UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

Larissa Souza Arantes

Filogeografia genômica e hibridização em tartarugas marinhas

BELO HORIZONTE

Agosto, 2019

Larissa Souza Arantes

Filogeografia genômica e hibridização em tartarugas marinhas

Tese apresentada ao programa de Pós-Graduação em Genética da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do Título de Doutor em Genética.

Orientador: Prof. Dr. Fabrício Rodrigues dos Santos

Co-orientadora: Dra. Camila Junqueira Mazzoni

Belo Horizonte

Agosto, 2019

043 Arantes, Larissa Souza.
Filogeografia genômica e hibridização em tartarugas marinhas [manuscrito] / Larissa Souza Arantes. – 2019.
122 f. : il. ; 29,5 cm.

Orientador: Fabrício Rodrigues dos Santos. Co-orientadora: Camila Junqueira Mazzoni.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética populacional. 2. Tartarugas. 3. Filogeografia. 4. Hibridização Genética. I. Santos, Fabrício Rodrigues dos. II. Mazzoni, Camila Junqueira. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 575

Ficha catalográfica elaborada pela Biblioteca do Instituto de Ciências Biológicas da UFMG

Pós-Graduação em Genética
Departamento de Biologia Geral, ICB
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627 - C.P. 486 - Pampulha - 31270-901 - Belo Horizonte - MG
e-mail: pg-gen@icb.ufmg.br FAX: (+31) - 3409-2570

"Filogeografia genômica e hibridização em tartarugas marinhas"

Larissa Souza Arantes

Tese aprovada pela banca examinadora constituída pelos Professores:

Fabrício Rodrigues dos Santos - Orientador
UFMG

Camila Junqueira Mazzoni - Coorientadora
Leibniz Institute for Zoo and Wildlife Research

Gisele Lobo-Hajdu
Universidade Estadual do Rio de Janeiro

Ligiane Martins Moras
Centro Universitário Izabela Hendrix

Fernando Perini
UFMG

Maíra Carneiro Proietti
Universidade Federal do Rio Grande

Belo Horizonte, 23 de agosto de 2019.

Aos meus pais, Oraldo e Dulce, por me darem asas para que eu alcançasse vôos altos.

AGRADECIMENTO

Minha eterna gratidão...

Ao professor Fabrício Rodrigues dos Santos pela orientação, pelos ensinamentos compartilhados, por ser um exemplo de ética na ciência, por abrir tantas portas pelas quais me aventurei e por me inserir no mundo das tartarugas marinhas.

À querida co-orientadora Camila Mazzoni, por me acolher em seu laboratório permitindo o meu crescimento profissional com desafios, incentivos e exemplos. Por me considerar, além de aluna, uma parceira de trabalho e amiga.

À equipe do Projeto TAMAR e do Parque Nacional Marinho de Abrolhos pela parceria nos trabalhos, conhecimentos compartilhados e amostras cedidas.

À todos os colegas do LBEM, pelas discussões de resultados enriquecedoras e por tornarem os meus dias de trabalho mais prazerosos com uma boa conversa acompanhada de um bom café. Em especial, agradeço às queridas Elisa e a Pâmela, por me inspirarem como mulheres e pesquisadoras e por serem amigas tão especiais.

Aos colegas do BeGenDiv, por me acolherem em seu instituto, por me introduzirem ao mundo da bioinformática e por me apresentarem o melhor da Alemanha.

Aos meus pais Oraldo e Dulce, por sempre me apoiarem e possibilitarem que eu buscasse alcançar meus sonhos. Por serem meus exemplos de honestidade, família, casal, pais. Pelo amor incondicional, conselhos e valores.

Aos meus irmãos, Aline e Vinícius, pelo carinho e amizade.

À minha família, avós, tios, primos e cunhados, por serem meu porto seguro.

Ao Ivan, por estar sempre ao meu lado me incentivando e vibrando comigo a cada descoberta e aprendizado. Por ser o melhor companheiro, mininizar todas as dificuldades do meu caminho e por preencher os meus dias com amor. Obrigada por acreditar mais em mim do que eu mesma!

Às minhas queridas Eneida e Carol, por me acolherem em BH e por tornarem meu dia-a-dia sempre agradável e aos amigos Henrique, Lulu, Lili, Cacá, Bruna, Cleide e Rafa por sempre estarem presentes, mesmo distantes. Pelos fortes laços de amizade incapazes de se romper.

Por fim agradeço à CAPES, FAPEMIG e CNPq pela concessão das bolsas e financiamentos que tornaram possível o desenvolvimento desta tese.

"Meu êxito como cientista (...) foi determinado (...) por qualidades e condições complexas e diversificadas. Destas as mais importantes foram o amor à ciência; uma paciência ilimitada para refletir longamente sobre qualquer assunto; zelo para observar e colecionar dados; uma boa porção de imaginação e senso comum."

Darwin, 1911

ÍNDICE

LISTA DE FIGURAS

INTRODUÇÃO

CAPÍTULO I

CAPÍTULO II

CAPÍTULO III

LISTA DE TABELAS

CAPÍTULO I

CAPÍTULO II

LISTA DE ABREVIATURAS

BSA - Bovine Serum Albumin

CCL - Curved Carapace Length

CCT - Taxonomic Collections Center

CCW - Curved Carapace Width

CS - Clutch Size

ddRADseq - Double digest RAD sequencing

Ei x *Cc* - Híbridos entre as espécies *Eretmochelys imbricata* e *Caretta caretta*

ERG - Elevação do Rio Grande

ESS - Effective Sample Sizes

ESU - Evolutionary Significant Units

F1 - Híbridos de primeira geração

GBS - Genotyping-by-Sequencing

HI - Hybrid Index

HPD - Highest Posterior Density

HxL - Hybrids between Hawksbill and Loggerhead turtles

HS - Hatchling Success

IP - Incubation Period

IUCN - International Union for Conservation of Nature and Natural Resources

MCC - Maximum Clade Credibility

MCMC - Markov Chain Monte Carlo

MSA - Mixed Stock Analyses

MU - Management Units

mtDNA – DNA mitocondrial

mya - million years ago

nDNA - nuclear DNA

NGS - Next Generation Sequencing

PAN - National Action Plan

PC - Principal Component

PCA - Principal Component Analysis

PCR - Polymerase Chain Reaction

RADseq - Restriction Associated DNA sequencing

RM-MJ - Reduced Median with Median Joining algorithm

RMU - Regional Management Units

SAMOVA - Spatial Analysis of MOlecular VAriance

TMRCA - Time to the Most Recent Common Ancestor

RESUMO

As tartarugas marinhas possuem ciclo de vida complexo, com fases de vida proximamente relacionadas aos seres humanos. Esta interação as expõe às ameaças causadas pelo impacto antrópico, o que levou a um declínio populacional global para todas as espécies. Além disso, a incidência extremamente alta de hibridização entre as tartarugas marinhas na costa brasileira é uma situação atípica que deve ser investigada. Assim, a compreensão da história evolutiva e da dinâmica das populações e a investigação do impacto do processo de hibridização entre as tartarugas marinhas se tornam essenciais para a conservação das espécies. Neste trabalho analisamos inicialmente os padrões filogeográficos globais da tartaruga-de-pente (*Eretmochelys imbricata*) através da compilação dos dados de haplótipos mitocondriais disponíveis na literatura. Importantes padrões demográficos globais estavam sendo ignorados devido ao uso de dados pouco representativos e de uma nomenclatura haplotípica não padronizada. Nossos resultados permitiram identificar conectividade entre algumas populações, reconhecer lacunas no conhecimento da espécie que devem ser priorizadas em pesquisas futuras e reforçar a necessidade de esforços internacionais focados na conservação da espécie. Em seguida, realizamos um estudo da filogeografia, demografia e hibridização das tartarugas marinhas na costa brasileira através de uma abordagem multilocus. Novos marcadores genéticos foram selecionados através da técnica de ddRADseq e utilizados para detecção de variação inter e intraespecífica para as cinco espécies de tartarugas marinhas que ocorrem na costa brasileira e para indivíduos híbridos. Nossos resultados sugerem que existe um fluxo gênico mediado pelos machos entre as colônias de desova da costa brasileira para as tartarugas de pente e cabeçuda (*Caretta caretta*). Também estimamos que o fenômeno de hibridização em alta frequência que ocorre na Bahia provavelmente é ainda mais recente do que previamente sugerido, pois apenas indivíduos F1 adultos foram encontrados. Por fim, realizamos a caracterização da população de desova de tartarugas-cabeçudas do Arquipélago de Abrolhos, através de dados genéticos e do monitoramento reprodutivo. Nossos resultados mostraram que Abrolhos é também uma área de ocorrência de desovas de fêmeas híbridas e que parecem apresentar um menor sucesso reprodutivo que fêmeas "puras". Em conclusão, esta tese de doutorado permitiu avançar no conhecimento acerca dos processos filogeográficos e filogenéticos das tartarugas marinhas e demonstrou como as análises genéticas em combinação com dados ecológicos podem contribuir para a compreensão de fenômenos importantes para a conservação das espécies.

ABSTRACT

Sea turtles have a complex life cycle, with life stages closely related to humans. This interaction exposes them to threats caused by anthropic impact, which lead to a global population decline for all species. In addition, an extremely high incidence of hybridization among sea turtles along the Brazilian coast is an atypical situation to be investigated. Thus, understanding the evolutionary history and population dynamics and investigating the impact of the hybridization process among sea turtles is essential for the species conservation. In this work we initially analyze the global phylogeographic patterns of hawksbill turtle (*Eretmochelys imbricata*) through the compilation of mitochondrial haplotype data available in the literature. Important global demographic patterns were being missed due to the use of non-representative data or non-standard haplotype nomenclature. Our results allowed us to identify the connectivity among populations, to recognize knowledge gaps that should be prioritized in future research, and to reinforce the need for international efforts focused on species conservation. Next, we performed a study of phylogeography, demography and hybridization of sea turtles along the Brazilian coast through a multilocus approach. New genetic markers were selected using the ddRADseq technique and used to detect inter and intraspecific variation for the five sea turtle species that occur along the Brazilian coast and for hybrid individuals. Our results suggest that there is a male-mediated gene flow between the rookeries along the Brazilian coast for hawksbill and loggerhead (*Caretta caretta*) turtles. We also estimate that the phenomenon of high frequency hybridization occurring in Bahia is probably even more recent than previously suggested, as only F1 hybrid adults were found. Finally, we investigate the Abrolhos Archipelago nesting population of loggerhead turtles, through genetic data and reproductive monitoring. Our results showed that Abrolhos is also a nesting area for hybrids, and female hybrids appear to present a lower reproductive success than “pure” females. In conclusion, this thesis allowed us to advance the knowledge about the phylogeographic and phylogenetic processes of sea turtles and demonstrated how the genetic analysis associated with ecological data can contribute to the understanding of important phenomena for the species conservation.

INTRODUÇÃO GERAL

As tartarugas marinhas no Brasil

As tartarugas marinhas compreendem sete espécies globalmente distribuídas, das quais cinco ocorrem no Brasil. Estas espécies são classificadas em duas famílias taxonômicas: Dermochelyidae, representada pela tartaruga-de-couro (*Dermochelys coriacea* (Vandelli, 1761)), e Cheloniidae, representada no Brasil pela tartaruga-verde (*Chelonia mydas* (Linnaeus, 1758)), tartaruga-cabeçuda (*Caretta caretta* (Linnaeus, 1758)), tartaruga-oliva (*Lepidochelys olivacea* (Eschscholtz, 1829)), e tartaruga-de-pente (*Eretmochelys imbricata* (Linnaeus, 1766)). Todas elas estão ameaçadas de extinção segundo a Lista Vermelha de Espécies Ameaçadas da IUCN, em diferentes níveis de ameaça (2019). As principais ameaças à conservação das tartarugas marinhas estão relacionadas à captura incidental na pesca, ao desenvolvimento costeiro, à poluição dos oceanos (esgoto, lixo, plástico, substâncias tóxicas), à patógenos e à comercialização de ovos, carne ou outros produtos de tartaruga (Wallace et al., 2011).

O ciclo de vida das tartarugas marinhas é longo e complexo (Figura 1). Após emergirem do ninho, os filhotes iniciam um longo período pelágico no qual realizam longas migrações e dificilmente são observados até seu retorno para águas costeiras como juvenis maiores (Putman et al., 2015). Esse período de vida é conhecido como “anos perdidos”, devido à falta de conhecimento sobre os padrões de dispersão e comportamento nesta fase de vida das tartarugas. Recentes evidências demonstraram que estes filhotes e juvenis também realizam dispersão ativa ao invés de distribuição apenas passiva influenciada basicamente pelas correntes marinhas, como era inicialmente hipotetizado (Putman et al., 2015; Briscoe et al., 2016).

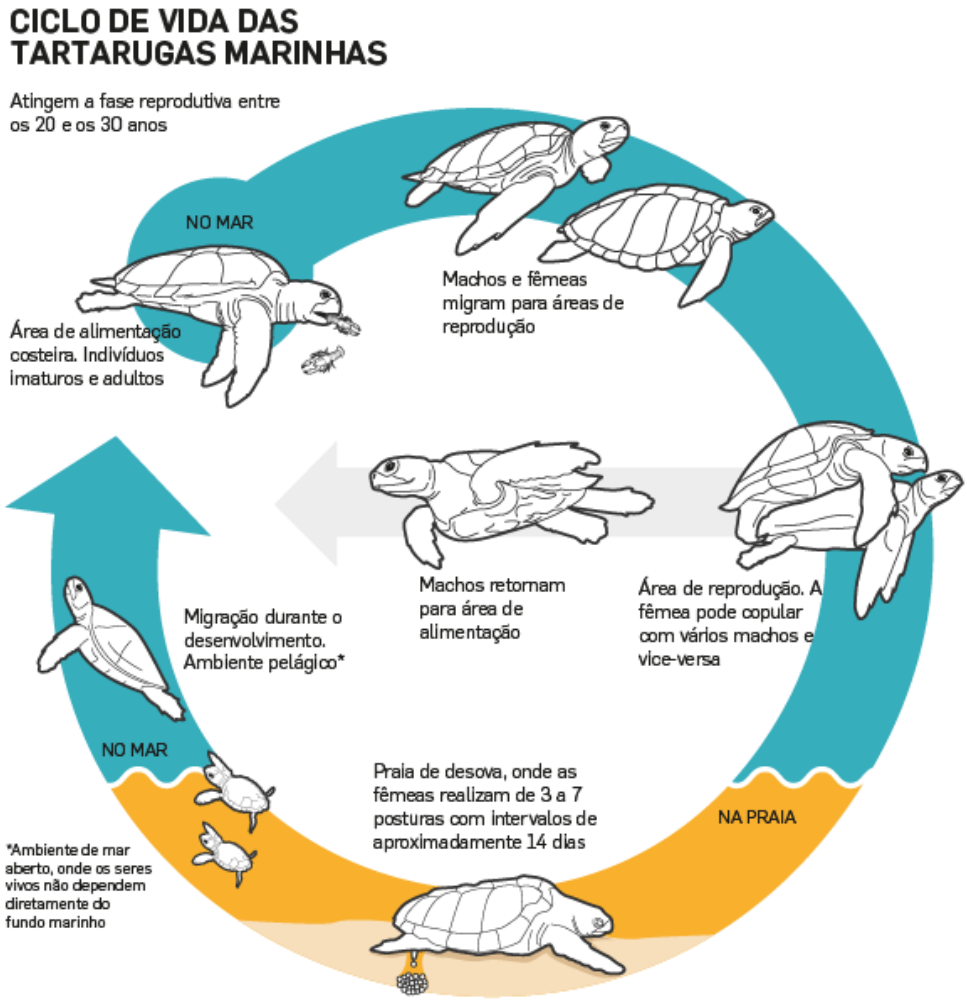

Figura 1 – Ciclo de vida das tartarugas marinhas.

Após a fase oceânica, os juvenis maiores (20 a 40 cm, dependendo da espécie e população) frequentam áreas costeiras e oceânicas de alimentação compostas por tartarugas advindas de diversos locais de desova do mundo (denominadas estoques mistos). Portanto, a diversidade genética encontrada nestes aglomerados é alta (Shamblin et al., 2014; Vilaça et al., 2013). Quando atingem a idade reprodutiva, as tartarugas marinhas retornam para a região onde nasceram para reproduzir e desovar, comportamento conhecido como filopatria natal (Carr e Ogren, 1960). Desta forma, as populações de tartarugas marinhas geralmente são geneticamente estruturadas entre as colônias de desova. Entretanto, o grau de fidelidade natal dos machos é ainda questionado, principalmente devido à menor quantidade de estudos que analisam machos em razão da dificuldade de coleta de amostras quando comparado às fêmeas. Para a tartaruga-verde foi demonstrado que os machos reprodutivos nas áreas de acasalamento apresentam os mesmos haplótipos mitocondriais que as fêmeas reprodutivas, o que demonstra que os machos

também exibem o comportamento filopátrico (FitzSimmons et al., 1997). Porém, machos e fêmeas da tartaruga-de-pente em Porto Rico mostraram diferença haplotípica significativa, revelando que o comportamento filopátrico dos machos é mais flexível (Velez-Zuazo et al., 2008).

A múltipla paternidade, quando mais de dois pais são identificados para os filhotes de um único ninho, já foi amplamente demonstrada para as tartarugas marinhas. Este comportamento, assim como os acasalamentos oportunistas em corredores migratórios e eventos de realocação, conferem vantagens aos indivíduos, uma vez que representam uma oportunidade de acasalar com mais parceiros e consequentemente aumentar a diversidade genética e o potencial adaptativo, além de evitar a endogamia (Encalada et al., 1998; Moore et al., 2002).

Filogeografia das tartarugas marinhas

Estudos filogeográficos utilizam dados genéticos e de distribuição das linhagens para investigar processos responsáveis pela origem, distribuição e manutenção da biodiversidade (Beheregaray, 2008). Essa abordagem tem sido especialmente importante para coletar informações atuais e históricas das populações de tartarugas marinhas, uma vez que os estudos com estas espécies são dificultados por seu comportamento altamente migratório, uso de diferentes habitats, ampla distribuição e longo ciclo de vida. Estes estudos permitem explorar a estruturação populacional e a história demográfica, avaliar a diversidade genética, investigar a origem local dos indivíduos das agregações de forrageamento, estimar o fluxo gênico e compreender as migrações oceânicas das tartarugas marinhas (Reis et al., 2010a, Vilaça et al., 2013, Proietti et al., 2014b, Shamblin et al., 2014, Vargas et al., 2016).

A atual distribuição da diversidade genética encontrada nas colônias e nos agregados de alimentação deve ser interpretada em um contexto espacial e temporal. Eventos históricos como a formação de barreiras físicas (istmos e pontes), os ciclos glaciais do Pleistoceno, as mudanças nas correntes marinhas e no nível do mar, e as diferenças na temperatura das águas oceânicas influenciaram os processos evolutivos das tartarugas (Vargas et al., 2016).

Para a maioria das linhagens das tartarugas marinhas, uma resposta demográfica simultânea, mas com diferentes magnitudes, foi vista após o Último Máximo Glacial. A maioria

das linhagens passou por um rápido processo de expansão populacional, demonstrando uma rápida resposta a mudanças no ambiente marinho, enquanto algumas permaneceram aparentemente estáveis (Reid et al., 2019). A estabilidade demográfica e a diversidade genética foi maior entre linhagens do Indo-Pacífico e para linhagens com ampla distribuição.

O padrão filogeográfico das tartarugas marinhas está relacionado à preferência por habitats e a tolerância termal de cada espécie. Espécies tropicais como as tartarugas verde, de pente e oliva exibem uma separação filogenética antiga e profunda entre as bacias oceânicas do Atlântico e do Indo-Pacífico, sugerindo que América e África são barreiras importantes para a dispersão das espécies devido a temperatura menor nas latitudes maiores (Bowen e Karl, 2007; Monzon-Arguello et al. 2011; Duchene et al., 2012). Já a tartaruga-de-couro, que é mais tolerante a temperaturas frias, apresenta evidências de colonização mais recentes do Atlântico por linhagens do Indo-Pacífico (Reid et al., 2019). A tartaruga-cabeçuda também apresenta uma divergência entre linhagens que está relacionada às bacias oceânicas, porém eventos mais recentes de colonização homogeneizaram suas populações. Assim, a divergência média entre as populações de tartaruga-cabeçuda do Atlântico é maior que a divergência entre populações que habitam o Atlântico e o Pacífico (Duchene et al. 2012).

Desta forma fica claro que as barreiras ao fluxo gênico não são as mesmas para todas as espécies de tartarugas marinhas, devido às diferentes capacidades de dispersão através dos oceanos e respostas evolutivas frente às mudanças ambientais (Duchene et al., 2012). A complexidade da dinâmica populacional das tartarugas marinhas ressalta a necessidade de mais estudos para compreensão dos padrões demográficos das espécies.

Hibridização entre tartarugas marinhas

Hibridização interespecífica pode ocorrer naturalmente entre muitas espécies de plantas e animais, sendo considerada um importante processo de diversificação na história evolutiva de alguns grupos taxonômicos. Porém, nas últimas décadas foi demonstrado um aumento nas taxas de hibridização interespecífica devido às modificações antrópicas no ambiente, como mudanças na abundância e distribuição das espécies e remoção de barreiras que isolavam ou restringiam as espécies (van Wyk et al., 2017).

As consequências do processo de hibridização variam entre diferentes táxons. Em alguns casos, a hibridização pode representar uma importante fonte de variação adaptativa por conferir um efeito vantajoso no *fitness* dos indivíduos, fenômeno chamado introgressão adaptativa (Hedrick, 2013). No entanto, a hibridização interespecífica também pode resultar em efeitos deletérios para a prole (F1, F2, etc.) devido às incompatibilidades resultantes da combinação genética das espécies parentais (Maheshwari e Barbash, 2011). Esta redução do valor adaptativo populacional é conhecida como depressão exogâmica, o qual pode conduzir ao declínio e à extinção local de populações (Allendorf et al., 2001; Todesco et al., 2016).

Casos de hibridização interespecífica entre tartarugas marinhas têm sido relatados entre cinco espécies da família Cheloniidae, as quais pertencem a grupos que se divergiram há aproximadamente 63 milhões de anos (Naro-Maciel et al., 2008). Baseados em dados morfológicos e genéticos, os seguintes tipos de cruzamentos já foram descritos para as tartarugas: *E. imbricata* × *C. caretta* (Conceição et al., 1990), *C. caretta* × *L. olivacea* (Reis et al., 2010b), *E. imbricata* × *L. olivacea* (Lara-Ruiz et al., 2006), *C. mydas* × *E. imbricata* (Karl et al., 1995), *C. mydas* × *C. caretta* (James et al., 2004), *C. mydas* × *L. olivacea* (Seminoff et al., 2003), *C. caretta* × *Lepidochelys kempii* (Barber et al., 2003) e híbridos provavelmente originados pelo cruzamento sequencial entre três espécies *C. mydas* × *E. imbricata* × *C. caretta* (Vilaça et al., 2012).

Estes casos de hibridização têm se restringido a poucos indivíduos ao redor do mundo (Soares et al., 2017). Entretanto, no Brasil o processo de hibridização interespecífica merece destaque visto que a frequência de híbridos é muito maior que em qualquer outra população analisada (Vilaça et al., 2012). Na costa norte da Bahia, local de sobreposição das maiores colônias das tartarugas de pente e cabeçuda do país, foi constatado uma frequência de hibridização de 42% para a população morfologicamente identificada como tartaruga-de-pente (Lara-Ruiz et al., 2006). No Sergipe, principal área de desova da tartaruga-oliva e uma importante colônia de cabeçudas, 27% dos indivíduos morfologicamente identificados como tartarugas cabeçudas eram na verdade híbridos entre olivas e cabeçudas (Reis et al., 2010b).

Os prováveis fatores que contribuem para este evento único no Brasil incluem: (i) a sobreposição espaço-temporal das desovas de diferentes espécies, que dão oportunidade para que o cruzamento interespecífico ocorra. Isto é corroborado pela observação de que o período de maior ocorrência de híbridos é intermediário ao pico reprodutivo das tartarugas de pente e cabeçuda na Bahia (Soares et al., 2017), (ii) a recente diminuição do tamanho populacional das

tartarugas marinhas, o que torna mais raros os encontros de indivíduos de cada sexo da mesma espécie (Vilaça et al., 2012), (iii) os tamanhos populacionais diferentes para as espécies envolvidas na hibridização, visto que a tartaruga-cabeçuda é mais abundante no país e a grande maioria dos casos de hibridização na costa brasileira envolvem fêmeas de tartaruga-cabeçuda (Vilaça et al., 2012; Proietti et al., 2014a), e (iv) a taxa sexual altamente desproporcional, uma vez que estima-se que aproximadamente 90% dos filhotes nascidos na Bahia sejam fêmeas (Marcovaldi et al., 1997, 2014, 2016).

Outras populações de tartarugas marinhas ao redor do mundo estão sujeitas aos mesmos fatores que as populações híbridas brasileiras, porém não apresentam o fenômeno de hibridização nesta mesma intensidade. No entanto, a identificação de híbridos através das características morfológicas não é tão óbvia e depende da experiência pessoal. Além disso, as análises genéticas são fundamentais para confirmação do processo de hibridização e para identificação das espécies e do número de gerações envolvidos. Assim, estudos genéticos devem ser expandidos para outras populações para investigar se o processo de hibridização em alta frequência é realmente restrito às populações do norte da Bahia e do Sergipe.

Existe um viés de gênero no processo de hibridização na costa brasileira, envolvendo na maioria dos casos fêmeas de tartaruga-cabeçuda e machos de tartaruga-de-pente. Além da maior abundância de tartarugas-cabeçudas no Brasil, a distribuição temporal das desovas também favorece este viés. O pico reprodutivo no Brasil ocorre entre 15 de outubro e 15 de dezembro para a tartaruga-cabeçuda, e entre 15 de dezembro e 15 de fevereiro para a tartaruga-de-pente (Marcovaldi et al. 1999). Os machos de tartaruga-de-pente que chegam as áreas costeiras de reprodução encontram um grande número de tartarugas-cabeçudas fêmeas para acasalar. Já o encontro entre machos de tartaruga-cabeçuda e fêmeas de tartaruga-de-pente acontece menos frequentemente, uma vez que os machos de tartaruga-cabeçuda deixam as áreas de acasalamento antes que um grande número de fêmeas de tartaruga-de-pente cheguem (Vilaça et al., 2012, Proietti et al., 2014).

A maioria dos híbridos analisados exibem um alelo de cada espécie parental, sendo assim híbridos de primeira geração (F1) (Vilaça et al., 2012). Porém, foram registrados em menor escala híbridos >F1, revelando que está ocorrendo introgressão com as tartarugas de pente e cabeçuda (Vilaça et al., 2012). A partir desses dados, estimou-se que o fenômeno de hibridização na costa brasileira ocorre há pelo menos duas gerações de tartarugas, entre 40 e 50

anos, o que reforça a hipótese associativa entre o grande declínio populacional das tartarugas marinhas e o processo de hibridização no Brasil (Vilaça et al., 2012).

Recentemente foi realizada uma análise comparativa de diversos parâmetros reprodutivos de fêmeas híbridas F1 da Bahia em relação às suas espécies parentais (Soares et al., 2017). Este estudo demonstrou que o sucesso de eclosão de filhotes dos ninhos de híbridos é menor, porém outros parâmetros como o tamanho e frequência das ninhadas não exibiram diferenças significativas em relação às espécies parentais, o que sugere que os híbridos podem persistir no local. Além disso, a viabilidade dos filhotes híbridos e não-híbridos foi indistinguível (Soares et al., 2018). Estes estudos demonstram a importância de integrar dados ecológicos e genéticos para compreender as consequências do processo de hibridização na costa brasileira. Estudos adicionais são necessários para confirmar o efeito na hibridização no *fitness* dos indivíduos híbridos de diferentes gerações e de diferentes cruzamentos.

O comportamento dos híbridos após o nascimento, suas rotas de migração, o tipo de alimentação e as preferências de acasalamento permanecem desconhecidos (Proietti et al., 2014a). Teoricamente, os híbridos podem apresentar comportamentos típicos de uma das espécies parentais ou intermediário. Um estudo com telemetria acompanhou fêmeas híbridas *Ei* × *Cc* após desovarem na Bahia e mostrou que a maioria dos indivíduos migrou para áreas de alimentação típicas da tartaruga-cabeçuda, com exceção de uma das híbridas que migrou para as áreas de recifes de corais no sul da Bahia típicas da tartaruga-de-pente (Marcovaldi et al., 2012). Híbridos imaturos *Ei* × *Cc* da Bahia foram encontrados em áreas de alimentação de águas temperadas do oeste do Atlântico Sul e no Ceará, onde ocorrem muitas cabeçudas (Proeitti et al., 2014a). Juvenis híbridos *Ei* × *Cc* também já foram reportados na costa da Argentina, outra área de alimentação típica de cabeçudas (Prodoscimi et al., 2014). Mais estudos envolvendo monitoramento do uso de áreas de alimentação dos híbridos podem contribuir para responder questões acerca da sua ecologia e comportamento.

Avanços nas técnicas genéticas aplicadas a estudos filogeográficos e de hibridização

A aplicação de ferramentas genéticas trouxe importantes avanços nos estudos evolutivos e ecológicos das tartarugas marinhas, revelando padrões de estruturação populacional e

permitindo compreender a conectividade entre áreas de desova e alimentação, além de determinar os padrões filogeográficos das espécies. Há aproximadamente três décadas, os estudos com *loci* mitocondriais, majoritariamente com a região controle ou D-loop, forneceram as primeiras evidências corroborando a hipótese de filopatria natal das fêmeas de tartarugas marinhas e determinando a estruturação entre diferentes estoques genéticos populacionais das colônias de desova (Komoroske et al., 2017).

A definição genética dos limites entre populações, as quais são importantes para a definição de unidades de manejo para a conservação, está condicionada ao tipo e à quantidade de marcadores genéticos utilizados no estudo. Por exemplo, o uso do mitogenoma completo (aproximadamente 16500 pb) oferece um maior poder de inferência de estruturação populacional quando comparado apenas com a região D-loop do DNA mitocondrial (mtDNA – aproximadamente 620 bp) (Shamblin et al., 2012; Duchene et al., 2012).

Devido à limitada capacidade de resolução genética e às características inerentes ao mtDNA, como a herança unicamente materna, estudos com marcadores nucleares também começaram a ser executados, melhorando a detecção de limites populacionais e permitindo avaliar a contribuição dos machos para o fluxo gênico entre populações (Komoroske et al, 2017).

Recentes avanços das técnicas de sequenciamento de nova geração (Next Generation Sequencing – NGS) estão propiciando um aumento na resolução dos níveis de estruturação populacional, assim como na compreensão das estratégias comportamentais e dos processos adaptativos das tartarugas marinhas (Komoroske et al., 2017). Além disso, várias estratégias NGS vêm sendo desenvolvidas com o objetivo de explorar o genoma de forma resumida, permitindo identificar e utilizar um subconjunto de sequências selecionadas em análises populacionais. Elas tornaram possível investigar em escala genômica um número maior de indivíduos de espécies modelo e não modelo, com um melhor custo benefício (Ekblom et al., 2011).

Existem diferentes abordagens de representação genômica reduzida (i.e. Restriction-Site Associated DNA sequencing - RADseq, double digest RADseq - ddRADseq, Genotyping-by-Sequencing - GBS, RNAseq, Target Capture, etc.), as quais apresentam benefícios específicos de acordo com o objetivo da pesquisa. Estas técnicas permitem a descoberta de polimorfismos (*SNP calling*) mais acurada, a identificação de variantes raras e o processamento simultâneo de uma maior quantidade de amostras (McCormack et al., 2013). Através destas

técnicas também é possível desenvolver marcadores moleculares específicos para a espécie de interesse, o que elimina o viés de averiguação associado à maioria dos estudos populacionais e permite explorar a variação intra e interpopulacional.

Embora recentemente alguns estudos envolvendo as tartarugas marinhas têm gerado dados através de abordagens genômicas, como o sequenciamento do transcriptoma (Bertley et al., 2017), o RAD-Capture (Komoroske et al., 2018) e a técnica de ddRADseq (Hurtado et al., 2016), estudos genômicos em larga escala são escassos e limitados a certas espécies de tartarugas marinhas. De fato, somente a tartaruga-verde tem o seu genoma parcialmente sequenciado (Wang et al., 2013).

A expansão do métodos genômicos para diferentes espécies e populações é altamente recomendada para aprofundar o conhecimento sobre os padrões demográficos e evolutivos das tartarugas marinhas, bem como os eventos de hibridização que ocorrem na costa brasileira, permitindo estimar o nível de introgressão, o número de gerações em que o processo está ocorrendo e o impacto do processo de hibridização na diversidade genética e conservação das espécies parentais.

Apresentação da tese

Diante do exposto, esta tese de doutorado teve como objetivos principais a compreensão da história evolutiva e da dinâmica das populações de tartarugas marinhas da costa brasileira e a investigação do processo de hibridização que ocorre em alta frequência exclusivamente no Brasil.

As análises filogeográficas da tartaruga-de-pente revelaram alguns desafios devido à falta de uma nomenclatura padronizada para os haplótipos da região controle do mtDNA, que são utilizados na maioria dos estudos que investigam a história evolutiva da espécie. Além disso, até hoje nenhum trabalho foi realizado para a espécie englobando todas as populações da tartaruga-de-pente de diferentes bacias oceânicas. Baseado nisso, no primeiro capítulo desta tese foi realizada uma compilação de todos os dados populacionais utilizando haplótipos de mtDNA para todas as populações da tartaruga-de-pente e diversas análises filogenéticas e filogeográficas globais representando toda a diversidade e distribuição da espécie.

No segundo capítulo desta tese foi realizado o desenvolvimento e a padronização de marcadores genéticos para estudar a filogeografia e os eventos de hibridização das tartarugas marinhas. Através do uso de *loci* nucleares identificados a partir de dados de ddRADseq, foram selecionados marcadores genéticos para identificar haplótipos específicos (diagnósticos) das espécies envolvidas no fenômeno de hibridização da costa brasileira. Estes dados foram utilizados para verificar e mensurar a hibridização interespecífica da costa brasileira. Os marcadores com variação intraespecífica foram utilizados para investigação dos padrões filogeográficos das espécies. Foram estudados os eventos demográficos históricos, a estruturação populacional, o fluxo gênico e o uso de áreas de alimentação por indivíduos de diferentes populações.

No terceiro capítulo, nosso objetivo foi utilizar uma abordagem multidisciplinar para investigar a suspeita de hibridização no Arquipélago de Abrolhos. Nós integramos dados genéticos obtidos através de análises multilocus e do método 3RAD e dados do monitoramento reprodutivo da população do arquipélago obtidos em quatro temporadas reprodutivas. Analisamos fêmeas em desova e filhotes, confirmando a identificação morfológica através de dados genéticos e comparando o sucesso reprodutivo de indivíduos “puros” e híbridos.

A introdução, metodologia, resultados e discussão específicas de cada capítulo serão apresentados no formato de artigos científicos em preparação para publicação ou submetidos a revistas científicas indexadas.

CAPÍTULO I

Global phylogeography of the critically endangered hawksbill turtle (*Eretmochelys imbricata*)

Larissa Souza Arantes, Sarah Maria Vargas, Fabrício Rodrigues Santos. Manuscrito submetido ao periódico Genetics and Molecular Biology

Global phylogeography of the critically endangered hawksbill turtle (*Eretmochelys imbricata*)

Larissa Souza Arantes[1], Sarah Maria Vargas[2], Fabrício Rodrigues Santos[1]

[1]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte - MG, 31270-010, Brazil.
[2] Departamento de Ciências Biológicas, Instituto de Ciências Humanas e Naturais, Universidade Federal do Espírito Santo (UFES), Vitória - ES, 29043–900, Brazil.

Abstract

The hawksbill turtle (*Eretmochelys imbricata*) is a broadly distributed, highly migratory and critically endangered sea turtle species. The lack of studies exploring the species natural history restricts the comprehension of its biology, behaviour and life history. In this work, we performed a global phylogeographic analysis using a compilation of previously published mitochondrial haplotype data to understand the species dynamics and diversity worldwide. Our results showed that a complex demographic pattern involving ancient and recent events could explain the phylogeography of hawksbill turtles. Isolation by distance is not enough to explain distinct demographic units of hawksbill turtles, which may also be explained by local factors as oceanic currents, reef coral distribution and nesting timing. The foraging aggregations are typically mixed stocks of individuals originating from multiple nesting areas, but the trend of remaining in the nearby regions of their natal beaches was also registered in some grounds. Phylogenetic analysis indicates two highly divergent lineages split between Atlantic and Indo-Pacific Ocean rookeries, and a more recent transoceanic colonization of the Atlantic Ocean. Long-distance dispersal events are likely responsible for homogenization between distant populations. Our findings provide new insights about population connectivity, identify gaps that should be prioritized in future research and highlight the need for international efforts aiming at hawksbill's conservation.

Introduction

Species that exhibit worldwide distribution and highly migratory behavior are rarely studied in a global context, covering all population diversity and connectivity. As a result, only fragmented information is generated on the evolutionary history of species. Sea turtles are cosmopolitan animals that can disperse among different ocean basins during their lifetime (Bolten *et al*., 1998; Monzón-Argüello *et al*., 2011). Species dynamics change according to life stages, varying from strongly structured nesting populations or rookeries to highly diverse foraging aggregations (Bowen *et al.*, 2005).

Phylogeographic studies use genetic and distribution data from populations to investigate processes related to their origin, distribution and dynamics (Beheregaray, 2008). This approach has been especially important to collect current and historical information of sea turtle populations, allowing to explore population structure and demographic history, to evaluate the genetic diversity, determine origins of individuals from foraging aggregations, estimate gene flow and comprehend oceanic migrations (Vilaça *et al.*, 2013; Proietti *et al.*, 2014; Shamblin *et al.*, 2014; Vargas *et al.*, 2016).

The majority of phylogeographic studies have used mitochondrial DNA (mtDNA) markers to study the genetic diversity of sea turtles (López-Barrera *et al.*, 2016), which allow recognizing different populations and tracking the relationships between rookeries. The control region (or D-loop) of the mtDNA is the baseline data for these analyses, but only local populations have been analyzed, most commonly limited to one or few ocean basins, undermining the access to the global population patterns. Moreover, diverse research groups have been performing genetic studies with sea turtles using independent methods, resulting in incomparable data.

Efforts to compile control region haplotypes have been performed for some sea turtle species, such as loggerhead (*Caretta caretta*) and green turtles (*Chelonia mydas*), in a collaboration between research groups (Shamblin *et al.*, 2014) and The Archie Carr Center for Sea Turtle Research (http://accstr.ufl.edu/resources/mtdna-sequences). For *C. mydas*, a recent study investigated the global phylogeography using a compiled worldwide haplotype dataset (Jensen *et al.*, 2019), and commitments to expand these models for other species are desired (Gaos *et al.*, 2016). In addition, international collaboration and the standardization of haplotype's nomenclature is critical to investigate the sea turtle biology in a global context.

The hawksbill turtle (*Eretmochelys imbricata* (Linnaeus, 1766)) is a critically endangered species that occurs in tropical and subtropical waters around the world (Mortimer

and Donnelly, 2008). The main threats to this species' conservation are anthropogenic impacts, such as nesting area degradation, pollution, fisheries bycatch and exploitation through commercialization of the unique hawksbill shell, meat and eggs (da Silva *et al.,* 2015). Threats are intensified since hawksbill turtles preferred feeding areas are coral reefs, a greatly endangered marine ecosystem (Mortimer and Donnelly, 2008).

Phylogeographic studies of hawksbill turtles have been performed apart for populations in Brazil (Lara-Ruiz *et al.,* 2006; Vilaça *et al.,* 2013; Proietti *et al.,* 2014), Caribbean (Bass *et al.,* 1996; Troëng *et al.,* 2005; Bowen and Karl, 2007; Velez-Zuazo *et al.,* 2008; Blumental *et al.,* 2009; Richardson *et al.,* 2009; Browne *et al.,* 2010; LeRoux *et al.,* 2012; Carreras *et al.,* 2013; Wood *et al.,* 2013; Gorham *et al.,* 2014; Trujillo-Arias *et al.,* 2014; Cazabon-Mannette *et al.,* 2016; Hill *et al.,* 2018; Labastida-Estrada *et al.,* 2018), Eastern Atlantic (Monzón-Argüello *et al.,* 2010, 2011; Putman *et al.,* 2014), Indo-Pacific (Vargas *et al.,* 2016), Persian Gulf (Tabib *et al*., 2011, 2014; Natoli *et al*., 2017), Eastern Pacific (Gaos *et al*., 2012, 2016, 2017, 2018; Zuniga-Marroquin and Monteros, 2017), Southeast Asia (Nishizawa *et al*., 2016) and Japan (Nishizawa *et al*., 2010). The first genetic studies used mtDNA control region data based on approximately 300 bp sequences (Bass *et al.,* 1999; Dias-Fernández *et al.,* 1999; Troëng *et al.,* 2005; Bowen and Karl, 2007; Nishizawa *et al.,* 2010). Then, longer sequences began to be used, increasing the detection of genetic variation and improving the inference power of population structure and connectivity among distant populations (LeRoux *et al.,* 2012).

In a global perspective, there is a deep genetic divergence between hawksbill's lineages from Atlantic and Indo-Pacific basins. The divergence time between lineages dates from the Pliocene, when the closing of the Isthmus of Panama occurs (Duchene *et al.,* 2012). The American and African continents are also considered important barriers to species migration, as the hawksbill turtle is adapted to tropical waters and likely does not reach such high latitudes. However, some gene flow has been reported from Indo-Pacific to Atlantic basins (Monzón-Argüello *et al.,* 2011; Vilaça *et al.,* 2013).

Within the ocean basins, hawksbill populations exhibit high levels of genetic diversity, but an evident population structure between rookeries was not observed (LeRoux *et al.,* 2012; Vargas *et al.,* 2016). Therefore, the phylogeographic pattern of the species is complex, likely involving weak isolation by distance and periodic long-distance colonization events.

The connectivity between populations is also influenced by oceanic currents, which

likely drive the formation of foraging aggregations (Blumental *et al.*, 2009). The mixed stocks in these areas comprise of a variety of individuals that come from distant rookeries under the influence of currents. Some cases of long-distance dispersals were registered, including transoceanic migrations (Monzón-Argüello *et al.*, 2010; Vilaça *et al.*, 2013; Putman *et al.*, 2014). These conclusions were achieved based on comparisons of mtDNA haplotype frequencies between rookeries and mixed stocks that have been genetically described through a Bayesian method of mixed stock analysis (MSA), but the lack of standardized data makes these analyses extremely difficult.

Management units (MU) have been recognized as populations with a significant divergence of allele frequencies considering current population structure and management issues (Moritz, 1994). Sea turtle studies suggest the establishment of MUs based on mtDNA variation found for each analyzed sample (Velez-Zuazo *et al.*, 2008; Vargas *et al.*, 2016; Gaos *et al.*, 2017). Wallace *et al.* (2010) developed the idea of regional management units (RMU) on a global scale that considers nesting sites, genetic stocks and geographic distributions. For hawksbills, 13 RMUs were identified worldwide, of which seven were designed as putative RMUs due to the lack of genetic or distribution data. Further studies were based on these delineations to identify knowledge gaps and to evaluate the conservation status of sea turtles (Mazaris *et al.*, 2017).

Therefore, further studies exploring all the species diversity and linking separated studies are essential to comprehend the biology, behavior and life history of hawksbill turtles. In this work, we analyzed a complete dataset compiled of all previously published mtDNA data with a standardized nomenclature to understand the hawksbill turtle dynamics and diversity worldwide.

Material and Methods

Data Collection

We compiled all control region mtDNA haplotypes data available in the literature for hawksbill turtles. The sequences were obtained from the Genbank Database (www.ncbi.nlm.nih.gov, accessed 11 February 2019) and from the Atlantic Ocean hawksbill haplotype database (A. Abreu-Gobrois, personal communication). According to the purpose of

this work, only haplotypes with associated population frequencies data were recovered, including females nesting in rookeries and juveniles in foraging aggregations. Populations with less than three individuals analyzed were grouped with the closest population (<300 km). We analyzed control region haplotypes spanning 739 bp, aiming to increase detection of variation and population structure. We standardized the nomenclature to EiA plus sequential number, when the haplotype was first detected in Atlantic Ocean, and EiIP plus sequential number, when the haplotype was first recorded in Indo-Pacific Ocean. The alignment of sequences was performed with ClustalX algorithm of the MEGA 7 software (Kumar *et al.,* 2016).

Phylogenetic analysis

The divergence time estimates between major mtDNA lineages described for hawksbill turtles was calculated using the software BEAST v2.4.3 (Bouckaert *et al.,* 2014). The choice of the best nucleotide substitution model was performed in jModelTest v. 2.1.7 (Posada, 2008) using a gamma distribution with four rate categories and Bayesian information criteria (BIC). The best model was TrN+G+I with gamma shape=0.71 and proportion invariable=0.75.

We assumed rate homogeneity among branches (strict molecular clock) under the Coalescent Bayesian Skyline Population tree model due to the intraspecific nature of the dataset (Drummond and Bouckaert, 2014). The times to the most recent common ancestors (TMRCAs) based on previous genetic studies (Duchene *et al.,* 2012; LeRoux *et al.,* 2012; Vargas *et al.,* 2016) were used as priors for tree calibration, assuming the monophyly of the group. The root age of the tree was 5.63 million years ago (mya) with a 95% confidence interval of 3.98–8.86 mya.

We conducted three independent runs for 200,000,000 generations, sampled every 1000 generations. Trace files were checked for chain convergence and sufficient effective sample sizes (ESS) in Tracer v. 1.6 (ESS>200 were considered acceptable) (Rambaut *et al.,* 2014). TreeAnnotator v2.4.2 was used to find the maximum clade credibility (MCC) tree within all trees generated.

Genetic Diversity and Population Structure Analyses

We estimated the number of haplotypes, haplotype diversity, nucleotide diversity (per site), number of variable sites and average number of nucleotide differences using the program DnaSP v5 program (Librado and Rozas, 2009). The parsimony relationship among haplotypes was represented in a network using the median-joining algorithm (Bandelt *et al.,* 1999) in the software PopART 1.7 (Leigh and Bryant, 2015).

Genetic structure of populations was defined using the population pairwise F_{ST} values based on haplotype frequencies and exact tests of population differentiation in the software Arlequin v3.5 (Excoffier and Lischer, 2010). We also used the spatial analysis of molecular variance (SAMOVA), which approach is based on the maximization of the proportion of total genetic variance due to the differences between groups (F_{CT}). The analysis was conducted for the complete dataset and separately for each ocean basin in the program SAMOVA 2 (Dupanloup *et al.,* 2002). The largest mean F_{CT} value is associated with the estimated number of simulated groups. This simulated annealing approach was performed testing the number of populations groups (K) from two to 15. Aldabra was not included in the statistical analyses due to the small number of individuals and distance from the other rookeries.

Results

Haplotype frequencies were obtained from the literature for 18 rookeries and 23 foraging aggregations from Atlantic Ocean and for 22 rookeries and 17 foraging aggregations from Indo-Pacific Ocean. We analyzed a total of 1983 individuals from worldwide rookeries that exhibited 88 haplotypes and 1577 individuals from foraging aggregations that displayed 79 haplotypes. The global dataset including all mtDNA control region haplotypes previously recorded for hawksbill turtles is presented in Table 1. The variable positions for control region haplotypes, Genbank access numbers, corresponding haplotypes based on shorter sequences (384 bp) (Bass *et al.,* 1996; Dias-Fernández *et al.,* 1999) and ambiguous names for haplotypes are available in Supplementary Materials Table S1.

Several ambiguities in the nomenclature of mtDNA haplotypes were found (Supplementary Materials Table S1). For example, haplotypes EiBR14 (Vilaça *et al.,* 2013), EiA67 (Proietti *et al.,* 2014) and EiIP33 (Vargas *et al.,* 2016) are equivalent to same 739 bp sequence haplotype. In this work they were treated as EiIP33, as it is related to the major Indo-Pacific lineage. The haplotypes EiBR7 (Vilaça *et al.,* 2013), EiA48 (Putman *et al.,* 2014) and

EiIP16 (Vargas *et al.,* 2016) are also equivalent and here they were named EiIP16, since it is also related to the major Indo-Pacific lineage. Haplotypes Ei_15 (Nishizawa *et al.,* 2016) and EiIP17 (Vargas *et al.,* 2016) correspond to the same 739 bp haplotype that is called EiIP17 in this work. We also found a miscalling in haplotypes EiA23 and EiA41 deposited in the Genbank database by LeRoux *et al.* (2012) (confirmed by personal communication). The haplotype EiA23 (Genbank accession number JN998521) corresponds to EiA41 (Genbank acc. no. EF210793) and EiA41 (Genbank acc. no. JN998517) corresponds to EiA23 (Genbank acc. no. EF210791) according to Velez-Zuazo *et al.* (2008) and the Atlantic Ocean hawksbill haplotype database (Supplementary Materials Table S1).

The complete dataset including worldwide rookeries and foraging aggregations of hawksbill turtles resulted in 126 control region haplotypes. Bayesian phylogenetic analysis indicated the presence of nine major mtDNA clades with the TMRCA for all lineages estimated at 5.02 mya (95% highest posterior density interval: 3.63–7.12 mya), when the Atlantic and Indo-Pacific lineages diverged (Figure 1). The first lineage in the Indo-Pacific basin to split was clade I at approximately 3.77 mya (95% HPD interval: 1.99–5.82 mya) in the early Pliocene. The other clades (IIA, IIB, III, IV and V) diverged in the late Pliocene or early Pleistocene, when the interglacial expansion was occurring. The clade Indo-Pacific II was split into two sub-clades (IIA and IIB) at about 0.41 mya (95% HPD interval: 0.16–0.75 mya), of which one was reported in foraging aggregations of Atlantic Ocean. The two main Atlantic lineages (clades I and II) diverged more recently, about 1.28 mya (95% HPD interval: 0.56–2.16 mya), followed by a more recent split in clade Atlantic II dated at 0.57 mya (95% HPD interval: 0.23–0.99 mya), when Pleistocene glacial cycles were occurring.

The genetic diversity of hawksbill turtle rookeries was compared between Atlantic and Indo-Pacific basins. The haplotype frequencies per rookeries are shown in Figure 2 and Supplementary Materials Table S2. The diversity parameters are higher in the Indo-Pacific basin in relation to the Atlantic basin (Table 1). Spatial analysis of molecular variance (SAMOVA) revealed substantial differentiation among Indo-Pacific and Atlantic population groups (F_{CT} = 0.635), with 63.5% of the variation partitioned between groups.

F_{ST} test (based on haplotype frequencies) indicated genetic differentiation among 22 groups of populations, while the exact test showed population differentiation among 25 rookery groups (Supplementary Materials Table S4). The genetic variance of Indo-Pacific populations was maximally differentiated when they were clustered in five geographical groups, when

65.4% of the variation was found among the following groups: (1) Persian Gulf, Western Australia, Solomon Islands and Eastern Pacific (Mexico, Nicaragua, El Salvador, Ecuador, Costa Rica and Panama), (2) Northern Territory and North Queensland, (3) Seychelles and Chagos Archipelago, (4) Peninsular Malaysia, and (5) East Malaysia (Supplementary Materials Table S5). However, the grouping of Atlantic populations was not consistent, suggesting different partitions for each K.

Based on exact tests of population differentiation, FST and SAMOVA analyses, we defined the most appropriated regional groups considering both genetic variation and geographical distribution. Considering that different groupings of Atlantic populations were proposed in different tests, we decided to treat each Atlantic population as independent, except for the two Tobago populations. For Indo-Pacific populations, a consistent grouping of 13 different regional groups was observed for all tests (Figure 2 and Supplementary Table S2).

The relationships among 739 bp haplotypes of hawksbill rookeries are represented in the mtDNA network (Figure 3). The haplotypes are grouped in seven main clades, which were named according to Vargas *et al.* (2016) and LeRoux *et al.* (2012). Two clades were reported in Atlantic Ocean and five in Indo-Pacific Ocean. Clades Indo-Pacific IV and V are represented each by only one haplotype and they were found exclusively in Australia and Persian Gulf, respectively. Clade II was found in rookeries of Seychelles and Chagos, two sets of islands in the middle of the Indian Ocean, and one individual was recorded in nesting area of East Malaysia. Haplotype EiIP33 was the most widely distributed haplotype in the Indo-Pacific Ocean, being present in 56% of rookeries. Haplotypes EiA01 and EiA11 were the most common haplotypes in Atlantic (78.9% and 63.1% of rookeries, respectively). The star shaped networks of the clades Atlantic I and II and Indo-Pacific I suggest that the groups have experienced a population expansion.

Some haplotypes are exclusive of unique populations or exhibited a very limited distribution. For example, Indian Ocean rookeries of Seychelles and Chagos present 13 exclusive haplotypes, Persian Gulf exhibits 11 exclusive haplotypes, and three of four haplotypes found in nesting areas along the Brazilian coast are also exclusive.

The frequencies of control region 739 bp haplotypes in foraging aggregations are shown in Supplementary Materials Table S3. Seventeen foraging aggregations were surveyed in Indo-Pacific Ocean, being 12 in the Eastern Pacific coast, four in Malaysian islands and one in North Queensland in Australia. In the Atlantic Ocean, 23 foraging aggregations were analyzed, being

six in Southwest, three in Eastern and 14 in Northwest Atlantic (see Supplementary Materials Table S3).

Eastern Pacific presents all the haplotypes from clade Indo-Pacific II. Southwest Atlantic exhibits a prevalence of haplotypes from clade Atlantic II, while haplotypes of both clades are present in Northwest Atlantic. Foraging aggregations in East and Southwest Atlantic present haplotypes without known origin (EiA49, EiA70, EiA75, EiA82 and EiA87) that belong to the clade Indo-Pacific IIA according to phylogenetic analysis. They were typically found in Seychelles and Chagos rookeries and with one record in Malaysia East rookery. The natal origins of 48.1% of the haplotypes registered in foraging aggregations could not be identified, as they had not been registered in rookeries. These 'orphan' haplotypes were found in 76 individuals (4.8% of the individuals analyzed).

Discussion

Genetic stocks of hawksbill turtles

Combining all the previously published mtDNA haplotypes, we were able to investigate the phylogeographic patterns of hawksbill turtles in a global context. We demonstrate that the genetic diversity based on control region of mtDNA among hawksbill turtle rookeries was higher in Indo-Pacific basin relative to Atlantic basin. Six out of the nine phylogenetic clades were found in Indo-Pacific rookeries, which correspond to 69.3% of all compiled haplotypes. It is important to note that only western Atlantic rookeries were analyzed, considering that the available data from eastern Atlantic rookeries comprises short sequences of 384 bp (EATL haplotype) preventing their inclusion in these analyses (Monzón-Argüello *et al.*, 2011).

Some rookeries contain haplotypes from several clades, while other rookeries only contain haplotypes belonging to a single clade. This pattern was observed in Atlantic (LeRoux *et al.*, 2012) and Indo-Pacific basins (Vargas *et al.*, 2016), which reveal a historic pattern of population divergence and subsequent secondary contact for the majority of rookeries.

The SAMOVA based on control region revealed that 63.5% of the variation is partitioned between Indo-Pacific and Atlantic groups. This differentiation was demonstrated by other studies, but only using limited haplotype representation. No haplotype is shared between Atlantic and Indo-Pacific rookeries, which indicates that the American and African continents

are important barriers directing the current genetic diversity distribution of hawksbill reproductive populations.

However, haplotypes were shared between distant intra-oceanic rookeries. No significant genetic differentiation was observed in F_{ST} and exact tests between Colombia and United States Virgin Islands, located about 1000 km apart. In the Indo-Pacific basin, widely distributed rookeries throughout the Persian Gulf and Eastern Pacific were clustered together in SAMOVA analysis even when a number of population groups of K=15 was simulated (Supplementary materials Table S5).

Previous studies demonstrated a positive but weak correlation between genetic and geographic distances of hawksbill rookeries in Atlantic (LeRoux *et al.*, 2012) and Indo-Pacific basins (Vargas *et al.*, 2016). They suggest that philopatric behavior is important to determine genetic stocks, but occurs at a variable scale and depends on the geographic location and the influence of oceanic currents. This finding makes the phylogeographic patterns more complex, refusing the isolation by distance model predominant in the literature.

The connectivity among geographically distant rookeries from Northwest and Southwest Atlantic suggests wide dispersal events. Brazilian populations present four haplotypes, being three exclusive haplotypes derived from EiA01 haplotype. Indeed, Vilaça *et al.* (2013) suggested that EiA01 is the ancestral haplotype for the Brazilian rookeries.

Populations of the Indian Ocean were clustered in two groups: Persian Gulf and Seychelles-Chagos, which present strong genetic differentiation with high number of exclusive haplotypes (11 and 13, respectively). Natoli *et al.* (2017) suggest that the Persian Gulf population originated from a single founder event, followed by population expansion.

Only over the last years, hawksbill rookeries from the Eastern Pacific were examined (Gaos *et al.*, 2016, 2018; Zuniga-Marroquin and Monteros, 2017). They exhibited low genetic diversity and 75% of the haplotypes are exclusive to the region. Haplotypes EiIP106 and EiIP108 were described as unique to the rookeries in mangrove estuaries, a particular hawksbill's behavior found in Pacific Central America (Gaos *et al.*, 2017). Nevertheless, Natoli *et al.* (2017) reported haplotype EiUAE08 in the United Arab Emirates, which matches with the haplotype EiIP106 based on the 739 bp sequence. The duplicity in the haplotype names precluded the detection of a possible origin of the Eastern Pacific haplotype, conducting to uncertain conclusions of species dynamics. The identification of exclusive haplotypes is

important to identify the origin of individuals captured in foraging aggregations, since an adequate compilation of all the haplotype diversity is fundamental to suggest endemic origin.

The star shaped network of the clades suggests that the groups have experienced a population expansion (Figure 3). Signals of demographic expansion have been previously found for turtles from the Persian Gulf (Vargas *et al.,* 2016; Natoli *et al.,* 2017), Caribbean (LeRoux *et al.,* 2012), Mexican Pacific (Zuniga-Marroquin and Monteros, 2017) and Brazil (Vilaça *et al.,* 2013). In the Persian Gulf and Eastern Pacific, this signal is attributed to a recent colonization event.

In a smaller geographical scale, the genetic differentiation between close hawksbill's rookeries is commonly observed. Two separate demographic units were recognized for Northern (Ceará and Rio Grande do Norte) and Southern (Bahia and Sergipe) Brazilian nesting areas (Vilaça *et al.,* 2013). In the Persian Gulf, a population boundary was found between Northern (Iran NW) and Southern regions (United Arab Emirates, Iran SE and Saudi Arabia) (Natoli *et al.,* 2017). In the Caribbean, hawksbill's rookeries on the leeward and windward side of Barbados, separated by 30 km, are genetically distinct (Browne *et al.,* 2010). Whilst Buck Island is not differentiated from Barbados rookery (750 km away), it is demographically distinct from Sandy Point in USVI, located 40 km away (Hill *et al.,* 2018). In short, it is not possible to establish distinct demographic units based simply on geographic distance.

Several particularities deserve additional attention in the understanding of the population structure and dynamics of hawksbill turtles. Foraging aggregations in Puerto Rico have revealed temporal variation in haplotype frequencies (Velez-Zuazo *et al.,* 2008). Australian rookeries located 800 km apart exhibited different nesting timing (Vargas *et al.,* 2016). A new reproductive habitat was recently reported in Eastern Pacific, as nesting sites were found in a mangrove estuary (Gaos *et al.,* 2017). These different characteristics highlight the complexity of species behavior that should be considered in the delimitation of effective management units and conservation programs.

Migration patterns

The lack of knowledge on migratory patterns and the origins of juvenile foraging aggregations due to the difficulties in tracking individuals, is a conservation issue for hawksbill turtles. Using a global dataset based on previously published mtDNA haplotypes, we were able

to identify new connections between populations and gaps that should be prioritized in order to clarify the species dynamics.

The foraging aggregations of Northwest and Southwest Atlantic are well characterized. Brazilian and Caribbean regions share haplotypes, but there is a prevalence of haplotypes from clade Atlantic II at the Brazilian coast, while haplotypes of both clades Atlantic I and II are present in the Caribbean. Foraging aggregations in the Southwest Atlantic present five haplotypes that belong to lineage Indo-Pacific according to phylogenetic analysis (Figure 1 and 4). Three of these haplotypes (EiA49, EiA70 and EiA75) are 'orphans', haplotype EiIP16 was reported in rookeries of Seychelles and Chagos Islands and EiIP33 is broadly distributed in Indo-Pacific basin (Vargas *et al.*, 2016). This is evidence of a transoceanic movement of individuals.

All the 739 bp haplotypes found in foraging aggregations of Cape Verde and Príncipe Island (EiA49, EiA82 and EiA87) were never registered in rookeries. According to the phylogenetic analysis, they belong to clade Indo-Pacific IIA, which is present in rookeries from Seychelles and Chagos Islands and with the record of one individual in the nesting area of East Malaysia. Interestingly, considering the mtDNA data of shorter sequences (384 bp), the clade Indo-Pacific IIA is also found in the rookeries of East Atlantic. Monzón-Argüello et al. (2011) reported the presence of females nesting in Príncipe Island with haplotype EATL, which matches with haplotypes EiA70, EiIP76 and EiIP16 of clade Indo-Pacific IIA. The authors suggested that Príncipe Island was probably colonized by migrants from the Indian Ocean via the Cape of Good Hope in southern Africa. Bowen et al. (2007) also proposed that a rare dispersal event of Indian Ocean into the Atlantic Ocean occurred in Late Pleistocene. Furthermore, high genetic distinctiveness is observed in the rookeries and foraging aggregations in the Eastern Atlantic regarding the Western Atlantic (Monzón-Argüello et al., 2011).

Shamblin *et al.* (2014) reanalyzed some individuals of loggerhead turtles that had shorter haplotypes available, generating longer sequences that yielded better resolution in population structure. The same is desired for some populations of hawksbill turtles, especially those from Eastern Atlantic, in order to include individuals from the Eastern Atlantic in a robust analysis and investigate their phylogenetic linkage with Indo-Pacific hawksbill lineages, as well as allow a better assignment of the origin of individuals from foraging aggregations

The foraging ground in Ascension Island is composed by sea turtles originated in Western Atlantic and Eastern Atlantic rookeries. Simulations of physical transport within an ocean circulation model showed that passive dispersal influencing Ascension Island ground is primarily from the East, involving rookeries along Western Africa and, potentially, the Indian Ocean (Putman *et al.,* 2014).

The integration of ocean drift models and genetic surveys have allowed identifying the influence of oceanic currents in the genetic diversity observed in foraging areas. Aggregations in Brazilian coast that are influenced by the South Equatorial/North Brazil Current present different and higher genetic diversity regarding aggregations influenced by the Brazil Current (Vilaça *et al.,* 2013; Proietti *et al.,* 2014). The dispersal patterns of hawksbill juveniles in Caribbean vary from regionally constrained groups to mixed stocks with broadly distributed individuals, depending on the local and regional current influences (Blumenthal *et al.,* 2009). For example, foraging areas at Caribbean Mexico are compound by a mix of individuals, while areas at Gulf of Mexico are dominated by self-recruited individuals, due to the effect of the Loop Current and its associated gyres (Labastida-Estrada *et al.,* 2018).

In the East Pacific Ocean, all the juvenile individuals sampled in foraging aggregations present haplotypes from clade Indo-Pacific I. Gaos *et al*. (2017) found that these individuals use foraging grounds in the region of their natal beaches, a behavior called natal foraging philopatry. In contrast, Vilaça *et al*. (2013) showed that the rookeries and foraging aggregations in Brazilian coast are distinct demographic units, due to a greater genetic diversity and Indo-Pacific haplotypes found in the foraging grounds. In general, there is a significant correlation between the contribution to foraging populations and proximity to the corresponding rookeries, but this pattern is not absolute and aggregates can connect individuals from distant rookeries (Bowen *et al.,* 2007).

A high proportion of haplotypes (48.1%) have no known source rookery, which may suggest that there are many rookeries not yet studied. Further sampling of "genetically unknown" rookeries is needed to identify haplotype origins and improving MSA resolution. For example, Labastida-Estrada *et al*. (2018) recently performed the genetic characterization of Mexican rookeries and revealed the origin of the haplotype EiA24, previously identified as ‘orphan’ by Pérez-Bermúdez *et al*. (2017). This finding changes the knowledge about the most important source of hawksbill contributions to Cuban foraging aggregations.

Satellite telemetry data can provide fine-scale information about breeding, foraging and migration of sea turtles in different life stages (Hart *et al.,* 2019). The tracking of female hawksbills showed relatively short post-nesting migrations towards foraging grounds in the Pacific and Atlantic basins (Cuevas *et al.,* 2008; Parker *et al.,* 2009; Gaos *et al.,* 2012; Marcovaldi *et al.,* 2012), but long-distance migrations were also recorded (van Dam *et al.,* 2007; Hart *et al.,* 2019). Acoustic telemetry monitoring of juvenile hawksbills in Belize showed high site fidelity over months to years with occasional wide range use of the atoll (Chevis *et al.,* 2017). However, tracking data is rare and restricted to a few regions worldwide (Marcovaldi *et al.,* 2012). There are several limitations to the remote tracking, as the high cost and brief periods coverage, as well as reduced sample size resulting in poor population-level inference.

Flipper tags and recapture data also contribute to track hawksbill migratory patterns. For example, Eastward transatlantic movements of juveniles were confirmed when a juvenile hawksbill tagged in a feeding ground at Atol das Rocas in Brazil was captured in Senegal (Marcovaldi and Filippini, 1991).

Hawksbill turtles feed on sponges and forage at or close to coral reef habitats. The migration patterns and the use of foraging aggregations is related to coral reef distribution. Areas with food availability allow turtles to remain resident for long periods (Marcovaldi *et al.,* 2012). Threats from pollution, overfishing and disease have caused coral reef collapse (Pandolfi *et al.,* 2003), which directly impact hawksbill turtles. Conservation strategies aiming to conserve both hawksbills and coral reefs should be broadly considered (Troëng *et al.,* 2005).

Our review and analyses suggest historical long-distance migration and transoceanic connectivity between hawksbill populations. This reinforces the need for extensive and continuous sampling to improve our understanding of the connections between rookeries and foraging aggregations in order to promote adequate protection for hawksbill turtles.

Phylogenetic history and divergence time

The hawksbill turtle exhibits a deep divergence between Atlantic and Indo-Pacific lineages, followed by a more recent transoceanic colonization event. The divergence of Indo-Pacific lineages of hawksbill turtles occurred during or subsequent to the early Pliocene. Vargas *et al.* (2016) hypothesized that this could be associated with changes in the Indo-Pacific oceanic

currents that happen at this age, as a result of continental movements (Australia and New Guinea), the Isthmus of Panama formation and oceanic temperature changes.

The clade Indo-Pacific I, currently the most widely distributed lineage, is the early diverging lineage in Indo-Pacific basin, having diverged from the remaining clades approximately 3.77 mya. In late Pliocene or early Pleistocene, the split of other clades (II, III, IV and V) occurred in association with the glacial cycles. The multiple radiations of the phylogenetic tree during the late Pleistocene suggest that hawksbill turtles have also experienced a rapid population expansion.

We found that the clade Indo-Pacific II is compound by two sub-clades (IIA and IIB). Based on short mtDNA sequence data, the sub-clade Indo-Pacific IIA is present in rookeries of Príncipe Island in Atlantic Ocean (Monzón-Argüello *et al.*, 2011). It is an evidence of a radiation from the Indian Ocean into Atlantic Ocean during the late Pleistocene.

The Atlantic lineages diverged more recently during the Pleistocene glaciation cycles. Climate change and sea level fluctuations (together possible changes of the currents) were likely responsible for processes of fragmentation and homogenization of populations, affecting population sizes and recolonization events (Reece *et al.*, 2005). Thus, contractions and expansions drove the demographic history of Caribbean populations (LeRoux *et al.*, 2012).

Conclusions

Phylogeographic pattern of hawksbill turtles can be interpreted as two highly divergent global lineages split between Atlantic and Indo-Pacific oceans during the early Pliocene, when they experienced contractions and expansions associated with the glaciation cycles that gave origin to new lineages. Migrations and long-distance dispersal events were likely responsible for the partial homogenisation of some distant populations. During the late Pleistocene, a population expansion and transoceanic dispersals from Indian to Atlantic Ocean were observed, establishing a recent separation between East Atlantic/Indo-Pacific and West Atlantic Ocean basins.

This work demonstrated the importance of analyses including the global distribution and diversity of hawksbill turtles. Global demographic patterns were being missed out due to the use of non-representative data or duplicate haplotype names. The systematic exploration of published data allowed achieving relevant conclusions about hawksbill's phylogeography and

phylogeny. The data compilation and standardization can be highly relevant for the future studies of species. Furthermore, this work highlighted the need to consolidate international cooperation to investigate the life history and demographic patterns of hawksbill turtle in a global perspective.

Acknowledgments

We would like to thank Alberto Abreu-Grobois, Robin LeRoux, Ariel Ruiz Urquiola, Emir Pérez-Bermúdez, Juanita Joseph, Elizabeth Labastida Estrada and Michael Jensen for suppling additional information about mtDNA haplotypes, and Ivan Belo Fernandes for the help with maps figures. We also thank to Projeto TAMAR, particularly Neca Marcovaldi and Luciano S. Soares for the invitation to participate in this research, starting in 2004. L.S.A. received a PhD fellowship from CAPES and F.R.S has a research fellowship from CNPq. This work was supported by FAPEMIG, CNPq and CAPES.

Reference

Bandelt HJ, Forster P and Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Bass AL, Good DA, Bjorndal KA, Richardson JI, Hillis ZM, Horrocks JA and Bowen BW (1996) Testing models of female reproductive migratory behaviour and population structure in the Caribbean hawksbill turtle, *Eretmochelys imbricata*, with mtDNA sequences. Mol Ecol 5:321–328.

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol Ecol 17:3754–3774.

Blumenthal JM, Abreu-Grobois FA, Austin TJ, Broderick AC, Bruford MW, Coyne MS, Ebanks-Petrie G, Formia A, Meylan PA, Meylan AB *et al.* (2009) Turtle groups or turtle soup: dispersal patterns of hawksbill turtles in the Caribbean. Mol Ecol 18:4841–4853.

Bolten AB, Bjorndal KA, Martins HR, Dellinger T, Biscoito MJ, Encalada SE and Bowen BW (1998) Transatlantic developmental migrations of loggerhead sea turtles demonstred by mtDNA sequence analysis. Ecol Appl 8:1–7.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A and Drummond AJ (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLOS Comput Biol 10, e1003537.

Bowen BW and Karl SA (2007) Population genetics and phylogeography of sea turtles. Mol Ecol 16:4886–4907.

Bowen BW, Clark AM, Abreu-Grobois, FA, Chaves A, Reichart HA and Ferl RJ (1997) Global phylogeography of the ridley sea turtles (Lepidochelys spp.) as inferred from mitochondrial DNA sequences. Genetics 101:179–189.

Bowen BW, Bass AL, Soares L and Toonen RJ (2005) Conservation implications of complex

population structure: lessons from the loggerhead turtle (*Caretta caretta*). Mol Ecol 14:2389–2402.
Browne DC, Horrocks JA and Abreu-Grobois FA (2010) Population subdivision in hawksbill turtles nesting on Barbados, West Indies, determined from mitochondrial DNA control region sequences. Conserv Genet 11:1541–1546.
Carreras C, Godley BJ, León YM, Hawkes LA, Revuelta O, Raga JA and Tomás J (2013) Contextualising the Last Survivors: Population Structure of Marine Turtles in the Dominican Republic. PLOS ONE 8, e66037.
Cazabon-Mannette M, Browne D, Austin N, Hailey A and Horrocks J (2016) Genetic structure of the hawksbill turtle rookery and foraging aggregation in Tobago, West Indies. J Exp Mar Biol Ecol 485:94–101.
Chevis MG, Godley BJ, Lewis JP, Jackson Lewis J, Scales KL and Graham RT (2017) Movement patterns of juvenile hawksbill turtles *Eretmochelys imbricata* at a Caribbean coral atoll: long-term tracking using passive acoustic telemetry. Endang Species Res 32:309-319.
Cuevas E, Abreu-Grobois FA, Guzmán-Hernández V, Liceaga-Correa MA and Van Dam RP (2008) Post-nesting migratory movements of hawksbill turtles *Eretmochelys imbricata* in waters adjacent to the Yucatan Peninsula, Mexico. Endang Species Res 10:123–133.
da Silva VRF, Mitraud SF, Ferraz MLCP, Lima EHSM, Melo MTD, Santos AJB, da Silva ACCD, de Castilhos JC, Batista JAF, Lopez GG, *et al.* (2016) Adaptive threat management framework: integrating people and turtles. Environ Dev Sustain 18:1541–1558.
Diaz-Fernandez R, Okayama T, Uchiyama T, Carrillo E, Espinosa G, Marquez R, Diez C and Koike H (1999) Genetic sourcing for the hawksbill turtle, *Eretmochelys imbricata*, in the northern Caribbean region. Chelonian Conserv Biol 3:296–300.
Dupanloup I, Schneider S and Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571-81.
Drummond AJ and Bouckaert RR (2014) Bayesian Evolutionary Analysis with BEAST 2. Cambridge University Press.
Duchene S, Frey A, Alfaro-Núñez A, Dutton PH, Gilbert MTP and Morin PA (2012) Marine turtle mitogenome phylogenetics and evolution. Mol Phylogenet Evol 65:241–250.
Dutton PH, Bowen BW, Owens DW, Barragan A and Davis SK (1999) Global phylogeography of the leatherback turtle (*Dermochelys coriacea*). J Zool 248:397–409.
Excoffier L and Heidi ELL (2010) Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. Mol Ecol Res 10: 564–567.
Gaos AR, Lewison RL, Wallace BP, Yañez IL, Liles MJ, Nichols WJ, Baquero A, Hásbún CR, Vasquez M, Urteaga J *et al.* (2012) Spatial ecology of critically endangered hawksbill turtles *Eretmochelys imbricata*: implications for management and conservation. Mar Ecol Prog Ser 450:181−194.
Gaos AR, Lewison RL, Liles MJ, Gadea V, Altamirano E, Henríquez AV, Torres P, Urteaga J, Vallejo F and Baquero A (2016) Hawksbill turtle terra incognita: conservation genetics of eastern Pacific rookeries. Ecol Evol 6:1251–1264.
Gaos AR, Lewison RL, Jensen MP, Liles MJ, Henriquez A, Chavarria S, Pacheco CM, Valle M, Melero D, Gadea V *et al.* (2017) Natal foraging philopatry of hawksbill turtles in the eastern Pacific Ocean. R Soc Open Sci 4:170153.
Gaos AR, Lewison RL, Jensen MP, Liles MJ, Henriquez S, Chavarria S, Pacheco CM, Valle M, Melero D, Gadea V *et al.* (2018) Rookery contributions, movements and

conservation needs of hawksbill turtles at foraging grounds in the eastern Pacific Ocean. Mar Ecol Prog Ser 586:203-216.

Gorham JC, Clark DR, Bresette MJ, Bagley DA, Keske CL, Traxler SL, Witherington BE, Shamblin BM and Nairn CJ (2014) Characterization of a subtropical hawksbill sea turtle (*Eretmocheyles imbricata*) assemblage utilizing shallow water natural and artificial habitats in the Florida Keys. PLOS ONE 9, e114171.

Hart KM, Iverson AR, Benscoter AM, Fujisaki I, Cherkiss MS, Pollock C, Lundgren I and Hillis-Starr Z (2019) Satellite tracking of hawksbill turtles nesting at Buck Island Reef National Monument, US Virgin Islands: Inter-nesting and foraging period movements and migrations. Biol Conserv 229:1–13.

Hill JE, King CM, Stewart KR, Paladino FV and Dutton PH (2018) Genetic Differentiation of Hawksbill Turtle Rookeries on St. Croix, US Virgin Islands. Chelonian Conserv Biol 17:303–308.

Jensen MP, FitzSimmons NN, Bourjea J, Hamabata T, Reece J and Dutton PH (2019) The evolutionary history and global phylogeography of the green turtle (*Chelonia mydas*). J Biogeogr 46:860– 870.

Kumar S, Stecher G and Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874.

Labastida-Estrada E, Machkour-M'Rabet S, Díaz-Jaimes P, Cedeño-Vázquez JR and Hénaut Y (2018) Genetic structure, origin, and connectivity between nesting and foraging areas of hawksbill turtles of the Yucatan Peninsula: A study for conservation and management. Aquatic Conserv: Mar Freshw Ecosyst 29:211–222.

Lara-Ruiz P, Lopez GG, Santos FR and Soares LS (2006) Extensive hybridization in hawksbill turtles (*Eretmochelys imbricata*) nesting in Brazil revealed by mtDNA analyses. Conserv Genet 7:773–781.

Leigh JW and Bryant D (2015) POPART: full-feature software for haplotype network construction. Methods Ecol Evol 6:1110–1116.

LeRoux RA, Dutton PH, Abreu-Grobois FA, Lagueux CJ, Campbell CL, Delcroix E, Chevalier J, Horrocks JA, Hillis-Starr Z, Troëng S *et al.* (2012) Re-examination of Population Structure and Phylogeography of Hawksbill Turtles in the Wider Caribbean Using Longer mtDNA Sequences. J Hered 103:806–820.

Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lopez-Barrera EA, Reis EC, Vilaça ST and Vargas SM (2016) Genetics and conservation of sea turtles. In: Advances in research techniques for the study of sea turtles. 1ed. New York: Nova Publishers 184:119-163.

Marcovaldi MA and Filippini A (1991) Trans-atlantic movement by a juvenile hawksbill turtle. Mar Turt Newsl 52: 3.

Marcovaldi MA, Lopez GG, Soares LS and López-Mendilaharsu M (2012) Satellite tracking of hawksbill turtles *Eretmochelys imbricata* nesting in northern Bahia, Brazil: turtle movements and foraging destinations. Endang Species Res 17:123−132.

Mazaris AD, Schofield G, Gkazinou C, Almpanidou V and Hays GC (2017) Global sea turtle conservation successes. Sci Adv 3, e1600730.

Monzón-Argüello C, Rico C, Marco A, López P and López-Jurado LF (2010) Genetic characterization of eastern Atlantic hawksbill turtles at a foraging group indicates major undiscovered nesting populations in the region. J Exp Mar Biol Ecol 387:9–14.

Monzón-Argüello C, Loureiro NS, Delgado C, Marco A, Lopes JM, Gomes MG and Abreu-Grobois FA (2011) Príncipe island hawksbills: Genetic isolation of an eastern Atlantic

stock. J Exp Mar Biol Ecol 407:345-354.
Moritz C (1994) Defining 'Evolutionarily Significant Units' for Conservation. Trends Ecol Evol 9:373–375.
Mortimer JA and Donnelly M. (IUCN SSC Marine Turtle Specialist Group) (2008) *Eretmochelys imbricata* - The IUCN Red List of Threatened Species 2008: e.T8005A12881238.
Natoli A, Phillips KP, Richardson DS and Jabado RW (2017) Low genetic diversity after a bottleneck in a population of a critically endangered migratory marine turtle species. J Exp Mar Biol Ecol 491:9–18.
Nishizawa H, Okuyama J, Kobayashi M, Abe O and Arai N (2010) Comparative Phylogeny and Historical Perspectives on Population Genetics of the Pacific Hawksbill (*Eretmochelys imbricata*) and Green Turtles (*Chelonia mydas*), Inferred from Feeding Populations in the Yaeyama Islands, Japan. Zool Sci 27:14–18.
Nishizawa H, Joseph J and Chong YK (2016) Spatio-temporal patterns of mitochondrial DNA variation in hawksbill turtles (*Eretmochelys imbricata*) in Southeast Asia. J Exp Mar Biol Ecol 474:164–170.
Pandolfi JM, Bradbury RH, Sala E, Hughes TP, Bjorndal KA, Cooke RG, McArdle D, McClenachan L, Newman MJH, Paredes G *et al.* (2003) Global trajectories of the long-term decline of coral reef ecosystems. Science 301:955-958.
Parker DM, Balazs GH, King CS, Katahira L and Gilmartin W (2009) Short range movements of post nesting hawksbill turtles (*Eretmochelys imbricata*) from nesting to foraging areas within the Hawaiian Islands. Pac Sci 63:371−382.
Pérez-Bermúdez E, Urquiola AR, Rebordinos-González L, Espinosa-López G, Lee-González I and Petric BLJ (2017) A new approach to mixed stock analysis that informs on the management and conservation of hawksbill in the wider Caribbean: the case of the legal fishery in Jardines del Rey, Cuba. Int J Mar Sci 7: 316-343.
Posada D (2008) jModelTest: Phylogenetic Model Averaging. Mol Biol Evol 25:1253–1256.
Proietti MC, Reisser J, Marins LF, Rodriguez-Zarate C, Marcovaldi MA, Monteiro DS, Pattiaratchi C and Secchi ER (2014) Genetic Structure and Natal Origins of Immature Hawksbill Turtles (*Eretmochelys imbricata*) in Brazilian Waters. PLOS ONE 9, e88746.
Putman NF, Abreu-Grobois FA, Broderick AC, Ciofi C, Formia A, Godley BJ, Stroud S, Pelembe T, Verley P and Williams N (2014) Numerical dispersal simulations and genetics help explain the origin of hawksbill sea turtles in Ascension Island. J Exp Mar Biol Ecol 450:98–108.
Reece JS, Castoe TA and Parkinson CL (2005) Historical perspectives on population genetics and conservation of three marine turtle species. Conserv Genet 6:235–251.
Richardson PB, Bruford MW, Calosso MC, Campbell LM, Clerveaux W, Formia A, Godley BJ, Henderson AC, McClellan K, Newman S *et al*. (2009) Marine Turtles in the Turks and Caicos Islands: Remnant Rookeries, Regionally Significant Foraging Stocks, and a Major Turtle Fishery. Chelonian Conserv Biol 8:192–207.
Shamblin BM, Bolten AB, Abreu-Grobois FA, Bjorndal KA, Cardona L, Carreras C, Clusa M, Monzón-Argüello C, Nairn CJ and Nielsen JT (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, e85956.
Tabib M, Zolgharnein H, Mohammadi M, Salari-Aliabadi MA, Qasemi A, Roshani S, Rajabi-Maham H and Frootan F (2011) mtDNA variation of the critically endangered hawksbill turtle (*Eretmochelys imbricata*) nesting on Iranian islands of the Persian Gulf. Genet Mol Res. 10:1499–1503.

Tabib M, Frootan F and Hesni MA (2014) Genetic diversity and phylogeography of hawksbill turtle in the Persian Gulf. J Biodivers Environ Sci. 4:51–57.

Troëng S, Dutton PH and Evans D (2005) Migration of hawksbill turtles *Eretmochelys imbricata* from Tortuguero, Costa Rica. Ecography. 28:394–402.

Trujillo-Arias N, Amorocho DF, López-Álvarez D and Mejía-Ladino LM (2014) Phylogeographic relations of some feeding and nesting of hawksbill turtle rookeries (*Eretmochelys imbricata*) in the Caribbean and Pacific of Colombia. Boletín de Investigaciones Marinas y Costeras 43:59–182.

van Dam RP, Diez CE, Balazs GH, Colón Colón LA, McMillan WO and Schroeder B (2008) Sex-specific migration patterns of hawksbill turtles breeding at Mona Island, Puerto Rico. Endang Species Res 4:85−94.

Vargas SM, Jensen MP, Ho SYW, Mobaraki A, Broderick D, Mortimer JA, Whiting SD, Miller J, Prince RIT, Bell IP *et al*. (2016) Phylogeography, Genetic Diversity, and Management Units of Hawksbill Turtles in the Indo-Pacific. J Hered 107:199–213.

Velez-Zuazo X, Ramos WD, Van Dam RP, Diez CE, Abreu-Grobois A and Mcmillan WO (2008) Dispersal, recruitment and migratory behaviour in a hawksbill sea turtle aggregation. Mol Ecol 17:839–853.

Vilaça ST, Lara-Ruiz P, Marcovaldi MA, Soares LS and Santos FR (2013) Population origin and historical demography in hawksbill (*Eretmochelys imbricata*) feeding and nesting aggregates from Brazil. J Exp Mar Biol Ecol 446:334–344.

Wallace BP, DiMatteo AD, Hurley BJ, Finkbeiner EM, Bolten AB, Chaloupka MY, Hutchinson BJ, Abreu-Grobois FA, Amorocho D, Bjorndal KA *et al.* (2010) Regional Management Units for Marine Turtles: A Novel Framework for Prioritizing Conservation and Research across Multiple Scales. PLOS ONE 5, e15465.

Wood LD, Hardy R, Meylan PA and Meylan AB (2013) Characterization of a hawksbill turtle (*Eretmochelys imbricata*) foraging aggregation in a high-latitude reef community in southeastern Florida, USA. Herpetol Conserv Biol 8:258–275.

Zuniga-Marroquin T and Monteros AE (2017) Genetic characterization of the critically endangered hawksbill turtle (*Eretmochelys imbricata*) from the Mexican Pacific region. Lat Am J Aquat Res 45:555–5.

Internet Resources

Rambaut A, Suchard MA, Xie D and Drummond AJ (2014) Tracer v. 1.6, http://beast.bio.ed.ac.uk/Tracer (15 Jan 2019)

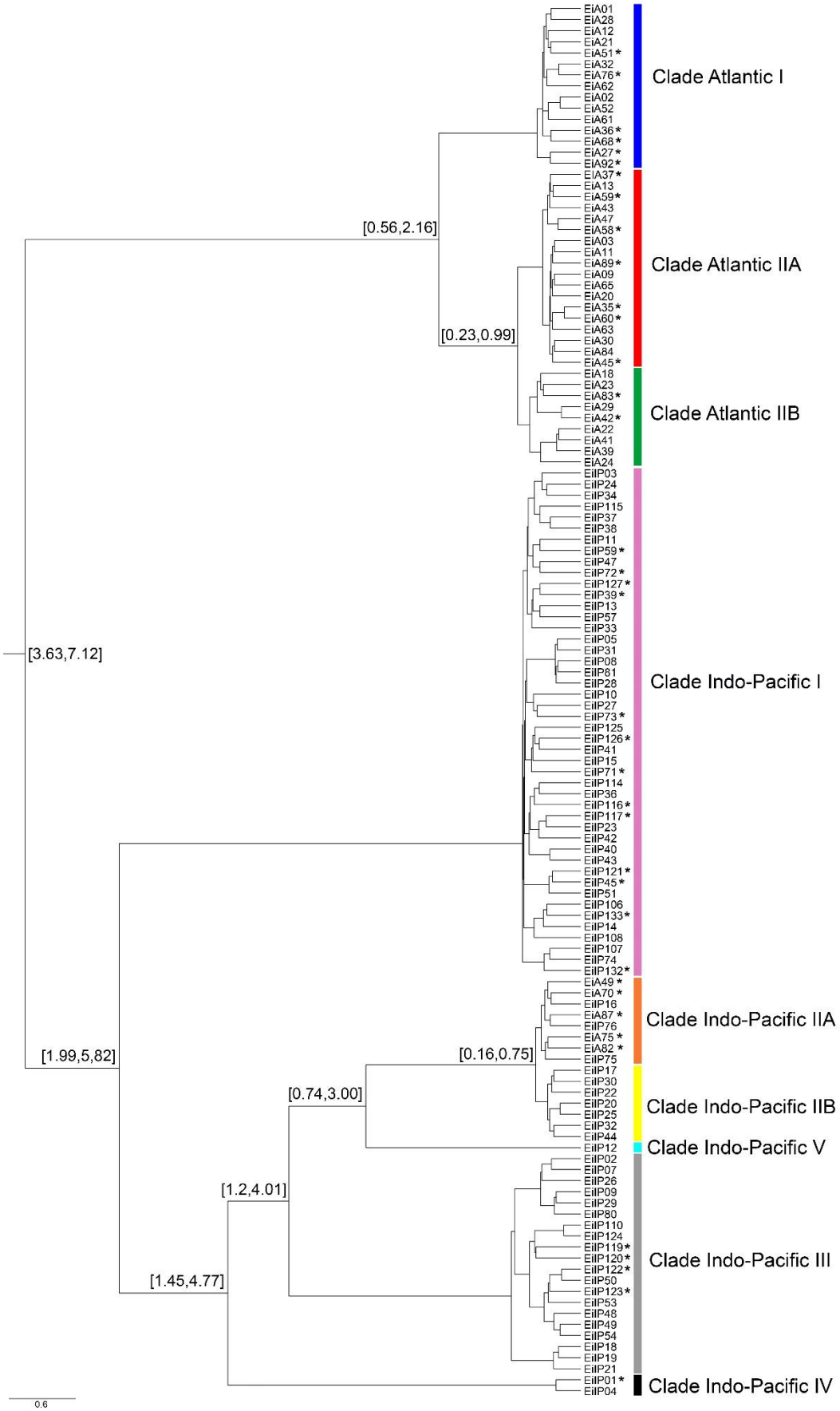
EiA01
EiA28
EiA12
EiA21
EiA51*
EiA32
EiA76*
EiA62
EiA02
EiA52
EiA61
EiA36*
EiA68*
EiA27*
EiA92*
Clade Atlantic I
EiA37*
EiA13
EiA59*
EiA43
EiA47
EiA58*
EiA03
EiA11
EiA89*
EiA09
EiA65
EiA20
EiA35*
EiA60*
EiA63
EiA30
EiA84
EiA45*
Clade Atlantic IIA
EiA18
EiA23
EiA83*
EiA29
EiA42*
EiA22
EiA41
EiA39
EiA24
Clade Atlantic IIB
EiIP03
EiIP24
EiIP34
EiIP115
EiIP37
EiIP38
EiIP11
EiIP59*
EiIP47
EiIP72*
EiIP127*
EiIP39*
EiIP13
EiIP57
EiIP33
EiIP05
EiIP31
EiIP08
EiIP81
EiIP28
EiIP10
EiIP27
EiIP73*
EiIP125
EiIP126*
EiIP41
EiIP15
EiIP71*
EiIP114
EiIP36
EiIP116*
EiIP117*
EiIP23
EiIP42
EiIP40
EiIP43
EiIP121*
EiIP45*
EiIP51
EiIP106
EiIP133*
EiIP14
EiIP108
EiIP107
EiIP74
EiIP132*
Clade Indo-Pacific I
EiA49*
EiA70*
EiIP16
EiA87*
EiIP76
EiA75*
EiA82*
EiIP75
Clade Indo-Pacific IIA
EiIP17
EiIP30
EiIP22
EiIP20
EiIP25
EiIP32
EiIP44
Clade Indo-Pacific IIB
EiIP12
Clade Indo-Pacific V
EiIP02
EiIP07
EiIP26
EiIP09
EiIP29
EiIP80
EiIP110
EiIP124
EiIP119*
EiIP120*
EiIP122*
EiIP50
EiIP123*
EiIP53
EiIP48
EiIP49
EiIP54
EiIP18
EiIP19
EiIP21
Clade Indo-Pacific III
EiIP01*
EiIP04
Clade Indo-Pacific IV
[0.56,2.16]
[0.23,0.99]
[3.63,7.12]
[1.99,5,82]
[0.16,0.75]
[0.74,3.00]
[1.2,4.01]
[1.45,4.77]
0.6

Figure 1 - Bayesian tree based on unique mitochondrial control region haplotypes (739 bp) of hawksbill turtles reported in literature from rookeries and foraging aggregations worldwide. The 95% highest posterior density (HPD) interval values calculated in BEAST are shown in each tree nodes. Branch lengths are proportional to time, with the horizontal axis given in millions of years. Asterisk: ‘orphan haplotypes’ (only found in foraging aggregations).

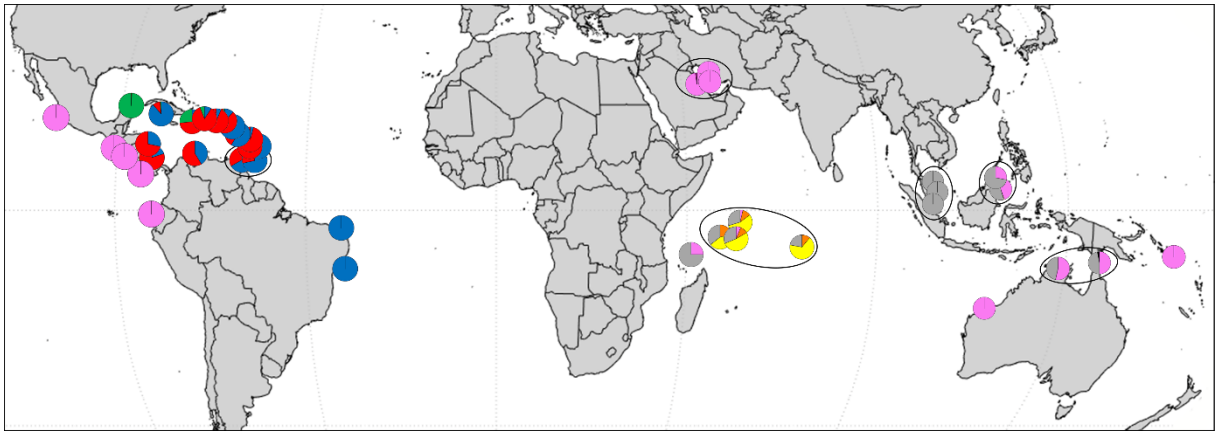

Figure 2 - Frequencies of control region haplotypes (739 bp) from each of nine mtDNA lineages in the hawksbill turtle rookeries. The haplotypes were grouped according to the genetic lineage identified by phylogenetic analysis (see Figure 1) and the pie charts show the proportion of individuals from each rookery that belong to each clade. Major geographical groups are surrounded by an ellipse on the map (see Supplementary Table S2 for geographical group names).

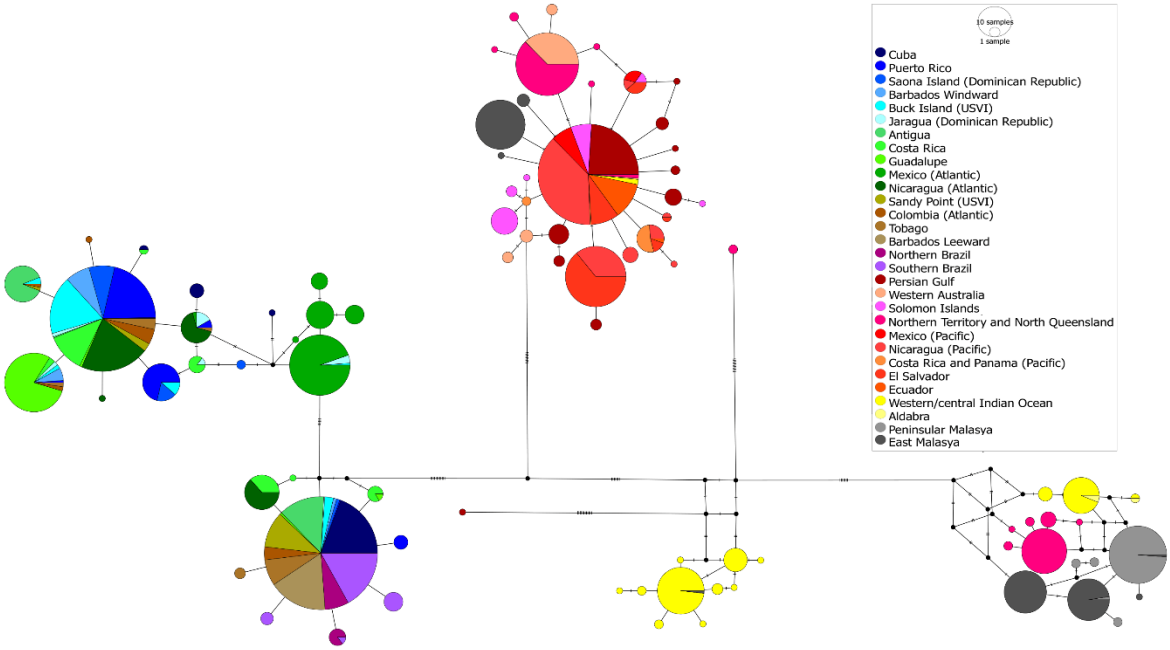

Figure 3 - Haplotype (Median Joining) network based on control region of mtDNA (739 bp) found in hawksbill turtle rookeries. The colors correspond to the 30 geographical groups. Small black dots represent median vectors.

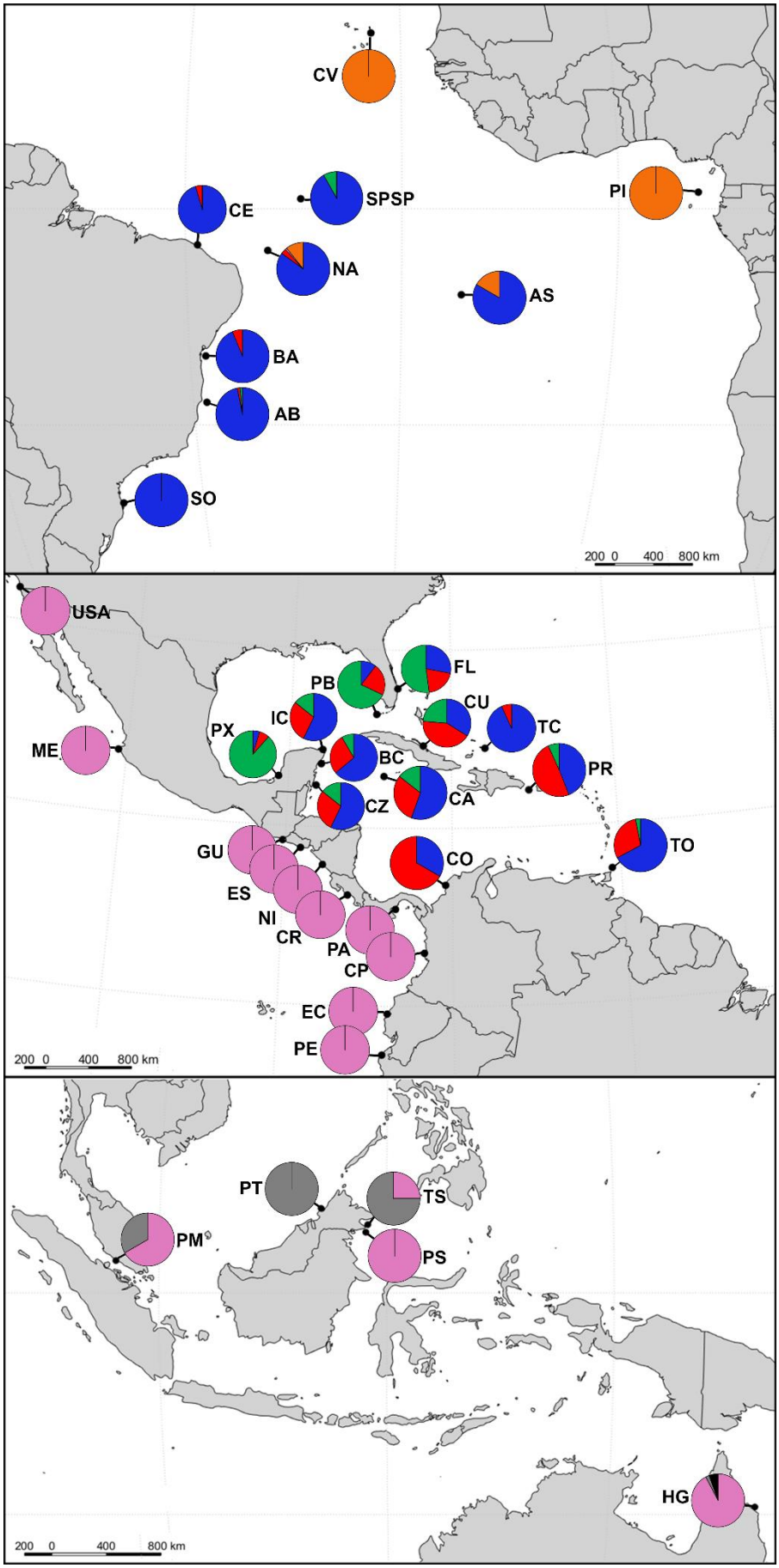
CV
CE
SPSP
PI
NA
AS
BA
AB
SO
200 0 400 800 km
USA
FL
PB
CU
IC
TC
PX
ME
BC
PR
CA
CZ
TO
GU
CO
ES
NI
CR
PA
CP
EC
PE
200 0 400 800 km
PT
TS
PM
PS
HG
200 0 400 800 km

Figure 4 - Frequencies of control region haplotypes (739 bp) in the foraging aggregations of Caribbean and Eastern Pacific (upper), South and Eastern Atlantic (middle) and Indo-Pacific basins (lower). The haplotypes were grouped and colored according to the genetic lineage identified by phylogenetic analysis (see Figure 1). See Supplementary Table S3 for location abbreviations.

Table 1. Dataset of haplotypes based on mitochondrial control region (739 bp) of the hawksbill turtles and genetic diversity of rookeries from Atlantic and Indo-Pacific Ocean. R: number of rookeries, FA: number of foraging aggregations, N: number of individuals, H: number of haplotypes. S: number of polymorphic sites. K: average number of nucleotide differences. π: nucleotide diversity (per site).

Ocean Basin	**Rookeries**						**Foraging aggregations**		
	R	**N**	**H**	**S**	**π**	**K**	**FA**	**N**	**H**
Atlantic	18	992	27	23	0.009	6.684	23	904	45
Indo-Pacific	22	991	61	69	0.022	16.144	17	673	35
Total	40	1983	88	-	-	-	40	1577	79

Supplementary material

Table S1. The variable positions for mtDNA control region haplotypes to hawksbill turtle, the access numbers of Genbank Database, the corresponding haplotypes based on shorter sequences (384 or 480 bp) (Bass et al., 1996; Dias-Fernández et al., 1999) and the ambiguous haplotype names.

Table S2. Compilation of mtDNA control region haplotype (739 bp) frequencies for Hawksbill turtle rookeries.

Table S3. Compilation of mtDNA control region haplotype (739 bp) frequencies for Hawksbill turtle foraging aggregations. Underlined haplotypes are only found in foraging aggregations ('orphan' haplotypes). Cells highlighted in grey correspond to individuals reported in Atlantic Ocean that present haplotypes belonging to Indo-Pacific lineage.

Table S4. Population pairwise F_{ST} based on haplotype frequencies (above diagonal) and P-values of exact tests of population differentiation method (below diagonal) between hawksbill turtle rookeries. (*) Indicate nonsignificant F_{ST} values ($p>0.05$). (+) Indicate significant P-value ($p<0.05$). (-) Indicate non-significant P-value ($p>0.05$).

Table S5. F-statistics (F_{SC}, F_{CT}, and F_{ST}) associated to the K groups calculated using mtDNA control region of hawksbill turtles from Indo-Pacific rookeries.

CAPÍTULO II

New genetic insights about hybridization and population dynamics of sea turtles from Brazil

Larissa Souza Arantes, Sibelle Torres Vilaça, Camila Junqueira Mazzoni, Fabrício Rodrigues Santos. Manuscrito submetido ao periódico Journal of Heredity.

New genetic insights about hybridization and population dynamics of sea turtles from Brazil

Larissa S. Arantes[1], Sibelle T. Vilaça[2,3], Camila J. Mazzoni[2,3], Fabrício R. Santos[1]

[1]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Avenida Antônio Carlos, 6627, Belo Horizonte - MG, Brazil, 31270-010. Phone: +55 31 3409 2581.
[2]Berlin Center for Genomics in Biodiversity Research (BeGenDiv), Königin-Luise-Straße 6-8, Berlin, Germany, D-14195. Phone: +49 (0)30 838 59961.
[3]Evolutionary Genetics Department, Leibniz Institute for Zoo and Wildlife Research (IZW), Alfred-Kowalke-Straße 17, Berlin, Germany, D-10315. Phone: +49 (0) 30 5168 0.

Abstract

An extremely high incidence of hybridization among sea turtles is found along the Brazilian coast. To understand this atypical phenomenon and its impact on sea turtle conservation, research focused in the evolutionary history of sea turtles is fundamental. We assessed high quality multilocus haplotypes of 143 samples of the five species of sea turtles that occur along the Brazilian coast to investigate the hybridization process and the population dynamics of hawksbill (*Eretmochelys imbricata*) and loggerhead turtles (*Caretta caretta*) of Brazil. The use of multilocus data provided a greater power to differentiate species and to characterize hybrid crossings and introgression levels. In contrast to prior results, only hatchlings of F1 female hybrids (hawksbill x loggerhead) were confirmed as backcrossed hybrids, indicating that the introgression was previously overestimated and hybridization may be a more recent phenomenon happening in Brazil. Phylogenetic analyses using nuclear markers recovered the mtDNA-based Indo-Pacific and Atlantic lineages for hawksbill turtles, demonstrating a deep genetic divergence dating from the early Pliocene. In contrast, loggerhead turtles that share a common feeding area and belong to distinct Indo-Pacific and Atlantic mtDNA clades present no clear genetic differentiation at the nuclear level. Finally, our data suggest that different nesting populations of hawksbill and loggerhead turtles along the Brazilian coast are likely connected by male-mediated gene flow.

Introduction

Sea turtles have complex life cycles, with life stages associated with different environments affected directly by human activities. This close interaction exposes them to several threats, which led to a global population decline of most species. The main threats are related to fisheries bycatch, coastal development, pollution (sewage, garbage, toxic substances), pathogens and exploitation of eggs, meat or other turtle products (Wallace et al. 2011). Thus, the monitoring of sea turtle populations followed by actions to mitigate the anthropogenic impact are essential to their conservation worldwide.

Understanding the population dynamics of sea turtles is challenging due to their highly migratory behavior, long lives and different levels of population structure associated with each life stage (Bowen and Karl 2007). Methodological advances in the last decades, such as satellite telemetry and molecular analyses, made important contributions to the comprehension of complex sea turtle behaviors, deepening the knowledge about the factors affecting the composition of foraging aggregations (Carreras et al. 2011; Proietti et al. 2014b), the migration route of pelagic juveniles that is known as "lost years" (Putman and Mansfield 2015; Briscoe et al. 2016), the frequency of occurrence of multiple paternity (Moore and Ball 2002; González-Garza et al. 2015), the level of gene flow among populations (Bowen et al. 2005; Monzón-Argüello et al. 2011; Clusa et al. 2018) and opportunistic mating systems (Stewart and Dutton 2011). Novel technologies have also allowed to investigate the arguable reproductive isolation of sea turtle species, as interspecific hybridization was detected among five out of seven extant species of sea turtles (Vilaça et al. 2012). Some hybridization cases involve crosses of species of the Cheloniidae family that diverged at 63 million years ago (Naro-Maciel et al. 2008), probably the most deeply divergent species group capable of producing viable hybrids in nature (Karl et al. 1995).

Hybrid zones of sea turtles may occur where there is an overlap of nesting areas and reproductive seasons of two or more species, which is for two coastal areas of Brazil (Soares et al. 2017). The hybridization process of sea turtles in the northeastern Brazilian coast is atypical, since the frequency of hybrids is much higher than in any other analyzed population worldwide. While nesting sites surveyed around the world presents no more than 2% of hybrids, the frequency of hybrids along the Brazilian coast was much larger than 20% in some nesting sites (Lara-Ruiz et al. 2006; Reis et al. 2010b; Vilaça et al. 2012).

In the northern coast of the Bahia state in Brazil, where the largest in-country rookeries of hawksbill (*Eretmochelys imbricata*) and loggerhead turtles (*Caretta caretta*) are found, 42% of female turtles morphologically identified as *E. imbricata* exhibited mitochondrial sequences of *C. caretta* (Lara-Ruiz et al. 2006). A more recent study confirmed that the incidence of hybrids is as high as 31.58% of the assumed *E. imbricata* population (Soares et al. 2018). The majority of the surveyed hybrids appear to have 50% of alleles of each parental species, thus being considered first generation (F1). However, backcrossings with either parental species were also detected, revealing the occurrence of introgression (Vilaça et al. 2012). Nevertheless, hybridization seems to be a relatively recent event, spanning about two generations (40 years), from a period that sea turtles' populations were heavily depleted due to anthropogenic impact.

In a nearby nesting site in the Sergipe state of Brazil, where there is a spatial and temporal overlapping of nesting of olive ridley (*Lepidochelys olivacea*) and loggerhead turtles, 27% of individuals morphologically assigned as loggerheads were shown to be hybrids between olive ridley and loggerhead turtles (Reis et al. 2010b), all of them classified as F1 hybrids (Vilaça et al. 2012).

This large frequency of interspecific hybrids in Brazil is an important conservation concern because it may result in outbreeding depression, which is a decrease of the fitness and/or reproductive viability of the local population (Allendorf et al. 2001; Maheshwari and Barbash 2011). Outbreeding depression can be observed in F1 hybrids, and also in F2 or later generation due to the disruption of coadapted gene complex as a result of meiosis and recombination during gametogenesis in F1 hybrids (Goldberg et al. 2005). However, other studies suggest that interspecific hybridization may eventually represent an important source of variation as it may confer an advantageous effect on fitness, also called adaptive introgression (Hedrick 2013). In any case, it is extremely necessary to carefully investigate the consequences of hybridization for the populations where it occurs particularly in high frequency, like the ones in Brazil.

Two previous studies (Soares et al. 2017, 2018) were designed to evaluate the likely outbreeding depression effects of hybridization by the comparison of several reproductive parameters between nesting sites of F1 hybrids and parental species, located at Bahia. Even though emergence success was shown to be lower for hybrid nests, other parameters such as the hatchling production per clutch and clutch frequency were similar to parental species, suggesting that hybrids may persist in this region (Soares et al. 2017). The initial viability of

hybrid hatchlings was also similar to non-hybrid hatchlings, revealing no significant evidence for hybrid breakdown at this early stage (Soares et al. 2018). However, once hybrid hatchlings achieve the sea, little is known about their survival until adulthood and reproductive fitness.

Other genetic studies with the Brazilian populations have investigated their demographic history (Bjorndal et al. 2006; Vargas et al. 2008; Molfetti et al. 2013), population structure (Reis et al. 2010a; Vilaça et al. 2013; Shamblin et al. 2014), mixed stocks at foraging aggregations (Proietti et al. 2009; Reis et al. 2010a; Vilaça et al. 2013; Proietti et al. 2014b;) and interspecific hybridization (Lara-Ruiz et al. 2006; Reis et al. 2010b; Vilaça et al. 2012; Proietti et al. 2014a; Soares et al. 2017, 2018). For example, phylogeographic analyses using mtDNA showed genetic divergence among three Brazilian populations of *C. caretta*, suggesting the recognition of three different management units (Shamblin et al. 2014), while two separate demographic units were recognized for *E. imbricata* in Brazilian nesting areas (Vilaça et al. 2013). The mixed stocks found at foraging aggregations along the Brazilian coast demonstrated connectivity among distant ocean basins (Reis et al. 2010a) and are highly influenced by oceanic currents (Vilaça et al. 2013; Proietti et al. 2014b).

All the above-cited studies have used data from mitochondrial DNA (mtDNA), microsatellites and/or other few nuclear (nDNA) markers, but recent advances in high-throughput sequencing technology (NGS) have opened new opportunities to survey genome wide data in multilocus approaches. Indeed, population genomic methods allow the survey of selected subsets of genetic markers in a large number of individuals (Harrisson et al. 2014). Furthermore, genome wide multilocus data is currently used in many ecological and evolutionary studies, providing inferences about the life history, population dynamics and demographic patterns of species, with important conservation implications.

In this work, we used multilocus data selected from a genomic survey to investigate the hybridization process and life history of sea turtles of the Brazilian coast, focusing mainly in populations of *E. imbricata* and *C. caretta*. We used the double-digest RADseq (ddRADseq) method (Peterson et al. 2012) to select informative loci/haplotypes to generate high quality phased sequences for characterization of species and hybrids for population and interspecific studies. We compared our results with previous studies to test the efficiency of RADseq-derived multilocus markers and to improve the knowledge about the hybridization phenomenon and population dynamics of sea turtles along the Brazilian coast.

Materials and Methods

Sampling

We analyzed 143 DNA samples from the five species of sea turtles and hybrid individuals from four different hybrid classes (Vilaça et al. 2012) that occur along the Brazilian coast (Table 1). The DNA samples were derived from individuals collected between 1999 and 2011 by Projeto TAMAR team, a consolidated and successful Brazilian Sea Turtle Conservation Program. Some samples have already been surveyed in previous studies of our research group using mtDNA and few autosomal markers (Lara-Ruiz et al. 2006; Reis et al. 2010a; Vilaça et al. 2012; Vilaça et al. 2013). The species, localities and number of individuals analyzed are shown in Table 1. Detailed information of hybrid individuals (locality, morphology, collected individual, mtDNA haplotype based on control region and previous classification by Vilaça et al. 2012) is available in Supplementary Table S1. All *C. caretta* x *E. imbricata* (*Cc* x *Ei*), *E. imbricata* x *L. olivacea* (*Ei* x *Lo*) and *E. imbricata* x *C. caretta* x *Chelonia mydas* (*Ei* x *Cc* x *Cm*) hybrids reported in rookeries were from Praia do Forte (Bahia), and *C. caretta* x *L. olivacea* (*Cc* x *Lo*) hybrids from Pirambu (Sergipe) rookeries. Other hybrid individuals were reported from foraging aggregations or bycatch in fisheries.

Discovery and standardization of nuclear markers

Sea turtle genomic structure is quite monotonous, presenting slow cytogenetic, morphologic and molecular divergence among species (FitzSimmons et al. 1995, Naro-Maciel et al. 2008, Wang et al. 2013). Indeed, our initial surveys of intra and interspecific variation for different sea turtles (Vilaça et al. 2012) revealed very few informative nuclear DNA (nDNA) markers. In this study, we developed a protocol to identify informative markers for both intra and interspecific variation using an initial genomic screening for variable anonymous loci.

The initial selection of nDNA markers was performed using a reduced genomic dataset produced via ddRADseq. As a preliminary analysis to help to establish a standardized ddRADseq protocol for sea turtles (manuscript in preparation), one individual of each parental species (*E. imbricata* and *C. caretta),* and one F1 hybrid individual were analyzed. The sequencing library was generated by digesting the genomic DNA using the restriction enzymes

*Nde*I and *Mlu*CI with subsequent ligation of Illumina adapters followed by a 10-cycle Polymerase Chain Reaction (PCR) for completeness of sequencing adapters, as described in Peterson et al. (2012). Libraries were pooled and size-selected between 500-600 bp using PippinPrep equipment (Sage Science). Sequencing was performed on an Illumina MiSeq machine using a 600-cycle kit with the 300 bp paired-end sequencing mode. Samples were demultiplexed and the three samples were run through the pyRAD (Eaton 2014) pipeline for homologous loci recognition and genotyping. Briefly, the 300 bp-long reads were merged using PEAR (Zhang et al. 2013) and aligned into single sequences and those were clustered at 85% identity with a minimum coverage of ten and a maximum of five heterozygous sites per locus. The selected loci were manually screened for interspecific variation between the two species, for which their sequences were extracted from the pyRAD output file and aligned using MUSCLE (Edgar 2004). We selected for subsequent primer design only loci found in all parental species and hybrids, showing a maximum of two indels and at least two interspecific differences between *E. imbricata* and *C. caretta*, which were confirmed as heterozygous in the hybrids. Primers were designed using the Primer3 (Untergasser et al. 2012) algorithm implemented in Geneious 8.1 (Kearse et al. 2012) using default parameters (Supplementary Table S2). Thus, we selected initially 24 anonymous nDNA markers with intra and inter-specific variation to be further validated for population studies.

The validation of RADseq-derived nDNA markers was made through PCR amplification and Sanger sequencing. PCR was done in a final volume of 20 µl using 200 µM dNTP, 0.5 units of Platinum™ Taq DNA polymerase (Invitrogen™ by Life Technologies™), 1.5 mM of $MgCl_2$, 0.5 µM primers forward and reverse and 10 ng genomic DNA in 1X reaction buffer. PCR conditions were performed with one initial denaturation cycle of 95ºC for 5 minutes, 35 cycles of denaturation of 95ºC for 30 seconds, variable annealing temperatures (Supplementary Table S2) for 40 seconds, extension at 72 ºC for 1 minute, and a final extension at 72 ºC for 7 minutes. PCR products were purified by precipitation using a solution of 20 mM polyethylene glycol and 2.5 mM NaCl.

The Sanger sequencing reaction was performed using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems™) following the manufacturer's standard protocol. Forward and reverse sequences were generated on the ABI 3130xl DNA sequencer (Applied Biosystems™). The SeqScape v2.6 software (Applied Biosystems™) was used to check the quality of the electropherograms. The heterozygous sites were verified for accuracy and coded

as ambiguous sites according to the IUPAC code. The alignment of high-quality consensus sequences was performed with the ClustalW algorithm in the MEGA 7 software (Kumar et al. 2016). The PHASE algorithm (Stephens et al. 2001) was used for the reconstruction of the gametic phase of the heterozygous sequences with the assistance of Seq-PHASE input/output interconversion tool (Flot 2010). The DnaSP v5 program (Librado and Rozas 2009) was used for the haplotype assignment. Heterozygous indels found in some sequences of locus 966 were phased using the Indelligent web tool (Dmitriev and Rakitov 2008). Finally, high quality phased sequences were verified again with the overlapping sequence chromatographs to edit for any inconsistencies.

Fifteen individuals of *E. imbricata*, 15 *C. caretta*, ten hybrids, as well as two individuals of each species green turtle (*Chelonia mydas*), olive ridley turtle (*L. olivacea*) and leatherback turtle (*Dermochelys coriacea*), were initially analyzed for the 24 selected loci. To decrease a likely ascertainment bias against intraspecific variation, we have we have characterized the intra and interspecific variation found on diploid (phased) Sanger sequences. Thus, 14 out of the 24 loci were selected to be analyzed for a greater number of individuals (Supplementary Table S2). The most variable loci were selected to intraspecific analyses for species *C. caretta* and *E. imbricata* (Table 2), while the loci with great power to distinguish different species were used to hybrid analysis (Supplementary Table S2).

All sequences generated in this study have been deposited in GenBank and this research is registered in the National System for Genetic Heritage and Associated Traditional Knowledge (SisGen) under number A03A2C2.

Analyses of hybrids

A phylogenetic network was built to represent the interspecific lineages admixture of hybrid individuals (Joly et al. 2015). We estimated genetic distances (Joly et al. 2015) using a distance matrix of alleles and converting it into a distance matrix of individuals using the program POFAD (Joly and Bruneau 2006). We used MEGA 7 software (Kumar et al. 2016) to generate the genetic distances using Kimura-2-parameters model for each of the 14 loci (Dataset 1 in Supplementary Table S2) and then generated a combined-locus distance matrix using POFAD. The resulting matrix was used to build a phylogenetic network (neighborNet) using the software SplitsTree 4 (Huson and Bryant 2006).

Bayesian clustering analysis was done in STRUCTURE software (Pritchard et al. 2000) for inference on population structure and assignment of individuals to populations using multilocus data. We assumed the admixture model where the individuals may have mixed ancestry in more than one of the K populations (species), allowing detection of the introgression level (Pritchard et al. 2000; Falush et al. 2003).

Five loci were excluded from the STRUCTURE analysis because they present either high-level of shared haplotypes between species or a considerable level of missing data. Two individuals of *D. coriacea* were excluded due to a large amount of missing data, likely due to the low level of homology in the selected primers originally designed from *E. imbricata* and *C. caretta* sequences. The final dataset was composed by haplotypic data inferred for nine nDNA loci (Dataset 2 in Supplementary Table S2). We also performed intraspecific analyses using the datasets including 11 loci for *C. caretta* and 14 loci for *E. imbricata* (Table 2). Twenty independent runs for each K value (from K=1 to K=7) were performed with 200,000 Markov Chain Monte Carlo (MCMC) repeats after a 100,000 burn-in period. The independent and correlated allele frequencies were tested. The best K was assessed using Evanno's methodology (Evanno et al. 2005) through the online tool STRUCTURE Harvester (Earl and VonHoldt 2012). We combined the replicate result files and visualized the estimated membership coefficients using CLUMPAK (Kopelman et al. 2015).

The posterior probability of each individual to belong to different hybrid classes was analyzed in NewHybrids v. 1.1 Beta3 (Anderson and Thompson 2002). Separate datasets combining different hybrid crossings were tested, since the NewHybrids only consider hybridization events involving two diploid species (Anderson 2008). Therefore, individuals resulted from crosses involving likely more than two species (R0264 and R0265) could not be analyzed. The analysis was done using the Jeffrey option, no priors, burn-in period of 100,000 and 500,000 MCMC sweeps. The following genotype classes were considered: pure parental (Pure 1 and Pure 2), first and second generation hybrids (F1 and F2 between F1 hybrids) and backcrosses between F1 and pure parental (BC1 and BC2). The R package HybridDetective was used to plot NewHybrids analysis (Wringe et al. 2017).

Genetic diversity and population structure

Population analyses were performed for *C. caretta* and *E. imbricata* using nDNA markers with larger intraspecific variation. Diversity indexes were generated using the Arlequin v3.5 (Excoffier and Lischer 2010), DnaSP and MEGA software. The summary statistics used were: number of haplotypes (H), haplotype diversity (k) and number of polymorphic sites (S). Principal Component Analysis (PCA) was performed using R package *adegenet* to evaluate the genetic diversity among the sampled individuals (Jombart and Ahmed 2011). The missing data was replaced by the mean allele frequency and the PCA of standardized allele frequencies at the individual level was calculated using multivariate methods without spatial components. The analyses were performed including individuals collected in rookeries and feeding areas along the Brazilian coast, or only sea turtles sampled in Brazilian rookeries.

To investigate the relationship of different lineages and to represent great part of the genetic diversity within species, mitochondrial control region haplotypes were compiled from literature and depicted in a network analysis. For *C. caretta*, the haplotypes (776 bp) were obtained from Shamblin et al. (2014), Nishizawa et al. (2014) and from the database of The Archie Carr Center for Sea Turtle Research (http://accstr.ufl.edu/resources/mtdna-sequences). For *E. imbricata*, the control region haplotypes (739 bp) were obtained from LeRoux et al. (2012), Vilaça et al. (2012), Vargas et al. (2016) and Gaos et al. (2018). The haplotype networks were constructed using the Reduced Median algorithm with reduction threshold 9 followed by Median Joining algorithm (RM-MJ network - Bandelt et al. 1995) using the software Network 5.0 (http://www.fluxus-engineering.com). The delimitation and nomenclature of the mtDNA clades were based on previous studies (LeRoux et al. 2012; Shamblin et al. 2014; Vargas et al. 2016) and are available in Supplementary Table S3.

To investigate the multilocus allelic variation, we built a phylogenetic network from a combined-locus genetic distance matrix. We used the dataset of 11 loci for *C. caretta* and 14 for *E. imbricata*, and performed the method using POFAD and SplisTree 4 software as described above.

Phylogenetic analysis

A phylogenetic relationship between sea turtle species was inferred using multilocus data in a Bayesian method implemented in BEAST v2.4.3 (Bouckaert et al. 2014). The sequences of 14 anonymous loci (Dataset 1 in Supplementary Table S2) were analyzed for five species of sea turtles.

The selection of partitioned models of molecular evolution was made using the PartitionFinder2 software (Lanfear et al. 2017). The best-fit model was selected by AICc criterion (Supplementary Table S4). The phylogenetic tree was inferred assuming a relaxed lognormal molecular clock under the Birth Death model. This model of diversification assumes the probability of speciation and extinction along the lineage. It was employed considering that sequences from different species were used and the species were sampled in different levels and presented very different branch length (Drummond and Bouckaert 2014). Time to most recent common ancestors (TMRCA) based on fossil evidence and previous genetic studies (Bowen et al. 1993; Duchene et al. 2012) were used as priors to tree calibration with a lognormal distribution, as follows: 1) split between Dermochelidae and Cheloniidae family as 115 million years ago (mya) with a 95% confidence interval of 106–130 mya (Hirayama 1998) and 2) Carettini and Chelonini tribe as 65 mya with a 95% confidence interval of 50–90 mya (Moody 1974; Cadena and Parham 2015). The monophyly of the ingroup and outgroup was assumed a priori. The estimated date should be interpreted as maximum age constraints of the nodes.

Three independent MCMC chains were run for 200,000,000 generations and sampled every 5,000 generations. Trace files were checked for chain convergence and sufficient effective sample sizes (ESS) in Tracer v. 1.6 (http://beast.bio.ed.ac.uk/Tracer), considering ESS>200 as acceptable. The maximum clade credibility (MCC) tree was summarized after a 50% burn-in in TreeAnnotator from the 20,000 trees.

Results

Analyses of hybridization

High quality multilocus data standardized in this work was used to identify hybrids and estimate the introgression level in sea turtles. Some nDNA markers were more informative to

characterize hybrids since they presented species-specific haplotypes that allowed us to identify the parental origin of the alleles with greater confidence. Loci 856, 3061, 76958 and 109472 were analyzed for a greater number of individuals and presented a larger number of diagnostic sites to identify *Cc* x *Ei* hybrids, while the loci 421, 3061 and 109472 have greater power to identify *Lo* x *Ei* hybrids, and the loci 421, 966, 67959 and 114650 to identify *Cc* x *Lo* hybrids (Supplementary Table S2).

Considering the combined data from all 14 nDNA loci selected with interspecific differences, the POFAD analysis produced a well-resolved dendrogram of all individuals. The five sea turtle species were recovered in different clusters and the hybrids were observed in an intermediate position between species involved in the hybridization process (Figure 1). This method allowed the characterization of the genomic admixture of hybrids using distance measures to estimate the contribution of parental genomes.

The clustering analysis generated by STRUCTURE using correlated allele frequencies model showed that the number of clusters that best fit the data according to the Evanno's statistics (Evanno et al. 2005) was five (Figure 2). The four sea turtle species included in this analysis were distinguished in different groups with high probability (99.9%) according to NewHybrids, and individuals of *C. caretta* were clustered in two different subgroups, one corresponding to individuals with mitochondrial haplotypes commonly found in Brazilian populations, and another corresponding to foraging individuals sampled at Elevação do Rio Grande (ERG) that present mitochondrial haplotypes found in rookeries of the Caribbean, Mediterranean and Indo-Pacific oceans. The same intraspecific subdivision was obtained when *C. caretta* individuals were analyzed separately (Supplementary Figure S1). This population structure found for *C. caretta* was not observed using independent allele frequency model (Supplementary Figure S2). All individuals of *E. imbricata* from Brazil were attributed to a single population in STRUCTURE analysis using both multi-species (Figure 2) and species-specific (Supplementary Figure S1) datasets.

Since the admixture model was assumed in STRUCTURE, the introgression level of hybrids could be inferred. F1 hybrids displayed intermediary genomic composition between parental species. All *Cc* x *Lo* and *Ei* x *Lo* hybrids were classified as F1 with a probability of 99.9% according to NewHybrids analysis. The *Cc* x *Ei* hybrids identified as F1 presented a posterior probability of 99.9% (NewHybrids) of belonging to this category. The parental *C.*

caretta population involved in hybridization cases is associated with mtDNA haplotypes typically found in Brazil.

Three individuals (R0069, R0072 and R0217) previously assigned as hybrids (F1 or >F1) by Vilaça et al. (2012) showed no evidence of admixture between species for all nine nDNA markers analyzed. Individuals R0069 and R0072 were identified as *E. imbricata* and R0217 as *C. caretta* with high posterior probability. This could have resulted from sample misidentification and/or the genetic markers used (microsatellites, RFLP and few nDNA markers) did not have enough diagnostic power to identify hybrids. The misidentification of these three samples was confirmed by re-sequencing the loci RAG1 and CMOS used by Vilaça et al. (2012), which reinforced that they are indeed 'pure' individuals (Supplementary Table S5). We have also re-sequenced the control region of mtDNA for individual R0072 and, in contrast to the previous work, it has haplotype of *E. imbricata*. For the individual R0217 "morphologically" identified as *E. imbricata*, all the genetic data suggest that it is a pure *C. caretta*, probably due to misidentification.

Previous work (Vilaça et al. 2012) identified 17 individuals as introgressed or >F1 hybrids, of which 15 were re-analyzed in this work with a multilocus approach. Using our nDNA dataset, we were able to recognize only six individuals with evidence of being >F1 generation hybrids (Supplementary Figure S3 and Table S1). Remarkably, they were all hatchlings collected in nests and showing characteristics of more than one sea turtle species (Vilaça et al. 2012). Individual R0025 is a hatchling of a *Cc* x *Ei* hybrid female (R0024) and it was attributed to the category of backcrossing with *E. imbricata* with a probability of 99.8% (NewHybrids). Individual R0196 was sampled with morphological evidence of hybridization and it was identified as backcrossing with *E. imbricata* with a probability of 96.8% (NewHybrids). The remaining four hatchlings (R0264, R0265, R0267 and R0268) are siblings derived from a single clutch. The genetic mixture of three species *E. imbricata* x *C. caretta* x *C. mydas* (*Ei* x *Cc* x *Cm*) was confirmed in two individuals (R0264 and R0265), although the posterior probability could not be estimated because NewHybrids only considers hybridization cases involving two species. The remaining siblings R0267 and R0268 were attributed to the category backcrossing with *E. imbricata* with a posterior probability of 99.8% (NewHybrids). This result is in accordance with Vilaça et al. (2012) which hypothesized that these hatchlings should have resulted of the crossing between one *Cc* x *Ei* F1 hybrid female with at least one *C.*

mydas male (evidenced by the R0264 and R0265) and another *E. imbricata* male (evidenced by the R0267 and R0268).

Population analyses

Intraspecific nuclear variation was analyzed for *C. caretta* and *E. imbricata*. A total of 4492 bp were sequenced from 14 nDNA markers for *E. imbricata* and 3592 bp were sequenced from 11 nDNA markers for *C. caretta* (Table 2). Different loci and sample numbers were surveyed for each species, making the total genetic diversity between species incomparable.

PCA of multilocus data was conducted to infer population structure assessing continuous axes of genetic variation of these species. First, we investigated nesting areas along the Brazilian coast for *C. caretta*. First axis explained only 33.8% of the total variation, while the second axis explained 29.8% of the variation, showing no relevant structure between Brazilian populations (Figure 3C). When individuals sampled in feeding areas were included in the analysis, the first principal component (PC1) explained 69.7% of the total variation and divided the samples into two clusters (Figure 3D). The first one corresponds to individuals from Brazilian rookeries and those captured in the feeding area in southern Brazil – Elevação do Rio Grande (ERG) – and present mtDNA haplotypes commonly found in Brazilian rookeries (CC-A4 and CC-A24). The second one corresponds to individuals from ERG that present mitochondrial haplotypes (CC-A11, CC-A2, CC-A33 and CC-A34) found in rookeries of the Caribbean, Mediterranean and Indo-Pacific Oceans. The second principal component (PC2) represents 29.7% of the variation.

A PCA was performed including individuals of *E. imbricata* sampled in Brazilian rookeries. PC1 and PC2 explained 46.12% and 34.87% of the total variation, showing no correlation between genetic variation and geographic distribution (Figure 3A). When sea turtles from feeding areas were included, PC1 and PC2 explained 49.77% and 40.44% of the total variation, respectively (Figure 3B). Five individuals sampled at feeding areas Fernando de Noronha and Atol das Rocas were slightly separated from other individuals. They presented mtDNA haplotypes either typically found in Indo-Pacific Ocean basin (EiIP16 and EiIP33) or 'orphan' haplotypes (EiA49, EiA75) which are differentiated from EiIP16 by one mutation step.

Population structure was analyzed comparing mtDNA and nDNA data of *C. caretta* (N=53) and *E. imbricata* (N=39) from the Brazilian coast. The 98 mtDNA haplotypes of *C. caretta* compilated from the literature were depicted in the network (Figure 4A), which showed three clades representing the main lineages of species (Shamblin et al. 2014). There was a large genetic divergence between mtDNA clades, a pattern not observed with multilocus data (Figure 4B). The neighborNet showed that some individuals from different mtDNA clades exhibited a close phylogenetic relationship when nDNA data was considered (highlighted with an ellipse in Figure 4B).

For *E. imbricata*, the relationship among 87 control region mtDNA haplotypes obtained from literature was depicted in a network shown in Figure 4C. They were clustered in seven main clades, two reported in the Atlantic Ocean and five in the Indo-Pacific region. The neighborNet built with multilocus data did not present large genetic distances between individuals from Atlantic and Indo-Pacific (Figure 4D). However, five of six individuals collected in foraging aggregations in northern Brazilian coast that belong to Indo-Pacific mtDNA clades were clustered in an end of the neighborNet (highlighted with an ellipse in Figure 4D), suggesting they come from another gene pool. Only individual R0242 from the Indo-Pacific mtDNA clade appeared more closely related to individuals that belong to the Atlantic mtDNA clade.

Phylogenetic analysis

The MCC tree obtained with multilocus data (Figure 5) showed the topology and dating congruent with previous phylogenetic studies of sea turtles (Bowen et al. 1993; Naro-Maciel et al. 2008; Duchene et al. 2012). The estimation of the TMRCA for the five species of sea turtles was 112.4 mya. The divergence between Carettini and Chelonini tribe (Cheloniidae) was estimated to have occurred at 65.9 mya. *Eretmochelys imbricata* separated from *C. caretta* and *L. olivacea* at 25 mya, followed by the split between *C. caretta* and *L. olivacea* at 21.6 mya.

The divergence between Atlantic and Indo-Pacific lineages of *E. imbricata* was estimated to have occurred at 5.93 mya, approximately the same date estimated using control region haplotypes (Vargas et al. 2016). Monophyly of clades based on mtDNA of *E. imbricata* (LeRoux et al. 2012; Vargas et al. 2016) was supported with nDNA in the Bayesian analysis using BEAST, except for one individual (R0242). This sea turtle belongs to Indo-Pacific clade

II according to mtDNA, but its nuclear composition showed that it is more similar to individuals from the Atlantic mtDNA clade. However, R0242 belongs to a lineage representing an early divergence in the Atlantic mtDNA clade, despite the low clade Bayesian posterior probability (0.41).

The earliest divergence between *C. caretta* lineages was estimated to have occurred 4.29 mya, similar to the date estimated using mitogenomes (Duchene et al. 2012) and control region haplotypes (Shamblin et al. 2014). One nDNA lineage gathers nesting and foraging individuals from Brazil that presents mtDNA haplotypes derived from CC-A4 and the other nDNA lineage presents individuals foraging in ERG that belong to three different mtDNA clades (IA, IB and II). Thus, the mtDNA based clades were only partially recovered with nuclear multilocus data, since individuals from different mtDNA clades were grouped in a single nDNA lineage.

Discussion

The interspecific hybridization phenomenon along the Brazilian coast

The use of a multilocus approach resulted in a powerful dataset based on haplotypes that allowed expanding our comprehension about the hybridization process of sea turtles. In this study, we re-analyzed 15 of 17 individuals previously identified as introgressed (>F1) hybrids by Vilaça et al. (2012), but we have only confirmed six backcrossed individuals, all of them were hatchlings. Seven F1 hybrids detected with our data were previously identified as introgressed (>F1) with information of a few nuclear markers, microsatellites or RFLP (Supplementary Table S1). Markers based on allele size differences, particularly microsatellites, may be not related to identity by descent due to their high level of homoplasy and other genotyping artifacts such as null alleles and allele dropouts (Zhang and Hewitt 2003). Here we used nDNA multilocus resequencing to characterize high quality haplotypes that supply a much higher level of resolution (Schlötterer 2004) at both inter and intraspecific analyses. The use of genetic markers randomly distributed throughout the genome, generated by high quality Sanger sequencing data, provided the highest genotyping accuracy with low ascertainment bias. Indeed, our multilocus dataset displayed a higher power to distinguish different hybrid crossings and introgression levels as compared to the previous methods.

Even though the initial NGS screening of variable locus was done by a ddRADseq approach with only two species (*C. caretta* and *E. imbricata*), we were able to validate informative nDNA loci with diagnostic alleles/haplotypes even for species displaying close phylogenetic relationship as *L. olivacea* and *C. caretta*. Considering the number of individuals analyzed and the number of diagnostic sites, we suggest the use of loci 856, 3061, 76958 and 109472 to characterize *Cc* x *Ei* hybrids, loci 421, 3061 and 109472 to characterize *Ei* x *Lo* hybrids, and loci 421, 966, 67959 and 114650 to characterize *Cc* x *Lo* hybrids (Supplementary Table S2). Future genetic studies investigating the hybridization between different species of sea turtles should be able to select more informative loci according to their target species.

According to Vilaça et al. (2012), there are introgressed (at least F2 hybrids) adult females nesting in Bahia (Brazil), and the first interspecific crossing could have occurred at two generations ago or a minimum of 40 years. In contrast, our data suggest that only hatchlings (newborns) were confirmed as introgressed hybrids. Considering the age at maturity from 20 to 40 years for *E. imbricata* (Meylan and Donnelly 1999) and from 22 to 29 years for *C. caretta* (Heppel 1998; Casale et al. 2011), we estimate that the minimum time for the first hybridization event was one generation ago (at least 20 years). Since the first hybrid female analyzed in this work was sampled in 2000 at Bahia, our data suggest that the high frequency hybridization event in Bahia may have started around 1980. This is supported by Conceição et al. (1990), which in 1989 first recorded hybrid juveniles in the state of Bahia. Bass et al. (1996) also support this hypothesis since they first reported, in 1992, a high incidence of *Cc* x *Ei* hybrid hatchlings (10 of 14 individuals) of females morphologically identified as *E. imbricata* at Praia do Forte, Bahia. This indicates that introgression was likely overestimated by Vilaça et al. (2012) and hybridization may be a more recent phenomenon happening in Brazil. However, another possible hypothesis is that the hybridization may be a recurrent event, but the introgressed hybrids (F2) are much less fertile or inviable.

Studies have reported that the reproductive output of hybrids and parental species in Bahia is similar (Soares et al. 2017) and the viability of hybrid and non-hybrid hatchlings is comparable (Soares et al. 2018). However, they only investigated F1 hybrids and their hatchlings. There is no information about the potential effects of hybridization in other life stages at the sea, as survivorship, growth rates and mating success. Indeed, if all (or the large majority) hybrid adults are first generation hybrids as our results indicate, thus a most likely explanation is that outbreeding depression (decrease of survival and/or reproductive fitness)

may occur mostly in the second and further generations of introgressed individuals. In this situation, the original parental gene combinations are broken up by recombination in >F1 hybrids, disrupting coadapted gene complex (Edmands et al. 1999; Goldberg et al. 2005).

The emergence of high-frequency hybridization event in Brazil coincides with the period of great population decline of sea turtles. The reduced chances of potential conspecific encounters may be associated with this unique event in the Brazilian coast (Vilaça et al. 2012). Reports of hybridization cases associated with human impact are increasing worldwide for other species (Allendorf et al. 2001). Human activities may lead to secondary contact between previously isolated populations also due to habitat disturbance and environmental changes that increase the hybrids rate (Todesco et al. 2016). Since 1988, sea turtle conservation in Brazil mostly relies on efforts of Projeto TAMAR, a consolidated and successful program aiming at environmental education and monitoring and research of sea turtles. Thereafter, the number of nesting females in monitored beaches has been increasing quickly (Marcovaldi and Chaloupka 2007), but in spite of this greater number of individuals, more recent hybridization events have been reported. A study of 2012 and 2013 nesting seasons showed that the incidence of hybridization in Bahia inferred from hatchlings of *C. caretta* females is 16.66% and for *E. imbricata* females is 8.15% (Soares et al. 2018).

Hybridization in Brazil is a local event with reports of fertile female hybrids in about 300 km of coastline between northern Bahia and Sergipe states. In this work, all female hybrids were sampled in rookeries of Bahia and Sergipe and pelagic individuals were sampled in coastal waters of Ceará, Bahia, Sergipe and São Paulo states. Other reports of hybrids in Brazil are juveniles from the states of Ceará and Rio Grande do Sul (Cassino Beach), which are two important feeding aggregations of *C. caretta* (Proietti et al. 2014). Further studies focusing on hybrids identification is recommended, mainly in nesting areas worldwide with the overlapping distribution of different sea turtle species.

We confirmed that all the *Cc* x *Ei* hybrids resulted from the crossing between *C. caretta* female and *E. imbricata* male, which indicate gender bias. This is probably associated with the prevalence of *C. caretta* along the Brazilian coast and the timing of reproductive seasons (Vilaça et al. 2012). *Eretmochelys imbricata* season begins around the *C. caretta* nesting peak (November and December), when the *E. imbricata* males encounter a higher number of *C. caretta* females to mate (Proietti et al. 2014). Conversely, the encounter between *C. caretta*

males and *E. imbricata* females happens less frequently, since *C. caretta* males leave the mating areas before a large number of *E. imbricata* females arrive (Vilaça et al. 2012).

Sea turtles present long and complex life cycles and monitoring the consequences of hybridization can be complicated, but it is extremely important to understand their impact on the management of sea turtle populations, particularly for parental species. Particular focus should be directed to hatchlings of F1 female hybrids to allow monitoring the future consequences of hybridization. We showed that improving the genetic resolution of studies is possible to better understand this local phenomenon in Brazil. New genomic approaches should be also able to elucidate the relation between genomic introgression and species-specific adaptive regions of the genome, in relation to their environment and behavior. Thus, we strongly recommend further studies to expand our comprehension of this particular evolutionary process of potential conservation impact.

Intraspecific studies of *C. caretta* and *E. imbricata*

Despite the lower mutation rate observed in nDNA markers (Zhang and Hewitt 2003) and the slow evolutionary rate of sea turtles (FitzSimmons et al. 1995), we obtained informative markers for intraspecific analyses that were validated after Sanger resequencing in some individuals of each sea turtle species. The intraspecific nuclear variation found in *C. caretta* and *E. imbricata* analyzed here (Table 2) allowed us to infer their population structure and dynamics.

Unlike mtDNA, the nuclear loci isolated in this study showed that variation within both species is not correlated to the geographic distribution along the Brazilian coast (Figure 3A and 3C, Supplementary Figure S1). Previous studies using mtDNA data showed significant differences in allelic frequencies between southern and northern Brazilian stocks for *C. caretta* (Reis et al. 2010a; Shamblin et al. 2014) and *E. imbricata* (Vilaça et al. 2013). For *C. caretta*, three genetically distinct stocks based on mtDNA were recognized along the Brazilian coast: northern coast (Bahia and Sergipe), Espírito Santo and Rio de Janeiro (Shamblin et al. 2014). For *E. imbricata*, two different mtDNA stocks, although closely related, were reported: Bahia and Rio Grande do Norte (Vilaça et al. 2013).

Some discrepant results between mtDNA and nDNA analyses are expected, as these markers have different mutation rates, inheritance patterns and effective sizes. mtDNA follows

a maternal inheritance, exhibits faster evolution rate and displays ¼ of the effective population size when compared with nDNA, allowing to investigate more recent demographic events (Cabanne et al. 2008; Brito and Edwards 2009). In contrast, nDNA reveals aspects from the biparental ancestry and more ancient history. Assessing the genetic diversity of nuclear markers is important for understanding the contribution of females and males to demographic patterns of sea turtles. However, regarding the definition of management units for conservation concerns, mtDNA data should be primordially considered since it characterizes relatively independent rookeries established by female philopatric recruitment (Shamblin et al. 2014).

Furthermore, these results can reflect an important sea turtle behavior. Lower population structure found in nuclear markers relative to mtDNA has been previously attributed to male-mediated gene flow (Bowen et al. 2005). Male sea turtles have uncertain philopatry and probably display greater flexibility in their choice of mating areas (FitzSimmons et al. 1997). Similar patterns were found in previous studies using nDNA and are indicative of lack of male philopatry (FitzSimmons et al. 1997; Bowen et al. 2005; Carreras et al. 2011; Vilaça et al. 2013; Clusa et al 2018). Thus, the apparently discrepant results for mtDNA and multilocus data could also be explained by gene flow between rookeries mediated by males.

Considering feeding areas, the analysis of pelagic individuals was important to provide a better understanding of the genetic diversity of turtles, and to investigate migration routes and the connectivity of distant rookeries. The Elevação do Rio Grande (ERG) area is located off the southern coast of Brazil and is an important feeding aggregation of *C. caretta*. Confirming previous studies (Reis et al. 2010a; Shamblin et al. 2014), this area seems to be visited by individuals from nesting sites in Brazil, northern Atlantic Ocean and Indo-Pacific Ocean, which is an evidence for transoceanic migrations for the species. Origin inference of the pelagic individuals was made based on the information of mtDNA, once individuals present typical haplotypes of specific rookeries. Two clusters were identified in the PCA of multilocus data, separating individuals of Brazilian rookeries, which present an exclusive mtDNA haplogroup CC-A4, from individuals originated in other continental rookeries (Figure 3C). The same clustering was observed for *C. caretta* in STRUCTURE analysis when correlated allele frequencies were used (Figure 2 and Supplementary Figure S1). In contrast, using independent allele frequencies model, this separation was not observed (Supplementary Figure S2). The correlated frequencies model provides greater power to identify distinct but closely related populations with recent shared ancestry (Porras-Hurtado et al. 2013).

For *E. imbricata*, all individuals were attributed to a single population in STRUCTURE analysis (Figure 2), even when the analysis was performed including only *E. imbricata* individuals and using correlated allele frequencies model (Supplementary Figure S1). However, some individuals from northern Brazilian feeding areas that exhibit mtDNA typically found in Indo-Pacific rookeries were slightly separated in PCA (Figure 3D). The presence of individuals from distant rookeries, indicated in this work by nDNA, confirms the occurrence of transoceanic migrations for this species. The high genetic diversity observed in the aggregations of Fernando de Noronha and Atol das Rocas is attributed to the influence of two great currents, the North Brazilian Current and the South Equatorial Current (Vilaça et al. 2013).

Population structure was analyzed comparing mtDNA and nDNA data. Considering great part of mtDNA haplotype diversity reported in the literature for *C. caretta*, it is possible to distinguish three clades with great genetic divergence. They correspond to two major lineages (clades I and II) of which the former passed by a more recent split (subclades IA and IB, Figure 4A). In this work, subclade IA is represented by individuals with haplotypes CC-A33 and CC-A34, which were registered in Australian rookeries. Subclade IB is represented by individuals with mtDNA haplotypes derived from CC-A4, which were recorded only in Brazilian rookeries, CC-A1.3, reported in Florida-USA, Mexico and Cape Verde, and CC-A11.6, reported in Oman (Indian Ocean) (Shamblin et al. 2014). Clade II is represented by haplotype CC-A2.1, recorded in rookeries from South Africa, Northwest Atlantic and Mediterranean, and display a star shaped network, which suggests that the group has experienced a population expansion.

Considering multilocus data, neighborNet analysis showed that nine individuals of *C. caretta* presented greater genetic divergence in relation to Brazilian individuals (highlighted with an ellipse in Figure 4B). These individuals were collected in the southern Brazilian feeding area (ERG) and belong to three different mtDNA clades (clades IA, IB and II). They also appear separately clustered in PCA (Figure 3D) and STRUCTURE (Supplementary Figure S1). Phylogenetic analysis also resulted in a MCC tree with two main lineages, of which one corresponds to a mix of individuals from three mtDNA clades (Figure 5).

Phylogeographic studies suggested that the two main mtDNA clades I and II of *C. caretta* were isolated by geographic and climatic factors into Atlantic and Indo-Pacific basins during the cooler periods of the Pleistocene (Bowen et al. 1994). As *C. caretta* is also adapted to temperate water, migrations via southern Africa, directed by the waters of the Agulhas Current, are possible. The phylogeographic scenario proposed for Shamblin et al. (2014)

suggests that mtDNA clade IA had an Indo-Pacific origin, where the earliest diverging lineages of *C. caretta* appear. The first colonization was likely from Indo-Pacific lineages invading the Atlantic Ocean. Brazilian haplotypes (CC-A4 and derived) seem to be the earliest diverging lineage within mtDNA clade IB. It was followed by a more recent colonization of the CC-A11.6 precursor from Atlantic to Indian Ocean, as it is closely related to Atlantic lineages. Therefore, transoceanic migration in both directions may be responsible for the gene flow between *C. caretta* populations. Furthermore, current geographic distribution of these lineages presents no phylogenetic concordance, as both lineages are found in both Atlantic-Mediterranean and Indo-Pacific basins (Reis et al. 2010a; Duchene et al. 2012).

Despite the small number of samples, this result can suggest a homogenization of *C. caretta* populations at a nuclear level for individuals sharing a common feeding area. This was previously reported for another *C. caretta* population and attributed to male-mediated gene flow (Bowen et al. 2005).

For *E. imbricata*, the relationship among mtDNA haplotypes previously reported revealed that there are seven main clades worldwide. Two of them were registered in rookeries from the Atlantic Ocean (LeRoux et al. 2012) and five in rookeries from the Indo-Pacific Ocean (Vargas et al. 2016). The neighborNet of nDNA data showed that 5 of 6 individuals of Indo-Pacific mtDNA clades are slightly more distant from individuals that belong to Atlantic mtDNA clades (Figure 4D). The same individuals belong to Indo-Pacific nDNA cluster according to the phylogenetic analyses (Figure 5). They were sampled in the Brazilian feeding aggregations, demonstrating long distance migrations for the species.

Despite the separation between *E. imbricata* individuals from Indo-Pacific and Atlantic was not observed in STRUCTURE, it was slightly observed in PCA and neighborNet, and strongly detected in the MCC tree. There is a deep genetic divergence between Indo-Pacific and Atlantic mtDNA lineages of *E. imbricata* dating from the early Pliocene, when the closing of the Isthmus of Panama occurs (Duchene et al. 2012). The geographic pattern of separation between ocean basins found with mtDNA was recovered with nuclear data, except for one individual (R0242). However, this individual is basal to Atlantic mtDNA clade.

Eretmochelys imbricata is strictly adapted to tropical waters and although some transoceanic migrations may occur, American and African continents are supposedly important barriers to species migration directing the current distribution of main lineages of sea turtles (Duchene et al. 2012). In contrast, *C. caretta* individuals were more divergent within the

Atlantic than between the Atlantic and Indo-Pacific, probably due to a transoceanic gene flow observed in this species more adapted to temperate water. The barriers to gene flow are not the same for all species of sea turtles likely due to their different ability of dispersion through the oceans and evolutionary responses to environmental changes (Duchene et al. 2012).

Regarding intraspecific phylogenetic analysis, the use of multilocus data resulted in similar topology and divergence times between species when compared to the previous studies that used mtDNA data (Duchene et al. 2012; LeRoux et al. 2012; Shamblin et al. 2014; Vargas et al. 2016). The divergences between main lineages within *E. imbricata* and *C. caretta* were estimated to have occurred about 5.93 mya and 4.29 mya, respectively. It is consistent with the age of formation of the Isthmus of Panama, associated to the deepest phylogenetic split of intraspecific lineages of different species of sea turtles (Naro-Maciel et al. 2008; Duchene et al. 2012). Besides, assessing multilocus data was possible to evaluate biparental ancestry, as well as to accommodate the stochasticity of the coalescent process combining information from multiple loci, instead of relying on inferences based on individual tree topologies (Edwards and Beerli 2000; Brito and Edwards 2009).

Concluding remarks

Next Generation Sequencing technologies allowed the initial identification of variable genome wide loci, which were selected for a Sanger sequencing validation step to characterize multilocus datasets useful for inter and/or intraspecific studies. The high quality multilocus data provided significant interspecific information for the inference of the phylogeny of sea turtles and characterization of hybrids. Besides, another multilocus dataset provided relevant intraspecific data for analyses of population dynamics, structure and demography. The presented results reveal important enhancements in the genetic resolution of the hybridization process and population dynamics of sea turtles.

Acknowledgments

We are grateful for the access of collection samples from Projeto TAMAR, a conservation program of the Brazilian Ministry of the Environment, affiliated to ICMBio (the Chico Mendes Institute for Conservation of Biodiversity), and is co-managed by Fundação Pró-

TAMAR and sponsored by Petrobras-Brazil. We are grateful for all members of Fundação Pró-TAMAR, particularly to M.A. Marcovaldi and L.S. Soares for their logistical support and and availability of samples. We also thank Gisele Lobo-Hajdu from UERJ (Universidade Estadual do Rio de Janeiro) and Sarah M. Vargas (Universidade Federal do Espírito Santos) for sharing of previously collected samples. We thank Christina Brown for helping with experiments and primer design. There are no samples collected for this particular study.

References

Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001. The problems with hybrids: setting conservation guidelines. *Trends Ecol Evol.* 16:613–622.

Anderson EC. 2008. Bayesian inference of species hybrids using multilocus dominant genetic markers. *Philos Trans R Soc B.* 363:2841–2850.

Anderson EC, Thompson EA. 2002. A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics.* 160:1217–1229.

Bandelt H-J, Forster P, Sykes BC, Richards MB. 1995. Mitochondrial Portraits of Human Populations Using Median Networks. *Genetics.* 141:743–753.

Bass AL, Good DA, Bjorndal KA. 1996. Testing models of female reproductive migratory behaviour and population structure in the Caribbean hawksbill turtle, *Eretmochelys imbricata*, with rntDNA sequences. *Mol Ecol.* 5:321–328.

Bjorndal KA, Bolten AB, Moreira L, Bellini C, Marcovaldi MA. 2006. Population Structure and Diversity of Brazilian Green Turtle Rookeries Based on Mitochondrial DNA Sequences. *Chelonian Conserv Biol.* 5:262–268.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol.* 10:e1003537.

Bowen BW, Bass AL, Soares L, Toonen RJ. 2005. Conservation implications of complex population structure: lessons from the loggerhead turtle (*Caretta caretta*). *Mol Ecol.* 14:2389–2402.

Bowen BW, Kamezaki N, Limpus CJ, Hughes GR, Meylan AB, Avise JC. 1994. Global Phylogeography of the Loggerhead Turtle (*Caretta caretta*) as Indicated by Mitochondrial DNA Haplotypes. *Evolution.* 48:1820.

Bowen BW, Karl SA. 2007. Population genetics and phylogeography of sea turtles. *Mol Ecol.* 16:4886–4907.

Bowen BW, Nelson WS, Avise JC. 1993. A molecular phylogeny for marine turtles: Trait mapping, rate assessment, and conservation relevance. *Proc Natl Acad Sci.* 90:5574–5577.

Briscoe DK, Parker DM, Balazs GH, Kurita M, Saito T, Okamoto H, Rice M, Polovina JJ, Crowder LB. 2016. Active dispersal in loggerhead sea turtles (*Caretta caretta*) during the 'lost years.' *Proc R Soc B Biol Sci.* 283:20160690.

Brito PH, Edwards SV. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica.* 135:439–455.

Cabanne G, d'Horta F, Sari E, Santos F, Miyaki C. 2008. Nuclear and mitochondrial phylogeography of the Atlantic forest endemic *Xiphorhynchus fuscus* (Aves:

Dendrocolaptidae): Biogeography and systematics implications. *Mol Phylogenet Evol.* 49:760–773.
Cadena EA, Parham JF. 2015. Oldest known marine turtle? A new protostegid from the Lower Cretaceous of Colombia. *PaleoBios.* 32:1–42.
Carreras C, Pascual M, Cardona L, Marco A, Bellido JJ, Castillo JJ, Tomás J, Raga JA, Sanfélix M, Fernández G, *et al*. 2011. Living Together but Remaining Apart: Atlantic and Mediterranean Loggerhead Sea Turtles (*Caretta caretta*) in Shared Feeding Grounds. *J Hered.* 102:666–677.
Casale P, Mazaris A, Freggi D. 2011. Estimation of age at maturity of loggerhead sea turtles *Caretta caretta* in the Mediterranean using length-frequency data. *Endanger Species Res.* 13:123–129.
Conceição MB, Levy JA, Marins LF, Marcovaldi MA. 1990. Electrophoretic characterization of a hybrid between *Eretmochelys imbricata* and *Caretta caretta* (Cheloniidae). *Comp Biochem Physiol Part B.* 97:275–278.
Clusa M, Carreras C, Cardona L, Demetropoulos A, Margaritoulis D, Rees A, Hamza A, Khalil M, Levy Y, Turkozan O, *et al*. 2018. Philopatry in loggerhead turtles *Caretta caretta*: beyond the gender paradigm. *Mar Ecol Prog Ser.* 588:201–213.
Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet.* 12:499–510.
Dmitriev DA, Rakitov RA. 2008. Decoding of Superimposed Traces Produced by Direct Sequencing of Heterozygous Indels. *PLoS Comput Biol.* 4:e1000113.
Drummond AJ, Bouckaert RR. 2014. *Bayesian evolutionary analysis with BEAST 2*. Cambridge University Press.
Duchene S, Frey A, Alfaro-Núñez A, Dutton PH, Thomas P Gilbert M, Morin PA. 2012. Marine turtle mitogenome phylogenetics and evolution. *Mol Phylogenet Evol.* 65:241–250.
Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* 4:359–361.
Eaton DAR. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics.* 30:1844–1849.
Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.
Edmands S. 1999. Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution.* 53:1757-1768.
Edwards SV, Beerli P. 2000. Perspective: gene divergence, population divergence, and the variance incoalescence time in phylogeographic studies. *Evolution.* 54:1839.
Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol.* 14:2611–2620.
Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 10:564–567.
Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics.* 164:1567–1587.
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.* 94:8912–8917.

FitzSimmons NN, 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 GeneFlow. *Genetics*. 147:1843–1854.

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.

Flot J-F. 2010. Seqphase: a web tool for interconverting phase input/output files and fasta sequence alignments: computer program note. *Mol Ecol Resour.* 10:162–166.

Goldberg TL, Grant EC, Inendino KR, Kassler TW, Claussen JE, Philipp DP. 2005. Increased Infectious Disease Susceptibility Resulting from Outbreeding Depression. *Conserv Biol.* 19:455–462.

González-Garza BI, Stow A, Sánchez-Teyer LF, Zapata-Pérez O. 2015. Genetic variation, multiple paternity, and measures of reproductive success in the critically endangered hawksbill turtle (*Eretmochelys imbricata*). *Ecol Evol.* 5:5758–5769.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P. 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. *Evol Appl.* 7:1008–1025.

Hedrick PW. 2013. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol.* 22:4606–4618.

Heppel SS. 1998. Application of life-history theory and population model analysis to turtle conservation. *Copeia*. 2:367–375.

Hirayama R. 1998. Oldest known sea turtle. *Nature.* 392: 705-708.

Huson DH, Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Mol Biol Evol.* 23:254–267.

Joly S, Bruneau A. 2006. Incorporating Allelic Variation for Reconstructing the Evolutionary History of Organisms from Multiple Genes: An Example from Rosa in North America. *Syst Biol.* 55:623–636.

Joly S, Bryant D, Lockhart PJ. 2015. Flexible methods for estimating genetic distances from single nucleotide polymorphisms. *Methods Ecol Evol.* 6:938–948.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics.* 27:3070–3071.

Karl SA, Bowen BW, Avise JC. 1995. Hybridization Among the Ancient Mariners: Characterization of Marine Turtle Hybrids With Molecular Genetic Assays. *J Hered.* 86:262–268.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, *et al*. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28:1647–1649.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak : a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour.* 15:1179–1191.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 33:1870–1874.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Mol Biol Evol*. 34:772-773.

Lara-Ruiz P, Lopez GG, Santos FR, Soares LS. 2006. Extensive hybridization in hawksbill turtles (*Eretmochelys imbricata*) nesting in Brazil revealed by mtDNA analyses. *Conserv Genet.* 7:773–781.

Leroux RA, Dutton PH, Abreu-Grobois FA, Lagueux CJ, Campbell CL, Delcroix E, Chevalier J, Horrocks JA, Hillis-Starr Z, Troëng S, *et al*. 2012. Re-examination of Population Structure and Phylogeography of Hawksbill Turtles in the Wider Caribbean Using Longer mtDNA Sequences. *J Hered.* 103:806–820.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 25:1451–1452.

Maheshwari S, Barbash DA. 2011. The Genetics of Hybrid Incompatibilities. *Annu Rev Genet.* 45:331–355.

Marcovaldi M, Chaloupka M. 2007. Conservation status of the loggerhead sea turtle in Brazil: an encouraging outlook. *Endanger Species Res.* 3:133–143.

Meylan AB, Donnelly M. 1999. Status Justification for Listing the Hawksbill Turtle (*Eretmochelys imbricata*) as Critically Endangered on the 1996 IUCN Red List of Threatened Animals. *Chelonian Conserv Biol.* 3:25.

Molfetti É, Vilaça ST, Georges J-Y, Plot V, Delcroix E, Le Scao R, Lavergne A, Barrioz S, Santos FR, de Thoisy B. 2013. Recent Demographic History and Present Fine-Scale Structure in the Northwest Atlantic Leatherback (*Dermochelys coriacea*) Turtle Population. *PLoS ONE.* 8:e58061.

Monzón-Argüello C, Loureiro NS, Delgado C, Marco A, Lopes JM, Gomes MG, Abreu-Grobois FA. 2011. Príncipe island hawksbills: Genetic isolation of an eastern Atlantic stock. *J Exp Mar Biol Ecol*. 407:345–354.

Moody RTJ. 1974. The Taxonomy and Morphology of *Puppigerus camperi* (Gray), an Eocene Sea Turtle from Northern Europe. *Geology*. 25:153–186.

Moore MK, Ball RM. 2002. Multiple paternity in loggerhead turtle (*Caretta caretta*) nests on Melbourne Beach, Florida: a microsatellite analysis. *Mol Ecol.* 11:281–288.

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.* 49:659–662.

Nishizawa H, Narazaki T, Fukuoka T, Sato K, Hamabata T, Kinoshita M, Arai N. 2014. Genetic composition of loggerhead turtle feeding aggregations: migration patterns in the North Pacific. *Endanger Species Res.* 24:85–93.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE.* 7:e37135.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo Á, Lareu MV. 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. *Front Genet.* 4.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics.* 155:945–959.

Proietti MC, Lara-Ruiz P, Reisser JW, Pinto L da S, Dellagostin OA, Marins LF. 2009. Green turtles (*Chelonia mydas*) foraging at Arvoredo Island in Southern Brazil: Genetic characterization and mixed stock analysis through mtDNA control region haplotypes. *Genet Mol Biol.* 32:613–618.

Proietti MC, Reisser J, Marins LF, Marcovaldi MA, Soares LS, Monteiro DS, Wijeratne S, Pattiaratchi C, Secchi ER. 2014a. Hawksbill × loggerhead sea turtle hybrids at Bahia, Brazil: where do their offspring go? *PeerJ.* 2:e255.

Proietti MC, Reisser J, Marins LF, Rodriguez-Zarate C, Marcovaldi MA, Monteiro DS, Pattiaratchi C, Secchi ER. 2014b. Genetic Structure and Natal Origins of Immature Hawksbill Turtles (*Eretmochelys imbricata*) in Brazilian Waters. *PLoS ONE.* 9:e88746.

Putman NF, Mansfield KL. 2015. Direct Evidence of Swimming Demonstrates Active Dispersal in the Sea Turtle "Lost Years." *Curr Biol.* 25:1221–1227.

Reis EC, Soares LS, Vargas SM, Santos FR, Young RJ, Bjorndal KA, Bolten AB, Lôbo-Hajdu G. 2010a. Genetic composition, population structure and phylogeography of the loggerhead sea turtle: colonization hypothesis for the Brazilian rookeries. *Conserv Genet.* 11:1467–1477.

Reis EC, Soares LS, Lôbo-Hajdu G. 2010b. Evidence of olive ridley mitochondrial genome introgression into loggerhead turtle rookeries of Sergipe, Brazil. *Conserv Genet.* 11:1587–1591.

Schlötterer C. 2004. The evolution of molecular markers—just a matter of fashion? *Nat Rev Genet.* 5:63–69.

Shamblin BM, Bolten AB, Abreu-Grobois FA, Bjorndal KA, Cardona L, Carreras C, Clusa M, Monzón-Argüello C, Nairn CJ, Nielsen JT. 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:e85956.

Stephens M, Smith NJ, Donnelly P. 2001. A New Statistical Method for Haplotype Reconstruction from Population Data. *Am J Hum Genet.* 68:978–989.

Stewart KR, Dutton PH. 2011. Paternal genotype reconstruction reveals multiple paternity and sex ratios in a breeding population of leatherback turtles (*Dermochelys coriacea*). *Conserv Genet.* 12:1101–1113.

Todesco M, Pascual MA, Owens GL, Ostevik KL, Moyers BT, Hübner S, Heredia SM, Hahn MA, Caseys C, Bock DG, *et al*. 2016. Hybridization and extinction. *Evol Appl.* 9:892–908.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40:e115–e115.

Vargas SM, Araújo FCF, Monteiro DS, Estima SC, Almeida AP, Soares LS, Santos FR. 2008. Genetic Diversity and Origin of Leatherback Turtles (*Dermochelys coriacea*) from the Brazilian Coast. *J Hered.* 99:215–220.

Vargas SM, Jensen MP, Ho SYW, Mobaraki A, Broderick D, Mortimer JA, Whiting SD, Miller J, Prince RIT, Bell IP, *et al*. 2016. Phylogeography, Genetic Diversity, and Management Units of Hawksbill Turtles in the Indo-Pacific. *J Hered.* 107:199–213.

Vilaça ST, Lara-Ruiz P, Marcovaldi MA, Soares LS, Santos FR. 2013. Population origin and historical demography in hawksbill (*Eretmochelys imbricata*) feeding and nesting aggregates from Brazil. *J Exp Mar Biol Ecol.* 446:334–344.

Vilaça ST, Vargas SM, Lara-Ruiz P, Molfetti É, Reis EC, LôBo-Hajdu G, Soares LS, Santos FR. 2012. Nuclear markers reveal a complex introgression pattern among marine turtle species on the Brazilian coast. *Mol Ecol.* 21:4300–4312.

Wallace BP, DiMatteo AD, Bolten AB, Chaloupka MY, Hutchinson BJ, Abreu-Grobois FA, Mortimer JA, Seminoff JA, Amorocho D, Bjorndal KA, *et al*. 2011. Global Conservation Priorities for Marine Turtles. *PLoS ONE.* 6:e24510.

Wang Z, Pascual-Anaya J, Zadissa A, Li W, Niimura Y, Huang Z, Li C, White S, Xiong Z, Fang D, *et al*. 2013. The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nat Genet.* 45:701–706.

Wringe BF, Stanley REE, Jeffery NW, Anderson EC, Bradbury IR. 2017. hybriddetective: A workflow and package to facilitate the detection of hybridization using genomic data in R. *Mol Ecol Resour.* 17:e275–e284.

Zhang D-X, Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Mol Ecol.* 12:563–584.
Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics.* 30:614–620.

Table 1. Sampling localities of sea turtles and hybrids and number of individuals per locality (N). *C. caretta* x *E. imbricata* (*Cc* x *Ei*), *E. imbricata* x *L. olivacea* (*Ei* x *Lo*), *C. caretta* x *L. olivacea* hybrids (*Cc* x *Lo*), *E. imbricata* x *C. caretta* x *C. mydas* (*Ei* x *Cc* x *Cm*).

Sea turtle species	**Sample Locality**		**N**
Loggerhead turtle (*Caretta caretta*)	Foraging Area	Elevação do Rio Grande	15
	Nesting Area	Bahia	7
		Rio Grande do Norte	1
		Sergipe	10
		Rio de Janeiro	10
		Espírito Santo	10
Hawksbill turtle (*Eretmochelys imbricata*)	Foraging Area	Fernando de Noronha	7
		Atol das Rocas	6
	Nesting Area	Bahia	13
		Rio de Janeiro	1
		Sergipe	1
		Rio Grande do Norte	11
Green turtle (*Chelonia mydas*)	Foraging Area	Fernando de Noronha	1
		Ilha do Arvoredo	1
Olive ridley turtle (*Lepidochelys olivacea*)	Nesting Area	Sergipe	7
	Foraging Area	Sergipe	2
Leatherback turtle (*Dermochelys coriacea*)	Foraging Area	Ceará	1
		Pesca	1
Hybrid *Cc* x *Ei*	Nesting Area	Bahia	17
	Foraging Area	Ceará	2
		Atol das Rocas	1
		Sergipe	1
Hybrid *Ei* x *Lo*	Nesting Area	Bahia	2
Hybrid *Cc* x *Lo*	Foraging Area	São Paulo	1
	Nesting Area	Sergipe	10
Hybrid *Ei* x *Cc* x *Cm*	Nesting Area	Bahia	4

Table 2 – Genetic diversity of nuclear markers to *C. caretta* and *E. imbricata*. Number of individuals (N), number of haplotypes (H), number of polymorphic sites (S) and haplotype diversity (k).

		421	**856**	**966**	**3061**	**9672**	**23712**	**30573**	**31476**	**42006**	**46208**	**67959**	**76958**	**109472**	**114650**	**267557**
Eretmochelys imbricata	N	39	39	39	35	-	39	39	34	39	39	35	35	39	39	39
	H	3	2	2	2	-	3	5	2	3	3	2	4	4	3	5
	S	3	1	3	1	-	2	3	4	2	3	2	4	3	2	4
	k	0.48	0.46	0.41	0.11	-	0.39	0.71	0.47	0.21	0.54	0.32	0.21	0.6	0.49	0.3
Caretta caretta	N	53	53	53	-	49	53	53	-	53	53	-	-	52	53	47
	H	2	3	6	-	4	3	2	-	5	4	-	-	4	3	7
	S	1	3	5	-	3	2	1	-	5	3	-	-	3	2	6
	k	0.17	0.12	0.47	-	0.44	0.07	0.05	-	0.67	0.23	-	-	0.54	0.51	0.59

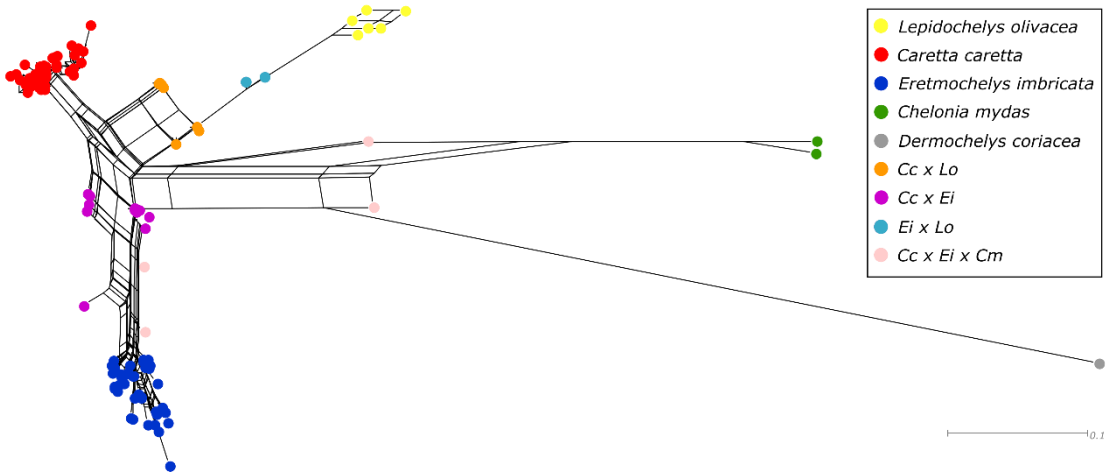

Figure 1. NeighborNet of organisms based on multilocus nuclear data for 5 species of sea turtles that occur along the Brazilian coast and hybrid individuals. The hybrids are observed intermediary between species involved in the hybridization process. Details of sampling (N=143) are described in Table 1. Tips of the neighborNet represent unique multilocus genotypes. Cc: *Caretta caretta*, Ei: *Eretmochelys imbricata*, Lo: *Lepidochelys olivacea*, Cm: *Chelonia mydas*.

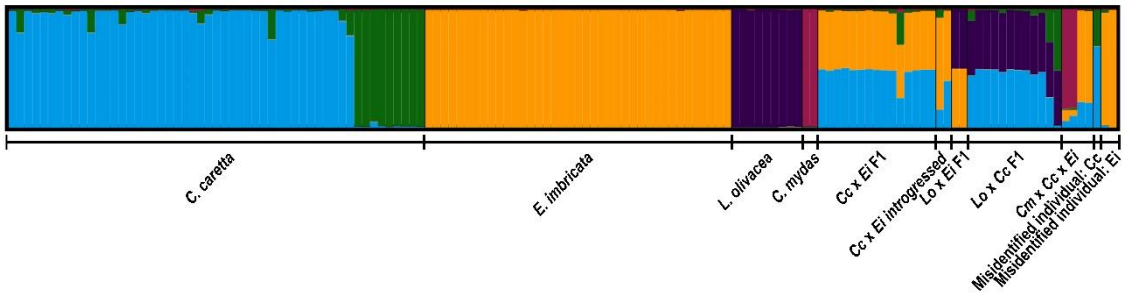

Figure 2. STRUCTURE bar plots representing K = 5 using correlated allele frequencies model. The x-axis represents each individual analyzed and the y-axis represents the estimated admixture proportions related to each parental species. The barplot was obtained with CLUMPAK. Cc: *Caretta caretta*, Ei: *Eretmochelys imbricata*, Lo: *Lepidochelys olivacea*, Cm: *Chelonia mydas*.

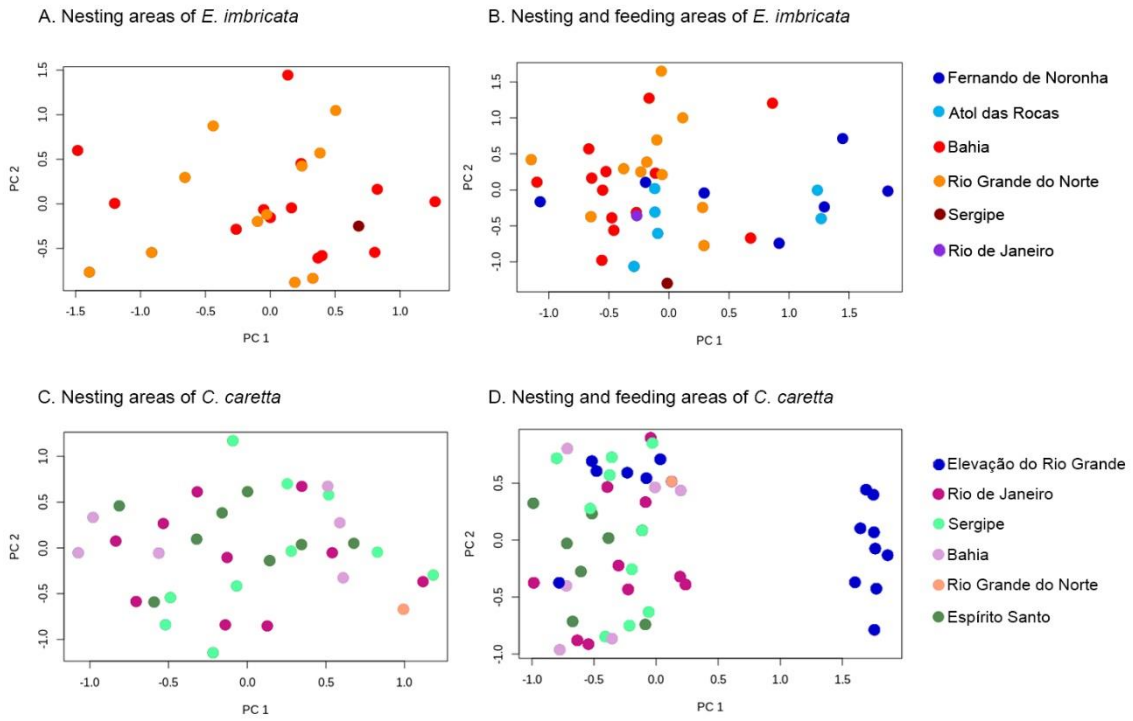

Figure 3. Principal component analysis of multilocus data for *C. caretta* and *E. imbricata* including only sea turtles sampled in Brazilian rookeries (A and C) and individuals collected in rookeries and feeding areas along the Brazilian coast (B and D). Color codes indicate the geographical location where the individuals were collected.

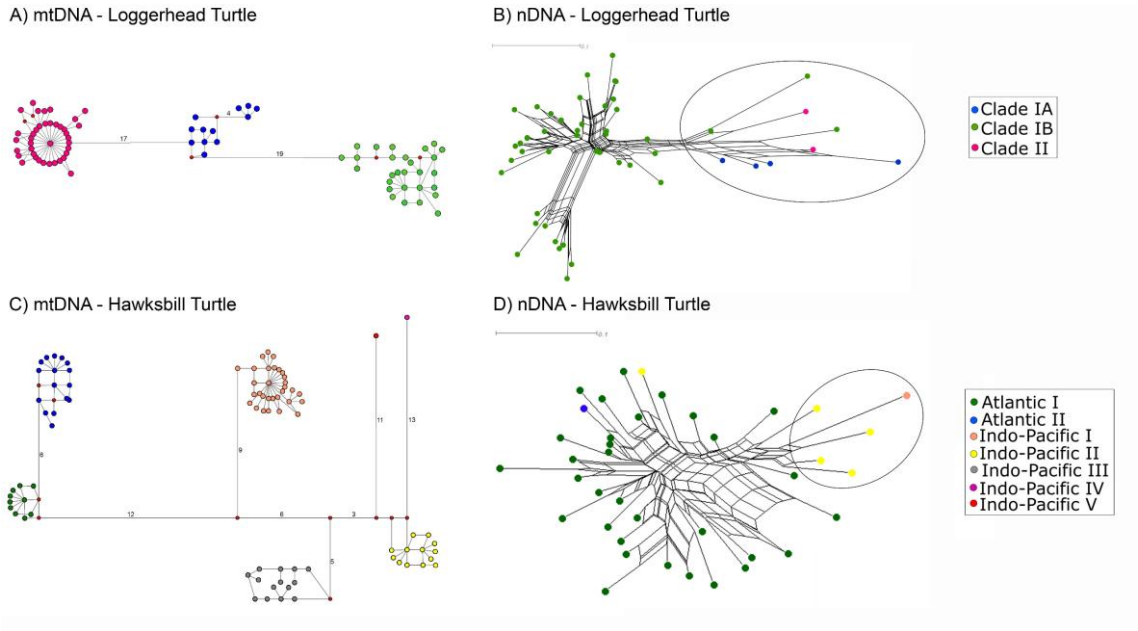

Figure 4. Haplotype network based on control region of the mtDNA data (A and C) and neighborNet of organisms based on multilocus nuclear data (B and D) for *C. caretta* and *E. imbricata*. The mitochondrial data was obtained from haplotypes based on control region previously published in literature. The nuclear data was obtained for 53 and 39 individuals of *C. caretta* and *E. imbricata*, respectively, using multilocus data. Tips of the neighborNet represent unique multilocus genotypes. The ellipses highlight the individuals of *C. caretta* more distantly related and supposed to have Indo-Pacific origin (B) and the individuals of *E. imbricata* that belong to the Indo-Pacific mtDNA clades and were grouped together (D).

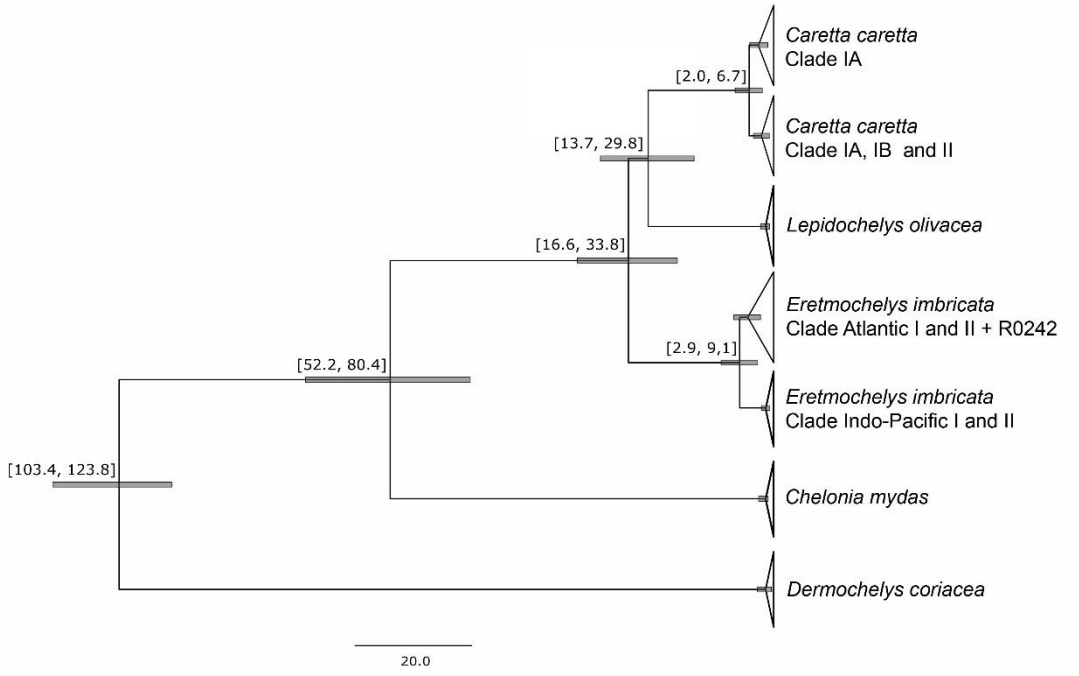

Figure 5. Bayesian chronogram of sea turtles from the Brazilian coast inferred from multilocus data. The horizontal axis indicates divergence times in million years before present. Horizontal bars and the numbers above branches correspond to the 95% highest posterior density (HPD) interval values estimated for all tree nodes with posterior probabilities above 0.8 calculated in BEAST. Clade names are based on mtDNA haplotypes as grouped by previous studies (LeRoux et al. 2012; Nishizawa et al. 2014; Shamblin et al. 2014; Vargas et al. 2016).

Supporting information

Table S1. Information of hybrid individuals: locality, morphology, collected individual, mtDNA haplotype based on control region, previous classification by Vilaça et al., 2012 and classification based on multilocus data investigated in this study.

Table S2. Primers, PCR specifications and number of diagnostic sites considering different hybrid crossings for each nuclear marker. Nuclear markers included in each dataset are specified with an X. Cc: *Caretta caretta*, Ei: *Eretmochelys. imbricata*, Lo: *Lepidochelys olivacea*, Cm: *Chelonia mydas*, Dc: *Dermochelys coriacea*.

Table S3. Mitochondrial haplotypes and its respective clades of each individual studied in population analyses. The clades were defined following the definition of Shamblin et al. (2014) for *Caretta caretta* and of Vargas et al. (2016) and LeRoux et al. (2012) for Eretmochelys imbricata. Legend: Rio de Janeiro (RJ), Sergipe (SE), Espírito Santo (ES), Rio Grande do Norte (RN), Fernando de Noronha (FN), Atol das Rocas (AR), Elevação do Rio Grande (ERG), data not available (NA).

Table S4. Best-fit partitioning model of molecular evolution of multilocus data.

Table S5. Resequencing of mtDNA control region, RAG1 and CMOS for the misidentified samples. The number correspond to the Genbank accession for the haplotypes.

Figure S1. STRUCTURE bar plots representing K = 2 using correlated allele frequencies model for *Eretmochelys imbricata* and *Caretta caretta* from different Brazilian populations. The x-axis represents each individual analyzed and the y-axis represents the estimated admixture proportions related to each population. This graphic was obtained with CLUMPAK.

Figure S2. STRUCTURE bar plots representing K = 5 using independent allele frequencies model. The x-axis represents each individual analyzed and the y-axis represents the estimated admixture proportions related to each parental species. This

graphic was obtained with CLUMPAK. Cc: *Caretta caretta*, Ei: *Eretmochelys imbricata*, Lo: *Lepidochelys olivacea*, Cm: *Chelonia mydas*.

Figure S3. NewHybrids analysis of 13 individuals previously identified as introgressed hybrids. Each vertical bar represents one individual and the y-axis represents its posterior probability of belonging to different classes: *C. caretta* (Pure 1), *E. imbricata* (Pure 2), F1 hybrid, F2 hybrid, backcross with *C. caretta* (BC1) or backcross with *E. imbricata* (BC2). This graphic was obtained with R package HybridDetective.

CAPÍTULO III

Genomic evidence of recent hybridization between sea turtles at Abrolhos Archipelago and its association to low reproductive output

Larissa Souza Arantes, Lucas Cabral Lage Ferreira, Maximillian Driller, Fernando Pedro Marinho Repinaldo Filho, Camila Junqueira Mazzoni, Fabrício Rodrigues Santos. Manuscrito submetido ao periódico Scientific Reports.

Genomic evidence of recent hybridization between sea turtles at Abrolhos Archipelago and its association to low reproductive output

Larissa Souza Arantes[1], Lucas Cabral Lage Ferreira[2], Maximillian Driller[3,4], Fernando Pedro Marinho Repinaldo Filho[2], Camila Junqueira Mazzoni[3,4], Fabrício Rodrigues Santos[1*]

[1]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Avenida Antônio Carlos, 6627, Belo Horizonte - MG, Brazil, 31270-010. Phone: +55 31 3409 2581.

[2]Parque Nacional Marinho dos Abrolhos, ICMBio, Ministério do Meio Ambiente, Brazil.

[3]Berlin Center for Genomics in Biodiversity Research (BeGenDiv), Königin-Luise-Straße 6-8, Berlin, Germany, D-14195. Phone: +49 (0)30 838 59961.

[4]Evolutionary Genetics Department, Leibniz Institute for Zoo and Wildlife Research (IZW), Alfred-Kowalke-Straße 17, Berlin, Germany, D-10315. Phone: +49 (0) 30 5168 0.

*Correspondence and requests for materials should be addressed to F.R.S (email: fsantos@icb.ufmg.br)

Abstract

Hybridization between sea turtle species is an important conservation concern, occurring particularly in high frequency in two hybrid zones in Brazil. We investigated the hypothesis of the existence of a new hybrid spot in the Abrolhos Archipelago, which is surrounded by the largest and richest coral reefs in the South Atlantic. We performed a multidisciplinary investigation of the sea turtles and their reproductive output in the Abrolhos beaches. Genetic data of mtDNA and six nuclear markers showed that there are first-generation hybrid females nesting in Abrolhos, resulted from the crossing between male hawksbill turtles (*Eretmochelys imbricata*) and female loggerhead turtles (*Caretta caretta*), and their progenies that are backcrossed with both parental species. The introgression level was characterized using genomic data obtained with 3RADseq method, which confirmed gene flow between F1 hybrids and loggerhead turtles. The reproductive output data of Abrolhos population strongly suggests a reproductive disadvantage of hybrids compared to loggerheads. Hatchling success for loggerhead

turtles in Abrolhos is also lower than other Brazilian populations, which may be associated with the Abrolhos' beach features and obstructions. We showed for the first time the association between hybridization process and low reproductive success (outbreeding depression), which can represent a threat to sea turtle conservation.

Introduction

The Abrolhos Bank is an expansion of the southern part of the Eastern Brazilian continental shelf that embraces the largest and richest coral reefs in the South Atlantic and the world's largest rhodoliths beds[1,2]. The area comprises mangroves, seagrass and algae bottoms, euphotic and mesophotic coral reefs with high levels of endemism, unique mushroom-shaped coralline pinnacles and a set of five volcanic islands that form the Abrolhos Archipelago[3,4]. There is a vast community of reef-associated species in Abrolhos, comprised of soft bodied sea anemones and zoanthids, sponges, polychaete worms, molluscs, crustaceans, echinoderms, sea birds, sea turtles and whales[5].

Regarding sea turtles, Abrolhos is an important feeding area for hawksbill (*Eretmochelys imbricata*) and green turtles (*Chelonia mydas*), attracted by the abundance and diversity of respectively invertebrates and seagrasses[6,7]. Furthermore, it is also a nesting area for loggerhead turtles (*Caretta caretta*), which lay their eggs on the short sandy beaches of some Abrolhos islands. These three sea turtle species are respectively listed in the IUCN Red List of Threatened Species as critically endangered, endangered and vulnerable to extinction[8]. Therefore, the Abrolhos region was defined as an important feeding area for sea turtles in the National Action Plan (PAN) for the Conservation of Sea Turtles[9].

The monitoring of juvenile sea turtles in Abrolhos reported behavioural patterns related to feeding, resting and cleaning symbiosis for hawksbill and green turtles[10]. Green turtles are mainly found in grounds rich in algae and seagrasses and hawksbill turtles are associated with shallow fringing reefs. Abrolhos is an important recruitment area for hawksbill turtles, as suggested by the high abundance of small turtles (mean curved carapace length = 37.9 cm)[7]. Proietti, et al.[11] performed the first genetic study of 65 hawksbill juveniles hand-captured by divers at Abrolhos. The estimated origin of the majority of these individuals was from Brazilian rookeries (Bahia, Sergipe and Rio Grande do Norte), with some contribution of Caribbean rookeries, likely associated with

the influence of Brazil Current. Some of the young tagged individuals were recently recaptured in Abrolhos, showing long period of residence in the archipelago[10].

The first nesting record of sea turtles in Abrolhos was in 1984, when a loggerhead turtle was observed nesting on the sandy beach of Redonda and Santa Bárbara Islands[12]. Recent nesting monitoring in the islands has identified a mean of 34 nesting events per reproductive season between the years 2015 and 2018[10,13]. In 2015, hawksbill turtles were also reported to be nesting in the archipelago, although the mixed morphology of some nesting females suggested that they could be hybrids[13]. Despite the hypothesis of the existence of hybrids in Abrolhos, no hybrid was detected in the above-cited population study of hawksbill juveniles[14].

The proportion of eggs that produced live hatchlings, called hatchling success (HS), ranged from 25.4% in 2015/2016[13] to 46.8% in 2016/2017 in Abrolhos[10]. The annual average HS of loggerhead turtles in Espírito Santo, Bahia and Rio de Janeiro states in Brazil was 79.9%[15], 73.1%[16] and 76.74%[17], respectively, which demonstrated that the HS in Abrolhos is very low. This could be associated with the features of Abrolhos' beaches, which are very short with a high number of big rocks and the presence of grasses, frequently taken by tides, besides being composed by a mixture of carbonate debris and rock fragments[5]. The difficulty to find appropriate nesting sites in Abrolhos is confirmed by the low success rate of nesting attempts, since 40% of sea turtles that emerged from the ocean onto the beach during the reproductive season of 2017/2018 returned to the ocean without nesting[10]. Alternatively, other biotic factors associated with the sea turtle population of Abrolhos may also affect the HS, such as interspecific hybridization[18].

Abrolhos Archipelago is located 70 km off shore from Bahia state, where the frequency of hybridization between loggerhead and hawksbill turtles has reached up to 42% of the identified hawksbill's population[19]. All the hybrid females nesting at Praia do Forte, Bahia, are first generation (F1) hybrids, which suggests that this event is recent and probably associated with environmental changes and global population decline of sea turtles[20]. The study on the reproductive output of nesting sites in Bahia indicated that the emergence success is lower for hybrids, but other parameters such as hatchling production per clutch, clutch frequency and breeding frequency were similar to parental species[21]. Furthermore, no negative effect of hybridization was found on hatchling viability[22]. However, little is known about sea turtle hybrids in other life stages in terms of survivorship, migratory behaviour, feeding preferences, reproductive fitness and mating choices[20].

In this paper, we tested the hypothesis of the existence of female sea turtle hybrids at Abrolhos Archipelago, through genetic analysis of mitochondrial DNA (mtDNA) and multilocus data. The hybridization process and introgression level were examined using genomic data, obtained by the 3RAD method. We also assessed the reproductive output of the Abrolhos population and tested whether interspecific hybridization is associated to low hatchling success in Abrolhos. These analyses provide new insights on whether this phenomenon represents a threat to different sea turtle species, and particularly to the Abrolhos population.

Methods

Study area

This study was performed in Abrolhos Archipelago (17º20′–18º10′S and 38º35′–39º20′W), a set of five small volcanic islands situated in the Abrolhos Bank, 70 km off shore from Bahia state (Fig. 1). The nesting area of sea turtles includes short sand stretches in Santa Bárbara and Redonda islands. Santa Bárbara is the largest island in the archipelago and presents two gravelly and sandy beaches in the northern and southern sides. Redonda island exhibits a longer coral sandy stretch strongly tide-influenced, with beach width ranging between 0 and 5 meters. Fringing reefs surround great part of the islands[1]. Data collection was conducted between September and April, characterized by hot summer temperatures of the tropical humid climate of Abrolhos. March is the warmest month with mean water surface temperature of 27°C[5]. The archipelago is part of the Abrolhos National Marine Park, created in 1983 and responsible for the conservation of its unique ecosystem.

Sampling and data collection

Sampling and data collection was performed by the Abrolhos National Marine Park team applying the methodology of Projeto TAMAR[16,23]. Sporadic patrols were conducted nightly along the beaches of the Santa Bárbara and Redonda islands during September and March from 2015 to 2019. Nesting females encountered were double tagged with Inconel metal tags on the front flippers and their curved carapace length

(CCL) and curved carapace width (CCW) measured[23]. During the 2015/2016 reproductive season, the females (N=6) found during oviposition were sampled through a small piece of skin taken from either the anterior flipper or the neck region using a sterilized scalpel. These samples were stored in EtOH 70% at -20°C.

Patrols during the morning were conducted daily on the beaches of Santa Bárbara Island and every two days in Redonda Island. Based on the tracks and nesting signs, nests were identified by carefully probing with a stick and then digging by hand to locate the eggs. All nests were marked with a unique number and monitored until hatchlings emergence, designated as successful nesting events. Once hatchling tracks emerging from the nest were detected, the hatched nests were excavated, the species were identified when possible, and the number of live and dead hatchlings and unhatched eggs were counted. The species were classified as loggerhead turtle or a putative hybrid, based on the size of the head, the shape of the beak, the number of lateral scutes and the shape of the shell. The loggerhead turtle has a heart-shaped carapace, a rather long and very broad head and five pairs of lateral scutes with the anterior pair touching the pre-central scute, as described by Conceição, et al.[24]. The hawksbill turtle has a medium-sized head, strong horny beak, depressed oval carapace, shell-imbricated scutes, and four pairs of lateral scutes on the carapace with the anterior pair not touching the pre-central scute. The putative hybrid hatchlings exhibit a different pattern of scutes in the shell, with medium-sized head, pointy beak and depressed oval carapace, similar to hawksbill turtles (Fig. S1).

During the 2017/2018 reproductive season, the nests identified as putative hybrids were sampled (N=20) by collecting anterior flipper fragments from dead hatchlings and embryos as described above. The samples were stored in EtOH 70% at -20°C and deposited in the Taxonomic Collections Center (CCT) of Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil. The species assignment was confirmed by genetic analysis.

The incubation period (IP) was calculated as the period between oviposition and hatchling emergence. The clutch size (CS) was obtained counting the total number of unhatched eggs, live and dead hatchlings per nest. The hatchling success (HS) was calculated through the proportion of eggs that produced live hatchlings reaching the beach surface. The female body size was calculated using the first measurement of CCL and CCW recorded for each animal. We compared these parameters between two different groups: (1) loggerhead turtles and (2) hybrids, including nesting females which were

encountered during the nesting and hatchlings whose species identification were confirmed by genetic analysis. To test for differences between groups, statistical analysis was performed with nonparametric t-test (Mann-Whitney U test) implemented in GraphPad Prism 5.0 software.

Genetic analysis

We extracted genomic DNA from the tissue samples with either a modified phenol-chloroform protocol[25] or DNeasy Blood and Tissue kit (Qiagen). We analysed one mtDNA marker and six nuclear DNA (nDNA) markers for each female and one hatchling of each putative hybrid nest. The control region of mtDNA was amplified using the primers LCM 15382 and H950[26]. The polymerase chain reaction (PCR) included 200 μM dNTP, 0.3 units (U) Platinum Taq DNA polymerase (Invitrogen by Life Technologies), 1.5 mM of $MgCl_2$, 0.5 mg/mL of Bovine Serum Albumin (BSA), 0.5 μM primers forward and reverse and 15 ng genomic DNA in 1X reaction buffer in a final volume of 10 μl. PCR cycling conditions were performed with one initial denaturation cycle of 95ºC for 5 minutes, 35 cycles of denaturation at 95ºC for 30 seconds, annealing at 50°C for 30 seconds, extension at 72 ºC for 1 minute, and a final extension at 72 ºC for 7 minutes.

The nDNA markers 421, 856, 3061, 64188, 76958 and 109472 were selected based on the species assignment accuracy according to Arantes, et al.[20]. These loci exhibited greater power to identify diagnostic alleles/haplotypes to species loggerhead and hawksbill turtles. We conducted PCR reactions as described in Arantes, et al.[20], with the following PCR cycling conditions: 95 °C for 5 min, 35 cycles of 95 °C for 30 s, primer-specific annealing temperature (57°C for locus 421, 59°C for 856, 56°C for 3061, 64°C for 76958, and 58°C for 64188 and 109472) for 30 s, 72 °C for 1 minute, and a final extension of 72 °C for 10 min. PCR reactions of 15 μL included 15 ng of genomic DNA, 1.5 mM $MgCl_2$, 0.5 μM of each primer, 200 μM dNTP, 0.3 U of Platinum Taq DNA polymerase and 1X Taq buffer (Invitrogen by Life Technologies).

PCR efficiency was checked by electrophoresis on 1% agarose gel and the products were cleaned up by precipitation using 20 mM polyethylene glycol and 2.5 mM NaCl. Amplicons were sequenced on the ABI 3130xl DNA sequencer (Applied Biosystems).

Genetic Data Analysis

Consensus sequences were generated using the SeqScape v2.6 software (Applied Biosystems), which was also used to check the quality of the electropherograms. Sequences alignments were performed with the ClustalW algorithm in the MEGA 7 software[27]. We included in the alignment for each marker all known loggerhead and hawksbill haplotypes previously determined by Arantes, et al.[20]. Gametic phases for nDNA markers were resolved with the PHASE algorithm[28] and the haplotypes inference was done with DnaSP v5 program[29].

For the control region mitochondrial marker, haplotypes were identified running a BLAST search implemented in NCBI database (http://www.ncbi.nlm.nih.gov). Haplotype identification was checked and named as assigned by Shamblin, et al[30].

For nDNA markers, the diagnostic sites were used to test the hybrid index (HI). We used the R package gghybrid[31], which calculates the proportion of allele copies coming from parental reference sets using a Bayesian algorithm[32,33]. We included as parental references five individuals each of loggerhead and hawksbill turtles from different populations along the Brazilian coast analysed by Arantes, et al.[20]. We ran HI estimations using 10,000 Markov Chain Monte Carlo (MCMC) iterations after a 100,000 MCMC burn-in period.

3RAD library construction and sequencing

We analysed female loggerhead turtles from Sergipe (N=3), foraging loggerheads from Elevação do Rio Grande (N=2), foraging hawksbill turtles from Fernando de Noronha (N=2) and Abrolhos (N=3) and 14 hatchlings from 6 different nests collected at Abrolhos. We used the 3RAD protocol described by Bayona-Vásquez, et al.[34]. Briefly, we digested 100 ng of genomic DNA with the enzymes *Mse*I, *Eco*RI and *Cvi*QI (New England Biolabs) at 10 U/μL and the specific combination of adapter for each sample for one hour at 37°C. We proceeded with the ligation of adapters immediately after the digestion. Barcoded samples were equimolarly pooled and cleaned with 0.8X CleanPCR magnetic beads (GC biotech). Subsequently, we performed the size selection of fragments between 350 and 450 bp with Blue Pippin equipment using a 1.5% cassette and the R2 marker (Sage Science). The libraries were submitted to a single-cycle PCR to add the Tru5-8N primer[35], followed by the ten-cycle indexing PCR, when we used the P5 and P7

primer to complete the library construction. The final library was characterized with a qPCR using the KAPA Library Quantification Kit (Kapa Biosystems) and checked with the Agilent 2100 Bioanalyzer High Sensitivity DNA. The fragments were sequenced using paired end reads on the HiSeq 4000 platform (Illumina) with the TruSeq 300-cycle Kit.

3RAD processing and variant detection

Illumina reads were demultiplexed based on P7 indexes and the raw sequence quality was checked with Fastqc[36] and Multiqc[37]. The software FLEXBAR[38] was used to demultiplex the P5 inline barcodes. To remove PCR replicates, we used the python script Filter_PCR_duplicates.py, which recognizes identical combinations of i5 sequences tags and the first 100 bp of each read pair sequence and keeps only a single representative copy of the pair in the output. We used the software PEAR[39] to merge reads pairs with overlapping regions of at least 30 bp. All merged reads with length less than 240 bp were considered as short fragments (i.e. out of the size selection range) and were removed from subsequent analysis. The unassembled reads were kept in the analysis as paired-end reads, filtered for quality (Q>30) and trimmed to a maximum length of 130 bp using Trimmomatic[40]. Read 2 presented low quality in the first 6 bp, which were therefore removed from the sequences. This probably occurred due to a low sequence diversity on the *Eco*RI recognition site, which was identical for virtually all reads. We used the Check_Restriction_Site.py script to select only read pairs digested by *Mse*I, filtering out any fragment digested by *Cvi*QI. Finally, read pairs containing internal complete restriction sites of *Mse*I, *Eco*RI and *Cvi*QI were removed using the Filter_Reads.py script.

3RAD sequences were analysed using the *Chelonia mydas* genome (GenBank accession number GCA_000344595.1) as a reference[41]. We used the Bowtie2 software[42] to perform the mappings with default parameters and the bowtie2 flags "--no-mixed", "--no-discordant" to ensure that only paired reads aligned to the same locus would be present in the SAM files. Aligned reads were then analysed with the Stacks reference-based pipeline[43]. The gstacks.pl pipeline was used to call variant sites within the population for each locus and genotype each individual at each identified SNP. The populations.pl pipeline was used to filter population parameters, allowing to include a locus in the final data set if the locus was genotyped in at least 40% of individuals within a population and was present in all 3 populations.

Since haplotypes increase the power for detecting population structure regarding SNPs[44], we used the Create_Haplotype_Structure.py script to code the haplotypes as multi-allelic loci to bayesian clustering analysis of STRUCTURE software[45]. This script uses Stacks output files and selects the minimum coverage needed to analyse a locus based on the number of mapped reads per locus in the SAM files (--readCOV), adjust the number of populations a locus must be present in (--popCOV) and the percentage of individuals in a population (--intraCOV) required to include the locus in the STRUCTURE output. It also filters out all the reads with undefined positions. We used the following parameters: readCOV=6, popCOV=3, --intraCOV=0.7.

The haplotype data was used to test the assignment of individuals to populations assuming the admixture model in the STRUCTURE software. We included individuals of loggerhead (N=5) and hawksbill (N=5) turtles as a control for species-specific haplotypes and hybrid diagnostic. Using independent allele frequencies, ten independent runs for each K value (from K=1 to K=5) were performed with 200,000 MCMC repeats after a 50,000 burn-in period. We summarized the replicates and visualized the STRUCTURE-estimated membership coefficients using CLUMPAK[46].

The parental species and the hybrid categories (F1, F2, backcrosses) of each hybrid individual were determined with the program NewHybrids v. 1.1 Beta3[47]. The analysis was done using a burn-in period of 10,000 followed by 50,000 MCMC iterations with Jeffrey option and no priors. We restricted the dataset to the first 300 loci, due to the NewHybrids' limited processing capacity. We also tested NewHybrids with a different set of 300 loci that were selected randomly to ensure the consistency of the results. The R package HybridDetective was used to plot NewHybrids analysis[48].

Results

We analysed a set of six nuclear markers with species-specific sites for six nesting females and for one hatchling in each of 20 distinct nests, which allowed us to identify the species, type of hybrid crosses and introgression level. Bi-allelic SNP data analysed through Bayesian hybrid-index showed that five out of the six females are loggerhead turtles and one female is a hybrid, confirming the morphological assignment (Fig. 2). This hybrid female presented one allele of hawksbill turtle and one allele of loggerhead turtle for all nuclear markers, as well as the mtDNA haplotype CC-A4.1, which is typically

found in loggerhead turtles of Brazilian rookeries. Thus, the results suggest that this individual is a first generation (F1) hybrid between a male hawksbill and a female loggerhead (HxL).

Out of the 20 nests analysed, the hybrid index showed that 11 hatchlings are loggerhead turtles, four hatchlings are backcrosses of HxL with loggerhead turtles and five hatchlings are backcrosses of HxL with hawksbill turtles (Fig. 2). It shows that the introgression process is occurring in both directions in the Abrolhos Archipelago. Five nests assigned as putative hybrids were indeed pure loggerhead turtles. Thus, we were able to confirm a species assignment based on morphology for 75% of the hatchlings.

3RAD data for 24 genotyped individuals yielded an average of 830k reads per individual (SD ± 460k) after quality filtering steps. The average per-sample locus coverage was 39.8x ± 22.2x. The datasets obtained with Stacks initially generated 23,406 loci and retained 4534 after the application of all filters. A total of 13,882 variant sites was combined in 2405 haplotypes, which were used in the analysis.

The genomic data analysis using STRUCTURE software confirmed the species assignments and introgression level for 6 nests (Fig. 3A). The hybrid individuals analysed with 3RAD method were identified as backcrosses with loggerhead (nests 11 and 17), refining the hybrid class assignment obtained with multilocus data. The results for the assignment to the different classes had 100% posterior probability support for all individuals according to NewHybrids analysis (Fig. 3B).

Analysis of the mtDNA control region revealed that all loggerhead and hybrid turtles of Abrolhos presented haplotype CC-A4.1, which was reported in rookeries in the state of Sergipe, Bahia, Espírito Santo and Rio de Janeiro[30]. Therefore, the Abrolhos population is not genetically distinct from coastal Brazilian populations.

A total of 31 records from 20 different nesting females were obtained in Abrolhos' beaches during the reproductive seasons between 2015 and 2019, being 15 loggerheads and 5 hybrids. Body size (CCL and CCW) was recorded for 18 females and showed that hybrids appear to be slightly larger and longer than pure loggerhead turtles, although no statistical significance was found (Fig. 4).

We compared the reproductive output of 31 nesting events from loggerhead turtles to 16 from hybrids, whose species assignment was confirmed by either female morphology or genetic analysis (Fig. 4). The clutch size was larger for hybrids (mean ± SD = 137±20.06) than for loggerheads (113.7±24.8), as expected due to the bigger body size of hybrids. Among the clutches, the number of unhatched eggs was greater for

hybrids (87.9±43.4) in comparison to loggerheads (40.5±36.1). Incubation period was significantly longer for hybrids (55±3.06) compared to loggerheads (52.6±3.5). Temporal distribution of nests analysed in this work revealed that the reproductive season for loggerhead females finish earlier than hybrids (Fig. 5).

The hatchling success was significantly greater for loggerheads (56.8±31.3) when compared to hybrids (27±26.97), which suggests that the low reproductive success (clutch size and hatching success) in Abrolhos Archipelago can be related to the hybridization process.

Discussion

Abrolhos: a new hybrid area along the Brazilian coast

There are two hybrid zones along the Brazilian coast: one involving mainly loggerhead and hawksbill turtles in the northern coast of Bahia state[19–22,49] and another involving loggerhead and olive ridley turtles in Sergipe state[50]. This work confirmed a new hybrid spot in Brazil, with hybrid F1 HxL females nesting at the Abrolhos Archipelago and their backcrossed progenies with loggerhead and hawksbill males.

The introgression process occurring in Abrolhos is bidirectional, i.e., there is gene flow between F1 hybrids and both parental species[49]. This is the first time that backcrosses with loggerhead turtles are confirmed by genomic analysis, since only backcrosses with hawksbills were previously confirmed by multi-locus analysis[20,48], and a backcross with loggerhead was suggested using only one nuclear marker[22].

It is well known that Abrolhos Archipelago is a breeding area for loggerhead turtles, as well as a foraging area for hawksbill turtles[6,7,13]. Thus, opportunities for interspecific mating with both species are possible in Abrolhos, as evidenced by our data. Our results suggest that the hybrid females are highly promiscuous in choosing the mating species.

Satellite telemetry studies tracking post-nesting females of loggerhead, hawksbill and hybrids from northern coast of Bahia (Praia do Forte) demonstrated that foraging areas of hawksbills are associated with coastal reef ecosystems[51] and loggerheads use foraging areas along the northern coast of Brazil[52]. Most of the tracked hybrids used the same migratory corridor as loggerheads to foraging areas along the northern coast of

Brazil, albeit one hybrid migrated southward, reaching habitats associated with reef in the southern coast of Bahia, close to the Abrolhos area[51]. No hybrid was detected among the hawksbill juvenile population of Abrolhos, which may indicate that they are not recruited to the same foraging area of "pure" hawksbills[14]. Nevertheless, hybrids are morphologically more similar to hawksbills (Fig. S1). Further studies better characterizing the Abrolhos population are necessary to understand hybrids behaviour.

The first hybrid crossing that originated F1 female hybrids currently nesting in Abrolhos beaches may have occurred between individuals either from the Abrolhos population, if the philopatric behaviour is also considered for hybrids, or from other nesting sites (like Praia do Forte), since only recently the "hawksbill turtles", here identified as hybrids, were registered in Abrolhos.

This is the first documentation of sea turtle hybrids nesting in islands. Indeed, the nesting of loggerhead turtles in Abrolhos Archipelago is also an atypical behaviour, since their nesting beaches typically are wide and open beaches fronted by a flat approach from the sea and backed by low dunes[53]. The nesting monitoring of sea turtle in Abrolhos' beaches have been surveyed since 2015 and will clarify important behaviours as the natal fidelity of females.

The high-frequency hybridization process that occurs in Brazil is probably associated to the human-mediated decline of sea turtle populations and the highly female-biased sex ratio, which may decrease the chance of con-specific mating encounters[49,54,55]. Furthermore, the timing of reproductive seasons and the abundance of species along the Brazilian coast also favour interspecific crossings. The loggerhead females nest earlier than hawksbills[21]. Hawksbill males arrive in the reproductive area of Bahia around the loggerhead nesting peak (November and December), having the opportunity to mate also with loggerhead females, which is the most abundant species in the Brazilian coast[14,49].

Given the recent effects of climate change and other anthropogenic pressures that are currently destabilizing populations[56], interspecific hybridization is an increasing conservation concern. Additional populations should be analysed to investigate the hybridization phenomenon between sea turtle species worldwide and its consequences for the species conservation.

Comparison of reproductive output of hybrid and loggerhead turtle

Hybrids presented larger body and clutch sizes relative to loggerhead turtles. Both parameters are interrelated, as the volume of eggs in a clutch is constrained by the volume within the hard shell of the female[21,57]. Our outcomes agree with previous results of Marcovaldi, et al.[51] and Soares, et al.[21], which found that HxL hybrid turtles are larger than hawksbill and loggerhead turtles. Soares, et al.[21] hypothesized that either hybrids reach sexual maturity later or have faster growth rates than the parental species.

The average HS for loggerhead turtles in Abrolhos Archipelago was 56.8%, which is lower than the loggerhead populations of Espírito Santo (79.9%), Bahia (73.1%) and Rio de Janeiro (76.74%)[15–17]. The Abrolhos' beaches features can be associated with lower reproductive success in the archipelago[10]. They consist of narrow and short stretches of coral sand with a high number of rocks and grasses coverage (Fig. 1). The beaches' widths range between 0 and 5 meters, depending on the tide, with reefs obstructing attempts of offshore approach[58]. Large sand grain size and low vegetation cover positively influence sea turtle hatchling success[59], while areas subject to tidal inundation lead to lower success[60]. Further studies should investigate the influence of sand structure and local conditions on the hatchling viability in Abrolhos.

The average HS for hybrids in Abrolhos was 27%, even lower than for loggerhead turtles. It demonstrated that the low overall reproductive success in Abrolhos is likely related to hybridization, with a reproductive disadvantage of hybrids. This outcome is related to the greater number of unhatched eggs for F1 HxL hybrids in comparison to loggerhead turtles. However, the presence of embryos in the unhatched eggs was not checked. Future research should evaluate whether the low HS is associated with low egg fertility or embryonic mortality, the two factors that can be responsible for this low success rate[61].

Crosses between genetically differentiated populations/species may conduct to the low reproductive viability of hybrids relative to their parents, referred as outbreeding depression, due to the disruption of local adaptation, the breakup of coadapted gene complexes, and/or the expression of hybrid incompatibilities[62]. The effect of outbreeding depression may be strongest in F1 hybrids, especially if the two hybridizing species have different karyotypes[18]. However, all sea turtle species have the same karyotype[63] and therefore F1 hybrids must carry a haploid set of homologous chromosomes from each parental lineage. In such cases, the effect of outbreeding depression is commonly delayed

until the F2 generation or later, when segregation and recombination begin to break apart coadapted genes from a single line[64]. The fitness decline of F2 hybrids has been showed for other taxa, such as tidepool copepod (*Tigriopus californicus*)[65] and largemouth bass (*Micropterus salmoides*)[66].

Previous studies analysing the reproductive output of F1 HxL hybrid females compared to the parental species in Praia do Forte in Bahia state revealed that the emergence success was the lowest for hybrids, but the hatchling production (product of clutch size and emergence success) per clutch was similar among all groups[21]. In addition, the clutch size, incubation period, observed clutch frequency and observed breeding frequency was similar among all groups, suggesting no reproductive advantage or disadvantage of hybrids relative to their parental species. The proportion of viable hybrid hatchlings was also similar to 'pure' hatchlings[22].

The reproductive success in other hybrid groups diverges greatly. Hybrids between the introduced brook trout (*Salvelinus fontinalis*) and the native white-spotted charr (*S. leucomaenis*) have experienced outbreeding depression in Hokkaido, Japan[67]. For hybrids between subspecies Yellow-shafted Flicker (*Colaptes auratus auratus*) and Red-shafted Flicker (*C. a. cafer*), no reproductive consequences were observed[68]. The first studies evaluating hybrids between Glaucous-winged Gulls (*Larus glaucescens*) and Western Gulls (*L. occidentalis*) showed greater reproductive performance for hybrids in relation to their parental species[69,70]. However, posterior studies found greatest breeding success for *L. occidentalis* among gulls in the hybrid zone[71,72]. Thus, interspecific hybridization may have destructive or constructive outcomes for groups. Our study highlights the importance of combining ecological and genetic data to provide important evidences about the effect of hybridization acting on sea turtle populations.

The incubation period (IP) was significantly greater for hybrids compared to loggerheads. IP is determined by the temperature at which a clutch develops, which varies during the nesting season. We also found that the nesting events for loggerhead females finish earlier than for hybrids. March is the warmest month in Abrolhos and only hybrids were observed nesting in this month. The IP and the temporal distribution of nesting may be related. However, the data collection of this work included hatchlings and females sampled in different reproductive seasons, with possible sampling bias. We recommend additional studies collecting seasonality data of temperature and reproductive output. Nevertheless, our temporal nesting distribution corroborates with Soares, et al.[21], who

showed that loggerhead females nest earlier than both hybrids and hawksbills and the hybrids have a temporal nesting distribution that overlaps with both parental species.

Considering that the estimated nesting frequency per season for loggerhead turtles is around four events[73,74] and that the average number of nesting events per reproductive season in Abrolhos is 34[10], we could infer that the number of females nesting in Abrolhos per season is about nine. Despite the small population size, it deserves thorough and continued research to understand the likely association between the hybridization process and reproductive success, which can impact sea turtles conservation in Abrolhos.

The Abrolhos National Marine Park was created in 1983 aiming to protect the largest marine biodiversity in Brazil. However, the conservation of this unique ecosystem is threatened by several factors, such as the illegal fishing activity, overfishing around the protected area, expanding tourism, dredging, invasive species, oil drilling and others[3,75]. More recently, a rupture of a tailings dam in Mariana (Minas Gerais, Brazil) caused a catastrophic environmental hazard in the Doce River and affected also its estuary, reaching the Abrolhos protection area and potentially contaminating the ecosystem[76]. These environmental stressors act in synergy reducing the resilience of the environment and species. Here we present another important conservation concern regarding sea turtle population of Abrolhos, where hatchling success is likely dependent on several individual-related as well as environmental variables.

References

1. Leão, Z. M. A. N. & Kikuchi, R. K. P. The Abrolhos Reefs of Brazil. in *Coastal Marine Ecosystems of Latin America* **144**, 83–96 (Springer Berlin Heidelberg, 2001).
2. Amado-Filho *et al.* Rhodolith beds are major CaCO3 bio-factories in the tropical South West Atlantic. *PLoS ONE* **7**, e35171 (2012).
3. Dutra, G. F., Allen, G. R., Werner, T. & McKenna, S. A. A Rapid Marine Biodiversity Assessment of the Abrolhos Bank, Bahia, Brazil. in *RAP Bulletin of Biological Assessment* **38** (Conservation International, 2005).
4. Moura, R. L. *et al.* Spatial patterns of benthic megahabitats and conservation planning in the Abrolhos Bank. *Cont. Shelf Res.* **70**, 109–117 (2013).
5. Leão, Z. M. A. N. Abrolhos -The south Atlantic largest coral reef complex. in *Geological and Paleontological Sites of Brazil* (1999).
6. Pedrosa, L. W. & Veríssimo, L. *Redução das Capturas Incidentais de Tartarugas Marinhas no Banco dos Abrolhos*. 233 (Relatório Final de Atividades - Parceria CBC/CI - FY04, 2006).
7. Proietti, M. C., Reisser, J. & Secchi, E. R. Foraging by Immature Hawksbill Sea Turtles at Brazilian Islands. *Mar. Turt. Newsl.* (2012).

8. IUCN. The IUCN Red List of Threatened Species. (2019). Available at: http://www.iucnredlist.org.
9. ICMBio. *Plano de Ação Nacional para Conservação das Tartarugas Marinhas*. (Instituto Chico Mendes de Conservação da Biodiversidade, 2011).
10. ICMBio. *Monitoramento das Tartarugas Marinhas no Parque Nacional Marinho dos Abrolhos - Temporada 2017/18*. (Instituto Chico Mendes de Conservação da Biodiversidade, 2018).
11. Proietti, M. C. *et al.* Genetic Structure and Natal Origins of Immature Hawksbill Turtles (*Eretmochelys imbricata*) in Brazilian Waters. *PLoS ONE* **9**, e88746 (2014).
12. IBAMA. *Plano de manejo: Parque Nacional Marinho dos Abrolhos*. (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - IBAMA, 1991).
13. ICMBio. *Monitoramento das Tartarugas Marinhas no Parque Nacional Marinho dos Abrolhos – Temporadas 2015/16 e 2016/17*. (Instituto Chico Mendes de Conservação da Biodiversidade, 2017).
14. Proietti, M. C. *et al.* Hawksbill × loggerhead sea turtle hybrids at Bahia, Brazil: where do their offspring go? *PeerJ* **2**, e255 (2014).
15. Baptistotte, C., Thomé, J. C. A. & Bjorndal, K. A. Reproductive Biology and Conservation Status of the Loggerhead Sea Turtle (*Caretta caretta*) in Espírito Santo State, Brazil. *Chelonian Conserv. Biol.* **4**, 523–529 (2000).
16. Marcovaldi, M. A. & Laurent, A. A Six Season Study of Marine Turtle Nesting at Praia do Forte, Bahia, Brazil, with Implications for Conservation and Management. *Chelonian Conservation and Biology* **2**, 55–59 (1996).
17. Lima, E. P., Wanderlinde, J., de Almeida, D. T., Lopez, G. & Goldberg, D. W. Nesting Ecology and Conservation of the Loggerhead Sea Turtle (*Caretta caretta*) in Rio de Janeiro, Brazil. *Chelonian Conserv. Biol.* **11**, 249–254 (2012).
18. Maheshwari, S. & Barbash, D. A. The Genetics of Hybrid Incompatibilities. *Annu. Rev. Genet.* **45**, 331–355 (2011).
19. Lara-Ruiz, P., Lopez, G. G., Santos, F. R. & Soares, L. S. Extensive hybridization in hawksbill turtles (*Eretmochelys imbricata*) nesting in Brazil revealed by mtDNA analyses. *Conserv. Genet.* **7**, 773–781 (2006).
20. Arantes, L. S., Vilaça, S. T., Mazzoni, C. J. & Santos, F. R. New genetic insights about hybridization and population dynamics of sea turtles from Brazil. *Paper submitted and under review* (2019).
21. Soares, L. S. *et al.* Comparison of reproductive output of hybrid sea turtles and parental species. *Mar. Biol.* **164**, (2017).
22. Soares, L. S. *et al.* Effects of hybridization on sea turtle fitness. *Conserv. Genet.* **19**, 1311–1322 (2018).
23. Marcovaldi, M. A. & dei Marcovaldi, G. G. Marine turtles of Brazil: the history and structure of Projeto TAMAR-IBAMA. *Biol. Conserv.* **91**, 35–41 (1999).
24. Conceição, M. B., Levy, J. A., Marins, L. F. & Marcovaldi, M. A. Electrophoretic characterization of a hybrid between *Eretmochelys imbricata* and *Caretta caretta* (Cheloniidae). *Comp. Biochem. Physiol.* **97**, 275–278 (1990).
25. Sambrook, J. & Russel, D. W. *Molecular cloning: a laboratory manual*. (Cold Spring Harbor Laboratory Press, 2001).
26. Abreu-Grobois, F. *et al.* New mtDNA Dloop primers which work for a variety of marine turtle species may increase the resolution of mixed stock analyses. in *Book of abstracts, twenty-sixth annual symposium on sea turtle biology and conservation.* (International Sea Turtle Society, 2006).
27. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

28. Stephens, M., Smith, N. J. & Donnelly, P. A New Statistical Method for Haplotype Reconstruction from Population Data. *Am. J. Hum. Genet.* **68**, 978–989 (2001).
29. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
30. Shamblin, B. M. *et al.* 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**, e85956 (2014).
31. Bailey, R. I. *gghybrid: Evolutionary Analysis of Hybrids and Hybrid Zones. R package version 0.0.0.9000.* (2018).
32. Buerkle, C. A. Maximum-likelihood estimation of a hybrid index based on molecular markers. *Mol. Ecol. Notes* **5**, 684–687 (2005).
33. Fitzpatrick, B. M. Alternative forms for genomic clines. *Ecol. Evol.* **3**, 1951–1966 (2013).
34. Bayona-Vásquez, N. J. *et al.* Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *bioRxiv* (2019). doi:10.1101/205799
35. Hoffberg, S. L. *et al.* RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Mol. Ecol. Resour.* **16**, 1264–1278 (2016).
36. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *http://www.bioinformatics.babraham.ac.uk/projects/fastqc* (2010).
37. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *In: Bioinformatics. http://multiqc.info/* (2016).
38. Dodt, M., Roehr, J., Ahmed, R. & Dieterich, C. FLEXBAR-Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology* **1**, 895–905 (2012).
39. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).
40. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
41. Wang, Z. *et al.* The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nat. Genet.* **45**, 701–706 (2013).
42. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
43. Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. H. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3 Genes Genomes Genet.* **1**, 171–182 (2011).
44. Willis, S. C., Hollenbeck, C. M., Puritz, J. B., Gold, J. R. & Portnoy, D. S. Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. *Mol. Ecol. Resour.* **17**, 955–965 (2017).
45. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
46. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
47. Anderson, E. C. & Thompson, E. A. A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics* **160**, 1217 (2002).

48. Wringe, B. F., Stanley, R. E. E., Jeffery, N. W., Anderson, E. C. & Bradbury, I. R. HYBRIDDETECTIVE: a workflow and package to facilitate the detection of hybridization using genomic data in R. *Mol. Ecol. Resour.* **17**, 275–284 (2017).
49. Vilaça, S. T. *et al.* Nuclear markers reveal a complex introgression pattern among marine turtle species on the Brazilian coast. *Mol. Ecol.* **21**, 4300–4312 (2012).
50. Reis, E. C., Soares, L. S. & Lôbo-Hajdu, G. Evidence of olive ridley mitochondrial genome introgression into loggerhead turtle rookeries of Sergipe, Brazil. *Conserv. Genet.* **11**, 1587–1591 (2010).
51. Marcovaldi, M. A., Lopez, G., Soares, L. & López-Mendilaharsu, M. Satellite tracking of hawksbill turtles *Eretmochelys imbricata* nesting in northern Bahia, Brazil: turtle movements and foraging destinations. *Endanger. Species Res.* **17**, 123–132 (2012).
52. Marcovaldi, M. A. *et al.* Satellite-tracking of female loggerhead turtles highlights fidelity behavior in northeastern Brazil. *Endanger. Species Res.* **12**, 263–272 (2010).
53. Miller, J. D., Limpus, C. J. & Godfrey, M. H. Nest Site Selection, Oviposition, Eggs, Development, Hatching, and Emergence of Loggerhead Turtles. in *The Biology of the Loggerhead sea turtles.* 125–143 (Smithsonian Institution Press, 2003).
54. Marcovaldi, M. A., Godfrey, M. H. & Mrosovsky, N. Estimating sex ratios of loggerhead turtles in Brazil from pivotal incubation durations. *Can. J. Zool.* **75**, 755–770 (1997).
55. Marcovaldi, M. A. *et al.* Spatio-temporal variation in the incubation duration and sex ratio of hawksbill hatchlings: Implication for future management. *J. Therm. Biol.* **44**, 70–77 (2014).
56. Komoroske, L. M. *et al.* A versatile Rapture (RAD-Capture) platform for genotyping marine turtles. *Mol. Ecol. Resour.* **19**, 497–511 (2019).
57. Ehrhart, L. M. A review of sea turtle reproduction. in *Biology and conservation of sea turtles* 29–38 (Smithsonian Institution Press, 1995).
58. IBAMA. *Plano de uso público para o Parque Nacional Marinho dos Abrolhos*. (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - IBAMA, 2003).
59. Ditmer, M. A. & Stapleton, S. P. Factors Affecting Hatch Success of Hawksbill Sea Turtles on Long Island, Antigua, West Indies. *PLoS ONE* **7**, e38472 (2012).
60. Kamel, S. J. & Mrosovsky, N. Repeatability of nesting preferences in the hawksbill sea turtle, *Eretmochelys imbricata*, and their fitness consequences. *Anim. Behav.* **70**, 819–828 (2005).
61. Bell, B. A., Spotila, J. R., Paladino, F. V. & Reina, R. D. Low reproductive success of leatherback turtles, *Dermochelys coriacea*, is due to high embryonic mortality. *Biol. Conserv.* **115**, 131–138 (2004).
62. Todesco, M. *et al.* Hybridization and extinction. *Evol. Appl.* **9**, 892–908 (2016).
63. Bickham, J. W. Two-Hundred-Million-Year-Old Chromosomes: Deceleration of the Rate of Karyotypic Evolution in Turtles. *Science* **212**, 1291–1293 (1981).
64. Marr, A. B., Keller, L. F. & Arcese, P. Heterosis and outbreeding depression in descendants of natural immigrants to an inbred population of song sparrows (*Melospiza melodia*). *Evolution* **56**, 131–142 (2002).
65. Edmands, S. Heterosis and Outbreeding Depression in Interpopulation Crosses Spanning a Wide Range of Divergence. *Evolution* **53**, 1757–1768 (1999).
66. Goldberg, T. L. *et al.* Increased Infectious Disease Susceptibility Resulting from Outbreeding Depression. *Conserv. Biol.* **19**, 455–462 (2005).

67. Fukui, S., May-McNally, S. L., Taylor, E. B. & Koizumi, I. Maladaptive secondary sexual characteristics reduce the reproductive success of hybrids between native and non-native salmonids. *Ecol. Evol.* **8**, 12173–12182 (2018).
68. Flockhart, D. T. T. & Wiebe, K. L. Absence of Reproductive Consequences of Hybridization in the Northern Flicker (*Colaptes auratus*) Hybrid Zone. *The Auk* **126**, 351–358 (2009).
69. Hoffman, W., Wiens, J. A. & Scott, J. M. Hybridization between Gulls (*Larus glaucescens* and *L. occidentalis*) in the Pacific Northwest. *Auk Ornithol. Adv.* **95**, 441–458 (1978).
70. Good, T. P., Ellis, J. C., Annett, C. A. & Pierotti, R. Bounded hybrid superiority in an avian hybrid zone: effects of mate, diet, and habitat choice. *Evolution* **54**, 1774–1783 (2000).
71. Bell, D. A. Hybridization and Reproductive Performance in Gulls of the *Larus glaucescens-occidentalis* Complex. *The Condor* **99**, 585–594 (1997).
72. Moncrieff, A. E., Megna, L. C., Hayward, J. L. & Henson, S. M. Mating Patterns and Breeding Success in Gulls of the *Larus glaucescens-occidentalis* Complex, Protection Island, Washington, USA. *Northwest. Nat.* **94**, 67–75 (2013).
73. Frazer, N. B. & Richardson, J. I. Annual Variation in Clutch Size and Frequency for Loggerhead Turtles, *Caretta caretta*, Nesting at Little Cumberland Island, Georgia, USA. *Herpetologica* **41**, 246–251 (1985).
74. Marcovaldi, M. A. & Chaloupka, M. Conservation status of the loggerhead sea turtle in Brazil: an encouraging outlook. *Endanger. Species Res.* **3**, 133–143 (2007).
75. Costa, T. J. F. *et al.* Expansion of an invasive coral species over Abrolhos Bank, Southwestern Atlantic. *Mar. Pollut. Bull.* **85**, 252–253 (2014).
76. Magris, R. A., Marta-Almeida, M., Monteiro, J. A. F. & Ban, N. C. A modelling approach to assess the impact of land mining on marine biodiversity: Assessment in coastal catchments experiencing catastrophic events (SW Brazil). *Sci. Total Environ.* **659**, 828–840 (2019).

Acknowledgements

We thank Abrolhos Park team (a special thanks to B. Barbosa and B. Figueiredo) for field support and Maíra Proietti from FURG (Federal University of Rio Grande) for providing some samples. We also thank Felix Heeger for support with Python scripting and Susan Mbedi and Sarah Sparmann from the Berlin Center for Genomics in Biodiversity Research for advising in NGS techniques and test sequencing runs. We are grateful to the team of Dr. Sascha Sauer for the sequencing performed at the Max Delbruck Center for Molecular Medicine in Berlin. We are also grateful to Projeto TAMAR, particularly Neca Marcovaldi and Luciano S. Soares, for the invitation to participate in this research, starting in 2004. This study was approved by the Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio), and conducted under SISBIO license 54358-3 and 47954-3.

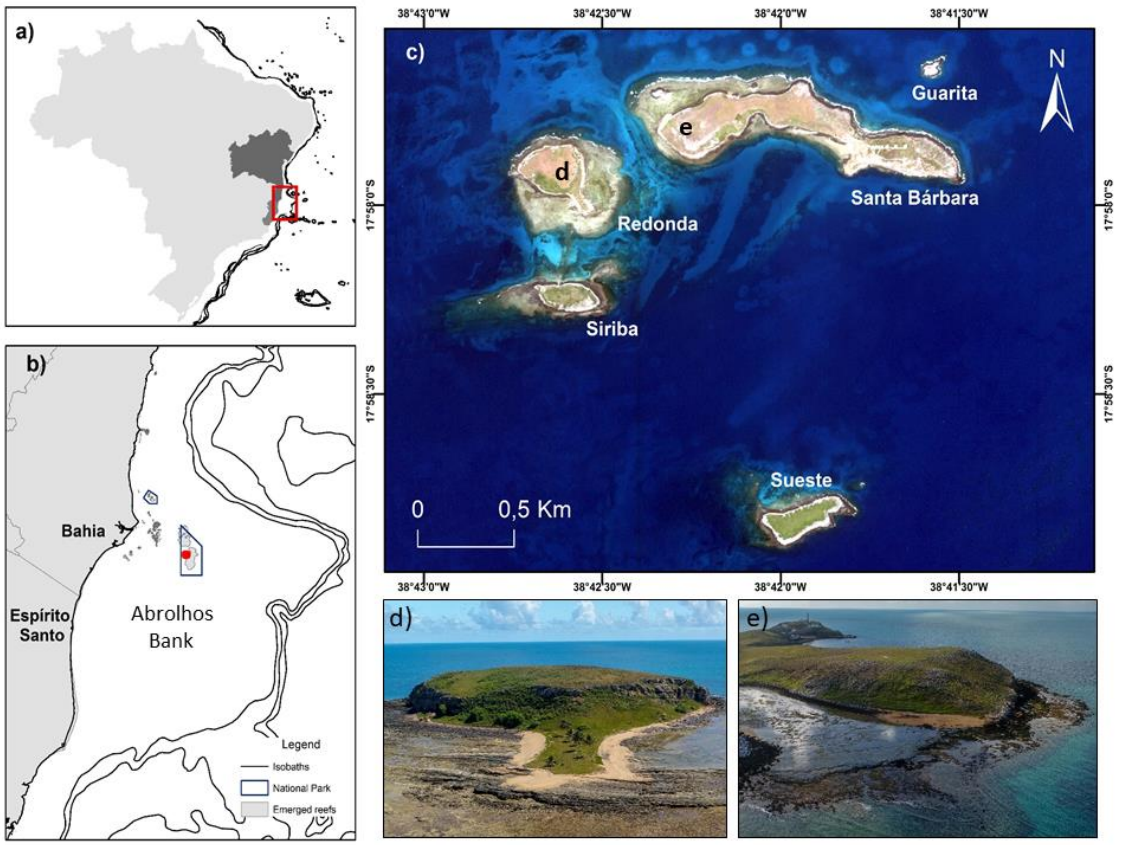

Figure 1. Nesting areas of sea turtles in Abrolhos. The archipelago encompasses five small islands (c) situated in the Abrolhos Bank (b). The nesting area of sea turtles include short sand stretches in Redonda (d) and Santa Bárbara (e) islands.

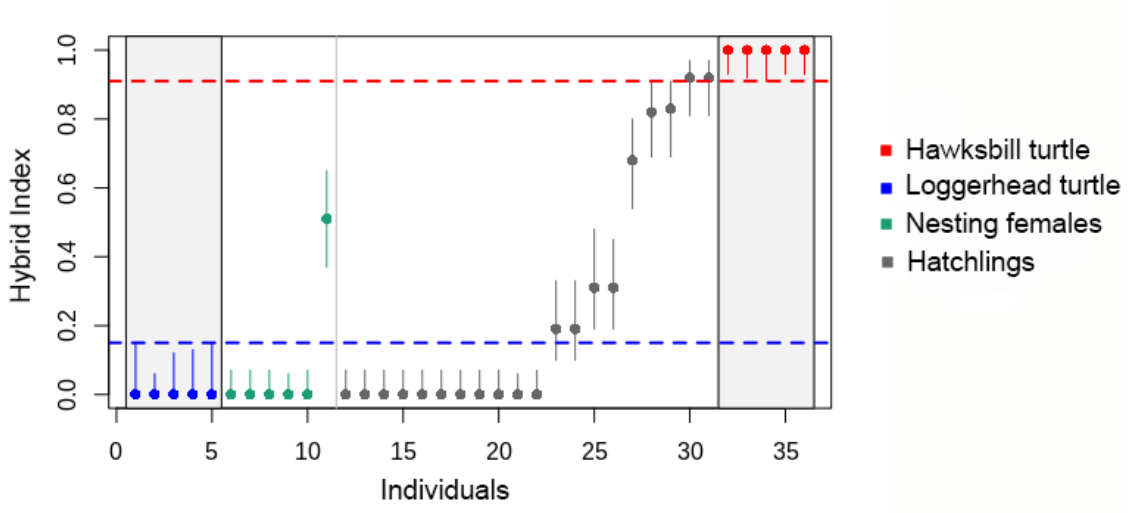

Figure 2. Bayesian estimates of the hybrid index (HI) for each individual analysed in Abrolhos Archipelago. Red and blue points represent the parental reference sets of loggerhead and hawksbill turtles, respectively. Cyan points represent nesting females and grey point are hatchlings from 20 nests from Abrolhos. The HI was estimated with gghybrid. HI estimated values equal 0.0 denotes pure loggerhead turtles and 1.0 denotes pure hawksbill individuals; lines represent 95% credibility intervals.

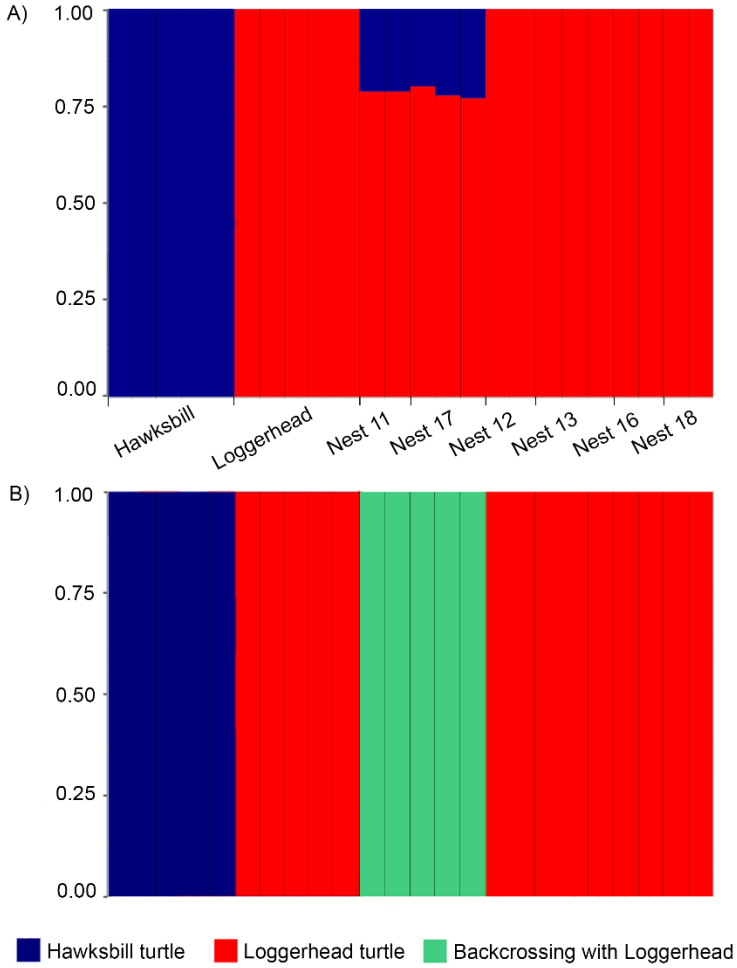

Figure 3. Genetic structure of parental species and their hybrids based on 3RAD data. Admixture proportions based on two clusters (K = 2) estimated by STRUCTURE (A) and assignment of hybrid class by NewHybrids (B). The plots include the 5 individuals of each hawksbill and loggerhead turtles and 14 hatchlings from 6 different nests from Abrolhos. The order of the individuals within the different groups is the same in both plots.

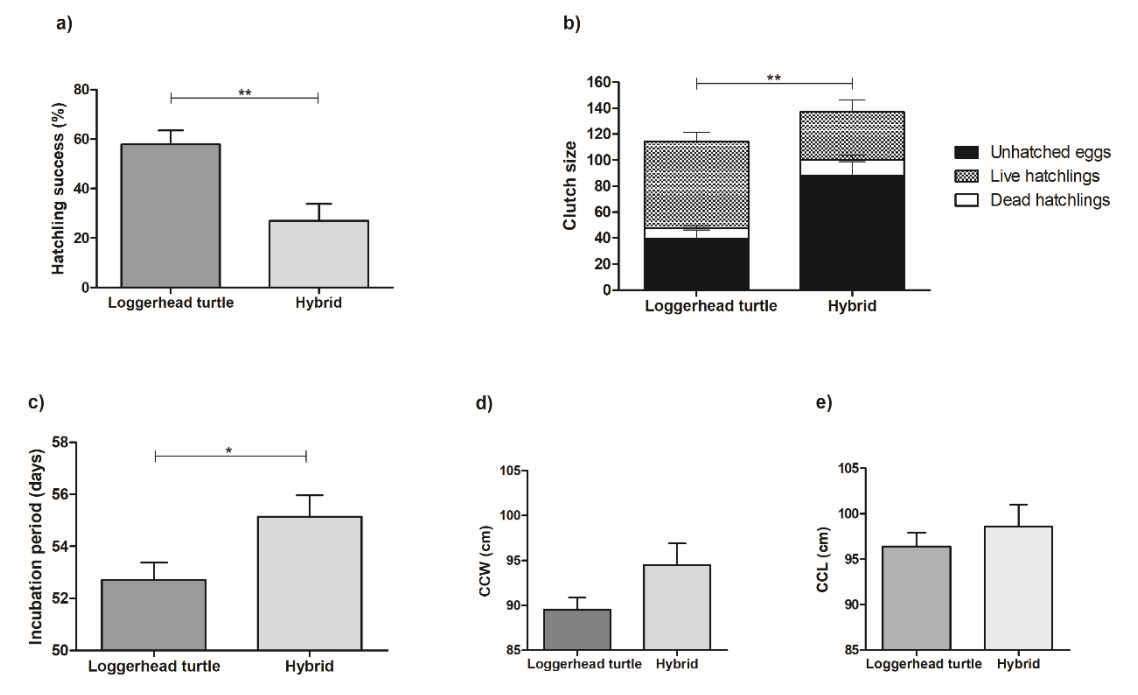

Figure 4. Graphical summary for morphological and reproductive parameters compared between loggerhead turtles and hybrids from Abrolhos Archipelago. Statistical analysis was performed by nonparametric t test (Mann-Whitney U test). **P<0.05; *P<0.01.

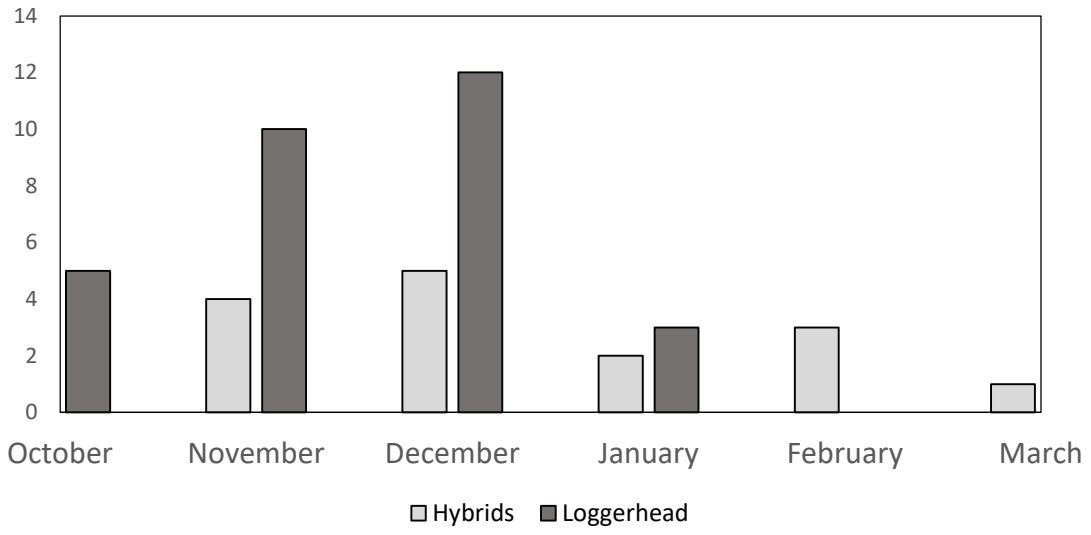

Figure 5. Nesting temporal distribution of hybrids and loggerhead turtles analysed in this study.

Supplementary information

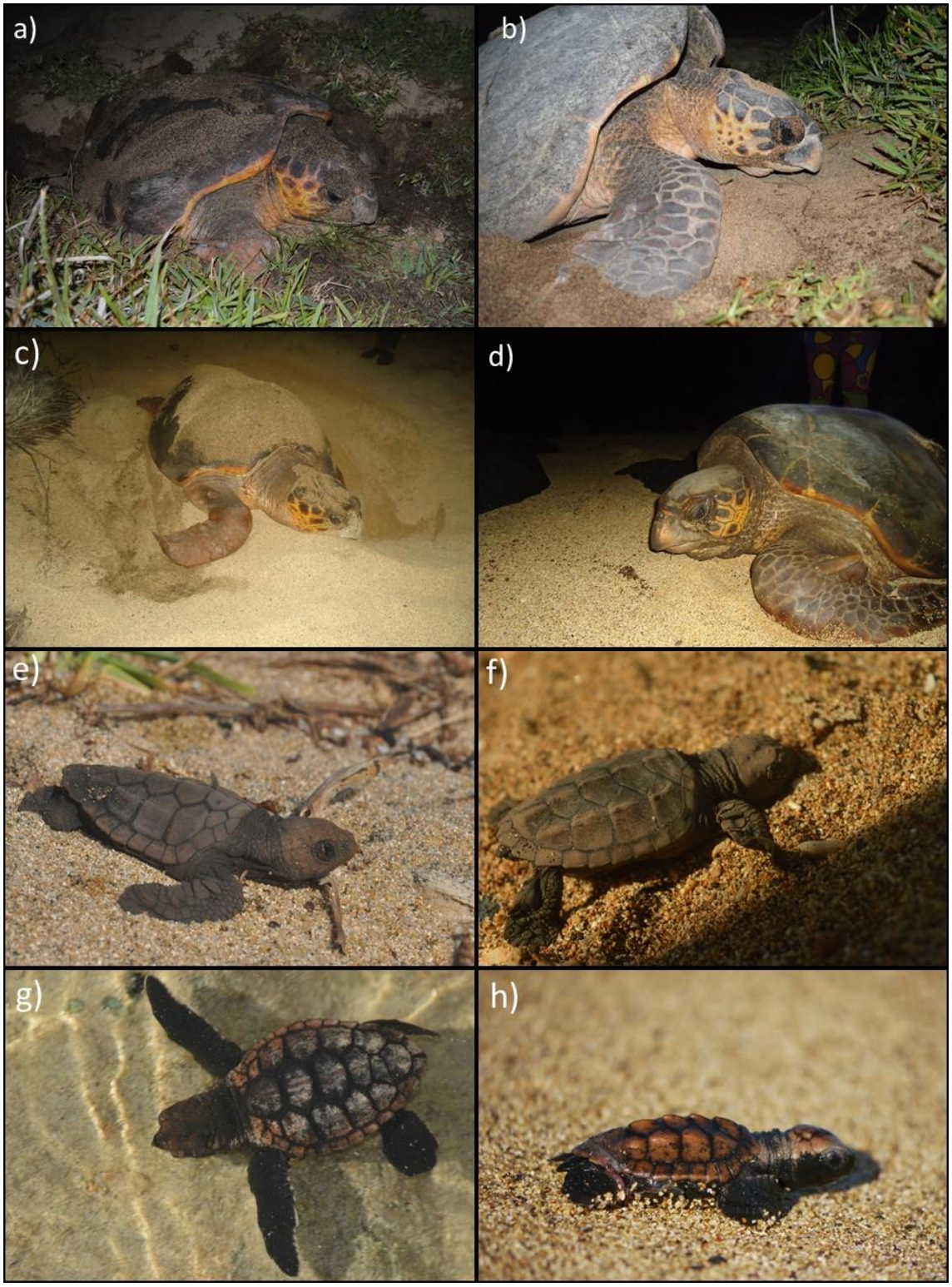

Figure S1 – Females and hatchlings of loggerhead turtles (left pictures – a, c, e, g) and loggerhead x hawksbill hybrids (right pictures – b, d, f, h). The pictures highlight the difference between "pure" and hybrid turtles in terms of the number of lateral scutes in the shell, the size of the head and the beak format.

CONSIDERAÇÕES FINAIS

Esta tese de doutorado permitiu avançar o conhecimento acerca dos processos filogeográficos, filogenéticos e de hibridização das tartarugas marinhas e demonstrou como as análises genéticas em combinação com dados ecológicos podem contribuir para a compreensão de fenômenos importantes para a conservação das espécies.

O nível de representação genômica foi aprofundado durante os três capítulos da tese, iniciando com análises de um gene mitocondrial, seguindo com uma abordagem multilocus e finalizando com uma metodologia de sequenciamento nova geração de representação reduzida do genoma.

A compilação de dados de haplótipos mitocondriais disponíveis na literatura e apresentados no primeiro capítulo da tese revelou como a utilização de uma amostragem global para uma espécie é importante para compreender padrões demográficos e evolutivos da espécie.

No segundo capítulo, demonstramos como o uso de uma abordagem multilocus com marcadores selecionados a partir da técnica de ddRADseq permitiu averiguar com maior confiabilidade o número de gerações de híbridos, resultando em uma estimativa ainda mais recente para o fenômeno de hibridização na Bahia do que previamente sugerido. Além disso, foi possível demonstrar que existe fluxo gênico mediado pelos machos entre as colônias de desova da costa brasileira para as tartarugas de pente e cabeçudas.

Os marcadores padronizados no segundo capítulo foram utilizados em conjunto com a técnica 3RADseq, proporcionando um conjunto de dados robustos a nível genômico, para averiguar a suspeita de hibridização no Arquipélago de Abrolhos, que foi apresentada no terceiro capítulo deste trabalho. Nós confirmamos a existência de híbridas desovando em Abrolhos. Além disso, o monitoramento reprodutivo das desovas no Arquipélago de Abrolhos revelou que o sucesso dos filhotes híbridos é menor que dos filhotes não-híbridos, demonstrando pela primeira vez que a hibridização pode conduzir à depressão exogâmica das tartarugas marinhas. Isto ressalta a preocupação com a conservação das populações envolvidas na hibridização.

Nossos resultados demonstraram como a expansão dos estudos genéticos, em termos de representatividade do genoma e de amostragem populacional, propiciam maior poder para realizar inferências biológicas, tão importantes para a conservação das tartarugas marinhas.

REFERÊNCIAS BIBLIOGRÁFICAS

Allendorf, F.W., Leary, R.F., Spruell, P., and Wenburg, J.K. (2001). The problems with hybrids: setting conservation guidelines. Trends in Ecology & Evolution 16, 613–622.

Barber, R.C., Fontaine, C.T., Flanagan, J.P., and Louis Jr., E.E. (2003). Natural hybridization between a Kemp's Ridley (*Lepidochelys kempii*) and Loggerhead sea turtle (*Caretta caretta*) confirmed by molecular analysis. Chelonian Conservation and Biology 4, 701–704.

Beheregaray, L.B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Molecular Ecology 17, 3754–3774.

Bentley, B.P., Haas, B.J., Tedeschi, J.N., and Berry, O. (2017). Loggerhead sea turtle embryos (*Caretta caretta*) regulate expression of stress response and developmental genes when exposed to a biologically realistic heat stress. Molecular Ecology 26, 2978–2992.

Bowen, B.W., and Karl, S.A. (2007). Population genetics and phylogeography of sea turtles. Molecular Ecology 16, 4886–4907.

Briscoe, D.K., Parker, D.M., Balazs, G.H., Kurita, M., Saito, T., Okamoto, H., Rice, M., Polovina, J.J., and Crowder, L.B. (2016). Active dispersal in loggerhead sea turtles (*Caretta caretta*) during the 'lost years.' Proceedings of the Royal Society B: Biological Sciences 283, 20160690.

Carr, A., and Ogren, L. (1960). The ecology and migrations of sea turtles: The green turtle in the Caribbean Sea. Bulletin of the American Museum of Natural History 121, 1–48.

Conceição, M.B., Levy, J.A., Marins, L.F., and Marcovaldi, M.A. (1990). Electrophoretic characterization of a hybrid between *Eretmochelys imbricata* and *Caretta caretta* (Cheloniidae). Comparative Biochemistry and Physiology Part B: Comparative Biochemistry 97, 275–278.

Duchene, S., Frey, A., Alfaro-Núñez, A., Dutton, P.H., Thomas P. Gilbert, M., and Morin, P.A. (2012). Marine turtle mitogenome phylogenetics and evolution. Molecular Phylogenetics and Evolution 65, 241–250.

Ekblom, R., and Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107, 1–15.

Encalada, S.E., Bjorndal, K.A., Bolten, A.B., Zurita, J.C., Schroeder, B., Possardt, E., Sears, C.J., and Bowen, B.W. (1998). Population structure of loggerhead turtle (*Caretta caretta*) nesting colonies in the Atlantic and Mediterranean as inferred from mitochondrial DNA control region sequences. Marine Biology 130, 567–575.

FitzSimmons, N.N., Limpus, C.J., Norman, J.A., Goldizen, A.R., Miller, J.D., and Moritz, C. (1997). Philopatry of male marine turtles inferred from mitochondrial DNA markers. Proceedings of the National Academy of Sciences 94, 8912–8917.

Hendry, A.P., Nosil, P., and Rieseberg, L.H. (2007). The speed of ecological speciation. Functional Ecology 21, 455–464.

Hurtado, L. (2016). Thousands of single nucleotide polymorphisms in the critically endangered kemp's ridley sea turtle (*Lepidochelys kempii*) revealed by double-digest restriction-associated DNA sequencing: opportunities for previously elusive conservation genetics research. Gulf of Mexico Science 214–218.

IUCN (2019) The IUCN Red List of Threatened Species. Available at: http://www.iucnredlist.org.

James, M.C., Martin, K., and Dutton, P.H. (2004). Hybridization Between a Green Turtle, *Chelonia mydas*, and Loggerhead Turtle, *Caretta caretta*, and the First Record of a Green Turtle in Atlantic Canada. The Canadian Field-Naturalist 118, 579.

Karl, S.A., Bowen, B.W., and Avise, J.C. (1995). Hybridization Among the Ancient Mariners: Characterization of Marine Turtle Hybrids with Molecular Genetic Assays. Journal of Heredity 86, 262–268.

Komoroske, L.M., Jensen, M.P., Stewart, K.R., Shamblin, B.M., and Dutton, P.H. (2017). Advances in the Application of Genetics in Marine Turtle Biology and Conservation. Frontiers in Marine Science 4.

Komoroske, L.M., Miller, M.R., O'Rourke, S.M., Stewart, K.R., Jensen, M.P., and Dutton, P.H. (2019). A versatile Rapture (RAD-Capture) platform for genotyping marine turtles. Molecular Ecology Resources 19, 497–511.

Lara-Ruiz, P., Lopez, G.G., Santos, F.R., and Soares, L.S. (2006). Extensive hybridization in hawksbill turtles (*Eretmochelys imbricata*) nesting in Brazil revealed by mtDNA analyses. Conservation Genetics 7, 773–781.

Maheshwari, S., and Barbash, D.A. (2011). The Genetics of Hybrid Incompatibilities. Annual Review of Genetics 45, 331–355.

Marcovaldi, M.A., Godfrey, M.H., and Mrosovsky, N. (1997). Estimating sex ratios of loggerhead turtles in Brazil from pivotal incubation durations. Canadian Journal of Zoology 75, 755–770.

Marcovaldi, M.A., Lopez, G., Soares, L., and 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, 123–132.

Marcovaldi, M.A., Santos, A.J.B., Santos, A.S., Soares, L.S., Lopez, G.G., Godfrey, M.H., López-Mendilaharsu, M., and Fuentes, M.M.P.B. (2014). Spatio-temporal variation in the incubation duration and sex ratio of hawksbill hatchlings: Implication for future management. Journal of Thermal Biology 44, 70–77.

Marcovaldi, M.A., López-Mendilaharsu, M., Santos, A.S., Lopez, G.G., Godfrey, M.H, Tognin, F., Baptistotte, C., Thomé, J.C., Dias, A.C.C., Castilhos, J.C., and Fuentes, M.M.P.B. (2016). Identification of loggerhead male producing beaches in the south Atlantic: Implications for conservation. Journal of Experimental Marine Biology and Ecology 477, 14-22.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., and Brumfield, R.T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66, 526–538.

Miller, J.D. (1996). Reproduction in Sea Turtles. In The Biology of Sea Turtle, (Boca Raton, Florida: Lutz PL & LA Musick), pp. 51–80.

Monzón-Argüello, C., Loureiro, N.S., Delgado, C., Marco, A., Lopes, J.M., Gomes, M.G., and Abreu-Grobois, F.A. (2011). Príncipe island hawksbills: Genetic isolation of an eastern Atlantic stock. Journal of Experimental Marine Biology and Ecology 407, 345–354.

Moore, M.K., and Ball, R.M. (2002). Multiple paternity in loggerhead turtle (*Caretta caretta*) nests on Melbourne Beach, Florida: a microsatellite analysis. Molecular Ecology 11, 281–288.

Naro-Maciel, E., Le, M., FitzSimmons, N.N., and Amato, G. (2008). Evolutionary relationships of marine turtles: A molecular phylogeny based on nuclear and mitochondrial genes. Molecular Phylogenetics and Evolution 49, 659–662.

Proietti, M.C., Reisser, J., Marins, L.F., Marcovaldi, M.A., Soares, L.S., Monteiro, D.S., Wijeratne, S., Pattiaratchi, C., and Secchi, E.R. (2014a). Hawksbill × loggerhead sea turtle hybrids at Bahia, Brazil: where do their offspring go? PeerJ 2, e255.

Proietti, M.C., Reisser, J., Marins, L.F., Rodriguez-Zarate, C., Marcovaldi, M.A., Monteiro, D.S., Pattiaratchi, C., and Secchi, E.R. (2014b). Genetic Structure and Natal Origins of Immature Hawksbill Turtles (*Eretmochelys imbricata*) in Brazilian Waters. PLoS ONE 9, e88746.

Prosdocimi, L., Bruno, I., Diaz, L., Carman, V.G.G., Albareda, D.A., and Remis, M.I. (2014). Southernmost reports of the Hawksbill sea turtle, *Eretmochelys imbricata*, in temperate waters of Argentina and evidence of a hybrid origin supported by mitochondrial DNA analysis. Herpetological Review 45, 1–5.

Putman, N.F., and Mansfield, K.L. (2015). Direct Evidence of Swimming Demonstrates Active Dispersal in the Sea Turtle "Lost Years." Current Biology 25, 1221–1227.

Reid, B.N., Naro-Maciel, E., Hahn, A.T., FitzSimmons, N.N., and Gehara, M. (2019). Geography best explains global patterns of genetic diversity and postglacial co-expansion in marine turtles. Molecular Ecology.

Reis, E.C., Soares, L.S., Vargas, S.M., Santos, F.R., Young, R.J., Bjorndal, K.A., Bolten, A.B., and Lôbo-Hajdu, G. (2010a). Genetic composition, population structure and phylogeography of the loggerhead sea turtle: colonization hypothesis for the Brazilian rookeries. Conservation Genetics 11, 1467–1477.

Reis, E.C., Soares, L.S., and Lôbo-Hajdu, G. (2010b). Evidence of olive ridley mitochondrial genome introgression into loggerhead turtle rookeries of Sergipe, Brazil. Conservation Genetics 11, 1587–1591.

Seminoff, J.A., Karl, S.A., Schwartz, T., and Resendiz, A. (2003). Hybridization of the green turtle (*Chelonia mydas*) and hawskbill turtle (*Eretmochelys imbricata*) in the Pacific Ocean: indication of an absence of gender bias in the directionality of crosses. Bulletin of Marine Science 73, 11.

Shamblin, B.M., Bjorndal, K.A., Bolten, A.B., Hillis-Starr, Z.M., Lundgren, I., Naro-Maciel, E., and Nairn, C.J. (2012). Mitogenomic Sequences Better Resolve Stock Structure of Southern Greater Caribbean Green Turtle Rookeries. Molecular Ecology 21(10), 2330–40.

Shamblin, B.M., Bolten, A.B., Abreu-Grobois, F.A., Bjorndal, K.A., Cardona, L., Carreras, C., Clusa, M., Monzón-Argüello, C., Nairn, C.J., Nielsen, J.T., et al. (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, e85956.

Soares, L.S., Bolten, A.B., Wayne, M.L., Vilaça, S.T., Santos, F.R., Marcovaldi, M.A., and Bjorndal, K.A. (2017). Comparison of reproductive output of hybrid sea turtles and parental species. Marine Biology 164.

Soares, L.S., Bjorndal, K.A., Bolten, A.B., Marcovaldi, M.A., Luz, P.B., Machado, R., Lo, R., McDaniel, S.F., Payton, A.C., Waltzek, T.B., et al. (2018). Effects of hybridization on sea turtle fitness. Conservation Genetics 19, 1311–1322.

Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G., et al. (2016). Hybridization and extinction. Evolutionary Applications 9, 892–908.

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., et al. (2016). Phylogeography, Genetic Diversity, and Management Units of Hawksbill Turtles in the Indo-Pacific. Journal of Heredity 107, 199–213.

Velez-Zuazo, X., Ramos, W.D., Van Dam, R.P., Diez, C.E., Abreu-Grobois, A., and Mcmillan, W.O. (2008). Dispersal, recruitment and migratory behaviour in a hawksbill sea turtle aggregation: Migratory behaviour of the hawksbill sea turtle. Molecular Ecology 17, 839–853.

Vilaça, S.T., Vargas, S.M., Lara-Ruiz, P., Molfetti, É., Reis, E.C., Lôbo-Hajdu, G., Soares, L.S., and Santos, F.R. (2012). Nuclear markers reveal a complex introgression pattern among marine turtle species on the Brazilian coast. Molecular Ecology 21, 4300–4312.

Vilaça, S.T., Lara-Ruiz, P., Marcovaldi, M.A., Soares, L.S., and 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.

Wallace, B.P., DiMatteo, A.D., Bolten, A.B., Chaloupka, M.Y., Hutchinson, B.J., Abreu-Grobois, F.A., Mortimer, J.A., Seminoff, J.A., Amorocho, D., Bjorndal, K.A., et al. (2011). Global Conservation Priorities for Marine Turtles. PLoS ONE 6, e24510.

Wang, Z., Pascual-Anaya, J., Zadissa, A., Li, W., Niimura, Y., Huang, Z., Li, C., White, S., Xiong, Z., Fang, D., et al. (2013). The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. Nature Genetics 45, 701–706.

van Wyk, A.M., Dalton, D.L., Hoban, S., Bruford, M.W., Russo, I.-R.M., Birss, C., Grobler, P., van Vuuren, B.J., and Kotzé, A. (2017). Quantitative evaluation of hybridization and the impact on biodiversity conservation. Ecology and Evolution 7, 320–330.