UNIVERSIDADE DO VALE DO RIO DOS SINOS
Unidade Acadêmica de Pesquisa e Pós-Graduação

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA
NÍVEL MESTRADO

VANESSA SOUZA SILVA

DNA Barcoding helping the identification of Brazilian cetaceans

São Leopoldo

2019

VANESSA SOUZA SILVA

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre, pelo Programa de Pós Graduação em Biologia da Universidade do Vale do Rio dos Sinos – UNISINOS.

Área de concentração: Conservação e Manejo de Ecossistemas e da Vida Silvestre.

Orientadora: Dra. Larissa Rosa de Oliveira

Coorientador: Dr. Victor Hugo Valiati

São Leopoldo

2019

VANESSA SOUZA SILVA

(Bolsista do programa PROSUC/CAPES)

Defesa em 28/11/2020

BANCA EXAMINADORA

_____

Dra. Larissa Rosa de Oliveira – Universidade do Vale do Rio dos Sinos Orientadora.

_____

Dr. Salvatore Siciliano - Fundação Oswaldo Cruz

_____

Dra. Flávia Pereira Tirelli Universidade Federal do Rio Grande do Sul

S586d Silva, Vanessa Souza.

DNA barcoding helping the identification of Brazilian cetaceans / Vanessa Souza Silva. – 2019.

48 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Biologia, 2019.

"Orientadora: Dra. Larissa Rosa de Oliveira
Coorientador: Dr. Victor Hugo Valiati."

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecária: Amanda Schuster – CRB 10/2517)

APRESENTAÇÃO

Este trabalho apresenta considerações sobre a utilização do DNA Barcode na identificação de carcaças de cetáceos encontradas ao longo da costa brasileira e águas interiores.

A presente dissertação é composta por dois capítulos, sendo o primeiro uma introdução do trabalho realizado e o segundo capítulo está estruturado na forma de artigo cientifico redigido em inglês, seguindo as normas do periódico “Marine Biodiversity”, com avaliação A1 no Qualis Capes. Neste capítulo são apresentadas considerações acerca da eficácia do gene mitocondrial Citocromo C Oxidase I (*COI*) na identificação a nível específico de amostras de carcaças de cetáceos encontrados na costa brasileira e também algumas amostras oriundas de biópsias, realizadas por instituições colaboradoras do projeto. Este capítulo será submetido ao periódico supracitado após as considerações dos coautores e também da banca examinadora.

Por fim, informo que este trabalho teve a contribuição de 13 instituições pesquisadoras e do Programa de Suporte à Pós-Graduação de Instituições Comunitárias de Educação Superior – PROSUC/ CAPES, a qual auxíliou no pagamento das taxas escolares

SUMÁRIO

RESUMO

A Citocromo C Oxidase Subunidade I (COI) é o gene mitocondrial reconhecido como o código de barras do DNA (DNA Barcoding) e utilizado para identificar diferentes espécies animais. No caso dos cetáceos (botos, baleias e golfinhos), o COI pode auxiliar na identificação mais precisa e/ou confirmar a identidade morfológica de espécimes encontrados em estado avançado de decomposição no litoral brasileiro, sem caracteres diagnósticos externos evidentes. Este estudo teve como objetivos testar a eficiência do COI na identificação específica da maioria das 48 espécies de cetáceos ocorrentes na costa brasileira e gerar suas respectivas sequências para depositá-las em bancos de dados públicos de sequências (BrBol, FishBol e GenBank), a fim de subsidiar novos trabalhos por pesquisadores de qualquer parte do mundo. Para tanto, o COI foi amplificado para 152 espécimes coletados por 13 instituições brasileiras e enviadas para análise. A comparação entre as sequências obtidas e as das espécies de cetáceos previamente depositadas no GenBank foi feita pela ferramenta Blast, a qual é baseada na similaridade entre as sequências. Além disso, cada amostra possuía um espécime depositado em coleção científica como material testemunho, o que permitiu sua identificação morfológica a priori. Como resultado, foram obtidas 152 sequências de COI de 33 espécies, as quais representam 70% da fauna dos 47 cetáceos registrados para a costa brasileira. Foram geradas sequências de COI das seguintes espécies: *Eubalaena australis* (n=5), *Balaenoptera acutorostrata* (n=4), *Balaenoptera bonaerensis* (n=2), *Balaenoptera brydei* (n=5), *Balaenoptera physalus* (n=1), *Megaptera novaeangliae* (n=9), *Physeter macrocephalus* (n=9), *Kogia breviceps* (n=2), *Kogia sima* (n=6), *Berardius arnuxii* (n=1), *Mesoplodon europaeus* (n=1), *Ziphius cavirostris* (n=5), *Delphinus delphis* (n=9), *Globicephala melas* (n=1), *Lagenodelphis hosei* (n=7), *Peponocephala electra* (n=4), *Pseudorca crassidens* (n=8), *Stenella attenuata* (n=4), *Stenella clymene* (n=6), *Stenella coeruleoalba* (n=5), *Stenella frontalis* (n=4), *Stenella longirostris* (n=5), *Steno bredanensis* (n=4), *Tursiops truncatus* (n=9), *Pontoporia blainvillei* (n=12), *Inia geoffrensis* (n=2), *Inia araguaiaensis* (n=3). Apenas nos gêneros *Stenella* e *Delphinus* o gene COI não foi capaz de identificar suas espécies de maneira consistente, quando comparado com os caracteres anatômicos do material depositado em coleção. Estes resultados sugerem a necessidade da continuidade deste estudo bem como reforça a importância da combinação de caracteres morfológicos e moleculares para identificação de cetáceos.

INTRODUÇÃO GERAL

O Brasil possui o mais extenso litoral inter e subtropical do mundo, com cerca de 8.000 Km de costa (Ab'saber 2001). Apesar de sua importância na região Neotropical em termos de diversidade biológica, as crescentes alterações causadas pelo homem na região marinha e costeira têm descaracterizado esses ambientes, colocando em risco a sobrevivência de muitas espécies e mesmo de comunidades inteiras (Amaral & Jablonski 2005). O conhecimento a respeito da diversidade existente nas regiões costeira e oceânica do Brasil é fundamental tanto para compreender o funcionamento dos diferentes ecossistemas associados, como para garantir o uso sustentável e a conservação de seus recursos vivos (e.g. Longo, Amado Filho 2014). Contudo, apesar da implementação de importantes programas de pesquisa nas regiões oceânicas brasileiras nas últimas décadas (e.g. REVIZEE, Programa Arquipélago e Ilhas Oceânicas), o conhecimento atual a respeito das comunidades existentes nessas regiões é ainda insuficiente para garantir a sua conservação, especialmente frente ao crescente interesse de exploração dessas áreas (Ott *et al*., 2009).

Um dos grupos taxonômicos que carece de informações essenciais tanto para uma melhor compreensão de seu papel ecológico no ecossistema oceânico, quanto para o estabelecimento de planos efetivos de conservação são os cetáceos (e.g. Zerbini *et al.,* 2004; Ott *et al*., 2009; Siciliano *et al*., 2012). Atualmente, existem registros confirmados de 47 cetáceos em águas brasileiras (Tabela 1), das quais seis são consideradas ameaçadas e, pelo menos, 15 insuficientemente conhecidas (DD) de acordo com a União Internacional para Conservação da Natureza (do inglês *International Union for Conservation of Nature* - IUCN) (MMA 2014; Hrbek *et al*., 2014; Cypriano-Souza *et al.,* 2016; ICMBio/MMA 2018).

Quanto à riqueza, existe a potencial subestimativa da real diversidade de espécies de cetáceos que seria encontrada ao longo da costa brasileira, isso em decorrência da dificuldade de identificação dos cetáceos e também do avançado grau de decomposição no qual muitos exemplares são encontrados nas praias, sendo que, na maioria das vezes, partes importantes como o crânio são perdidas (Zerbini *et al.,* 2004).

Essa questão está bem representada na família Ziphiidae, comumente chamadas de baleias-bicudas, com raros registros de encalhe na costa brasileira, mesmo em regiões com longo histórico de amostragem (*e.g.* Prado *et al*., 2016; Vianna *et al*. 2016). O comportamento discreto dessas baleias,

que inclui movimentação lenta, pouca exposição na superfície e distância de embarcações tornam as informações sobre esta família ainda mais difíceis de serem obtidas (Reeves *et al*. 2002; Dalebout *et al.* 1998). De maneira geral, a identificação das espécies é realizada, primordialmente, pela análise da morfologia do crânio, posição e formato dos dentes de espécimes encalhados, porém, a correta identificação depende, dentre outras coisas, do estágio de decomposição em que as carcaças chegam até à costa e também da maturidade do animal encalhado, sendo que não é incomum ocorrência de identificações errôneas (Dalebout et al. 1998). Além disso, a dificuldade de identificar caracteres diagnósticos, principalmente dentre as seis espécies que compõem o gênero *Mesoplodon*, e a forte sobreposição da área de ocorrência das espécies, tornam ainda mais complexa a identificação (Dalebout et al. 1998). Assim, em muitas situações, apenas a utilização de caracteres morfológicos não fornece informações suficientes para a identificação a nível de espécie.

Neste contexto, a inclusão de técnicas de identificação molecular e a posterior comparação com banco de dados moleculares com amostras validadas por materiais testemunhos contribuiu para facilitar a identificação das espécies desta família, especialmente daquelas espécies crípticas ou pouco amostradas, ou ainda corrigir identificações realizadas de maneira errônea e também contribuir na descrição de novas espécies (Yamada et. al., 2019; Dalebout et al., 2004; Dalebout et al., 1998). Alguns casos notáveis no qual a genética foi preponderante na determinação da espécie, ou mesmo, na sua validação (ou revalidação) taxonômica e mesmo na descrição, são a baleia-bicuda-de-Perrin (*Mesoplodon perrini*), a baleia-bicuda-de-Longman (*Indopacetus pacificus*) e recém descrita baleia-bicuda-mínima (*Berardius minimus*) (Dalebout *et al*. , 2002; Yamada et. al., 2019)

Tabela 1. Lista das espécies (nomes populares) da ordem Cetartiodactyla ocorrentes no litoral brasileiro e seus respectivos status de conservação conforme IUCN.

IUCN: Dados Insuficientes (DD); Pouco Preocupante (LC); Quase Ameaçada (NT);Vulnerável (VU) e Em Perigo (EN).

Espécie	Nome comum	Categoria de Ameaça na IUCN
Família Balaenidade		
Eubalaena australis	Baleia-franca-austral	LC
Família Balaenopteridae		
Balaenoptera acutorostrata	Baleia-minke-anão	LC
Balaenoptera bonaerensis	Baleia-minke-antártica	NT
Balaenoptera borealis	Baleia-sei	EN
Balaenoptera brydei	Baleia-de-Bryde	DD
Balaenoptera musculus	Baleia-azul	EN
Balaenoptera physalus	Baleia-fin	VU

Espécie	Nome comum	Categoria de Ameaça na IUCN
Balaenoptera omurai	Baleia-de-Omura	DD
Megaptera novaeangliae	Baleia-jubarte	LC
Família Physeteridae		
Physeter macrocephalus	Cachalote	VU
Família Kogiidae		
Kogia breviceps	Cachalote-pigmeu	DD
Kogia sima	Cachalote-anão	DD
Família Ziphiidae		
Berardius arnuxii	Baleia-bicuda-de-Arnuxi	DD
Hyperoodon planifrons	Baleia-bicuda-nariz-de-garrafa	LC
Mesoplodon densirostris	Baleia-bicuda-de-Blainvillei	DD
Mesoplodon europaeus	Baleia-bicuda-de-Gervais	DD
Mesoplodon grayi	Baleia-bicuda-de-Gray	DD
Mesoplodon hectori	Baleia-bicuda-de-Hector	DD
Mesoplodon layardii	Baleia-bicuda-de-Layard	DD
Mesoplodon mirus	Baleia-bicuda-de-True	DD
Ziphius cavirostris	Baleia-bicuda-de-Cuvier	LC
Família Delphinidae		LC
Cephalorhynchus commersonii	Golfinho-de-Commersoni	LC
Delphinus delphis	Golfinho-comum	LC
Feresa attenuata	Orca-pigméia	DD
Globicephala macrorhynchus	Baleia-piloto-de-peitorais-curtas	LC
Globicephala melas	Baleia-piloto-de-peitorais-longas	LC
Grampus griseus	Golfinho-de-Risso	LC
Lagenodelphis hosei	Golfinho-de-Fraser	LC
Lagenorhynchus australis	Golfinho-de-Peale	DD
Lissodelphis peronii	Golfinho-liso-do-sul	LC
Orcinus orca	Orca	DD
Peponocephala electra	Golfinho-cabeça-de-melão	LC
Pseudorca crassidens	Falsa-orca	NT
Sotalia fluviatilis	Tucuxi	DD
Sotalia guianensis	Boto-cinza	NT
Stenella attenuata	Golfinho-pintado-pantropical	LC
Stenella clymene	Golfinho-de-Clymene	LC
Stenella coeruleoalba	Golfinho-listrado	LC
Stenella frontalis	Golfinho-pintado-do-Atlântico	LC
Stenella longirostris	Golfinho-rotador	LC
Steno bredanensis	Golfinho-de-dentes-rugosos	LC
Tursiops truncatus	Golfinho-nariz-de-garrafa ou Boto	LC
Família Phocoenidae		
Phocoena dioptrica	Boto-de-óculos	LC
Phocoena spinipinnis	Boto-de-Burmeister	NT
Família Pontoporiidae		
Pontoporia blainvillei	Toninha	VU
Família Iniidae		
Inia geoffrensis	Boto-vermelho; Boto-cor-de-rosa	EN
Inia araguaiaensis	Boto-do-Araguaia	Não avaliada

Dessa forma, a identificação inequívoca das espécies depende, muitas vezes, da coleta e análise de material osteológico (e.g. Pinedo *et al.,* 2002) ou ainda da identificação molecular (e.g. Hebert *et al*., 2003; Sholl *et al*., 2013; Cypriano-Souza *et al*., 2016). Contudo, quando partes

diagnósticas como o crânio são perdidas e a coloração original já não está mais presente, a identificação anatômica é praticamente impossível. Este foi justamente o cenário apresentado no primeiro registro da baleia-de-Omura (*Balaenoptera omurai*) para a costa do Brasil e Atlântico Sul Ocidental, realizado por Cypriano-Souza *et al*., (2016). Os autores só conseguiram chegar à identificação inequívoca do espécime após gerar informações de três segmentos mitocondriais (região controladora, citocromo b e citocromo oxidase subunidade I) para posterior comparação com sequências destes mesmos seguimentos depositadas em bancos de dados moleculares. A partir desse resultado, demonstrou-se a potencial diversidade críptica dos cetáceos na costa brasileira que se encontra "escondida" em decorrência da falta de uso rotineiro das técnicas moleculares como ferramentas diagnósticas.

Diversos marcadores moleculares auxiliam na identificação das espécies e na caracterização da estruturação genética de suas populações (Avise 2004). A utilização de determinadas regiões do DNA mitocondrial (DNAmt), como a região controladora e citocromo b, já vem sendo realizado desde a década de 90 no controle da caça e comércio ilegal de produtos oriundos de cetáceos (e.g. Baker & Palumbi 1994, Sholl *at al*., 2008, Siciliano *et al*., 2018). Contudo, uma abordagem diferente para a identificação de espécies foi trazida por Hebert *et al*., (2003), que recomendaram o uso do gene Citocromo C Oxidase Subunidade I (*COI*) para determinar espécies em filos de todo o reino animal. Por meio da utilização desta técnica foi possível identificar carnes processadas de mamíferos marinhos comercializados de forma ilegal e ainda contribuir na identificação de duas espécies de golfinhos ameaçados (Chang *at al.,* 2014, Kandu *at al.,* 2019).

O COI é parte de um complexo gênico codificante de proteínas transmembranas, envolvidas no transporte elétrico e catálise da cadeia respiratória de organismos eucariotos. O uso do *COI* ficou conhecido como DNA Barcoding ou "código de barras de DNA" (Hebert *et al*., 2003). A técnica é um método relativamente preciso para identificação das espécies e pode ser muito útil em casos específicos, quando outras técnicas de identificação taxonômica não são efetivas (e.g. quando um cetáceo morto, em avançado estado de decomposição e sem partes diagnósticas como crânio, é encontrado na costa). A utilização do *COI* tem como fundamentos (i) a formação de uma base de dados de sequências geradas a partir de amostras com identidade conhecida; (ii) a identificação de caracteres diagnósticos a nível específico; e (iii) a comparação de sequências obtidas a partir de amostras

desconhecidas (p.ex. ovos apreendidos, carne de caça, fragmentos de peles ou ossos, alimentos industrializados) com a base de dados.

No entanto, a precisão do DNA Barcoding tem sido questionada para alguns táxons, inclusive para os cetáceos (e.g. Viricel & Rosel 2011), principalmente em função do grau de separação entre variação genética intraespecífica e interespecífica do *COI*, chamado de *gap barcoding* (Meyer & Paulay, 2005). Em alguns grupos de cetáceos, não há essa separação bem definida, ocorrendo sobreposição entre variação genética intraespecífica e interespecífica (Viricel & Rosel, 2011). É importante salientar que a crítica do uso do *COI* em função de problemas com o *gap barcoding* é válida apenas para alguns cetáceos de determinadas famílias porque foram os únicos testados (Viricel & Rosel, 2011), não tendo sido avaliado algumas espécies, incluindo as endêmicas da América do Sul e do Brasil (*e.g.* Skueresky, 2014; Falcão *et al.,,* 2017).

Nesse contexto, esta abordagem combinada do uso do *COI* com demais segmentos do DNAmt, aliada à comparação de material morfológico *voucher* (e.g. crânios e dados de anatomia externa) torna-se um método mais robusto e muito útil que deve ser testado para a correta identificação específica de cetáceos encontrados em avançado estado de decomposição na costa brasileira. Para tanto, a manutenção de um banco de dados atualizado (com a identificação morfológica aliada à molecular) permitiria a resolução de questões taxonômicas cotidianas através de um sistema rápido e eficiente de identificação de espécimes para diferentes fins, com base nos códigos de barra de DNA, o que seria feito por meio da comparação dos dados genéticos para esse marcador molecular obtido a partir destas carcaças dos cetáceos.

Desta forma, o presente estudo testou a eficiência do uso do Citocromo C Oxidase Subunidade I (*COI*) na identificação de espécimes de cetáceos encontrados na costa brasileira, os quais, na sua maioria, apresentam material testemunho como fonte alternativa de confirmação da identificação molecular. A partir dessas análises moleculares dos espécimes previamente identificados anatomicamente, pretendia-se a disponibilizar mais uma ferramenta para facilitar a identificação desses indivíduos.

Como objetivo secundário, o presente estudo pretendeu ampliar o conhecimento da variabilidade intraespecífica e interespecífica dos cetáceos encontrados na costa brasileira, com a geração de 152 novas sequências da região *COI* para banco de sistema BOLD, incluindo 33 das 47 espécies de cetáceos registrados na costa do Brasil.

REFERÊNCIAS

Ab'saber, Aziz Nacib. Litoral do Brasil. São Paulo: Meta Livros, 2001. 287 p.

Amaral AZ, Jablonski S (2005) Conservation of Marine and Coastal Biodiversity in Brazil. Coxazanservation Biology, [s.l.], v. 19, n. 3, p.625-631

Avise JC (2004). Molecular markers, natural history and evolution. Chapman & Hall

Baker CS, Palumbi SR (1994) Which Whales Are Hunted? A Molecular Genetic Approach to Monitoring Whaling. Science, [s.l.], v. 265, n. 5178, p.1538-1539. American Association for the Advancement of Science (AAAS)

BRASIL. Ministério do Meio Ambiente. Portaria MMA nº 444, de 17 de dezembro de 2014.

Chang CH, Yao CJ, Yu HY, Liao YC, Jang-Liaw N H, Tsai C L, Shao KT (2014) A molecular forensic method for identifying species composition 339 of processed marine mammal meats. Journal of Forensic and Legal Medicine, 340 23, 65-69.

Chaves PB, et al (2011) DNA barcoding meets molecular scatology: short DNAmt sequences for standardized species assignment of carnivore noninvasive samples. Molecular Ecology Resources, [s.l.], v. 12, n. 1, p.18-35

CyprianoALS et al (2016) Rare or cryptic? The first report of an Omura's whale (Balaenoptera omurai) in the South Atlantic Ocean. Marine Mammal Science, [s.l.], v. 33, n. 1, p.80-95

Dalebout ML, Baker CS, Mead JG, Cockcroft VG, Yamada TK (2004) A comprehensiveand validated molecular taxonomy of beaked whales, family Ziphiidae. Journal of Heredity 95:459–473 DOI 10.1093/jhered/esh054

Dalebout ML, Mead JG, Baker CS, Baker NA, Van Helden AL (2002) A new species of beaked whale *Mesoplodon perrini* sp. n. (Cetacea: Ziphiidae) discovered through phylogenetic analyses of mitochondrial DNA sequences. Marine Mammal Science 18

Dalebout ML, van Helden A, Van Waerebeek K, Baker CS (1998) Molecular genetic identification of southern hemisphere beaked whales (Cetacea: Ziphiidae). Mol Ecol 7:687–69

Falcão LHO, Campos AS, Freitas, JEP, Furtado-Neto MAA, Faria VV (2017) Molecular identification of cetaceans from the West Atlantic using the E3-I5 region of *COI*. Genetics and Molecular Research 16 (2): gmr16029607

Hebert PDN et al (2003) Biological identifications through DNA barcodes Proceedings Of The Royal Society B: Biological Sciences, [s.l.], v. 270, n. 1512, p.313-321. The Royal Society

IUCN 2019. The IUCN Red List of Threatened Species. Version 2019-2. <http://www.iucnredlist.org>. Acesso em 15 de março de 2019.

Kundu S, Sharma G, Balakrishnan S, Tyagi K, Chandra K, Kumar V (2019) DNA barcoding identified two endangered dolphins: threats on living aquatic mammals in India, Mitochondrial DNA Part B, 4:1, 77-80, DOI: 10.1080/23802359.2018.1536467

Longo LL, Amado F, Gilberto M (2014) O conhecimento da fauna marinha bentônica brasileira através dos tempos. História, Ciências, Saúde-manguinhos, [s.l.], v. 21, n. 3, p.995-1010, FapUNIFESP (SciELO)

Meyer CP, Paulay G (2005) DNA Barcoding: Error Rates Based on Comprehensive Sampling. Plos Biology, [s.l.], v. 3, n. 12, p.422-423, Public Library of Science (PLoS).

Palumbi SR, Kessing, Bailey D(1991) Population biology of the trans-arctic exchange: DNAmt sequence similarity between pacific and atlantic sea urchins. Evolution, [s.l.], v. 45, n. 8, p.1790-1805

Pinedo MC, Barreto AS, Lammardo MP, Andrade ALV, Geracitano L (2002) Northernmost records of the spectacled porpoise, Layard's beaked whale, Commerson's dolphin, and Peale's dolphin in the southwestern Atlantic Ocean. Aquat. Mamm., v. 28, n. 1, p. 32-37

Prado HF, Mattos PH, Silva KG, Secchi ER (2016) Long-Term Seasonal and Interannual Patterns of Marine Mammal Strandings in Subtropical Western South Atlantic. PLoS ONE 11(1) doi: 10.1371/journal.pone.0146339

Reeves RR, Brent SS, Clapham PJ, Powell JÁ, Folkins P A (2002) National Audubon Society guide to marine mammals of the world. Chanticleer Press, New York, NY

Sholl TGC et al (2008) Taxonomic identification of dolphin love charms commercialized in the Amazonian region through the analysis of cytochrome b DNA. Journal Of The Marine Biological Association Of The United Kingdom, [s.l.], v. 88, n. 06, p.1207-1210

Skueresky NS (2014) Código de barra de dna de cetáceos brasileiros: Uma ferramenta molecular como subsídio para a construção de uma rede nacional que busca a conservação de espécies. 2014. 35 f. TCC (Graduação) - Curso de Ciência BiolÓgicas, Universidade do Vale do Rio dos Sinos, São Leopoldo

Siciliano S et al (2018) Dolphins, Love and Enchantment: Tracing the Use of Cetacean Products in Brazil. Frontiers in Marine Science, v.5, Article 107, 10p

Viricel A, Rosel PE (2011) Evaluating the utility of cox1 for cetacean species identification. Marine Mammal Science, [s.l.], v. 28, n. 1, p.37-62.

Yamada TK, Kitamura , Abe S,Tajima Y, Matsuda A, Mead JG, Matsuishi TF (2019) Description of a new species of beaked whale (Berardius) found in the North Pacific. *Scientific Reports*, 9 (1) DOI: 10.1038/s41598-019-46703-w

DNA Barcoding helping the identification of Brazilian cetaceans

ABSTRACT

Cytochrome C Oxidase Subunit I (*COI*) is the mitochondrial gene recognized as the DNA Barcoding and used to identify different animal species. In the case of cetaceans (porpoises, whales and dolphins), the *COI* can assist in more accurate identification and/or confirm the morphological identity of specimens found in advanced state of decomposition on the Brazilian coast, without evident external diagnostic characters. This study aimed to test the efficiency of the *COI* in the specific identification of most of the 47 species of cetaceans off the Brazilian coast and to generate their respective sequences for depositing them in public sequence databases, in order to subsidize new work by researchers from any country. Part of the world. To this end, the *COI* was amplified to 152 specimens collected by 13 Brazilian institutions and sent for analysis. The comparison between the sequences obtained and those of cetacean species previously deposited in GenBank was made by the Blast tool, which is based on the similarity between the sequences, which was performed on the *Barcode Of Life Data System* platform. In addition, most samples had a specimen deposited in a scientific collection as witness material, or specimen images, which allowed their *a priori* morphological identification in cases where morphological identification diverged from molecular identification. As a result, *COI* sequences were obtained from 33 species, which represent 70% of the 47 cetacean fauna recorded for the Brazilian coast. Of the 152 sequences analyzed, the *COI* gene was inefficient in identifying only two species: *Stenella coeruleoalba and S. clymene* due to the absence of the so-called “gap barcode”, i.e., the absence of well-established inter and intraspecific limits. These results suggest that barcode DNA was efficient in identifying the great majority of cetacean specimens (~93%) studied. However, for species of Delphinidae family the identification should be integrated with other methods, such as Cyt b, nuclear DNA and morphological character analysis whenerver possible.

Keywords: cetaceans, stranding, morphological, identification, molecular

1. INTRODUCTION

Brazil has the most extensive coastline in the world with almost 8,000 km of coastline (Ab'Saber 2001). Despite their representativeness for the Neotropics in terms of biological diversity, the increasing anthropogenic alterations in the marine and coastal regions have disrupted these environments, threatening the survival of many species and even entire communities (Amaral & Jablonski 2005).

The knowledge about the existing diversity in the coastal and oceanic regions of Brazil is essential to understand the functioning of the different associated ecosystems, as well as to ensure the sustainable use and conservation of their living resources (e.g. Longo & Amado Filho 2014). The current knowledge about the marine communities in these regions is still insufficient to guarantee their conservation, especially in view of the growing economic interest in exploring these areas, even with the implementation of important research programs in the Brazilian oceanic regions in the last decades (e.g. REVIZEE, Archipelago Program and Oceanic Islands). Cetaceans (i.e. whales and dolphins) are one of the taxonomic groups that have lack of essential information mainly regarding their ecological role in the ocean ecosystem as well as for the establishment of effective plans for conservation (Zerbini et al. 2004; Ott et al. 2009 and Siciliano et al. 2012).

Currently, there are confirmed records of 47 cetaceans in Brazil of 90 species described in the world (Pinedo et al. 1992; Zerbini et al. 1997; Zerbini & Secchi 2001, Pinedo et al. 2002; Bastida et al. 2007; ICMBio 2011b; MMA 2014, Hrbek et al. 2014 and Cypriano-Souza et al. 2016), which six of these are considered as threatened and, at least, 29 categorized as data deficient (MMA 2014; Hrbek et al. 2014; Cypriano-Souza et al. 2016 and; IUCN 2019). Moreover, Brazil has two species of mustelids and two of manatees, besides records of eight pinniped species (Antarctic seals, fur seals, sea lions and elephant seals, for review see Milmann et al. 2019).

The records of this remarkable biodiversity of aquatic mammals are usually based on specimens found dead or stranded along the Brazilian coast, many times related to anthropogenic actions (Van Bressem et al. 2007; Lemos et al. 2013; Ott et al. 2013). In this context, Brazilian stranding network of Aquatic Mammals (REMAB) was created in 2011. This initiative includes other four regional networks: the Northeast Aquatic Mammals Stranded Network (REMANE), the Southern Aquatic Mammals Stranded Network (REMASUL), the Northern Aquatic Mammals Stranded Network (REMANOR) and the Southeast Aquatic Mammals (REMASE). REMAB is coordinated by the National

Aquatic Mammals Center (*Centro de Mamíferos Aquáticos* - CMA) and operates throughout the national territory. The purpose of this network is to exchange of information obtained during field work, research and monitoring, among institutions that work with aquatic mammals in Brazil.

However, the real richness of aquatic mammal species along the Brazilian coast could be underestimated, mainly due to factors: 1) the difficult to identify the species at the sea only by few exposed parts of the body and due to the morphological similarity between some species, and 2) the advanced stage of decomposition frequently presented by the cetacean specimens found along the coastline (Sholl et al. 2013). Therefore, unambiguous species identification often depends on the analysis of osteological material collected (e.g. Pinedo et al. 2002) or molecular identification (e.g. Sholl et al. 2013; Siciliano et al. 2016 and Cypriano-Souza et al. 2016). It is important to highlight that when diagnostic parts, such as the skull, are lost and the original color of the species is no longer present, anatomical identification is virtually impossible. This was precisely the scenario presented in the first record of the Omura's whale (*Balaenoptera omurai*) to the coast of Brazil and the Southwestern Atlantic, published by Cypriano-Souza et al. (2016).

The authors were only able to reach unambiguous identification of the specimen after generating information from three mitochondrial segments (control region, cytochrome b and cytochrome oxidase subunit I) and subsequent comparison with sequences of these same segments deposited in molecular databases. Based on this result, it was demonstrated the potential of cryptic diversity of cetaceans in the Brazilian coast that is "hidden" due to the lack of use of molecular techniques as diagnostic tool for these taxa (Sholl et al. 2008 and Cypriano-Souza et al. 2016). Same situation occurred when Hrebk et al. (2014) and afterwards Siciliano et al. (2016) found great molecular divergence in mtDNA genes to support the split of the Brazilian river dolphin of the genus *Inia* in two species: *I. geoffrensis* and the new species *I. araguaiaensis,* which is the only cetacean species endemic to the Brazilian waters.

This issue is well represented in the Ziphiidae family, commonly called beaked whales, with rare strandings on the Brazilian coast, even in regions with a long history of sampling (e.g. Prado et al 2016, Vianna et al. 2016). The discreet behavior of these whales, which includes slow movement, little exposure on the surface and distance from vessels makes information about this family even more difficult to obtain (Reeves et al. 2002, Dalebout et al 1998). In general, the identification of species is carried out, primarily, by analyzing the morphology of the skull, position and shape of the stranded

specimens' teeth, however, the correct identification depends, among other things, on the decomposition stage in which the carcasses reach up to to the coast and also the maturity of the stranded animal, and it is not uncommon for erroneous identifications to occur (Dalebout et al. 1998). In addition, the difficulty of identifying diagnostic characters, especially among the six species that make up the Mesoplodon genus, and the strong overlap of the species' area of occurrence, make identification even more complex (Dalebout et al. 1998). Thus, in many situations, the use of morphological characters alone does not provide sufficient information for identification at the species level.

In this context, the inclusion of molecular identification techniques and the subsequent comparison with a molecular database with samples validated by testimony materials contributed to facilitate the identification of species in this family, especially those cryptic or poorly sampled species, or to correct identifications carried out in a erroneous and also contribute to the description of new species (Yamada et al. 2019, Dalebout et al. 2004, Dalebout et al. 1998). Some notable cases in which genetics was preponderant in determining the species, or even in its taxonomic validation (or revalidation) and even in the description, are the Perrin's beaked whale (Mesoplodon perrini), the beaked whale -Longman (Indopacetus pacificus) and newly described minimal-beaked whale (Berardius minimus).

In order to stablish the correct specific identification of these mammals, which are frequently found stranded/dead and usually in advanced stage of decomposition, DNA barcoding becomes a very useful tool for such identifications (Hebert et al. 2003 and Alfonsi et al. 2013). This powerful tool of biological identification, called "DNA barcodes" (DNA barcoding), is based on a fragment Cytochrome C Oxidase I (COI) of the mitochondrial genome to determine individuals at the species level (Hebert et al. 2003).

Alfonsi et al. (2013), suggested that DNA barcoding could be useful for the monitoring of marine mammal strandings at three levels: i) by providing a confirmation or an additional degree of taxonomic determination of rare species identified by field researchers, mainly in uncommon stranding events of exotic or deep living species (Thompson et al. 2012); ii) by helping the identifications at species level when it is not possible to identify the animal by the external morphology due to highly degraded carcasses or even when morphology-based identification only reaches the genus or family levels, due to incomplete skeleton or skull, and iii) by offering intraspecific genetic variability, which allows genetic structure analysis, and eventually monitoring population movements (Pauls et al. 2012).

However, there are criticisms on DNA barcoding studies on cetaceans due to application of a single gene, which eventually does not have sufficient resolution to settle on inter and intraspecific positions (Amaral et al. 2007; Wiemers & Fiedler 2007; Viricel & Rosel 2012; Alfonsi et al. 2013 and Sholl et al. 2013). Thus, it has been suggested that the complementary use of multiple loci of mitochondrial gene, for example, cytochrome b and control region (Sholl et al., 2013), or even, if it is possible, the comparison with osteological material for the correct identification at species level.

In this context, the integrative taxonomy, which includes genetic and morphological/anatomical diagnoses (Padial et al. 2010), is the best option to properly evaluate the biodiversity. There are few studies integrating *COI* as a taxonomic marker and morphology, in order to identify cetaceans species (Amaral et al. 2007, Viricel & Rosel, 2012 and Alfonsi et al. 2013), and fewer studies including morphological voucher material, such as skulls, to compare with *COI* results, probably due to the need of a great number of cetacean species with DNA samples and bones collected.

According to Galimberti et al. (2015) it is clear that there is a hidden biodiversity within the mammal record. The BOLD System (http://www.boldsystems.org) has been barcoded at the end of May 2015 about 2850 mammal species, at least 300 unnamed clusters (i.e. not assigned taxonomic rank). Currently, in the MammaliaBoL, there are, recognized on approximately, 3,587 species with barcodes, which 75 are identified as cetaceans.

Recently, Falcão et al. (2017) published a barcoded DNA study on marine mammal species from Brazil and Canada. However, this study covered only a small portion of the Northeast Brazilian coast and only four individuals of four different species.

In this context, we evaluate the contributions of the DNA barcoding for the monitoring of the cetacean biodiversity along the coast of Brazil and its inner waters. Based on the establishment of a consortium of 13 institutions from the Brazilian stranding network , included in the project “Tetrapoda DNA Barcodes: Building an Integrated Network DNA barcoding of amphibians, reptiles, birds and mammals”, tissue samples were collected from stranded cetaceans and few biopsies taken from live animals along the Brazilian coast. Moreover, most of DNA samples have voucher material deposited in scientific collections (e.g. skull and skeletons), allowing the comparison between morphological and molecular identifications with *COI* gene.

This study evaluated the quality and reproducibility of the cetacean taxonomic identification performed by the consortium field correspondents, by identifying degraded carcasses, determining intraspecific variations for some dolphin species and by evaluating the hypothesis that *COI* could be an efficient molecular marker to support species limits (Taylor et al. 2018 and Hebert et al. 2003). Here, we discuss the results under the light on the method validation and its potential inconsistencies due to morphological and molecular mismatches (Viricel & Rosel 2011).

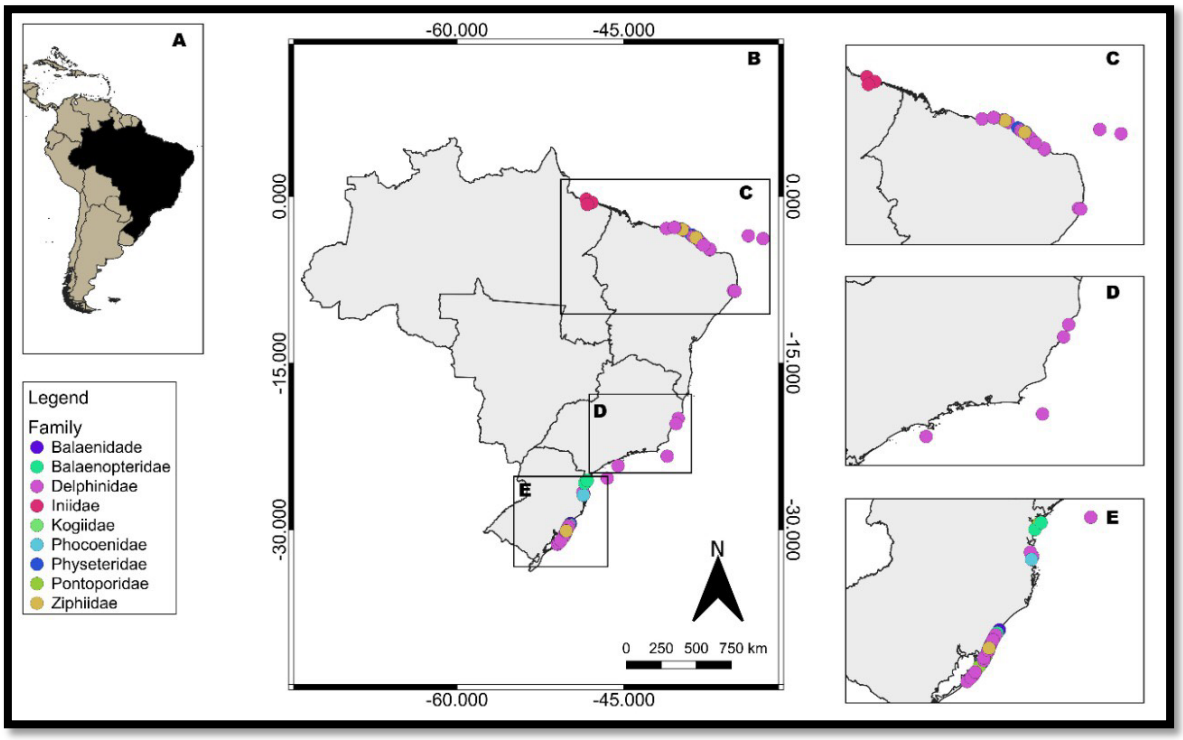
2. MATERIAL AND METHODS

2.1 Samples

A collaborative stranding network of 13 research institutions working along the Brazilian coast and inner waters obtained tissue samples from 145 cetacean carcasses, collected during beach surveys from 1989 to 2018 in four regions: south, southeast, northeast and north (see Fig 1 and Table 1). Additionally, we also included seven biopsied samples from oceanic waters retrieved during cetacean-sighting cruises. This samples were collected in surrounded waters of the Saint Paul's Rocks (00°56'S; 29°22'W) and Campos and Santos Basins (from 21°40′S to 27°00'S).

All specimens sampled have additional material information, such as skull and skeleton, or even photographs records deposited in scientific collections. The exception is the samples from biopsied animals that only have images used for photo identification by onboard specialists. This is the first cetacean barcoding study that includes voucher material, allowing reproducibility of species identification made by for field corresponds.

Fig. 1 Sampling sites of the 152 cetacean specimens analyzed along the Brazilian coast. The specimens were categorized by family.



The specimens were firstly identified by experienced researchers or trained field correspondents, following the standard procedures suggested by the American Society of Mammalogists published in 1987 in the protocol *Acceptable Field Methods in Mammalogy: Preliminary Guidelines Approved by the American Society of Mammalogists* (ad hoc Committee on Acceptable Field Methods in Mammalogy 1987, http://mammalogy.org/uploads/committee_files/ACUC1987.pdf) and by Geraci & Lounsbury (1999), both protocols were adopted by Brazilian stranding marine mammal network. The identification of each specimen was through external morphology or skull anatomy diagnosis. Moreover, information related to total length, sex and the condition of each carcass, including its degree of decomposition (Geraci & Lounsbury 1993) were also recorded whenever was possible. All tissue samples collected were stored in 70% alcohol or DMSO saturated with NaCl 20% and sent to Laboratory of Genetics and Molecular Biology (LGBM) at the University of Vale dos Rio dos Sinos.

DNA was extracted using a phenol chloroform protocol, and the quality and concentration of DNA were verified in 1% agarose gel electrophoresis. The concentrations of genomic DNA were estimated with Nanodrop UV spectrophotometry (Thermo Scientific Wilmington, DE). The DNA samples were diluted in deionized water until reaching a concentration of approximately 100 ng/ul when necessary.

We amplified COI fragments with polymerase chain reactions (PCRs) by applying three primer pairs, VF1d, VF1i,VR1 and VR1d, which targeted approximately 800 bp (see Supplementary Material 1 for details). PCR results were verified on 1% agarose gels stained with GelRed (Biotium, Hayward, CA, USA). PCR products were purified using Shrimp Alkaline Phosphatase (SAP) and exonuclease I (New England Biolabs), following the manufacturer's recommendation. Amplicons were sequenced in both directions using universal primers (M13-FP and M13R-pUC, see Supplementary Material 1) at Macrogen Inc. (Seoul, Korea).

2.2 Analysis

We manually selected only the high-quality COI sequences, with high and clear peaks for each nucleotide, based on the observation of electropherograms with ChromasPro 2.6.6 (http://www.technelysium.com.au). Following the DNA Barcoding criteria, to maximize the accuracy of this method, all deposited sequences are associated with samples linked to institutional voucher material (e.g. skulls), and furthermore, the samples contain data regarding the date and place of collection and primers used in PCR (Hanner, 2009).

A total of 152 consensus sequences were automatically aligned (with minor manual correction) in ClustalW implemented in MEGA 7 (Tamura et al. 2016), with later edition in BioEdit 5.0.9 (Hall 1999). After the alignment, we compared the COI sequences with those available in GenBank (www.ncbi.nlm.nih.gov) and Boldsystems (www.boldsystems.org) using the Basic Local Alignment Search Tool (BLAST) (blast.ncbi.nlm.nih.gov). The molecular identification suggested by GenBank and BOLD was based on the percentage of similarity among sequences. Here, we highlight that for those species with no COI sequence available in the databases the similarity outcome was that related to the closest taxon or no result was returned, when the species does not exist in the database.

The cases of molecular-morphology mismatch, due to incongruence between the species identification suggested by COI sequences (from GenBank or BOLD systems) and the morphological identification informed by the field collaborating institutions, were further investigated. Whenever it was possible, a revision of the species identification was conducted by requesting skull or carcass images to the field correspondents. External traits or anatomical diagnostic characters of the skull were analyzed to confirm the identification. In cases of uncertainties, marine mammal specialists were consulted. This

procedure was conducted for the species of the genus Balaenoptera, Eubalaena, Stenella and Pseurdorca. Moreover, field notes on the specimen collected were also double-checked in the catalog book of the scientific collection, mainly on its decomposition stage (including images from the sampling), which could explain some of the mismatch results (see discussion).

Genetic divergences (dA) (intraspecific and interspecific) were calculated using the K2P model (Kimura 1980) for those species that do not have gap barcode to establish the interval of genetic divergence between them (e.g some delphinideos). According to Hebert et al. (2003), the lower limit for genetic divergence (dA) between species is around 3%.This limit was considered in the present study as the lower level for cetacean species delimitation using *COI* marker.

To test the hypothesis that all *COI* sequences correspond to a monophyletic cluster and belong to the same cetacean species, we perform two types of phylogenetic reconstruction: a Neighbor-Joining tree (NJ) using the Kimura 2-parameter (K2P) model implemented in the software MEGA 7 (Tamura et al. 2016); and the maximum likelihood tree (ML) recovered in the program RAxML 8.2 (Stamatakis 2014). For the latter we used GTR+4G as evolutionary model estimated by the jmodeltest2 (Darriba et al. 2012). To infer the phylogenies with the different approaches we assembled and aligned the 152 *COI* consensus sequences generated with the 72 available on the BOLD platform, totalizing 224 *COI* sequences representing 33 cetacean taxa. The species *Hippopotamus amphibius* (GBMA2411-09), available on the Bold platform was used as outgroup.

3. RESULTS

We recovered *COI* sequences of 644 to 847 bp obtained from 152 samples representing 33 species. The recorded species were distributed in nine families of cetaceans (Table 1), the molecular identification was in accordance with external morphology identification made by the correspondents of the stranding network in 93% of the specimens (Table 2).

Table 1. Species, popular name and number of sequences

Specie	Popular name	Nº sequences
Delphinidae		
common dolphin	*Delphinus delphis*	9
common bottlenose dolphin	*Tursiops truncatus*	9
rough-toothed dolphin	*Steno bredanensis*	4
Atlantic spotted dolphin	*Stenella frontalis*	4
striped dolphin	*Stenella coeruleoalba*	5

Specie	Popular name	Nº sequences
Clymene dolphin	*Stenella clymene*	6
spinner dolphin	*Stenella longirostris*	5
pantropical spotted dolphins	*Stenella attenuata*	4
Fraser's dolphin	*Lagenodelphis hosei*	7
Guiana dolphin	*Sotalia guianensis*	7
long-finned pilot whale	*Globicephala melas*	1
false killer whale	*Pseudorca crassidens*	8
Risso's dolphin	*Grampus griseus*	3
melon-head whale	*Peponocephala electra*	4
Kogiidae		
pygmy sperm whale	*Kogia breviceps*	2
dwarf sperm whale	*Kogia sima*	6
Physeteridae		
sperm whale	*Physeter macrocephalus*	9
Pontoporidae		
franciscana dolphin	*Pontoporia blainvillei*	12
Phocoenidae		
spectacled porpoise	*Phocoena dioptrica*	1
Burmeister's porpoise	*Phocoena spinipinnis*	3
Iniidae		
Araguaian River dolphin	*Inia araguaiaensis*	3
Amazon River	*Inia geoffrensis*	2
Ziphiidae		
Arnoux's beaked whale	*Berardius arnuxii*	1
dolphin Gevais' beaked whale	*Mesoplodon europaeus*	1
Cuvier's beaked whale	*Ziphius cavirostris*	5
Balaenidae		
southern right whale	*Eubalaena australis*	5
Balaenopteridae		
humpback whale	*Megaptera novaeangliae*	9
common minke whale	*Balaenoptera acutorostrata*	4
Antarctic minke whale	*Balaenoptera bonaerensis*	2
sei whale	*Balaenoptera borealis*	3
Bryde's whale	*Balaenoptera brydei*	5
fin whale	*Balaenoptera physalus*	1
Omura's whale	*Balaenoptera omurai*	1

Table 2. Sample: sample identification, Instituition: acronym of the institution responsible for collecting, MID: Morphological identification, NCBI/GenBank ID: molecular identification suggested by NCBI, BOLD ID: molecular identification suggested by BOLD SYSTEM, Locality: sampling site.

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
GEMARS469	GEMARS	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	99.05	99.04	Mostardas- RS
GEMARS 1042	GEMARS	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*			Palmares do Sul- RS
GEMARS 1468	GEMARS	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	99.15	99.13	
LEC#119	UFPR	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	*Balaenoptera acutorostrata*	*98.90*	98.9	Matinhos- PR
GEMARS 1269	GEMARS	*Balaenoptera bonaerensis*	*Balaenoptera bonaerensis*	*Balaenoptera bonaerensis*	100	100	Torres- RS
MINK FURG	FURG	*Balaenoptera bonaerensis*	*Balaenoptera bonaerensis*	*Balaenoptera bonaerensis*	99.58	99.65	Rio Grande- RS
CT1050 F63	FURG	*Balaenoptera borealis*	*Balaenoptera brydei*	*Balaenoptera brydei*	99.23	99.22	Rio Grande- RS
CT1130 F61	FURG	*Balaenoptera borealis*	*Balaenoptera borealis*	*Balaenoptera borealis*	100	99.35	Rio Grande- RS
CT9,5M F62	FURG	*Balaenoptera borealis*	*Balaenoptera borealis*	*Balaenoptera borealis*	99.70	99.7	Rio Grande- RS
MPEG 39691	GEMAM	*Balaenoptera borealis*	*Balaenoptera borealis*	*Balaenoptera borealis*	100	100	Viseu - PA
GEMARS 1406	GEMARS	*Balaenoptera brydei*	*Balaenoptera brydei*	*Balaenoptera brydei*	100	100	Mostardas- RS
GEMARS 1425	GEMARS	*Balaenoptera brydei*	*Balaenoptera brydei*	*Balaenoptera brydei*	99.71	99.81	Quintão- RS
GEMARS 1694	GEMARS	*Balaenoptera brydei*	*Balaenoptera brydei*	*Balaenoptera brydei*	99.54	99.39	Cidreira- RS
LEC#154	UFPR	*Balaenoptera brydei*	*Balaenoptera brydei*	*Balaenoptera brydei*	100	*99.34*	Ilha do Mel- PR
02C0100/421	AQUASIS	*Balaenoptera omurai*	*Balaenoptera omurai*	*Balaenoptera omurai*	100.00	100	São Gonçalo- CE
GEMARS 0825	GEMARS	*Balaenoptera physalus*	*Balaenoptera physalus*	*Balaenoptera physalus*	97.81	97.79	
GEMARS 1155	GEMARS	*Berardius arnuxii*	*Berardius bairdii*	*Berardius bairdii*	99.70	99.69	Pinhal- RS
#57	Fiocruz-GEMM-Lagos	*Delphinus delphis*	*Delphinus delphis*	*Delphinus delphis*	99.56	99.69	Bacia de Campos
BC 04 (RJ Salvatore)	Fiocruz-GEMM-Lagos	*Delphinus delphis*	*Delphinus delphis*	*Delphinus delphis*	99.68	99.67	Bacia de Campos

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
GEMARS 0221	GEMARS	*Delphinus delphis*	*Delphinus delphis*	*Delphinus delphis*			Mostardas
GEMARS 0419	GEMARS	*Delphinus delphis*	*Delphinus delphis*	*Delphinus delphis*	99.68	99.68	Imbé
PA288	IO -USP	*Delphinus delphis*	*Delphinus delphis*	*Delphinus delphis*	100	99,84	
FURG 07 CT208	FURG	*Delphinus sp*	*Delphinus delphis*	*Delphinus delphis*	100	99,68	Rio Grande- RS
FURG 08	FURG	*Delphinus sp*	*Delphinus delphis*	*Delphinus delphis*	99	99,23	Rio Grande- RS
FURG 6 CT 197	FURG	*Delphinus sp*	*Delphinus delphis*	*Delphinus delphis*	98.32	98.57	Rio Grande- RS
Franca_RJ	Franca_RJ	*Eubalaena australis*	*Eubalaena australis*	*Eubalaena glacialis*	99	99,69	RJ
GEMARS 1456	GEMARS	*Eubalaena australis*	*Eubalaena australis*	*Eubalaena australis*	100	100	Capão da Canoa- RS
GEMARS 1467	GEMARS	*Eubalaena australis*	*Eubalaena australis*	*Eubalaena australis*	100	100	Torres- RS
MN60458	Museu nacional RJ	*Eubalaena australis*	*Eubalaena australis*	*Eubalaena glacialis*	100	100	RJ
FURG38	FURG	*Globicephala melas*	*Globicephala melas*	*Globicephala melas*	100	99,68	Rio Grande- RS
02C1812/588	AQUASIS	*Grampus griseus*	*Grampus griseus*	*Grampus griseus*	99.38	99.37	Icapuí-CE
G1236	GEMARS	*Grampus griseus*	*Grampus griseus*	*Grampus griseus*	99.22	99.52	Capão da Canoa- RS
GEMAM 070	GEMAM	*Grampus griseus*	*Grampus griseus*	*Grampus griseus*	100	99.83	Marapanim- PA
GEMAM 110	GEMAM	*Inia geoffrensis*	*Inia araguaiaensis*	*Inia araguaiaensis*	100	99.81	Salvaterra- PA
GEMAM 334	GEMAM	*Inia geoffrensis*	*Inia araguaiaensis*	*Inia araguaiaensis*	100	99.81	Soure- PA
GEMAM 368	GEMAM	*Inia geoffrensis*	*Inia araguaiaensis*	*Inia araguaiaensis*	100	99.81	Curuçá- PA
MPEG 42179	GEMAM	*Inia geoffrensis*	*Inia geoffrensis*	*Inia geoffrensis*			Curuçá- PA
MPEG 42180	GEMAM	*Inia geoffrensis*	*Inia geoffrensis*	*Inia geoffrensis*			Curuçá- PA
FURG 33	FURG	*Kogia breviceps*	*Kogia breviceps*	*Kogia breviceps*	100	99,37	
G1496	GEMARS	*Kogia breviceps*	*Kogia breviceps*	*Kogia breviceps*	99.68	99.68	Mostardas- RS
02C0511/703	AQUASIS	*Kogia sima*	*Kogia sima*	*Kogia sima*	100	99.83	Rio Grande- RS
02C0511/726	AQUASIS	*Kogia sima*	*Kogia sima*	*Kogia sima*	99.33	99.33	Itarema-CE
02C0512/585	AQUASIS	*Kogia sima*	*Kogia sima*	*Kogia sima*	99.33	99.32	
GEMARS 1311	GEMARS	*Kogia sima*	*Kogia sima*	*Kogia sima*	100		Imbé- RS
GEMARS 1407	GEMARS	*Kogia sima*	*Kogia sima*	*Kogia sima*	99.06	99.84	Cidreira- RS

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
GEMARS 1421	GEMARS	*Kogia sima*	*Kogia sima*	*Kogia sima*	100	99.85	Palmares do Sul- RS
02C0212/342	AQUASIS	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	100	99.83	Cascavel- CE
02C2512/389	AQUASIS	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	99.69	99.84	Amontada- CE
FURG 22	FURG	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*			Rio Grande- RS
FURG23	FURG	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	100	98.89	Rio Grande- RS
GEMARS 0467	GEMARS	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	100	99.85	Mostardas- RS
GEMARS 0488	GEMARS	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	100	99.85	Mostardas- RS
GEMARS 1453	GEMARS	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	100	99,84	Imbé- RS
GEMARS 0435	GEMARS	*Lagenodelphis hosei*	*Lagenodelphis hosei*	*Lagenodelphis hosei*	100	100	Cidreira- RS
02C0211/418	AQUASIS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	100	100	
02C0212/645	AQUASIS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	100	100	
GEMARS 0597	GEMARS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	99.84	99.84	Nova Tramandaí- RS
GEMARS 1409	GEMARS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	99.84	99.84	
GEMARS 1451	GEMARS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	100	99.85	Pinhal- RS
GEMARS 1683	GEMARS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	100	100	Cidreira- RS
GEMARS 1684	GEMARS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	100	100	Arroio Do Sal- RS
GEMARS 1685	GEMARS	*Megaptera novaeangliae*	*Megaptera novaeangliae*	*Megaptera novaeangliae*	100	100	Nova Tramandaí- RS
MN 84736		*Mesoplodon europaeus*	*Mesoplodon europaeus*	*Mesoplodon europaeus*	99	99,41	Quissamã- RJ
CT169 F45	FURG	*Orcinus orca*	*Pseudorca crassidens*	*Pseudorca crassidens*	100	100	Rio Grande- RS
02C1511/783	AQUASIS	*Peponocephala electra*	*Peponocephala electra*	*Peponocephala electra*	100	100	Cruz-CE
02C1511/784	AQUASIS	*Peponocephala electra*	*Peponocephala electra*	*Peponocephala electra*	100	100	Cruz-CE
02C1512/669	AQUASIS	*Peponocephala electra*	*Peponocephala electra*	*Peponocephala electra*	100	100	Beberibe- CE
CEUNES-UFES#6	UFES	*Peponocephala electra*	*Peponocephala electra*	*Peponocephala electra*	100	100	Vila Velha - ES

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
II47907	UNIVALI	*Phocoena dioptrica*	*Phocoena spinipinnis*	*	98	*	Santa Catarina - SC
FURG 68 - CT146,2	FURG	*Phocoena spinipinnis*	*Phocoena spinipinnis*	*Phocoena spinipinnis*	100	98.99	Rio Grande- RS
FURG 69 - CT175	FURG	*Phocoena spinipinnis*	*Phocoena spinipinnis*	*Phocoena spinipinnis*	100	100	Rio Grande- RS
FURG 70 - CT161	FURG	*Phocoena spinipinnis*	*Phocoena spinipinnis*	*Phocoena spinipinnis*	98.92	98.91	Rio Grande- RS
02C0410/308	AQUASIS	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	99.84	100	Fortaleza - CE
02C0410/338	AQUASIS	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	100	100	
02C0411/542	AQUASIS	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	100	100	
02C0411/809	AQUASIS	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	100	101	Paracuru- CE
02C0412/792	AQUASIS	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	99.85	99.85	Barroquinha- CE
FURG 17 CT213	FURG	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	100	100	Rio Grande- RS
FURG 18	FURG	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	99	100	Rio Grande- RS
FURG 19	FURG	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	100	100	Rio Grande- RS
FURG 20	FURG	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	98	99,7	Rio Grande- RS
GEMARS 0941	GEMARS	*Physeter macrocephalus*	*Physeter macrocephalus*	*Physeter macrocephalus*	100	100	Palmares do Sul- RS
GEMARS 0215	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	98	99.85	Tavares
GEMARS 0424	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	99	99,84	
GEMARS 0530	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	99	99,85	RS
GEMARS 0550	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	100	100	RS
GEMARS 0634	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	99	100	Imbé- RS
GEMARS 0745	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	100	100	RS
GEMARS 0748	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	100	100	RS
GEMARS 0749	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	99.82	100	Tramandaí- RS

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
GEMARS 1487	GEMARS	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	100	99,84	RS
LEC#01	UFPR	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	99.85	100	Pontal do Paraná- PR
LEC#71	UFPR	*Pontoporia blainvillei*	*Pontoporia blainvillei*	*Pontoporia blainvillei*	99.69	99.69	Pontal do Paraná- PR
FURG1_	FURG	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	100	99,5	Rio Grande- RS
FURG2	FURG	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	99	99,4	Rio Grande- RS
FURG4	FURG	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	98	99.83	Rio Grande- RS
FURG5	FURG	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	100	100	Rio Grande- RS
G1659	GEMARS	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	100	99,85	Terra De Areia- RS
G1665	GEMARS	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	99.53	99.53	Cidreira- RS
CEUNES-UFES#1	UFES	*Pseudorca crassidens*	*Pseudorca crassidens*	*Pseudorca crassidens*	*99.87*	99,87	Santa Cruz - ES
02C1412/290	AQUASIS	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	100	99.69	Cascavel - CE
02C1412/406	AQUASIS	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	92	99,82	Caucaia- CE
02C1412/508	AQUASIS	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	99.26	99.55	F
02C1412/523	AQUASIS	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	98	100	Fortaleza - CE
LEC#92	UFPR	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	*99.54*	*99.82*	
PA186	IO -USP	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	92	99.69	
PA226	IO -USP	*Sotalia guianensis*	*Sotalia guianensis*	*Sotalia guianensis*	100	100	
#39	GEMM Lagos	*Stenella attenuata*	*Stenella attenuata*	*Stenella attenuata*	99.56	99.55	
02C1121/614	AQUASIS	*Stenella attenuata*	*Stenella attenuata*	*Stenella attenuata*	100	100	Barroquinha-CE
BC02	Fiocruz-GEMM-Lagos	*Stenella attenuata*	*Stenella attenuata*	*Stenella attenuata*	99	99,41	Bacia de campos
GEMARS_BC03_	GEMARS	*Stenella attenuata*	*Stenella attenuata*	*Stenella attenuata*	100	101	Bacia de Campos- SP
02C1151/531	AQUASIS	*Stenella clymene*	*Stenella clymene*	*Stenella clymene*	99.35	99.83	Beberibe-CE
02C1151/543	AQUASIS	*Stenella clymene*	*Stenella clymene*	*Stenella clymene*	100	99,83	
02C1152/333	AQUASIS	*Stenella clymene*	*Stenella clymene*	*Stenella frontalis*	100	100	Itarema- CE
02C1152/733	AQUASIS	*Stenella clymene*	*Stenella clymene*	*Stenella clymene*	100	99,85	Cascavel- CE
GEMARS0795	GEMARS	*Stenella clymene*	*Stenella clymene*	*Stenella clymene*	100	99.71	Tramandaí- RS

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
CEUNES-UFES 01C1152/99	UFES	*Stenella clymene*	*Stenella clymene*	*Stenella clymene*	99.69	100	Fernando de Noronha - PE
02C1142/295	AQUASIS	*Stenella coeruleoalba*	*Stenella coeruleoalba*	*Stenella coeruleoalba*	100	100	Icapuí - CE
GEMARS 0047	GEMARS	*Stenella coeruleoalba*	*Stenella coeruleoalba*	*Stenella coeruleoalba*	100	100	Pinhal- RS
GEMARS 1240	GEMARS	*Stenella coeruleoalba*	*Delphinus delphis*	*Delphinus delphis*	100	99.85	
GEMARS 1346	GEMARS	*Stenella coeruleoalba*	*Stenella coeruleoalba*	*Stenella coeruleoalba*	99,7	99.69	Cidreira- RS
GEMARS 1416	GEMARS	*Stenella coeruleoalba*	*Stenella coeruleoalba*	*Stenella coeruleoalba*	100	100	Quintão
GEMARS 1478	GEMARS	*Stenella coeruleoalba*	*Stenella coeruleoalba*	*Stenella coeruleoalba*	100	100	Pinhal- RS
FURG 29	FURG	*Stenella frontalis*	*Stenella frontalis*	*Stenella frontalis*	100	100	
FURG 30	FURG	*Stenella frontalis*	*Stenella frontalis*	*Stenella frontalis*	99	100	
GEMARS 1174	GEMARS	*Stenella frontalis*	*Stenella frontalis*	*Stenella frontalis*	99.69	99,68	Palmares do Sul- RS
GEMARS 1488	GEMARS	*Stenella frontalis*	*Stenella frontalis*	*Stenella frontalis*	99	99.69	
BC_009	Fiocruz-GEMM Lagos	*Stenella frontalis*	*Stenella frontalis*	*Stenella frontalis*	100	100	Bacia de Campos- SP
BC_051	Fiocruz-GEMM Lagos	*Stenella frontalis*	*Stenella frontalis*	*Stenella frontalis*	99.85	99.84	Bacia de Campos- SP
02C1131/226	AQUASIS	*Stenella longirostris*	*Stenella longirostris*	*Stenella longirostris*	100	100	Fortaleza- CE
02C1131/672	AQUASIS	*Stenella longirostris*	*Stenella longirostris*	*Stenella longirostris*	100	100	OCEANO
02C1131/681	AQUASIS	*Stenella longirostris*	*Stenella longirostris*	*Stenella longirostris*	100	100	Icapuí - CE
02C1132/403	AQUASIS	*Stenella longirostris*	*Stenella longirostris*	*Stenella longirostris*	100	100	Fortaleza- CE
G1317	GEMARS	*Stenella longirostris*	*Stenella longirostris*	*Stenella longirostris*	100	100	Xangrilá- RS
02C1210/601	AQUASIS	*Steno bredanensis*	*Steno bredanensis*	*Steno bredanensis*	100	99,85	Caucaia- CE
FURG 11	FURG	*Steno bredanensis*	*Steno bredanensis*	*Steno bredanensis*			Rio Grande- RS
GEMARS 0512	GEMARS	*Steno bredanensis*	*Steno bredanensis*	*Steno bredanensis*	96	99,56	Mostardas- RS
GEMARS 1621	GEMARS	*Steno bredanensis*	*Steno bredanensis*	*Steno bredanensis*	99	100	Mostardas- RS
02C1312/696	AQUASIS	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	100	100	Fortaleza- CE

SAMPLE	INSTITUTION	MID	NCBI/ GenBank ID	BOLD ID	NCBI	BOLD	LOCALITY
GEMARS 1485 CERAM 2828)	GEMARS	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	98,96	99,54	Cidreira- RS
GEMARS 1633	GEMARS	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	99	100	Cidreira- RS
GEMARS_ASPSP_C	GEMARS	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	97.22	97.21	Ipojuca- PE
GEMARS_ASPSP_E	GEMARS	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	97.06	96.92	Ipojuca- PE
GEMARS_ASPSP_N	GEMARS	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	100	99,85	Ipojuca- PE
II 2061	UNIVALI	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	100	99,85	Barra Velha- SC
II1850	UNIVALI	*Tursiops truncatus*	*Tursiops truncatus*	*Tursiops truncatus*	100	100	Penha- SC
02C0000/683	AQUASIS	*Ziphius cavirostris*	*Ziphius cavirostris*	*Ziphius cavirostris*	100	100	Itarema - CE
02C0812/305	AQUASIS	*Ziphius cavirostris*	*Ziphius cavirostris*	*Ziphius cavirostris*	100	100	Fortaleza - CE
AB02	ANDRE BARRETO	*Ziphius cavirostris*	*Ziphius cavirostris*	*Ziphius cavirostris*	100	100	
GEMARS 1484	GEMARS	*Ziphius cavirostris*	*Ziphius cavirostris*	*Ziphius cavirostris*	99.85	99.85	Cidreira- RS
UNIVALI_2	UNIVALI	*Ziphius cavirostris*	*Ziphius cavirostris*	*Ziphius cavirostris*	99,37	99,37	

* Molecular database did not recognize any similar sequence.

Regarding the mismatches found between molecular and morphological indetifications of cetacean carcasses, we can separate the results in four types: 1) incorrect morphological identification, 2) incomplete molecular databases; 3) incorrect species identification in molecular databases, and 4) absence of gap barcode between species.

3.1 Incorrect morphological identification

The first mismatch case between morphological and molecular identifications was found in the FURG 45 specimen. It was identified during field work as an orca (*Orcinus orca*), but both molecular databases indicated greater similarity with the false killer whale (*Pseudorca crassidens*). The pictures from the head of the dead specimen were sent to two marine mammal specialists, who concluded based on teeth countings that FURG 45 specimen was in fact *P. crassidens*. The determinant characteristic for the identification was the number of teeth, which varies from 7 to 12 teeth pairs in *P. crassidens* and from 10 to 12 in *O. orca* (Fig. 2a) (Jefferson et al., 2008).

Other two molecular morphological mismatches due to incorrect morphological identification was reported: a) GEMARS 1491 identified during sampling activity as a humpback whale (*Megaptera novaeangliae*), but in both molecular databases the identification suggested was southern right whale (*Eubalaena australis*), and b) FURG 63 specimen, morphologically identified as sei whale (*Balaenoptera borealis*) but molecularly, as a Bryde's whale (*Balaenoptera brydei*). In both cases, the specimens were found in advanced state of decomposition, categories 4 and 5, respectively, according to the classification determined by Geraci & Lounsbury (1993). Based on the carcasses conditions, we believe that evaluation of the species identidy were very difficult, and probably the collectors misidentified the two specimens.

Fig. 2A. FURG 45 specimen: a *Pseudorca crassidens* found stranded in advanced decomposing stage. The species identification was based on teeth counting as well as DNA barcoding analysis. Photo: FURG. 2B. Gemars 1491 specimen found on the coast of Rio Grande do Sul and identified as *Megaptera novaeangliae* cf. but genetically it was *Eubalaena australis.* Photo: GEMARS.

3.2 Incomplete molecular databases

The GEMAM110 was identified by the field researchers as *Inia geoffrensis,* but the molecular identification was *I. araguaiaensis.* This mismatch is explained by the fact that at the time the sample was collected and deposited in the molecular database, the species *I. araguaiaensis* was not formally decribed yet (for details see Hrbek et al. 2014). Until the publication of this new species and the deposit of its sequences in the molecular databases, all Amazon river dolphins were uniquely identified as *I. geoffrensis*.

The GEMARS 1155 specimen was morphologically identified as Arnoux's beaked whale (*Berardius arnuxii*) and in the molecular databases as Baird's beaked whale (*Berardius bairdii*). These species also cause uncertainties in their morphological differences, since the validity of these species had already been questioned (Balcomb, 1989 in Jefferson et al., 2008). However, studies based on the mitochondrial cytochrome b gene have proven molecular differences and recognized them as distinct species (Dalebout et al., 2004). On the other hand, it is noteworthy that the species *B. bairdii* occurs only in the northern hemisphere (Perrin et al., 2009). Thus, we can state that the molecular identification provided by the molecular platforms do not match with morphological identifcation probably because there are no sequences of *B. arnuxii* for the *COI* region in them, which makes *B. bairdii* the closest species sequence suggested by blast tool. Same result was observed with II47907 specimen, which was morphologically identified as spectacled porpoise (*Phocoena dioptrica)* based on a fresh carcass, but molecularly as Burmeister's porpoise (*Phocoena spinipinnis*), because there are no *COI* sequences of *P. dioptrica* in both GenBank and BOLD databases.

3.3 Incorrect species indetification in molecular databases

We found only two cases of potential incorrect species indetification in molecular databases for the same species. According to external morphology, samples MN60458 and FrancaRJ were identified as *Eubalaena australis*, these identifications were confirmed by Blast tool from Genbank. However, the Bold platform identified both samples as *Eubalaena glacialis,* a species that only occurs in North Atlantic waters (Cummings, 1985). The result points to an erroneous deposit of sequences in the Bold platform.

3.4 Absence of gap barcode between species of Delphinidae family

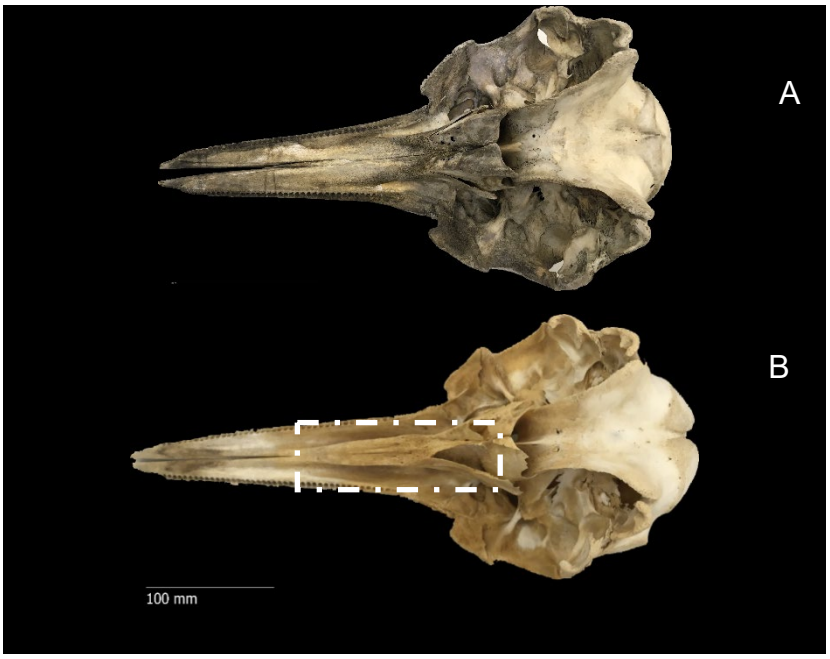
The sample AQUASIS 02C1152/333 was morphologically identified as Clymene dolphin (*Stenella clymene*). This identification was confirmed by GenBank, but not by BOLD System database, which reported 100% similarity with *Stenella frontalis* and *Stenella clymene*. This specimen was live rescued, but very debilitated according to the Marine Mammal Rehabilitation Center (CRMM), and it died a few hours after of treatment. Due to fresh conditions of the external morphology traits, there was no doubt about the identification of the specimens by field collectors as *S. clymene* (Fig. 3). Since the species identification suggested by GenBank database and by members of the stranding network, that use morphological diagnosis, agrees, and in BOLD system exists 12 *COI* sequences for *S. clymene* the only possible explanation is the absence of gap barcode between the species *S. frontalis* and *S. clymene*.

Fig 3. AQUASIS 02C1152/333 specimen of Clymene dolphin (*Stenella clymene*) still alive during its treatment.

The blast results suggested that the GEMARS1240 specimen was a short-beaked common dolphin (*Delphinus delphis)* by both molecular databases, but the specimen was morphologically

identified during field work and also after the skull examination as striped dolphin (*Stenella coeruleoalba*). Taking into account the morphological diagnosis of *D. delphis* skull, based on the presence a prominent trapezius palatal groove (Fig.4 a,b), and the fact that in G1240 specimen this trait was absent, we agree with the morphological identification as *S. coeruleoalba.*

Fig. 4. Ventral view of skulls A. Striped dolphin (*Stenella coeruleoalba*) (G1240). B. common dolphin (*Delphinus delphis),* highlighting the prominence of the palatal sulcus (traced).



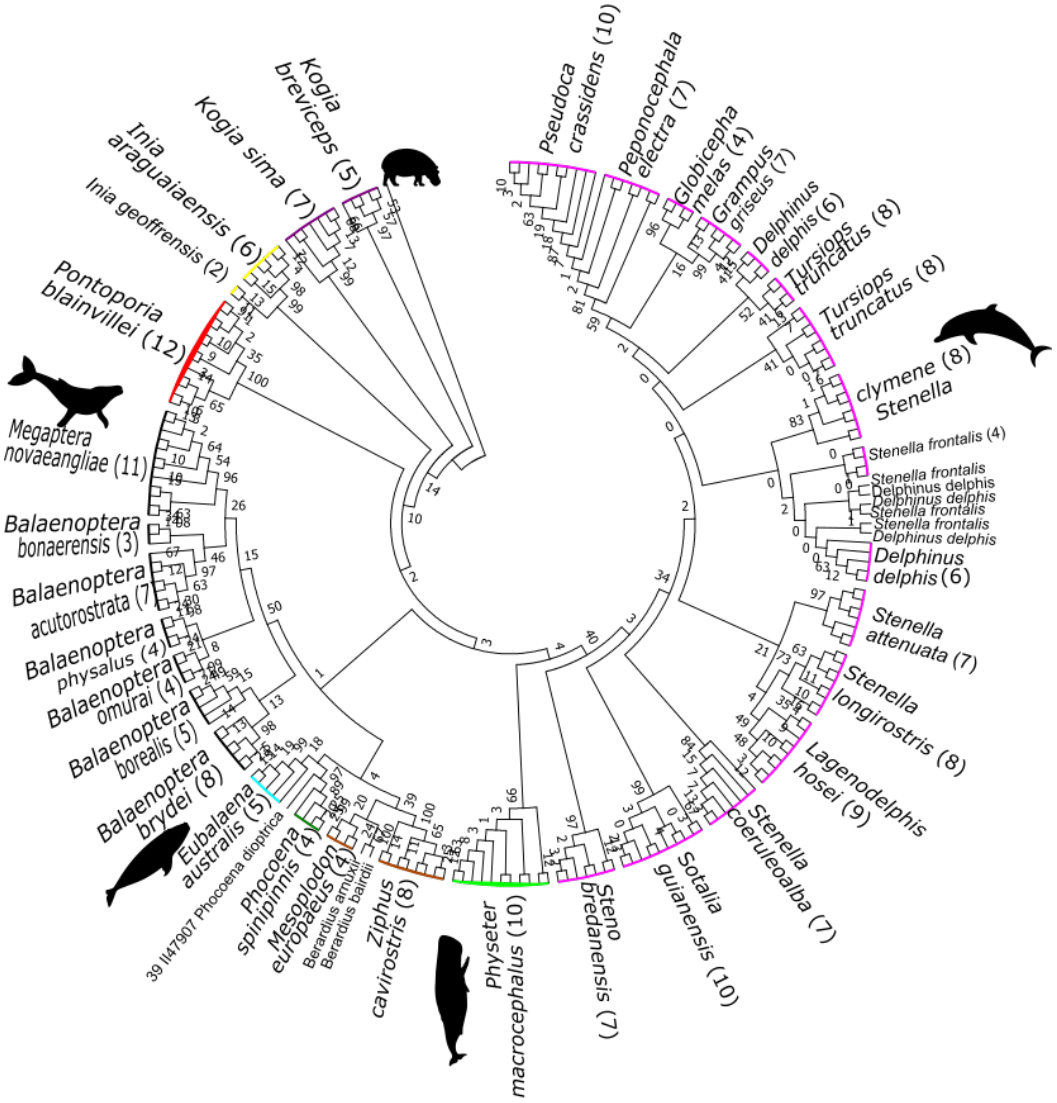
3.5 Inter and Intraspecific distances of *COI*

Intra and interspecific genetic divergences for *Delphinus delphis*, *Stenella clymene*, *Stenella coeruleoalba, Stenella frontalis* and *Tursiops truncatus* of the family Delphinidae are detailed in Table II. Measurements of intra-specific variation ranged from 0 to 0.56% while interspecific variation ranged from 0.38% to 2.56%, with a mean divergence of 1.5%. The Neighbour-Joining (NJ) tree correctly distinguished all the cetaceans analyzed (Fig.5), except the species of the Delphinidae family, which presented very small intra-specified genetic divergences. However, some species of this same family formed clades with high bootstrap support values (> 90): *Sotalia guianensis, Steno bredanensis, Grampus griseus, Stenella attenuata and Globicephala melas.*

Table II. Species whose *COI* marker was not efficient to identify of species level taking into account pairwise inter and intraspecific distances

		Genetic divergences (%)					
		Between species					Within species
Species		1	2	3	4	5	
1	*Delphinus delphis*						0.56
2	*Stenella clymene*	1.56					0.00
3	*Stenella coeruleoalba*	1.73	2.56				0.27
4	*Stenella frontalis*	0.38	1.16	1.33			0.00
5	*Tursiops truncatus*	1.03	1.97	2.15	0.70		0.28

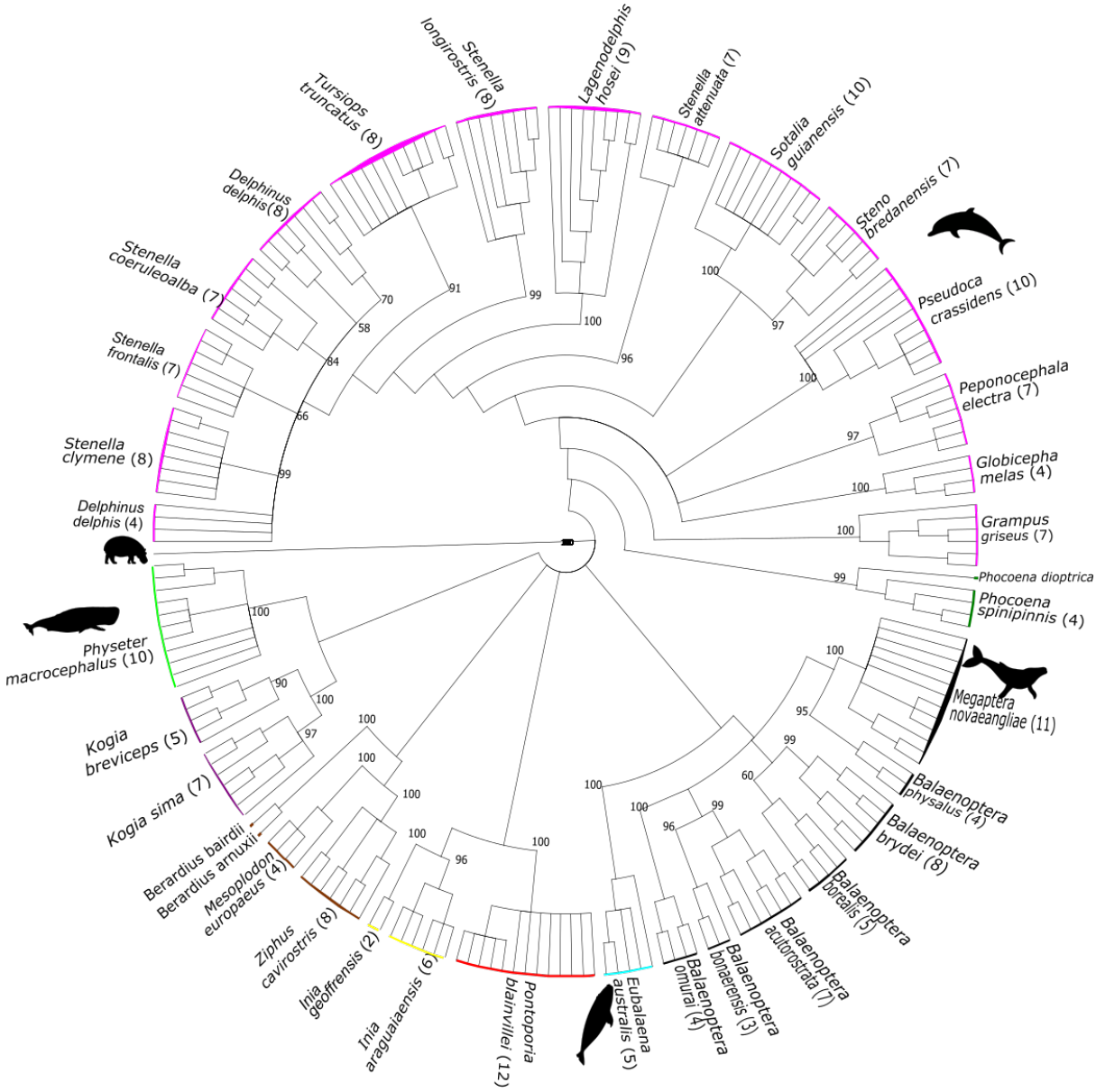
Fig. 5 NJ tree generated from pre-existing sequences in BoldSystem and the 152 sequences generated in this study. The number within the parentheses represents the number of individuals analyzed, the colors indicate different families of cetaceos.



3.6 Phylogenetic reconstruction through Maximum-likelihood

Although there are problems with determining intraspecific limits for some species of delphinidae, maximum likelihood (ML) shows a tree with species in well-defined clades (Fig. 6), supporting the *COI* as a plausible marker for species identification cetaceans, except *for Delphinus delphis*, the only species one not to form monophilic clade.

Fig.6 Maximum-likelihood tree generated from pre-existing sequences in BoldSystem and the 152 sequences generated in this study. The number within the parentheses represents the number of individuals analyzed, the colors indicate different families of cetaceos.



4. DISCUSSION

This study generated 152 sequences from the *COI* region, representing 33 cetacean species, that reflects 70% of the Brazilian diversity of this taxon. The molecular identification was in accordance with external morphology identification in 93% of the specimens. We detected 11 cases of molecular-morphological mismatch identifications, all were solved in favor of the molecular identification, excepting three cases of absence of gap barcoding between delphinid species (genus Stenella and Delphinus). It indicates a great reliability to the DNA barcoding data generated here and its relevance for any a cetacean stranding network, in terms of biological sample bank.

Alfonsi et al. (2013) analyzing an average of 150 marine mammals per year, along the Brittany coast in France (including cetaceans and pinnipeds carcasses), also observed similar results. They concluded that barcoding DNA, even with certain constrains, is still very useful for the French stranding network. Morever, they were able to amplify *COI* sequences of good quality from highly degraded carcasses, which allowed them to correct identified all the expected specimens with no exceptions, that represents around 16% of the specimens every year.

In general, most of species molecularly identified was very common in coastal region with no challenging identification, such as common bottlenose dolphin (*Tursiops truncatus*) and franciscana dolphins (*Pontoporia blainvillei*). However, there were also records of rare species from Ziphiidae and Phocoenidae families (Pinedo et al. 2002), besides oceanic and deep diving species such as sperm whales (including pygmy, dwarf and *Physeter macrocephalus*), and the only endemic cetacean species for Brazil, the recently described the Araguaian River dolphin *Inia araguaiaensis*.

It is worth mentioning that the present research is one of the few studies involving DNA barcoding sequences of 152 specimens with voucher materials/ deposited in a scientific collection, enabling morphological checking whenever necessary, and thus providing greater reliability of the use of the molecular marker used as barcode. According to Hanner (2009), part of BOLD quality control, barcode must be associated with specimen records linked to institutional (e.g., museum) material making them the most valuable among putative reference accessions. This is particularly important in cases of rare species, which usually have no sequences deposited in molecular platforms. Gaubert and co-workers (2014) highlighted that the accuracy of DNA barcoding species assignment relies upon the level of taxonomic representation for each taxonomic group and the amount of intraspecific genetic

diversity represented in the databases barcode sequences in BOLD. It is important to mention that the present research contributed with the inclusion of the first *COI* sequences of spectacled porpoise (*Phocoena dioptrica*) and Arnoux's beaked whale (*Berardius arnuxii*), in both GenBank and Bold databases, besides the inclusion of 150 new *COI* sequences of cetaceans.

Taking into account the requirement of *COI* sequences associated to voucher material, Galimberti et al. (2015) emphasized that the standardized molecular reexamination of museum-deposited voucher specimens and the comparison with other reference information allows the fast detection of misidentification, that typically occurring during field surveys.

This was particularly true for the case GEMARS 1491 specimen found on the coast of Rio Grande do Sul and identified as humpback whale (*Megaptera novaeangliae* c.f.), but genetically it was a southern right whale (*Eubalaena australis*). As mentioned earlier, when we examined the field notes present in the catalog book of the scientific collection, we found that the specimen was in an advanced stage of decomposition and that there was a highlighted note saying "c.f," which means "need to confirm" or "need to compare with" Latin, which supports the care referred to by Galimberti et al. (2015). Moreover, the GEMARS 1491 specimen has some cyamids known as whale-lice which crustacean parasites attached on it, which were *a posteriori* identified as belonging exclusively from to southern right whale (Iwasa-Arai et al. 2017), corroborating the barcoding identification.

Francis et al. (2010), pointed out that field determinations for many mammal species are difficult, because they require the analysis of internal morphology such as skull or dentition (including cetaceans)), which makes so important the existence of comparative material in scientific collection as well as DNA samples, in order to confirm the identification of specimens through DNA barcodes (or other gene) allied to a voucher specimen. In the case of marine mammals, one of the main challenges for a taxonomist relies on the fact that the largest reference collections are scattered among museums. The present study has the privilege to count with 13 institutions in Brazilian territory with vast scientific collections, that allow us to detect cases of incorrect morphological identification on field, as was the case with sample GEMARS 1491.

The FURG 45 specimen was misidentified during field work as an orca (*Orcinus orca*), but both molecular databases indicated greater similarity with the false killer whale (*Pseudorca crassidens*). The final identification was based on internal morphology taking into account the number of teeth, which

varies from 7 to 10 teeth pairs in *P. crassidens* and from 10 to 12 in *O. orca* (Jefferson et al., 1993). This morphological diagnosis agreed with molecular suggestion that the specimen was a *P. crassidens*.

However, other case involving the combination of barcoding information and the reexamination of the internal morphology of skull anatomy from two dolphin specimens revealed an unsolved morphological-molecular mismatch: both molecular databases suggested that the GEMARS 1240 specimen was a short-beaked common dolphin (*Delphinus delphis*), but the presence a prominent trapezius palatal groove indicated morphological identification as *S. coeruleoalba* (Jefferson et al., 1993). Moreover, this case also has an overlap between intra- and interspecific *COI* genetic variation, suggesting that *COI* is an imperfect barcode for these species. In fact, all three unsolved mismatches between the morphological and molecular identifications found in the present study occurred with species belonging to the Delphinidae family, as already reported in other studies (Amaral et al., 2007; Viricel and Rosel 2011; Alfonsi et al., 2013), that demonstrated the limited efficiency of this marker in identifying species this group. Moreover, NJ analysis showed that *D. delphis*, *S. frontalis* and *T. truncatus* species do not form monophyletic groups, probably due to introgression processes, as reported for *D. delphis* in *S. coeruleoalba* species (Kessler, unpublished data) or due to insufficient time of divergence between the taxa (Zhou et al. 2011). On the other hand, through methods that use an a priori evolutionary model, such as maximum likelihood (ML), it was possible to recover a greater number of monophyletic groups that corresponded to the sequences identified at species level for all cetaceans in this study, except for the species. *D. delphis.*

Moreover, *D. delphis*, *T. truncatus*, *S. coeruleoalba*, *S. frontalis* and *S. clymene* presented very low interspecific *COI* distances, which ranged from 2.56% (*S. clymene vs*. *S. coeruleoalba*) to 0.38% (*D. delphis* vs. *S. frontalis*), these values are below the 3% expected to delimit species-level taxa using the *COI* as a marker, as proposed by Hebert et al. (2003). Due to the difficulty of species separation, the Delphinidae family has been the target of studies that seek to solve evolutionary relationships among its members using information regarding morphology, genetics and, recently, acoustics and historical biogeography (Amaral et al. 2007 and Vollmer et al., 2019). On the other hand, the *COI* proved to be an efficient tool to identify other families, obtaining high support values for all groups in the NJ analysis, demonstrating the efficiency of this marker for non-Delphinidae species.

In summary, the main concerns regarding the identification of cetaceans using the *COI* gene are related to: i) cetaceans seem to do not have the “barcoding gap”, which is a lack of overlap between

intraspecific and interspecific nucleotide divergence in the investigated taxa (Viricel & Rosel, 2012); and it is crucial to higher identification success rates; ii) the hybrids generate taxonomic questions, since COI gene is maternal inherited and requires the use of biparental nuclear genes to establish the identification of the species; iii) the taxonomic updates are not simultaneous, and iv) the existence of mistakenly identified aquatic mammalian sequences deposited in Bold and GenBank, as the case of the species *Eubalaena australis*, deposited as *Eubalaena glacialis* . In these last cases, to minimize the possibility of errors in the correct identification, it is important to have knowledge about the geographical region where the study material was collected.

All these concerns are relevant, but the according to Galimbert et al. (2015) "...reference sequences constitute the main core of the DNA barcoding initiative and their absence or the lack of control of the correct identification of the source specimens by expert taxonomists, can irremediably affect the assignment of newly generated query sequences". This is why is so important the existence of voucher material related to every *COI* sequence generated.

Although DNA barcoding still generated controversies, when it is considered as a "taxonomic service" it become a very interesting tool, able to contribute to the knowledge of mammal diversity, providing information on the biology, distribution and conservation of mammals, mainly on rare or poorly investigated taxa (Galimberti et al. 2015). It is clear that there is a hidden biodiversity within the mammal record, including in great whales, which have their last species described in 2003 as *Balaenoptera omurai*, based on comparisons of external morphology, osteology and mitochondrial DNA data, and raising the number of known living *Balaenoptera* species to eight (Wada et al. 2003). This same species has its distribution recently expanded to Brazilian waters based on stranded specimen identified by cytochrome b and COI sequences (Cypriano et al. 2016), since identification through its external morphology had been compromised due to decomposition process. Moreover, DNA barcoding proved to be more effective in discriminating cryptic or morphologically similar species, such as the species of Amazon River dolphins, genus *Inia*. DNA barcoding approach recognized the existence of new lineage that are confined to Araguaia and Tocantins basins (Hrebk et al. (2014) as well as Marajó Bay (Siciliano et al., 2016), in north of Brazil.

The generation of barcode segment for the spectacled porpoise (*Phocoena dioptrica*) in this study provides the first COI sequence of the species, which is a rare dolphin found out of their known distributional range. It is a small cetacean with circumpolar distribution in Antarctic and sub Antarctic

waters, which only has one previous record published to the Brazilian coast in 1994 (Pinedo et al. 2002). There are another unpublished record in August 2016 for Cassino beach, in Rio Grande do Sul coast (Fruet unpublished data). The specimen analyzed in the present study was collected at Navegantes Beach, Santa Catarina State (26º53'40.4"S 48º38'32.3"W) in July 2017, and represents the northernmost record of this species in Atlantic Ocean (Barreto in prep.).

Another record of a poorly known and elusive cetacean that we provide new COI sequences is the Fraser's dolphin (*Lagenodelphis hosei*). There was a mass stranding event of 10 dolphins along 156 km of sandy beaches in the Rio Grande do Sul state coast, between September and November 1997 (Pinedo et al. 2001; Moreno et al. 2003), and four of these specimens were analyzed in this study. This stranding was not an isolated event, other stranded animals were reported to Uruguay as well as Rio de Janeiro state coast. As a final counting, around 100 specimens were reported to the Southwestern Atlantic coast in 1997 (for a review see Moreno et al. 2003).

Fifteen cetaceans registered for Brazilian coast are classified by IUCN as "Data Deficient", mainly due to the lack of taxonomic or ecological information about these animals (MMA 2014, Hrbek et al., 2014, Cypriano-Souza et al., 2016, IUCN 2019). This scarcity of data and the accelerated process of degradation and pollution that suffers the marine and fresh water environments occupied by these species, reinforce the need for studies that can help and optimize the production of knowledge about this group, enabling the elaboration of action plans aimed at the conservation of these species (MMA 2014).

Considering the given scenario, stranded cetacean carcasses can provide valuable information about the richness and pattern of occurrence of this group off the Brazilian coast (Sholl et al. 2008), since correctly identified. Therefore, despite some recognized limitations (Galimberti et al. 2015), our results demonstrate that the DNA barcode, when properly used, can be an exceptional tool for the scientific community related to the stranding network.

5. ACKNOWLEDGEMENTS

- Laboratório de Estudo de Mamiferos da Unisinos, Laboratório de Biologia Molecular da UNISINOS
- Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS,
- Grupo de Estudos de Mamíferos Aquáticos do Rio Grande do Sul – GEMARS,
- Aquasis, Universidade Federal do Espírito Santo – UFES,
- Universidade de São Paulo – USP,
- Universidade Federal de Rio Grande – FURG,
- Grupo de Estudos de Mamíferos Aquáticos da Amazônia, ligado ao Museu Paraense Emilio Gueld GEMMAM
- Universidade do Vale do Itajaí – UNIVALI,
- Fundação Osvaldo Cruz – FIOCRUZ,
- Associação de Pesquisa e Preservação de Ecossistemas Aquáticos,
- Museu Nacional do Rio de Janeiro,
- Universidade Federal do Paraná UFPR,
- Grupo de Estudos de Mamíferos, Aves e Répteis Marinhos e Costeiros da Região dos Lagos – GEMM/Lagos

6. REFERENCES

Ab'saber, Aziz Nacib (2001). Litoral do Brasil. São Paulo: Meta Livros, 287 p.

Amaral AZ, Jablonski S (2005) Conservation of Marine and Coastal Biodiversity in Brazil. Coxazanservation Biology, [s.l.], v. 19, n. 3, p.625-631

Avise JC (2004). Molecular markers, natural history and evolution. Chapman & Hall

Baker CS, Palumbi SR (1994) Which Whales Are Hunted? A Molecular Genetic Approach to Monitoring Whaling. Science, [s.l.], v. 265, n. 5178, p.1538-1539. American Association for the Advancement of Science (AAAS)

BRASIL. Ministério do Meio Ambiente. Portaria MMA nº 444, de 17 de dezembro de 2014.

Chang CH, Yao CJ, Yu HY, Liao YC, Jang-Liaw N H, Tsai C L, Shao KT (2014) A molecular forensic method for identifying species composition 339 of processed marine mammal meats. Journal of Forensic and Legal Medicine, 340 23, 65-69.

Chaves PB, et al (2011) DNA barcoding meets molecular scatology: short DNAmt sequences for standardized species assignment of carnivore noninvasive samples. Molecular Ecology Resources, [s.l.], v. 12, n. 1, p.18-35

CyprianoALS et al (2016) Rare or cryptic? The first report of an Omura's whale (Balaenoptera omurai) in the South Atlantic Ocean. Marine Mammal Science, [s.l.], v. 33, n. 1, p.80-95

Cummings WC (1985) Right whales Eubalaena glacialis (Müller, 1776) and Eubalaena australis (Desmoulins, 1822) . In “ Handbook of Marine Mammals” (S. H. Ridgway, and R. Harrison , eds) Vol. 3 , pp. 275 – 304 . Academic Press , San Diego, CA

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9: 772

Dalebout ML, Baker CS, Mead JG, Cockcroft VG, Yamada TK (2004) A comprehensiveand validated molecular taxonomy of beaked whales, family Ziphiidae. Journal of Heredity 95:459–473 DOI 10.1093/jhered/esh054.

Dalebout ML, van Helden A, Van Waerebeek K, Baker CS (1998) Molecular genetic identification of southern hemisphere beaked whales (Cetacea: Ziphiidae). Mol Ecol 7:687–69

Dalebout ML, Mead JG, Baker CS, Baker NA, Van Helden AL (2002) A new species of beaked whale *Mesoplodon perrini* sp. n. (Cetacea: Ziphiidae) discovered through phylogenetic analyses of mitochondrial DNA sequences. Marine Mammal Science 18

Jefferson TA, Leatherwood S, Webber MA (1993) FAO Species Identification Guide: Marine Mammals of the World. Roma, Food and Agriculture Organization, 320p

Galimbertia A, Sandionigia A, Brunoa A, Bellatib A, Casiraghi M (2015) DNA barcoding in mammals: what’s new and where next? Hystrix, It. J. Mamm

Gaubert P, Njiokou F, Olayemi A, Pagani P, Dufour S, Anquah E, Nutsuakor MEK, Ngua G, Missoup AD, Tedesco P, Dernat R, Antunes A (2014) Bushmeat genetics: setting up a reference framework for the DNA typing of African forest bushmeat. Mol. Ecol. Resour. 15(3): 633–651

Hall TA (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95-98

Hanner R (2009) Data Standards for BARCODE Records in INSDC (BRIs). Available from http://barcoding.si.edu/PDF/DWG_data_standards-Final.pdf Accessed 15.05.15

Hebert PDN et al (2003) Biological identifications through DNA barcodes Proceedings Of The Royal Society B: Biological Sciences, [s.l.], v. 270, n. 1512, p.313-321. The Royal Society

Hrbek T, Silva VMF, Dutra N, Gravena W, Martin AR, Farias IP (2014) A New Species of River Dolphin from Brazil or: How Little Do We Know Our Biodiversity. PLoS ONE 9(1): e83623. https://doi.org/10.1371/journal.pone.0083623

ICMBio (2011b) Plano de Ação para a Conservação dos mamíferos aquáticos – Grandes Cetáceos e Pinípedes. In: Rocha-Campos et al. (Org). Brasília: Instituto Chico Mendes de Conservação da Biodiversidade. 156 p

ICMBio/MMA (2018) Livro Vermelho da Fauna Brasileira Ameaçada de Extinção: Volume I / -- 1. ed. -- Brasília, 492 p

IUCN (2019) The IUCN Red List of Threatened Species. Version 2019-2. <http://www.iucnredlist.org>. Acesso em 03 de novembro de 2019

Iwasa-Arai T, Sicilianao S, Serejo CS, Rodríguez-Rey GT (2017) Life history told by a whale-louse: a possible interaction of a southern right whale *Eubalaena australis* calf with humpback whales *Megaptera novaengliae*. Helgoland Marine Reserach 71: 6.

Kimura MJ (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences Journal of Molecular Evolution 16: 111. https://doi.org/10.1007/BF01731581

Lemos LS, Moura JF, Hauser-Davis RA, De Campos R C, Salvatore S (2013). Small cetaceans found stranded or accidentally captured in southeastern Brazil: Bioindicators of essential and non-essential trace elements in the environment. Ecotoxicology and Environmental Safety. 97: 166-175

Longo LL, Amado F, Gilberto M (2014) O conhecimento da fauna marinha bentônica brasileira através dos tempos. História, Ciências, Saúde-manguinhos, [s.l.], v. 21, n. 3, p.995-1010, FapUNIFESP (SciELO)

Meyer C P, Paulay G (2005) DNA Barcoding: Error Rates Based on Comprehensive Sampling. Plos Biology, [s.l.], v. 3, n. 12, p.422-423, Public Library of Science (PLoS).

Moreno IB, Zerbini AN, Danilewicz D, Santos MOS, Simões-Lopes PC, Lailson-Brito J, Azevedo AF (2005) Distribution and habitat characteristics of dolphins of the genus *Stenella* (Cetacea: Delphinidae) in the southwest Atlantic Ocean. Marine Ecology Progress Series. Vol. 300: 229–240

Ott P H, Tavare M, Moreno IB, Oliveira LR, Danilewicz D (2009) Os cetáceos do Arquipélago de São Pedro e São Paulo, Brasil PAGES Capítulo 11. in: Mohr, L. V.; Wagner, J. C.; Costa, P. M. and Válka, R. J. (Orgs.). Ilhas Oceânicas Brasileiras: da Pesquisa ao Manejo. Volume II. Brasília: Ministério do Meio Ambiente

Ott PH, Tavares M, Secchi ER, Di TullioJC (2013) Cetacea. P. 457-550. In: Weber, M. M.; Roman, C; Cáceres N. C. (Org.). Mamíferos do Rio Grande do Sul. 1ed. Santa Maria: UFSM

Palumbi SR, Kessing BD (1991) Population biology of the trans-arctic exchange: DNAmt sequence similarity between pacific and atlantic sea urchins. Evolution, [s.l.], v. 45, n. 8, p.1790-1805

Pauls SU, Nowak C, Bálint M, Pfenninger M (2012) The impact of global climate change on genetic diversity within populations and species. Molecular Ecology 22: 925–946 doi: 10.1111/mec.12152

Perrin WF (2009) Encyclopedia of Marine Mammals. Second edition. Academic Press: London

Pinedo MC, Barreto AS, Lammardo MP, Andrade ALV, Geracitano L (2002) Northernmost records of the spectacled porpoise, Layard's beaked whale, Commerson's dolphin, and Peale's dolphin in the southwestern Atlantic Ocean. Aquat. Mamm., v. 28, n. 1, p. 32-37

Prado HF, Mattos PH, Silva KG, Secchi ER (2016) Long-Term Seasonal and Interannual Patterns of Marine Mammal Strandings in Subtropical Western South Atlantic. PLoS ONE 11(1) doi: 10.1371/journal.pone.0146339

Reeves RR, Brent SS, Clapham PJ, Powell JÁ, Folkins P A (2002) National Audubon Society guide to marine mammals of the world. Chanticleer Press, New York, NY

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, Bioinformatics, Vol 30, Issue 9, 1312–1313, doi: 10.1093/bioinformatics/btu033

Sholl TGC et al (2013) Cytochrome b sequencing for the species identification of whale carcasses washed ashore in Brazil. Marine Biodiversity Records, Vol. 6; e30

Sholl TGC et al (2008) Taxonomic identification of dolphin love charms commercialized in the Amazonian region through the analysis of cytochrome b DNA. Journal Of The Marine Biological Association Of The United Kingdom, [s.l.], v. 88, n. 06, p.1207-1210

Skueresky NS (2014) Código de barra de dna de cetáceos brasileiros: Uma ferramenta molecular como subsídio para a construção de uma rede nacional que busca a conservação de espécies. 2014. 35 f. TCC (Graduação) - Curso de Ciência BiolÓgicas, Universidade do Vale do Rio dos Sinos, São Leopoldo

Siciliano S, Moura JF, Filgueiras HR, Rodrigues PP, Leite Jr NO (2012) Sightings of humpback whales on the Vitória-Trindade chain and around Trindade Island, Brazil. Brazilian Journal of Oceanography , 60: 455-459

Siciliano S et al (2018) Dolphins, Love and Enchantment: Tracing the Use of Cetacean Products in Brazil. Frontiers in Marine Science, v.5, Article 107, 10p

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets Molecular Biology and Evolution 33:1870-1874

Thompson K, Baker CS, Van Helden A, Patel S, Millar C, Constantine R (2012) The world's rarest whale. Current Biology 22, R905–R906. doi: 10.1016/j.cub.2012.08.055

Van Bressem MF, Van Waerebeek K, Reyes J, Félix F, Echegaray M, Siciliano S, Di Beneditto AP, Flach L, Viddi F, Avila IC, Herrera JC, Tobon IC, Bolanos-Jimenez J, Moreno IB, Ott PH, Sanino GP, Castineira E, Montes D, Crespo E, Flores PAC, Haase B, Souza SMFM, Laeta M, Fragoso AB (2007) A preliminary overview of skin and skeletal diseases and traumata in small cetaceans from South American waters. Latin American Journal of Aquatic Mammals, 6: 7-42.

Vianna TS, Loch C, Castilho, PV, Cremer MJ,Simões-Lopes CP (2016) Review of thirty-two years of toothed whale strandings in Santa Catarina, southern Brazil (Cetacea: Odontoceti). Zoologia v33:5 https://doi.org/10.1590/S1984-4689zool-20160089

Vollmer NL, Ashe E, Brownell RLJr, Cipriano F, Mead JG, Reeves RR, Soldevilla MS, Williams R (2019) Taxonomic revision of the dolphin genus Lagenorhynchus Marine Mammal Science. 35:957-1057.

Viricel A, Rosel PE (2011) Evaluating the utility of cox1 for cetacean species identification. Marine Mammal Science, [s.l.], v. 28, n. 1, p.37-62.

Wada S, Oishi M, Yamada TK (2003) A newly discovered species of living baleen whale. Nature, 426:278-281

Yamada TK, Kitamura , Abe S,Tajima Y, Matsuda A, Mead JG, Matsuishi TF (2019) Description of a new species of beaked whale (Berardius) found in the North Pacific. *Scientific Reports*, 9 (1) DOI: 10.1038/s41598-019-46703-w

Zerbini AN, Secchi ER, Bassoi M, Dalla Rosa L, Higa A, Sousa L, Moreno IB, Moller L, Caon G (2004) Distribuição e Abundância Relativa de Cetáceos na Plataforma Continenal Externa e Talude da Zona Econômica Exclusiva no Sul e Sudeste do Brasil: Resultados Preliminares e Perspectivas Futuras. Sao Paulo: Editora da Universidade de Sao Paulo, v. 1. 40p.

7. SUPPLEMENTARY MATERIAL

Supplementary Table 1. Description of primer sequences[1], concentration and specific parameters used in all performed PCRs and primer sequences[2] used in sequencing.

Gene	Primers[1]	Primers sequence (5' - 3')	PCR reaction (25 µl final volume)	Cycling conditions	References
COI	VF1d	TTCTCAACCAACCACAARGAYATYGG	1 µM MgCl2; 0.3 µM dNTPs; 0.2 µM of each primer; 0.2 µM Taq DNA Polymerase; 2.5 µM 1X Buffer; 1,25 µM of bovine serum albumin (BSA); 0.2 µM sense primer and 0.2 µM antisense for a 25 µL of PCR **volume**	After the initial 5 min denaturation step at 95 ° C, the cycling parameters were: 1 cycle of 95 °C for 40 s, 56 °C for 30 s, 54 °C for 30 s, 72 °C, followed by 34 cycles at 95 °C for 40 s each, reactions were terminated by a final 5 min extension step at 72 °C.	Ivanova et al., 2007
	VF1i	TTCTCAACCAACCAIAAIGAIATIGG			
	VR1	TAGACTTCTGGGTGGCCAAAGAATCA			
	VR1d	TAGACTTCTGGGTGGCCRAARAAYCA			
	Primers[2]				
	M13-FP	TGTAAAACGACGGCCAGT	Amplicons were submitted to direct sequencing at Macrogen (Macrogen Inc., Seoul, Korea),		
	M13R-pUC	CAGGAAACAGCTATGAC			