

Supplementary Material

Conservation issues using discordant taxonomic and evolutionary units: a case study of the American manatee (*Trichechus manatus*, Sirenia)

Camilla S. Lima^A, Rafael F. Magalhães^B and Fabricio R. Santos^{C,D}

^APrograma de Pós-Graduação em Genética, Departamento de Genética, Ecologia e Evolução, ICB, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, CP 486, 31270-010 Belo Horizonte, Brazil.

^BDepartamento de Ciências Naturais, Campus Dom Bosco, Universidade Federal de São João del-Rei, Pç. Dom Helvécio 74, 36301-160 São João del-Rei, Brazil.

^CLaboratório de Biodiversidade e Evolução Molecular, Departamento de Genética, Ecologia e Evolução, ICB, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, CP 486, 31270-010 Belo Horizonte, MG, Brazil.

^DCorresponding author. Email: fsantos@icb.ufmg.br

DNA extraction, amplification and Sanger sequencing

The DNA extraction of a new sample, from North of Brazil, was performed following the Phenol-Chloroform protocol (Sambrook *et al.* 1989). The mtDNA control region was amplified using the primers L15926 and H16498 (Kocher *et al.* 1989). The amplification cycle and PCR mix followed Vianna *et al.* (2006). PCR products were purified by the polyethylene glycol (PEG) method (20% PEG 8000, 2.5 M NaCl) with modifications (Santos Júnior *et al.* 2015) and sequenced on the ABI 3130xl Genetic Analysis (Applied Biosystems) using the BigDye Termination v3.1 Cycle Sequencing kit. The same primers used for amplification were used for DNA sequencing. The Chromatograms were analysed in SeqScape 2.6 and the consensus sequence was aligned with the others using the ClustalX (Larkin *et al.* 2007) algorithm in the software MEGA 6 (Tamura *et al.* 2013). The haplotype was inferred in software DNAsp 5 using the algorithm Phase (Rozas *et al.* 2003), confirming that this sample has a new haplotype (M5).

The summary statistics of the *T. manatus* haplotypes were calculated also in DNAsp 5 (Rozas *et al.* 2003).

GMYC analysis

We estimated a mtDNA gene tree using 410 bp sequences of the control region (Garcia-Rodriguez *et al.* 1998; Vianna *et al.* 2006), including the new haplotypes G3, J4, M4, and M5 that were recently found in Brazil (Saticzábal *et al.* 2012; Luna *et al.* 2012; Vilaça *et al.* 2019), and using *T. inunguis* as outgroup. The best substitution model and the ultrametric gene trees of unique haplotypes were estimated in BEAST v2.5 software (Bouckaert *et al.* 2019), assuming the lognormal uncorrelated clock and the Birth-Death tree prior. For the model selection, we used the package bModelTest v1.1.2 under the “allreversible” option (Bouckaert and Drummond, 2017). The analysis was run twice, 1×10^6 pre burnin, 1×10^7 generations and thinning of 2×10^4 each [CL1]. The proper parameter mixing [AT2] of each replicate and convergence between the runs were evaluated visually in Tracer v 1.7 (Rambaut *et al.* 2018) through inspection of traces of each parameter and minimum adequate effective sample size (ESS > 200). The replicates were combined in LogCombiner and the maximum clade credibility (MCC) tree was annotated using the TreeAnnotator post-processing tools (Bouckaert *et al.* 2019).

The mtDNA lineage delimitation was made in two steps. In the first step, we submitted the MCC tree to a maximum likelihood version of the GMYC model (Pons *et al.* 2006), using the single threshold algorithm implemented in the splits package of the R software (Fujisawa and Barraclough 2013; R Core Team 2017). In this analysis, the null hypothesis of a single lineage on the tree was rejected (LR test = 18.3617.23; $p < 0.001$) and three lineages were recovered, with a confidence interval from three to five. In the second step, we submitted a random sample of 2.5×10^4 trees retained in BEAST to a Bayesian version of the GMYC model, implemented in the bGMYC package (Reid and Carstens 2012; R Core Team 2017), in order to account for phylogenetic uncertainties associated with differences in sampled topologies and variation in branch lengths.

Table S1. Genbank' accession numbers and definitions from all sequences used in GMYC analysis.

Accession	Definition	Countries of occurrence
AY963840	<i>T. manatus</i> haplotype A1	Colombia ⁴ , Cuba ⁵ , Dominican Republic ^{1,2} , Mexico ⁵ , Puerto Rico ^{1,2,3} , United States ^{1,2}
AY963841	<i>T. manatus</i> haplotype A2	Puerto Rico ^{2,3}
AY963842	<i>T. manatus</i> haplotype A3	Belize ⁶ , Cuba ⁵
AY963843	<i>T. manatus</i> haplotype A4	Belize ^{2,6} , Mexico ²
AY963844	<i>T. manatus</i> haplotype A5	Colombia ^{2,4}
AY963845	<i>T. manatus</i> haplotype B1	Colombia ⁴ , Dominican Republic ^{1,2} , Puerto Rico ^{1,2,3}
JX564997	<i>T. manatus</i> haplotype B2	Puerto Rico ³
AY963846	<i>T. manatus</i> haplotype C1	Colombia ^{1,2,4}
AY963847	<i>T. manatus</i> haplotype D1	Colombia ^{1,2}
AY963848	<i>T. manatus</i> haplotype E1	Colombia ^{1,2,4}
AY963849	<i>T. manatus</i> haplotype G1	Colombia ^{1,2,4}
AY963850	<i>T. manatus</i> haplotype G2	Colombia ^{2,4}
JX982651	<i>T. manatus</i> haplotype G3	Colombia ⁴
AY963851	<i>T. manatus</i> haplotype H1	Colombia ^{1,2}
AY963852	<i>T. manatus</i> haplotype I2	Colombia ² , Venezuela ²
AY963853	<i>T. manatus</i> haplotype J1	Belize ^{2,6} , Colombia ^{1,2,4} , Mexico ^{1,2} , Venezuela ^{1,2}
MK673553	<i>T. manatus</i> haplotype J4	French Guiana ⁷
AY963854	<i>T. manatus</i> haplotype K1	Venezuela ^{1,2}
AY963855	<i>T. manatus</i> haplotype L2	Guyana ²
AY963856	<i>T. manatus</i> haplotype M1	Brazil ^{1,2,7,8} , Guyana ^{1,2}
AY963857	<i>T. manatus</i> haplotype M2	Brazil ²
AY963858	<i>T. manatus</i> haplotype M3	Brazil ^{2,8}
JX171295	<i>T. manatus</i> haplotype M4	Brazil ⁸
MK673554	<i>T. manatus</i> haplotype M5	Brazil (This study)
AY963859	<i>T. manatus</i> haplotype N1	Guyana ^{1,2}
AY963860	<i>T. manatus</i> haplotype O1	Guyana ^{1,2}
AY963861	<i>T. inunguis</i> haplotype P1	Guyana ^{1,2}

¹ Garcia-Rodriguez *et al.* 1998; ² Vianna *et al.* 2006; ³ Hunter *et al.* 2012; ⁴ Satizábal *et al.* 2012; ⁵ Hernandez-Martínez *et al.* 2013; ⁶ Hunter *et al.* 2010; ⁷ Vilaça *et al.* 2019; ⁸ Luna *et al.* 2012;

Table S2. Summary statistics of the *T. manatus*, calculated do each ESUs and both.

	Number of haplotypes	Number of variable sites	Nucleotide diversity
Caribbean ESU	17	36	0.03818
Atlantic ESU	9	25	0.01518
Both ESUs	26	46	0.04164

Migrate-N analysis

We made an independent test of gene flow using the Bayesian algorithm of MIGRATE-N v.4.2.14 (Beerli, 2006; Beerli and Palczewski, 2010) implemented in the CIPRES Science gateway online portal (Miller *et al.*, 2010). We tested 12 different models: (1) Panmixia (i.e., null model); (2) isolation between Florida, Caribbean and Brazil sequences; (3) considering Florida and Caribbean sequences a single population, isolated from Brazil; (4) considering Brazil and Caribbean as a single population, isolated from Florida; (5) Caribbean as a drain, between Florida and Brazil; (6) bidirectional gene flow between Florida and Caribbean, and Brazil isolated; (7) unidirectional gene flow from Florida to Caribbean, and Brazil isolated; (8) unidirectional gene flow from Caribbean to Florida, and Brazil isolated; (9) bidirectional gene flow between Brazil and Caribbean, and Florida isolated; (10) unidirectional gene flow from Brazil to Caribbean, and Florida isolated; (11) unidirectional gene flow from Caribbean to Brazil, and Florida isolated; and (12) isolation between Florida, Caribbean, and Brazil. To estimate each model, we used sequences of mtDNA region control from each location (Florida = 56, Caribbean = 428 and Brazil = 105), excluding samples from the hybrid zone. We set an empirical transition/transversion ratio ($R_{\text{mtDNA}} = 20.292$) estimated under Kimura 2-parameter model in MEGA X v.10.0.5 (Kimura, 1980; Kumar *et al.*, 2018) and estimated the migration parameters in terms of mutation-scaled effective immigration rate M (m/μ) and mutation-scaled effective population sizes Θ ($N_e\mu$). Initial values for M and θ were derived from F_{ST} estimates, and the prior distributions were adjusted after exploratory analyses. Each analysis was made in two parallel replicates, using a static heating scheme with one long and 4 short chains in which “temperatures” increased according to the program’s suggested range of values. For each run, we set a burn-in of 100,000 steps for each chain and 1,000,000 recorded chain, with a sampling increment of 100. The convergence between the replicates was accessed through the pairs strategy, and the mixing was evaluated using effective sample sizes (ESS) plus update rates of the chains. The model selection was made comparing the marginal likelihood of each model, using the Bezier’s approximation scores calculated (Beerli and Palczewski, 2010).

Table S3. Frequencies of each haplotype in each region.

Florida (N = 56)				
A1 = 56				
Caribbean (N = 428)				
A1 = 67	A2 = 7	A3 = 9	A4 = 91	A5 = 4
B1 = 114	B2 = 1	C1 = 14	D1 = 1	E1 = 26
G1 = 10	G2 = 2	G3 = 3	H1 = 2	I2 = 1
J1 = 75	J4 = 2	K1 = 1	L2 = 1	M1 = 1
N1 = 1	O1 = 1			
Brazil (N = 105)				
M1 = 98	M2 = 1	M3 = 2	M4 = 3	M5 = 1

Table S4. Model selection with Migrate-N.

Model	Bezier's approximation score	PP	Weight
<i>Fullmodel</i>	-2376.16	0.000	<0.001
<i>Panmixia</i>	-2589.19	0.000	<0.001
<i>Florida + Caribbean / Brazil</i>	-2104.69	0.000	<0.001
<i>Brazil + Caribbean / Florida</i>	-2441.46	0.000	<0.001
<i>Florida → Caribbean ← Brazil</i>	-2406.79	0.000	<0.001
<i>Florida ↔ Caribbean / Brazil</i>	-2012.75	0.332	0.524
<i>Florida → Caribbean / Brazil</i>	-2012.11	0.633	1.000
<i>Florida ← Caribbean / Brazil</i>	-2051.32	0.000	<0.001
<i>Brazil ↔ Caribbean / Florida</i>	-2457.25	0.000	<0.001
<i>Brazil → Caribbean / Florida</i>	-2301.37	0.000	<0.001
<i>Brazil ← Caribbean / Florida</i>	-2292.50	0.000	<0.001
<i>Isolation</i>	-2015.01	0.035	0.055

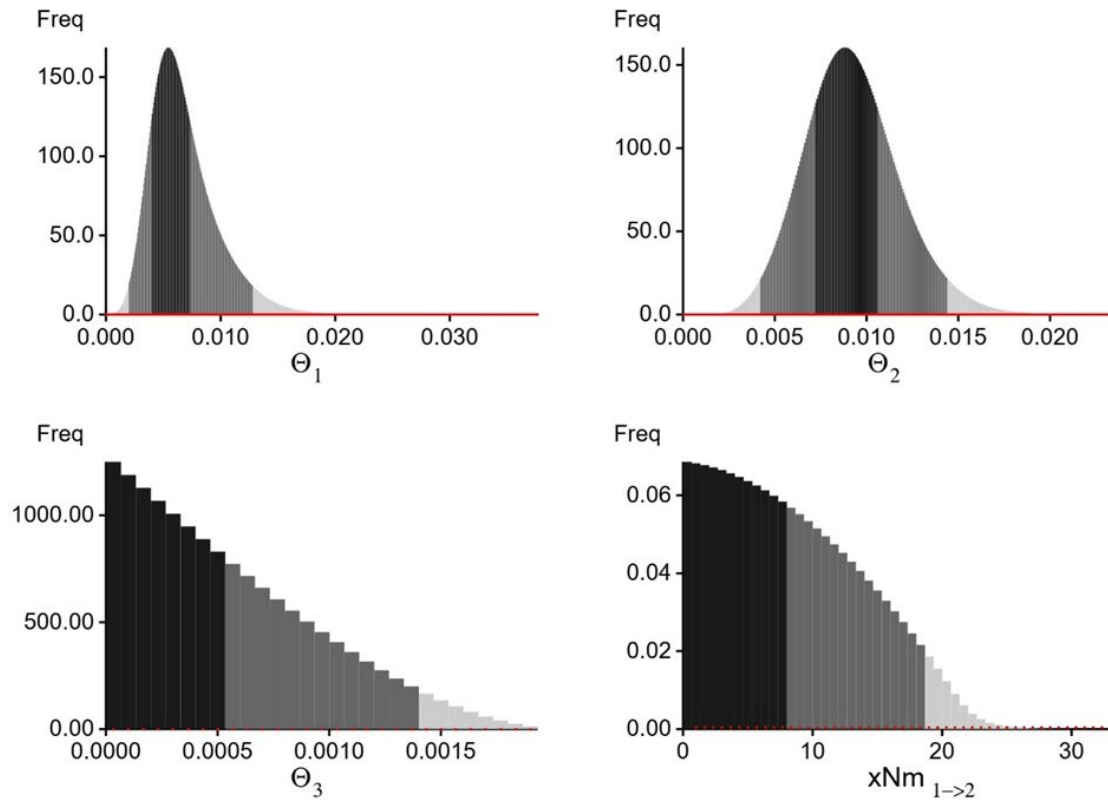


Figure S1. Histogram of posterior distribution to estimated parameters for the best model selected (unidirectional gene flow from Florida to Caribbean), where $\Theta_1 = \Theta_{\text{Florida}}$, $\Theta_2 = \Theta_{\text{Caribbean}}$, $\Theta_3 = \Theta_{\text{Brazil}}$, and $xNm_{1 \rightarrow 2} = xNm_{\text{Florida} \rightarrow \text{Caribbean}}$.

References

- Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22**, 341-345. doi:10.1093/bioinformatics/bti803
- Beerli, P., and Palczewski, M. (2010). Unified Framework to Evaluate Panmixia and Migration Direction Among Multiple Sampling Locations. *Genetics* **185**, 313-332. doi:10.1534/genetics.109.112532
- Bouckaert, R., and Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology* **17**, 1–11. doi: 10.1186/s12862-017-0890-6
- Bouckaert, R.; Vaughan, T. G.; Barido-Sottani, J.; Duchêne, S.; Fourment, M.; Gavryushkina, A.; Heled, J.; Jones, G.; Kühnert, D.; De Maio, N.; Matschiner, M.; Mendes, F. K.; Müller, N. F.; Ogilvie, H. A.; du Plessis, L.; Popinga, A.; Rambaut, A.; Rasmussen, D.; Siveroni, I.; Suchard, M. A.; Wu, C.; Xie, D.; Zhang, C.; Stadler, T.; Drummond, A. J. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **15**, e1006650. doi:10.1371/journal.pcbi.1006650
- Fujisawa, T., and Barraclough, T. G. (2013). Delimiting Species Using Single-Locus Data and the

- Generalized Mixed Yule Coalescent Approach: A Revised Method and Evaluation on Simulated Data Sets. *Systematic Biology* **62**, 707–724. doi:10.1093/sysbio/syt033
- Garcia-Rodriguez, A. I., Bowen, B. W., Domning, D. P., Mignucci-giannoni, A. A., Marmontel, M., Montoya-Ospina, R. A., Morales-Vela, B., Rudin, M., Bonde, R. K., and McGUIRE, P. M. (1998). Phylogeography of the West Indian manatee (*Trichechus manatus*): how many populations and how many taxa? *Molecular Ecology* **7**, 1137–1149. doi:10.1046/j.1365-294x.1998.00430.x
- Hernandez-Martínez, D. H., Álvarez-Alemán, A. A., Bonde, R. K., Powell, J. A., García-Machado, E. (2013). Diversidad haplotípica em el manatí *Trichechus manatus* em Cuba: Resultados preliminares. *Revista de investigaciones marinas* **33**, 58-61. doi: 10.13140/2.1.2922.5601
- Hunter, M. E., Auil-Gomez, N. E., Tucker, K. P., Bonde, R. K., Powell, J., and McGuire, P. M. (2010). Low genetic variation and evidence of limited dispersal in the regionally important Belize manatee. *Animal Conservation* **13**, 592–602. doi:10.1111/j.1469-1795.2010.00383.x
- Hunter, M. E., Mignucci-Giannoni, A. a., Tucker, K. P., King, T. L., Bonde, R. K., Gray, B. a., and McGuire, P. M. (2012). Puerto Rico and Florida manatees represent genetically distinct groups. *Conservation Genetics* **13**, 1623–1635. doi:10.1007/s10592-012-0414-2
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–120. doi:10.1007/BF01731581
- Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X., and Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences* **86**, 6196–6200. doi:10.1073/pnas.86.16.6196
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35**, 1547–1549. doi:10.1093/molbev/msy096
- Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Thompson, J., Gibson, T., and Higgins, D. (2007). ClustalW and ClustalX version 2. *Bioinformatics* **23**, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Luna, F. O., Bonde, R. K., Attademo, F. L. N., Saunders, J. W., Meigs-Friend, G., Passavante, J. Z. O., and Hunter, M. E. (2012). Phylogeographic implications for release of critically endangered manatee calves rescued in Northeast Brazil. *Aquatic Conservation: Marine and Freshwater Ecosystems* **22**, 665–672. doi:10.1002/aqc.2260
- Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 2010, 1-8. doi:10.1109/GCE.2010.5676129
- Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., Vogler, A. P. (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55**. 595-609. doi: 10.1080/10635150600852011

- R Core Team (2017). R: A language and environment for statistical computing.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*. 1–5. doi:10.1093/sysbio/syy032
- Reid, N. M., and Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology* **12**, 1–11. doi:10.1186/1471-2148-12-196
- Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497. doi: 10.1093/bioinformatics/btg359
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning. *Society* **68**, 1232–1239. doi:10.1128/AEM.68.3.1232
- Santos Júnior, J. E., Santos, F. R., and Silveira, F. A. (2015). Hitting an Unintended Target: Phylogeography of *Bombus brasiliensis* Lepeletier, 1836 and the First New Brazilian Bumblebee Species in a Century (Hymenoptera: Apidae). *PLoS ONE* **10**, e0125847. doi:10.1371/journal.pone.0125847
- Satizábal, P., Mignucci-Giannoni, A. A., Duchêne, S., Caicedo-Herrera, D., Perea-Sicchar, C. M., García-Dávila, C. R., Trujillo, F., Caballero, S. J. (2012). Phylogeography and sex-biased dispersal across riverine manatee populations (*Trichechus inunguis* and *Trichechus manatus*) in South America. *PLoS ONE* **7**, e52468. doi: 10.1371/journal.pone.0052468
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729. doi:10.1093/molbev/mst197
- Vianna, J. A., Bonde, R. K., Caballero, S., Giraldo, J. P., Lima, R. P., Clark, A., Marmontel, M., Morales-Vela, B., De Souza, M. J., Parr, L., Rodríguez-Lopez, M. a, Mignucci-Giannoni, A. a, Powell, J. a, and Santos, F. R. (2006). Phylogeography, phylogeny and hybridization in trichechid sirenians: implications for manatee conservation. *Molecular ecology* **15**, 433–447. doi:10.1111/j.1365-294X.2005.02771.x
- Vilaça, S. T., Lima, C. S., Manzoni, C. J., Santos, F. R., Thoisy, B. (2019) Manatee genomics supports a special conservation area in the Guianas coastline under the influence of the Amazon River plume. *Estuarine, Coastal and Shelf Science* **226**, 106286. doi: 10.1101/552919