatellite markers is variable among taxa, but is usually successful among sea turtles [18–20].

To date, 36 microsatellite markers have been developed specifically for *C. mydas* [21–23], and another 13 which were developed for the loggerhead, the hawksbill, the olive ridley and the Kemp's ridley turtles, have been used on *C. mydas* in population structure and multiple paternity analyses [10, 11, 14, 24–27]. These studies typically used between 2 and 13 markers. For the loggerhead turtle, *Caretta caretta*, 42 markers are available [18, 19, 21]; and 39 have been developed for the hawksbill turtle, *Eretmochelys imbricata* [20, 21, 28, 29]. The aim of the present study was to develop a new set of polymorphic microsatellite markers specific for *C. mydas*, and to test their amplification on two additional

species: the hawksbill turtle, *E. imbricata*, and the loggerhead turtle, *C. caretta*.

Materials and methods

Sample collection

Biopsies were performed on 107 adult specimens of *Chelonia mydas*. Approximately 0.5 cm³ of skin and muscle tissues were collected from the posterior fin of nesting females on Tetiaroa Atoll, French Polynesia, between 2010 and 2021. Samples were stored in 90% ethanol and kept at 4 °C or -20 °C until processing. All the samples were collected by the local NGO *Te mana o te moana* based in Moorea, French Polynesia with authorizations from the Direction of Environment of French Polynesia (DIREN). For cross species amplification, hawksbill and loggerhead turtle were tested for transferability of microsatellite markers. Seventeen individual hawksbill turtles (*E. imbricata*) from French Polynesia corresponding to stranding and seized poaching were provided by DIREN, and 16 individual loggerhead turtles (*C. caretta*) were provided by the Réseau Tortues Marines de Méditerranée Française (RTMMF) stranding network. Loggerhead samples correspond to injured individuals rescued by the RTMMF and strandings from the Mediterranean Sea.

DNA extraction and microsatellite marker design

Total genomic DNA was extracted using the QIAamp 96 DNA QIAcube HT Kit and the QIAcube HT DNA extraction robot (QIAGEN GmbH, Hilden, Germany) following the manufacturer's protocol. The first step was modified as follows: 1 mm² of tissue was digested in 200 µL of digestion buffer with 20% Proteinase K (QIAGEN), and left at 56 °C overnight. Total genomic DNA was quantified using an Epoch BioTek spectrophotometer (Agilent, Santa Clara, US) and an equimolar pool of 8 samples of *C. mydas* (total quantity 3 µg) was sent to GenoScreen (Lille, France) for microsatellite library preparation and sequencing. Samples were sequenced on an Illumina MiSeq platform using a Nano v2 2 × 250 cycles chip. A total of 3361 primer pairs were designed. Among these, 50 pairs were selected and tested based on their repeat number (≥ 6), motif, and PCR product size (> 100 bp). The selected pairs included 23 dinucleotide (DRM), 15 trinucleotide (TRM), and 12 tetranucleotide (TeRM) repeat motifs. For each motif, various ranges of product size were selected in order to minimize overlapping size ranges and facilitate fragment analysis while multiplexing.

Molecular analyses

The 50 selected primer pairs were tested on 107 *C. mydas* individuals at four annealing temperatures (53°C, 57°C, 60°C, and 63°C). PCR amplifications were performed on 4 DNA samples of *C. mydas* for each primer pair and temperature. PCR amplifications were performed using Type-it Microsatellite PCR kit (Qiagen) in 11 µL total volume reactions containing 4 µL Type-it Multiplex PCR Master Mix 2X (contains HotStarTaq® Plus DNA Polymerase, Type-it Microsatellite PCR Buffer with 6 mM MgCl₂ and dNTPs), 5 µL RNase-free water, 1 µL primers (2 µM forward and reverse primers diluted in 1xTE pH 8 buffer) and 1 µL of DNA template at 10–20 ng/µL. PCR cycles consisted of: 5 min at 95°C, followed by 45 cycles of 30 s at 95°C, 1 min 30 s at the annealing temperature, 30 s at 72°C, and a final extension step of 30 min at 60°C. Amplification success was detected on 1.5% agarose gel and visualized with ethidium bromide. Out of these 50 loci, 39 were successfully amplified for at least one temperature including 21 DRM, 12 TRM, and 6 TeRM, and 25 were multiplexed for further characterization on all samples from each of the three species. For multiplexing, forward primers were labeled with a fluorescent dye on the 5' end (ATTO565, ATTO550, FAM, YAKYE; Eurofins Genomics, Ebersberg, Germany) (Table 1). PCR products were sent to GenoScreen and allele sizes were assessed using an Applied Biosystems 3730 Sequencer. For accurate sizing, an internal size ladder (GeneScan 500 LIZ, Applied Biosystems) was used. The 25 selected markers were then tested at optimal annealing temperature on 17 *E. imbricata* individuals and 16 *C. caretta* individuals for cross-species amplification.

Data analysis

Allele sizes were visually assessed using GENEMAPPER software v.5 (Applied Biosystems) on the 107 *C. mydas* individuals. Allele size call consistency over all samples was checked twice, and approximately 5% of the total dataset was read by a second person to compare size calls. All ambiguous peak profiles were considered missing data. MICROCHECKER v.2.2.3 [30] was used to identify null alleles, stuttering errors, and large allele dropout. For each locus, allele frequencies, total number of alleles (Na), private allele number (PA), observed and expected heterozygosities (H_o and H_e), and divergence from Hardy-Weinberg equilibrium were calculated with GenAIEX v.6.503 [31]. The inbreeding coefficient (F_{IS}) and linkage disequilibrium (LD) were calculated using GENETIX v.4.05.2 [32]. LD was calculated between pairs of loci, and the percentage of LD per locus was defined as the percentage of combinations showing significant LD for each locus. To detect

potential siblings in the 107 *C. mydas* individuals from Tetiroa, French Polynesia, the software COLONY v.2.0.6.6 [33] was run 3 times with different starting seeds with a model that allowed for inbreeding and polygamy for both sexes. A full-likelihood analysis method was used with high precision on long runs and no sibship prior. Dyads were considered full-siblings or half-siblings if their probability was greater than 0.95 for the 3 runs.

For *E. imbricata* and *C. caretta* samples, allele sizes were assessed with GENEMAPPER software v.5, and the total number of alleles (Na) and number of private alleles (PA) were calculated with GenAIEX v.6.503. In order to explore whether this set of markers was able to detect genetic variance among species, a Principal Coordinates Analysis (PCoA) was computed in GenAIEX and the Nei unbiased genetic distance was calculated between the three species with GenAIEX v.6.503. Samples that had more than 16% missing data were removed from the PCoA analysis and Nei distance calculation.

Results

Genetic diversity

A panel of 25 polymorphic microsatellite loci was successfully developed and showed clear amplification profiles. However, one locus (CMY12) presented stuttering errors and 19% null alleles and was therefore removed for further analysis. The 24 remaining loci were all polymorphic with a number of alleles per locus ranging from 2 to 17, an average of 8 alleles per locus and a total of 191 alleles (Table 1). 16 loci exhibited a dinucleotide repeat motif (DRM), 7 showed a trinucleotide repeat motif (TRM), and 1 contained a tetranucleotide repeat motif (TeRM). DRM loci showed 4 to 17 alleles (mean: 9, total: 146), while TRM loci displayed 2 to 10 alleles (mean: 4, total: 31), and the TeRM locus had 14 alleles.

MICROCHECKER analyses only revealed the likely occurrence of null alleles for locus CMY07 (6.57%), and no evidence of stuttering errors or large allele dropout were detected on any of the 24 loci. Expected heterozygosity (H_e) ranged from 0.174 to 0.886 (mean: 0.649 ± 0.039), while observed heterozygosity (H_o) was overall slightly lower, ranging from 0.187 to 0.860 (mean: 0.631 ± 0.037) (Table 1). The inbreeding coefficient F_{IS} ranged from -0.088 to 0.196 and was significantly divergent from zero for 6 loci (CMY07, CMY08, CMY16, CMY19, CMY25) (Table 1). Total F_{IS} was also significant (F_{IS} = 0.034, p-value < 0.001). 10 loci (CMY07, CMY09, CMY10, CMY11, CMY14, CMY15, CMY18, CMY19, CMY26, CMY33) deviated significantly from Hardy-Weinberg equilibrium (p-value < 0.05). P-value

Table 1 Characterization of the 24 microsatellites developed for *Chelonia mydas* in this study. Microsatellite markers were included in five multiplexes corresponding to the annealing temperature T_a , number of samples; N_s , number of alleles; H_o , observed heterozygosity; H_e , expected heterozygosity; HWE, deviation from Hardy-Weinberg equilibrium (p-values shown); F_{is} , fixation index; LD, linkage disequilibrium; PA, private alleles, Null alleles: Yes/No.

Locus	Primer Sequence	Repeat motif	Ta (°C)	Multiplex number	Dye	Size range (bp)	N	Na	H_o	H_e	HWE	F_{is}	% LD	Null alleles	PA	GenBank Accession #
CMY03	F: GCTTCTCTACCGCCTCTTG R: TAGTCCACTCCAGCCTCACA	(AC) ₁₇	60	4	FAM	217–255	104	17	0.856	0.886	0.563	0.039	9%	N	4	OQ162049
CMY04	F: TGCAAAATCAGGTGGGTGTA R: ACCAAGTCAAGTGCCTTACAGA	(AC) ₁₆	60	4	ATTO 550	153–171	95	9	0.726	0.757	0.881	0.046	13%	N	5	OQ162050
CMY07	F: ACTATGCTTTGTGCGTTGCC R: GCTCATCACAGTCCCTCCAT	(AC) ₁₄	53	2	YAKYE	109–225	107	12	0.383	0.474	0.000***	0.196***	17%	Y	7	OQ162051
CMY08	F: CCTCTGTGGTCTCAACTGGAC R: GGATCACTTTCCCTGCCACT	(AC) ₁₃	60	4	FAM	117–145	106	7	0.66	0.726	0.221	0.095*	13%	N	3	OQ162052
CMY09	F: AACTTGGCATGTGCACTCAC R: TGTACCTGTGTACAAGACGGT	(AC) ₁₃	57	5	ATTO 550	126–182	104	15	0.731	0.75	0.01**	0.030	13%	N	13	OQ162053
CMY10	F: TGTCCGTAATCAGTACTTCCGTC R: TCCTTCTGTGGTAAACGCCA	(AC) ₁₃	60	3	FAM	145–177	107	11	0.86	0.869	0.004**	0.016	9%	N	4	OQ162054
CMY11	F: TCTGTGTCTAATTTCTGACATTTGT R: ATCATGGCAGTGCACCTGAA	(AC) ₁₃	53	1	YAKYE	231–251	100	8	0.85	0.796	0.026*	-0.062	13%	N	7	OQ162055
CMY14	F: TGCATCCAGTAGATCAGCAGTT R: TCAGCCCAGTCAGCATACAG	(AC) ₁₁	60	3	ATTO 550	135–169	106	14	0.83	0.881	0.045*	0.062	17%	N	9	OQ162057
CMY15	F: CTTCAGGGACAAAGTGCAAT R: GGAACCAACTGCAGAGGAAT	(AG) ₁₁	53	1	FAM	152–164	107	4	0.477	0.491	0.024*	0.034	0%	N	1	OQ162058
CMY16	F: TGGACACAAGGCTTCCAATT R: TACCCGTGAACAAGCCGTGAG	(AC) ₁₁	53	2	ATTO 565	169–213	106	10	0.717	0.81	0.421	0.120**	13%	N	3	OQ162059
CMY17	F: TGGTATCTTGCACGCACA R: GGCAAATGCTCAGAAATAAGTGCA	(AC) ₁₁	60	4	FAM	167–203	106	11	0.717	0.759	0.990	0.060	22%	N	4	OQ162060
CMY18	F: TCCTTCGTTTATAGAAATGCCA R: TGAGATACTCACTGCTCACCC	(AC) ₁₁	60	4	ATTO 565	176–194	92	5	0.641	0.635	0.000***	-0.005	0%	N	3	OQ162061
CMY19	F: GGGCGTGGATGCATATAAGC R: CTGATTTGGCTCAGAATGTTTGT	(AC) ₁₁	53	1	FAM	193–209	105	7	0.667	0.757	0.001**	0.124*	9%	N	4	OQ162062
CMY20	F: TGTGGTCTAGCTCAAAGGGC R: GGAATTGACTTCACAGCAGGG	(AC) ₁₁	60	3	FAM	233–249	106	7	0.642	0.671	0.753	0.048	0%	N	4	OQ162063
CMY21	F: ACTTGAATGCATATTTGACCTGTAT R: GCACATACATCATGCCTGAGC	(AG) ₁₁	53	2	ATTO 565	256–274	106	5	0.689	0.633	0.779	-0.083	4%	N	2	OQ162064
CMY22	F: AACTGGGAGATCTTGACGT R: TGCACCTGCAATCCAAAGAACA	(AC) ₁₁	57	5	ATTO 565	266–278	101	4	0.584	0.558	0.988	-0.042	0%	N	2	OQ162065
CMY25	F: ACTGCCATTCAGAGACCATCA R: ATACTGCAACTTGTGACCT	(AAT) ₁₁	53	1	ATTO 565	173–200	101	5	0.644	0.738	0.123	0.133***	0%	N	1	OQ162066
CMY26	F: ACGGCCTCAGGTAGATGGAT R: AACCATACAACATGCAGCCA	(ATC) ₉	60	3	ATTO 565	198–243	104	10	0.75	0.706	0.000***	-0.057	17%	N	5	OQ162067

Table 1 (continued)

Locus	Primer Sequence	Repeat motif	Ta (°C)	Multiplex number	Dye	Size range (bp)	N	Na	H ₀	H _e	HWE	F _{is}	% LD	Null alleles	PA	GenBank Accession #
CMY27	F: TCCAAGGGCAGAAATTTGGGT R: ACGGGTTATACAGATATGIGGGA	(AAT) ₈	53	1	ATTO 550	118–136	103	2	0.398	0.4	0.955	0.010	0%	N	0	OQ162068
CMY29	F: TGCTTATCAATGATGGTGCC R: TGGATCAAACCCCTCAGTCAGG	(ATC) ₇	57	5	FAM	120–132	107	4	0.607	0.63	0.704	0.040	0%	N	2	OQ162069
CMY32	F: GCGGCATGACTAGGTGTAGG R: GCAITTCCTCTTACATGTCCA	(AAC) ₇	57	5	YAKYE	243–255	106	3	0.274	0.258	0.381	-0.056	0%	N	1	OQ162070
CMY33	F: AAGCTGCTTGATCCTGAGCA R: TGAAGCCTCCAAGGAGCAG	(AAC) ₆	53	2	ATTO 550	114–153	107	4	0.477	0.436	0.000***	-0.088	9%	N	0	OQ162071
CMY35	F: TCTCCCTGGTGTTCATTT R: AGTAGGCACCGACAAACCCAC	(AAC) ₆	60	3	YAKYE	181–190	107	3	0.187	0.174	0.768	-0.072	4%	N	1	OQ162072
CMY45	F: TCCTTACCGGCACTTCTGC R: TGCTCCAAGGTGAAATCTCCA	(AGAT) ₉	53	2	FAM	133–289	104	14	0.769	0.79	0.624	0.031	17%	N	4	OQ162073
CMY12	F: CCAGCTTTGCCACTGAGGTA R: CCCAAACGGAGATCAGGTCC	(AG) ₁₂	57	5	YAKYE	180–190	-	-	-	-	-	-	-	O	-	OQ162056

indicates significant p values (: p < 0.05; **: p < 0.01; ***: p < 0.001).

Table 2 Fullsib and Halfsib dyads of *C. mydas* samples from Tetiaroa. FS: full-siblings; HS: half-siblings. The probability is calculated as the mean probability of the three runs with Colony software

Offspring 1	Sampling date 1	Offspring 2	Sampling date 2	Relationship	Probability
CMY0309	06/01/2015	CMY2742	19/11/2019	FS	1
CMY1100	07/10/2016	CMY3467	30/09/2020	FS	1
CMY2740	17/11/2019	CMY2744	19/11/2019	FS	1
CMY3466	29/09/2020	CMY3468	07/10/2020	FS	1
CMY3472	12/10/2020	CMY3495	11/12/2020	FS	1
CMY0010	01/12/2011	CMY2735	02/11/2019	FS	0.999
CMY1100	07/10/2016	CMY3466	29/09/2020	FS	0.990
CMY2745	28/11/2019	CMY3465	24/09/2020	FS	0.986
CMY1100	07/10/2016	CMY3468	07/10/2020	FS	0.971
CMY1114	03/11/2016	CMY2736	07/11/2019	FS	0.969
CMY2416	12/12/2018	CMY2421	03/01/2019	HS	0.964
CMY3470	11/10/2020	CMY3489	05/12/2020	HS	0.960

was the lowest (<0.001) for CMY07, CMY18, CMY26, and CMY33. Significant linkage disequilibrium was also identified as 8.3% of the pairwise loci combinations showed significant disequilibrium after sequential Bonferroni correction (Online Resource 1). 8 loci (CMY15, CMY18, CMY20, CMY22, CMY25, CMY27, CMY29, CMY32) were not linked with any of the others. The rest of the loci showed a percentage of linkage disequilibrium ranging from 4% (CMY21) to 22% (CMY17). Using all of the loci, Colony revealed 10 pairs of full-siblings and 2 pairs of half-siblings with a probability greater than 0.95 in the *C. mydas* samples from Tetiaroa (Table 2), corresponding to 20 different individuals out of the 107.

Cross-species amplification

The 25 loci were tested for amplification on *E. imbricata* and *C. caretta* samples (Table 3). CMY12 was included as it showed a clear peak profile despite the detection of null alleles and stuttering errors on *C. mydas*. All loci were successfully amplified on both species, but rate of success was dependent on both locus and species. For *C. caretta*, amplification success ranged from 50% (CMY09) to 100% of samples (CMY04, CMY15, CMY17, CMY19, CMY25, CMY35), with an average of 87%. For *E. imbricata*, amplification success was lower (72%) on average, and ranged from 35% (CMY18) to 100% (CMY26). Allele polymorphism was variable across species and locus, ranging from 1 to 10 alleles per locus (Table 3). It was higher in *E. imbricata* with 110 alleles in total, compared to 92 for *C. caretta*. *Chelonia mydas* presented a large number of private alleles (89) with an average of 3.71 per locus. *Eretmochelys imbricata* and *C. caretta* revealed 23 and 25 private alleles, respectively, with an average of 0.92 and 1 private alleles per locus. Five loci were monomorphic for *C. caretta*, one

of which was also monomorphic for *E. imbricata* (CMY11) (Table 3).

For the PCoA analysis and Nei distance calculation, 107 samples of *C. mydas*, 11 samples of *C. caretta*, and 9 samples of *E. imbricata* were conserved as they showed less than 16% missing data. The discarded samples had between 20% and 88% missing data. The PCoA clearly discriminated between the three species, and *C. mydas* appeared distant from the two others on the first axis, which covered the majority of the variance (Fig. 1). Nei unbiased genetic distance was the greatest for *C. mydas* and *C. caretta* (1.575), followed by *C. mydas* and *E. imbricata* (1.221). The distance between *C. caretta* and *E. imbricata* was only 0.646 (Table 4).

Discussion

This study successfully developed 24 microsatellite markers specific to the green sea turtle, *Chelonia mydas*. A total of 191 alleles were retrieved in the dataset. This set is composed of 16 di- (DRM), 7 tri- (TRM), and 1 tetranucleotide repeat motives (TeRM). These markers add to the 36 markers previously developed for *C. mydas* that included 5 DRM, 4 TRM, and 27 TeRM [21–23]. TRM and TeRM are usually sought after due to the larger number of base pair differences between alleles, which decreases the risk of stuttering errors [9]. TeRM are usually less polymorphic than TRM and DRM [9]. However, in this study and in accordance with Dutton & Frey [22], TeRM showed the highest level of polymorphism (14 alleles per locus), followed by DRM (9 alleles per locus) and TRM (4 alleles per locus). With an average of 8 alleles per locus, the level of polymorphism of these loci is robust enough for structure and parentage analyses of *C. mydas* populations [16, 34]. It is comparable with the one found by Dutton & Frey [22] (8.33 alleles per locus), and lower than those found by FitzSimmons

Table 3 Cross-species amplification of the 25 developed loci for two species of marine turtles (*Caretta caretta* and *Eretmochelys imbricata*). N/N_{tot}: number of individuals successfully amplified/number of individuals tested; Na: number of alleles, PA: number of private alleles

Locus	<i>C. caretta</i>				<i>E. imbricata</i>			
	Size	N/N _{tot}	Na	PA	Size	N/N _{tot}	Na	PA
CMY03	225–239	13/16	7	1	235–257	11/17	9	1
CMY04	145–163	16/16	3	1	147–159	13/17	5	2
CMY07	149–181	13/16	6	4	103–203	13/17	6	1
CMY08	123–171	10/16	7	3	123–165	13/17	8	4
CMY09	132–142	8/16	2	0	132–144	9/17	4	1
CMY10	141–155	15/16	5	2	145–165	13/17	10	2
CMY11	235	13/16	1	0	235	11/17	1	0
CMY12	180	15/16	1	0	178–188	16/17	4	3
CMY14	143–163	14/16	5	1	145–153	13/17	2	0
CMY15	158–160	16/16	2	0	158–162	11/17	3	0
CMY16	151–203	11/16	9	4	165–185	13/17	5	2
CMY17	173–185	16/16	4	1	175–191	14/17	7	0
CMY18	180–192	13/16	3	0	180–192	6/17	2	0
CMY19	193–199	16/16	3	1	197–199	11/17	2	0
CMY20	237–239	14/16	2	0	231–241	13/17	4	1
CMY21	256–268	15/16	3	0	256–268	14/17	5	1
CMY22	272	13/16	1	0	272–278	13/17	3	1
CMY25	179–185	16/16	3	0	167–179	12/17	4	2
CMY26	204–225	15/16	4	2	207–228	17/17	4	0
CMY27	133–145	13/16	4	2	127–136	13/17	4	0
CMY29	123	15/16	1	0	123–129	9/17	3	0
CMY32	243	14	1	0	240–249	14/17	3	2
CMY33	132–150	12/16	3	1	126–138	9/17	3	0
CMY35	190–193	16/16	2	1	184–190	11/17	2	0
CMY45	209–265	14/16	10	1	217–245	13/17	7	0

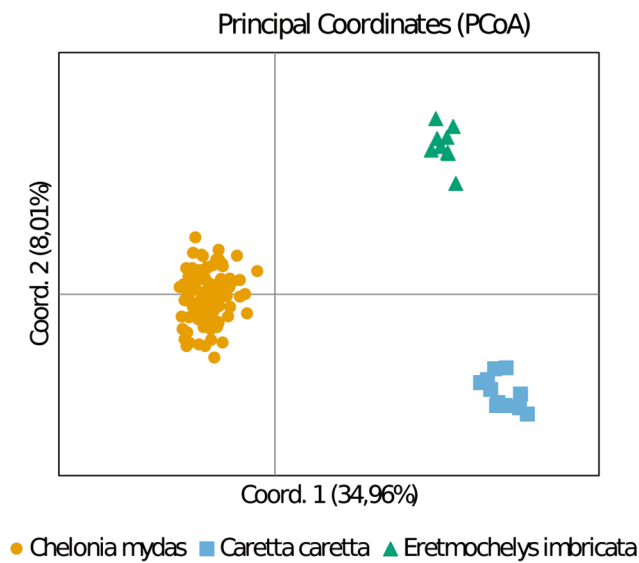


Fig. 1 Principal coordinates analysis on 3 marine turtle species analyzed with 24 newly developed microsatellite markers. Variance explained by each axis is shown in brackets

Table 4 Nei unbiased genetic distance between the three species *C. mydas*, *C. caretta*, and *E. imbricata*

	<i>C. mydas</i>	<i>C. caretta</i>
<i>C. caretta</i>	1.575	-
<i>E. imbricata</i>	1.221	0.646

et al. [21] with 4 DRM (18.5 alleles per locus) and Shamblin et al. [23] with 20 TeRM (12.5 alleles per locus). For other sea turtle species, the level of polymorphism ranges from 5.25 in hawksbill turtles [20] to 11.18 alleles per locus in loggerhead turtles [18].

Levels of observed and expected heterozygosity (H_o and H_e) fall in the lower range of heterozygosity levels published for *C. mydas* populations from other regions [10, 11, 14, 22, 24]. They were closer to those found in the Mediterranean populations (H_o : 0.652–0.671 / H_e : 0.645–0.671) [14]. They were also lower than levels reported in populations of other sea turtle species, such as the loggerhead turtle [18, 35].

Ten out of 24 loci showed a significant departure from the Hardy-Weinberg equilibrium, although the pattern of departure was variable across loci, with 6 loci showing a deficit in heterozygosity and 4 loci exhibiting an excess of heterozygosity (Table 1). Heterozygosity deficiency can be due to selection, population substructure leading to a

Wahlund effect, null alleles, or inbreeding [36]. A spatial Wahlund effect is unlikely because all of the samples were collected on females nesting on the same island of French Polynesia, Tetiaroa. As green turtles exhibit strong natal homing, we can reasonably assume that all of these females are from the same population. Samples were pooled across several nesting seasons between 2010 and 2021, and thus a temporal Wahlund effect is plausible. However, a Wahlund effect would affect all loci [37]. Null alleles were detected with MICROCHECKER at only one locus (CMY07). As this locus also showed a departure from the Hardy-Weinberg equilibrium and a significant F_{is} , it was removed from further analyses on *C. mydas* populations. Null alleles were not detected on any other locus, however CMY19 also presented significant F_{is} and departure from the Hardy-Weinberg equilibrium. This can be a sign of a genotyping artefact, such as null alleles [38], which is why we recommend using this locus with caution and performing the analyses with and without it to rule out any bias. Finally, inbreeding is a possible cause of heterozygosity deficiency in this population, as the total F_{is} is significantly deviant from zero (0.034, p -value < 0.001). The population size is estimated at around 1000 breeding females annually in French Polynesia [39]. Although more recent assessments are needed, the present estimation is coherent with the estimation from annual surveys of Tetiaroa Atoll's nesting population where 20 to 940 nests were recorded annually between 2008 and 2019 [40, 41]. This small estimated population size, coupled with the philopatric behavior of green turtles on reproductive and nesting sites, can lead to inbreeding [42]. The sibship analysis among our samples revealed that 18% of the samples (20 specimens) are engaged in a relationship with at least one full- or half-sibling, thus confirming that significant F_{is} values may, in part, be explained by inbreeding and family structure of this French Polynesian population.

On the other hand, heterozygosity excess is generally associated with missing data or genotyping errors [43]. It can also be due to associative overdominance, if the neutral microsatellite locus is linked with a locus under selection favoring heterozygosity [44]. Missing data were present in CMY11 (6%) and CMY18 (14%). CMY33 and CMY26 might present genotyping errors, and similar to CMY19, we recommend using these loci with caution.

Levels of linkage disequilibrium (LD) are variable across species and populations, and are the result of many forces such as selection, genetic drift, mutation, gene conversion, epistasis, and recombination [45]. This population of *C. mydas* had 8.3% of the pairwise loci combinations showing significant LD, which is moderate compared to other species [45, 46]. Other species of sea turtles however, showed no LD with microsatellite markers in many of their populations [18, 19].

Sibship analysis revealed that most adult female full-siblings were sampled either with a 4-year gap or in the same year, and one dyad was sampled 8 years apart (Table 2). This provides a first indication of the reproduction frequency of green turtle females on Tetiaroa, which seems to be around 4 years. This is consistent with general knowledge on the green sea turtle that defines the reproductive frequency between 2 and 4 years [47]. Although this conclusion is preliminary, a more in-depth parentage analysis on a larger sample size that includes hatchlings will help reveal the reproductive behavior in both males and females. This demonstrates that this set of markers is promising for parentage analysis, which, coupled with field data, will give precious insights into individual behavior.

Cross-species amplification was successful on the two species that were tested, *C. caretta* and *E. imbricata*, with 87% and 72% amplification success, respectively. Cross-species amplification is commonly used for microsatellite analysis in all the sea turtle species with a high rate of success [18–22]. Other reptilian taxa are also known to show a high rate of cross-species amplification success [48], while other marine taxa such as fishes and bivalves show a low rate of cross-species amplification success [38]. The success of cross-species amplification indicates that the flanking regions of microsatellite loci are conserved across sea turtle species, in line with the findings of FitzSimmons et al. [21]. This new set of loci could thus be useful for further studies on other sea turtle species, such as *C. caretta* and *E. imbricata*, but is also likely to amplify successfully on the closely related olive ridley (*Lepidochelys olivacea*) and Kemp's ridley (*L. kempii*) turtles.

Furthermore, all loci revealed one or several private alleles for at least one of the species. The largest number of private alleles was found for *C. mydas*, due to the specific development of this microsatellite marker set for this species. The occurrence of private alleles, although it can be artificially increased by a low sample size, shows that this set of markers can be used to distinguish between the three sea turtle species.

Additionally, the PCoA showed a clear genetic differentiation between the three species. The green turtle appears to be more distant from the two other species, which is confirmed with the Nei unbiased genetic distance. This is in accordance with the phylogenetic distance between the species [49]. The hawksbill and the loggerhead turtles are more closely related and belong to the Carettini tribe, with a split between the species about 29 million years ago. In contrast, their common ancestor with *C. mydas*, which belongs to the Chelonini tribe, is distant from 63 million years.

In conclusion, 23 of the microsatellite loci developed here can be used to assess the genetic variability of *Chelonia mydas* populations and two other species of sea turtle. Most

importantly, this marker set can help to unravel the behavior of both males and females by providing a high number of loci and alleles which are required for robust parentage analyses [17, 34]. This aspect of sea turtle biology, which to date is largely unknown, is critical for the conservation of populations, as increasing global temperatures are already driving a massive feminization within the world's largest green turtle rookery [50].

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11033-023-08341-4>.

Acknowledgements This research was funded by the Direction of Environment of French Polynesia. We would like to thank the volunteers of the NGO Te mana o te moana, coordinated by Cecile Gaspar, for providing *Chelonia mydas* and *Eretmochelys imbricata* samples for this study. We also thank the volunteers of the French Mediterranean Stranding Network for Sea Turtles (Réseau Tortues Marines de Méditerranée Française, coordinated by Claude Miaud) for turtle surveillance and collection of the *Caretta caretta* samples. Thank you to Jeanine Almany for the english corrections.

Author Contribution Sample collection of *Chelonia mydas* was coordinated by Miri Tatarata. Sample processing and analysis was performed by Violaine Dolfo, reviewed by Emilie Boissin, and supervised by Emilie Boissin and Serge Planes. The first draft of the manuscript was written by Violaine Dolfo, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the Direction of the Environment of French Polynesia.

Data Availability The microsatellite sequences generated during the current study are available in the GenBank repository under accession numbers OQ162049–OQ162073. All the other data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval *Chelonia mydas* and *Eretmochelys imbricata* sample collection was authorized and coordinated by the Direction of the Environment of French Polynesia. Samples were exported to France for processing with CITES permits n° FR1298700118-E and n° FR2098700187-E. *Caretta caretta* sample collection was allowed by the decree of December 31, 2012 on the renewal of the environmental protection approval of the Société Herpétologique de France (SHF), and the decree of October 24, 2016 on the collection of biological data in the event of sea turtle stranding or bycatch on French metropolitan coasts (NOR: DEVL1500415N) for the mainland sea turtle observatory program of the National Museum of Natural History (MNHN).

References

1. FitzSimmons NancyN, Moritz C, Limpus CJ, Pope L, Prince R (1997) Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian Green Turtle populations and male-biased Gene Flow. *Genet* 147(4):1843–1854
2. Jensen MP, FitzSimmons NN, Dutton PH, Wyneken J, Lohmann KJ, Musick JA (2013) Boca Raton, Florida: CRC Press; p 135–62
3. Seminoff JA (2004) *Chelonia mydas*. [cité 17 janv 2020]. Disponible sur <https://doi.org/10.2305/IUCN.UK.2004.RLTS.T4615A11037468.en>. The IUCN Red List of Threatened Species. [Internet]. IUCN Red List of Threatened Species
4. Broderick A, Patricio A *Chelonia mydas* (South Atlantic subpopulation). The IUCN Red List of Threatened Species. [Internet]. IUCN Red List of Threatened Species. 2019 [cité 17 janv 2020]. Disponible sur: <https://doi.org/10.2305/IUCN.UK.2019-2.RLTS.T142121866A142086337.en>
5. Chaloupka MY, Pilcher NJ *Chelonia mydas* (Hawaiian subpopulation). The IUCN Red List of Threatened Species. [Internet]. IUCN Red List of Threatened Species. 2019 [cité 17 janv 2020]. Disponible sur: <https://doi.org/10.2305/IUCN.UK.2019-2.RLTS.T16285718A142098300.en>
6. Dutton PH, Jensen MP, Frutchey K, Frey A, LaCasella E, Balazs GH et al (2014) Genetic Stock structure of Green Turtle (*Chelonia mydas*) nesting populations across the Pacific Islands. *Pac Sci* nov 68(4):451–464
7. Ruiz-Urquiola A, Riverón-Giró F, Pérez-Bermudez E, Abreu-Grobois A, González M, James-Petric BL et al (2010) Population genetic structure of Greater Caribbean green turtles (*Chelonia mydas*) based on mitochondrial DNA sequences, with an emphasis on rookeries from southwestern Cuba. *Rev Investig Mar* 8 avr 31:33–52
8. Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol Ecol* 11(12):2453–2465
9. Jones AG, Small CM, Paczolt KA, Ratterman NL (2010) A practical guide to methods of parentage analysis. *Mol Ecol Resour* 10(1):6–30
10. Hancock JM, Vieira S, Taraveira L, Santos A, Schmitt V, Semedo A et al (2019) Genetic characterization of green turtles (*Chelonia mydas*) from São Tomé and Príncipe: insights on species recruitment and dispersal in the Gulf of Guinea. *J Exp Mar Biol Ecol* 1 sept 518:151181
11. Nishizawa H, Abe O, Okuyama J, Kobayashi M, Arai N (2011) Population genetic structure and implications for natal philopatry of nesting green turtles *Chelonia mydas* in the Yaeyama Islands, Japan. *Endanger Species Res* 7 juill 14(2):141–148
12. Roberts MA, Schwartz TS, Karl SA (2004) Global population genetic structure and male-mediated gene flow in the green sea turtle (*Chelonia mydas*): analysis of microsatellite loci. *Genet* avr 166(4):1857–1870
13. Roden SE, Morin PA, Frey A, Balazs GH, Zarate P, Cheng IJ et al (2013) Green turtle population structure in the Pacific: new insights from single nucleotide polymorphisms and microsatellites. *Endanger Species Res* 20(3):227–234
14. Bradshaw PJ, Broderick AC, Carreras C, Fuller W, Snape RTE, Wright LI et al (2018) Defining conservation units with enhanced molecular tools to reveal fine scale structuring among Mediterranean green turtle rookeries. *Biol Conserv* 1 juin 222:253–260
15. FitzSimmons NN, Limpus CJ, Norman JA, Goldizen AR, Miller JD, Moritz C (1997) Philopatry of male marine turtles inferred from mitochondrial DNA markers. *Proc Natl Acad Sci* 5 août 94(16):8912–8917
16. Wang J (2016) Individual identification from genetic marker data: developments and accuracy comparisons of methods. *Mol Ecol Resour* 16(1):163–175

17. Harrison HB, Saenz-Agudelo P, Planes S, Jones GP, Berumen ML (2013) Relative accuracy of three common methods of parentage analysis in natural populations. *Mol Ecol* 22(4):1158–1170
18. Shamblin BM, Faircloth BC, Dodd MG, Bagley DA, Ehrhart LM, Dutton PH et al (2009) Tetranucleotide markers from the loggerhead sea turtle (*Caretta caretta*) and their cross-amplification in other marine turtle species. *Conserv Genet* 10(3):577–580
19. Monzón-Argüello C, Muñoz J, Marco A, López-Jurado LF, Rico C (2008) Twelve new polymorphic microsatellite markers from the loggerhead sea turtle (*Caretta caretta*) and cross-species amplification on other marine turtle species. *Conserv Genet* 9(4):1045–1049
20. Lin G, Chang A, Yap HW, Yue GH (2008) Characterization and cross-species amplification of microsatellites from the endangered hawksbill turtle (*Eretmochelys imbricata*). *Conserv Genet* 9(4):1071–1073
21. FitzSimmons NN, Moritz C, Moore SS (1995) Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Mol Biol Evol* 12:432–440
22. Dutton PH, Frey A (2009) Characterization of polymorphic microsatellite markers for the green turtle (*Chelonia mydas*). *Mol Ecol Resour* 9(1):354–356
23. Shamblin BM, Berry BE, Lennon DM, Bagley DA, Ehrhart LM, Nairn CJ (2012) Tetranucleotide microsatellite loci from the endangered green turtle (*Chelonia mydas*). *Conserv Genet Resour* 4(3):783–785
24. Bagda E, Bardakci F, Turkozan O (2012) Lower genetic structuring in mitochondrial DNA than nuclear DNA among the nesting colonies of green turtles (*Chelonia mydas*) in the Mediterranean. *Biochem Syst Ecol* 43:192–199
25. Alfaro-Núñez A, Jensen MP, Abreu-Grobois FA (2015) ;3:e880
26. Wright LI, Stokes KL, Fuller WJ, Godley BJ, McGowan A, Snape R et al (2012) Turtle mating patterns buffer against disruptive effects of climate change. *Proc R Soc B Biol Sci* 79(1736):2122–2127
27. Lara I, Oyama K, Cano-Camacho H, Zavala Paramo M, Vázquez-Marrufo G, Chassin-Noria O (2010) Detecting patterns of fertilization and frequency of multiple paternity in *Chelonia mydas* of Colola (Michoacán, México). *Hidrobiol Rev Dep Hidrobiol* 1 avr 20:85–89
28. Shamblin BM, Berry BE, Lennon DM, Meylan AB, Meylan PA, Outerbridge ME et al (2013) Tetranucleotide microsatellite loci from the critically endangered hawksbill turtle (*Eretmochelys imbricata*). *Conserv Genet Resour* 5(1):23–26
29. Miro-Herrans AT, Velez-Zuazo X, Acevedo JP, Mcmillan WO (2008) Isolation and characterization of novel microsatellites from the critically endangered hawksbill sea turtle (*Eretmochelys imbricata*). *Mol Ecol Resour* 8(5):1098–1101
30. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4(3):535–538
31. Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6(1):288–295
32. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX4. 05, logiciel sous Windows TM pour la génétique des populations. *Lab Génome Popul Interact CNRS UMR* 1 janv 5000:1996–2004
33. Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* 10(3):551–555
34. Bernatchez L, Duchesne P (2000) Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Can J Fish Aquat Sci - CAN J Fish AQUAT SCI* 1 janv 57:1–12
35. Loisier A, Savelli MP, Arnal V, Claro F, Gambaiani D, Sénégas JB et al (2021) Genetic composition, origin and conservation of loggerhead sea turtles (*Caretta caretta*) frequenting the french Mediterranean coasts. *Mar Biol* 26 mars 168(4):52
36. Nei M (1987) *Molecular evolutionary genetics*. Columbia university press
37. Sharma R, Kumar B, Arora R, Ahlawat S, Mishra AK, Tania MS (2016) Genetic diversity estimates point to immediate efforts for conserving the endangered tibetan sheep of India. *Meta Gene* 19 janv 8:14–20
38. Peyran C, Planes S, Tolou N, Iwankow G, Boissin E (2020) Development of 26 highly polymorphic microsatellite markers for the highly endangered fan mussel *Pinna nobilis* and cross-species amplification. *Mol Biol Rep* 1 avr 47(4):2551–2559
39. Seminoff JA, Balazs GH, Dutton PH, Eguchi T, Haas HL, Hargrove SA et al (2015) p Status review of the green turtle (*Chelonia mydas*) under the U.S. Endangered Species Act. NOAA;571. Report No.: NOAA-NMFS-SWFSC-539.
40. Tournon M, Genet Q, Gaspar C (2018) Final report on the green sea turtle egg-laying season of 2017–2018 (*Chelonia mydas*) on the atoll of Tetiaroa. *Assoc Te Mana O Te Moana* ; 54
41. Laloë JO, Monsinjon J, Gaspar C, Tournon M, Genet Q, Stubbs J et al (2020) Production of male hatchlings at a remote South Pacific green sea turtle rookery: conservation implications in a female-dominated world. *Mar Biol* 1 mai 167(5):70
42. Shields WM (1982) Philopatry, inbreeding, and the evolution of sex. State University of New York Press, p 260
43. Chen B, Cole JW, Grond-Ginsbach C (2017) Departure from Hardy Weinberg Equilibrium and Genotyping Error. *Front Genet* [Internet]. [cité 22 sept 2022];8. Disponible sur: <https://www.frontiersin.org/articles/https://doi.org/10.3389/fgene.2017.00167>
44. Stoeckel S, Grange J, Fernández-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S (2006) Heterozygote excess in a self-incompatible and partially clonal forest tree species -- *Prunus avium* L. *Mol Ecol* 15(8):2109–2118
45. Schaeffer SW, Walthour CS, Toleno DM, Olek AT, Miller EL (2001) Protein variation in *Adh* and *Adh*-related in *Drosophila pseudoobscura*. Linkage disequilibrium between single nucleotide polymorphisms and protein alleles. *Genet* 159(2):673–687
46. Tardy C, Planes S, Jung JL, Ody D, Boissin E (2020) Characterization of 25 new microsatellite markers for the fin whale (*Balaenoptera physalus*) and cross-species amplification in other cetaceans. *Mol Biol Rep* 1 sept 47(9):6983–6996
47. Broderick AC, Glen F, Godley BJ, Hays GC (2003) Variation in reproductive output of marine turtles. *J Exp Mar Biol Ecol* 25 mars 288(1):95–109
48. Miles LG, Lance SL, Isberg SR, Moran C, Glenn TC (2009) Cross-species amplification of microsatellites in crocodylians: assessment and applications for the future. *Conserv Genet* 10(4):935–954
49. Naro-Maciel E, Le M, FitzSimmons NN, Amato G (2008) Evolutionary relationships of marine turtles: a molecular phylogeny based on nuclear and mitochondrial genes. *Mol Phylogenet Evol* 49(2):659–662
50. Jensen MP, Allen CD, Eguchi T, Bell IP, LaCasella EL, Hilton WA et al (2018) Environmental warming and feminization of one of the Largest Sea Turtle populations in the World. *Curr Biol* 8 janv 28(1):154–159e4

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.