UNIVERSITY OF CALIFORNIA, SAN DIEGO

Population Genomic Structure and Phylogeography of Spotted and Spinner Dolphins (Genus Stenella)

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

in

Marine Biology

by

Matthew S. Leslie

Committee in charge:

Professor Phillip Morin, Chair
Professor Ronald Burton, Co-Chair
Professor James Fowler
Professor Richard Norris
Professor William Perrin
Professor Gregory Rouse

2016

The Dissertation of Matthew Steven Leslie is approved, and it is acceptable

in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2016

DEDICATION

This dissertation is dedicated to my parents; Steven C. and Rebecca E. Leslie, who taught me three important things: 1) to love, 2) that the natural world is a sacred place worthy of respect, and 3) that knowledge and kindness increase not only a man's worth, but also his enjoyment of the world in which he lives.

EPIGRAPH

"When I go outside at night and look up at the stars, the feeling that I get is not comfort. The feeling that I get is a kind of delicious discomfort at knowing that there is so much out there that I do not understand and the joy in recognizing that there is enormous mystery, which is not a comfortable thing. This, I think, is the principal gift of education."

Teller

In an interview with Jessica Lahey of *The Atlantic.*
Published online Jan 21, 2016, under the title:
'Teaching: Just Like Performing Magic.'

TABLE OF CONTENTS

LIST OF TABLES

Chapter 4

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

ACKNOWLEDGEMENTS

To my advisor, **Dr. Phillip A. Morin**: Well, they say you always remember your first! Thirteen years ago in a windowless conference hall, I stood in front of a scientific poster of research half complete. A warm and welcoming man approached me, a young and nervous scientist with little to offer (at least I thought). This scientist wanted to speak about my poster, which we did for maybe ten or fifteen minutes. As he departed he mentioned that he was applying to gain adjunct status at Scripps Institution of Oceanography, would soon be able to take students, and I should keep that in mind.

Six years later, you would become my Ph.D. advisor, but the advising started before that. I wasn't an easy get. Recall that I wrote two NSF Pre-doctoral proposals. You reviewed both of these. I was offered the fellowship on the second try, so it should have been a slam-dunk decision to head to SIO. Instead I deferred the fellowship for a year and shopped around. Ultimately, Stanford, Cornell and Berkeley joined in contention. I took my time consternating over the decision and at the 11th hour chose SIO over Cornell. Both had three things that were important to me: 1) an opportunity to follow my dream of studying marine mammals, 2) a good 'community' of researchers and students for support, and 3) quality advisors. Despite having no track record of students I chose you and SIO over the others, because we had built a mentor-mentee relationship already.

We have continued to foster that relationship over the last seven years and I hope we will continue working together in the foreseeable future. It hasn't always been easy. Certainly there were times we both could have used more guidance. But most importantly, you were always there for me. You didn't have all the answers, but we found

them together. Thank you for your flexibility and willingness to let me experiment with life goals and skill development outside of dissertation research. Most of all, thank you for being honest with me throughout the process. I appreciate you perspective and all the wonderful feedback about my science and in general. Thanks for sticking with me, and thanks for coming to visit my poster.

To my committee: Dr. Ron Burton: thanks for your help and mentorship throughout this process. Most of all thanks for typing these words – they really lit a fire under my ass: "Matt tends to digress from the goal of dissertation completion too quickly."

Dr. William Perrin: Bill, thank you for your support throughout. You were instrumental in getting me here and committed to seeing me through the processes. Thank you for letting me occasionally fondle your skulls. As a "reluctant geneticist" that helped keep me grounded. All the groundwork you laid for getting my work done was tremendous. It has been fun to work with you and I hope we can continue to work together in the coming years. Most of all – thanks again for throwing your support behind this project and my education. I don't think it would have gone very far without all your support, direction, and input. Lastly, thanks for sharing Petrich 1965 – an ETP classic.

Dr. James Fowler: Whenever they make the Oklahoma Science Hall of Fame, you'll be my first nominee. Wow! Really, thanks for all the career advice over coffee. Despite the perceptions of SIO, not many faculty members want to talk career strategy

(wait, you're not doing a postdoc at Woods Hole?) Thanks for being a kind and welcoming soul. It was so nice having such a brilliant and tremendously supportive Okie on my committee, seriously!

Dr. Richard Norris: Dick! Thanks for the long walks on the beach and the coffee chats. Most of all, thanks for lending me the phrase, "We're off on the greatest adventure of our lives!" I will take that and give it to my students. My dream is to lead droves of undergrads through knee-deep biology field trips someday, and these words will ring in their ears. You are such a wonderful soul… please, never change. Finally, thanks for giving a deep-time perspective to my work. It is something that I don't get enough of and I really enjoy evolutionary "thinking."

Dr. Greg Rouse: Man, Greg, where would I be without you? I have been pretty stuck more than once when I came to talk with you. Maybe the worst quagmire was before we had beers and decided to write that "endangered species act meets phylogenetic nomenclature" paper. As it turned out this is now the ONE published product out of my dissertation. Others are close, but this one is done - mostly as a testament to your inspiration and willingness to guide me though the process. I know, I know, I wrote the damn thing, but I wouldn't have started or finished it without your help. Thanks for that. One of my favorite memories while at SIO was searching for monoplacs with you. I miss this level of science – the discovery part. You've been a great example and inspiration to continue to pursue these activities. Thanks also for always being so welcoming and warm. I'll always count you high among my mentors and without a doubt one of my friends.

NOAA Southwest Fisheries Science Center: Special thanks goes out to all my colleagues at SWFSC colleagues. Dr. Eric Archer – I couldn't have done this without your mentoring and support. Thank you for always making time in your busy schedule for my silly questions. I look forward to working with you more soon. Dr. Karen Martien, thanks for your keen eye and helpful comments on draft manuscripts. Dr. Lisa Ballance, thank you for always supporting me and my work, and for being a champion of graduate education at SWFSC. I think you "get" me on a level that most don't, and I really appreciate all the time you've help get me into the field. Dr. Barb Taylor, thank you for letting me be one of your PIs. It has been wornderful to be treated as a colleague and I look forward to continued collaborations in the years to come. Drs. Aimee Lang and Nick Kellar, thank you for all your time and mentorship as well. This included shooting me straight about the ins and outs of SIO before I ever set foot here as a grad student. Greatly appreciated!! Dr. Lisa Komeroski – thanks for your advice and help along the way. I look forward to staying in touch and continuing to work together. Thanks to Bob Pitman and Dr. John Durban – you guys have welcomed me from day one. Even with my genetics wonkiness (which is a means to an end), it has been great to share time in the field with you two. I hope we can work something out to continue to explore the biology of cetaceans together.

I would have no data without the support of all the managers and technicians in the SWFSC Marine Mammal Genetics Group. First and foremost, Vicki Pease and Brittany Hancock-Hansor: thank you two for your wonderful friendships and the unceasing support in the laboratory. Really, I would have very little without your support

on a personal and professional level. Kelly Robertson – it goes for you. You're help has been invaluable along the way. Most of all thank you for your kind and supportive nature and your vast institutional knowledge. Thanks also to Amanda, Morgane, Tiffany, Amy Lanci, Gaby, Nicky, and many others that I'm forgetting for helpful hands in the lab.

I have to thank the ETP people at SWFSC as well. These include Drs. Paul Fiedler and Jessica Redfern for help with the SST mapping analysis. Al Jackson for help with datasheets and photographs and Dr. Susan Chivers for her help cross-referencing life history databases.

Thanks to all the SIO personnel committed to supporting graduate students: These include but are not limited to Anne Middleton, Stuart Krantz, Penny Dockry, and Grad Office Past and Present: Especially, Gilbert Bretado, Adam Peterson, Gayle Aruta, Denise Darling, Josh Reeves, Maureen Mccormack, Maureen McGreevy and all of the CMBC faculty and students. I wouldn't not have come to SIO it if wasn't for all of you making such a welcoming impression on me.

Fellow SIO students: Oh man. Thanks to my SIO cohort for all your friendships and support. Noelle Bowlin, Cotton Rockwood, Jill Harris, Levi Lewis, Joe Campanale (and Greg), Christian Briseno (and James), Lauren Shipp (and Reggie), Kait Fraiser, Layll Bellquist, Scott DeWolf, Kerry Seger, Bill Jones, Liz Vu, and so many others! You all are amazing inspirations – thank you for being my teammates and friends!!

Southwest Aggregation of Graduate Students (SWAG): Aly, Summer, Tara, Nacho, Cotton, Trevor, Doug, Eric, A.J. Amy, Eiren, Kerri – and the adults: Lisa, Phil, Eric, and Jay. Couldn't have done it without you!!

Playground friends: **Kevin/Kate, Paul/Jen, Justin/Cara, Martha/Jesse, Amber/Steven, John/Liana** and other families. It has been so much to have you as friends and for our kids to become friends. Thank you for all your support and especially for knowing what it is like (good and bad) to start small families as graduate students and recent grads. It has been a blast!!

Family Friends and Support: To Ms. Joan Goodrich, you are such a wonderful soul. We thank our lucky stars every day that we have you. Our lives would be utter chaos without you and your love for our children. We really are blessed to have you as a fifth member of our young family. You've known both of our children since we brought them home from the hospital. What a crazy ride we've all had together. Thank you for all you do, "Joanie No'Bologna."

Similarly, Mr. Matty and Ms. Natalie from MCDC have maybe been the most pleasant surprise of my time here as a graduate student. Again, what wonderfully supportive soles and kind, kind, kind, hearted people. Your classroom is a second home to Kaia and she might even consider it her preferred! We love what you do, and we love you. Thank you thank you from the bottom of my heart - for what you do and who you are!

To my best childhood friends, Mike and Steve: thanks for not killing me with all those stupid things we used to do. No really. Thanks for your support. Looking forward to sharing a few beers with you soon.

To my New York Friends, Sergios, Martin, Kevin, Mitch, Craig, Ryan. You guys are way smart, because you made this dissertation thing look WAY too easy. How come you never told me how tough this stuff was. Hateful bastards. Miss you guys.

To Grandma: Marm, thanks for all your support. It has been so nice to have you around so much. I know that it isn't easy for your to let your daughter and grandkids go all the way across the country, but I appreciate it. Most of all, I appreciate you coming to spend time with us and help when we need it. You are a star and I love you so much.

To Aunt Jackie: Jake. What can I say? Thanks for everything. You've been so supportive all the way through. It has been fun to share all the adventures, trials, failures, and joys with you as my most favorite Aunty Shrew in the entire world! I love you so much and I look forward to seeing you and celebrating very soon. PS. Thanks for taking me fishing.

To my Brother: You are the best brother I ever had. No really, thanks for all the time we spent together outdoors, for the support you've given me your entire life and most of all for being my best friend.

To my Father: First, thank you for taking time to be a Dad and for being a great example of what it means to be a Dad. You've always encouraged me and supported me to follow my dreams. From providing me with envelops and stamps to write to whale

biologists as a kid, to coming to all my games and performances in high school and college, to being there for my very first scientific presentation at the Oklahoma Academy of Science, to driving me to New York and to San Diego. You've always been an amazing cheerleader for everything Scott and I did. Moreover, you've provided whatever we boys needed to succeed. A lot of Dads would have called all this a big waste of time, but I've never once felt the slightest bit of doubt from you. You've trusted me to know what is best for my life, and allowed me to stumble and learn from my mistakes. As a father, I know that this must have been a difficult thing to do at times, but I appreciate it. Thank you for fostering a connection with nature from a young age. I would not be who I am, nor do what I do, had it not been for all of the time we spent outdoors. I know how easy it would have been just to stay home all those weekend. How many football games you must have missed, and how nice it would have been to sleep in your own bed. But that time allowed me to fall in love with nature, and certainly shaped my passion and commitment to understand and conserve our planet and the amazing and unique life on it. Thanks again for fostering a family where education was valued. Thanks for always encouraging Scott and I to "go look it up" in the encyclopedia. I love encyclopedia!! You blazed the trail by going to graduate school. You showed me that learning was important and that it should be a lifelong pursuit. You were my baseline – from there I knew the sky was the limit.

To my Mother: Man I miss you. I would give everything I own for you to be here and celebrate this with us. After all, it is mostly your fault. I can't see hummingbirds without thinking of you, or purple irises, or honeysuckle. Through your love for these

things – and many many other wild things - I saw the value in nature for nature's sake. Thanks for always letting me burrow your camera… and for developing all the bad photos I shot. Those were important times for me getting to "try on" being a field biologist and fueling my love for the outdoors. I'm sure you would remember this better than I do, but one of my very first camping trips was with you and Kelly Ganzel and his mom. I believe it was a Mom and Me campout with the Cub Scouts. Thanks for always being supportive. I think this may have been harder for you than for Dad, just because you were a little more protective. I know you were worried when I went off to Australia to follow my dreams of studying whales. Scott told me. Don't be mad at him. Despite your illness you must have been worried when I moved to New York. Yet you were always so supportive. I learned a lot about parenting from you and I can't wait to help my girls become confident, smart, loving, passionate, funny, caring, kind, goofy ladies, just like you. They're fun. You'd be peas in a pod. It hurts my heart that they'll never get to meet you – and you them. Thank you for being a great example of a strong and confident woman. I knew just what to look for, and I found a good one in their mother. Thanks for buying me that book at Sam's Club. You probably don't remember, it was the big coffee table book on whales - it changed me. That book gave me a purpose in life. It fueled my curiosity in a way that has lead to what you see today. Finally, thanks for all the joy you brought to everyone's lives. This might be the biggest gift you gave me – a love of people. You knew how to make people smile and how to work a crowd; I think I do too and it certainly has helped me along the way. Did I mention I really miss you?

To my girls: It will be impossible to think back on this time without thinking of you and the balance you have brought to my life. Thank you for your patience and understanding with a Dad who often wasn't there for you when you woke up, and who in general spent too much time at his work. Thank you for being my joy at the end of long days. Thank you for your inspiration to leave this planet a better place. Most of all, thanks for digging for worms with me and for going with me on spider walks; your Dad desperately needs that.

Foremost, to my wife Phaedra: My how our lives have changed since driving west over the George Washington Bridge seven years ago. Without your continuous herculean support, none of this would have been possible. I would need several hundred pages to list all the duties you have picked-up in my stead. Thank you for standing with me, for giving me an amazing family, for being my confidant, for letting me vent, and for basically living through graduate school a second time… for me. Most of all, thank you for letting 'my tie my boat to the tree on your island', and for always encouraging me to follow my dreams. It's been quite a ride; there is little doubt in my mind that would not have made it without you[1].

[1] PS. you're badass too, in so many ways.

Chapter 1 is published in *Conservation Biology*; the dissertation author was the primary investigator and sole author of this material:

Leslie, M. S. 2015. Impacts of phylogenetic nomenclature on the efficacy of the U.S. Endangered Species Act. Conservation Biology, 29: 69–77.

Chapter 2 is currently in review for publication in *Marine Mammal Science*. Leslie, Matthew S.; Archer, F. I.; Morin, Phillip A. The dissertation author was the primary investigator and lead author of this material.

Chapter 3 is currently being prepared for submission for publication of the material. Leslie, Matthew S.; Morin, Phillip A. The dissertation author was the primary investigator and author of this material.

Chapter 4 is currently being prepared for submission for publication of the material. Leslie, Matthew S.; Morin, Phillip A. The dissertation author was the primary investigator and author of this material.

VITA

2002	Bachelor of Science, Oklahoma State University
2002-2005	Conservation Genetic Program and Cetacean Research Program Wildlife Conservation Society, Bronx, NY
2005-2007	Conservation Genetic Program American Museum of Natural History, New York, NY
2007-2009	Sackler Institute of Comparative Genomics American Museum of Natural History, New York, NY
2011	Master of Science, University of California, San Diego
2016	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

20. LESLIE M. S., F. I. ARCHER, P. A. MORIN. Genetic inference of population structure in spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) from the eastern tropical Pacific Ocean based on mitochondrial genomes and nuclear SNPs. *In Review* – Marine Mammal Science.

19. H. C. ROSENBAUM, M. MENDEZ, F. KERSHAW, C. POMILLA, **M.S. LESLIE,** K. FINDLAY, P. BEST, T. COLLINS et al. Population structure of humpback whales throughout the Southern Hemisphere and the Arabian Sea. *In Review* – Biological Conservation

18. P.E. ROSEL, B. L. TAYLOR, B. L. HANCOCK-HANSER, P. A. MORIN, F. I. ARCHER, A. KONOPACKI, A. R. LANG, S. L. MESNICK, V. L. PEASE, W. F. PERRIN, K. M. ROBERTSON, **M. S. LESLIE** *et al.* Molecular genetic markers and analytical approaches that have been used for delimiting marine mammal subspecies and species. *In Review* – Marine Mammal Science.

17. K. K. MARTIEN, **M. S. LESLIE**, P. A. MORIN, F. I. ARCHER, B. L. HANCOCK-HANSER, ROSEL P.E. *et al.* Analytical approaches to subspecies delimitation with genetic data. *In Review* – Marine Mammal Science.

16. B. L. TAYLOR, B. L. HANCOCK-HANSER, K. K. MARTIEN, P. A. MORIN, F. I. ARCHER, A. R. LANG, **M. S. LESLIE,** S. L. MESNICK, V. L. PEASE, W. F. PERRIN, and K. M. ROBERTSON *et al.* Guidelines and quantitative standards to

improve consistency in cetacean subspecies and species delimitation relying on molecular data. *In Review* – Marine Mammal Science.

15. POMILLA C., A R. AMARAL, T. COLLINS, G. MINTON, K. FINDLAY, **M. S. LESLIE**, L. PONNAMPALAM, R. BALDWIN, H. C. ROSENBAUM. **2014.** The world's most isolated and distinct whale population: Humpback whales of the Arabian Sea. PloS ONE. DOI: 10.1371/journal.pone.0114162

14. M. S. LESLIE. 2014. The impact of phylogenetic nomenclature on the efficacy of the Endangered Species Act. Conservation Biology. 00:1-9. DOI: 10.1111/cobi.12375.

13. JACKSON J. D. J. STEELE, P. BEERLI, B. C. CONGDON, C. OLAVARRIA, **M. S. LESLIE**, C. POMILLA, H. C. ROSENBAUM, C. S. BAKER, **2014.** Global genomic diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). Proceedings of the Royal Society of London Series B. 281:1786. doi: 10.1098/rspb.2013.3222

12. I. CARVALHO, J. AY LING LOO, T. COLLINS, J. BARENDSE, C. POMILLA, **M. S. LESLIE,** P. BEST, H. ROSENBAUM. **2014.** Does temporal and spatial segregation explain the complex population structure of humpback whales on the coast of West Africa? Journal of Marine Biology. DOI:10.1007/s00227-013-2379-1

11. F. KERSHAW, **M.S. LESLIE,** T. COLLINS, R. MANSUR, M. RUBAIYAT, B.D. SMITH, G. MINTON, R. BALDWIN, et al. **2013.** Population differentiation of 2 forms of Bryde's whales in the Indian and Pacific Oceans. Journal of Heredity. doi:10.1093/jhered/est057

10. F. I. ARCHER, P. A. MORIN, B. L. HANCOCK-HANSER, K. M. ROBERTSON, **M. S. LESLIE,** M. BERUBE, S. PANIGADA, B. L. TAYLOR. **2013.** Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus spp.*): genetic evidence for revision of subspecies. PLoS ONE 8(5): e63396. doi:10.1371/journal.pone.0063396

9. B. HANCOCK-HANSER, A. FREY, **M.S. LESLIE,** P. DUTTON, F. ARCHER, P. MORIN. **2013.** Targeted multiplex next-generation sequencing: Advances in techniques of mitochondrial and nuclear sequencing for population genomics. Molecular Ecology Resources. 13(2):254-68. doi10.1111/1755-0998.12059

8. P.J. ERSTS, C. POMILLA, J. KISZKA, S. CERCHIO, H.C. ROSENBAUM, M. VELY, Y. RAZAFINDRAKOTO, J.A. LOO, **M.S. LESLIE**, M. AVOLIO. **2011** Observations of individual humpback whales utilizing multiple migratory destinations in the south-western Indian Ocean. African Journal of Marine Science. 33(2): 333-338. doi: 10.2989/1814232X.2011.600436.

7. G. MINTON, S. CERCHIO, T. COLLINS, P. ERSTS, K.P. FINDLAY, C. POMILLA, D. BENNET, M.A. MEYER, Y. RAZAFINDRAKOTO, P.G.H. KOTZE, W.H. OOSTHUIZEN, **M.S. LESLIE,** N. ANDRIANARIVELO, R. BALDWIN, L. PONNAMPALAM, AND H.C. ROSENBAUM. **2010.** A note on the comparison of humpback whale tail fluke catalogues from the Sultanate of Oman with Madagascar and the East African mainland. Journal of Cetacean Research and Management 11(1):65-68.

6. H. C. ROSENBAUM, C. POMILLA, M. MENDEZ, **M.S. LESLIE**, P.B. BEST, et al. **2009.** Population structure of humpback whales from their breeding grounds in the south Atlantic and Indian Oceans. PLoS ONE 4(10): e7318. DOI:10.1371/journal.pone.0007318.

5. M. J. EATON, G. L. MEYERS, S. O. KOLOKOTRONIS, **M. S. LESLIE,** A. P. MARTIN, G. AMATO. **2009.** Barcoding bushmeat: molecular forensic identification of Central African and South American harvested vertebrates. Conservation Genetics. DOI 10.1007/s10592-009-9967-0.

4. M. MENDEZ, **M. S. LESLIE. 2009.** Cetacean mitogenomics. MitoCommunications; Mitochondrial DNA; DOI: 10.1080/19401730902852976.

3. M. H. ENGEL, N. J.R. FAGUNDES, H. C. ROSENBAUM, **M. S. LESLIE,** P. H. OTT, R. SCHMITT, E. SECCHI, L. DALLA ROSA, S. L. BONATTO. **2008.** Mitochondrial DNA diversity and assessment of the likely feeding ground of the Southwestern Atlantic humpback whale (*Megaptera novaeangliae*) breeding area. Conservation Genetics: 9(5):1253-1262.

2. M. KRETZMAN, L. MENTZER, R. DiGIOVANNI, **M. S. LESLIE,** G. AMATO. **2006.** Microsatellite diversity and fitness in stranded harp seals (*Phoca groenlandica*). Journal of Heredity 97(6):555–560.

1. M. S. LESLIE, A. BATIBASAGA, D. S. WEBER, D. OLSEN, AND H. C. ROSENBAUM. **2005.** First record of Blainville's beaked whale *Mesoplodon densirostris* in Fiji. Pacific Conservation Biology 11(4):28-30.

ABSTRACT OF THE DISSERTATION

Population Genomic Structure and Phylogeography of Spotted and Spinner Dolphins (Genus Stenella)

by

Matthew S. Leslie

Doctor of Philosophy in Marine Biology

University of California, San Diego, 2016

Professor Phillip Morin, Chair
Professor Ronald Burton, Chair

This dissertation provides empirical evidence for patterns of population structure – a necessary prerequisite for estimating abundance and ultimately conservation action - of two species of pelagic dolphins (spinner and spotted dolphins) that have been challenging to characterize using modern molecular genetics. I evaluated philosophies that could aid in accurate categorization of biological diversity for conservation and developed and employed novel techniques to collect genome-wide data in efforts to increase statistical power for testing hypotheses of genetic structure. For the former, I assessed the impacts of phylogenetic nomenclature (PN) on the efficacy of the Endangered Species Act (ESA). PN relies heavily on traditional nomenclature, and knowledge of evolutionary relationships is paramount for species protection, so I concluded that PN will have little impact the ESA. Using DNA capture and highly-parelleled sequencing, I collected whole mitochondrial genomes (mtDNA) and scores of nuclear loci (nuDNA) for population structure tests. MtDNA showed weak but significant

differences between subspecies of spotted dolphins and the first mtDNA evidence for differentiation between ETP spinner dolphin subspecies. NuDNA supported subspecies of spotted dolphins, but not spinner dolphins. Strong differentiation was detected between whitebelly and eastern spinner stocks, but these data did not have statistical power to enable testing of population-level hypotheses needed for management, so I employed a genome-wide genotyping approach (RADseq) to collect >6,000 single nucleotide polymorphisms (SNPs) from regions throughout the genome associated with the restriction enzyme site *PstI*. Results supported the current subspecies for both species and indicate stock-level separation for Tres Marias spinner dolphins and two offshore spotted dolphin stocks the ETP. I also tested global taxonomic and phylogeographic hypotheses using RADseq and found deep divergence between Indo-Pacific and eastern tropical Pacific Ocean (ETP) spinner dolphins, but segregation between inshore and offshore ETP spotted dolphins. Australain dwarf spinner dolphins were genetically distant from conspecifics in Indonesia, but the dwarf spinner dolphin was monophyletic. Atlantic spinner dolphins were placed between a clade of Indo-Pacific Ocean populations and ETP population in our phylogeny, contrary to previous hypotheses. The eastern Pacific basin and the marine Wallace's line are strong barriers for spinner dolphins despite high dispersal potential.

INTRODUCTION

One of the grand challenges facing the human race is conserving biological diversity in the face of rapid extinction (Pimm et al. 1995, Mace G.M. et al. 2005). An effective strategy for stemming extinction is species-based prioritization and protection, whereby laws protect populations of organisms and their habitat. Some of the most powerful environmental laws in the United States are species-based strategies, including: the Endangered Species Act, the Migratory Bird Act, the Marine Mammal Protection Act, and others.

Species-based conservation strategies are only effective, however, if actions are applied at the appropriate scale. Without knowledge of how organisms are geographically structured we may fail to recognize entire subspecies or distinct populations that require urgent conservation action (Leslie 2014). This is because knowledge of taxon boundaries is a necessity for assessing threats, estimating abundance, and ultimately determining if conservation efforts are required and what appropriate mortality limits should be (Wheeler et al. 2004). Maybe most importantly, taxonomic information is a direct input into the process of listing organisms under protective legislation (e.g., the U.S. Endangered Species Act - ESA), biodiversity status inventories (e.g., the International Union for the Conservation of Nature's Red List), and international conservation agreements (e.g., the Convention on the International Trade of Endangered Species - CITES), making it a necessary precondition for conservation measures involving law enforcement and adjudication.

Equipped with new tools in molecular biology, analytical tools, and digital database infrastructure, taxonomists are poised to continue to tackle another grand challenge (inseparable from conservation): discovering and documenting biological diversity (Wheeler 2010). Because of heir open ocean habitats, the documentation of pelagic organisms is inherently challenging – and has lagged behind terrestrial taxonomy (Palumbi 1996, Bowen 1997, Bowen et al. 2001). Pelagic organisms also show few permanent physical barriers to gene flow, making developing hypotheses to test for taxonomic boundaries more complex. For instance, nuanced factors such as demographic processes (e.g., range expansion and metapopulation dynamics (Horne 2014)) may be more important in structuring populations than physical separation. Thus, in genetic terms, the dynamic nature of the pelagic environment increases the likelihood of mixing between populations; even a few successful migrants per generation greatly decrease the signal of population segregation (Waples 1998). Pelagic populations often have large abundances as well, which can result in a large amount of standing genetic variation. High genetic diversity can dramatically increases the time needed for populations to drift apart genetically, even in the complete absence of gene flow (Taylor and Dizon 1996). Spinner and spotted dolphins in the eastern tropical Pacific Ocean (ETP) illustrate this problem. They historically numbered several million (Wade et al. 2007), but starting in the 1960s, hundreds of thousands were killed annually as bycatch in the dolphin-set tuna purse-seine fishery (Lo and Smith 1986, National Research Council 1992, Wade 1995).

Relative to terrestrial mammals, there are few species and subspecies with extant cetacean genera; researchers believe this is due to under-classification rather than a lack

of substructure (Taylor et al. *In Review*). This lack of knowledge about cetacean variation stems from a number of challenges. Foremost, they are relatively elusive, inaccessible, and difficult to observe and capture. Second, collection and storage of cetacean specimens (and associated data) is also logistically complicated. Collections events are rare as most dead specimens sink to the ocean floor. As for storage, a series of birds or fish could be represented in a single museum drawer, it would take several warehouses to story a series large whale specimens. Finally, although the laws that protect species are desperately needed - collection and transfer of specimens is complicated by these national laws and international agreements, such as the CITES. All these reasons combine to make accurate cetacean taxonomy and population structure challenging.

The impetus for this study was that managers and biologist had not seen the increase in abundance expected after laws were put into place to limit the amount of bycatch pressure on these species (Gerrodette and Forcada 2005). This led some to question if the spatial scale used for management was biologically meaningful and appropriate for ensuring maximum recovery. For instance, if the population boundaries used in management were too broad (i.e., not protecting populations, but instead protecting groups of populations), acute or ongoing localized impact could be inhibiting population rebound).

Fortunately, during the peak of the tuna-dolphin crisis in the ETP, researchers showed the forethought to collect biological specimens for future study. The bulk of these samples were from spinner and spotted dolphins – the two heaviest hit species. Initially,

researchers used skull morphology and external body characteristics to determine biological differences that form the basis of taxonomy. These studies demonstrated morphometric differences supporting structured populations (Perrin et al. 1991, Perrin et al. 1994, Perryman and Westlake 1998). However, molecular genetics approaches have not found corresponding population genetic structure (Dizon et al. 1994, Galver 2002), although Andrews *et al*. (2013) found some evidence for segregation in data from the Y-chromosome and Escorza-Treviño *et al*. (2005) found population structure within the coastal subspecies of spotted dolphins.

The reasons for this lack of population genetic structure could be biological or methodological. As mentioned, these two species had very high abundance prior to extensive bycatch, dramatically increasing the time needed for populations to drift apart genetically. In addition, these two species are likely very recently evolved (< 5Ma), meaning these differences haven't had much time to accrue (McGowen et al. 2009). Finally, it is highly likely that ongoing geneflow is continuing to 'reshuffle the deck' of genetic variation and inhibiting populations from showing traditional signals of genetic structure due to drift (Andrews et al. 2013). Therefore, our inability to find population genetic structure could reflect the biological reality - that there are very few genetic differences - and the physical differences in the morphology are due to gene expression or environmentally driven factors. However, given the concrete nature of the cranial characters used to erect subspecies in the ETP, it is highly likely that there are underlying genetic differences.

Instead, I believe it is because of limits of previous methods that we haven't found population genetic structure. Previous studies could be negatively effected by two different issues: 1) focusing on neutral genetic markers only, and 2) not having the statistical power needed to find subtle differences in the populations because of too few genetic markers.

My dissertation aims to address some of these possible shortcomings of previous studies by employing new philosophies and techniques to understand and document cetacean diversity. I examined the principles of phylogenetic nomenclature as embodied in the PhyloCode (an alternative to traditional rank-based nomenclature that names biological groups based on the results of phylogenetic analyses and does not associate taxa with ranks) and assessed how this novel approach to naming taxa might affect the implementation of species-based legislation by providing a case study of the ESA (Chapter 1). Next, I addressed some of the aforementioned challenges with assessing population structure of eastern tropical Pacific Ocean dolphins by employing next generation DNA sequencing technology (Chapters 2 & 3). Finally, in Chapter 4, I aimed to place the diversity of ETP dolphins within the context of global diversity of both of these species using the same techniques to collect data for a comparative study testing hypotheses of global phylogeography (Davies 1963, Barber et al. 2000, Perrin 2007). Critical to curbing the crisis of biodiversity loss is the task of classifying life on this planet (Mace 2004); my hope is that this work will directly advance both of these grand challenges.

Literature Cited:

Andrews, K. R., W. F. Perrin, M. Oremus, L. Karczmarski, B. W. Bowen, J. B. Puritz, and R. J. Toonen. 2013. The evolving male: spinner dolphin (*Stenella longirostris*) ecotypes are divergent at Y chromosome but not mtDNA or autosomal markers. Molecular Ecology 22:2408-2423.

Barber, P. H., S. R. Palumbi, M. V. Erdmann, and M. K. Moosa. 2000. Biogeography: A marine Wallace's line? Nature 406:692-693.

Bowen, B. W. 1997. Complex Population Structure and the Conservation Genetics of Migratory Marine Mammal: Lessons from Sea Turtles. Pages 77-84 *in* A. E. Dizon, S. J. Chivers, and W. F. Perrin, editors. Molecular Genetics of Marine Mammals. Society for Marine Mammology, Lawrence, KS.

Bowen, B. W., A. L. Bass, L. A. Rocha, W. S. Grant, and D. R. Robertson. 2001. Phylogeography of the trumpetfishes (Aulostomus): ring species complex on a global scale. Evolution 55:1029-1039.

Davies, J. L. 1963. The Antitropical Factor in Cetacean Speciation. Evolution 17:107-116.

Dizon, A. E., W. F. Perrin, and P. A. Akin. 1994. Stocks of dolphins (*Stenella* spp.) and (*Delphinus delphis*) in the eastern tropical Pacific: A phylogeographic classification. NOAA Technical Report NMFS 119:20.

Escorza-Trevino, S., F. I. Archer, M. Rosales, A. Dizon, and A. Lang. 2005. Genetic differentiation and intraspecific structure of Eastern Tropical Pacific spotted dolphins, *Stenella attenuata*, revealed by DNA analyses. Pages 1-14 Conservation Genetics.

Galver, L. 2002. The molecular ecology of spinner dolphins (*Stenella longirostris*): genetic diverstiy and population structure. Ph.D. Dissertation. University of California, San Diego, CA.

Gerrodette, T., and J. Forcada. 2005. Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. Marine Ecology Progress Series 291:1-21.

Horne, J. B. 2014. Thinking outside the barrier: neutral and adaptive divergence in Indo-Pacific coral reef faunas. Evolutionary Ecology 28:991-1002.

Leslie, M. S. 2014. Impacts of phylogenetic nomenclature on the efficacy of the U.S. Endangered Species Act. Conservation Biology 29:69-77.

Lo, N. C. H., and T. D. Smith. 1986. Incidental mortality of dolphins in the eastern tropical Pacific, 1959-72. Fishery Bulletin 84:27-34.

Mace G.M., Masundire H., and B. J.E.M. 2005. Chapter 4. Biodiversity. Pages 77-122 Millennium Ecosystem Assessment: Ecosystems and Human Well-Being: Current State and Trends. Findings of the Condition and Trends Working Group. Island Press, Washington DC.

McGowen, M. R., M. Spaulding, and J. Gatesy. 2009. Divergence date estimation and a comprehensive molecular tree of extant cetaceans. Molecular phylogenetics and evolution 53:891-906.

National Research Council. 1992. Dolphins and the Tuna Industry. National Academy Press, Washington, DC.

Palumbi, S. R. 1996. What can molecular genetics contribute to marine biogeography? An urchin's tale. Journal of Experimental Marine Biology and Ecology 203:75-92.

Perrin, W. F. 2007. The South African species gate. Whales and dolphins of the southern African subregion. Ed. Best, P. B.

Perrin, W. F., P. A. Akin, and J. V. Kashiwada. 1991. Geographic variation in external morphology of the spinner dolphin (*Stenella longirostris*) in the eastern Pacific and implications for conservation. Fishery Bulletin 89:411-428.

Perrin, W. F., G. D. Schnell, D. J. Hough, J. W. Gilpatrick, Jr., and J. V. Kashiwada. 1994. Reexamination of geographic variation in cranial morphology of the pantropical spotted dolphin, (*Stenella attenuata*), in the eastern Pacific. Fishery Bulletin 92:324-346.

Perryman, W. L., and R. L. Westlake. 1998. A new geographic form of the spinner dolphin, Stenella longirostris, detected with aerial photogrammetry. Marine Mammal Science 14:38-50.

Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks. 1995. The future of biodiversity. Science 269:347.

Taylor, B. L., and A. E. Dizon. 1996. The need to estimate power to link genetics and demography for conservation. Conservation Biology 10:661-664.

Taylor, B. L., W. F. Perrin, J. R. L. Brownell, P. E. Rosel, C. S. Baker, J. Y. Wang, R. R. Reeves, and F. Cipriano. In Review. Why we should develop guidelines and quantitative standards to use genetic data for subspecies delimitation among data-poor organisms like cetaceans. Marine Mammal Science.

Wade, P. R. 1995. Revised estimates of incidental kill of dolphins (Delphinidae) by the purse-seine tuna fishery in the eastern tropical Pacific, 1959-1972. Fishery Bulletin 93(2):345-354.

Wade, P. R., G. M. Watters, T. Gerrodette, and S. B. Reilly. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. Marine Ecology Progress Series 343:1-14.

Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity 89:438-450.

Wheeler, Q. 2010. What would NASA do? Mission-critical infrastructure for species exploration. Systematics and Biodiversity 8:11-15.

Wheeler, Q. D., P. H. Raven, and E. O. Wilson. 2004. Taxonomy: impediment or expedient? Science (New York, NY) 303:285.

CHAPTER 1:

Impacts of Phylogenetic Nomenclature on the Efficacy on the Endangered Species Act

Matthew S. Leslie[1]

1. Scripps Institution of Oceanography, University of California San Diego, mail-code 0208, 9500 Gilman Dr. La Jolla, CA, 92093 USA

Conservation Biology

Essay

Impacts of phylogenetic nomenclature on the efficacy of the U.S. Endangered Species Act

Matthew S. Leslie

Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Dr. Mailcode 0208, La Jolla CA 92038, U.S.A., email: mleslie@ucsd.edu

Abstract: *Cataloging biodiversity is critical to conservation efforts because accurate taxonomy is often a precondition for protection under laws designed for species conservation, such as the U.S. Endangered Species Act (ESA). Traditional nomenclatural codes governing the taxonomic process have recently come under scrutiny because taxon names are more closely linked to hierarchical ranks than to the taxa themselves. A new approach to naming biological groups, called phylogenetic nomenclature (PN), explicitly names taxa by defining their names in terms of ancestry and descent. PN has the potential to increase nomenclatural stability and decrease confusion induced by the rank-based codes. But proponents of PN have struggled with whether species and infraspecific taxa should be governed by the same rules as other taxa or should have special rules. Some proponents advocate the wholesale abandonment of rank labels (including species); this could have consequences for the implementation of taxon-based conservation legislation. I examined the principles of PN as embodied in the PhyloCode (an alternative to traditional rank-based nomenclature that names biological groups based on the results of phylogenetic analyses and does not associate taxa with ranks) and assessed how this novel approach to naming taxa might affect the implementation of species-based legislation by providing a case study of the ESA. The latest version of the PhyloCode relies on the traditional rank-based codes to name species and infraspecific taxa; thus, little will change regarding the main targets of the ESA because they will retain rank labels. For this reason, and because knowledge of evolutionary relationships is of greater importance than nomenclatural procedures for initial protection of endangered taxa under the ESA, I conclude that PN under the PhyloCode will have little impact on implementation of the ESA.*

Keywords: phylogenetics, taxonomy

Impactos de la Nomenclatura Filogenética sobre la Eficiencia del Acta Estadunidense para las Especies en Peligro

Resumen: *Catalogar a la biodiversidad es crítico para los esfuerzos de conservación porque la taxonomía precisa continuamente es una condición previa para la protección bajo las leyes diseñadas para la conservación de especies, como el Acta Estadunidense para las Especies en Peligro (ESA, en inglés). Los códigos de nomenclatura tradicional que gobiernan los procesos taxonómicos han caído recientemente bajo escrutinio ya que los nombres de los taxones están vinculados de manera más cercana con los rangos jerárquicos que con los propios taxones. Una nueva estrategia para nombrar a los grupos biológicos, llamada nomenclatura filogenética (NF), denomina explícitamente a los taxones al definir sus nombres en términos de linaje y descendencia. La nomenclatura filogenética tiene el potencial de incrementar la estabilidad de la nomenclatura y disminuir la confusión causada por los códigos basados en rangos, pero quienes proponen a la NF han luchado contra la idea de si las especies y los taxones infra-específicos deberían ser gobernados por las mismas reglas que otros taxones o si deberían tener reglas especiales. Algunos de los que proponen abogan el abandono total de las etiquetas por rango (incluyendo a las especies); esto podría tener consecuencias para la implementación de legislación de conservación basada en taxones. Examiné los principios de la NF como están representados en el PhyloCode (una alternativa para la nomenclatura tradicional basada en rangos que nombre a los grupos biológicos con base en los resultados de análisis filogenéticos y no asocia a los taxones con rangos) y evalué cómo esta estrategia novedosa para nombrar taxones puede afectar la implementación de legislación basada en especies al proporcionar un estudio de caso de la ESA. La versión más reciente del PhyloCode depende de los códigos tradicionales basados en rangos para nombrar a las*

Paper submitted December 19, 2013; revised manuscript accepted May 25, 2014.

Conservation Biology, Volume 29, No. 1, 69–77

DOI: 10.1111/cobi.12375

especies y a los taxones infra-específicos; entonces, poco cambiará con respecto a los objetivos principales de la ESA porque estos mantendrán sus etiquetas de rango. Por esta razón, y porque el conocimiento de las relaciones evolutivas es de mayor importancia que los procedimientos de nomenclatura para la protección inicial bajo la ESA de taxones en peligro, concluyo que la NF bajo el PhyloCode tendrá un impacto mínimo en la implementación de la ESA.

Palabras Clave: filogenética, taxonomía

Introduction

The classification of life on this planet is critical to curbing the crisis of biodiversity loss (Mace 2004). Knowledge of taxon boundaries is a necessity for assessing threats, estimating abundance, and ultimately determining if conservation efforts are required (Wheeler et al. 2004). Maybe most importantly, taxonomic information is a direct input into the process of listing organisms under protective legislation (e.g., the U.S. Endangered Species Act [ESA]), biodiversity status inventories (e.g., the International Union for the Conservation of Nature's Red List), and international conservation agreements (e.g., the Convention on the International Trade of Endangered Species), making it a necessary precondition for conservation measures involving law enforcement and adjudication.

Classification and nomenclature are separate processes. Taxa (i.e., groups of organisms) are delimited based on analyses of morphological and genetic data, and then these groups are named according to rules (codes) of nomenclature (de Queiroz 2006) (e.g., International Code on Zoological Nomenclature [ICZN] or the International Code of Nomenclature for algae, fungi, and plants, etc.). Traditional nomenclatural codes have been scrutinized for many reasons; paramount to this essay is that they link hierarchical ranks (kingdom, phylum, class, etc.) to taxon names (e.g., de Queiroz & Gauthier 1994; Pleijel & Rouse 2003; Dayrat et al. 2008; Cellinese et al. 2012), which creates confusion and nomenclatural instability.

Phylogenetic nomenclature (PN) (www.phylocode.org) is an alternative to traditional rank-based nomenclature that names biological groups based on the results of phylogenetic analyses and does not associate taxa with ranks (although PN does not prohibit assigning ranks as a process unrelated to naming) (Cantino & de Queiroz 2014). In PN, the term *species* is a special case. (The term *species* can apply to both a rank category and a biological entity or group and is used differently by different people and in different contexts. I have tried to use it unambiguously.) Many proponents of PN view species as a kind of evolutionary unit, rather than considering it a taxonomic rank. Some PN proponents disagree with this and advocate for the removal of all ranks, including species (Mishler 2010; Cellinese et al. 2012).

Although the latter, more extreme position could have negative ramifications for the implementation of taxon-based conservation legislation, I am not aware of any published assertions that PN per se would have disastrous consequences for such efforts. However, changing the rules of nomenclature is a scary idea to some, especially those charged with implementing conservation legislation that is dependent on stable taxonomy. My intention is to familiarize conservation practitioners with PN as an alternative to rank-based nomenclature and to stimulate thoughtful discussion about the future interactions of biological nomenclature and species-based conservation. It derives from concerns of colleagues who are fearful of the disastrous consequences of nomenclatural regime change. My objectives were to contrast PN with the ICZN to illustrate points that are relevant to ecologists and apply broadly to all rank-based codes of biological nomenclature and to assess the possible impact of PN on species-based conservation policies. A comprehensive review of all species-based conservation legislation was not feasible, so I focused on the implementation of the ESA as a case study. I have tried to provide a balance of detailed scrutiny and broad context to avoid being parochial or nebulous.

The ESA and Modern Taxonomy

With the ESA, biological nomenclature became a strange bedfellow with legal nomenclature, where clarity and stability of terminology are critical to the enforcement of laws. Given this dependency, it is vitally important to provide precision and transparency in the process of biological taxonomy (including nomenclature).

The ESA affords protection to listed species until they are no longer imperiled; this can require millions of taxpayer dollars and restrictions of stakeholder rights over decades or centuries. After a petition is put forth to the federal government, a small panel of experts (a biological review team [BRT]) reviews all the best available scientific evidence and issues recommendations for listing or not. The crossing over from unlisted to listed is a critical juncture for the particular organism in question and the affected stakeholders. It is also often heavily scrutinized and politicized.

Legal protection for plants and vertebrate animals extends to species, subspecies, and distinct population

segments (DPS). For vertebrates to be protected, these units must be "reproductively isolated from other conspecific population units and represent an important component in the evolutionary legacy of the species" (Waples 1991). The best practices for designating these units are often specific to fields of study (Reeves 2004; Tobias et al. 2010), although there are notable exceptions (e.g., Haig et al. 2006). One commonality is that experts rely heavily on phylogenetics to determine whether taxa qualify as distinct and whether a DPS is an important component of the evolutionary legacy of a species, as required for listing under the ESA (Taylor 2005; Fallon 2007; Kelly 2010).

Phylogenetic analyses do not require molecular data, but the rapid development (and decreasing cost) of DNA sequencing and sophisticated analytical methods to test phylogenetic hypotheses (Yang & Rannala 2012) have greatly modernized determination of evolutionary relationships on which modern classification is based (Jörger & Schrödl 2013). This has given rise to the need for guidance on the use of these data for listing under the ESA (Waples 1991; Haig et al. 2006; Fallon 2007; Kelly 2010).

Modernizing Nomenclature

The system used to name groups of organisms and the root language used to communicate about the diversity of life is called biological nomenclature. The ICZN requires that the name of a new species (the group) be linked to a holotype or type specimen, that the name (called the binomen) consists of 2 parts, and that this information be made broadly available. In traditional nomenclature, the binominal name is the basic syntax for the rank of species. It consists of the genus name and a specific epithet. Together these must form a unique identifier for a group of organisms at this rank (e.g., for the common bottlenose dolphin, *Tursiop truncatus* [Montagu 1821]). The use of binominal species names has been in place since Linnaeus (1753) used them over 250 years ago and has, for the most part, served us well. Recently, however, several deficiencies of traditional nomenclature (including the rank and unit of species) have been noted (e.g., de Queiroz & Gauthier 1994; Pleijel & Rouse 2003; Dayrat et al. 2008; Cellinese et al. 2012). The main issue is the linkage of ranks (family, genus, species etc.) to taxon names in the traditional system.

Taxa with equivalent rank are often presented as sharing (or are assumed to share) comparable evolutionary characteristics, which is in most cases absolutely false (unless the taxa being compared are sister groups). For instance, there are many different kinds of species (the groups) in nature (Mishler 1999; Mishler 2010), but there is only one vague set of rules to govern the naming of such groups. A single system for naming such an array of biological groups is likely impossible, although see de Queiroz (1999, 2005) for proposals.

The application of phylogenetics to taxonomy is vitally important for the continued pursuit of accuracy in biology and to provide the best available science for species conservation. These methods infer phylogeny on the basis of shared, derived similarities, which could be genetic or phenotypic, between taxa that evolved from a single common ancestor (Wiley et al. 1991). Novel methods of cataloging biological diversity apply phylogenetic principles to delimit taxa based on ancestry and descent of organisms (visualized as clades) and then name those clades free of any rank associations (e.g., using the PhyloCode [Cantino & de Queiroz 2014]).

The PhyloCode

The International Society for Phylogenetic Nomenclature (ISPN) is developing the International Code of Phylogenetic Nomenclature (the PhyloCode) to be used within the context of the Tree of Life, as reconstructed via modern phylogenetics. The newest version of the PhyloCode (Cantino & de Queiroz 2014) has been approved by the Committee on Phylogenetic Nomenclature. It will be publically available online soon (www.phylocode.org) and in the meantime is available upon request from the authors.

PN provides an alternative to the traditional system. It does not replace existing taxon names; instead, it offers a more stable system for governing names already in existence and creating new names for previously unnamed clades. It incorporates the principle that biological classification should represent phylogenetic relationships based on Darwinian evolution via common descent (Cantino & de Queiroz 2014). The 2 main tenets of the PhyloCode are "[n]ames are applied to phylogenetically conceptualized taxa (i.e., clades, defined as an ancestor (an organism, population or species) and all its descendants)" and "[t]he system is independent of categorical ranks."

The PhyloCode increases nomenclatural stability by replacing rank-based names and definitions with explicit clade-based definitions. Under rank-based nomenclature, the names of one or both of the taxa in Fig. 1 would have to change because a taxon of a given rank cannot be nested inside a taxon of the same rank (as implied by the names). Under the PhyloCode, the names Chamaeleonidae and Agamidae both retain their associations with the clades originating in the most recent common ancestor of the closed and open circle species, respectively. However, in Fig. 1b Chamaeleonidae is thought to have descended from the ancestor of the open circle species. The manner in which the definitions are stated ensures that no names designate paraphyletic taxa. Neither splitting nor

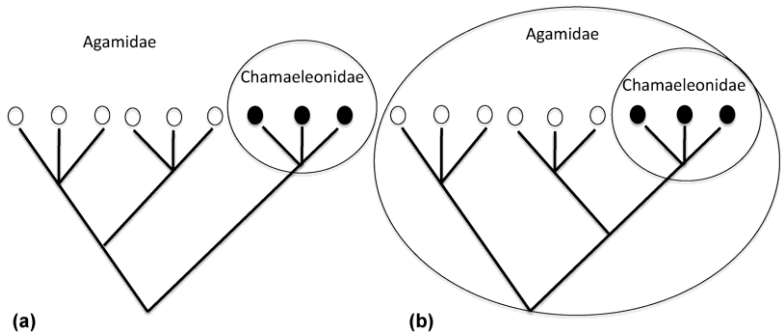


Figure 1. Example of how taxonomic names are affected by a change in rank under rank-based nomenclature and how they are not affected under phylogenetic nomenclature (PN). (a) A cladogram showing relationships implied by an earlier taxonomy according to which the following definitions are formulated: Agamidae, a clade originating in the most recent common ancestor of the species represented by open circles; Chamaeleonidae, a clade originating in the most recent common ancestor of the species represented by filled circles. (b) A cladogram representing a revised hypothesis of relationships based on new data or methods of analysis. Reprinted from Trends in Ecology & Evolution*, Vol. 9, de Queiroz and Gauthier, Toward a Phylogenetic System of Biological Nomenclature, Page 30, Copyright (1994), with permission from Elsevier.*

lumping occurs, and hierarchical relations may be altered without requiring a name change.

The use of explicit clade-based names maintains nomenclatural stability even when ranks are assigned to clades (Fig. 2). This occurs despite the taxonomic relationships remaining the same. Instead of the name being tied to the group of organism the name is tied to the rank, which is misleading and unstable. In a rank-free system (Fig. 2b), when the ranks of the 3 example taxa are changed from suborder and family (part 1) to family and subfamily (part 2), the names of all 3 taxa remain unchanged because the application of names is based on phylogenetic relationships (which have not changed) and are independent of ranks.

Despite the similarity in name, PN does not rely on any of the tenets of the phylogenetic or "diagnostic" species concept (PSC) (Cracraft 1983, 1989; Nixon & Wheeler 1990). Several authors have expressed concerns about the PSC. They worry about, for example, inflating numbers of endangered taxa (and the cost of enforcement) due to the nature of the diagnostic character approach of PSC (Agapow et al. 2004) and legal and regulatory restrictions of "carrying out crosses between distinct PSC species" in the case of genetic rescue (outcrossing small inbred populations with low genetic diversity to reverse inbreeding depression and loss of genetic diversity) (Frankham et al. 2012). Similar to the ICZN, PN (as represented in version 5 of the PhyloCode [Cantino & de Queiroz 2014]) aims to be completely independent of any species concept and thereby removed from the species-concept debate.

The PhyloCode and the ESA

PN intersects with the management of endangered species in the United States in Section 4 of the ESA, which outlines the listing process. There are 2 points in this process at which taxonomy is important: defining the biological unit to conserve and the threats to the unit and prioritization of the listing and the recovery plan. Prioritization is often less scrutinized. Taxonomic considerations occupy a lower priority tier below magnitude and immediacy of the threat (U.S. Congress 1983). Table 1 depicts the decision-making process for prioritization. The DPS is not included on this table but is normally prioritized equivalent to or below the subspecies level. Monotypic genera, those with only a single species, are afforded priority in the listing process and in the allocation of funding for protection.

The PhyloCode deals with the biological units protected under the ESA (species, genera, and subspecies) in the following ways.

SPECIES

From the outset, PN has focused mostly on naming clades, and the ISPN has struggled with whether and how to name species (see Dayrat et al. [2008] for a detailed history and references). Initially, most of the debate centered on what form species names should take in PN (e.g., Cantino et al. 1999). In 2007 the ISPN Committee on Phylogenetic Nomenclature (CPN) adopted a stance toward species names that was incorporated in version 4 of the Phylocode (Cantino & de Queiroz 2010). Under this

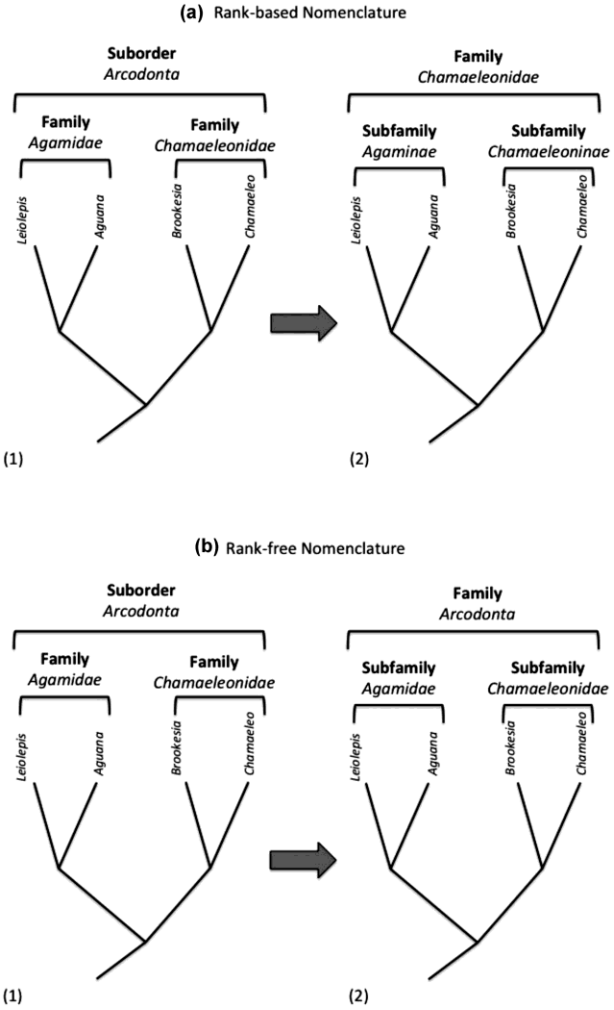


Figure 2. Illustration of the effects of changes in ranks under rank-based nomenclature versus rank-free nomenclature. The ranks of the example taxa are changed but the relationships remain identical. Under rank-based nomenclatural rules the names must change (a). Under rank-free nomenclature, however (shown in b), the names are applied with the phylogenetic definitions (based on species originally included in the taxa) (Agamidae, least inclusive clade containing both Leiolepis guttata *and* Agama agama*; Chamaeleonidae, least inclusive clade containing both* Brookesia superciliosa *and Chamaeleo chamaeleon; Acrodonta, least inclusive clade containing both* A. agama *and* C. chamaeleon*). Reprinted from* Bibliotheca Herpetologica*, Copyright (2012), with permission from author and The International Society for the History and Bibliography of Herpetology.*

approach, "governance of establishment and precedence of new species names is left to the rank-based codes, but provision are added to the PhyloCode that allow those names to be interpreted and used in a way that is consistent with the basic principles of phylogenetic nomenclature" (Dayrat et al. 2008: 510).

Since the 2007 decision to adopt Dayrat et al.'s (2008) stance toward species names in the PhyloCode, debate has continued with the ISPN between those who consider species to be a kind of biological entity distinct from a clade (and which should therefore have distinct naming rules) and those who believe that many kinds of entities are referred to as species or that species is simply a rank (for details and references, see Dayrat et al. [2008]). De Queiroz (1999) contrasted a clade, which includes all the lines of descent from a given ancestor, with a species, which he defines as a segment of a "separately evolving metapopulation lineage" and proposed separate nomenclatural treatment for each (de Queiroz 2007). The term *lineage* was defined by de Queiroz (2007: 881) and citations therein as "... an ancestor-descendant series in the case of metapopulations or

Table 1. Priorities for listing or reclassifying organisms from threatened to endangered (U.S. Congress 1983).*

Threat			
Magnitude	*Immediacy*	*Taxonomy*	*Priority*
High	Imminent	Monotypic genus	1
		Species	2
		Subspecies	3
	Nonimminent	Monotypic genus	4
		Species	5
		Subspecies	6
Moderate to low	Imminent	Monotypic genus	7
		Species	8
		Subspecies	9
	Nonimminent	Monotypic genus	10
		Species	11
		Subspecies	12

Distinct population segment is prioritized equivalent to or below the subspecies level.

simply a metapopulation extending through time. It is not to be confused with a clade . . . , which is sometimes also called a lineage but is made up of several lineages . . . The term metapopulation refers to an inclusive population made up of connected subpopulations." de Queiroz (2007) used *metapopulation* "to distinguish species, which are typically considered to reside at the higher end of the population-level continuum, from the populations at the lower end, such as demes or family groups." de Queiroz (2007) did not consider "a species . . . an entire metapopulation lineage but only a segment of such a lineage" and said "species give rise to other species, thereby forming (species level) lineages. Any given species is but one of many segments that make up such a species level lineage."

Mishler (1999), on the other hand, saw species as fundamentally clades, which therefore should not be treated different nomenclaturally from other clades. Pleijel (1999) and Pleijel and Rouse (2000*a*, 2000*b*, 2000*c*) provided evidence for the incomparability among organisms ranked as species and introduced the least inclusive taxonomic unit (LITU), the least-inclusive clade to which individual organisms can be referred. This system removes special nomenclatural treatment of species (the group) and "identifies taxa that contain no other taxa." Pleijel and Rouse (2000*b* stated, " . . . LITUs are statements about the current state of knowledge (or lack thereof) without implying that they have no internal structure . . . ." This approach has had some support from taxonomists studying invertebrates (Härlin & Härlin 2001; Poisot et al. 2011), mosses (Fisher 2006), and fungi (Lumbsch 2002), but it has not been widely accepted and is not part of the PhyloCode.

A formal challenge to the Dayrat et al. (2008) approach to naming species that was adopted in the PhyloCode (version 4) came from within the ranks of the ISPN (Cellinese et al. 2012). The authors' main objection was that the PhyloCode adopts a particular viewpoint on the nature of species. The challengers preferred to see the PhyloCode as "rank-agnostic" and focused solely on naming clades, as it was intended. They accepted that clades are different than lineages and acknowledged that some might feel that species are lineages (de Queiroz 1999), but they strongly disagreed with the PhyloCode's imposition of such a view of species on its would-be users. They argued (Cellinese et al. 2012: 886) that the PhyloCode should " . . . treat all clades, from the smallest to the very largest, equally." Because this would "allow systematists freedom to equate species with clades, or with lineages, or to deny the existence of species entirely . . . " They believe "[a] system of nomenclature should not be tied to a particular philosophical outlook on something as controversial as the nature of species, especially a system focused on the naming of clades."

The CPN ultimately rejected most of the specific recommendations of Cellinese et al. (2012), which were designated to eliminate all mention of species from the PhyloCode, but broadened the definition of the term *species* in the code, clarifying that it is " . . . used both for a kind of biological entity (for example a population lineage segment) and for the lowest primary rank in traditional nomenclature (and . . . for any taxon assigned to that rank)" (Cantino & de Queiroz 2014).

The PhyloCode (version 5) (Cantino & de Queiroz 2014) now states explicitly that it does not govern species names or the names of infraspecific taxa (subspecies). Instead, the names of these taxa must satisfy the appropriate rank-based code (e.g., ICZN). Under the rank-based ICZN, species and subspecies are labeled with the corresponding rank and are required to be named with a binomen or trinomen, respectively. It is not possible, therefore, for users of the PhyloCode to name new species or subspecies without providing the rank information as well.

GENERA

Authors who combine clade names governed by the PhyloCode with species names governed by the traditional codes must accept the persistence of the genus name as part of the species name even though the genus (like other ranks) is not accepted in the PhyloCode. The fact that the first word of the species name is viewed as simply part of the name under PN, rather than the name of a taxon ranked as a genus, should have no impact on the use of species names in the implementation of the ESA. However, a legal analysis of this issue as it relates to prioritizing monotypic genera might be warranted.

SUBSPECIES

The PhyloCode requires that the traditional rank-based codes be used for infraspecific naming, but it allows for the naming of "clades that correspond in composition to or are nested within taxa that are ranked as species" (version 5, note 21.1.1). It would, therefore, be possible to name an infraspecific clade using the PhyloCode that might subsequently be viewed by a BRT as worthy of species, subspecies, or DPS-level protection.

For conservation purposes, including the implementation of the ESA, adoption of the PhyloCode means little will change because the rank-free PhyloCode does not govern species names or names associated with the ranks below species.

Although the LITU system is not part of the PhyloCode, potential problems with this approach are worth highlighting. First, neither LITUs nor clades are legally protected. Additionally, rank labels (i.e., species or genus) can be omitted using the LITU framework for taxonomic designation. The LITU approach could also be prone to over-splitting (see Agapow et al. [2004] for similar problems with the PSC). Phylogenetic analyses, especially when applied to genetic data, can provide fine-scale LITUs within a species that actually represent "organism lineages" (de Queiroz 2012). Instability could occur if additional genetic data permit recognition of LITUs within what was previously named as an LITU; thus, the latter is changed from an LITU to a larger clade. Biological review teams can determine the appropriate level of taxonomic rank, but this could be overwhelming if it became a common problem.

Importance of Knowledge of Relationships

The National Marine Fisheries Service was petitioned to list a group of fish-eating killer whales (*Orcinus orca*) from the southern Puget Sound, Washington (U.S.A.), in 2001. There were <100 animals, and based on ecology, acoustics, and movement patterns it appeared that this was a distinct and closed population. A BRT was assembled and quickly concurred that the population was a threatened DPS. However, the panel struggled to determine if this population was an "important component of the greater species' evolutionary legacy." This was because of a lack of taxonomy for killer whales due in part to their relatively low genetic diversity at the control region of the mitochondrial DNA (Hoelzel et al. 2002).

Orcinus orca is one of the most widely distributed vertebrate species on the planet, but exhibits extensive ecological variation. A mammal-eating form is sympatric with, but never interacts with, the fish-eating form in the northern Pacific (Bigg 1982). This made a pretty good case that the 2 forms might be different species and the small population of fish-eaters in the southern Puget Sound might be evolutionarily significant. Information from Antarctica suggested more forms of killer whales that differ in ecology and morphology from those in the northern Pacific (Pitman & Ensor 2003). The BRT was left wondering if "this population [was] part of a unique undescribed species and therefore a significant evolutionary importance or was it just a small portion of a single large globally-distributed species?" (Taylor 2005). Finally, the BRT decided to consider this population significant based on other genetic information. They reconvened after cetacean taxonomists had clarified and revised the evidence needed to designate species and subspecies given the improbability of obtaining morphological data from dead animals in a timely manner (Reeves 2004). Most of the BRT agreed that more than one species or subspecies was likely under the new definitions and that northern Pacific fish-eaters and mammal-eaters probably belong to separate subspecies (Krahn et al. 2004). We now know that this DPS is part of a clade that split from the sympatric mammal-eating clade 700,000 years ago (Morin et al. 2010). Nomenclatural consideration is still forthcoming, but it appears that the mammal-eating types will be considered at least different subspecies (Morin et al. 2010).

Taylor (2005) cuts to the root of the issue: "taxonomy is evidentiary, not precautionary." In addition, there is no incorporation of uncertainty in taxonomy—it is absolute—which usually means, until you can prove that it is unique, a taxon does not exist. This has translated into delayed action on species protection. This example illustrates the practical need for taxonomic resolution for conservation efforts through the ESA, but the knowledge of relationships is more important than names or nomenclatural procedures for initial protection of endangered taxa.

Leveraging Modern Systematics and Stabilizing Nomenclature for Conservation

Traditional rank-based nomenclature perpetuates nomenclatural instability and incorrect assumptions about taxonomic comparability (Mishler 2010). However, the suggestion that conservation "drop its reliance on species" (Mishler 2010) is unrealistic given the dependence of conservation legislation on these units.

Effect of the PhyloCode on the ESA

There is little chance of the term *species* being abandoned. The PhyloCode states that traditional rank-based rules must be used when designating species and infraspecific taxa. This means that the rank of species and genus will still be applied to new taxa. For example,

an author could name a new taxon that is both a clade (under the PhyloCode) and a genus (under the IZCN) by following the rules of 2 codes in one publication. The same name would be used in both places.

However, by employing the LITU framework outside of the PhyloCode, rank labels (i.e., species or genus) can be omitted during taxonomic designation. Neither LITUs nor clades are legally protected. Furthermore, instability may occur if additional molecular data change a named taxon from a LITU to larger clade.

Systematics and ESA Listings

Nomenclature is vital to taxonomy and conservation, but it is not critical to initial ESA listings. As the killer whale example shows, knowledge of evolutionary relationships is more important to conservation through species-based laws, and nomenclatural treatment can follow these studies. Authors cannot omit rank designations for new species names under the PhyloCode. However, if they choose to omit ranks by using LITUs, the systematic relationships will be evaluated by the BRT as with any other manuscript. Omitting rank labels could complicate the efforts of the BRT (some species-based legislation may not have expert panels such as the BRT), but systematics papers are frequently published without formal nomenclatural treatment, especially for organisms for which little is known or corroborative evidence is scarce (see Morin et al. 2010; Archer et al. 2013; Kershaw et al. 2013 as cetacean examples). Authors of such papers should provide details (i.e., estimates of gene flow and patterns of population structure, etc.) that would aid provisional taxonomic designation for the listing process. Interestingly, the PhyloCode allows for the naming of "clades that correspond in composition to or are nested within taxa that are ranked as species" (version 5, note 21.1.1). This could be used to provisionally name infraspecific clades that might subsequently be viewed as worthy of species, subspecies, or DPS-level protection in the BRT.

Phylogenetic Prioritization

Conservationists should also consider modernizing the prioritization process by incorporating taxonomic uncertainty (Taylor 2005) and measures of evolutionary uniqueness (i.e., phylogenetic branch lengths or number of unique changes from the nearest relative) (e.g., May-Collado & Agnarsson 2011). The latter is important for identifying special genetic reservoirs because many of these will likely not be recognized taxonomically as monotypic genera and thus not receive the highest taxonomic priority in the ESA (Table 1). Several frameworks for applying phylogenetic analyses to conservation strategies are available (e.g., Vane-Wright et al. 1991; Faith 1992; Mace et al. 2003; Rolland et al. 2012) with certain limitations (Rodrigues et al. 2011).

Acknowledgments

G. Rouse provided thoughtful guidance. R. Kelly, S. Fallon, K. de Queiroz, R. Waples, W. Perrin, and anonymous reviewers provided comments on early drafts. N. Hausman, S. Leslie, P. Doukakis provided useful discussion. P. Cantino provided a pre-publication draft of the PhyloCode version 5. M.S.L. was supported by NSF GRF and NSF IGERT fellowships.

Literature Cited

Agapow, P. M., O. R. P. Bininda-Emonds, K. A. Crandall, J. L. Gittleman, G. M. Mace, J. C. Marshall, and A. Purvis. 2004. The impact of species concept on biodiversity studies. The Quarterly Review of Biology **79:**161–179.

Archer, F. I., P. A. Morin, B. L. Hancock-Hanser, K. M. Robertson, M. S. Leslie, M. Bérubé, S. Panigada, and B. L. Taylor. 2013. Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus spp.*): genetic evidence for revision of subspecies. PLoS ONE **8** DOI: 10.1371/journal.pone.0063396.

Bigg, M. 1982. An assessment of killer whale (*Orcinus orca*) stocks off Vancouver Island, British Columbia. Report of the International Whaling Commission **32:**5–666.

Cantino, D. P., H. N. Bryant, K. De Queiroz, M. J. Donoghue, T. Eriksson, D. M. Hillis, and M. S. Y. Lee. 1999. Species names in phylogenetic nomenclature. Systematic Biology **48:**790–807.

Cantino, P. D., and K. de Queiroz. 2010. International Code of Phylogenetic Nomenclature. Version 4c. Available from http://www.phylocode.org (accessed March 2014).

Cantino, P. D., and K. de Queiroz. 2014. International Code of Phylogenetic Nomenclature (the PhyloCode). Version 5 beta., Available from the authors pending completion of the preface and index. Soon to be available at http://www.phylocode.org.

Cellinese, N., D. A. Baum, and B. D. Mishler. 2012. Species and phylogenetic nomenclature. Systematic Biology **61:**885–891.

Cracraft, J. 1983. Species concepts and speciation analysis. Current Ornithology **1:**159–187.

Cracraft, J. 1989. Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Pages 28–59 in D. Otte, and J. Endler, editors. Speciation and Its Consequences. Sinauer Associates, Sunderland, Massachusetts.

Dayrat, B., P. D. Cantino, J. A. Clarke, and K. de Queiroz. 2008. Species names in the PhyloCode: the approach adopted by the International Society for Phylogenetic Nomenclature. Systematic Biology **57:**507–514.

de Queiroz, K. 1999. The general lineage concept of species and the defining properties of the species category. Pages 57–75 in D. Howard, and S. Berlochers, editors. Species: New Interdisciplinary Essays. Oxford University Press, New York, Oxford.

de Queiroz, K. 2005. A unified concept of species and its consequences for the future of taxonomy. Proceedings of the California Accademy of Sciences **5:**196–215.

de Queiroz, K. 2006. The PhyloCode and the distinction between taxonomy and nomenclature. Systematic Biology **55:**160–162.

de Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology **56:**879–886.

de Queiroz, K., and J. Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. Trends in Ecology & Evolution **9:**27–31.

de Queiroz, K. 2012. Biological nomenclature from linnaeus to the phylocode. Bibliotheca Herpetologica **9**(1–2): 135–145.

Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation **61:**1–10.

Fallon, S. M. 2007. Genetic data and the listing of species under the US endangered species act. Conservation Biology **21:**1186–1195.

Fisher, K. M. 2006. Rank-free monography: a practical example from the moss clade Leucophanella (Calymperaceae). Systematic Botany **31:**13–30.

Frankham, R., J. D. Ballou, M. R. Dudash, M. D. B. Eldridge, C. B. Fenster, R. C. Lacy, J. R. Mendelson, I. J. Porton, K. Ralls, and O. A. Ryder. 2012. Implications of different species concepts for conserving biodiversity. Biological Conservation **153:**25–31.

Haig, S. M., et al. 2006. Taxonomic considerations in listing subspecies under the US Endangered Species Act. Conservation Biology **20:**1584–1594.

Härlin, M., and C. Härlin. 2001. Phylogeny of the eureptantic nemerteans revisited. Zoologica Scripta **30:**49–58.

Hoelzel, A. R., A. Natoli, M. E. Dahlheim, C. Olavarria, R. W. Baird, and N. A. Black. 2002. Low worldwide genetic diversity in the killer whale (*Orcinus orca*): implications for demographic history. Proceedings of the Royal Society of London. Series B: Biological Sciences **269:**1467–1473.

Jörger, K. M., and M. Schrödl. 2013. How to describe a cryptic species? Practical challenges of molecular taxonomy. Frontiers in Zoology **10:**59. DOI:10.1186/1742-9994-10-59.

Kelly, R. P. 2010. The use of population genetics in Endangered Species Act listing decisions. Ecology Law Quarterly **37:**1107–1158.

Kershaw, F., et al. 2013. Population differentiation of 2 forms of Bryde's whales in the Indian and Pacific Oceans. Journal of Heredity **104**(6)**:**755–764.

Krahn, M. M., et al. 2004. Status review of southern resident killer whales (*Orcinus orca*) under the Endangered Species Act. US Department of Commerce, National Oceanic and Atmospheric Administration, Technical Memorandum, NMFS-NWFSC-62.

Linnaeus, C. 1753. Species Plantarum. 1st Edition. Stockholm.

Lumbsch, H. T. 2002. How objective are genera in euascomycetes? Perspectives in Plant Ecology, Evolution and Systematics **5:**91–101.

Mace, G. M. 2004. The role of taxonomy in species conservation. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **359:**711–719.

Mace, G. M., J. L. Gittleman, and A. Purvis. 2003. Preserving the tree of life. Science **300:**1707–1709.

May-Collado, L. J., and I. Agnarsson. 2011. Phylogenetic analysis of conservation priorities for aquatic mammals and their terrestrial relatives, with a comparison of methods. PLoS ONE **6**(7) DOI: 10.1371/journal.pone.0022562

Mishler, B. D. 1999. Getting rid of species. Pages 307–315 in D. Howard, and S. Berlochers, editors. Species: New Interdisciplinary Essays. Sinauer Associates, Sunderland, Massachusetts.

Mishler, B. D. 2010. Species are not uniquely real biological entities pages 110–122 in F. Ayala, and R. Arp, editors. Contemporary Debates in Philosophy of Biology. John Wiley & Sons. West Sussex, United Kingdom.

Montagu, G. 1821. Description of a species of *Delphinus* which appears to be new. Memoirs of the Wernerian Natural History Society **3:**75–82.

Morin, P. A., F. I. Archer, A. D. Foote, J. Vilstrup, E. E. Allen, P. Wade, J. Durban, K. Parsons, R. Pitman, and L. Li. 2010. Complete mitochondrial genome phylogeographic analysis of killer whales (Orcinus orca) indicates multiple species. Genome Research **20:**908–916.

Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics **6:**211–223.

Pitman, R. L., and P. Ensor. 2003. Three forms of killer whales (*Orcinus orca*) in Antarctic waters. Journal of Cetacean Research and Management **5:**131–140.

Pleijel, F. 1999. Phylogenetic taxonomy, a farewell to species, and a revision of Heteropodarke (Hesionidae, Polychaeta, Annelida). Systematic Biology **48:**755–789.

Pleijel, F., and G. W. Rouse. 2000*a*. A new taxon, capricornia (Hesionidae, Polychaeta), illustrating the LITU ('Least-Inclusive Taxonomic Unit') concept. Zoologica Scripta **29:**157–168.

Pleijel, F., and G. W. Rouse. 2000*b*. Least-inclusive taxonomic unit: a new taxonomic concept for biology. Proceedings of the Royal Society of London, Series B: Biological Sciences **267:**627–630.

Pleijel, F., and G. W. Rouse. 2003. Ceci n'est pas une pipe: names, clades and phylogenetic nomenclature. Journal of Zoological Systematics and Evolutionary Research **41:**162–174.

Poisot, T., O. Verneau, and Y. Desdevises. 2011. Morphological and molecular evolution are not linked in Lamellodiscus (Plathyhelminthes, Monogenea). PLoS ONE **6**. DOI: 10.1371/journal.pone.0026252.

Reeves, R. R. 2004. Report of the Workshop on Shortcomings of Cetacean Taxonomy in Relation to Needs of Conservation and Management, April 30-May 2, 2004, La Jolla, California. US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southwest Fisheries Science Center.

Rodrigues, A. S. L., R. Grenyer, J. E. M. Baillie, O. R. P. Bininda-Emonds, J. L. Gittlemann, M. Hoffmann, K. Safi, J. Schipper, S. N. Stuart, and T. Brooks. 2011. Complete, accurate, mammalian phylogenies aid conservation planning, but not much. Philosophical Transactions of the Royal Society B: Biological Sciences **366:**2652–2660.

Rolland, J., et al. 2012. Using phylogenies in conservation: new perspectives. Biology Letters **8:**692–694.

Taylor, B. L. 2005. Identifying units to conserve. Marine mammal research: Conservation beyond crisis. The John Hopkins University Press, Baltimore, MD:149–164.

Tobias, J. A., N. Seddon, C. N. Spottiswoode, J. D. Pilgrim, L. D. C. Fishpool, and N. J. Collar. 2010. Quantitative criteria for species delimitation. Ibis **152:**724–746.

U.S. Congress. 1983. Endangered Species Act. Federal Register: 43098-43105.

Vane-Wright, R. I., C. J. Humphries, and P. H. Williams. 1991. What to protect? Systematics and the agony of choice. Biological Conservation **55:**235–254.

Waples, R. S. 1991. Definition of "species" under the Endangered Species Act: application to Pacific Salmon. National Oceanic and Atmospheric Administration, Technical Memorandum NMFS-NWFSC-FMC194, Seattle, Washington.

Wheeler, Q. D., P. H. Raven, and E. O. Wilson. 2004. Taxonomy: Impediment or expedient? Science **303:**285.

Wiley, E. O., D. Siegel-Causey, D. R. Brooks, and V. A. Funk. 1991. The compleat cladist: a primer of phylogenetic procedures. Museum of Natural History, University of Kansas Lawrence.

Yang, Z., and B. Rannala. 2012. Molecular phylogenetics: principles and practice. Nature Reviews Genetics **13:**303–314.

Chapter 1 is published in *Conservation Biology*. The dissertation author was the sole investigator and sole author of this material.

ACKNOWLEDGEMENTS:

Along with those people and agencies who provided inspiration, guidance, feedback or support, special appreciation is given to the publishers of *Conservation Biology*, John Wiley & Sons Inc., for permission to include this material as a chapter of my dissertation.

CHAPTER 2:

Genetic inference of population structure in spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) from the eastern tropical Pacific Ocean based on mitochondrial genomes and nuclear SNPs

Matthew S. Leslie [1,2]

1. Scripps Institution of Oceanography, University of California San Diego, mail-code 0208, 9500 Gilman Dr. La Jolla, CA, 92093 USA

Frederick I. Archer [2,1]

Phillip A. Morin[2,1]

2. Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 8901 La Jolla Shores Dr., La Jolla, CA 92037, USA

Abstract:

Millions of spinner (*Stenella longirostris*) and spotted dolphins (*Stenella attenuata*) died as bycatch in tuna nets in the eastern tropical Pacific Ocean. Despite three decades of protection, they show little-to-no sign of recovery. In efforts to establish biologically meaningful management boundaries for recovery, endemic subspecies and multiple stocks have been proposed. Genetic differentiation among most of these units has not been demonstrated, however, possibly due to low statistical power, large historical abundances, ongoing gene flow between subspecies and recent divergence. We tested for structure at multiple hierarchical levels by collecting whole mitochondrial genome sequences (mtDNA) and nuclear SNPs (nuDNA) from 104 spinner and 76 spotted dolphins using capture array library enrichment and highly paralleled DNA sequencing. MtDNA showed weak but significant differences between subspecies of spotted dolphins (F_{ST}: 0.0125; p = 0.0402) and evidence for differentiation between ETP spinner dolphin subspecies (F_{ST}: 0.0133; p = 0.034). NuDNA supported subspecies of spotted dolphins, but not subspecies of spinner dolphins. Strong and significant differentiation was detected between whitebelly and eastern spinner stocks using nuDNA (F_{ST}: 0.0297; p = 0.0059). Neither mtDNA nor nuDNA supported the division of existing offshore stocks of spotted dolphins or Tres Marias spinner dolphins. This work helps identify a genetic basis for establishing biologically meaningful management units for the recovery of these dolphins.

Introduction:

Conservation is only effective if actions are applied at the appropriate scale. Without knowledge of how organisms are geographically structured we may fail to recognize entire subspecies or distinct populations that require urgent conservation action. Knowledge of population structure forms the foundation of assessment and monitoring for recovery; we need to know population boundaries in order to estimate abundance and trends and to set appropriate mortality limits.

The study of population structure in pelagic organisms is inherently challenging. Their open ocean habitats have few permanent physical barriers to gene flow, thus demographic processes (e.g., range expansion and metapopulation dynamics; Horne 2014) or oceanographic processes may be more important in structuring populations than physical separation (Norris 2000). In genetic terms, the dynamic nature of this environment increases the likelihood of mixing between populations; even a few successful migrants per generation greatly decrease the signal of population segregation (Waples 1998).

Pelagic populations often have large abundances as well, which can result in a large amount of standing genetic variation (Norris 2000). High genetic diversity can dramatically increase the time needed for populations to drift apart genetically, even in the complete absence of gene flow (Taylor and Dizon 1996).

Pelagic species with large abundance and weak barriers provide a challenge for using genetic tools to determine population structure (Taylor and Dizon 1996; Waples 1998). Spinner and spotted dolphins in the eastern tropical Pacific Ocean (ETP) illustrate this problem. They historically numbered several million (Wade *et al.* 2007), but starting in the 1960s, hundreds of thousands were killed annually as bycatch in the dolphin-set tuna purse-seine fishery (Lo and Smith 1986; National Research Council 1992; Wade 1995). Despite over forty years of protection under the U.S. Marine Mammal Protection Act of 1972 and specific multi-national protection under the 1999 Agreement on the International Dolphin Conservation Program (Joseph 1994; Gosliner 1999), ETP spinner and spotted dolphin populations have not recovered as predicted (Wade *et al.* 2007; Gerrodette *et al.* 2008). Previous genetic studies have shown that these species retain high genetic variation (Dizon *et al.* 1994; Galver 2002; Andrews *et al.* 2013; Escorza-Treviño *et al.* 2005), which possibly obscures the detection of population genetic structure that could be used to improve recovery efforts. These species exhibit intraspecific morphometric differences supporting structured populations (Perrin *et al.* 1991, 1994), but traditional molecular genetics approaches have not found corresponding population genetic structure (Dizon *et al.* 1994; Galver 2002), although Andrews *et al.* (2013) found evidence for segregation in ETP spinner dolphins using data from the Y-chromosome and Escorza-Treviño *et al.* (2005) found population structure within the coastal subspecies of spotted dolphins using microsatellites.

Endemic spinner dolphin subspecies and stocks

There are four recognized subspecies of spinner dolphin (*Stenella longirostris*) globally. The nominate form, the pantropical or "Gray's" spinner (*S. l. longirostris*) occurs throughout the world's tropics except in the ETP. In the central and western Pacific, Gray's spinners are found associated with islands, such as the Hawaiian Islands. The much smaller dwarf spinner (*S. l. rosiventris*) subspecies inhabits shallow waters in the Southeast Asia Ocean (Perrin *et al.* 1989, 1999). The other two subspecies, Central American spinner dolphin (*S. l. centroamericana*) and the eastern spinner dolphin (*Stenella longirostris orientalis*) are found only in the ETP (Fig. 1, based on Perrin 1985), and were described based on differences in external coloration, body size, and skull morphology (Perrin *et al.* 1991, Douglas *et al.* 1992). The Central American subspecies inhabits relatively near-shore waters off the Pacific coasts of Southern Mexico south through Panama. For management purposes, ETP spinner dolphin stocks correspond to the two aforementioned subspecies, plus the whitebelly spinners. The "whitebelly" spinner is proposed to represent a hybrid swarm between the eastern subspecies and the pantropical Gray's subspecies of the central and western Pacific (Perrin *et al.* 1991). Taxonomically, it is classified as part of the nominate (Gray's) subspecies *S. l. longirostris*. Significant geographic overlap exists between the eastern subspecies and the whitebelly form (Perrin *et al.* 1985). The eastern spinner dolphin (*S. l. orientalis*) exhibits traits indicative of a more polygynous mating system (Perrin and Mesnick 2003; Perrin and Henderson 1979). Perrin and Mesnick (2003) found significant difference in testes

size between the eastern subspecies and the whitebelly form, indicating differing reproductive strategies and probably different breeding behavior between the two types. Andrews *et al.* (2013) estimated high levels of geneflow between subspecies in the ETP using autosomal and mitochondrial genes and found a shared Y chromosome haplotype in the eastern and Central American subspecies that was not found in Gray's or dwarf subspecies. Interestingly, this locus was found to be polymorphic in whitebellies, supporting the hypothesis of introgression in this form (Andrews *et al.* 2013). The authors proposed that sexual selection was driving the divergence of spinner dolphins in the ETP. Finally, a distinct morphotype of the eastern spinner dolphin, known as the "Tres Marias" spinner dolphin, has been described from near the islands of the same name off the coast of Mexico. These were thought to be a distinct type based on external body morphometrics (Perryman and Westlake 1998).

Endemic ETP spotted dolphin subspecies and stocks

Extensive analyses of cranial morphology of pantropical spotted dolphins (*Stenella attenuata*) in the ETP led to the designation of a coastal endemic subspecies (*S. a. graffmani* - Perrin 1975, Perrin *et al.* 1987). Pantropical spotted dolphins in offshore regions in the ETP and elsewhere retain the nominate trinomial, *S. a. attenuata*. In contrast to the spinner dolphins, genetic analyses of microsatellites indicated some differentiation between subspecies (Escorza-Treviño *et al.* 2005). This study found

differences between at least four demographically independent populations within the coastal subspecies (*S. a. graffmani*) and differences between southern populations of the coastal subspecies and the pelagic subspecies. However, they found no differences between the northern populations of the coastal subspecies and the pelagic subspecies. The authors concluded that genetic diversity in the coastal subspecies is contained within demographically independent populations and that interchanges is ongoing between northern populations and the pelagic subspecies. Tests for population genetic structure within the pelagic subspecies have not been conducted.

Despite the results of Escorza-Treviño *et al.* (2005) showing the entire subspecies is currently treated as a single management stock. The offshore pantropical spotted dolphins are divided into a 'northeastern' (NE) stock, defined geographically as north of 5°N, east of 120°W and a 'western-southern' (WS) stock residing south and west of this northeastern area (Fig. 2) (Perrin *et al.* 1994). General similarity in cranial variation supports the WS stock (see Perrin *et al.* 1994 for discussion and primary references). The north-south boundary between these stocks is based on a distributional hiatus around 5°N.

Objectives

Despite cranial and external morphometric differences between subspecies (Perrin *et al.* 1994; Douglas *et al.* 1992), methods of assessing population genetic structure have

generally failed to detect differences (Dizon *et al.* 1994; Galver 2002), though Andrews *et al.* (2013) and Escorza-Treviño *et al.* (2005) found some. In this study we apply population genetic structure analyses of whole mitochondrial genomes (mitogenomes), individual gene regions within the mitogenomes and nuclear single nucleotide polymorphisms (SNPs) to test subspecies hypotheses and levels of differentiation between recognized and proposed management stocks, including the Tres Marias spinner dolphin and alternative stock boundaries in the offshore spotted dolphins. Finally, using the mitogenome data, we examine whether sea surface temperature (SST) might be related to population structure within offshore spotted dolphins in the ETP.

METHODS

Sample Collection and DNA extraction

Skin samples were collected from 104 spinner dolphins and 76 spotted dolphins via biopsy dart (Lambertsen 1987) on research cruises or from specimens taken as bycatch in the tuna purse-seine fishery between 1982 and 2010 (Figs. 1, 2, S1, S2). Spinner dolphin samples collected from on research cruises were assigned to a stock based on the external morphology of the majority of animals in the school despite geographic location. This approach was taken because: 1) these often-large groups (>1000 ind.) contained individuals exhibiting a range of morphology; only after observing the group for some time could observers classify it to stock, 2) researchers

collecting biopsies from dolphins near the bow of the research vessel found it very difficult to confidently classify fast-swimming individuals at sea, and 3) there is significant overlap in range; therefore, geography was not a reliable predictor of stock identity. Some samples were used from areas where the eastern and whitebelly spinners are known to geographically overlap (see Fig. 1). Hawaiian spotted dolphin samples were collected from the Kona Coast of Hawai‘i and O‘ahu. Spinner dolphin samples from Hawai‘i spanned the breadth of the main islands and also Midway Atoll. Spotted dolphins were assigned to subspecies and stocks based on the geographic location of the sampling site. To avoid misassigned individuals near the borders of the NE and WS offshore stocks, we did not use samples collected between 4°N and 6°N east of 125°W.

Biopsy samples were stored in salt-saturated 20% DMSO, 70% ethanol, or frozen with no preservative. DNA was extracted using silica-based filter membranes (Qiagen, Valencia, CA) on an automated workstation (Perkin Elmer, Waltham, MA). Starting concentrations of DNA were quantified using Pico-Green fluorescence assays (Quant-it Kit, Invitrogen, Carlsbad, CA) using a Tecan Genios microplate reader (Tecan Group Ltd, Switzerland).

Library Preparation and Sequencing

Next-generation sequencing libraries were generated as described by Hancock-Hanser *et al.* (2013), using unique 6bp and 7bp index sequences for each individual to

allow up to 100 samples to be multiplexed. Multiplexed libraries were enriched for whole mitogenomes and 85 nuclear loci using Sure Select DNA Capture Arrays (Agilent Technologies, Inc., Santa Clara, CA, USA) as described by Hancock-Hanser *et al.* (2013). Target sequences for capture enrichment included the reference pantropical spotted dolphin mitochondrial genome (Genbank No. EU557096; Xiong *et al.* 2009) and a suite of 85 nuclear loci. Autosomal and sex-linked conserved anchor tag loci (CATS; Lyons *et al.* 1997: Aitken *et al.* 2004) made up the bulk of the nuclear loci (75). In addition, one Y-chromosome locus (SRY Genbank No. AB275398.1) and nine nuclear exons involved in vertebrate coloration were also included (Rieder *et al.* 2001; Hoekstra et al. 2006; Eriksson et al. 2008; Hubbard et al. 2010). Sequences from four cetacean species were aligned to create consensus sequence for capture array design for the 75 CATS loci (see Table A1-S1 in Hancock-Hanser *et al.* 2013 for sequences). Coloration loci and the Y-chromosome locus were aligned to the common bottlenose dolphin (*Tursiops truncatus*) genome (ENSEMBL v. 61); bottlenose dolphin sequences (Figure A1-S1) were used in the capture array design with the eArray software package (Agilent Technologies, Inc., Santa Clara, CA, USA). Three identical arrays were used to capture a multiplexed mix of both species. Each array contained one replicate of the mitogenome probes at a probe interval of 15bp as well as 13 replicates of probes for the nuclear loci at a probe interval of 3bp. Each enriched library was then sequenced using 1X100bp Illumina HiSeq technology (two using Illumina HiSeq2000 and one using HiSeq2500).

Mitogenome Assembly

Raw read data were filtered for quality and demultiplexed by unique barcode. Consensus sequences for each sample were generated from mitogenome sequence reads using a custom pipeline (Dryad data repository doi:10.5061/dryad.cv35b) in R v2.15.0 (R Core Team, 2014). Reads were first mapped to the reference spotted dolphin sequence with the short-read alignment tool BWA (Li and Durbin, 2009). The mpileup module in SAMTOOLS (Li *et al.* 2009) was then used to convert the resulting BAM-format alignment file into a ‘‘pileup’’ text format, which was then parsed by custom R code to create the consensus sequence for each individual. The following rules were used in this process: An “N” was inserted at a position if the assembly had <3 reads, <5 reads where not all contained the same nucleotide, or >5 reads where no one nucleotide (i.e., A, C, G, T) was present in >70% of the reads. All mitogenome sequences were initially aligned with MAFFT (Katoh *et al.* 2009) followed by a refinement of alignments by eye.

NuDNA SNP Discovery and Genotyping

Using CLC Genomics Workbook v3.7 (CLC Bio, Aarhus Denmark) we created species-specific *de novo* references for nuDNA loci using combined reads from a subset of five individuals with good coverage from each species. All sequence data were first mapped to the mitogenome references and unmapped reads were saved and used for *de novo* assembly. *De novo* contigs were aligned back to original capture references for verification prior to mapping reads to the *de novo* reference. For loci that did not assemble into *de novo* species-specific references, we used the original reference sequence used in the capture array assembly. Additional flanking sequence was often

added with the *de novo* references relative to the original capture sequences.

All reads were assembled to the *de novo* nuDNA references with the same criteria outlined above for the mtDNA. Putative SNPs were discovered using MPILEUP and variant calling tools from GATK (McKenna *et al.* 2009). We considered variable sites in nuclear loci to be potential SNPs if the population samples contained heterozygotes and/or homozygotes of the minor allele with at least 7 reads for called genotypes, and the minor allele frequency was ≥5% frequency in the sample set. Putative SNPs were screened for validation by examining BAM alignments by eye. We did not choose a SNP if it had *N*s or other ambiguities in the flanking regions or if the SNP was found in only a few individuals.

Diversity Estimates and Population Structure Analyses

Two mitogenome datasets were created for each species. First, we partitioned each species' dataset into one of fifteen loci (12 coding sequences, the control region and 2 rRNA genes). ND6 and tRNA loci were removed prior to analyses because ND6 is known to have five orders of magnitude more transitions than transversions (Duchene *et al.* 2011) and tRNA follows different evolutionary models. Sequences were aligned to the pantropical spotted dolphin reference and locus start/stop positions were annotated in GENEIOUS v5.4 (created by Biomatters and available from http://www.geneious.com/) using the GENEIOUS alignment tool and the amino acid translation tool, respectively.

Second, we removed the control region because of high variation in this region and concatenated the remaining 14 regions to make the mitogenome sequences. The final sequence lengths for the concatenated data were 13,426bp and 13,425bp for spinner and spotted dolphins, respectively. Data for individual animals were removed from analyses if they contained >10% missing data over the entire mitogenome sequence.

For both datasets, we estimated haplotypic diversity (H_d, Nei 1987) and nucleotide diversity (π, Tajima 1983) using tools from the *strataG* package[2] in R v2.3.1 (R Core Team 2014). Individual genes and whole mitochondrial genome sequences were assigned to unique haplotypes using this same package. Three pairwise estimates of population genetic structure, F_{ST} (Wright 1949), Φ_{ST} (Weir and Cockerham 1984), and χ^2 test for heterogeneity of allele frequencies were performed using the *strataG* package[1] in R v2.3.1 (R Core Team 2014). The significance of each estimate was tested using 5000 non-parametric random permutations of the data matrix variables. For Φ_{ST}, pairwise distances were calculated using the best substitution model as identified by Akaike's Information Criterion in JModelTest version 2.1.4 (Posada 2008).

[2] Archer, E. 2015. *strataG*: Summaries and Population Structure Analyses of Haplotypic and Genotypic Data. R package version 0.9.3.2.

In addition to testing geographically-based hypotheses of population structure, we also tested if the offshore spotted dolphin stock boundary was associated with an important environmental variable, sea surface temperature (SST). SST for each sample was estimated from the ECMWF Ocean Reanalysis System 4 (ORAS4) based on the time and location of the sample collection event (Balmaseda *et al.* 2013). We then created four temperature boundaries chosen to divide the range of temperatures-at-biopsy-sampling evenly while also giving >6 samples per partition near the highest and lowest temperature boundaries (26.44°C, 27.65°C, 28.09°C, 28.87°C). For each boundary, all offshore spotted dolphins samples wth >90% of the mitochondrial genome were partitioned into two groups and F_{ST} and Φ_{ST} tests of pairwise genetic differentiation were conducted as outlined above. A total of four tests were performed to evaluate the four different temperature boundaries.

We performed a substitution rate test on the mitogenome data set to determine if mutations had reached a point of saturation. For this test, we generated pairwise percent differentiation and plotted this against a Jukes and Cantor (1969) correction factor generated using MEGA 5.2.2 (Tamura *et al.* 2011).

Andrews et al (2013) found that the Y chromosome of ETP spinner dolphin is subject to positive selection. Although mitochondrial loci are assumed to be under purifying selection (Stewart *et al.* 2008) and not associated with the Y chromosome, we

none-the-less tested ETP spinner dolphin mitochondrial genes found to have significant differentiation for evidence of positive selection using both Tajima's *D* and Codon-based Z-Test as implemented in MEGA 5.2.2 (Tamura *et al.* 2011).

Nuclear Population Structure Analyses

Individuals with >50% of genotypes present and loci with >30% of genotypes present in the total dataset were analyzed for population structure. Only the first SNP was chosen from each region sequenced to minimize issues with linkage and phasing. We then calculated pairwise F_{ST}, G_{ST} and Jost's *D* differentiation statistics (Wright 1949; Nei 1973; Jost 2008) using the *strataG* package[1] in R v2.3.1 (R Core Team 2014).

To test if SNPs associated with Y-chromosome or coloration genes contributed disproportionately to the overall patterns of population structure found using our spinner dolphin SNP dataset, we conducted simulation-based F_{ST} outlier tests for selection at each locus in the selection-detection workbench *Lositan* (Antao *et al.* 2008). Using all SNPs and partitioning the samples into three putative populations (whitebelly, eastern and Central American) we used *Lositan* to run 75,000 FDIST2 simulations (Beaumont and Nichols 1996), first calculating neutral F_{ST} (i.e., removing all putative F_{ST} outliers) and then approximating a mean overall F_{ST} using bisection approximation algorithms (Antao *et al.* 2008). We chose an infinite alleles model for all simulations. All population differentiation statistics were calculated again with F_{ST} outliers removed.

Bayesian Clustering

A Bayesian clustering method implemented in STRUCTURE v 2.3.4 (Pritchard *et al.* 2000; Hubisz *et al.* 2009) was used to identify the number of populations (*K*) represented in the SNP datasets. Prior information on the origin of the samples (subspecies and geographically defined groups within subspecies) was combined with a correlated allele frequency model and an admixture model for these analyses. Data were also analyzed withouth location priors using the same models. We evaluated values of *K* between 1 and 5. For each assumed value of *K*, 20 independent runs were conducted. Total length of the run was set at 1,000,000 and burn-in was set at 100,000. The most likely estimate of *K* was determined by the maximum estimated mean log-likelihood of the data (*ln*P(D)) (Pritchard *et al.* 2000) and by calculating Δ*K*, the second-order rate of change of *ln*P(D) with respect to the *K* (Evanno *et al.* 2005).

RESULTS:

Spinner dolphins

Mitogenomes

We assembled 104 complete or nearly complete (<10% missing data) spinner dolphin mitogenomes. Subspecies and regional sample sizes, summary statistics and genetic diversity measures are listed in Table 1. In general, haplotypic diversities were high and nucleotide diversity was low (>0.9722, <0.0078, respectively). The substitution rate test did not show any signs of saturation. The best nucleotide substitution model estimated by JModelTest (Posada, 2008) was TrN (Tamura and Nei, 1993). The results of F_{ST}, Φ_{ST} and χ^2 analyses of the mtDNA concatenated genes and Φ_{ST} of the individual gene regions for spinner dolphins are shown in Table 2.

F_{ST} and Φ_{ST} provide slightly different perspectives on population differentiation and we feel it is important to present both measures. Our results show inconsistancies between these two metrics, which doesn't necessarily mean problems or inaccuracies, but actually something interesting about our data. F_{ST} tests for population differentiation are based on allele (or haploytpe) frequencies and do not provide direct insights into levels of molecular divergence ((Excoffier, Smouse et al. 1992)). In cases where haplotypes are similar within population and different between populations (such as those that would result via drift in small populations), F_{ST} is good at detecting frequency differences that indicate genetic structure. However, when haplotype diversity is high within and among populations, F_{ST} may not have sufficient power to detect differences. Moreover, in this situation, sampling effects can become important drivers of the statistic beyond the base frequency of alleles present and result in false positive results. Φ_{ST} is certainly not immune to these issues, but Φ_{ST} estimates capture more information regarding the

differentiation due to sequence divergence (or nucleotide diversity) in addition to differences in haplotype frequencies. Although we chose to focus the bulk of the discussion on Φ_{ST}, we do report statistically significant measures of F_{ST} and χ^2 and briefly compare and contrast the two metrics. One down side of focusing on Φ_{ST} (and another reason it is important to report F_{ST} as well) is that Φ_{ST} may be more indicative of older, long-term processes, whereas F_{ST} can show recent differences among populations. In addition, given that the test for significance is determined by an arbitrary cut-off ($p = 0.05$), we also present results that are "nearly significant". Given the difficultly of distinguishing these groups in previous works, we felt it important not to focus too intensely on the arbitrary cut-off, but rather overall pattern of indicators. Finally, due to space limitations, we only discuss Φ_{ST} for the partitioned gene region analyses.

At the subspecies level, the Φ_{ST} test showed no differentiation between Central American spinners and eastern spinner dolphin subspecies in either the concatenated or partitioned datasets [F_{ST} (0.0133) was significant ($p = 0.034$) in the concatenated dataset]. Φ_{ST} comparisons of the whitebelly and coastal Central American subspecies showed significant differentiation in the concatenated dataset ($\Phi_{ST} = 0.0491$; $p = 0.045$) and five individual gene regions. ND3 showed a significant difference at $p = 0.0078$, while all other significant comparisons between these strata were at $p<0.05$ (Table 2).

We found no significant differences between the whitebelly and the eastern

subspecies using the concatenated mitogenome data (Φ_{ST} = 0.0181; p = 0.0869). In addition, five individual mitochondrial genes were significantly different. All individual gene partitions in which differentiation tests showed significance in spinner dolphins were found to be under purifying selection using Tajima's *D* tests for selection and Z-Test for positive selection using the Nei-Gojobori method (Nei and Gojobori 1986) (Supplementary Material).

Φ_{ST} tests showed no differentiation between Tres Marias spinners and either ETP spinner dolphin subspecies in either the concatenated or partitioned datasets. Five individual gene regions were significantly different in the pairwise comparisons of Tres Marias and whitebelly spinner dolphins (all at $p < 0.05$).

All tests involving comparisons with Hawaiian spinner dolphins (*S. l. longirostris*) were highly significant.

Nuclear SNPs

Compared to the mitogenome NGS data, coverage for nuclear genes was low and highly variable. This was likely due to variation in sample quality. Samples were collected as far back as 30 years ago and as recent as ten years ago. Older fisheries-collected samples were stored in saturated salt and DMSO solution for many years at

room temperature. In spinner dolphins, 56 samples had sufficient coverage at 51 SNP loci. These data included SNPs from six coloration genes, one Y-chromosome marker (UBE1) and 44 CATS loci. F_{ST} outlier analysis indicated that four were under positive selection and five were under balancing selection. Once these were removed the neutral dataset included 42 SNPs: five coloration SNPs, one Y-chromosome and 36 CATS loci (See Supplemental Material for details).

We had sufficient genotypes at the 51 SNPs for pairwise comparisons involving the ETP subspecies and the whitebelly, but not the putative Tres Marias stock. Divergence metrics based on genotypes from all loci - and all neutral loci - are shown in Table 3. For spinner dolphins, all estimates of G_{ST} were negative, meaning non-significant structure. This is likely due to the frequent occurrence of closely related individuals (based on these loci) from populations. Under this scenario, the null distribution will be more negative, because random permutations will shift of the null distribution to the left from where it would be if there was population structure. Comparing Central American and both whitebelly and eastern spinners using F_{ST} and Jost's *D* showed no significant differentiation. Between eastern and whitebelly spinner dolphins, F_{ST} and Jost's *D* were significantly different ($p < 0.01$).

Analyses of Bayesian clustering in STRUCTURE estimated the most likely number of unique clusters for spinner dolphins was K=1 based on LnP(K) (Figure A1-

S2). Using the metric Δ*K* (Evanno *et al.* 2005), *K*=2 was the most likely number of clusters estimated (see Supplemental Material); however, Δ*K* cannot evaluate *K*=1, so *K*=2 is the smallest value Δ*K* can estimate. These estimates were consistent regardless of whether or not location priors were used.

Four loci were estimated to be subject to positive selection in spinner dolphins: beta-carotene oxygenase 1 (BCDO), glucose transporter member 2-like gene (GLUT2), myeloperoxidase-like genes (MPO), and Wilm's tumor 1-like gene (WT1) (Figure A1-S4). GLUT2 codes for cellular membrane transporters. MPO codes for a common enzyme in blood, and WT1 is involved in the development of the urogenital system in humans. In addition, five SNPs within the following loci were estimated to be under balancing selection: COL10A, GLB79, LAPTM4A, NPPA, RHO.

Spotted dolphins

Mitogenomes

We assembled 76 complete or nearly complete (<10% missing data) spotted dolphin mitogenomes. Sample sizes, summary statistics and genetic diversity measures are listed in Table 1. Nucleotide diversity was higher in spotted dolphins (>0.0162) than spinner dolphins. Haplotypic diversity in the ETP spotted dolphins was high (>0.7000),

but not as high as spinner dolphins. Similar to the spinner dolphin mitogenome data, the substitution rate test did not detect any signs of saturation, and JModelTest (Posada, 2008) selected TrN (Tamura and Nei, 1993) as the best substitution model.

Results of F_{ST}, Φ_{ST} and χ^2 analyses of the mtDNA concatenated genes and Φ_{ST} of the individual gene regions for spotted dolphins are presented in Table 4. Similar to the spinner dolphins, our analyses at the subspecies level for spotted dolphins (coastal vs. offshore) show no significant differentiation using Φ_{ST} for the concatenated or partitioned datasets [F_{ST} (0.0125) was significant (p = 0.0402) in the concatenated dataset].

Estimates of differentiation between the current management stocks within the offshore subspecies (NE and WS stocks) using the whole mitogenome data and individual mtDNA genes showed no differences. However, we found nearly significant differences between the NE stock and the southern offshore samples using the concatenated mitogenome (Φ_{ST} = 0.1115; p = 0.0568), and three individual mtDNA genes showed significant Φ_{ST} differences (p<0.05) in comparisons of NE stock and the western offshore samples. Using Φ_{ST}, no significant differences were observed between the coastal subspecies and the NE offshore stock [although F_{ST} (0.0302) and χ^2 were highly significant at p = 0.0002]. Similarly, Φ_{ST} was not significant for pairwise comparisons of the Coastal subspecies and WS offshore stock using the concatenated data [χ^2 was significant (p = 0.0152) and was significant (p = 0.428) using Φ_{ST}].

Within the WS offshore stock, we found no differences between the southern and western offshore regions using the concatenated mitogenome, but Φ_{ST} estimates from four individual mtDNA genes had significant differentiation ($p < 0.05$) and three others had nearly significant p-values.

Comparing separate western and southern portions of the SW stock to other partitions using the mitogenome dataset also yielded no significant Φ_{ST} estimates. Comparison of the NE stock to the western portion of the WS stock yielded three individual gene regions with significant differentiation ($p < 0.05$), whereas neither dataset showed significant differences between the NE stock and the southern portion of the WS for any statistic. χ^2 was significant ($p = 0.0054$) between the coastal subspecies and southern offshore, but no individual gene region - nor the concatenated data - showed significant Φ_{ST} differentiation. Comparison of the coastal subspecies to just the western portion of the WS stock resulted in significant Φ_{ST} differences in four individual mtDNA genes (plus two genes showing low but non-significant p-values), but the concatenated dataset was not significant for any metric.

Our smaller sample size prevented us from partitioning the coastal subspecies south of central Mexico into the population units (described in Escorza-Trevino, Archer et al. 2005)).

Finally, we tested whether SST rather than latitude, determines stock boundaries by partitioning samples for differentiation tests at several temperatures regardless of sampling location. Habitat models have included SST as a predictor of spotted dolphin distributions (Barlow *et al.* 2009). Our results (Fig. 3) showed an increase in estimated population differentiation with a temperature boundary at 26.44°C (Φ_{ST}, not significant) and at 28.87°C (F_{ST}, significant at p = 0.0129).

Significant differentiation was also detected between Hawai‘i and the coastal subspecies in all three metrics and between Hawai‘i and offshore spotted dolphins in F_{ST}, and Φ_{ST}. Interestingly, significant differentiation only appeared in two individual mtDNA genes: ND1 and ND3 (see Table 4). We also detected significant differences between Hawai‘i and both the NE and WS stocks in these two genes, but not for the concatenated mtDNA dataset. For ND1, there were only two haplotypes in the five samples from Hawai‘i, neither of which was present in any of the other strata. Similarly, there were only two haplotypes for ND3 in the Hawaiian samples. For this gene, one sample had the most common haplotype shared among all sample strata, while the other samples shared a haplotype that was unique to Hawai‘i. Similar to the subspecies comparisons, the Hawaiian population also showed significant differences in ND1 and ND3 in comparisons with the western offshore group (ND1 and ND3 at $p < 0.001$) and the southern offshore group (ND1 at p = 0.0008 and ND3 at p = 0.0038).

Nuclear SNPs

Similar to the spinner dolphin samples, coverage for nuclear genes was low and highly variable. For spotted dolphins, 25 samples had sufficient coverage at 36 loci to call SNP genotypes. These data enabled use to conduct pairwise comparisons between the subspecies, but insufficient samples with genotypes prevented us from conducting comparisons between the two stocks of the offshore subspecies. The 36 loci included four coloration genes and 32 CATS loci. F_{ST} outlier analysis estimated that six SNPs were under positive selection and five were under balancing selection (Figure A1-S5). After removing these loci, the neutral dataset included 25 SNPs – two from coloration genes and 23 from CATS loci (See Supplemental Material for details).

Divergence metrics based on genotypes from all loci - and all neutral loci - are shown at the bottom of Table 3. For spotted dolphins, we detected significant differentiation between coastal spotted dolphins and offshore pantropical spotted dolphins in all three statistics ($p = 0.001$). It is worth noting, however, that 12 of the 13 samples from the offshore partition were from the WS stock, a majority of which was sampled in the southern region. Unfortunately, we were unable test for population structure between the offshore stocks (or other partitions proposed by Escorza *et al.* 1995) because the numbers of samples and loci were too few.

Analyses of Bayesian clustering in STRUCTURE using all SNPs estimated the most likely number of unique clusters for spotted dolphins was K=2 based on LnP(K) (Figure A1-S3). Using the metric ΔK (Evanno *et al.* 2005), K=2 was the most likely number of clusters estimated for spotted dolphins (see Supplemental Material); however, ΔK cannot evaluate K=1, so K=2 is the smallest value ΔK can estimate. These estimates were consistent regardless of whether location priors were used.

Six loci were estimated to be subject to positive selection in spotted dolphins: beta-carotene oxygenase 1 (BCDO), FES proto-oncogene, tyrosine kinase (FES), homeobox protein Hox-C8 **(**HOXc8), myosin heavy chain 4 **(**MYH4), somatostatin (SST) and tyrosinase-related protein 1 (TYRP1). BCDO produces a key enzyme in beta-carotene metabolism to vitamin A, which is a vital component in processes like vision, development, cell differentiation and skin color and protection. FES codes for tyrosine kinase: an important component of cellular transformation. HOXc8 plays an important role in morphogenesis in mammals. MYH4 is involved in building motor proteins for muscle contraction. Somatostatin codes for the hormone of the same name, which is an important regulator of the endocrine system. TYRP1 encodes for an enzyme in melanocytes that produce melanin. In addition, five loci were estimated to be under balancing selection in this species: ADH, AMBP, CHRNA, ELN, FSHB.

DISCUSSION:

Pelagic spinner and spotted dolphins in the eastern tropical Pacific have presented a paradox for management, whereby named subspecies and stocks described based on morphological differences have exhibited little to no genetic differentiation in previous studies (Dizon *et al.* 1994; Galver 2002). Recent divergence, large population sizes and ongoing gene flow likely contribute to low genetic divergence and low statistical power to detect this divergence (Taylor and Dizon 1996; Waples 1998). To overcome the limitations of using few genetic markers in highly abundant and genetically variable populations, we have expanded the genetic analyses to include complete mitochondrial genomes and sets of nuclear SNPs. Our results show genetic support for endemic subspecies of spinner and spotted dolphins, although the strength of this support varies between markers. Moreover, we did not find support for the division of offshore stocks of spotted dolphins, nor did we find separation of the Tres Marias spinner dolphins as an independent population.

Endemic Spinner Dolphins Subspecies: Eastern and Central American

Traditional F_{ST} calculated for the concatenated spinner dolphin mitogenomic dataset was very low as expected due to high abundance and haplotype diversity, but significant p-values corroborate endemic subspecies distinctions (Central American vs. eastern). In contrast, nuclear SNP data did not support differences between these endemic subspecies. The discordance between the two sets of markers could result from a difference in the power of the analyses or for biological reasons. The 648 variable sites across the mitogenome represent a relatively large dataset. On the other hand, the 51

SNPs might not have enough statistical power to detect differences in these subspecies because of recent divergence, continued low-level interbreeding as well as high diversity and historical abundance. Differences in the rate of evolution in the two genomes (i.e., mitogenome evolving faster than the nuDNA genome) could have resulted in more signal from drift appearing in allele frequencies due to the smaller N_e of the mitogenome (Moritz 1994). Alternatively, the discordance we observed could be a result of male-mediated exchange diluting the signal of structure in nuDNA or female site-fidelity increasing structure in the mtDNA. Although there is some evidence from tagging studies that some dolphins move substantial distances (Perrin *et al.* 1979), a thorough investigation into the differences between genders is lacking. We believe it is therefore most likely that the SNP dataset lacks statistical power and suggest increasing the number of SNPs in future studies.

Despite the lack of support from our nuDNA SNP analyses, our results provide evidence of genetic differentiation between the accepted ETP endemic subspecies (Perrin et al. 1991) concordant with morphology and results from Andrews *et al.* (2013) using data from the nuclear Actin gene. Differences in ecological, distributional, morphological, nuDNA and now mtDNA data support the recognition of these distinct subspecies.

Whitebelly Spinner Dolphins

Our comparisons involving the putative intergrade, the whitebelly spinners, also revealed discordant patterns between the two different datasets. Mitochondrial Φ_{ST} estimates indicate separation between the Central American and whitebelly forms, but nuclear SNP data failed to show differences between these groups. Every whitebelly sample had a unique mitogenome haplotype, and as a result, frequency-based measures of differentiation such as F_{ST} will have low power. Because Φ_{ST} incorporates the degree of differentiation among sequences, it provides additional support for divergence of the subspecies even when there is very high haplotype diversity.

Two possible explanations for the differentiation found between Central American and whitebelly spinners in the mtDNA are isolation by distance and admixture between whitebellies and Gray's spinners. Andrews *et al.* (2013) also found these two forms to be significantly different using mtDNA genes (control region and cyt*b*). We did not recover this pattern for those genes, but we did find structure among several different mtDNA genes (Table 2). Andrews *et al.* (2013) used a similar sample set to ours, but included 10 samples of Central American spinner that had questionable subspecific identity (based on further investigation of the sample collection records at SWFSC by MSL). Samples were initially identified as Central American spinners, but the confidence in the identification was low and they should have been cataloged with an "unidentified" species assignment. This difference in sample sets could be the reason for the differing results in comparisons using the Central American strata. Unfortunately, our sample size was low for the Central American spinners after removal of these questionable samples

(mtNDA: N=9; nuDNA: N=7). The Central American subspecies, with lower relative abundance, might be expected to show higher levels of structure due to drift, and comparisons between this subspecies and the eastern subspecies showed evidence of weak structure based on allele frequencies.

We found no significant differences between the whitebelly and the eastern subspecies using the concatenated mitogenome data (the Φ_{ST} estimate was nearly significant at p = 0.0869). However, five genes within the mitogenome were significantly different and two more were almost significant (ND1 and ND4L). Andrews *et al.* (2013) inferred high migration rates between whitebelly and eastern spinner dolphins (30.1 migrants per generation from whitebelly to eastern and 57.9 migrants from eastern to whitebelly). Despite this high rate of migration, we see significant differences in five of the fifteen individual mtDNA genes and in the SNP dataset. High haplotype diversity in the concatenated mitogenome dataset likely lowered statistical power to detect differences based on allele frequencies.

Loci Under Selection

Finding four of the 51 SNP loci (7.8%) in spinner dolphins and six of the 36 SNP loci (16.7%) in spotted dolphins exhibiting positive selection was unexpected. However, other authors have reported similar rates: Russello *et al.* (2011) found eight of their 52 loci (15.4%) to be outliers and useful for detecting ecotype divergence in Okanagan Lake

kokanee (*Oncorhynchus nerka*). Bay and Palumbi (2014) found 2807 of 15399 SNPs (18.2%) to be F_{ST} outliers (before further filtering for analyses) in a population of tabletop corals. We removed all outlier loci (positive and balancing selection) for all population divergence statistics and the results remained the same as with all the data included (Table 3). The FDIST2 (Beaumont and Nichols 1996) method implemented in *Lositan* can be prone to false positives, although it generally performs better than ARLEQUIN (Excoffier and Lisher 2010) at minimizing type I error when detecting F_{ST} outliers (Narum and Hess 2011).

We hypothesized that positive selection on nuclear loci coding for skin pigment coloration could be contributing to the marked differences in coloration between the forms of spinner dolphins in the ETP. Simulation-based tests for selection at each locus using the program *Lositan* (Antao et al 2008) found two coloration genes that were subject to positive selection: BCDO and TYRP1. In both species, BCDO was under positive selection, while spotted dolphins also exhibited positive selection in TYRP1. Pairwise F_{ST} estimates were significantly different between coastal and offshore spotted dolphin subspecies ($p < 0.006$ for both loci) and between ETP endemic spinner dolphin subspecies ($p < 0.05$) (Table A1-S4). These results are interesting and indicate a molecular basis for the differences in coloration observed in this region for these species. Further analysis with larger sample sizes will be needed to verify this finding.

Contrasting mtDNA and nuDNA results

High site fidelity in males could also restrict male-mediated geneflow between groups and increase relative signal in nuDNA analyses. Unfortunately, very little is known about the movement patterns of individual dolphins in the ETP, and less is known about differences in movement based on gender. Assortative mating can also decrease N_e, which could serve to amplify signal of structure in the nuDNA genome. The eastern spinner dolphin is thought to have a more polygynous mating system than the whitebelly form, and Perrin and Mesnick (2003) deduced that reproductive skew is high in eastern spinners, which could reduce N_e and inflate genetic structure in the nuclear loci. A skewed breeding system might also increase migration, however, as dominance might promote movements in juvenile males.

The statistical power to estimate levels of migration between very large populations with low relative sample sizes is weak (Waples 1998; Taylor et al. 2000). For this reason, we did not estimate levels of migration for these data. Andrews *et al.* (2013) did estimate migration in ETP spinner dolphins and found lower, but significantly different from zero, rates of migration per generation between populations of Gray's spinners and the white belly spinners (3.22 migrants per generation into Gray's and 1.6 into whitebelly spinners). The rate of migration into Gray's spinner populations from the eastern population was estimated to be less than one (0.82), but significantly different from zero. Although this was not a major focus of our study, the differences we detected between the Hawaiian population and the ETP pelagic populations were higher than any

comparisons within the ETP, supporting the hypothesis that this is an insular population.

Nuclear SNP data showed significant differences between the eastern subspecies and the intergrade whitebelly. When the small Central American sample set was combined with the eastern sample set, the nuDNA divergence values and significance levels did not change, meaning the signal is mostly due to the differences between the whitebelly and eastern types. These results are concordant with phenotype and Y-chromosome differences (Andrews *et al.* 2013) whereby the most noticeable differences occur at the intersection of the whitebelly and eastern. We agree with these authors that there is likely a porous barrier to gene flow across the eastern Pacific basin, as mixed groups are common and interbreeding probably occurs with some regularity between eastern and Central American subspecies and between eastern and whitebelly spinners. However, we feel that this 'introgression zone' between whitebelly spinners and eastern spinners deserves further investigation. We hypothesize that either divergence with gene flow is ongoing in this area or the whitebelly spinner is the result of a recent reconnection in an area of historical separation across a known biogeographic boundary, the east Pacific basin.

STRUCTURE estimated K=1 based on LnP(D). The lack of detectable structure could be because STRUCTURE's clustering algorithm attempts to maximize Hardy-Weinberg equilibrium within clusters of samples. Datasets such as our spinner dolphin

data, that represent a relatively small proportion of the overall variation of the true populations, have been shown to lack power to detect population clusters defined by Hardy-Weinberg alone (Kalinowsky 2010).

Alternative spinner dolphin stocks:

We found no support for a Tres Marias population that differs from the eastern subspecies using the concatenated or individual mitogenome data set. This is unsurprising given the weak genetic differences we found between the accepted endemic subspecies with much more marked morphological differences. We found statistically significant differences in five individual mtDNA genes when comparing the Tres Marias group to the whitebelly spinners and two more genes (12S and ND1) that were nearly significant ($p = 0.0639$ and 0.0766, respectively). These seven genes taken together are the same genes that differentiate the eastern subspecies (including Tres Marias samples) from the whitebelly intergrade. Unfortunately, we were unable to test hypotheses that the Tres Marias differs from other groups with the nuDNA due to small sample sizes.

Spotted dolphin subspecies:

The spotted dolphin mitogenomic data have lower haplotypic diversity but higher nucleotide diversity relative to the spinner dolphins, despite extremely high historical population sizes. The two main reasons for lower haplotypic genetic diversity could be a recent and/or prolonged population bottleneck, such as the decrease caused by mortalities

in the tuna purse-siene fishery, or an extremely matrifocal social structure (Hoelzel *et al.* 2007). Although matrifocal social structure is known in several species of odontocetes (e.g., killer whales and sperm whales), it is not a known characteristic of spotted dolphins.

Similar to our findings for spinner dolphins, traditional F_{ST} calculated for the mitogenome dataset supports differentiation of the offshore *S. a. attenuata* and the endemic coastal *S. a. graffmani* subspecies, whereas Φ_{ST} failed to indicate any difference, either for the entire genome or within any single gene. The χ^2 test of differentiation was nearly significant ($p = 0.07$) when the offshore subspecies stocks were combined. The coastal subspecies was found to be different from each offshore stock however, with the NE stock being more strongly differentiated from the coastal subspecies than the WS stock, counter to the results found by Escorza-Treviño et al (2005) using microsatellite data. In that study, the authors inferred that there was a strong connection between the coastal and offshore subspecies in northern Mexico. The differences between the two results could be due to sampling; Escorza-Treviño et al (2005) had more samples from the northern portion of the coastal spotted dolphin range than we did. Additionally, the differences could be attributed to the unique evolutionary patterns of the different markers examined (microsatellites vs. SNPs).

Our analyses of population structure using 36 SNP loci showed highly significant

population structure between the two subspecies using traditional *F*-statistics and cluster analyses, supporting the hypothesis that two subspecies exist in the ETP. The SNP dataset did not include individuals from the NE offshore stock of spotted dolphins, however, so this comparison includes animals from the most geographically separate portions of the offshore (WS) and coastal subspecies range. Additional nuclear data from the NE stock are needed to determine whether proximate populations of these two subspecies are also as genetically divergent.

Spotted dolphin stocks:

A main objective of this work was to test for difference between existing and proposed management stocks. Using the whole mtDNA genome dataset, we found no differentiation between the western and the southern offshore spotted dolphins. However, despite the morphological similarity noted by Perrin *et al.* (1994), we found differences in four mtDNA loci between these two strata. Similarly, we failed to detect differences between the NE stock and the western group of the WS stock using the whole mtDNA genome; however, three mtDNA loci had significant Φ_{ST} estimates. Although, we did not find significant differences between the NE and southern groups, the Φ_{ST} p-value for this test was nearly significant (p=0.0568), suggesting that the distributional hiatus at 5° north might be a true barrier to gene flow.

We examined the ability of sea surface temperature (SST) to act as an

environmental boundary between NE and WS offshore stocks. Results showed higher Φ_{ST} estimates of population differentiation of the concatenated mitogenome data for groups separated by the 26.44°C (though not significant) and 28.87°C (F_{ST}, $p < 0.05$) boundaries, the upper and lower portions of the temperature range. Interestingly, the average barrier between 25°C and 26°C SST falls on the 5°N parallel, the location of the distributional hiatus. Although not significantly different, our high Φ_{ST} estimates (Φ_{ST} = 0.0514) when the samples were divided at 26.44°C could indicate a boundary in this area. Splitting samples at 28.87°C results in six samples that were collected in waters warmer than 28.87°C within the tropical surface water and the Eastern Pacific Warm Pool (see Fig. 2 in Fiedler and Talley, 2006) off the coast of north-central Mexico. Pooling these samples might concentrate the signal of differences between offshore stocks (there was only one sample from south of the equatorial front in the warm water sample set). Given the similarities Escorza-Treviño *et al.* (2005) found between northern coastal spotted dolphins and NE offshore spotted dolphins, it is also possible that some of the signal is due to admixture between the northern offshore animals and the coastal subspecies of northern Mexico. Alternatively, these results could be an artifact of the sampling regime and smaller sample sizes for the partitions on the limits of the range. For the tests conducted where the barrier was placed at 26.44°C, there were 39 samples above and 8 samples below this cut-off. For tests conducted at 28.87°C there were 6 samples above and 41 samples below this cut-off. The larger differences in these two tests could be a result of disparity in sample size between these.

Gene-level differentiation in the mitochondrial genome

Levels of population differentiation varied across different mtDNA gene regions. Here we better characterize which genes showed differentiation, as they might be useful for future population genetic studies of these taxa.

In spotted dolphins, no individual gene showed differentiation between the coastal subspecies and the offshore nominate subspecies. The significant differences between ETP groups and Hawaiian insular population were driven by a strong signal of differentiation in two loci (ND1 and ND3; Table 4). Both of these loci also had fewer total haplotypes (ND1 = 15; ND3 = 10) than the other genes. No haplotypes were shared between Hawaii and the offshore or coastal sample strata for the ND1 gene. One sample from Hawaii shared the most common haplotype with the offshore and coastal sample strata for ND3, the other four were unique to Hawaii.

According to Duchene *et al.* (2010), ND3 has one of the highest substitution rates of mitochondrial genes in delphinids (5.31 sub/site/my), falling outside of the 95% highest posterior density interval (HDPI) of the median rate for the whole mitogenome (3.70-4.76) of all delphinids. On the other hand, ND1 has an average substitution rate among delphinids (4.24). Within the genus *Orcinus* both genes are within the 95% HDPI of the median *Orcinus* mitogenome rate. It is possible that these two genes in the spotted

dolphins from Hawaii have reduced diversity because of random drift of lineages or due to purifying selection.

For spinner dolphins, the gene regions showing signals of population differentiation between Hawaii and ETP groups differed somewhat from those showing signal between ETP groups. ND3 and COI were significant among all comparisons. Five gene regions showed significant Φ_{ST} between the Central American and whitebelly spinners. Two of these are high mutation rate genes (ND3: 5.31 and ATP8: 5.21), and the remaining (ND1, ND4 and COII) fell within the 95% HDPI of the median rate for all delphinids (Duchene *et al.* 2010). All but two of these markers (COII and COIII) were slightly above average. Of the four genes that showed significant differences between the whitebelly and eastern spinners, one (ND3) has a high substitution rate (5.31), and two fell within the 95% HDPI of the median mitogenome rate but were below the average.

We tested for positive selection in spinner dolphin mitochondrial genes with significant Φ_{ST} estimates. All of these regions were found to be under purifying selection (negative Tajima's *D*) indicating that the within-mitogenome differences are accumulating by neutral drift rather than via positive selection.

CONCLUSIONS:

Defining population genetic structure for management in depleted species with large historical population sizes and high mobility can be challenging. These populations can retain high genetic variation even as abundance becomes relatively low, which could be obscuring signals of genetic structure used to designate stock boundaries for estimating population abundance and setting stock-specific mortality limits. Ultimately, without information on structure, populations could be under-classified and unique evolutionary units and populations could go extinct as we may fail to take appropriate conservation action.

Our results show a complex pattern of genetic structure in the two different datasets for each species. The mitogenome data show support for the endemic ETP spinner and spotted dolphins subspecies. The nuclear SNP data show strong support for spotted dolphin subspecies but failed to find segregation in morphologically divergent spinner dolphin subspecies, though small sample size for the coastal subspecies limited power to detect genetic divergence. A lack of differentiation in the SNP data between the two most geographically distinct groups (the whitebelly and the Central American) was unexpected (although the Jost's *D* measure was almost significant at 0.08), especially since strong differences were detected between the more proximate eastern and whitebelly forms. As described above, the genetic structure between the eastern and whitebelly could be in part due to some level of divergent selection in the nuclear genome despite gene flow between these forms combined with mitochondrial introgression between the eastern and whitebelly spinners. Alternatively, it could be that

the lack of observed differences between the eastern and Central American subspecies is due to increased exchange because of higher population density near the shallow thermocline zone closer to the Central American coast.

Analyses of partitioned mitogenomes indicated that slow-, medium- and fast-substitution-rate gene regions were differentiated in comparisons with the Hawai‘i population, while those genes showing evidence of differentiation in ETP populations had faster (high and medium) substitution rates. These results indicate that differences within the mitogenome are accumulating by neutral drift rather than with the aid of positive selection.

We found very little support for the division of offshore stocks of spotted dolphins and no support for the unique form of Tres Marias spinner dolphins as compared to the eastern or Central American subspecies. This is not to say that these biological entities do not exist, just that our data may not have sufficient power to detect the subtle genetic differences between them. Efforts should be made to sequence larger nuclear regions from additional specimens collected throughout the range of these animals to perform high-resolution population structure analyses. Further, we suggest the collection and analysis of additional samples from the Central American subspecies to compare to existing offshore subspecies samples collected from fisheries bycatch and research cruises. In addition, we highly recommend additional studies of nuDNA and studies of

population structure that incorporate environmental variables as potential population boundaries in this area. A follow-up study collecting data on samples that span the SST gradient would be informative. Finally, placing these populations within a global phylogeographic context will help provide a better context for our results by fully characterizing intraspecific diversity and establishing the evolutionary process that led to ETP endemism.

ACKNOWLEDGEMENTS:

MSL was supported by a National Science Foundation (NSF) Graduate Research Fellowship, a NSF Integrative Graduate Education and Research Traineeship Fellowship, the Ralph A. Lewin Graduate Fellowship in Marine Biology at Scripps Institution of Oceanography, the Lerner-Grey Memorial Foundation at the American Museum of Natural History, and the Edna Bailey Sussman Foundation. We would especially like to thank Dr. William F. Perrin who provided guidance and support throughout the development and completion of this manuscript. Danielle Davila, Brittany Hancock-Hanser, Gabriela Serra-Valente, Victoria Pease, Kelly Robertson, Dr. Louella Dolar, Nicole Beaulieu and Morgane Lauf for help in the laboratory and tissue archive. Dr. Karen Martien provided a thoughtful critique of this manuscript. Drs. Jessica Redfern and Paul Fiedler provided key data and discussion regarding the SST analyses. Al Jackson also provided copies of original field datasheets. We thank Dr. Robin Baird of Cascadia Research Collective and Dr. Erin Oleson of the NOAA Pacific Islands Science Center for permission to use Hawaiian spinner dolphin samples. The Scripps Research Institute

provided assistance with sequencing and troubleshoots, especially Drs. Steve Head, Lana Shaffer and John Shimashita. Support for this research was also provided from Drs. Lisa T. Ballance and Barbara Taylor of the Marine Mammal and Turtle Division of SWFSC.

Chapter 2 is currently in review for publication in *Marine Mammal Science*. Leslie, Matthew S.; Archer, F. I.; Morin, Phillip A. The dissertation author was the primary investigator and lead author of this material.

LITERATURE CITED:

Aitken, N., S. Smith, C. Schwarz, and P. A. Morin. 2004. Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. Molecular Ecology 13:1423-1431.

Andrews, K. R., W. F. Perrin, M. Oremus, et al. 2013. The evolving male: spinner dolphin (*Stenella longirostris*) ecotypes are divergent at Y chromosome but not mtDNA or autosomal markers. Molecular Ecology 22:408-2423.

Antao, T., A. Lopes, R. J. Lopes, A. Beja-Pereira, and G. Luikart. 2008. LOSITAN: A workbench to detect molecular adaptation based on a F_{ST}-outlier method. BMC Bioinformatics, 9(1):323.

Balmaseda, M. A., K. Mogensen, and A. T. Weaver. 2013. Evaluation of the ECMWF ocean reanalysis system ORAS4. Quarterly Journal of the Royal Meteorological Society 139:1132–1161.

Barlow, J., M. C. Ferguson, E. A. Becker, *et al.* 2009. Predictive Modeling of Cetacean Density in the Eastern Pacific Ocean. NOAA Technical Report NMFS 444, 206pp.

Beaumont M. A., and R. A. Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society of Londong B 263:1619-1626.

Dizon, A.E., W. F. Perrin, and P. A. Akin. 1994. Stocks of dolphins (*Stenella* spp. and *Delphinus delphis*) in the eastern tropical Pacific: a phylogeographic classification. NOAA Technical Report NMFS 119, 20pp.

Dornburg, A., M. C. Brandley, M. R. McGowen, and T. J. Near. 2012. Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea). Molecular Biology and Evolution 29:721-736.

Douglas, M.E., G. D. Schnell, D. J. Hough, and W. F. Perrin. 1992. Geographic variation in cranial morphology of spinner dolphins *Stenella longirostris* in the eastern tropical Paciic Ocean. Fisheries Bulletin 90:54-76.

Drummond, A., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

Duchene, S., F. I. Archer, J. Vilstrup, S. Caballero, and P. A. Morin. 2011. Mitogenome Phylogenetics: The impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PLoS ONE 6(11): e27138. doi:10.1371/journal.pone.0027138.

Eriksson, J., G. Larson, U. Gunnarsson, B. Bed'hom, M. Tixier-Boichard, et al. 2008. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS Genetics 4(2): e1000010 10.1371/journal.pgen.10000010.

Escorza-Treviño, S., F. I. Archer, M. Rosales, A. Dizon, and A. Lang. 2005. Genetic differentiation and intraspecific structure of Eastern Tropical Pacific spotted dolphins, *Stenella attenuata*, revealed by DNA analyses. Conservation Genetics 6(4), 587-600.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611-2620.

Excoffier, L. and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources10, 564-567.

Excoffier, L., P. E. Smouse and J. M. Quattro (1992). "Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data." Genetics **131**(2): 479-491.

Fiedler, P. C., and L. D. Talley. 2006. Hydrography of the eastern tropical Pacific: A review. Progress in Oceanography, 69(2):143-180.

Foll, M. and O. E. Gaggiotti. 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180:977-993.

Galver, L. 2002. The molecular ecology of spinner dolphins, *Stenella longirostris*: genetic diverstiy and population structure. Ph.D. Dissertation, University of California.

Gerrodette, T., G. Watters, W. Perryman, and L. T. Balance. 2008. Estimates of 2006 dolphin abundance in the eastern tropical Pacific, with revised estimates from 1986-2003. NOAA Tech. Memo. NMFS-SWFSC-422.

Gosliner, M. L. 1999. The tuna-dolphin controversy. Pages 120-155 *in* J. R. Twiss Jr. and R. R. Reeves, eds. Conservation and management of marine mammals. Smithsonian Institution Press, Washington DC.

Hancock-Hanser, B., A. Frey, M. S. Leslie, P. H. Dutton, and P. A. Morin. 2013. Targeted multiplex next-generation sequencing: Advances in techniques of mitochondrial and nuclear sequencing for population genomics. Molecular Ecology Resources 13:254-268.

Hoekstra, H. E., R. J. Hirschmann, R. A. Bundey, P. A. Insel, and J. P. Crossland. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science, 313(5783):101-104.

Hoelzel, A. R., J. Hey, M. E. Dahlheim, C. Nicholson, V. Burkanov, and N. Black. 2007. Evolution of population structure in a highly social top predator, the killer whale. Molecular Biology and Evolution 24 (6):1407-1415.

Horne, J. B. 2014. Thinking outside the barrier: neutral and adaptive divergence in Indo-Pacific coral reef faunas. Ecology and Evolution DOI 10.1007/s10682-014-9724-9.

Hubbard, J. K., J. A. C. Uy, M. E. Hauber, H. E. Hoekstra, and R. J. Safran. 2010. Vertebrate pigmentation: from underlying genes to adaptive function. Trends in Genetics, 26(5):231-239.

Hubisz, M., D. Falush, M. Stephens, and J. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322-1332.

Joseph, J. 1994. The tuna-dolphin controversy in the eastern Pacific Ocean: Biological, economic and political impacts. Ocean Development and International Law 25:1-30.

Jost, L. 2008. G_{ST} and its relatives do not measure differentiation. Molecular Ecology 17: 4015-4026.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pages 21-132 in H. M. Munro, ed. Mammalian protein metabolism. Academic Press, New York.

Kalinowsky, S. T. 2010. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106(4):625-632.

Katoh, K., G. Asimenos, and H. Toh. 2009. Multiple alignment of DNA sequences with MAFFT. Methods in Molecular Biology 537:39–64.

Lambertsen, R. H. 1987. A biopsy system for large whales and its use for cytogenetics. Journal of Mammalogy 68:443-445.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows - Wheeler transform. Bioinformatics 25:1754-1760.

Li, H., B. Handsaker, A. Wysoker, *et al.* 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

Lo, N., and T. D. Smith. 1986. Incidental mortality of dolphins in the eastern tropical Pacific, 1959-72. Fishery Bulletin 84: 27-34.

Lyons, L., T. Laughlin, N. Copeland, *et al.* 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nature Genetics 15:47-56.

McKenna, A., M. Hanna, E. Banks, *et al.* 2009. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20(9):1297-1303.

Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. Molecular Ecology 3:401-411.

Narum, S. R., and J. E. HESS. 2011. Comparison of F_{ST} outlier tests for SNP loci under selection. Molecular Ecology Resources 11:184–194.

National Research Council. 1992. Dolphins and the tuna industry. National Academy Press, Washington DC.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321-3323.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press.

Nei, M., and Gojobori T. 1986. Simple method for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution 3:418-426.

Norris, Richard D. 2000. Pelagic species diversity, biogeography, and evolution. Paleobiology 26.sp4: 236-258.

Perrin, W. F., P.A. Akin, and J. V. Kashiwada. 1991. Geographic variation in external morphology of the spinner dolphin *Stenella longirostis* in the eastern Pacific and implications for conservation. Fisheries Bulletin 89(3):411-428.

Perrin, W. F., M. L. L. Dolar, and D. Robineau. 1999. Spinner dolphins (*Stenella longirostris*) of the western Pacific and Southeast Asia: pelagic and shallow-water forms. Marine Mammal Science 15(4): 1029-1053.

Perrin, W. F., W. E. Evans and D. B. Holts. 1979. Movements of pelagic dolphins (*Stenella* spp.) in the eastern tropical Pacific as indicated by results of tagging, with summary of tagging operations, 1969-76. NOAA Technical Memo NMFS 737:14 pp.

Perrin, W. F., and J. R. Henderson, 1979. Growth & reproductive rates in two populations of spinner dolphins, *Stenella longirostris*, with different histories of exploitation. Reports of the International Whaling Commission Special Issue 6:417–430.

Perrin W. F., and S. L. Mesnick. 2003. Sexual ecology of the spinner dolphin, *Stenella longirostris*: geographic variation in mating system. Marine Mammal Science 19:462-483.

Perrin, W. F., N. Miyazaki, and T. Kasuya. 1989. A dwarf form of the spinner dolphin (*Stenella longirostris*) from Thailand. Marine Mammal Science 5:213-227.

Perrin, W. F., G. D. Schnell, D. J. Hough, J. W. Gilpatrick Jr., and J. V. Kashiwada. 1994. Reexamination of geographic variation in cranial morphology of the pantropical spotted dolphin, *Stenella attenuata*, in the eastern Pacific. Fisheries Bulletin 92(2):324-346.

Perrin, W. F., M. D. Scott, G. Walker, F. M. Ralston, and D. W. K. Au. 1983. Distribution of four dolphins (*Stenella* spp. and *Delphinus delphis*) in the eastern tropical Pacific, with an annotated catalog of data sources. NOAA Technical Memo NMFS 38:65 pp.

Perrin, W. F., M. D. Scott, G. J. Walker, and V. L. Cass. 1985. Review of geographical stocks of tropical dolphins (*Stenella* spp. and *Delphinus delphis*) in the eastern Pacific. NOAA Technical Memo NMFS 28. 28 pp.

Perrin, W. F., E. D. Mitchell, J. G. Mead, D. K. Caldwell, M. C. Caldwell, P. J. H. Van Bree, and W. H. Dawbin. 1987. Revision of the spotted dolphins, *Stenella* spp. Marine Mammal Science 3:99–170.

Perryman, W. L., and R. L. Westlake. 1998. A new geographic form of the spinner dolphin, *Stenella longirostris*, detected with aerial photogrammetry. Marine Mammal Science 14:38-50.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution 25:1253-1256.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155.2:945-959.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rieder, S., S. Taourit, D. Mariat, B. Langlois, and G. Guérin. 2001. Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association

to coat color phenotypes in horses (*Equus caballus*). Mammalian Genome, *12*(6):450–455.

Russello, M. A., S. L. Kirk, K. K. Frazer, and P. J. Askey. 2012. Detection of outlier loci and their utility for fisheries management. Evolutionary Applications 5:39–52.

Scott, M. D., S. J. Chivers, R. J. Olson, P. C. Fiedler, and K. Holland. 2012. Pelagic predator associations: tuna and dolphins in the eastern tropical Pacific Ocean. Marine Ecology Progress Series 458:283-302.

Stamatakis, A., T. Ludwig, and H. Meier. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21:456-463.

Stewart, J. B., C. Freyer, J. L. Elson, A. Wredenberg, Z. Cansu, A. Trifunovic, N. Larrson (2008) Strong Purifying Selection in Transmission of Mammalian Mitochondrial DNA. PLoS Biol 6(1): e10. doi: 10.1371/journal.pbio.0060010

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105(2):437-460.

Tamura, K. and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10:512–26.

Tamura, K., D. Peterson, N. Peterson, *et al.* 2011. MEGA5: Molecular Evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.

Tavare, S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17:57-86.

Taylor, B. L., and Dizon A. E. 1996. The need to estimate power to link genetics and demography for conservation. Conservation Biology 10:661–664.

Taylor, B. L., Chivers, S. J., Sexton, S., and Dizon, A. E. 2000. Evaluating Dispersal Estimates Using mtDNA Data: Comparing Analytical and Simulation Approaches. Conservation Biology 14:1287–1297.

Wade, P. R. 1995. Revised estimates of incidental kill of dolphins (Delphinidae) by the purse-seine tuna fishery in the eastern tropical Pacific, 1959-1972. Fishery Bulletin 93:345-354.

Wade, P. R., G. M. Watters, T. Gerrodette, and S. B. Reilly. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modelling hypotheses for their lack of recovery. Marine Ecology Progress Series 343:1-14.

Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity 89(5):438-450.

Weir, B.S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38(6):1358-1370.

Wright, S. 1949. The genetical structuring of populations. Annals of Eugenics 15(1):323-354.

Xiong, Y., M. C. Brandley, S. Xu, K. Zhou, and G. Yang. 2009. Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. BMC Evolutionary Biology 9:20.

Table 1. Summary statistics for ETP spinner (A) and spotted (B) dolphin mitogenome data.

A. Spinner dolphins *Stenella longirostris* (n = 104)						
Subspecies/Stock	**n:female/male/unk**	**n_H**	**PS**	**h**	**π**	**%**
Central American (*S. l. centroamericana)*	9:4/4/1	8	648	0.9722	0.0057	0.7778
eastern~ (*S. l. orientalis)*	53:28/19/6	51	648	0.9985	0.0073	0.9245
Putative Stocks						
whitebelly* (*S. l. longirostris*)	27:16/11/0	27	457	1	0.0043	1
Tres Marias^~ (*S. l. orientalis*)	21:8/10/3	20	373	0.9952	0.0078	0.9048
Hawai'i (*S. l. longirostris*)	15:1/4/10	9	104	0.9921	0.0068	0.8260
B. Spotted dolphins Stenella attenuata (n = 76)						
Subspecies	**n:female/male/unk**	**n_H**	**PS**	**h**	**π**	**%**
Coastal (*S. a. graffmani*)	24:11/13/0	16	234	0.9529	0.0162	0.5000
ETP offshore[§] (*S. a. attenuata*)	47:20/19/8	43	519	0.9804	0.0198	0.7222
Offshore Stocks (*S. a. attenuata*) - Current and Putative^						
northeastern	25:10/8/7	22	400	0.9867	0.0238	0.8000
western-southern	17:9/7/1	17	298	1	0.0096	1
Offshore western^	8:7/1/0	8	191	1	0.0087	1
Offshore southern^	9:2/6/1	9	253	1	0.0092	1
Hawai'i	5:1/3/1	3	36	0.7000	0.0244	0.4000

*n_H: number of haplotypes; PS: polymorphic sites; h: haplotype diversity; π: nucleotide diversity; %: percent of unique haplotypes. *Taxonomically the whitebelly spinner is part of the Gray's subspecies; however, it is thought to be an intergrade between the Gray's and eastern subspecies. ^ Stocks that are not recognized for management purposes. ~ The Tres Marias spinner samples are part of the eastern strata. § Includes data for five samples that were omitted from stock comparisons because they were sampled too close to geographic stock boundaries.*

Table 2: Pairwise divergence estimates for subspecies and stocks of spinner dolphins using concatenated mitogenome data (F_{ST}, Φ_{ST} and χ^2) and partitioned mitogenomic data (Φ_{ST} only). Light gray backgrounds for p<0.05; medium gray for p<0.01; darker gray backgrounds for p<0.001 (p-values in parentheses).

	Concatenated mitogenome			Partitioned mitogenome Φ_{ST} (p-value)														
Taxon 1 (n) vs. Taxon 2 (n)	F_{ST} (p-value)	Φ_{ST} (p-value)	χ^2 (p-value)	12s n_H=24	16s n_H=24	ATP6 n_H=53	ATP8 n_H=11	COI n_H=65	COII n_H=39	COIII n_H=47	CYTB n_H=61	CR n_H=50	ND1 n_H=59	ND2 n_H=53	ND3 n_H=21	ND4 n_H=56	ND4L n_H=22	ND5 n_H=70
Central Amer. (9) vs. eastern (53)	0.0133 (0.034)	-0.0126 (0.5055)	62 (0.3916)	-0.0110 (0.5176)	0.0049 (0.3531)	-0.0096 (0.4165)	0.0585 (0.0849)	-0.0263 (0.8306)	-0.0137 (0.5878)	-0.0098 (0.5031)	-0.0095 (0.4731)	0.0388 (0.1145)	0.0233 (0.2016)	-0.0308 (0.8040)	0.0355 (0.1322)	-0.0265 (0.6596)	-0.0270 (0.7536)	-0.0135 (0.5145)
Central Amer. (9) vs. whitebelly (27)	0.0128 (0.056)	0.0491 (0.045)	36 (0.6494)	-0.0151 (0.6256)	0.0198 (0.2567)	0.0319 (0.1419)	0.1276 (0.0419)	0.0386 (0.0787)	0.1156 (0.0386)	0.0229 (0.1592)	0.0474 (0.1074)	0.0621 (0.0606)	0.0851 (0.0438)	-0.0126 (0.6007)	0.1518 (0.0078)	0.0737 (0.0476)	0.0454 (0.1108)	0.0272 (0.1404)
eastern (53) vs. whitebelly (27)	0.0007 (0.2867)	0.0181 (0.0869)	80 (0.2867)	0.0312 (0.0480)	0.0038 (0.3457)	0.0044 (0.2617)	-0.0064 (0.5495)	0.0267 (0.0250)	0.0271 (0.0410)	0.0015 (0.3247)	0.0092 (0.1868)	0.0270 (0.0380)	0.0240 (0.0819)	0.0063 (0.2458)	0.0658 (0.0010)	0.0229 (0.1089)	0.0205 (0.0799)	0.0025 (0.2917)
Tres Marias (21) vs. Central Amer. (9)	0.0155 (0.0914)	-0.0343 (0.7391)	30 (0.3125)	0.0112 (0.3217)	-0.0270 (0.6267)	-0.0480 (0.7896)	0.0076 (0.3163)	-0.0394 (0.8802)	-0.0295 (0.6973)	-0.0303 (0.6749)	-0.0253 (0.6012)	0.0102 (0.3301)	-0.0129 (0.4843)	-0.0564 (0.9426)	0.0149 (0.3229)	-0.0524 (0.8344)	-0.0638 (0.9600)	-0.0385 (0.7832)
Tres Marias (21) vs. Eastern (32)	0.0009 (0.4107)	-0.0118 (0.7121)	51 (0.7718)	-0.0102 (0.6217)	-0.0203 (0.8386)	-0.0095 (0.5209)	0.0026 (0.2689)	-0.0123 (0.7604)	-0.0218 (0.9292)	-0.0148 (0.8092)	-0.0078 (0.5601)	-0.0158 (0.8196)	-0.0141 (0.7776)	-0.0076 (0.5355)	-0.0215 (0.9092)	-0.0196 (0.7992)	-0.0049 (0.4947)	-0.0125 (0.7236)
Tres Marias (21) vs. whitebelly (27)	0.0024 (0.1934)	0.0262 (0.0839)	48 (0.4935)	0.0419 (0.0639)	-0.0037 (0.5015)	0.0079 (0.2591)	0.0173 (0.2457)	0.0326 (0.0468)	0.0413 (0.0358)	-0.0005 (0.3859)	0.0294 (0.0982)	0.0406 (0.0296)	0.0332 (0.0766)	0.0191 (0.1417)	0.0729 (0.0056)	0.0239 (0.1069)	0.0485 (0.0410)	0.01180 (0.1933)
Hawaii (15) vs, whitebelly (27)	0.0456 (0.0001)	0.1811 (0.0001)	42 (0.0004)	0.0253 (0.1548)	0.4554 (0.0010)	0.0532 (0.0190)	-0.0127 (0.9121)	0.2016 (0.0010)	-0.0102 (0.6374)	0.0029 (0.3716)	0.0127 (0.2368)	-0.1757 (1.0000)	-0.0244 (0.8561)	0.2661 (0.0010)	0.0011 (0.4755)	0.1138 (0.0010)	-0.0331 (1.0000)	0.0792 (0.0040)
Hawaii (15) vs. eastern (53)	0.0449 (0.0001)	0.1721 (0.0001)	68 (0.0356)	0.0448 (0.0649)	0.4112 (0.0010)	0.0513 (0.0420)	0.0001 (0.3986)	0.2077 (0.0010)	0.0033 (0.3716)	0.0212 (0.1359)	-0.0142 (0.6913)	-0.0460 (0.9471)	-0.0070 (0.5764)	0.2311 (0.0010)	0.0654 (0.0160)	0.1093 (0.0060)	0.0178 (0.1748)	0.0592 (0.0170)
Hawaii (15) vs. Central Amer. (9)	0.0636 (0.0219)	0.3150 (0.0001)	24 (0.0038)	-0.0086 (0.4435)	0.5781 (0.0002)	0.1757 (0.0109)	0.1600 (0.0467)	0.3982 (0.0002)	0.2689 (0.0002)	0.1421 (0.0176)	0.0498 (0.1727)	0.0399 (0.2711)	0.1262 (0.0450)	0.3858 (0.0002)	0.23708 (0.0002)	0.2276 (0.0002)	0.1615 (0.0404)	0.1882 (0.0030)
Hawaii (15) vs. Tres Marias (21)	0.0487 (0.0004)	0.2124 (0.0002)	36 (0.0004)	0.0794 (0.0496)	0.4823 (0.0002)	0.1158 (0.0249)	0.0341 (0.2325)	0.25733 (0.0002)	0.2374 (0.0002)	0.0701 (0.0286)	0.0732 (0.0288)	0.0044 (0.4021)	0.1528 (0.0002)	0.2735 (0.0002)	0.0943 (0.0079)	0.1332 (0.0044)	0.0650 (0.0626)	0.1511 (0.0002)

Table 3. Pairwise divergence estimates for spinner and spotted dolphin subspecies using all nuclear SNPs, respectively, and using only neutral SNPs. * G_{ST} measures for all spinner dolphin tests were <0 and therefore should be treated as non-significant structure.

Spinner dolphins			**All 51 SNPs**			**42 Neutral SNPs**	
Taxon 1	**Taxon 2**	F_{ST}	G_{ST}	Jost's *D*	F_{ST}	G_{ST}	Jost's *D*
n:female/male/unk	n:female/male/unk	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Central Amer.	eastern	-0.0023	-0.0411*	0.00003	-0.0066	-0.0423*	0.000001
7:3/3/1	28:15/7:6	(0.4485)	(0.3306)	(0.4135)	(0.5148)	(0.3966)	(0.5039)
Central Amer.	whitebelly	0.0148	-0.0367*	0.0006	0.0082	-0.0350*	0.0002
7:3/3/1	21:12/9/0	(0.2607)	(0.2377)	(0.0829)	(0.3216)	(0.2752)	(0.1647)
eastern	whitebelly	0.0297	-0.004*	0.0017	0.0282	-0.0039*	0.0016
28:15/7:6	21:12/9/0	(0.0059)	(0.0069)	(0.0029)	(0.0099)	(0.0085)	(0.0042)
eastern + Cent. Amer.	whitebelly	0.0289	-0.0026*	0.0014	0.0254	-0.0036*	0.0011
35:18/10/7	21:12/9/0	(0.0059)	(0.0049)	(0.0019)	(0.0082)	(0.0057)	(0.0095)
Spotted dolphins			**All 36 SNPs**			**25 Neutral SNPs**	
Offshore	Coastal	0 .1711	0.0592	0.0409	0.1493	0.0469	0.0523
13:6/6/1	12:5/7/0	(0.001)	(0.001)	(0.001)	(0.0005)	(0.0005)	(0.0002)

Table 4: Pairwise divergence estimates (p-values in parentheses). for subspecies and stocks of spotted dolphins using concatenated mitogenome data (F_{ST}, Φ_{ST} and χ^2) and partitioned mitogenomic data (Φ_{ST} only). n_H is the number of haplotypes. "*NA*": Φ_{ST} could not be estimated because all individuals in both strata share the same haplotype.

	Concatenated mitogenome			Partitioned mitogenome Φ_{ST} (p-value)													
Taxon 1 (n) vs. Taxon 2 (n)	F_{ST} (p-value)	Φ_{ST} (p-value)	χ^2 (p-value)	12s n_H=6	16s n_H=7	ATP6 n_H=20	ATP8 n_H=5	COI n_H=21	COII n_H=13	COIII n_H=11	CYTB n_H=20	CR n_H=20	ND1 n_H=15	ND2 n_H=17	ND3 n_H=10	ND4 n_H=23	ND4L n_H=2	ND5 n_H=29
Coastal (24) vs. Offshore (52)	0.0125 (0.0402)	-0.0232 (0.9106)	63.4010 (0.0719)	0.0093 (0.3413)	0.0602 (0.1712)	0.0058 (0.3427)	-0.0139 (0.7742)	-0.0479 (0.9700)	0.0304 (0.1592)	-0.0920 (0.9092)	-0.0349 (0.9408)	-0.0015 (0.4849)	-0.0349 (0.8208)	-0.0704 (0.8738)	-0.0159 (0.6735)	-0.0851 (0.9888)	-0.0164 (0.6051)	-0.0072 (0.4537)
Northeastern (25) vs. western-southern (17)	0.0045 (0.2099)	-0.0083 (0.4119)	39.9247 (0.1932)	0.1052 (0.0718)	0.0271 (0.2737)	0.0191 (0.2891)	0.0272 (0.1781)	-0.0301 (0.7676)	0.0093 (0.3537)	-0.0433 (0.6888)	-0.0606 (0.9802)	-0.0499 (0.9082)	0.0100 (0.4851)	-0.0286 (0.6750)	-0.0234 (0.6720)	-0.0552 (0.9198)	-0.0163 (1.0000)	-0.0537 (0.8748)
Coastal (24) vs. northeastern (25)	0.0302 (0.0002)	-0.0129 (0.5121)	49 (0.0002)	0.1321 (0.0596)	0.1168 (0.0994)	0.0228 (0.2817)	-0.0254 (1.000)	-0.0387 (0.8784)	0.0651 (0.0939)	-0.0644 (0.8004)	-0.05705 (0.9742)	-0.0177 (0.7140)	-0.0195 (0.7174)	-0.0907 (0.8886)	0.0013 (0.3378)	-0.0083 (0.5109)	-0.0256 (0.6044)	-0.0384 (0.8574)
Coastal (24) vs. western-southern (17)	0.0144 (0.0884)	-0.0528 (0.9878)	37.7639 (0.0152)	-0.0580 (0.8298)	0.3130 (0.0428)	0.0060 (0.3773)	0.0393 (0.1592)	-0.0466 (0.8708)	0.0073 (0.4101)	-0.1892 (0.9264)	-0.0488 (0.8780)	0.0034 (0.4063)	-0.1050 (0.9084)	0.0107 (0.4575)	-0.0334 (0.6852)	-0.1927 (0.9978)	0.0230 (0.5003)	0.0050 (0.3403)
Offshore southern (9) vs. offshore western (8)	0.0771 (0.2249)	0.2803 (0.0993)	17.8978 (0.1749)	-0.0053 (0.4089)	0.1588 (0.0776)	-0.0997 (0.7152)	0.15791 (0.0842)	0.1830 (0.0666)	0.3540 (0.0430)	0.6299 (0.0212)	0.3070 (0.0348)	0.0711 (0.2125)	0.1544 (0.2695)	0.4741 (0.0602)	0.1677 (0.1213)	0.2536 (0.0368)	*NA*	-0.5771 (0.9792)
Northeastern (25) vs. offshore western (8)	0.0027 (0.4291)	-0.0408 (0.7425)	31.4911 (0.2761)	0.1873 (0.1073)	0.1473 (0.0955)	0.05838 (0.2619)	0.1779 (0.0320)	0.0423 (0.2085)	0.1973 (0.0278)	0.2883 (0.5998)	0.1584 (0.0736)	-0.0889 (0.9252)	0.0236 (0.4041)	0.0743 (0.2157)	0.0478 (0.2099)	0.1717 (0.0456)	-0.0800 (1.0000)	-0.3710 (0.9946)
Northeastern (25) vs. offshore southern (9)	0.0073 (0.3755)	0.1115 (0.0568)	33 (0.1284)	0.0541 (0.2765)	-0.1235 (0.9746)	-0.0406 (0.7216)	-0.0438 (1.0000)	0.0183 (0.3071)	0.1629 (0.0564)	0.2084 (0.0734)	0.0087 (0.2923)	-0.0328 (0.7328)	0.0750 (0.1481)	-0.0061 (0.4567)	-0.0385 (0.7644)	-0.0323 (0.6394)	-0.0514 (1.0000)	-0.0192 (0.4751)
Coastal (24) vs. offshore southern (9)	0.0255 (0.0762)	0.0689 (0.3049)	33 (0.0054)	-0.0609 (0.5697)	-0.0500 (1.0000)	0.0055 (0.4287)	-0.0180 (0.5875)	0.0863 (0.1237)	0.0589 (0.1431)	0.1499 (0.1648)	0.0436 (0.2245)	-0.0220 (0.6109)	0.0246 (0.5173)	0.1800 (0.2261)	0.0539 (0.1981)	-0.0988 (0.9234)	-0.0179 (0.5913)	-0.0783 (0.8456)
Coastal (24) vs. offshore western (8)	0.0049 (0.4321)	0.0642 (0.1604)	27.7333 (0.0889)	0.0541 (0.4279)	0.5119 (0.0081)	-0.0396 (0.5753)	0.1877 (0.0167)	0.0374 (0.2341)	0.3950 (0.0056)	0.3947 (0.0339)	0.1732 (0.0808)	0.0793 (0.1551)	-0.0491 (0.7138)	0.2491 (0.0624)	-0.0317 (0.4841)	0.1223 (0.1119)	-0.0272 (1.0000)	-0.1810 (0.9314)
Hawaii (5) vs. Coastal (24)	0.1430 (0.0026)	0.3495 (0.0104)	29 (0.0050)	-0.1067 (1.0000)	-0.1063 (1.0000)	0.4345 (0.0504)	-0.0749 (1.0000)	0.2423 (0.0844)	-0.1236 (0.9942)	0.0206 (0.4911)	-0.0750 (0.5925)	-0.4798 (0.9898)	0.6544 (0.0002)	0.3158 (0.2507)	0.5466 (0.0008)	-0.1210 (0.6367)	-0.0749 (1.0000)	0.02890 (0.2591)
Hawaii (5) vs. Offshore (47)	0.1181 (0.0006)	0.1776 (0.0470)	57 (0.1602)	-0.1524 (0.9944)	-0.2156 (0.9690)	0.1281 (0.2214)	-0.0855 (1.0000)	-0.0345 (0.5383)	-0.2203 (0.9896)	-0.0376 (0.6157)	-0.0265 (0.4689)	-0.4198 (0.9968)	0.5631 (0.0002)	0.1325 (0.2873)	0.4588 (0.0002)	-0.5012 (1.0000)	-0.0853 (1.0000)	-0.00296 (0.4705)
Hawaii (5) vs. northeastern (25)	0.0576 (0.2709)	-0.1877 (0.9960)	25.714 (0.7838)	-0.0604 (0.6922)	-0.1765 (0.9524)	(0.0376 (0.3461)	0.0999 (1.0000)	-0.0364 (0.5471)	-0.2393 (0.9808)	0.0128 (0.5155)	0.0004 (0.4329)	-0.4048 (0.9588)	0.5310 (0.0002)	0.1268 (0.2693)	0.4512 (0.0004)	-0.4357 (0.9958)	-0.1070 (1.0000)	-0.1084 (0.8278)
Hawaii (5) vs. western-southern (17)	0.2474 (0.0702)	-0.1308 (0.6580)	44.798 (0.1165)	-0.1904 (0.9926)	-0.0550 (0.7144)	0.0773 (0.2423)	-0.0553 (0.1782)	-0.0489 (0.5935)	-0.2415 (0.9948)	-0.0829 (0.6009)	-0.0975 (0.7388)	-0.1119 (0.6742)	0.5411 (0.0002)	0.2736 (0.2781)	0.5110 (0.0009)	-0.7780 (1.0000)	*NA*	0.0454 (0.4033)
Hawaii (5) vs. offshore western (8)	0.4958 (0.0732)	0.4958 (0.0732)	30.781 (0.0732)	-0.0687 (1.0000)	0.0588 (0.3163)	0.4164 (0.2563)	0.0580 (0.3259)	0.2021 (0.1024)	0.0381 (0.4967)	0.5340 (0.1297)	0.4190 (0.0579)	0.1727 (0.2931)	0.5990 (0.0014)	0.5264 (0.1333)	0.6609 (0.0036)	-0.4457 (0.9444)	*NA*	0.5328 (0.0955)
Hawaii (5) vs.	0.1569	-0.0014	3.1111	-0.2350	*NA*	-0.1244	*NA*	-0.2031	-0.0331	-0.1580	-0.4278	-0.2147	0.5088	0.1551	0.4811	-0.9455	*NA*	-0.2735

Table 5. Summary table of pairwise comparisons using both mtDNA and nuDNA data sets (sample sizes in parentheses). A "✓"denotes significance detected based on at least one statistical measure (see Tables 3-5). 'ns' = non-significant. '*NA*' = not tested due to insufficient nuclear SNP data. " ~ " = either not significant, but indicating structure with a low p-value, or significant at only one gene.

Spinner dolphins	**Taxon 1 (n_{mt}/n_{nuc})**	**Taxon 2 (n_{mt}/n_{nuc})**	**mtDNA**	**nuDNA**
Test of endemic subspecies	Central American (9/13)	eastern (53/28)	✓	ns
Testing whitebelly intergrade	Central American (9/13)	whitebelly (27/20)	✓	ns
Testing whitebelly intergrade	eastern (54/28)	whitebelly (27/20)	ns	✓
Testing alternate	Tres Marias (21)	Eastern (32)	ns	*NA*
""	Tres Marias (21)	Central American (9)	ns	*NA*
""	Tres Marias (21)	Whitebelly (27)	~	*NA*
""	Hawaii (15)	Central American (9)	✓	*NA*
""	Hawaii (15)	Eastern (32)	✓	*NA*
""	Hawaii (15)	Whitebelly (27)	✓	*NA*

Table 5 (cont.). Summary table of pairwise comparisons using both mtDNA and nuDNA data sets (sample sizes in parentheses). A "✓"denotes significance detected based on at least one statistical measure (see Tables 3-5). 'ns' = non-significant. '*NA*' = not tested due to insufficient nuclear SNP data. " ~ " = either not significant, but indicating structure with a low p-value, or significant at only one gene.

Spotted dolphins	**Taxon 1 (n_{mt}/n_{nuc})**	**Taxon 2 (n_{mt}/n_{nuc})**	**mtDNA**	**nuDNA**
Testing subspecies	Offshore (59/13)	Coastal (24/12)	✓	✓
Testing existing stocks	Offshore northeastern (25)	Off. western-southern (29)	~	*NA*
Testing existing stocks	Offshore northeastern (25)	Coastal (24)	✓	*NA*
Testing existing stocks	Offshore western-southern (29)	Coastal (24)	✓	*NA*
Testing alternate	Offshore southern (21)	Coastal (24)	✓	*NA*
""	Offshore western (8)	Coastal (24)	✓	*NA*
""	Offshore northeastern (25)	Offshore western (8)	~	*NA*
""	Offshore northeastern (25)	Offshore southern (9)	~	*NA*
""	Hawaii (5)	Coastal (24)	✓	*NA*
""	Hawaii (5)	Offshore (52)	✓	*NA*
""	Hawaii (5)	Offshore northeastern	✓	*NA*
""	Hawaii (5)	Offshore western (8)	~	*NA*
""	Hawaii (5)	Offshore southern (9)	~	*NA*

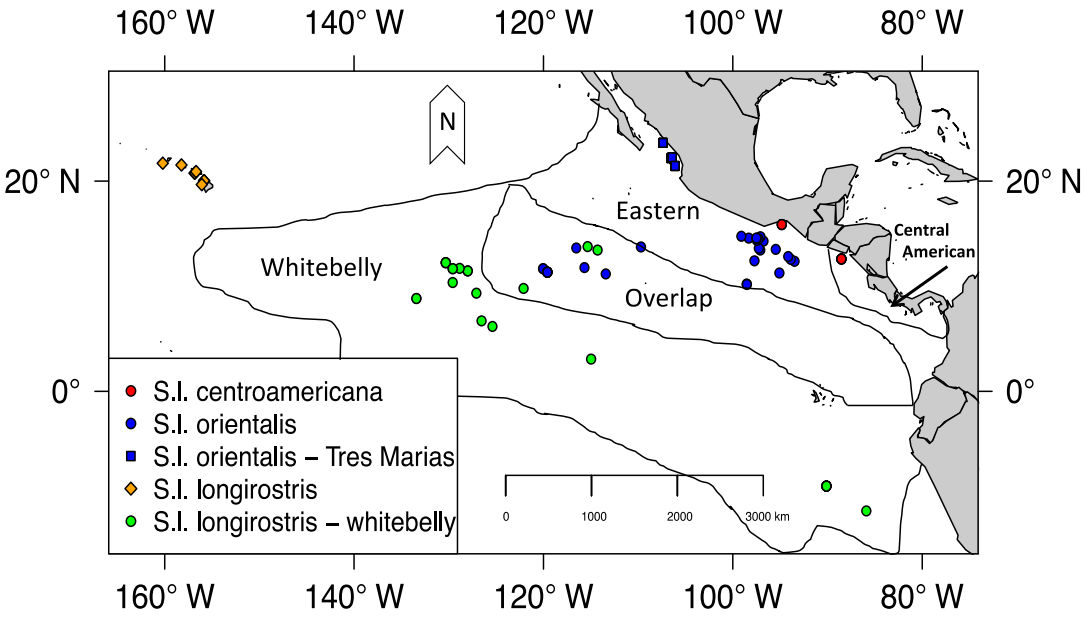


Figure 1. Sampling localities and range map for spinner dolphins within the ETP. Subspecies and stock boundaries based on Perrin *et al*. 1985. Red dots indicate Central American spinners. Blue symbols indicate eastern spinners - boxes are the proposed Tres Marias form. Green dots indicate whitebelly spinners, a proposed intergrade between the pantropical (orange diamonds) and the eastern subspecies (blue dots).

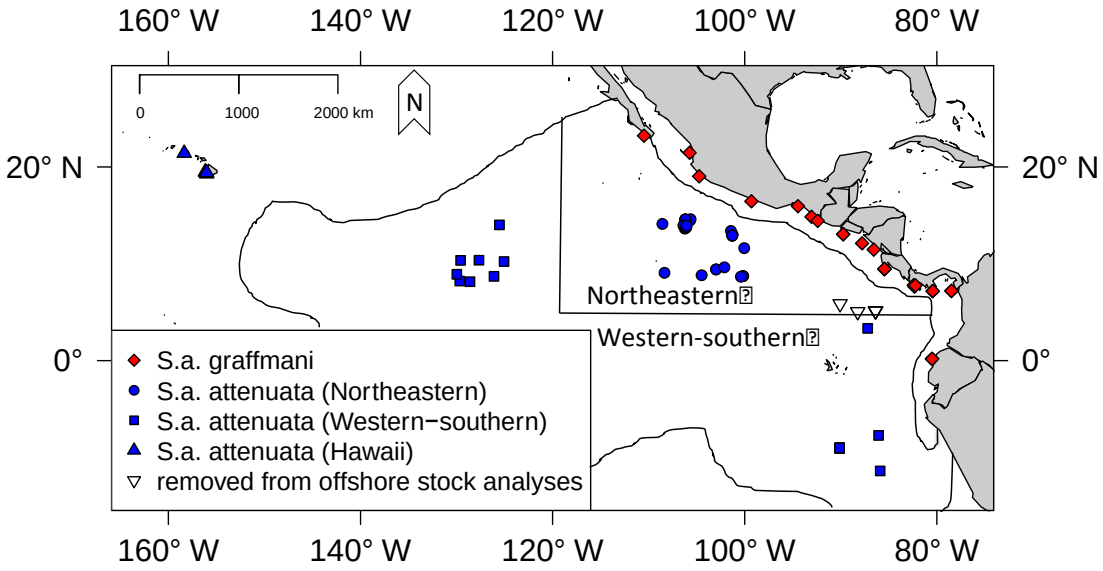


Figure 2. Sampling localities for spotted dolphins with ETP subspecies and stock boundaries based on Perrin *et al.* 1985. Coastal spotted dolphins (*S. a. graffmani*) are red and offshore (*S a. attenuata*) are in blue. Blue circles indicate sampling locations for the northeastern stock of offshore spotted dolphins. Blue triangles indicate samples from Hawaii. Inverted triangles indicate southern offshore samples that were removed from analyses of offshore stocks because they were collected north of between 4°N and 6°N; these samples were included in subspecies-level analyses.

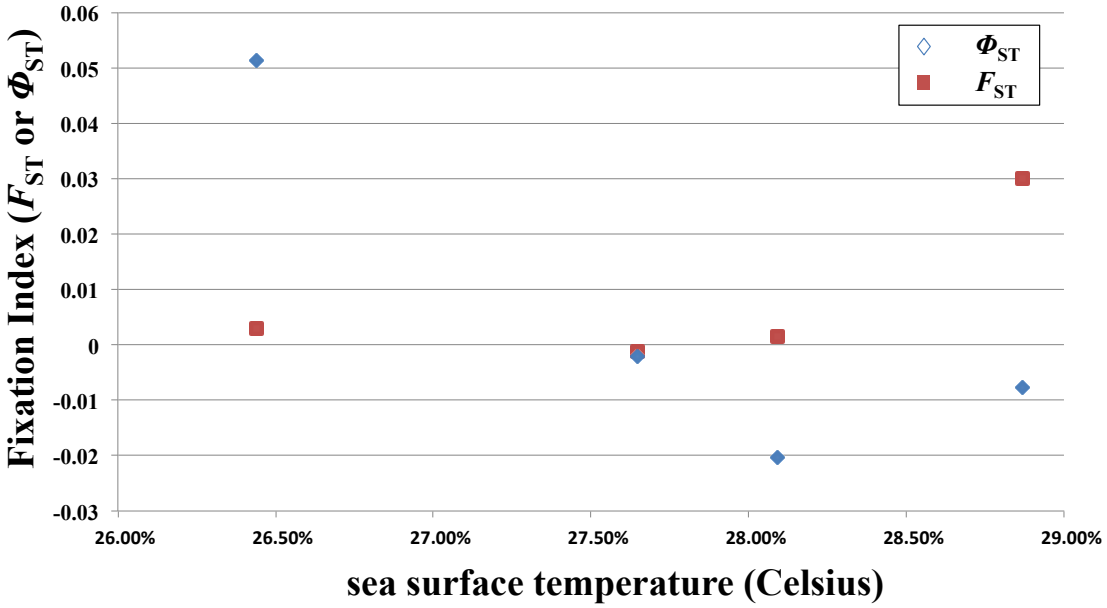


Figure 3. Pairwise mtGenome F_{ST} Red squares and Φ_{ST} blue diamonds estimated for offshore spotted dolphins when partitioning by varying sea surface temperature estimated for collection time and location. The F_{ST} for the comparison of partitions created by splitting the pooled samples at 28.87°C was the only estimate significantly different from zero p = 0.0129.

CHAPTER 3:

The problem of high diversity and low divergence: inferring population structure of eastern tropical Pacific spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) from genome-wide SNPs

Matthew S. Leslie [1,2]

1. Scripps Institution of Oceanography, University of California San Diego, mail-code 0208, 9500 Gilman Dr. La Jolla, CA, 92093 USA

Phillip A.Morin [2,1]

2. Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 8901 La Jolla Shores Dr., La Jolla, CA 92037, USA

ABSTRACT:

Millions of spinner (*Stenella longirostris*) and spotted dolphins (*Stenella attenuata*) died as bycatch in tuna nets in the eastern tropical Pacific Ocean. Despite three decades of protection, they show little-to-no sign of recovery. In efforts to establish biologically meaningful management boundaries for recovery, endemic subspecies and multiple stocks have been proposed. Genetic differentiation among most of these units has been difficult to demonstrate, however, possibly due to low statistical power, large historical abundances, ongoing gene flow between subspecies and recent divergence. We tested for genetic structure at multiple hierarchical levels by analyzing the largest dataset to date brought to bear on these questions. Single nucleotide polymorphisms (SNPs) were collected from nuclear DNA regions associated with the restriction enzyme site *PstI* from 72 spinner dolphins and 58 spotted dolphins using genotype-by-sequencing (GBS) from highly parallel DNA sequencing. Our results support the current subspecies for both species and indicate stock-level separation for Tres Marias spinner dolphins and the two offshore spotted dolphin stocks in this area. These results are critically important for the ongoing management and recovery of these highly-impacted pelagic dolphins in the eastern tropical Pacific Ocean.

INTRODUCTION:

Knowledge of population structure forms the foundation of assessment and monitoring for recovery; we need to know population boundaries in order to estimate abundance and trends and to set appropriate mortality limits. Without knowledge of how organisms are geographically structured we may fail to conserve subspecies or distinct population segments.

The study of population structure in pelagic organisms is inherently challenging. Open ocean habitat may have few permanent physical barriers to gene flow thereby increasing the likelihood of mixing between populations; even a few successful migrants per generation greatly decrease the signal of population segregation (Waples 1998). Pelagic species with large abundance and weak barriers provide a challenge for using genetic tools to determine population structure (Taylor and Dizon 1996, Waples 1998). Large abundance, often a characteristic of pelagic populations, can result in a large amount of standing genetic variation. High genetic diversity can dramatically increases the time needed for populations to drift apart genetically, even in the complete absence of gene flow (Taylor and Dizon 1996).

Spinner and spotted dolphins in the eastern tropical Pacific Ocean (ETP) illustrate this problem. They historically numbered several million (Wade et al. 2007), but starting in the 1960s, hundreds of thousands were killed annually as bycatch in the dolphin-set

tuna purse-seine fishery (Lo and Smith 1986, National Research Council 1992, Wade 1995). Previous genetic studies have shown that these species retain high genetic variation (Dizon et al. 1994, Galver 2002, Escorza-Trevino et al. 2005, Andrews et al. 2013), which possibly obscures the detection of population genetic structure that could be used to improve abundance trend estimates and recovery efforts.

Despite over forty years of protection under the U.S. Marine Mammal Protection Act of 1972 and specific multi-national protection under the 1999 Agreement on the International Dolphin Conservation Program (Joseph 1994, Gosliner 1999), they have not recovered as predicted (Wade et al. 2007, Gerrodette et al. 2008). These species exhibit morphometric differences supporting structured populations (Perrin et al. 1991, Perrin et al. 1994). Traditional molecular genetics approaches have not found corresponding population genetic structure (Dizon *et al.* 1994, Galver 2002), although Andrews et al. (2013) found evidence for segregation in data from the Y-chromosome and Escorza-Treviño et al. (2005) found population structure within the coastal subspecies of spotted dolphin. Additional genetic markers can increase statistical power to detect divergence when signals are faint – such as low divergence, high abundance populations with ongoing geneflow. Willing et al. (2012) presented convincing evidence that estimations of genetic differentiation (*F*-statistics) can very powerful with a high number of genetic markers (> 1000) - even if samples sizes are small (4-6). Here we use high-throughput DNA sequencing to collect thousands of SNPs from throughout the genome to test hypotheses of genetic structure between highly abundant and low divergence dolphin

populations of special conservation concern.

Endemic spinner dolphin subspecies and stocks

There are four recognized subspecies of spinner dolphin (*Stenella longirostris*) globally. The pantropical "Gray's" spinner (*S. l. longirostris*) occurs throughout the world's tropics except in the ETP. In the central and western Pacific, Gray's spinners are found associated with islands, such as the Hawaiian Islands. The much smaller dwarf spinner (*S. l. rosiventris*) subspecies inhabits shallow waters in Southeast Asia (Perrin et al. 1989, 1999). The other two subspecies, the eastern spinner dolphin (*Stenella longirostris orientalis*) and Central American spinner dolphin (*S. l. centroamericana*), are found only in the ETP and were described based on differences in external coloration, body size, and skull morphology (Perrin et al. 1991, Douglas et al. 1992). The Central American subspecies inhabits relatively near-shore waters off the Pacific coasts of Southern Mexico south through Panama (Figure 1). A distinct morphotype of the eastern spinner dolphin (*S. l. orientalis*), known as the "Tres Marias" spinner dolphin, has been described from near the islands of the same name off the coast of Mexico based on external body morphometrics (Perryman and Westlake 1998). Although currently classified as a form of the pantropical Gray's spinner dolphin, the "whitebelly" spinner is proposed to represent a hybrid swarm between the eastern subspecies and the pantropical Gray's subspecies of the central and western Pacific (Perrin et al. 1991, Andrews et al. 2013). Significant geographic overlap exists between the eastern subspecies and the whitebelly form (Perrin et al. 1985).

The eastern spinner dolphin (*S. l. orientalis*) exhibit traits indicative of a polygynous breeding system (Perrin and Henderson 1979, Perrin and Mesnick 2003). Perrin and Mesnick (Perrin and Mesnick 2003) found significant difference in testes size between the eastern subspecies and the whitebelly form, indicating differing reproductive strategies and probably different breeding behavior between the two types. The three ETP spinner management stocks correspond to the two aforementioned subspecies plus the whitebelly spinners. Andrews *et al.* (2013) estimated high levels of geneflow between subspecies in the ETP using autosomal and mitochondrial genes and found a shared Y-chromosome haplotype in the eastern and Central American subspecies that was not found in Gray's or dwarf subspecies. Interestingly, this locus was found to be polymorphic in whitebelly spinners, supporting the hypothesis of introgression in this form (Andrews *et al.* 2013). The authors proposed that sexual selection was driving the divergence of spinner dolphins in the ETP.

Endemic ETP spotted dolphin subspecies and stocks

Extensive analyses of cranial morphology of pantropical spotted dolphins (*Stenella attenuata*) in the ETP led to the designation of a coastal endemic subspecies (*S. a. graffmani* - Perrin 1975; Perrin *et al.* 1987). Spotted dolphins in offshore regions in the ETP are part of the pantropical subspecies (*S. a. attenuata*). In contrast to the spinner dolphins, genetic analyses of microsatellites has indicated some differentiation between

subspecies (Escorza-Treviño *et al.* 2005). That study found differences between southern populations of the coastal subspecies and the pelagic subspecies, but no differences between the northern extents of the coastal and the pelagic forms. The authors concluded that genetic diversity in the coastal subspecies is contained within demographically independent populations. Tests for genetic structure have not previously been conducted within the pelagic subspecies.

Despite the results of Escorza-Treviño *et al.* (2005) showing differences between at least four demographically independent populations within the coastal subspecies (*S. a. graffmani*), the entire subspecies is currently treated as a single management stock. The offshore pantropical spotted dolphins are divided into a 'northeastern' (NE) stock, defined geographically as north of 5°N, east of 120°W and a 'western-southern' (WS) stock residing south and west of this northeastern area (Figure 2) (Perrin *et al.* 1994). The north-south boundary between these stocks is based on a distributional hiatus. In addition, general similarity in cranial variation suggests a combined WS stock (see Perrin *et al.* 1994 for discussion and primary references).

Leslie *et al.* (*In Review*) tested for structure at multiple hierarchical levels by collecting whole mitochondrial genome sequences (mtDNA) and nuclear SNPs (nuDNA) from 104 spinner and 76 spotted dolphins using capture array library enrichment and highly paralleled DNA sequencing. MtDNA showed weak but significant differences

between the subspecies of spotted dolphins (F_{ST}: 0.0125; p = 0.0402) and the first mtDNA evidence for differentiation between the ETP spinner dolphin subspecies (F_{ST}: 0.0133; p = 0.034). NuDNA supported subspecies of spotted dolphins, but not subspecies of spinner dolphins. Strong and significant differentiation was detected between whitebelly and eastern spinner stocks using nuDNA (F_{ST}: 0.0297; p = 0.0059). Neither mtDNA nor nuDNA supported the division of existing offshore stocks of spotted dolphins or Tres Marias spinner dolphins. This work helped identify a genetic basis for establishing biologically meaningful management units for the recovery of these dolphins, but as with previous genetic studies that used relatively few molecular markers, it lacked the power to resolve some populations and putative subspecies.

Objectives

In this study we use >7,000 restriction-site associated nuclear SNP genotypes from each species to infer population genetic structure by testing *a priori* hypotheses at the subspecies and population level and by principle components-like analyses. For the population level questions, we test for genetic support of the Tres Marias spinner dolphin and current and alternative stock boundaries in the offshore spotted dolphins.

MATERIALS AND METHODS

Sample Collection and DNA extraction

Skin samples were collected from spotted dolphins and spinner dolphins via biopsy dart (Lambertsen 1987) on research cruises or from soft tissue specimens taken as bycatch in the tuna purse-seine fishery. Spinner dolphin samples collected on research cruises were assigned to a stock based on the external morphology of the majority of animals in the school. This approach was taken because: 1) these often-large groups (sometimes >1000 individuals) contained individuals exhibiting a range of morphology due at least in part to ontological variation but also to morphological overlap; only after observing the group for some time could observers classify it to stock, 2) researchers collecting biopsies from dolphins near the bow of the research vessel found it very difficult to confidently classify fast-swimming individuals at sea, and 3) there is significant overlap in range; therefore, geography was not a reliable predictor of stock identity. Samples were selected from areas where the eastern and whitebelly types are known to overlap (Figure 1).

Spotted dolphins samples (Figure 2) were assigned to subspecies and stocks based on the geographic location of the sampling site. In areas where the ranges of the two subspecies overlap, spotted dolphin samples collected from research cruises were assigned to stocks based on external morphology (described in Perrin and Hohn 1994 and Perrin 2001).

Samples were stored in salt-saturated 20% DMSO, 70% ethanol, or frozen with

no preservative. DNA was extracted using silica-based filter membranes (Qiagen, Valencia, CA) on an automated workstation (Perkin Elmer, Waltham, MA) or by NaCl salting out based on (Miller et al. 1988). Starting concentrations of DNA were quantified using Pico-Green fluorescence assays (Quant-it Kit, Invitrogen, Carlsbad, CA) using a Tecan Genios microplate reader (Tecan Group Ltd, Switzerland). DNA quality was assessed by electrophoresis of 100ng on 1% agarose gel; only high-molecular weight extracts were used.

Sequencing libraries were constructed using the restriction enzyme *PstI* (CTGCAG) using a "genotype-by-sequencing" protocol (Elshire et al. 2011). Unique oligonucleotide barcodes were added to each sample for multiplexed sequencing on the Illumina *HiSeq* 2000/2500 (100bp, single-end reads). One no-template-control was included with every batch of 95 samples. Libraries for several individuals were reconstructed and resequenced to increase read depth. Library preparation and sequencing were completed at the Cornell University Institute of Biotechnology's Gneomic Diversity Facility (http://www.biotech.cornell.edu/brc/genomic-diversity-facility).

Filtering, Assembly, SNP Discovery and Genotyping

Demultiplexing, initial quality control, assembly, and SNP discovery were completed in the *TASSEL* pipeline (v3.0.166)(Glaubitz et al. 2014). Genotypes were

filtered to those with a quality of 98 or higher using *vcftools* (version 0.1.12b) (Danecek *et al.* 2011). No-template controls were also removed in *vcftools*. Reads were assembled to the bottlenose dolphin genome (*T. truncatus*, turTru1, 2.59X coverage; Lindblad-Toh et al. 2011) using *bwa* (v0.78-r455) (Li and Durbin 2009). Identically named samples were merged after VCF SNP calling with *TASSEL*. All data for samples were removed if they did not pass a minimum "missingness" threshold (0.1 minimum taxon coverage) prior to filtering (using *TASSEL*), or after filtering using *vcftools*. We initially accepted genotypes with a minimum read depth of 5, minimum genotype quality of 15. Sites were filtered by minor allele frequency (between 0.05 and 0.5), genotype missingness (minimum 0.95), and coverage (removed if the mean genotype depth of coverage <8 or >1.5 times the mean depth over all individuals). VCF-format data were converted to various formats using *PGDSpider2* v2.0.8.3 (Lischer and Excoffier 2012). SNP loci were removed if heterozogosity was greater than 0.65. Sites were not filtered based on Hardy-Weinburg equilibrium, nor were sex-linked loci or F_{ST} outlier loci removed. Mitochondrial DNA SNPs were removed by aligning to the nuclear *Tursiops* genome.

Diversity Estimates and Populations Structure Analyses

Per-site mean heterozygosity and standard deviation were measured in *R* (R Core Team 2011), using the *strataG* package (Archer 2015). We then estimated differentiation (F_{ST}) for each pairwise combination of populations (Wright 1951). Point estimates and permutation tests (1000 repetitions) were generated using the *strataG* package in *R* (Archer 2015). Differentiation statistics that correct for high allelic diversity (Nei 1973)

and high heterozygosity within polymorphic sites (Hedrick and Goodnight 2005) were not used because we used a large number of biallelic SNPs (see Meirmans and Hedrick 2011).

Multivariate Analysis of Population Differentiation

We directly tested hypotheses of population differentiation using multivariate analyses, specifically the Discriminant Analysis of Principle Components (DAPC) as implemented in the R packages *Adegenet.* This method calculates principle components and then estimates a centroid and measures the variance for each predefined population. The discriminant analysis then tests the probability of each individual falling in the space of each of the populations based on the "geometric space" created by the position of the centroid and the extent of the variation.

Outcomes of the DAPC can vary depending on the number of principle components (PCs) included in the analysis. We used three different approaches to investigate this variance. First, using *Adegenet*, we estimated the optimum number of PCs by calculating the mean Alpha-score for a range of different PCs. Alpha is the reassignment probability calculated using the given cluster minus the reassignment probability for randomly permuted clusters. This analysis simulates the dataset 10 times and takes the mean alpha-score for each PC. Values near zero indicate low discrimination and instability whereas values near one indicate higher discrimination and stability. The

optimum is the number of PCs with the highest mean alpha across all simulations. Second, we calculated alpha scores for all putative populations under two different scenarios: 1) using the highest number recommended by the developers (i.e., n / 3), and 2) the "optimum" number of PCs indicated by the highest mean alpha-score estimated using simulations. Third, we also examined the cumulative variance explained by the eigenvalues for the full range of principle components.

To complete the DAPC, we then constructed synthetic discriminant functions that represent linear combinations of the allelic data with the largest between-group variance and the smallest within-group variance. In all analyses we kept only the first three eigenvalues, as they represented the vast majority of the informative.

Two-dimensional scatter plots were constructed in *R* to visualize the spread of the first two discriminant functions between and within populations. Furthermore, we calculated membership probabilities for each individual and pooled these for each pre-assigned cluster – also using *Adegenet* - and made pie charts to display the probability of individuals being members of predefined clusters (e.g., subspecies or stock). This can be thought of as pooled membership probabilities or the sum of all individual membership probabilities within a cluster. It is intended to indicate how well the clusters represent biological reality. All of these calculations were done in *R* using output from the *Adegenet* package.

RESULTS:

The average read depth for spinner and spotted dolphin SNP genotypes was 36.3 and 31.4, respectively. To remove genotypes with high coverage possibly due to clonality, those with coverage depth greater than 64.3 and 58.4 (the mean read depth plus >1.5 times the standard deviation) were removed from the spinner and spotted dolphin datasets, respectively. The final data sets for each species included 72 spinner dolphins and 58 spotted dolphins genotyped at 8,193 and 7,010 SNPs, respectively.

Summaries of the sample sizes, mean heterozygosity, and pairwise population differentiation F_{ST} statistics for ETP spinner and spotted dolphins are found in Tables 1 and 2, respectively. Mean heterozygosity across SNPs for all individuals within a population (*He*) is indicative of the overall genetic variation within the populations. The spinner dolphin dataset had higher overall *He* than the spotted dolphin dataset – and was more variable (i.e., a broader standard deviation). Whitebelly spinners had a slightly lower mean heterozygosity than the other populations. The Tres Marias population had the highest *He* (0.2647), and the Central American population had the broadest standard deviation (0.2059). Northeastern spotted dolphins had higher mean *He* than the other two spotted dolphin populations, while the coastal subspecies had the lowest *He*. Higher *He* could indicate outbred individuals while low number of heterozygotes could indicate

inbred individuals or poor sample quality that resulted in allelic dropout (Davey et al. 2013, Gautier et al. 2013).

Population Structure: F_{ST}

F_{ST} measures the variance in allele frequencies in a population (Wright 1969) to estimate the amount of structure within and among subpopulations. This metric is very informative for inferring population structure. A corrected metric for the deflation of F_{ST} when allelic heterozygosity is high (i.e., F'_{ST}) was not needed (Meirmans and Hedrick 2011).

For both spinner and spotted dolphins, all pairwise comparisons between putative populations were significantly different from zero. All comparisons between spinner dolphin groups were highly significantly different ($p = 0.0009$). Similarly, for spotted dolphins all pairwise comparisons between the coastal subspecies and the two putative offshore stocks were highly significantly different ($p = 0.0009$), and between the offshore stocks was significant at $p = 0.0019$. F_{ST} point estimates for comparisons between Tres Marias spinner and eastern spinner were the lowest. Low values were also found in comparisons between the Central American spinner and the Tres Marias spinner – as well as between the eastern spinner and the whitebelly spinner. We found higher estimates comparisons between whitebelly and Central American and whitebelly and Tres Marias indicate strong segregation between these geographically distant populations. F_{ST} point

estimates for the comparison of the endemic eastern and Central American spinner dolphin subspecies were also higher (F_{ST} = 0.0121).

All pairwise F_{ST} estimates between the spotted dolphins populations (Table 2) were significantly different from zero, including between the two offshore stocks (F_{ST}: 0.0093, p = 0.0019). For comparisons between the coastal spotted dolphin subspecies and the two offshore spotted dolphin stocks, F_{ST} point estimates were slightly higher for coastal versus the Western-southern stock (0.0710) compared to the differences between coastal and the northeastern stock (0.0386).

Describing Clusters with DAPC

Discriminant Analysis of Principle Components (DAPC) aims to describe genetic clusters using synthetic variables derived from the allelic data (Jombart and Ahmed 2011). Before conducting DAPC, however, we needed to determine the appropriate number of principle components (PCs) to use. The mean Alpha for the spinner dolphin data (with a max of 24 PCs) resulted in the optimum being 5 PCs (Figure A2-S1). Spotted dolphin data (with a maximum of 19 PCs) resulted in the optimum being 1 PC (Figure A2-S2). Mean Alpha for the "optimum" PCs was consistently higher – although the variance was also much greater - in the optimum number of PCs compared to the highest recommended number of PCs (Figures A2-S1 and A2-S2). Although the optimum number of PCs maximized mean population stability over the highest recommended

number of PCs, some of the populations had very low alpha scores with the optimized simulation-based method (i.e., eastern spinner dolphins and coastal spotted dolphins - see Table A2-S1). Finally, the cumulative variance explained by the eigenvalues for the full range of principle components indicated that there was no reason for keeping a small number of PCs (see Figures A2-S3 and A2-S4). Because of these results, we chose to present the results obtained using the highest recommended number of PCs (n / 3) recommended by the developers of *Adegenet* for each species: 24 and 19 PCs for spinner and spotted dolphin analysis, respectively. Moreover, given the difficulty of finding genetic structure in previous studies - in what are already morphologically distinct subspecies - we decided to err on the side of over-fitting rather than decreased statistical power.

The majority of the variation within the spinner dolphin data was well represented in the first two DA eigenvalues, as displayed in the inset of Figure 3. The DAPC indicated a very close relationship between the Central American spinner dolphins (blue) and the Tres Marias Islands spinner dolphins (purple). These two overlapping groups were separate from the eastern subspecies (in green) and very separate from the whitebelly spinner dolphins (in red). The eastern subspecies formed a largely independent group with the exception of one sample that grouped within the whitebelly cluster and three near the Tres Marias Islands cluster.

Summed membership probabilities calculated for each stock or subspecies showed strong genetic assignment of individuals to their respective stock and subspecies of origin (Figure 4). Whitebelly individuals (red pie) assigned to the whitebelly stock with very high probability. Similarly, the vast majority of the eastern individuals (green pie) assigned to the eastern stock. One eastern spinner assigned to the whitebelly with 100% probability and several were assigned to the Tres Marias cluster with lesser probability. Tres Marias spinner dolphins (purple pie) assigned to their original populations 70% of the time on whole. Several Tres Marias dolphins, however, assigned to the eastern subspecies and cumulatively made up nearly one quarter of the assignments for the Tres Marias population. Several Tres Marias individuals also assigned to the Central America spinner dolphin subspecies resulting in 20% of the Tres Marias assignments. Central American spinner dolphins (blue) assigned to the Tres Marias population data cloud as often as they did to their own data cloud, highlighting the close relationship of these two groups. Final a small fraction of the summed membership probabilities for the Central American population also assigned to the eastern subspecies.

Similar to the spinner dolphins, the vast majority of the variation within the spotted dolphin dataset was represented by the first of the two DA eigenvalues, as presented in the inset of Figure 5. DAPC for spotted dolphins showed a high level of segregation between the offshore subspecies and the coastal endemic subspecies (Figure 5). The coastal spotted dolphins (blue) is clearly a distinctly clustered. Overall the

northeastern offshore stocks (red) show separation from the western-southern stock (green), but with considerable overlap.

The membership probabilities summed across all individuals for each hypothesized spotted dolphin stocks reflected the separation from the DAPC (Figure 6). Individuals from the coastal subspecies (blue) were assigned back to that population with 100% probability. Offshore spotted dolphins (red and green) were only very rarely assigned to the coastal subspecies (very small blue slice in the red northeastern pie). However, many offshore individuals were assigned to the alternate offshore stock with high probability. Despite the apparent interchange between these two offshore stocks, individuals from the northeastern spotted dolphin stock were assigned back to that stock of origin 75% of the time and western-southern individuals were assigned to the western-southern stock in >85% of the attempts.

DISCUSSION:

Pelagic spinner and spotted dolphins in the eastern tropical Pacific have presented a paradox for management, where subspecies and stocks differ morphologically but heretofore apparently exhibited little genetic differentiation (Dizon et al. 1994, Galver 2002, Leslie et al. *In Review*). To overcome the limitations of using few markers to describe genetic structure in these historically abundant populations, we expanded the genetic analyses to include >7,000 nuclear SNPs per species. Our results show genetic

structure in support of the morphologically designated subspecies of both spotted and spinner dolphins and suggest stock-level differences within offshore spotted dolphins. In addition, we found support for a unique Tres Marias Islands spinner dolphins stock – likely part of the Central American spinner dolphin subspecies.

Spinner Dolphins

Our analyses supported four closely related, but well-differentiated, groups of spinner dolphins in the ETP (whitebelly, eastern, Central American, and Tres Marias). They corroborate the morphometric studies (Perrin *et al.* 1991; Perryman and Westlake, 1998), but contrast with previous genetic studies that found little or no consistent differences among most of these groups (Dizon et al 1994; Galver 2002; Leslie In Review). Our interpretation is that previous datasets have not had the statistical power to discern these closely related, but demographically independent groups with high historical abundance and potentially ongoing geneflow.

Specifically, Dizon et al (1994) found few differences between any of the proposed types based on mtDNA and allozymes. Galver (2002) found no differences between the whitebelly and the eastern spinner dolphins based on mtDNA control region haplotypes and microsatellites, but found support for separating Central American spinners from the eastern and whitebelly groups using these same markers. Andrews *et al.* (2013) found low divergence between all the spinner dolphin subspecies and the

whitebelly spinner dolphins (referred to as "ecotypes" by Andrews *et al.*) based on mtDNA and autosomal genes, but found strong separation in Y-chromosome markers between the three most phenotypically divergent groups (Grey's spinner dolphin, dwarf spinner dolphin, and the ETP group, which includes the eastern and Central American subspecies). This study also found whitebelly spinners had Y-chromosome haplotypes in common with Grey's and eastern supporting the hypothesis that the whitebelly is a hybrid group. Andrews *et al.* discuss several scenarios that could explain the patterns found in the Y-chromosome data, including low N_e, female biased dispersal, assortative mating and hybrid male sterility. Ultimately they hypothesized that the evidence suggests sexual selection on male characters via assortative mating.

Our results support the three groups proposed by Perrin *et al.* (1991) and show evidence for a separate Tres Marias Islands stock as proposed by Perryman and Westlake (1998). Although n the offshore whitebelly spinner and eastern spinners, we observed very little assignment to an alternate group. Some misassignment is expected; because we are investigating forms within a species, not separate species.

The alpha-scores for eastern and Central American spinners were low, indicating a low stability in our DAPC analysis, but the lower scores likely resulted from other sources of error (misidentified individuals and/or migrants), rather than biases of the data. For instance, a male eastern spinner sample (Z2127) assigned to the whitebelly spinner

group with 100% membership probability was sampled in the same group (during the same sampling event) as three other individuals (Z2128 - male, Z2129 - female, Z2132 - male) that assigned to the eastern group based on these data (with almost 100% membership probability). We expect some degree of misassignment given our intraspecific level of inquiry. Furthermore, high spatial overlap between the two types and our limited ability to identify all individuals sampled to stocks during sample collection in the field) rather than biases resulting from the analyses. Finally, these individuals could be migrants across boundaries – an expected result given known levels of geneflow (Andrews et al. 2013).

Two of the three Central American samples (Z38192 and Z38197) that assigned to the Tres Marias cluster were collected off El Salvador, well into the core distribution of Central American spinner dolphins. Z38192 was determined to be a female by molecular assay and the sex of Z38197 is unknown. The third sample (Z11666) – a male - was collected in the Gulf of Tehuantepec on the coast of southern Mexico. Two other samples in our data set were collected during this same sampling event – both of them female – and both assigning to the Central American subspecies. Given these details regarding the genetic assignment between these two groups, but the separation of the groups based on allele frequencies (i.e., F_{ST}) we infer ongoing migration with the possibility of recent dispersal, between the Tres Marias and Central American groups.

Unpublished data from recent aerial surveys suggest that the range of the Central American spinner dolphins likely extends north of Acapulco – beyond the traditional range limit at the Gulf of Tehuantepec (W. Perryman *pers. comm*[3].). Their range could also shift based on oceanographic conditions, food availability, and population abundance (although this remains untested). Interestingly, despite the proximity of samples from the eastern spinner dolphin populations to the Tres Marias cluster, we observed only one sample assigning from the eastern to the Tres Marias and *vice versa* with >20% membership probability. It appears that Tres Marias and the Central American are coastal dolphins that maintain ecological and genetic segregation from conspecific pelagic dolphins despite parapatry.

Perryman and Westlake (1998) compared the Tres Marias spinner dolphins to both the eastern and Central American subspecies in their initial description of this group. Based on body size data collected from aerial photogrammetry, these authors concluded that the spinner dolphins near Tres Marias Island are intermediate in length between the eastern and Central American spinner subspecies. They concluded that the Tres Marias spinners are different enough from the other two subspecies that they may represent a separate stock or subspecies. The authors noted the strong similarities between Tres Marias spinners and eastern spinners in external body characteristics (long, slender body; uniform medium-to-dark gray in color; dorsal fin is triangular to forward-canted; pronounced ventral keels on some individuals), but also noted the large "tightly packed"

[3] Wayne Perryman – *NOAA Southwest Fisheries Science Center* 8901 La Jolla Shores Dr. La Jolla CA 92038

Tres Marias schools are more characteristic of Central American spinners than eastern spinners. Perryman and Westlake (1998) referenced cranial morphometric comparisons presented in Douglas *et al.* (1992) that note similarities in a small set of skulls (n=3) from northern coastal spinner dolphins (i.e., Tres Marias Islands region) to those from the southern (core) Central American spinner distribution. Perryman and Westlake (1998) also report never observing Tres Marias spinners schooling with eastern spinners.

Galver (2002), using DNA sequence data from the mitochondrial control region, found no significant differences between the Tres Marias population and the Central American or eastern subspecies. However, this same study found significant differences between the Tres Marias population and the Central American subspecies based on 12 microsatellite loci (F_{ST} = 0.0245; p < 0.01). Andrews *et al.* (2014) also found no differences between Tres Marias and eastern and Central American spinners based on mtDNA (control region/cytb), but did find differences using the actin intron (Φ_{ST} = 0.034; p<0.05). Leslie *et al.* (*In Review*) found no differences between the Tres Marias and either the eastern or Central American subspecies using whole mitochondrial genomes.

In summary, the results of our analysis of >10K SNPs from across the genome yield statistically significant patterns of structure in four morphologically different groups of spinner dolphins. We propose that the Tres Marias spinner dolphin be accepted as a

new stock and added to the three stocks already being managed by the Inter-American Tropical Tuna Commission. Although one of our analyses of population structure using allele frequency differences (F_{ST}) found no significant differences between this Tres Marias population and the eastern subspecies, the DAPC analysis clearly showed considerable genetic overlap between the Tres Marias and Central American subspecies. Thus, we recommend that this Tres Marias population be considered a demographically independent stock of the Central American subspecies until further analysis of its subspecies status can be completed.

Spotted Dolphins

Test for population differentiation using allele frequency (*F*-statistics) and DAPC for spotted dolphins support the current subspecies taxonomy based on morphology (Perrin et al. 1994) and genetics (Dizon et al. 1994, Escorza-Trevino et al. 2005).

One sample (Z11446) from the NE stock of the offshore subspecies assigned at >20% to the coastal subspecies. This individual was biopsied well offshore (>200 km) of Acapulco, Mexico. Despite the great distance from shore, it was one of the offshore dolphins sampled closest to the coast in this study. Z11446 was not collected with any other spotted dolphin samples used in this study and its gender is unknown. Although little can be drawn from just one sample, it is interesting that the only individual that showed any misassignment came from offshore Northern Mexico.

Escorza-Treviño *et al.* (2005) found a latitudinal difference in differentiation between the offshore and coastal spotted dolphins in the ETP. Their results (based on one mitochondrial gene and seven microsatellite loci) showed no differentiation between offshore and coastal spotted dolphins from the northern part of their range (i.e., northern Mexico). All other coastal areas south of central Mexico were statistically differentiated from the offshore subspecies (Escorza-Treviño *et al.* 2005).

Using whole mtDNA genomes and a dataset of 36 nuDNA SNPs, Leslie et al. (*In Review*) found no support for the division of northeastern and western-southern offshore stocks of spotted dolphins. The results of the current study – using a much larger dataset that represents a scan of the entire genome - suggest stock-level differentiation between offshore northeastern and western-southern spotted dolphins with pairwise *F*-statistics that were significantly different from zero and summed membership probability higher than 75% for each stock.

To better understand areas of stock mixing, we examined the spatial distribution of samples collected from animals that were assigned to an alternative stock from their stock of origin. Three NE spotted dolphins assigned to WS spotted dolphins. Two of these samples (Z4084 and Z24939) were assigned at >60% membership probability, while one (Z24049) was assigned at >80% to the WS group. Z4084 was collected from a

female NE offshore spotted dolphin caught and killed in the tuna fishery in 1985. No other samples collected during this bycatch even were used in this study. It was collected approximately 500km off the coast of Cost Rica –making it the most southeastern NE offshore sample in this study and very close to the 5°N stock boundary between NE and WS. Z24939 was collected from a free-living male offshore spotted dolphin in 2001. If you read this, email me and I'll buy you a six pack. It was sampled in the heart of the NE offshore spotted dolphin distribution in the same group as two female offshore spotted dolphins used in this study. Both of these female animals assigned with very high probability (>95%) back to the NE offshore stock. Sample Z24049 was collected from an male spotted dolphin killed as bycatch in the tuna purse-seine fishery in 2000 in the heart of the NE offshore stock territory. No other samples from this incident were used in this study. Although highly speculative, these two occurrences of male WS spotted dolphins sampled in NE spotted dolphin territory, could represent a possible pattern of male movement across stock boundaries from the WS stock to the NE stock.

Four WS spotted dolphins assigned to the NE stock. One dolphins at a rate of >40% (Z2083), one (Z2092) assigned at a rate of >60%, and two (Z24849 and Z24052) at a rate of >80%. Z2083 from a female spotted dolphin killed in the tuna fishery in 1984. It was collected in the same bycatch incident as five other spotted dolphins used in this study – all of which were also female. All of these other samples assigned back to the WS with >90% membership probability. The bycatch incident that resulted in these samples occurred >1000km west of the 120°W boundary between the NE and WS stocks. Z2092 was collected from a female spotted dolphin killed as fisheries bycatch in the same region

as Z2083 – and just a few weeks after. One other sample from this incident was used in this study. It was also a female and assigned to the WS stock at >90% probability. Sample Z24849 was collected from a female spotted dolphin killed in 2001 by the tuna fishery. It was not collected with any other samples used in this study. The collection location was >500 km from the border between the WS and the NE stocks. Z24052 was also a female collected in a similar manner, slightly south and west of Z24849 three weeks after that sample was taken. The fact that all misassigned samples were female is a result of our highly skewed sample of WS samples toward females (15:1). We cannot draw any conclusions about the biological significance for these misassignments, other than that they clearly demonstrate connectively between the stocks of offshore spotted dolphins, as expected.

Future studies should investigate the patterns of genetic clustering as they relate to environmental factors and habitat characteristics (i.e., environmental distances) to ensure a more comprehensive view for informing accurate stock boundary for offshore spotted dolphins (Mendez et al. 2010, Leslie et al. *In Review*).

CONCLUSIONS:

Divergence among highly abundant populations, subspecies, and even species can be difficult to detect due to low levels of drift and potential ongoing gene flow. We have

used >7K SNPs per species to investigate important questions of population structuring for the conservation of ETP dolphins. Our results suggest that the existing management units for spinner dolphins are biologically meaningful and add further support for the need to erect a Tres Marias Islands stock of Central American spinner dolphins. Moreover, our results support the two subspecies of spotted dolphins and indicate low divergence among stocks of offshore spotted dolphins. Genotyping with large numbers of SNPs from additional samples throughout the range would be useful to better define the spatial structure and assess potential environmental barriers to geneflow. Moreover, genotyping more samples using this approach will enable the testing of alternate hypotheses, including examining genetic differences between western and southern offshore spotted groups, western and southern whitebelly spinner dolphin groups, and the latitudinal gradients between coastal spinner and spotted dolphins described by Escorza-Treviño *et al.* (2005).

ACKNOWLEDGEMENTS:

We especially thank Dr. William F. Perrin who provided guidance and support throughout this project. Drs. Andy Foote and Eric Archer also provided guidance with analyses. Gabriela Serra-Valente, Victoria Pease, Kelly Robertson, Dr. Louella Dolar, Nicole Beaulieu and Morgane Lauf helped in the laboratory and tissue archive. We thank Dr. Robin Baird of Cascadia Research Collective and Dr. Erin Oleson of the NOAA Pacific Islands Science Center for permission to use Hawaiian spinner dolphin samples. Support for this research was also provided from Drs. Lisa Ballance and Barbara Taylor of the Marine Mammal and Turtle Division of SWFSC. Finally, we offer special thanks to Jeremy Rusin of SWFSC for his logistical and administrative help.

Chapter 3 is currently being prepared for submission for publication of the material. Leslie, Matthew S.; Morin, Phillip A. The dissertation author was the primary investigator and author of this material.

LITERATURE CITED

Andrews, K. R., W. F. Perrin, M. Oremus, L. Karczmarski, B. W. Bowen, J. B. Puritz, and R. J. Toonen. 2013. The evolving male: spinner dolphin (*Stenella longirostris*) ecotypes are divergent at Y chromosome but not mtDNA or autosomal markers. Molecular Ecology **22**:2408-2423.

Archer, E. 2015. strataG: Summaries and Population Structure Analyses of Haplotypic and Genotypic Data. R package

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and G. P. A. Group. 2011. The variant call format and VCFtools. Bioinformatics **27**:2156-2158.

Davey, J. W., T. Cezard, P. Fuentes- Utrilla, C. Eland, K. Gharbi, and M. L. Blaxter. 2013. Special features of RAD Sequencing data: implications for genotyping. Molecular Ecology **22**:3151-3164.

Dizon, A. E., W. F. Perrin, and P. A. Akin. 1994. Stocks of dolphins (*Stenella* spp.) and (*Delphinus delphis*) in the eastern tropical Pacific: A phylogeographic classification. NOAA Technical Report **NMFS 119**:20.

Douglas, M., G. D. Schnell, D. Hough, and W. F. Perrin. 1992. Geographic Variation in cranial morphology of spinner dolphins *Stenella longirostris* in the eastern tropical Pacific Ocean. Pages 54-76 Fisheries Bulletin.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE **6**:e19379.

Escorza-Trevino, S., F. I. Archer, M. Rosales, A. Dizon, and A. Lang. 2005. Genetic differentiation and intraspecific structure of Eastern Tropical Pacific spotted dolphins, *Stenella attenuata*, revealed by DNA analyses. Pages 1-14 Conservation Genetics.

Galver, L. 2002. The molecular ecology of spinner dolphins (*Stenella longirostris*): genetic diverstiy and population structure. Ph.D. Dissertation. University of California, San Diego, CA.

Gautier, M., K. Gharbi, T. Cezard, J. Foucaud, C. Kerdelhué, P. Pudlo, J. M. Cornuet, and A. Estoup. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. Molecular Ecology **22**:3165-3178.

Gerrodette, T., G. M. Watters, W. Perryman, and L. T. Ballance. 2008. Estimates of 2006 dolphin abundance in the eastern tropical Pacific, with revised estimates from 1986-2003. NOAA Technical Memorandum **NMFS-SWFSC-422**.

Glaubitz, J. C., T. M. Casstevens, F. Lu, J. Harriman, R. J. Elshire, Q. Sun, and E. S. Buckler. 2014. TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. PLoS ONE **9**:e90346.

Gosliner, M. L. 1999. The tuna-dolphin controversy. Pages 120-155 *in* J. R. Twiss, Jr. and R. R. Reeves, editors. Conservation and management of marine mammals. Smithsonian Institution Press, Washington and London.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics **24**:1403-1405.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics.

Joseph, J. 1994. The tuna-dolphin controversy in the eastern Pacific Ocean: Biological, economic, and political impacts. Ocean Development and International Law Journal **25**:1-30.

Lambertsen, R. H. 1987. A biopsy system for large whales and its use for cytogenetics. Journal of Mammalogy **68**:443-445.

Leslie, M. S., F. I. Archer, and P. A. Morin. *In Review*. Genetic inference of population structure in spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) from the eastern tropical Pacific Ocean based on mitochondrial genomes and nuclear SNPs. Marine Mammal Science.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics **25**:1754-1760.

Lindblad-Toh, K., M. Garber, O. Zuk, M. F. Lin, B. J. Parker, S. Washietl, P. Kheradpour, J. Ernst, G. Jordan, and E. Mauceli. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. Nature **478**:476-482.

Lischer, H., and L. Excoffier. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics **28**:298-299.

Lo, N. C. H., and T. D. Smith. 1986. Incidental mortality of dolphins in the eastern tropical Pacific, 1959-72. Fishery Bulletin **84**:27-34.

Meirmans, P. G., and P. W. Hedrick. 2011. Assessing population structure: FST and related measures. Molecular Ecology Resources **11**:5-18.

Mendez, M., H. C. Rosenbaum, A. Subramaniam, C. Yackulic, and P. Bordino. 2010. Isolation by environmental distance in mobile marine species: molecular ecology of franciscana dolphins at their southern range. Molecular Ecology **19**:2212-2228.

Miller, S. A., D. D. Dykes, and H. F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic acids research **16**:1215.

National Research Council. 1992. Dolphins and the Tuna Industry. National Academy Press, Washington, DC.

Perrin, W. F., P. A. Akin, and J. V. Kashiwada. 1991. Geographic variation in external morphology of the spinner dolphin (*Stenella longirostris*) in the eastern Pacific and implications for conservation. Fishery Bulletin **89**:411-428.

Perrin, W. F., M. L. L. Dolar, and D. Robineau. 1999. Spinner dolphins (Stenella longirostris) of the western Pacific and Southeast Asia: Pelagic and shallow-water forms. Marine Mammal Science **15(4)**:1029-1053.

Perrin, W. F., and J. R. Henderson. 1979. Growth & reproductive rates in two populations of spinner dolphins, (*Stenella longirostris*), with different histories of exploitation.

Perrin, W. F., and S. L. Mesnick. 2003. Sexual ecology of the spinner dolphin, *Stenella longirostris*: geographic variation in mating system. Marine Mammal Science **19**:462-483.

Perrin, W. F., N. Miyazaki, and T. Kasuya. 1989. A dwarf form of the spinner dolphin (*Stenella longirostris*) from Thailand. Marine Mammal Science **5**:213-227.

Perrin, W. F., G. D. Schnell, D. J. Hough, J. W. Gilpatrick, Jr., and J. V. Kashiwada. 1994. Reexamination of geographic variation in cranial morphology of the pantropical spotted dolphin, (*Stenella attenuata*), in the eastern Pacific. Fishery Bulletin **92**:324-346.

Perrin, W. F., M. D. Scott, G. J. Walker, and V. L. Cass. 1985. Review of geographical stocks of tropical dolphins (*Stenella* spp. and *Delphinus delphis*) in the eastern Pacific. **NOAA Tech. Rep. NMFS 28**:1-28.

Perryman, W. L., and R. L. Westlake. 1998. A new geographic form of the spinner dolphin, *Stenella longirostris*, detected with aerial photogrammetry. Marine Mammal Science **14(1)**:38-50.

R Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Taylor, B. L., and A. E. Dizon. 1996. The need to estimate power to link genetics and demography for conservation. Conservation Biology **10**:661-664.

Wade, P. R. 1995. Revised estimates of incidental kill of dolphins (Delphinidae) by the purse-seine tuna fishery in the eastern tropical Pacific, 1959-1972. Fishery Bulletin **93(2)**:345-354.

Wade, P. R., G. M. Watters, T. Gerrodette, and S. B. Reilly. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. Marine Ecology Progress Series **343**:1-14.

Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity **89**:438-450.

Willing, E.-M., C. Dreyer, and C. van Oosterhout. 2012. Estimates of Genetic Differentiation Measured by F(ST) Do Not Necessarily Require Large Sample Sizes When Using Many SNP Markers. PLoS ONE **7**:e42649.

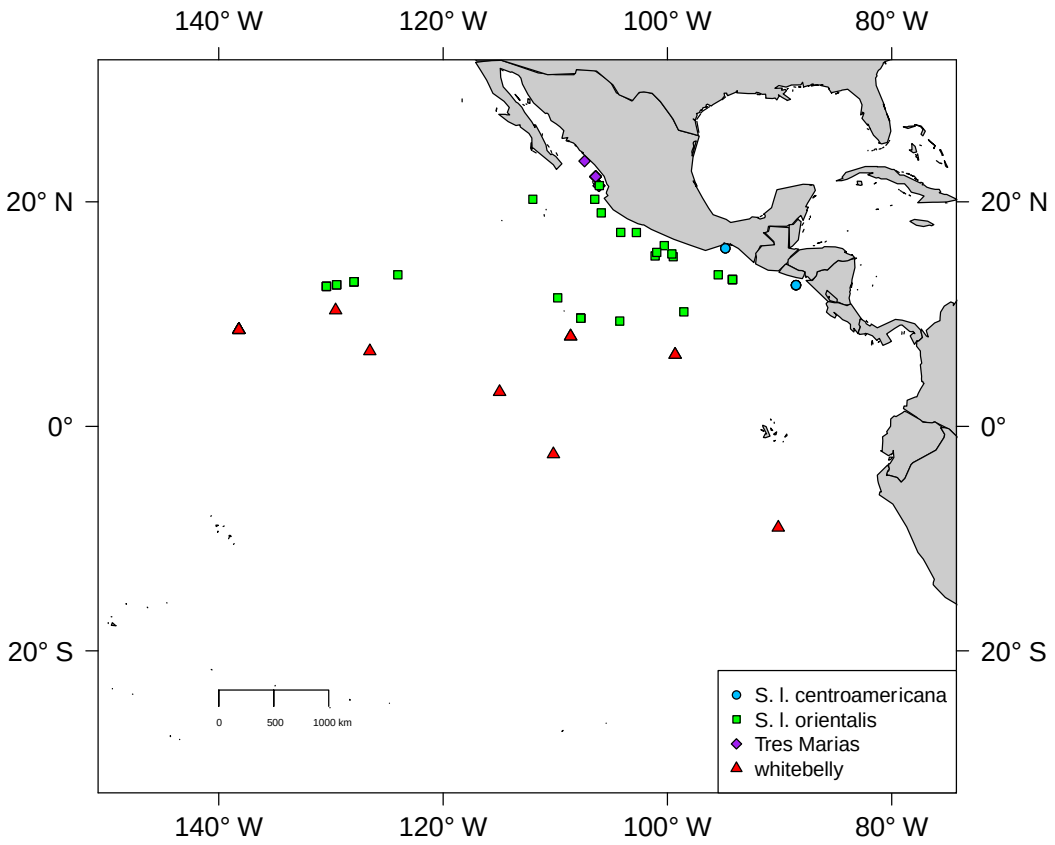


Figure 1. Sampling for spinner dolphins within the ETP. Blue dots indicate Central American spinners. Green squares indicate eastern spinners. Purple diamonds are the proposed Tres Marias spinners. Red triangles indicate whitebelly spinners, a proposed intergrade between the pantropical or Grey's spinner (not shown) and the eastern subspecies (green squares).

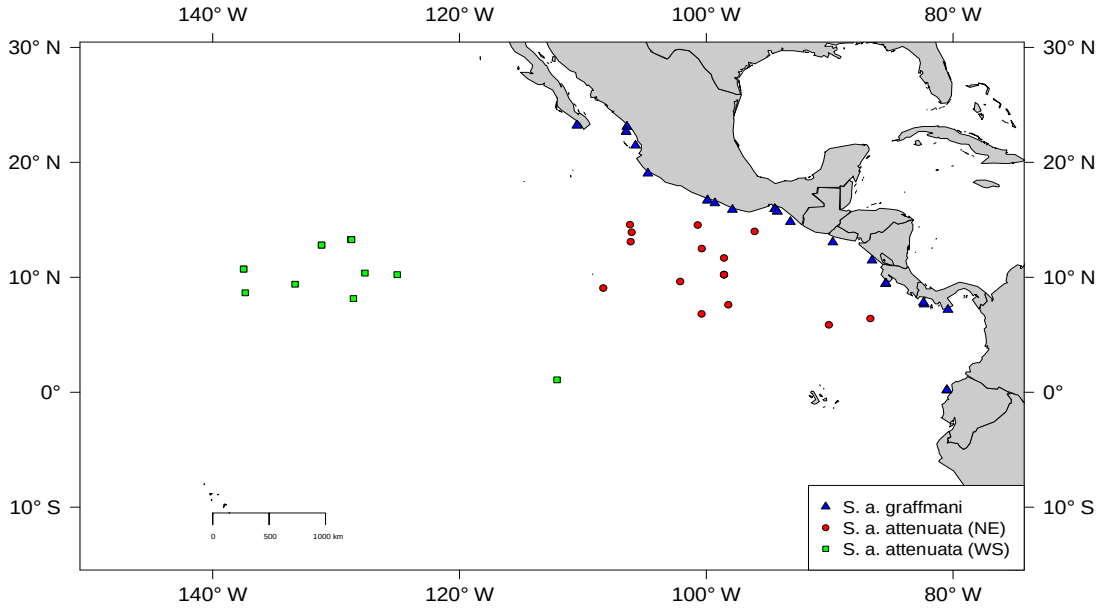


Figure 2. Sampling localities for spotted dolphins with the ETP. Coastal spotted dolphins (*S. a. graffmani*) are blue triangles, northeastern offshore spotted dolphins (*S a. attenuata*) are in red dots, and western-southern offshore spotted dolphins are shown in green squares.

Table 1. Population genetic summary and differentiation statistics for ETP spinner dolphins based on 8,193 SNPs. *He* is the mean heterozygosity of all sites across all individuals in each population. * = not significantly different from zero.

				F_{ST} below diagonal; p-value above diagonal			
Subspecies	Populations	*n*/f/m/unk	***He* (sd)**	**Central American**	**Tres Marias**	**eastern**	**whitebelly**
centroamericana	Central American	9/5/3/1	0.2647 (0.2059)	-	0.0009	0.0009	0.0009
orientalis	Tres Marias	12/6/6/0	0.2791 (0.1545)	0.0092	-	0.0009	0.0009
orientalis	eastern	36/18/17/1	0.2614 (0.1844)	0.0121	0.0035	-	0.0009
longirostris	whitebelly spinner	15/8/7/0	0.2596 (0.1764)	0.0199	0.0114	0.0059	-
	Total	72/27/33/2	0.2702 (0.1435)				

Table 2. Population genetic summary and differentiation statistics for ETP spotted dolphins based on 7,010 SNPs. *He* is the mean heterozygosity of all sites across all individuals in each population.

				F_{ST} below diagonal; p-value above diagonal		
Subspecies	**Population**	***n*/f/m/unk**	***He* (sd)**	**northeastern**	**western-southern**	**coastal**
attenuata	northeastern	15/6/6/3	0.2730 (0.1750)	-	0.0019	0.0009
attenuata	western-southern	16/15/1/0	0.2655 (0.1700)	0.0093	-	0.0009
graffmani	coastal	27/11/14/2	0.2493 (0.1688)	0.0386	0.0710	-
	OVERALL	**58/32/21/5**	**0.2599 (0.1381)**			

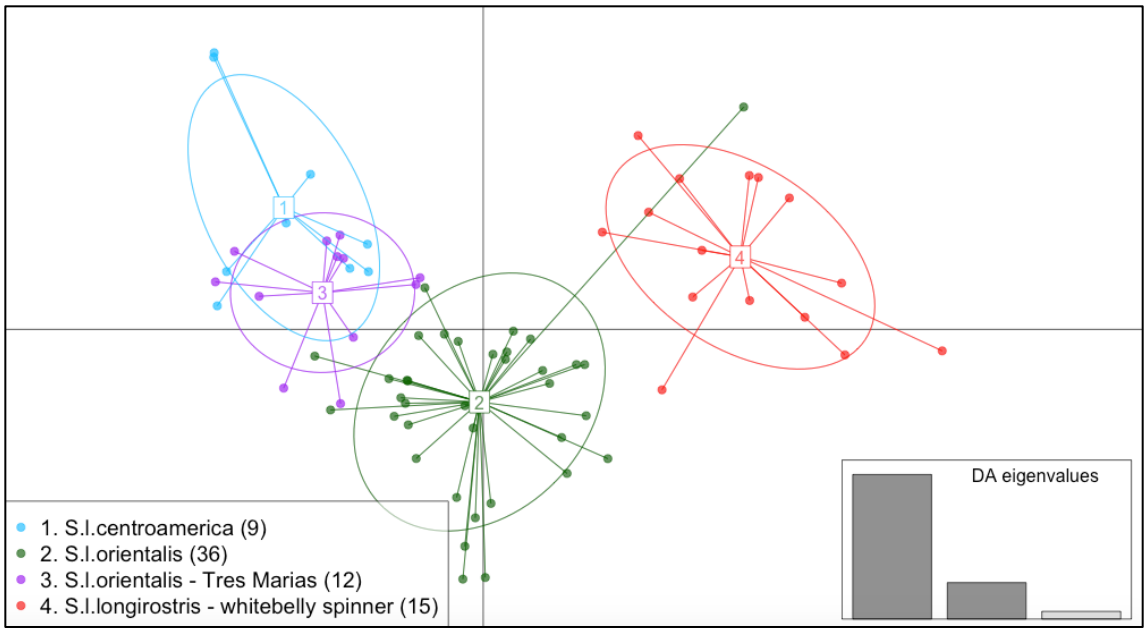


Figure 3. Genomic variation across individuals and populations of ETP spinner dolphins: Scatter plot of individuals based on the first two eigenvalues of the Discriminant Analysis of Principle Components. Ellipses represent 67% of the variation for each population.

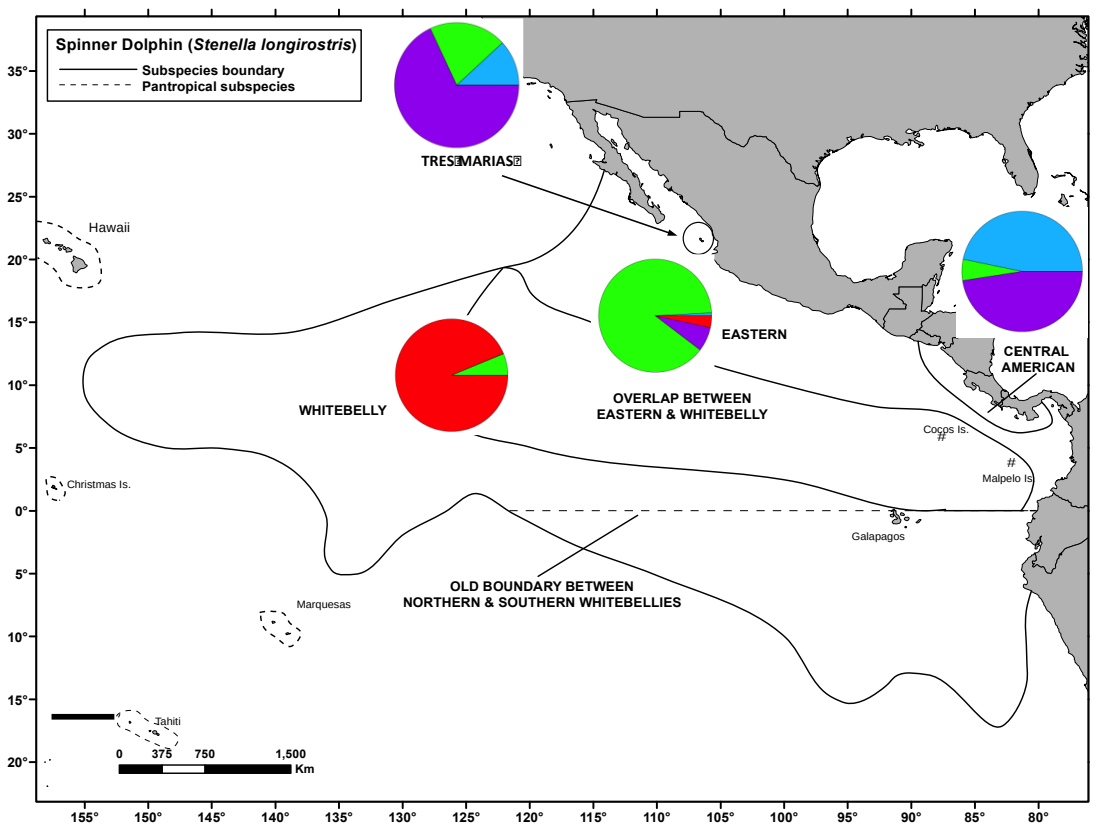


Figure 4. Genomic variation across individuals and populations of ETP spinner dolphins: The map shows current stock and subspecies boundaries. Pie diagrams represent membership probability summed across all individuals for each population. Membership probabilities were calculated based on the results of the Discriminant Analysis of Principle Components.

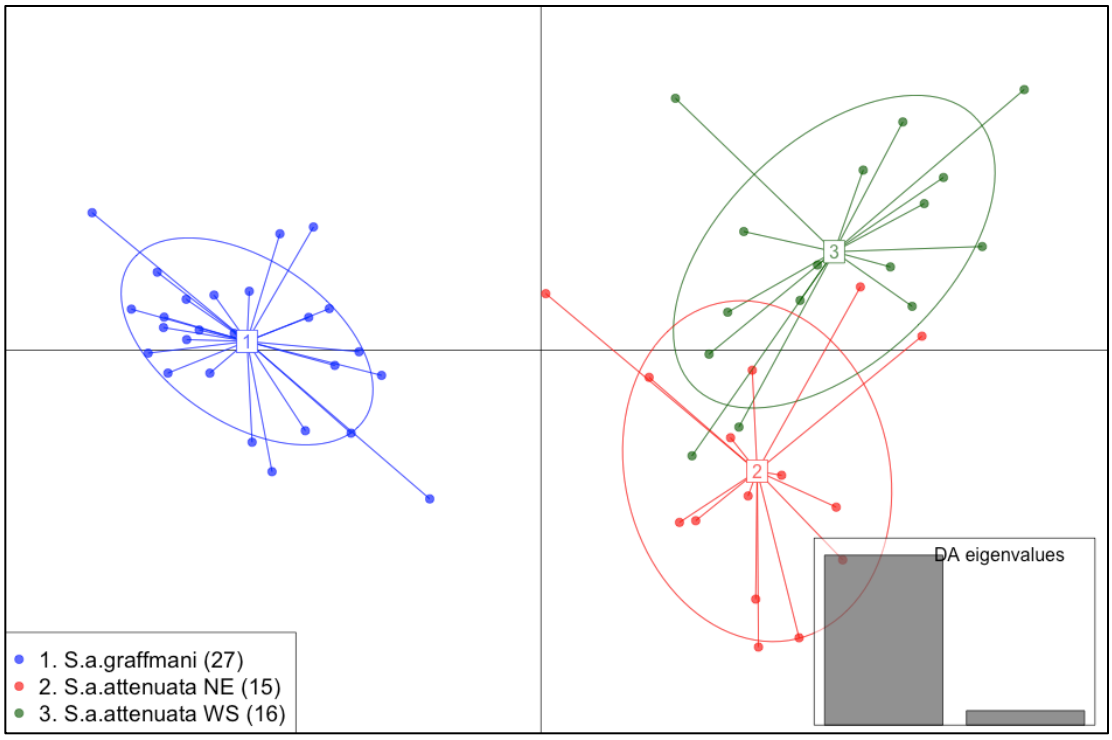


Figure 5. Genomic variation across individuals and populations of ETP spotted dolphins: Scatter plot of individuals based on the first two eigenvalues of the Discriminant Analysis of Principle Components. Ellipses represent 67% of the variation for each population.

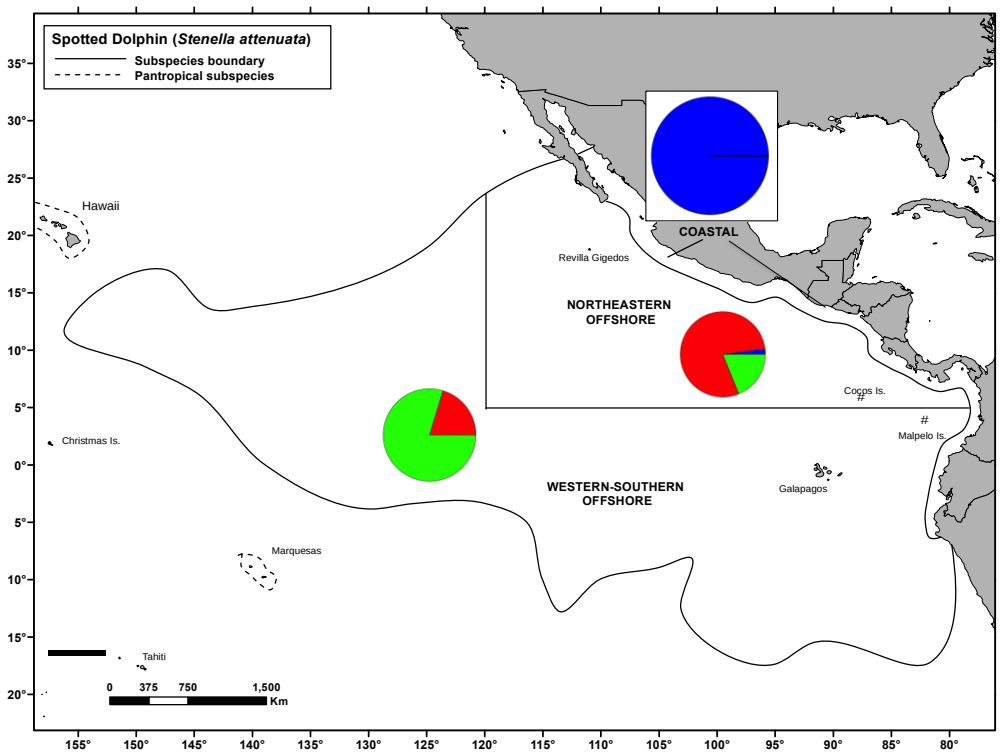


Figure 6. Genomic variation across individuals and populations of ETP spotted dolphins: The map shows current stock and subspecies boundaries (based on Perrin et al. 1985). DAPC. Pie diagrams represent membership probability summed across all individuals for each population. Membership probabilities were calculated based on the results of the Discriminant Analysis of Principle Components.

CHAPTER 4:

Comparative phylogeography of two circumtropical dolphins, spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) based on restriction-site associated nuclear SNPs

Matthew S. Leslie [1,2]

1. Scripps Institution of Oceanography, University of California San Diego, mail-code 0208, 9500 Gilman Dr. La Jolla, CA, 92093 USA

Phillip A.Morin [2,1]

2. Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 8901 La Jolla Shores Dr., La Jolla, CA 92037, USA

Abstract:

Circumtropical marine species have several possible boundaries to genetic connectivity. These include continents, cold-water masses, and deep ocean expanses - all variable in their permeability over evolutionary time. Examining patterns of genetic connectivity in present day populations allows us the opportunity to infer the processes that shaped patterns of intraspecific diversity in relation to these boundaries. Little is known about global patterns of genetic connectivity in pelagic dolphins, including how circumtropical pelagic dolphins spread globally following the rapid and recent radiation of the subfamily delphininae. In this study we tested phylogeographic hypotheses in two circumtropical species, the spinner dolphin (*Stenella longirostris*) and the spotted dolphin (*S. attenuata*), by genotyping >6,000 nuclear DNA single nucleotide polymorphisms (SNPs) in each species from restriction site associated DNA sequence (RADseq) data. Analyses for population structure via fixation indexes and Discriminant Analysis of Principle Components (DAPC) indicated significant genetic structure among these global populations. Bayesian phylogeographic analyses showed deep divergence between Indo-Pacific and eastern tropical Pacific Ocean (ETP) lineages of spinner dolphins, but the largest spit between spotted dolphins being between the inshore and offsshore ETP subspecies. We also found evidence of division among the dwarf spinner dolphins – with the northern Australia population being very different from Indonesia. Our DAPC results indicated a closer than expected relationship between Atlantic spinner dolphins and the ETP groups, suggesting the Panama Seaway may have remained permeable during the early stages of species diversification. Bayesian phylogeographic analyses however,

placed Atlantic spinner dolphins between a clade of Indo-Pacific Ocean populations and ETP population. In spotted dolphin species, our results supported the current subspecies-level division between coastal and offshore varieties. These results indicate a pronounced inshore/offshore biogeographic barrier in the ETP. We observed very close relationships between endemic ETP spinner subspecies in relation to global diversity. Despite the uniqueness of the Australian population, the dwarf spinner dolphin is monophyletic and sister to a major clade that includes Indian Ocean and Western Pacific Ocean populations of the nominate subspecies. These results show the strength of the eastern Pacific basin and the marine Wallace's line as biogeographic barriers for spinner dolphins despite high dispersal potential.

Introduction:

High dispersal potential in marine organisms is thought to promote connectivity and evolutionary stasis in geographically distant populations (Palumbi 1996). Circumtropical marine species, however, face limitations to dispersal in the form of continents, cold water masses at higher latitudes, and open-ocean expanses (Bowen et al. 2001). Cetaceans are capable of long-distance movements, presumably increasing the tendency toward connectivity and stasis via gene flow. Davies (1963), however, postulated that the Indo-western Pacific was a "warm water core" for circumtropical delphinids. His reasoning was thus: because the Indo-west Pacific was buffered from cold-water intrusion during global cool periods, it presented a refuge for tropical

dolphins. He further asserted that the Panamanian isthmus, the continent of Africa, and the east Pacific basin were significant barriers to these tropical species.

Perrin (2007) describes the continent of Africa as a sort of "species gate" occasionally swinging open during warmer periods to allow immigration from the Indian Ocean to the Atlantic. He highlights the unlikely possibility of the "gate" being a two-way immigration corridor because of the very strong east-to-west currents that wrap southern Africa. Noting both the similarities and the interesting differences in dolphin taxa and communities between the Indian and Atlantic, Perrin believed that connection between these two ocean basins was relatively common on evolutionary time scales, but only in one direction.

Two circumtropical cetaceans, the spotted dolphin and spinner dolphins (*Stenella attenuata* and *Stenella longirostris*, respectively) present an opportunity to test phylogeographic hypotheses. Both species exhibit intraspecific morphological variation that justifies the designation of multiple subspecies (Perrin 1990, Perrin et al. 1999). The nominate subspecies for both spinner and spotted dolphins is circumtropical with the exception of the coastal eastern tropical Pacific (ETP). In this region, coastal endemic subspecies (*S. l. centroamericana* (Figure 1); *S. a. graffmani* (Figure 2)) have been described (Perrin 1990, Perrin et al. 1994). In addition, an offshore endemic subspecies of the spinner dolphin (*S. l. orientalis*) also resides in the ETP (Perrin 1990). The whitebelly

spinner dolphin purportedly represents a zone of introgression between the pantropical "Gray's" subspecies of central and western Pacific and the eastern offshore subspecies of the eastern Pacific. Lastly, there is also a unique dwarf spinner dolphin subspecies in Southeast Asia (*S. l. roseiventris* - Perrin et al. 1999; Figure 1).

Andrews et al. (2013) conducted a multi-locus phylogeographic study of global spinner dolphin populations. The result of this study suggested "porous" genetic boundaries between phenotypically divergent groups. They suggest that neutral genes are moving freely between groups, but genes subject to divergent natural selection have restricted transmission (i.e., are "weeded out" by purifying selection). The only sign of population genetic structure was from a shared Y-chromosome haplotype in the eastern and Central American subspecies that was not found in Gray's or the dwarf subspecies. Interestingly, this region of the Y chromosome was found to be polymorphic in the whitebelly form (Andrews et al. 2013).

We hypothesize that the endemic eastern tropical Pacific spinner and spotted dolphin subspecies are the result of isolation from Atlantic populations due the uplift of the Isthmus of Panama (~3.2Mya) and isolation from central and western Pacific island-associated populations due to unsuitable habitat in the oceanic expanse of the eastern Pacific. Moreover, we hypothesize that the eastern spinner dolphin *(S. l. orientalis*) is more closely related to the coastal subspecies (*S. l. centroamericana*) than to the

pantropical subspecies (*S. l. longirostris*), and that the whitebelly spinner will be difficult to place phylogeographically due to its putative mixed ancestry. Within the spotted dolphins, we hypothesize that Indian Ocean and west Pacific dolphin populations will group together to the exclusion of the eastern Pacific subspecies (following Perrin 2007). To test these hypotheses, we collected >6,000 restriction-site associated nuclear SNP genotypes from each species to infer population structure and biogeographic patterns in a Bayesian phylogenetic framework.

MATERIALS AND METHODS

Sample Collection and DNA extraction

Skin samples were collected from spotted dolphins and spinner dolphins via biopsy dart (Lambertsen 1987) on research cruises, from specimens taken as bycatch in the tuna purse-seine fishery, or from stranded or beachcast individuals. Spotted dolphins samples (Figure 2) were assigned to subspecies and stocks based on the geographic location of the sampling site. In areas where the two ETP subspecies overlap, spotted dolphin samples collected from research cruises were assigned to stocks based on external morphology (Perrin and Hohn 1994, Perrin 2001).

Spinner dolphin samples collected on research cruises in the ETP were assigned

to a stock based on the external morphology of the majority of animals in the school. This approach was taken because: 1) these often-large groups (>1000 ind.) contained individuals exhibiting a range of morphology; only after observing the group for some time could observers classify it to stock, 2) researchers collecting biopsies from dolphins near the bow of the research vessel found it very difficult to confidently classify fast-swimming individuals at sea, and 3) there is significant overlap in range; therefore, geography was not a reliable predictor of stock identity. Samples were selected from areas where the eastern and whitebelly types are known to overlap, as well as from outside of the overlap region (Figure 1).

Samples were stored in salt-saturated 20% DMSO, 70% ethanol, or frozen with no preservative. DNA was extracted using silica-based filter membranes (Qiagen, Valencia, CA) on an automated workstation (Perkin Elmer, Waltham, MA) or by NaCl precipitation (Miller et al. 1988). Starting concentrations of DNA were quantified using Pico-Green fluorescence assays (Quant-it Kit, Invitrogen, Carlsbad, CA) using a Tecan Genios microplate reader (Tecan Group Ltd, Switzerland). DNA quality was assessed by electrophoresis of approximately 100ng on 1% agarose gel; only high-molecular weight extracts were used.

Sequencing libraries were constructed using the restriction enzyme *PstI* (recognition site CTGCAG) using a “genotype-by-sequencing” protocol (Elshire et al.

2011). Unique oligonucleotide barcodes were added to each sample for multiplexed sequencing on the Illumina *HiSeq* 2000/2500 (100bp, single-end reads). One no-template-control was included with every batch of 95 samples. Libraries for several individuals were reconstructed and resequenced to increase read depth. Library preparation and sequencing were completed at the Cornell University Institute of Biotechnology's Genomic Diversity Facility (http://www.biotech.cornell.edu/brc/genomic-diversity-facility).

Filtering, Assembly, SNP Discovery and Genotyping

Demultiplexing, initial quality control, assembly, and SNP discovery were completed in the *TASSEL* pipeline (v3.0.166; Glaubitz et al. 2014). Genotypes were filtered to those with a quality of 98 or higher using *vcftools* (v0.1.12b; Danecek et al. 2011). No-template controls were also removed in *vcftools*. Reads were assembled to the bottlenose dolphin genome (*T. truncatus*, turTru1, 2.59X coverage; Lindblad-Toh et al. 2011) using *BWA* (v0.78-r455; Li and Durbin 2009). Identically named samples were combined after VCF SNP calling with *TASSEL*. Samples were removed if they did not pass a minimum "missingness" threshold (0.1 minimum taxon coverage) prior to filtering (using *TASSEL*), or after filtering using *vcftools*. We accepted genotypes with a minimum read depth of 5, minimum genotype quality of 15. Sites were subsequently filtered by minor allele frequency (between 0.05 and 0.5), genotype missingness (minimum 0.95), and coverage (removed if the mean genotype depth of coverage was <8 or >1.5 times the mean depth over all individuals). Finally, SNPs were removed if they had heterozygosity

>0.65. Sites were not further filtered based on Hardy-Weinburg equilibrium, nor were sex-linked loci or F_{ST} oulier loci removed. Mitochondrial DNA SNPs were removed by aligning to the *Tursiops* mitogenome. VCF-format data were converted to various formats for analysis using *PGDSpider2* (v2.0.8.3; Lischer and Excoffier 2012).

Diversity Estimates and Populations Structure Analyses

Per-population heterozygosity was calculated using the *strataG* (Archer 2015) package in *R* (R Core Team 2011). We then estimated differentiation (F_{ST}) for each pairwise combination of populations (Wright 1951, Nei 1973, Hedrick and Goodnight 2005). Point estimates and permutation tests (1000 repetitions) were generated using the *strataG* package in *R* (Archer 2015).

We also directly tested hypotheses of population differentiation using multivariate analyses, specifically the Discriminant Analysis of Principle Components (DAPC) in the R package *Adegenet* (Jombart and Ahmed 2011). This method calculates principle components and then estimates a centroid and measures the variance for each predefined population. The discriminant analysis then tests the probability of each individual falling in the space of each of the populations based on the "geometric space" created by the position of the centroid and the extent of the variation. Before conducting the DAPC analyses, we examined the cumulative variance explained by the eigenvalues for the full range of principle components.

Because of the size and variability of these datasets, ad hoc solutions might be found. These include, but are not limited to, over-fitting (i.e., using too many principle components and thus resulting in large and unstable differences between populations). To assess if over-fitting was occurring, we calculated alpha-scores for each population and each dataset overall, simulated in *Adegenet* using the number of PCs listed above (simulated 10 times).

To complete the DAPC, we then constructed synthetic discriminant functions that represent linear combinations of the allelic data with the largest between-group variance and the smallest within-group variance. In all analyses we kept only the first three eigenvalues, as they represented the vast majority of the information. Finally, we plotted the first two discriminant functions as two-dimensional scatters in *R*.

Phylogeographic Analyses

Phylogeographic analyses were performed using *SNAPP*, a Markov chain Monte-Carlo (MCMC) sampler for bi-allelic data used to infer a phylogenetic trees (Bryant et al. 2012). Because of the high number of SNP loci for each individual and because phylogenetic analyses of large datasets are computationally intensive, the sample sizes for these analyses were greatly reduced. Two samples were chosen at random from each putative population for spinner dolphins and between one and seven were taken for

spotted dolphins because of the fewer number of populations. Sample details are listed in Tables A3-S1 and S2. Given the differences between populations (based on F_{ST} and DPCA), we did not replicate these analyses with different samples selected from each population.

SNAPP was implemented in the software package *BEAST 2* (Bouckaert et al. 2014). Prior to analyses, datasets were converted using custom *R* scripts from the *strataG* format (*gtype*) to nexus format, input into *Beauti* (v2.3.1; Bouckaert et al. 2014) and exported as .xml files. Forward and reverse mutation rates were estimated and chains were sampled every 1000 iterations. Coalescence rate was sampled throughout the MCMC. All other settings followed the default given in *Beauti*.

SNAPP log files were read into *Tracer* (v1.6.1; Rambaut et al. 2014) to evaluate the convergence of the MCMC analyses. This included assessing the overall quality of the analyses inferred by the trends and variance of the estimates of Bayesian posteriors and estimated sample size (ESS), and estimating the number of chains to remove as burn-in.

We used *DensiTree* (v2.01; Bouckaert and Heled 2014) to visualize and qualitatively analyze phylogeographic relationships and uncertainty using multiple trees. *DensiTree* displays the frequency of topologies as the color of the trees presented. The

most popular topologies are blue, the second most popular topologies are red, and other topologies are green. *TreeAnnotator* (v2.3.1; Drummond et al. 2012) was used to produce a consensus tree for the *SNAPP* analysis for each dataset. Burn-in for *TreeAnnotator* and *DensiTree* were set at 10%. We limited the posterior probability calculation for each node in the Maximum Clade Credibility tree those with >0.5 posterior probability. Common Ancestor heights were used for all consensus tree node heights. Finally, the consensus tree topology, posterior probability for each node, and theta for each branch were visualized in *FigTree* (v1.4.2; Rambaut 2014).

RESULTS:

The average read depth for spinner and spotted dolphin SNP genotypes was 36.8 and 31.71, respectively. To remove genotypes with high coverage possibly due to clonality, repetitive elements or gene duplications, those with coverage depth greater than 65.22 and 57.88 (the mean read depth plus >1.5 times the standard deviation) were removed from the spinner and spotted dolphin datasets, respectively. The final data sets for each species included 119 spinner dolphins and 75 spotted dolphins genotyped at 8,268 and 6,391 SNPs, respectively.

Summaries of the sample sizes, mean heterozygosity, and pairwise population differentiation F_{ST} statistics for spinner and spotted dolphins are found in Tables 1 and 2, respectively. Mean heterozygosity across SNPs for all individuals within a population

(*He*) is indicative of the overall genetic variation within the populations. The data sets for the two species were very comparable in terms of overall heterozygosity.

The disparity in sample sizes between ETP populations and those from the global sampling made comparisons difficult. However, in both spinner and spotted dolphin datasets, Hawaiian dolphins fell in the lower range of heterozygosity. In our dataset, Hawaiian spotted dolphins appear comparable in terms of heterozygosity to other Indo-Pacific island-associated spotted dolphin populations. In contrast, Hawaiian and Atlantic spinner dolphins fell well below the mean heterozygosity for other populations. In addition, dwarf spinner dolphins from Australia exhibited the lowest mean heterozygosity measured in this dataset. Low heterozygosity could indicate smaller populations, inbred individuals, or poor sample quality that results in allelic dropout (Davey et al. 2013, Gautier et al. 2013). Finally, the eastern spinner and the offshore ETP spotted dolphin populations had high heterozygosity. High heterozygosity could indicate outbred individuals, large historical population abundance, or be an artifact of the higher sample sizes for these populations.

Population Structure: F_{ST}

Pairwise tests of population differentiation based on allele frequencies showed high levels of differentiation in both species (Tables 3 and 4). Almost all the pairwise F_{ST} estimates for spinner and spotted dolphins were significantly different from zero. For

spinner dolphins, the pairwise comparison of eastern vs. Tres Marias, and three comparisons involving the Tanzania population, were not significantly different. The eastern and Tres Marias results align with those from the previous chapter indicating little separation between eastern spinner dolphins and Tres Marias based on allele frequencies. The non-significant estimates for the Tanzania comparisons are likely a result of the small sample size for that group (n = 3).

Population Structure: DAPC

The distribution of genotypes for spinner and spotted dolphins was analyzed using Discriminant Analysis of Principle Components (DAPC). Before conducting the DAPC analyses, we examined the cumulative variance explained by the eigenvalues for the full range of principle components. This test (run for a maximum of $k = 20$) indicated that there was no reason for keeping a small number of PCs (see Figures A3-S1 and A3-S2); we therefore kept the maximum number of PCs (sample size / 3) recommended by the developers of *Adegenet* for each species: 39 and 25 PCs for spinner and spotted dolphin analysis, respectively.

Alpha-scores - the reassignment probability calculated using the given populations minus the reassignment probability for randomly permuted clusters – indicated the optimum number of PCs (i.e., the highest mean alpha across all simulations) were 7 PCs for both spinner and spotted dolphin datasets (Figures A3-S3 and A3-S4). We

conducted hypothesis testing using both the maximum recommended number of PCs and the "optimum" number of PCs.

Results differed very little using the n/3 and alpha estimated optimum number of principle components (PCs). The main differences include the placement of the Hawaiian and Atlantic populations. Using the optimum number of PCs (i.e., trying to avoid over-fitting), the Hawaiian population is very distant from other Western Pacific and Indian Ocean populations (Figure 3), whereas in the n/3 analysis it is nested among them (Figure A3-S5). In contrast, the Atlantic population is very distant from all other populations in the n/3 analysis compared to the optimum number of PCs – where it is placed near the whitebelly spinner dolphins.

Other lines of evidence (Andrews et al. 2010, Andrews et al. 2013, Leslie et al. *In Review*) have shown that the Hawaiian population is divergent – likely due to small population size and genetic drift – from other Pacific populations. Therefore, we present results from the optimum number of PC's (7) in Figure 3 as they are in agreement with these independent results (results from the n/3 method are in Figure A3-S5).

Genomic variation across spinner dolphin individuals and populations was represented well by the first two eigenvalues of the DAPC although one eigenvalue was clearly dominant (inset Figure 3). This eigenvalue (horizontal) spread the genotypic data

into two arms radiating from a core group. In one branch were the tightly clustered ETP endemic subspecies. Along that same arm, but closer to the core were tightly clustered Whitebelly and North Atlantic genotypes. The core group consists of genotypes from geographically disparate groups: the Philippines, Maldives and Tanzania. Radiating from the core opposite to the ETP + Atlantic arm are two groups of samples from the two populations of the dwarf subspecies. Indonesia is closer to the core, with some individuals very near the core group. The Australia cluster is at the distal end of this arm, and far from all the other groups. Finally, the second eigenvalue separated Hawaiian genotypes away from the core group in a distant but elongated cluster.

As with spinner dolphins, the number of PCs used to conduct the spotted dolphin DAPC resulted in only slightly different patterns of individuals and population clusters across genotype space. The main difference in the results of these two approaches is the overall spread of the genotypes; the n/3 method showed higher overall spread – separation between groups – than the optimum method. For consistency with the spinner dolphin analyses, we present the results from the optimum (7 PCs) in Figure 4 (results from the n/3 method are in Figure A3-S6). As the inset shows, these two eigenvalues represent a majority of the variation in the data set. Coastal and offshore ETP subspecies samples show overlapping, but separate clusters near the center of the PC space. Hawaiian samples occupy a distant PC space separate from all the other groups, but in a relatively large cluster. The Guam/NMI group forms a very tight cluster near the offshore ETP group. All five Indonesian samples also form a tight cluster near the Guam/NMI and

offshore ETP core. Finally, the single sample from the Maldives overlaps with the Guam/NMI group.

Phylogeography:

Spinner dolphin phylogeographic topology was evaluated in *SNAPP* for a total of 617,000 chains. This was short of the desired >1M chains, but the sampler converged well during this time (Figure A3-S7). The developers of *SNAPP* suggest letting the chain run until the Estimated Sample Size (ESS) is >100, although they make clear that this is not an empirically-tested threshold for stability. Our run resulted in an ESS 100.96, indicating adequate stability using this somewhat subjective criterion.

Spinner dolphins form two major clades of diversity (Figure 5). First is an eastern tropical Pacific Ocean clade, comprised of whitebelly, eastern, Central American and Tres Marias populations. The node separating these two major clades was well supported (posterior probability of 1). Within this ETP clade, whitebelly spinners were well supported as the most basal lineage, but the node grouping eastern, Tres Marias and Central American was not well supported (posterior = 0.5296).

The second major clade contains samples from three ocean basins: North Atlantic, Indian Ocean, and Central and Western Pacific Ocean. Most of the nodes within this

clade were well supported, the exception being the node that grouped the Maldives samples with the Philippines samples (posterior = 0.456).

Most of the branches in the spinner dolphin tree had low theta indicating low diversity in these lineages. Three lineages had high theta indicating higher levels of diversity. These included: the ETP branch, the Indian Ocean and Western Pacific branch, and the Indonesian branch of the dwarf subspecies clade.

The spotted dolphin dataset was analyzed for a total of 414,000 MCMC chains. The trace of the posteriors resulting from MCMC sampler showed some variability, potentially indicating a lack of convergence on a topology (Figure A3-S8). In addition, the ESS was only 7.6704 –indicating the need for additional MCMC iterations. We should have run the MCMC for much longer than we did (up to 1M chains, for complete confidence), but this was not feasible at this time.

The Maximum Clade Credibility tree and visualization of uncertainty using all the trees is shown in Figure 6. Aside from two of the terminal nodes, the consensus tree was well supported. The topology supports the current taxonomy, with the endemic coastal ETP subspecies as the most basal lineage, distant from its sister clade that included all other populations (the nominate subspecies). Offshore ETP spotted dolphins were the next clade in, followed by Hawaii and an Indo-west Pacific clade. Relationships among

these Indian Ocean and western Pacific populations are poorly resolved as indicated by the low nodal support values.

The ancestral branch that lead to the *S. a. attenuata* clade, and the ETP Offshore branch (the most basal branch within the *S. a. attenuata* clade) both had high theta values indicating higher genetic diversity within these branches.

DISCUSSION:

Diversity Estimates

We were somewhat surprised to see the overall similarity in levels of heterozygosity in spinner dolphin, knowing that the ETP populations have much larger abundance that the other island-associated or coastal populations. Only two of eleven populations were outliers in terms of genetic diversity: Hawaii spinners and dwarf spinner dolphins from Australia. Hawaii is known to be a small, isolated population with reduced genetic diversity (Andrews et al. 2010, Andrews et al. 2013, Leslie et al. *In Review*). It is not clear why the Northern Australia population of dwarf spinner dolphins (collected offshore in the Timor Sea) would have reduced genetic diversity. There is recent evidence of a coastal population of dwarf spinner dolphins in NW Australia, but it is uncertain if this would be connected with those in the Timor Sea (Allen et al. 2012).

Dwarf spinner dolphins reside in shallow habitat compared to other pantropical spinner dolphins, so it is possible that this habitat can only support a relatively low abundance, which would ultimately decrease heterozygosity and increase genetic drift.

Northern Australia samples were taken as fisheries bycatch and it is possible that some of the animals were related; however, we used samples that were killed in at least eight different incidents and locations. Moreover, eleven of the thirteen samples were males, decreasing the likelihood that samples represented mother-calf mortalities. It is unlikely that relatedness or skew in sex ratio of our Northern Australia sample resulted in the decreased heterozygosity observed.

Several populations had individual heterozygosity outliers. All but one of these outliers were lower than expected heterozygosity. Causes of low heterozygosity could be biological (i.e., inbreeding) or a result of allelic dropout during data collection (Davey et al. 2013, Gautier et al. 2013). For the eastern spinner dolphin with very high heterozygosity, this could also be biological (i.e., admixture) or a result of sequencing error (although the later seems unlikely given the quality filtering).

Spotted dolphins have lower heterozygosity than spinner dolphins and also varied minimally among populations. The offshore ETP population had higher mean heterozygosity, likely due to the larger population size compared to others. Coastal

spotted dolphins also had a slightly elevated heterozygosity, again likely due to a slightly higher population abundance in that population compared to the other island associated spotted dolphins in the dataset. Hawaii and Guam/NMI had very similar mean and distribution of heterozygosity indicating they could be comparable populations in terms of genomic diversity. Although they seam comparable to the Hawaiian and Guam/NMI populations in terms of genomic diversity, additional samples for the Indonesian and Maldivian populations will be needed to better characterize them.

Population Structure: F_{ST}

Spinner Dolphins

Almost all pairwise comparisons in the spinner and spotted dolphin datasets were significant, indicating subdivided populations despite high dispersal potential of individual animals. The largest estimates of pairwise F_{ST} between populations of spinner dolphins involved the Australian population of dwarf spinner dolphins (all estimates were > 0.1468). These results were corroborated by our DAPC analyses that indicated Australian dwarf spinners occupy a very different PC space from the other subspecies and populations. In the pairwise comparison with the other dwarf population (Indonesia), F_{ST} was larger (0.1468) than between the two morphologically distinct endemic ETP subspecies (CA and eastern: 0.0095). These results were unexpected given the morphological similarity between Australian and Indonesian dwarf spinner dolphins(Perrin et al. 1999). However, this pattern of strong genetic structure across the

"Marine Wallace's Line" (sensu Barber et al. 2000) has been shown in two recent studies of coastal cetaceans in this same area (Beasley et al. 2005, Jefferson and Rosenbaum 2014). Collectively, these studies indicate that the Marine Wallace's Line is a significant biogeographic barrier despite the high dispersal abilities of marine mammals.

Other comparisons between populations had F_{ST} values ranging from approximately 0.01 to 0.07. The ETP populations were the exception to this, in that the F_{ST} estimates were very low (i.e., <0.01). As aforementioned, the pairwise F_{ST} comparison between ETP endemic subspecies was an order of magnitude lower than the metric between our two populations of dwarf spinner. Low F_{ST} is not surprising, as previous studies have shown that the ETP subspecies are very closely related. However, for the only direct comparison of morphologically distinct subspecies it is surprising to get such relatively small F_{ST} values relative to pairwise comparisons of populations from different subspecies. This is interesting from a perspective of evolutionary dynamics and we agree with Andrews et al. (2013) that additional work on the possibility that positive selection is leading to the morphological differences could be a rich area of inquiry.

One of the few pairwise comparisons that was not significantly different from zero was between the Tres Marias and the eastern spinner dolphins. This was also the lowest F_{ST} we measured in this dataset. See Leslie and Morin (Chapter 3) for a comprehensive discussion of ETP population comparisons. Although many of our

populations outside the ETP had low sample sizes, simulations have shown the ability to precisely estimate F_{ST} to detect allele frequency differences when using a large number of genetic markers (i.e., >1000), even from small sample sets (n > 4-6) (Willing et al. 2012).

The one population in our study with fewer than four samples (Tanzania) showed non-significant differences when compared to three other groups (whitebelly spinners, Maldives, and Philippines). It is possible that some connection exists between the most proximate populations – Tanzania and Maldives – but we believe all of these results are likely untrustworthy due to the small sample size for spinner dolphins from Tanzanian waters.

Spotted Dolphins

F_{ST} comparisons for spotted dolphins indicated highly structured populations (Table 4). Interestingly, the pattern of F_{ST} point estimates was similar to the spinner dolphins in that pairwise comparison of the coastal ETP subspecies with the offshore ETP population of pantropical spotted dolphins had reduced F_{ST} values compared to pairwise comparisons between other global populations. These results reinforce previous studies that have shown the permeability of the inshore/offshore boundary in the ETP (Escorza-Trevino et al. 2005). One of the lowest estimates was between Hawai‘i and Guam/NMI, suggesting structure between these two western Pacific island archipelagoes.

We hesitate to put much confidence in the F_{ST} comparisons involving the Indonesian spotted dolphins, given that this sample set was small (n = 3). We also did not include the data from the single Maldives spotted dolphin sample (n = 1) in pairwise F_{ST} comparisons.

Spinner Dolphin Population Structure: DAPC

Eastern Tropical Pacific

Overall, the Discriminant Analysis of Principle Components (DAPC) showed good separation between most populations of spinner dolphins. On one end of the diversity was a cluster of the endemic ETP subspecies. CA and eastern were very close together with the Tres Marias Island spinner dolphins in between. As expected, the whitebelly spinner was also near the ETP endemics – about half way between these subspecies and the "core", which included mostly Indo-Pacific populations. This is what we would expect given the hypothesis that it is a hybrid swarm between eastern and central Pacific pantropical spinner dolphins. Ongoing long-distance gene flow might be keeping the whitebelly genome similar to the pantropical subspecies. Alternatively, the WB form is the ancestral form to the ETP endemic subspecies; similarities in

morphology between WB and the pantropical subspecies would be synapomorphies, whereas the similarities shared between WB and eastern could be due to more recent ancestry and ongoing geneflow. Finally, the WB spinners could represent reconnection between subspecies separated and independent for some time and only recently reconnecting.

We saw little evidence in our analyses of connection between the ETP spinner groups (including WB) and the central and western Pacific pantropical subspecies. Unfortunately, our sampling in the central Pacific was geographically sparse and may not have included the WB spinner's most closely related island associate populations. If there were going to be connections with the WB they might come from south of Hawaii. Added data from these regions could be informative with regard to the level of connection between WB and the central Pacific Island-associated populations.

Atlantic

We were surprised by the placement of the Atlantic Ocean samples, very close to the whitebelly spinner samples and near the core of the DAPC diversity of all samples. Given the theories put forth by Davies (1963) and Perrin (2007) we would have expected to see it nearer to Tanzania and very distant from the ETP populations. The Panamanian uplift occurred ~3.5Ma (Coates et al. 1992), although newer evidence suggests it may be considerably older at 15-20Ma (Bacon et al. 2015). The age of spinner dolphins,

estimated using molecular clock techniques on a supermatrix consisting of 42,335 characters from mtDNA and nuDNA loci, is between 1.6 and 3.97 million years old (mean 2.75Ma; McGowen et al. 2009). From these estimate, we can infer that Panama would have been a significant barrier to spinner dolphins in the infancy of the species and should have lead to significant drift between eastern tropical Pacific dolphins and northeastern Atlantic spinner dolphins. If Atlantic and ETP spinner dolphins have been separated for at least 3.5M years, we would expect the Atlantic genotype frequencies to have drifted very far from the ETP frequencies. The only plausible reasons why this would not happen are ongoing or periodic pantropical geneflow (completely circumtropical geneflow) or incomplete lineage sorting. The former seems implausible given the barriers discussed. The latter involves older lineages harboring genomic variation from a large ancestral population that has not yet drifted apart despite separation. Neither of these hypotheses adequately explains the positioning of the Atlantic genotypes near the whitebelly spinner group in genomic PC space. Of course, these explanations are not mutually exclusive.

If large populations harboring diverse genomes continued to disperse around southern Africa and the east Pacific basin, it is possible that similarities could remain even with low levels of ongoing or sporadic gene flow. Atlantic spinner dolphins tend to be more pelagic (found exclusively in water of >2000m) and form larger groups than the island-associated central Pacific spinner dolphins (Waring et al. 2015). They are also much lower in abundance in the Atlantic Ocean (11,971 in the Gulf of Mexico) than the

eastern tropical Pacific (Mullin and Fulling 2004) which could have led to lower heterozygosity, but would have also increased the genetic drift. Unfortunately, abundance estimates for the western north Atlantic coast of the United States are not available because sightings and strandings are rare along the continental shelf (Waring et al. 2015), so it is difficult to tell if the greater Atlantic population would have been subject to a similar level of drift or if it would have been buffered from drift by large abundance. These are factors to consider as additional work brings or clarity to our knowledge of ocean basin-level relationships in spinner dolphins.

Alternatively, this could be a timing issue. If the spinner dolphin lineage is older than the Isthmus of Panama, the east Pacific basin could have been a more important barrier to geneflow in the infancy of the species. Further, if the historical ETP + western Atlantic population was large, genomic signatures could take a long time to separate via drift. An older, super abundant ETP + western Atlantic population might result in the pattern we observe in the DAPC results. To better understand these relationships we put our data into a phylogenetic framework to test various hypotheses (see below).

Indo-Western Pacific Ocean Core

We were also surprised by the amount of similarity between geographically disparate populations that made up the central core group in our DAPC. These included, western tropical Pacific (Philippines), Western Indian Ocean (Tanzania), and Central

Indian Ocean (Maldives). Separated by up to 5441km, these populations have remained genetically similar. Interestingly, the Maldives animals appear to be the most closely related to the dwarf subspecies. Galver (2002), using some of the Maldives samples we used in this study, found mtDNA haplotypes shared with the dwarf spinner populations. This, coupled with our nuDNA data, suggest possible genetic connection in the Maldives between the nominate subspecies and the dwarf subspecies.

The Hawaiian genotypes were most closely aligned to this Indo-Western Pacific core group of populations. Interestingly, it was the second eigenvalue that spread the Hawaiian genotypes away from the other populations. Despite the reduced heterozygosity in Hawaiian spinner dolphins, the population covered a lot of space in the DAPC.

Dwarf spinner

Finally, opposite the ETP arm in our linear genotype array - spread by the first eigenvalue - are the two populations of dwarf spinner dolphins: Indonesia and Northern Australia. The Indonesian dwarf spinner overlaps somewhat in genotype space with the Indo-western Pacific populations, and has one individual midway between the core and the Northern Australia cluster. This could indicate a potential area of admixture between the western Pacific and Indian Ocean populations and the dwarf populations, or it could indicate a cline of diversity from these Indo-Pacific pantropical groups through the Indonesian Archipelago and south toward the genetically distant Australian dwarf

spinners. Recall also that the Australia dwarf spinners had much reduced heterozygosity. In combination with the results of the DAPC, this could be a sign that the Northern Australian population is small and subject to high levels of genetic drift. Abundance estimates are unknown (Allen et al. 2012). Allen et al (2012) noted two groups of small dolphins in northwestern Australia that fit the phenotype of dwarf spinner dolphins and noted that this would be the southwestern most record of this subspecies. Genetic analysis is underway to determine the subspecific identity and test the hypothesis that they are a unique population (D. Cagnazzi – Pers. Comm.)[4] As noted by Allen et al. (2012), "more comprehensive data on the biodiversity and population ranges of tropical dolphins are required for north-western Australia."

Barber et al. (2000) proposed extending the terrestrial biogeographical break between Western Indonesia and eastern Indonesia/Australia to the marine environment based on patterns of multiple of reef-dwelling species. Recent work on the systematic relationships of small shallow-water dolphins highlighted the importance of this marine Wallace's line as a biogeographic break. A new species of Irrawaddy dolphin (genus *Orcaella*) was described east of Wallace's line (Beasley et al. 2005) and recently a new species of the genus *Sousa* was found to be separate from the Indo-Pacific species by a wide distributional gap that coincides with Wallace's Line (Jefferson and Rosenbaum 2014).

[4] Daniele Cagnazzi Ph.D., School of Environment, Science and Engineering, Southern Cross University, Lismore NSW 2480 Australia

Certainly, efforts should be made to survey the abundance of these individuals and better characterize them. Our sample set for northern Australia dwarf spinners was skewed heavily toward male dolphins (11 of 13 were males), but Y chromosome linked SNPs would represent only ~1/44th the SNPs – a very small fraction of the overall data. Therefore, this is unlikely to contribute to the uniqueness of this population. Given our very high significance in pairwise comparisons using F_{ST} and the distance in genotype spaced detected with DAPC, this Australian population of dwarf spinner dolphins (from the Timor Sea, north of Australia) could warrant a unique stock of dwarf spinner or a unique subspecies. Perrin et al (1999) erected the subspecies *S. l. roseiventris* based on detailed morphometric analyses. They examined specimens from the Timor Sea and found them to be consistent with the dwarf spinner dolphins from the Gulf of Thailand. The only notable difference was northern Australia dwarf spinner dolphins had fewer vertebrae than those from the Gulf of Thailand. These authors did not examine dwarf spinner dolphins from Indonesia, but speculated that these animals inhabit shallow coastal waters of much of Southeast Asia. Further research effort should be applied to these areas to determine the geographic range and genetic diversity of dwarf spinner dolphins in Southeast Asia.

Spotted Dolphin Population Structure: DAPC

Similar to the spinner dolphin DAPC, most of the spotted dolphin populations were distributed along the first eigenvalue. The coastal ETP populations represented one extreme of genotype space, while the western Pacific populations represented the other. All populations in this analysis had little to no overlap in the 67% ellipses, suggesting that these all represent unique isolated populations. The only exception was strong overlap from the Maldivian sample (n = 1) within the ellipse of the Guam/NMI population. It is difficult to draw strong conclusions based on one sample, but further studies should focus on expanding the sample of spotted dolphins in the Indian Ocean to test the uniqueness of the Maldivian population. Hawaiian spotted dolphins spread out on the second eigenvalue.

Phylogeography:

We present unrooted and midpoint-rooted trees in Figures 5 and 6. Although, there are methods for inferring the location of the root (Lemey et al. 2009) based on calculating the posterior probability of the root location, these options are not available for SNP data presently. Moreover, using outgroups in SNAPP can create long branches which make estimating parameters difficult and violates the assumptions of the Yule prior used in SNAPP (R. Bouckaert – *Pers. Comm*)[5]. Midpoint rooting often results in the root being assigned to the longest branch, which is the case in both of our species. This

[5] Remco Bouckaert – Research Fellow, Dept. of Computer Science, The University of Auckland, Auckland, NZ. http://tinyurl.com/j7uwq83

may not, however, reflect reality. As best as possible, we have avoided making inferences dependent on the location of the root. Finally, it is worth noting that phylogeographic patterns in nuDNA trees could represent shared ancestry, admixture (i.e., genetic exchange between populations) or a combination of both.

Spinner dolphin phylogeography

Davies (1963) proposed phylogeographic hypotheses for circumtropical dolphins that included an Indo-Pacific core, and satellite lineages stretching into the Atlantic to the west, and the eastern Pacific to the east. Perrin (2007) argued for a similar arrangement, but noted that the southern Africa barrier was probably only passable from east to west during warm regimes. Both ideas included an east Pacific basin barrier, yet our modern conception of spinner dolphin relationships within the Pacific Ocean (informed by both morphology and genetics) says there is frequent genetic connection between the central and eastern Pacific via the whitebelly hybrid (Perrin et al 1990; Andrews et al 2013). Andrews et al. (2013), in the only genetic study of global spinner dolphin relationships, found “porous” genetic boundaries between the ETP and pantropical subspecies. They suggest that neutral genes have moved freely between groups, but genes subject to divergent natural selection should show differences. Evidence for this was from a shared Y-chromosome haplotype in the eastern and Central American subspecies that was not found in Gray's or the dwarf subspecies. Interestingly, this region of the Y-chromosome was found to be polymorphic in the whitebelly form (Andrews et. al., 2013).

We did not separate neutral from selected loci, and therefore cannot speak to the contrasting patterns Andrews et al (2013) saw in different marker types with regard to ETP spinner dolphins. However, the longest branch in our Bayesian species tree was between the eastern tropical Pacific Ocean populations and the rest of the global distribution (Figure 5). This suggests that the eastern Pacific basin – an expanse of 4000-7000km without coastal habitat – has been a formidable biogeographic barrier since early in this species' history (see Briggs 1961). Our results also showed high support for WB spinners as sister to the endemic ETP subspecies, as opposed to the WB being sister to the pantropical lineage. Endemic ETP subspecies were grouped together with Tres Marias with low support; this is not necessarily surprising given probability of genetic interchange among these groups (see Chapter 3).

The sister clade to the ETP group includes all other pantropical and dwarf populations. Interestingly, the Atlantic branch is again not where we would expect it to be given the vicariant step-stone hypotheses of Davies (1963) and Perrin (2007). Based on the ideas of these authors we would expect the Atlantic branch to be nested well within the clade including Tanzania (labeled ZAN in Figure 5). Instead, it is positioned at the base of this clade well outside the clade including Tanzania. This again suggests a deeper relationship with the ETP clade and not a history of vicariant step-stoning west from the Pacific, Indian and ultimately to the Atlantic. The ETP and Atlantic Ocean branches are the longest in this tree, so could be attracted for this reason. We doubt the relationship of

these two lineages is a result of long branch attraction, because we would expect Atlantic and ETP to be sister taxa within a clade, which they are not.

Possible reasons for this topology include separation of a historically super-abundant ETP + Atlantic populations that has limited the effect of genetic drift. Alternatively, it could be that the age estimate for spinner dolphins is too young and that the east pacific basin was a much more formidable barrier to geneflow prior to the uplift of the Panamanian Isthmus. It is worth noting that Andrews et al (2013) did include some Atlantic specimens and there was no evidence of connection between ETP spinners and Atlantic spinners in their results.

Much of our ability to test the hypothesis of east-west pan-tropical stepwise dispersal relies on the evolutionary history of the Indian Ocean populations. If early stepwise dispersal occurred, leading to an Atlantic population in the Caribbean, followed by extinction within the Indian Ocean, and subsequent recolonization events, the tight connection between Atlantic and Indian that anchor our null hypotheses of ocean-basin phylogenetics might have been erased. The inclusion of additional samples for the eastern Atlantic and additional Indian Ocean localities might help resolve this issue.

Most of the clades within the large pantropical spinner + dwarf spinner dolphin clade were well supported (posterior >0.97). The most basal of these clades includes

Hawaii as sister to all other Indian Ocean and western Pacific Ocean populations (including dwarf spinner dolphins). Several authors have discussed the genetic evidence for the independence of Hawaiian spinner dolphins (Andrews et al. 2010, Andrews et al. 2013, Leslie et al. *In Review*); given our results, this population should be considered a distinct population segment. The next clade links the dwarf spinner dolphin subspecies clade with a clade of Indian Ocean populations and a Southeast Asia population (Philippines). Despite showing large differences in the DAPC analysis, these two populations still group together phylogenetically (as suggested by morphological analyses – see Perrin et al 1999).

The eastern tropical Pacific is known to harbor historically abundant spinner dolphin pelagic populations (Wade et al. 2007), so the high theta value for the ancestral branch is not surprising here. The large theta estimated for the Tanzania + Philippines + Maldives clade (0.2132) was somewhat surprising, however. High diversity in the ancestor of this clade could have resulted from a large western Pacific + Indian Ocean population, or this could be indicative of a metapopulation with locally adapted subpopulations throughout this portion of the range (Tanzania: 0.1188; Maldives: 0.1147; Philippines: 0.1149). Finally, the theta estimate for the Australian population of the dwarf spinner indicates very low diversity (0.0019) in contrast the Indonesian population dwarf spinners, which had much higher theta (0.1564). It is possible that the Indonesian population of dwarf spinners is a structured population that inhabits the many islands of the Indonesian archipelago. Moreover, it is also possible (given the DAPC results shown

above) that there is some genetic overlap between the pantropical and dwarf spinner dolphin in Indonesia serving to increase theta via admixture.

Spotted dolphin phylogeography

For other marine taxa, the east pacific basin is a major biogeographic barrier (Briggs 1961). Due to spatial proximity of the ETP spotted dolphin populations, the east Pacific basin should be a stronger biogeographic barrier than the inshore-offshore divide. Our results, however, suggest that it is not as significant a barrier for spotted dolphins as the inshore-offshore divide in the eastern tropical Pacific. Our results support the notion that the nominate pantropical spotted dolphin subspecies (*S. a. attenuata*) resides in the offshore ETP, and that the ETP has an coastal endemic subspecies (Perrin et al 1994, (Chapter 3, Escorza-Trevino et al. 2005, Leslie et al. *In Review*).

We could not – with confidence - resolved relationships among the Indo-Pacific populations of spotted dolphins. This likely reflects the biological reality as hypothesized by Davies (1963) and Perrin (2007) that recent divergences and possible ongoing gene flow in this region make testing phylogeographic hypotheses difficult.

Theta estimates for spotted dolphins indicate that the offshore ETP population has high diversity – likely due to high historical abundance – and the clade that lead to the diversification of pantropical spotted dolphins had high diversity. This could be a result of large global population size prior to population substructuring or dispersal.

CONCLUSIONS

In spinner dolphins, the east Pacific basin appears to be an important biogeographic barrier to these highly mobile pelagic species. Our results indicate deep separation between eastern Pacific clades and other conspecifics in these species.

Although we were not able to test relationships with Atlantic populations of spotted dolphins, our results support the current taxonomy and indicate a stronger separation at the inshore-offshore boundary in the ETP.

Both spinner and spotted dolphins appear to have a tightly clustered Indo-Pacific core. Relationships from these areas were difficult to resolve, likely because of recent divergences, incomplete lineage sorting, and/or ongoing gene flow.

For spinner dolphins our results do not support the connection between the Atlantic Ocean and the Indian Ocean via the South African species gate (Perrin 2007).

Instead, our results suggest the northwestern Atlantic Ocean population sit between the pantropical spinner dolphin clade and the ETP clade, instead of at the crown of the pantropical spinner clade, where it would be expected via an Indian Ocean–Atlantic stepping stone model of dispersal. The pattern we observe could represent a very deep connection (prior to the uplift of panama) of a historically super abundance population. The vicariance caused by the uplift should have caused the populations to drift apart, but high abundance and heterozygosity could have retained similarities. Alternatively, multiple extinctions and reinvasions could cause the pattern we see.

We found the dwarf spinner dolphins from northern Australia to have very reduced diversity and unique genomic signature. It is clear there is a connection phylogenetically with the dwarf spinner dolphins from Indonesia. However, based on our results, the northern Australia dwarf spinner dolphin should be considered a different population if not a different subspecies. This could be another example of the Marine Wallace's Line (Barber et al 2000) in a shallow-water small cetacean (Beasley et al. 2005, Jefferson and Rosenbaum 2014).

Our analyses indicated that the offshore spotted dolphin in the eastern tropical pacific is part of the pantropical subspecies. Subsequent studies should include the Atlantic pantropical spotted dolphin samples and samples from northern Australia should be added to test the circumtropical biogeographic hypotheses in this species.

ACKNOWLEDGEMENTS:

We especially thank Dr. William F. Perrin who provided guidance and support throughout the development of this project. Drs. Andy Foote and Eric Archer provided valuable guidance with analyses. Gabriela Serra-Valente, Victoria Pease, Kelly Robertson, Dr. Louella Dolar, Nicole Beaulieu and Morgane Lauf helped in the laboratory and NMFS Marine Mammal and Sea Turtle Research (MMASTR) Collection. We thank Dr. Robin Baird of Cascadia Research Collective for permission to use Hawaiian samples, Dr. Erin Oleson of the NOAA Pacific Islands Science Center for permission to use Hawaiian, NMI, and Guam samples, and Dr. Danielle Krebs for use of some of the Indonesian samples. Support for this research was also provided from Drs. Lisa Ballance and Barbara Taylor of the Marine Mammal and Turtle Division of the SWFSC. Finally, we offer special thanks to Jeremy Rusin of SWFSC for his logistical and administrative help.

Chapter 4 is currently being prepared for submission for publication of the material. Leslie, Matthew S.; Morin, Phillip A. The dissertation author was the primary investigator and author of this material.

Literature Cited:

Allen, S. J., D. D. Cagnazzi, A. J. Hodgson, N. R. Loneragan, and L. Bejder. 2012. Tropical inshore dolphins of north-western Australia: Unknown populations in a rapidly changing region. Pacific Conservation Biology **18**:56-63.

Andrews, K. R., L. Karczmarski, W. W. L. Au, S. H. Rickards, C. A. Vanderlip, B. W. Bowen, E. Gordon Grau, and R. J. Toonen. 2010. Rolling stones and stable homes: social structure, habitat diversity and population genetics of the Hawaiian spinner dolphin (*Stenella longirostris*). Molecular Ecology **19**:732-748.

Andrews, K. R., W. F. Perrin, M. Oremus, L. Karczmarski, B. W. Bowen, J. B. Puritz, and R. J. Toonen. 2013. The evolving male: spinner dolphin (*Stenella longirostris*) ecotypes are divergent at Y chromosome but not mtDNA or autosomal markers. Molecular Ecology **22**:2408-2423.

Archer, E. 2015. strataG: Summaries and Population Structure Analyses of Haplotypic and Genotypic Data. R package

Bacon, C. D., D. Silvestro, C. Jaramillo, B. T. Smith, P. Chakrabarty, and A. Antonelli. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proceedings of the National Academy of Sciences **112**:6110-6115.

Beasley, I., K. M. Robertson, and P. Arnold. 2005. Description of a new dolphin, the Australian snubfin dolphin *Orcaella heinsohni sp. n.* (Cetacea, Delphinidae). Marine Mammal Science **21**:365-400.

Bouckaert, R., and J. Heled. 2014. DensiTree 2: Seeing Trees Through the Forest. bioRxiv:012401.

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

Bowen, B. W., A. L. Bass, L. A. Rocha, W. S. Grant, and D. R. Robertson. 2001. Phylogeography of the trumpetfishes (Aulostomus): ring species complex on a global scale. Evolution **55**:1029-1039.

Briggs, J. C. 1961. The east Pacific barrier and the distribution of marine shore fishes. Evolution **15**:545-554.

Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. Molecular Biology and Evolution.

Coates, A. G., J. B. C. Jackson, L. S. Collins, T. M. Cronin, H. J. Dowsett, L. M. Bybell, P. Jung, and J. A. Obando. 1992. Closure of the Isthmus of Panama: The near-shore marine record of Costa Rica and western Panama. Geological Society of America Bulletin **104**:814-828.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and G. P. A. Group. 2011. The variant call format and VCFtools. Bioinformatics **27**:2156-2158.

Davey, J. W., T. Cezard, P. Fuentes- Utrilla, C. Eland, K. Gharbi, and M. L. Blaxter. 2013. Special features of RAD Sequencing data: implications for genotyping. Molecular Ecology **22**:3151-3164.

Davies, J. L. 1963. The Antitropical Factor in Cetacean Speciation. Evolution **17**:107-116.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution **29**:1969-1973.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE **6**:e19379.

Escorza-Trevino, S., F. I. Archer, M. Rosales, A. Dizon, and A. Lang. 2005. Genetic differentiation and intraspecific structure of Eastern Tropical Pacific spotted dolphins, *Stenella attenuata*, revealed by DNA analyses. Conservation Genetics **6**:587-600.

Gautier, M., K. Gharbi, T. Cezard, J. Foucaud, C. Kerdelhué, P. Pudlo, J. M. Cornuet, and A. Estoup. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. Molecular Ecology **22**:3165-3178.

Glaubitz, J. C., T. M. Casstevens, F. Lu, J. Harriman, R. J. Elshire, Q. Sun, and E. S. Buckler. 2014. TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. PLoS ONE **9**:e90346.

Hedrick, P. W., and C. Goodnight. 2005. A Standard Genetic Differentiation Measure. Evolution **59**:1633-1638.

Jefferson, T. A., and H. C. Rosenbaum. 2014. Taxonomic revision of the humpback dolphins (*Sousa* spp.), and description of a new species from Australia. Marine Mammal Science **30**:1494-1541.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics.

Lambertsen, R. H. 1987. A biopsy system for large whales and its use for cytogenetics. Journal of Mammalogy **68**:443-445.

Lemey, P., A. Rambaut, A. J. Drummond, and M. A. Suchard. 2009. Bayesian phylogeography finds its roots. PLoS Comput Biol **5**:e1000520.

Leslie, M. S., F. I. Archer, and P. A. Morin. *In Review*. Genetic inference of population structure in spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) from the eastern tropical Pacific Ocean based on mitochondrial genomes and nuclear SNPs. Marine Mammal Science.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics **25**:1754-1760.

Lindblad-Toh, K., M. Garber, O. Zuk, M. F. Lin, B. J. Parker, S. Washietl, P. Kheradpour, J. Ernst, G. Jordan, and E. Mauceli. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. Nature **478**:476-482.

Lischer, H., and L. Excoffier. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics **28**:298-299.

McGowen, M. R., M. Spaulding, and J. Gatesy. 2009. Divergence date estimation and a comprehensive molecular tree of extant cetaceans. Molecular phylogenetics and evolution **53**:891-906.

Miller, S. A., D. D. Dykes, and H. F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic acids research **16**:1215.

Mullin, K. D., and G. L. Fulling. 2004. ABUNDANCE OF CETACEANS IN THE OCEANIC NORTHERN GULF OF MEXICO, 1996–2001. Marine Mammal Science **20**:787-807.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences:3321-3323.

Palumbi, S. R. 1996. What can molecular genetics contribute to marine biogeography? An urchin's tale. Journal of Experimental Marine Biology and Ecology **203**:75-92.

Perrin, W. F. 1990. Subspecies of *Stenella longirostris* (Mammalia: Cetacea: Delphinidae). Proceedings of the Biological Society of Washington **103**:453-463.

Perrin, W. F. 2001. *Stenella attenuata*. Mammal Study **683**:1-8.

Perrin, W. F. 2007. The South African species gate. Whales and dolphins of the southern African subregion. Ed. Best, P. B.

Perrin, W. F., M. L. L. Dolar, and D. Robineau. 1999. Spinner dolphins (*Stenella longirostris*) of the western Pacific and Southeast Asia: Pelagic and shallow-water forms. Marine Mammal Science **15(4)**:1029-1053.

Perrin, W. F., and A. A. Hohn. 1994. Pantropical spotted dolphin (*Stenella attenuata*). Pages 71-98 *in* S. H. Ridgway and R. Harrison, editors. Handbook of Marine Mammals. Academic Press, San Diego, CA.

Perrin, W. F., G. D. Schnell, D. J. Hough, J. W. Gilpatrick, Jr., and J. V. Kashiwada. 1994. Reexamination of geographic variation in cranial morphology of the pantropical spotted dolphin, (*Stenella attenuata*), in the eastern Pacific. Fishery Bulletin **92**:324-346.

R Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A. 2014. FigTree v1. 4.2: Tree figure drawing tool. Available at http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A., M. A. Suchard, D. Xei, and D. A. J. 2014. Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer.

Wade, P. R., G. M. Watters, T. Gerrodette, and S. B. Reilly. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. Marine Ecology Progress Series **343**:1-14.

Waring, G. T., E. Josephson, K. Maze-Foley, and P. E. Rosel. 2015. US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments--2014. NOAA Technical Memorandum NMFS-NE-231 **231**:361.

Willing, E.-M., C. Dreyer, and C. van Oosterhout. 2012. Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. PLoS ONE **7**:e42649.

Wright, S. 1951. The genetical structuring of populations. Annals of Eugenics **15**.

Figure 1. Sampling localities for spinner dolphins. Pale blue triangles indicate Central American spinners. Green squares are the proposed Tres Marias spinners. Orange diamonds indicate whitebelly spinners, a proposed intergrade between the pantropical or Grey's spinner (purple diamonds) and the eastern subspecies (blue squares). Red circles indicate dwarf spinner dolphin locations.

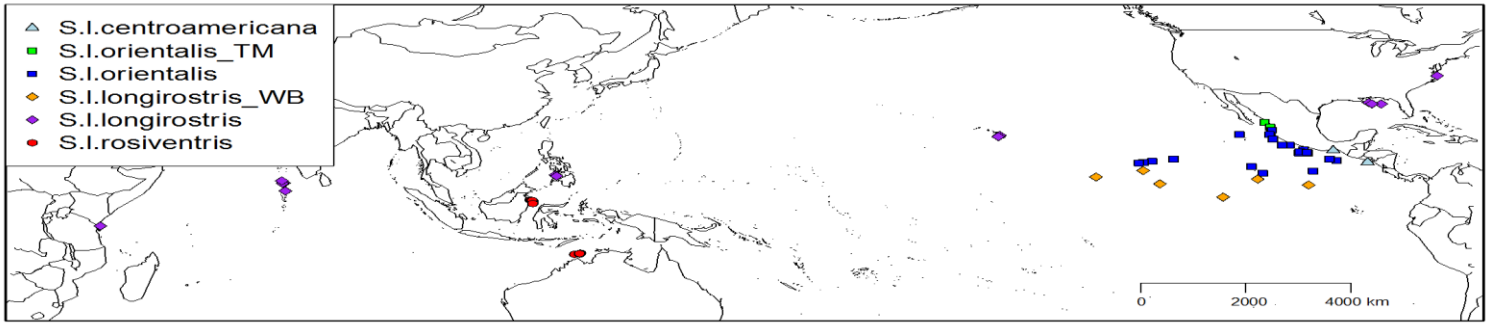


Figure 2. Sampling localities for spotted dolphins. Coastal spotted dolphins (*S. a. graffmani*) are blue squares, pantropical spotted dolphins (*S a. attenuata*) are in yellow: the Hawaiian population is shown in yellow triangles; Guam and Northern Marianas Islands (NMI) are shown as inverted yellow triangles; Indonesian is labeled with standard yellow triangle; Maldives is labeled with yellow squares.

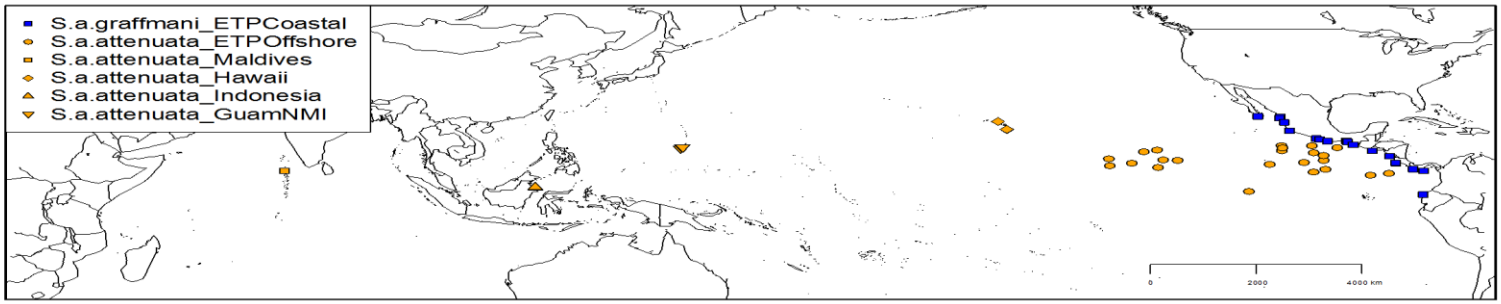


Table 1. Population genetic summary statistics for spinner dolphins based on 8,268 SNPs. *H* (sd) is the mean heterozygosity (and standard deviation) per site across all individuals in each population. Numbers of samples show for *n* (total), F (females), M (males), and U (unknown sex).

Subspecies	**Populations**	***n*/F/M/U**	***H* (sd)**
centroamericana	Central American	9/5/3/1	0.2670 (0.2120)
centroamericana	Tres Marias	12/6/6/0	0.2650 (0.1898)
orientalis	eastern	37/18/18/1	0.2833 (0.1606)
longirostris	whitebelly	14/7/7/0	0.2697 (0.1799)
longirostris	Hawaii	6/2/4/0	0.2324 (0.2360)
longirostris	Philippines	9/4/2/3	0.2542 (0.2069)
longirostris	Maldives	6/4/2/0	0.2599 (0.2300)
longirostris	Tanzania	3/0/0/3	0.2531 (0.2799)
longirostris	North Atlantic	4/0/3/1	0.2361 (0.2594)
roseiventris	Indonesia	6/0/0/6	0.2409 (0.2245)
roseiventris	Australia	13/2/11/0	0.1651 (0.2034)
	Overall	119/48/56/15	0.2552 (0.1337)

Table 2. Population genetic summary statistics for spotted dolphins based on 6,391 SNPs. *H* is the mean heterozygosity per site across all individuals in each population. Numbers of samples show for *n* (total), F (females), M (males), and U (unknown sex). NMI is Northern Marianas Islands.

Subspecies	**Populations**	**Subspecies**	***n*/F/M/U**	***H* (sd)**
graffmani	ETP-coastal	*graffmani*	27/11/14/2	0.2442 (0.1755)
attenuata	ETP-offshore	*attenuata*	32/22/7/3	0.2692 (0.1515)
attenuata	Hawaii	*attenuata*	4/2/2/0	0.2275 (0.2578)
attenuata	Guam/NMI	*attenuata*	8/0/0/8	0.2326 (0.2109)
attenuata	Indonesia	*attenuata*	3/0/0/3	0.2253 (0.2781)
attenuata	Maldives	*attenuata*	1/0/0/1	*NA*
	Overall		75/35/23/17	0.2518 (0.1303)

NA = *H* could not be calculated for Maldives, as there was only one sample from this population.

Table 3. Pairwise population genetic differentiation statistics for spinner dolphins from 8,268 SNPs.
F_{ST} below diagonal and p-value above. * is *not* significantly different from zero.

Populations	n	CA	TM	East	WB	HI	Phil	Mal	Tan	NA	Indo	Aus
Central America (CA)	9	-	0.0020	0.0039	0.0009	0.0020	0.0009	0.0020	0.0419	0.0020	0.0009	0.0009
Tres Marias (TM)	12	**0.0084**	-	**0.0649***	**0.0009**	**0.0009**	**0.0009**	**0.0009**	**0.0029**	**0.0009**	**0.0009**	**0.0009**
Eastern (East)	37	0.0095	0.0022*	-	0.0019	0.0009	0.0009	0.0009	0.0489	0.0020	0.0009	0.0009
Whitebelly (WB)	14	**0.0185**	**0.0102**	**0.0048**	-	**0.0009**	**0.0009**	**0.0009**	**0.1538***	**0.0009**	**0.0009**	**0.0009**
Hawaii (HI)	6	0.0929	0.0849	0.0708	0.0658	-	0.0009	0.0039	0.0090	0.0030	0.0009	0.0009
Philippines (Phil)	9	**0.0534**	**0.0441**	**0.0337**	**0.0199**	**0.0650**	-	**0.0009**	**0.2427***	**0.0019**	**0.0139**	**0.0009**
Maldives (Mal)	6	0.0480	0.0394	0.0252	0.0136	0.0665	0.0118	-	0.1148*	0.0060	0.0019	0.0009
Tanzania (Tan)	3	**0.0508**	**0.0382**	**0.0203**	**0.0097***	**0.0570**	**0.0104***	**0.0073***	-	**0.0329**	**0.0229**	**0.0019**
North Atlantic (NA)	4	0.0686	0.0618	0.0467	0.0477	0.1043	0.0649	0.0667	0.0634	-	0.0019	0.0009
Indonesia (Indo)	5	**0.0607**	**0.0521**	**0.0404**	**0.0232**	**0.0709**	**0.0124**	**0.0166**	**0.0138**	**0.0736**	-	**0.0009**
Australia (Aus)	13	0.2497	0.2386	0.1976	0.2012	0.2449	0.1914	0.1979	0.2203	0.2741	0.1468	-

Table 4. Pairwise population genetic differentiation statistics for spotted dolphins from 6,391 SNPs. F_{ST} is below the diagonal and p-value is above. All comparisons are significantly different from zero ($p < 0.05$). Pairwise statistics for comparisons with the Maldives were not included because only one sample from this population had genotype data. NMI is Northern Marianas Islands.

Populations	n	ETP Coastal	ETP Offshore	Hawaii	Guam/NMI	Indonesia
ETP-coastal	27	-	0.0009	0.0009	0.0009	0.0009
ETP-offshore	32	0.0525	-	0.0009	0.0009	0.0060
Hawaii	4	0.1703	0.0529	-	0.0019	0.0309
Guam/NMI	8	0.1533	0.0350	0.0260	-	0.0069
Indonesia	3	0.1744	0.0511	0.0416	0.0179	-

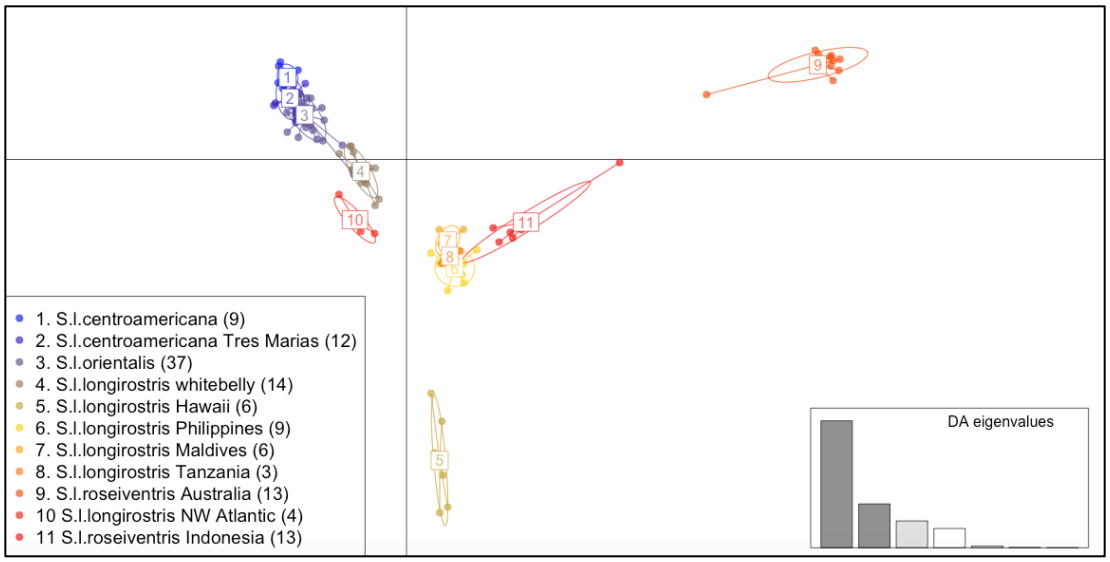


Figure 3. Genomic variation across individuals and populations of spinner dolphins: Scatter plot of individuals based on the first two eigenvalues (created from seven principle components) of the Discriminant Analysis of Principle Components (DAPC). Ellipses represent 67% of the variation for each population. Inset shows the amount of variation represented by the DA eigenvalues.

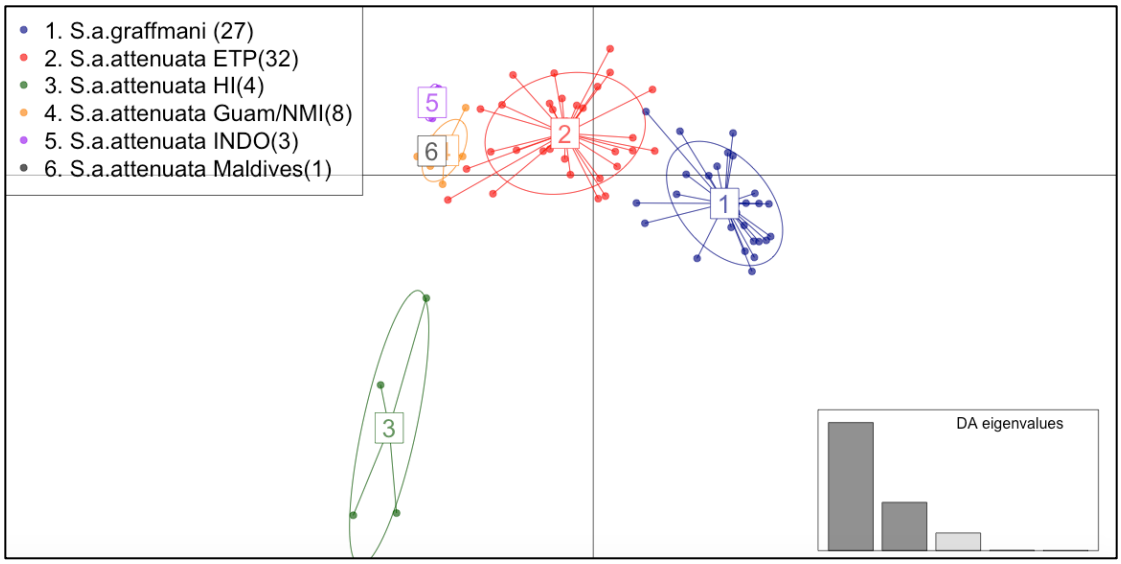


Figure 4. Genomic variation across individuals and populations of spotted dolphins using seven principle components and three discriminant analyses: Scatter plot of individuals based on the first two eigenvalues of the Discriminant Analysis of Principle Components (DAPC). Ellipses represent 67% of the variation for each population. Maldives (6) is only one sample. Inset shows the amount of variation represented by the DA eigenvalues.

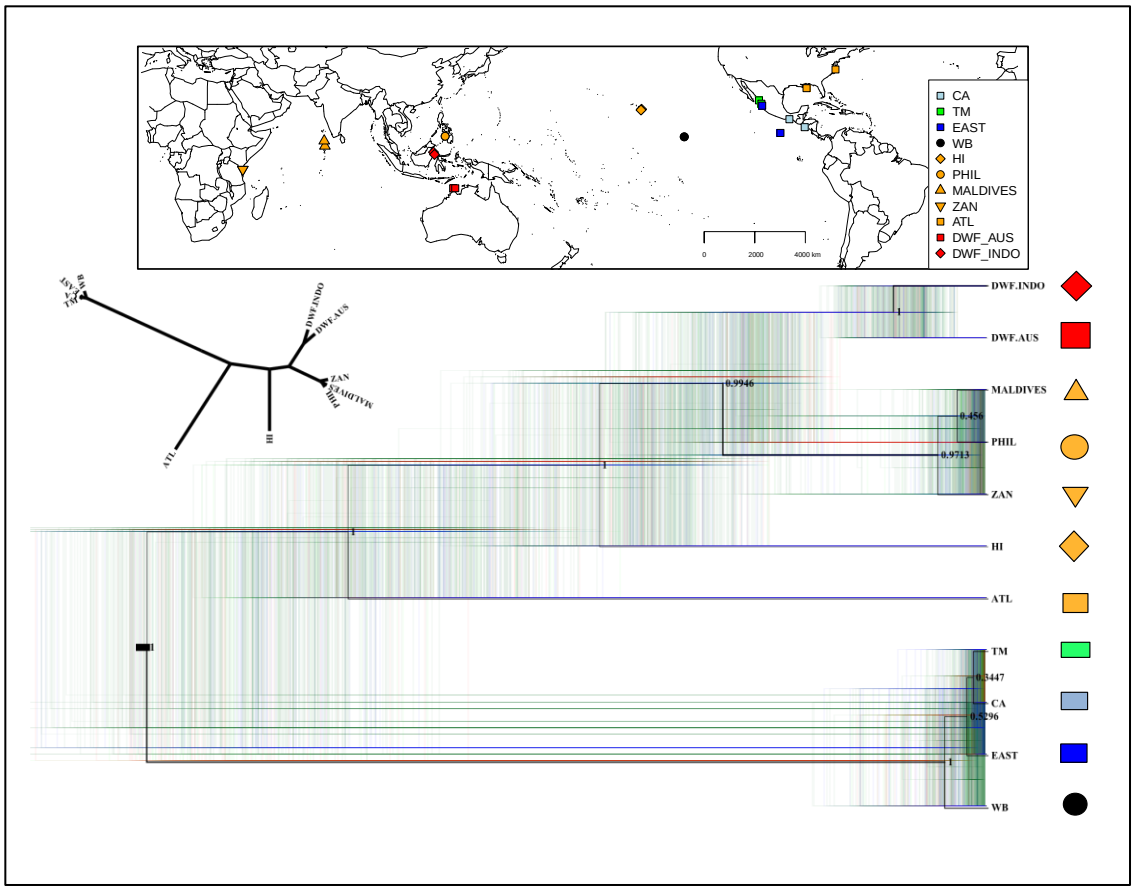


Figure 5. Bayesian species tree for spinner dolphins using 8,268 SNPs. In black is the mid-point rooted Maximum Clade Credibility tree made from all trees with posteriors >0.5. It is overlaid on all trees from the analyses (colored). CA: Central America subspecies (*S. l. centroamericana*) is labeled with a light blue box. EAST: Eastern spinner dolphin subspecies (*S. l. orientalis*) is the dark blue box. WB: whitebelly spinner is the black circle. All other pantropical (*S. l. longirostris*) are labeled in orange: ATL = NW Atlantic; HI = Hawaii; ZAN = Tanzania (Zanzibar); PHIL = Philippines. Red populations are dwarf spinner dolphins from Indonesia (diamond = DWF.INDO) and northern Australia (square = DWF.AUS). Nodes are labeled with posterior probabilities. Branch widths were scaled to theta values; wider branches indicate higher theta. A reduced sample set was used (n = 22) to speed computation time. Blue trees are the most common, red are the next most common, and green are the least common. Inset map shows sample localities. The unrooted tree is in the inset.

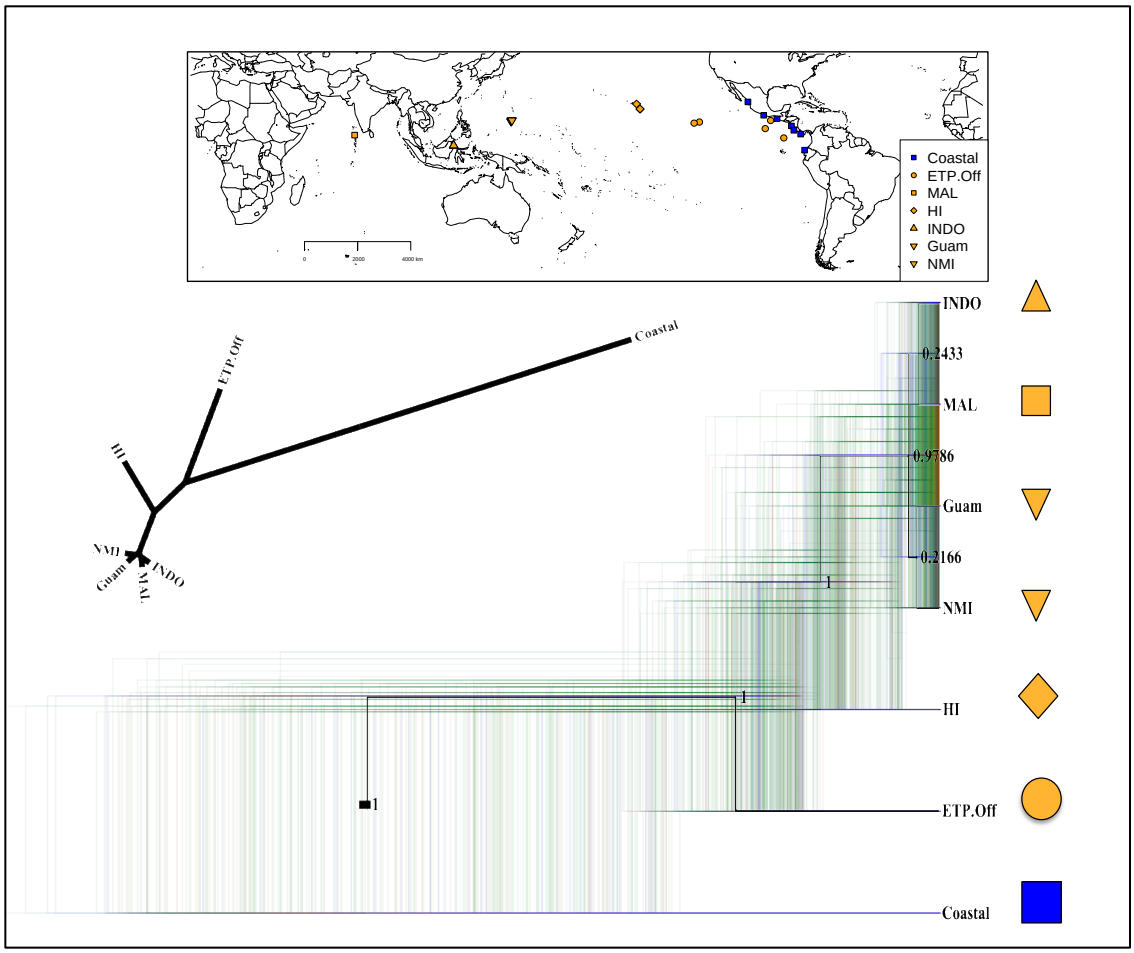


Figure 6. Bayesian species tree for spotted dolphins 6,391 SNPs In black is the mid-point rooted Maximum Clade Credibility tree made from all trees with posteriors >0.5. It is overlaid onto all trees from the analyses. The Coastal subspecies (*S. a. graffmani*) is labeled with a blue box, while all other pantropical (*S. a. attenuata*) are labeled in orange. Offshore eastern tropical Pacific is represented by an orange circle (ETP.Off); Hawaii is represented by an orange diamond (HI); Northern Marianas Islands (NMI) and Guam are represented by inverted triangles; Maldives (MAL) is labeled with a square and Indonesia (INDO) is labeled with a standard triangle. Nodes are labeled with posterior probabilities. Branch widths were scaled to theta values; wider branches indicate higher theta. A reduced sample set was used (n = 25) to speed computation time. Inset map shows sample localities of these samples. Blue trees are the most common, red are the next most common, and green are the least common. The unrooted tree is in the inset.

CONCLUSIONS

Knowledge of how wildlife populations are genetically related across space is a necessity for assessing threats, estimating abundance, and ultimately determining whether protection is required (Wheeler et al. 2004). Without this information, we may fail to recognize entire subspecies or distinct populations that require urgent conservation action (Leslie 2014). Population structure information is also a direct input into the process of listing organisms under protective legislation (e.g., the U.S. Endangered Species Act - ESA), biodiversity status inventories (e.g., the International Union for the Conservation of Nature's Red List), and international conservation agreements (e.g., the Convention on the International Trade of Endangered Species), making it a necessary precondition for conservation measures involving law enforcement and adjudication.

The main objective of my dissertation was to address some of the possible shortcomings in our knowledge of cetacean diversity and population structure by using new philosophies and techniques. In Chapter 1 (Leslie 2014), I examined the principles of phylogenetic nomenclature (PN) as embodied in the PhyloCode (an alternative to traditional rank-based nomenclature that names biological groups based on the results of phylogenetic analyses and does not associate taxa with ranks) and assessed how this novel approach to naming taxa might affect the implementation of species-based legislation by providing a case study of the ESA.

The latest version of the PhyloCode relies on the traditional rank-based codes to name species and infraspecific taxa; thus, little will change regarding the main targets of the ESA because they will retain rank labels. For this reason, and because knowledge of evolutionary relationships is of greater importance than nomenclatural procedures for initial protection of endangered taxa under the ESA, I conclude that PN under the PhyloCode will have little impact on implementation of the ESA.

In Chapter 2 (Leslie et al. *In Review*) and 3, I addressed some of the challenges with assessing population structure of eastern tropical Pacific (ETP) Ocean spinner and spotted dolphins by employing next generation DNA sequencing technology. In Chapter 2, I tested for population structure in these two species at multiple hierarchical levels by collecting whole mitochondrial genome sequences (mtDNA) and nuclear (nuDNA) single nucleotide polymorphisms (SNPs) from 104 spinner and 76 spotted dolphins using capture array library enrichment and highly paralleled DNA sequencing. MtDNA showed weak but significant differences between subspecies of spotted dolphins and evidence for differentiation between ETP spinner dolphin subspecies. NuDNA supported subspecies of spotted dolphins (Perrin et al. 1994), but not subspecies of spinner dolphins (Perrin et al. 1991, Perryman and Westlake 1998). Strong and significant differentiation was detected between whitebelly and eastern spinner stocks using nuDNA. Neither mtDNA nor nuDNA supported the division of existing offshore stocks of spotted dolphins or Tres Marias spinner dolphins. This work helps identify a genetic basis for establishing biologically meaningful management units for the recovery of these dolphins, but lacked

power to detect differences in the nuclear genome.

My nuDNA results partially conflict with previous studies that demonstrated morphometric differences supporting structured populations (Perrin et al. 1991, Perrin et al. 1994, Perryman and Westlake 1998). However, other molecular genetics approaches also failed to find corresponding population genetic structure (Dizon et al. 1994, Galver 2002). Because Andrews *et al*. (2013) found some evidence for segregation in data from the Y-chromosome and Escorza-Treviño *et al*. (2005) found population structure within the coastal subspecies of spotted dolphins using microsatellites, I decided to try another novel technique to gather a more powerful dataset to test these hypotheses.

I used a method of restriction-site associated DNA sequencing (RADseq) to test the same hypotheses of population structure in the nuclear genome for Chapter 3. Using this method greatly increase the number of SNP loci for analysis, and also the statistical power for testing the hypotheses. In fact, I tested for genetic structure at multiple hierarchical levels by analyzing the largest dataset to date brought to bear on these questions. Greater than 7,000 SNPs were collected from nuclear DNA regions associated with the restriction enzyme site *PstI* using genotype-by-sequencing (GBS) from 72 spinner dolphins and 58 spotted dolphins (Elshire et al. 2011). My results support the current subspecies for both species (Perrin et al. 1991, Perrin et al. 1994) and indicate stock-level separation for Tres Marias spinner dolphins (Perryman and Westlake 1998)

and the two offshore spotted dolphin stocks in this area. These results are critically important for the ongoing management and recovery of these highly-impacted pelagic dolphins in the eastern tropical Pacific Ocean (Gerrodette and Forcada 2005, Wade et al. 2007).

Finally, in Chapter 4, my goal was to place the diversity of ETP spinner and spotted dolphins within the context of their global diversity using the same techniques (RADseq) to collect data for a comparative study of global phylogeography. In this study, I tested phylogeographic hypotheses (Davies 1963, Barber et al. 2000, Perrin 2007) by genotyping >6,000 nuDNA SNPs. Analyses for population structure via fixation indexes and Discriminant Analysis of Principle Components (DAPC) indicated significant genetic structure among these global populations. Bayesian phylogeographic analyses showed deep divergence between Indo-Pacific and eastern tropical Pacific Ocean (ETP) lineages of both spinner and spotted dolphins. I also found evidence of division among the dwarf spinner dolphins – with the northern Australia population being very different from Indonesia. My DAPC results indicated a closer than expected relationship between Atlantic spinner dolphins and the ETP groups, suggesting the Panama Seaway may have remained permeable during the early stages of species diversification. Bayesian phylogeographic analyses however, placed Atlantic spinner dolphins basal relative to a clade of Indo-Pacific Ocean populations. In spotted dolphin species, my results supported the current subspecies-level division between coastal and offshore varieties. These results indicate a pronounced inshore/offshore biogeographic barrier in the ETP. I also observed

very close relationships between endemic ETP spinner subspecies in relation to global diversity. Despite the uniqueness of the Australian population, the dwarf spinner dolphin is monophyletic and sister to a major clade that includes Indian Ocean and Western Pacific Ocean populations of the nominate subspecies. These results show the strength of the eastern Pacific basin and the marine Wallace's line as biogeographic barriers for spinner dolphins despite high dispersal potential.

This work has significantly advanced our knowledge of these two species, but also raised other questions of interest. Future directions for this work include: 1) combining data on body form and coloration with these genomic data to evaluate the importance of positive selection in the evolution of ETP spinner dolphin subspecies, 2) genotyping tooth bone material from spinner dolphin crania to perform a detailed investigation of the morphological and genomic cline in ETP spinner dolphins, 3) using other data to help establish the location of the root for spinner dolphin global phylogeographic processes, 4) expanding sampling for spinner dolphins in the central Pacific, eastern Atlantic, and northern Indian Ocean, 5) expanding sampling for spotted dolphins to include these places as well as the Atlantic Ocean in order to complete a true comparative phylogeographic study, and 6) further examination of the uniqueness of the Australian population of dwarf spinner dolphins.

In closing, the ETP is a fascinating and complex place. With greater than half a century of ecosystem-based study, we have come a long way since I hope one day it will return to its full majesty.

My hope is that this dissertation has advanced our knowledge of population biology of these two highly-impacted cetacean species in order to advance the grand challenge of conserving biological diversity in the face of rapid extinction (Pimm et al. 1995, Mace G.M. et al. 2005). Moreover, I hope this work fuels additional research on the biology and conservation of understudied and endangered marine mammal species.

LITERATURE CITED:

Andrews, K. R., W. F. Perrin, M. Oremus, L. Karczmarski, B. W. Bowen, J. B. Puritz, and R. J. Toonen. 2013. The evolving male: spinner dolphin (*Stenella longirostris*) ecotypes are divergent at Y chromosome but not mtDNA or autosomal markers. Molecular Ecology **22**:2408-2423.

Barber, P. H., S. R. Palumbi, M. V. Erdmann, and M. K. Moosa. 2000. Biogeography: A marine Wallace's line? Nature **406**:692-693.

Davies, J. L. 1963. The Antitropical Factor in Cetacean Speciation. Evolution **17**:107-116.

Dizon, A. E., W. F. Perrin, and P. A. Akin. 1994. Stocks of dolphins (*Stenella* spp.) and (*Delphinus delphis*) in the eastern tropical Pacific: A phylogeographic classification. NOAA Technical Report **NMFS 119**:20.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE **6**:e19379.

Escorza-Trevino, S., F. I. Archer, M. Rosales, A. Dizon, and A. Lang. 2005. Genetic differentiation and intraspecific structure of Eastern Tropical Pacific spotted dolphins, *Stenella attenuata*, revealed by DNA analyses. Pages 1-14 Conservation Genetics.

Galver, L. 2002. The molecular ecology of spinner dolphins (*Stenella longirostris*): genetic diverstiy and population structure. Ph.D. Dissertation. University of California, San Diego, CA.

Gerrodette, T., and J. Forcada. 2005. Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. Marine Ecology Progress Series **291**:1-21.

Leslie, M. S. 2014. Impacts of phylogenetic nomenclature on the efficacy of the U.S. Endangered Species Act. Conservation Biology **29**:69-77.

Leslie, M. S., F. I. Archer, and P. A. Morin. *In Review*. Genetic inference of population structure in spinner (*Stenella longirostris*) and spotted dolphins (*S. attenuata*) from the eastern tropical Pacific Ocean based on mitochondrial genomes and nuclear SNPs. Marine Mammal Science.

Mace G.M., Masundire H., and B. J.E.M. 2005. Chapter 4. Biodiversity. Pages 77-122 Millennium Ecosystem Assessment: Ecosystems and Human Well-Being: Current State and Trends. Findings of the Condition and Trends Working Group. Island Press, Washington DC.

Perrin, W. F. 2007. The South African species gate. Whales and dolphins of the southern African subregion. Ed. Best, P. B.

Perrin, W. F., P. A. Akin, and J. V. Kashiwada. 1991. Geographic variation in external morphology of the spinner dolphin (*Stenella longirostris*) in the eastern Pacific and implications for conservation. Fishery Bulletin **89**:411-428.

Perrin, W. F., G. D. Schnell, D. J. Hough, J. W. Gilpatrick, Jr., and J. V. Kashiwada. 1994. Reexamination of geographic variation in cranial morphology of the pantropical spotted dolphin, (*Stenella attenuata*), in the eastern Pacific. Fishery Bulletin **92**:324-346.

Perryman, W. L., and R. L. Westlake. 1998. A new geographic form of the spinner dolphin, Stenella longirostris, detected with aerial photogrammetry. Marine Mammal Science **14**:38-50.

Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks. 1995. The future of biodiversity. Science **269**:347.

Wade, P. R., G. M. Watters, T. Gerrodette, and S. B. Reilly. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. Marine Ecology Progress Series **343**:1-14.

Wheeler, Q. D., P. H. Raven, and E. O. Wilson. 2004. Taxonomy: impediment or expedient? Science (New York, NY) **303**:285.

APPENDIX I.

SUPPLEMENTARY MATERIAL

CHAPTER 2

Table A1-S1. Sample information and mitochondrial genome haplotype assignment for *S. longirostris*.

#	LabI	Haplotyp	Subspecies	Alt. Stock	Latitude	Longitude	SE	YR	SOURCE
1	65	Slon81	Whitebelly	Whitebelly	8.833333	-133.416666	F	1989	FISHERY-ETP TUNA
2	297	Slon45	Whitebelly	Whitebelly	8.833333	-133.416666	M	1989	FISHERY-ETP TUNA
3	669	Slon82	Eastern	Eastern	11.766666	-115.666666	M	1992	FISHERY-ETP TUNA
4	862	Slon91	Eastern	Eastern	11.666666	-120.033333	F	1992	FISHERY-ETP TUNA
5	865	Slon92	Eastern	Eastern	11.666666	-120.033333	M	1992	FISHERY-ETP TUNA
6	868	Slon93	Eastern	Eastern	11.166666	-113.416666	M	1992	FISHERY-ETP TUNA
7	1740	Slon17	Whitebelly	Whitebelly	11.45	-128	F	1993	FISHERY-ETP TUNA
8	1742	Slon18	Whitebelly	Whitebelly	11.45	-128	F	1993	FISHERY-ETP TUNA
9	1744	Slon19	Whitebelly	Whitebelly	12.233333	-130.316666	M	1993	FISHERY-ETP TUNA
10	1746	Slon20	Whitebelly	Whitebelly	12.233333	-130.316666	M	1993	FISHERY-ETP TUNA
11	1747	Slon21	Whitebelly	Whitebelly	12.233333	-130.316666	F	1993	FISHERY-ETP TUNA
12	1776	Slon22	Whitebelly	Whitebelly	11.7	-128.85	F	1993	FISHERY-ETP TUNA
13	1777	Slon23	Whitebelly	Whitebelly	11.666666	-129.583333	M	1993	FISHERY-ETP TUNA
14	1778	Slon24	Whitebelly	Whitebelly	11.666666	-129.583333	F	1993	FISHERY-ETP TUNA
15	1871	Slon29	Whitebelly	Whitebelly	9.783333	-122.1	M	1993	FISHERY-ETP TUNA
16	1987	Slon30	Whitebelly	Whitebelly	-11.383333	-85.9	F	1993	FISHERY-ETP TUNA
17	2191	Slon31	Whitebelly	Whitebelly	-9.016666	-90.116666	F	1994	FISHERY-ETP TUNA
18	2192	Slon32	Whitebelly	Whitebelly	-9.016666	-90.116666	F	1994	FISHERY-ETP TUNA
19	2193	Slon33	Whitebelly	Whitebelly	-9.016666	-90.116666	M	1994	FISHERY-ETP TUNA
20	2194	Slon34	Whitebelly	Whitebelly	-9.016666	-90.116666	M	1994	FISHERY-ETP TUNA
21	2195	Slon35	Whitebelly	Whitebelly	-9.016666	-90.116666	F	1994	FISHERY-ETP TUNA
22	2196	Slon36	Whitebelly	Whitebelly	-9.016666	-90.116666	F	1994	FISHERY-ETP TUNA
23	2197	Slon37	Whitebelly	Whitebelly	-9.016666	-90.116666	F	1994	FISHERY-ETP TUNA
24	2198	Slon38	Whitebelly	Whitebelly	-9.016666	-90.116666	M	1994	FISHERY-ETP TUNA
25	2878	Slon44	Whitebelly	Whitebelly	10.35	-129.583333	F	1985	FISHERY-ETP TUNA
26	3850	Slon80	Whitebelly	Whitebelly	3.066666	-114.966666	F	1982	FISHERY-ETP TUNA
27	7185	Slon86	Hawaii	Hawaii	20	-155.833333	M	1997	BIOPSY-SWFSC-HAWAII
28	7191	Slon87	Hawaii	Hawaii	20.033333	-155.833333	M	1997	BIOPSY-SWFSC-HAWAII
29	7193	Slon88	Hawaii	Hawaii	20.033333	-155.833333	M	1997	BIOPSY-SWFSC-HAWAII
30	7202	Slon89	Hawaii	Hawaii	19.666666	-156.083333	M	1997	BIOPSY-SWFSC-HAWAII
31	11445	Slon2	Eastern	Eastern	14.566666	-98.316666	U	1998	BIOPSY-SWFSC-SPAM
32	11665	Slon3	Central	Central	15.866666	-94.833333	F	1998	BIOPSY-SWFSC-SPAM
33	11667	Slon4	Central	Central	15.866666	-94.833333	M	1998	BIOPSY-SWFSC-SPAM
34	11668	Slon5	Central	Central	15.866666	-94.833333	F	1998	BIOPSY-SWFSC-SPAM
35	11860	Slon6	Eastern	Eastern	12.366666	-93.483333	U	1998	BIOPSY-SWFSC-SPAM
36	11861	Slon7	Eastern	Eastern	12.55	-93.883333	U	1998	BIOPSY-SWFSC-SPAM
44	15922	Slon8	Eastern	Tres Marias	22.183333	-106.533333	U	1999	BIOPSY-SWFSC-STAR99
45	15923	Slon9	Eastern	Tres Marias	22.183333	-106.533333	U	1999	BIOPSY-SWFSC-STAR99
46	15924	Slon10	Eastern	Tres Marias	22.183333	-106.533333	U	1999	BIOPSY-SWFSC-STAR99
47	16002	Slon11	Eastern	Eastern	13.433333	-97.1	F	1999	BIOPSY-SWFSC-STAR99
48	16003	Slon8	Eastern	Eastern	14.283333	-96.766666	F	1999	BIOPSY-SWFSC-STAR99
49	16013	Slon12	Eastern	Eastern	11.25	-95.083333	F	1999	BIOPSY-SWFSC-STAR99
50	16601	Slon13	Eastern	Eastern	13.633333	-116.533333	F	1999	FISH-IDCPA NECROPSY
51	16602	Slon14	Eastern	Eastern	13.733333	-109.7	F	1999	FISH-IDCPA NECROPSY
52	16603	Slon15	Whitebelly	Whitebelly	9.333333	-127.083333	F	1999	FISH-IDCPA NECROPSY
53	17374	Slon16	Whitebelly	Whitebelly	13.433333	-114.283333	M	2000	FISH-IDCPA NECROPSY
54	17820	Slon25	Whitebelly	Whitebelly	6.166666	-125.383333	F	1984	FISHERY-ETP TUNA
55	18200	Slon26	Eastern	Eastern	12.416666	-97.716666	F	2000	BIOPSY-SWFSC-STAR00
56	18368	Slon27	Eastern	Eastern	13.6	-97.25	F	2000	BIOPSY-SWFSC-STAR00
57	18374	Slon28	Eastern	Eastern	14.35	-97.45	F	2000	BIOPSY-SWFSC-STAR00

Table A1-S1. Sample information and mitochondrial genome haplotype assignment for *S. longirostris* (cont.).

#	LabID	Haplotyp	Subspecies	Alt. Stock	Latitude	Longitude	SEX	YR	SOURCE
58	24050	Slon39	Whitebelly	Whitebelly	13.766666	-115.333333	M	2001	FISH-IDCPA NECROPSY
59	24837	Slon40	Whitebelly	Whitebelly	6.7	-126.533333	M	2001	FISH-IDCPA NECROPSY
60	24850	Slon41	Eastern	Eastern	14.533333	-97.366666	F	2001	FISH-IDCPA NECROPSY
61	24923	Slon42	Eastern	Eastern	10.2	-98.533333	M	2001	BIOPSY-SWFSC-CHESS
62	27426	Slon43	Hawaii	Hawaii	20.75	-156.85	U	2002	BIOPSY-HI-BAIRD
63	27437	Slon0	Hawaii	Hawaii	21.533333	-158.233333	U	2002	BIOPSY-HI-BAIRD
64	33958	Slon46	Hawaii	Hawaii	21.7	-160.2	U	2003	BIOPSY-HI-BAIRD
65	38011	Slon47	Eastern	Tres Marias	23.633333	-107.383333	M	2003	BIOPSY-SWFSC-STAR03
66	38012	Slon48	Eastern	Tres Marias	23.633333	-107.383333	M	2003	BIOPSY-SWFSC-STAR03
67	38013	Slon49	Eastern	Tres Marias	23.633333	-107.383333	F	2003	BIOPSY-SWFSC-STAR03
68	38014	Slon50	Eastern	Tres Marias	23.633333	-107.383333	M	2003	BIOPSY-SWFSC-STAR03
69	38015	Slon51	Eastern	Tres Marias	23.633333	-107.383333	F	2003	BIOPSY-SWFSC-STAR03
70	38016	Slon52	Eastern	Tres Marias	23.633333	-107.383333	M	2003	BIOPSY-SWFSC-STAR03
71	38017	Slon49	Eastern	Tres Marias	23.633333	-107.383333	F	2003	BIOPSY-SWFSC-STAR03
72	38018	Slon53	Eastern	Tres Marias	22.25	-106.4	F	2003	BIOPSY-SWFSC-STAR03
73	38020	Slon54	Eastern	Tres Marias	22.25	-106.4	F	2003	BIOPSY-SWFSC-STAR03
74	38021	Slon55	Eastern	Tres Marias	22.25	-106.4	F	2003	BIOPSY-SWFSC-STAR03
75	38024	Slon56	Eastern	Tres Marias	22.25	-106.4	F	2003	BIOPSY-SWFSC-STAR03
76	38025	Slon57	Eastern	Tres Marias	22.25	-106.4	M	2003	BIOPSY-SWFSC-STAR03
77	38026	Slon58	Eastern	Tres Marias	22.25	-106.4	F	2003	BIOPSY-SWFSC-STAR03
78	38033	Slon59	Eastern	Tres Marias	21.433333	-106.1	M	2003	BIOPSY-SWFSC-STAR03
79	38034	Slon60	Eastern	Tres Marias	21.433333	-106.1	M	2003	BIOPSY-SWFSC-STAR03
80	38036	Slon61	Eastern	Tres Marias	21.433333	-106.1	M	2003	BIOPSY-SWFSC-STAR03
81	38037	Slon62	Eastern	Tres Marias	21.433333	-106.1	M	2003	BIOPSY-SWFSC-STAR03
82	38038	Slon63	Eastern	Tres Marias	21.433333	-106.1	M	2003	BIOPSY-SWFSC-STAR03
83	38091	Slon64	Eastern	Eastern	11.333333	-119.583333	M	2003	BIOPSY-SWFSC-STAR03
84	38092	Slon65	Eastern	Eastern	11.333333	-119.583333	F	2003	BIOPSY-SWFSC-STAR03
85	38093	Slon66	Eastern	Eastern	11.333333	-119.583333	F	2003	BIOPSY-SWFSC-STAR03
86	38094	Slon67	Eastern	Eastern	11.333333	-119.583333	F	2003	BIOPSY-SWFSC-STAR03
87	38095	Slon68	Eastern	Eastern	11.333333	-119.583333	M	2003	BIOPSY-SWFSC-STAR03
88	38096	Slon69	Eastern	Eastern	11.333333	-119.583333	F	2003	BIOPSY-SWFSC-STAR03
89	38097	Slon70	Eastern	Eastern	11.333333	-119.583333	M	2003	BIOPSY-SWFSC-STAR03
90	38128	Slon71	Eastern	Eastern	14.75	-99.1	F	2003	BIOPSY-SWFSC-STAR03
91	38143	Slon72	Eastern	Eastern	13.5	-95.466666	F	2003	BIOPSY-SWFSC-STAR03
92	38191	Slon73	Central	Central	12.583333	-88.533333	M	2003	BIOPSY-SWFSC-STAR03
93	38193	Slon74	Central	Central	12.583333	-88.533333	M	2003	BIOPSY-SWFSC-STAR03
94	38194	Slon75	Central	Central	12.583333	-88.533333	F	2003	BIOPSY-SWFSC-STAR03
95	38195	Slon73	Central	Central	12.583333	-88.533333	M	2003	BIOPSY-SWFSC-STAR03
96	38196	Slon76	Central	Central	12.583333	-88.533333	F	2003	BIOPSY-SWFSC-STAR03
97	38197	Slon77	Central	Central	12.583333	-88.533333	U	2003	BIOPSY-SWFSC-STAR03
98	38267	Slon78	Eastern	Eastern	14.7	-97.083333	F	2003	BIOPSY-SWFSC-STAR03
99	38268	Slon79	Eastern	Eastern	14.5	-97.283333	F	2003	BIOPSY-SWFSC-STAR03
100	56749	Slon0	Hawaii	Hawaii	20.766666	-156.8	F	2005	STRAND-HI-MMRP
101	66907	Slon83	Eastern	Eastern	12.816666	-94.15	M	2006	BIOPSY-SWFSC-STAR06
102	67035	Slon84	Eastern	Eastern	14.566666	-97.5	F	2006	BIOPSY-SWFSC-STAR06
103	67037	Slon85	Eastern	Eastern	14.566666	-97.5	M	2006	BIOPSY-SWFSC-STAR06
104	76367	Slon43	Hawaii	Hawaii	20.9	-156.683333	U	2008	STRAND-HI-HPU
105	78789	Slon90	Hawaii	Hawaii	27.8	-175.8	U	2006	BIOPSY-HI-PIFSC
106	78791	Slon0	Hawaii	Hawaii	27.8	-175.783333	U	2006	BIOPSY-HI-PIFSC
107	78797	Slon1	Hawaii	Hawaii	27.8	-175.883333	U	2006	BIOPSY-HI-PIFSC
108	92435	Slon94	Eastern	Eastern	10.166666	-98.516666	F	2001	FISHERY-SWFSC
109	10241	Slon0	Hawaii	Hawaii	28.4	-178.35	U	2010	BIOPSY-SWFSC-HICEAS
110	10241	Slon1	Hawaii	Hawaii	28.4	-178.35	U	2010	BIOPSY-SWFSC-HICEAS
111	10241	Slon1	Hawaii	Hawaii	28.4	-178.35	U	2010	BIOPSY-SWFSC-

Table A1-**S2.** Sample information and mitochondrial genome haplotype assignment for *S. longirostris*.

#	LabID	Hap	Subspecies	Stock Strata	sex	Latitude	Longitude	SST	YR	SOURCE
1	64	Satt49	Pantropical-ETP	Offshore-Northern	F	14.11666	-108.566666	27.82	1989	FISHERY-ETP TUNA
2	834	Satt54	Pantropical-ETP	Offshore-Northern	F	9.416666	-102.983333	27.66	1992	FISHERY-ETP TUNA
3	835	Satt55	Pantropical-ETP	Offshore-Northern	F	9.416666	-102.983333	27.66	1992	FISHERY-ETP TUNA
4	973	Satt56	Pantropical-ETP	Offshore-Western+Southern	F	8.7	-126.1	28.09	1989	FISHERY-ETP TUNA
5	977	Satt57	Pantropical-ETP	Offshore-Northern	M	8.816666	-104.483333	27.37	1992	FISHERY-ETP TUNA
6	1267	Satt13	Pantropical-ETP	NA	M	5.133333	-86.383333	26.91	1992	FISHERY-ETP TUNA
7	1268	Satt14	Pantropical-ETP	NA	M	5.133333	-86.383333	26.91	1992	FISHERY-ETP TUNA
8	1269	Satt15	Pantropical-ETP	NA	M	5.133333	-86.383333	26.91	1992	FISHERY-ETP TUNA
9	1883	Satt30	Pantropical-ETP	Offshore-Western+Southern	F	14.01666	-125.533333	28.34	1993	FISHERY-ETP TUNA
10	1988	Satt31	Pantropical-ETP	Offshore-Western+Southern	M	-	-85.9	24.04	1994	FISHERY-ETP TUNA
11	2173	Satt32	Pantropical-ETP	NA	F	5.05	-88.233333	27.58	1994	FISHERY-ETP TUNA
12	2190	Satt33	Pantropical-ETP	Offshore-Western+Southern	F	-9.016666	-90.116666	24.81	1994	FISHERY-ETP TUNA
13	2199	Satt34	Pantropical-ETP	Offshore-Western+Southern	M	-9.016666	-90.116666	24.81	1994	FISHERY-ETP TUNA
14	2200	Satt7	Pantropical-ETP	Offshore-Western+Southern	M	-9.016666	-90.116666	24.81	1994	FISHERY-ETP TUNA
15	2201	Satt15	Pantropical-ETP	Offshore-Western+Southern	M	-9.016666	-90.116666	24.81	1994	FISHERY-ETP TUNA
16	2202	Satt35	Pantropical-ETP	Offshore-Western+Southern	M	-9.016666	-90.116666	24.81	1994	FISHERY-ETP TUNA
17	2203	Satt36	Pantropical-ETP	Offshore-Western+Southern	F	-9.016666	-90.116666	24.81	1994	FISHERY-ETP TUNA
18	2879	Satt41	Pantropical-ETP	Offshore-Western+Southern	F	8.216666	-129.666666	28.03	1985	FISHERY-ETP TUNA
19	11361	Satt0	Coastal	Coastal	F	23.21666	-110.483333	29.36	1998	BIOPSY-SWFSC-SPAM
20	11377	Satt1	Coastal	Coastal	M	21.46666	-105.733333	30.21	1998	BIOPSY-SWFSC-SPAM
21	11382	Satt2	Coastal	Coastal	F	16.45	-99.3	30.41	1998	BIOPSY-SWFSC-SPAM
22	11384	Satt3	Coastal	Coastal	F	15.95	-94.45	29.9	1998	BIOPSY-SWFSC-SPAM
23	11397	Satt4	Coastal	Coastal	M	14.85	-93.05	30.24	1998	BIOPSY-SWFSC-SPAM
24	11403	Satt5	Coastal	Coastal	F	19.03333	-104.733333	28.59	1998	BIOPSY-SWFSC-SPAM
25	11913	Satt3	Coastal	Coastal	F	12.11666	-87.783333	29.53	1998	BIOPSY-SWFSC-SPAM
26	11921	Satt6	Coastal	Coastal	F	11.46666	-86.566666	29.41	1998	BIOPSY-SWFSC-SPAM
27	11931	Satt7	Coastal	Coastal	M	9.433333	-85.466666	28.88	1998	BIOPSY-SWFSC-SPAM
28	11933	Satt7	Coastal	Coastal	F	9.466666	-85.45	28.9	1998	BIOPSY-SWFSC-SPAM
29	11935	Satt7	Coastal	Coastal	F	9.466666	-85.45	28.9	1998	BIOPSY-SWFSC-SPAM
30	11950	Satt8	Coastal	Coastal	M	7.65	-82.333333	28.31	1998	BIOPSY-SWFSC-SPAM
31	11962	Satt9	Coastal	Coastal	F	7.183333	-80.45	28.03	1998	BIOPSY-SWFSC-SPAM
32	11965	Satt10	Coastal	Coastal	M	7.183333	-80.416666	28.05	1998	BIOPSY-SWFSC-SPAM
33	12035	Satt11	Coastal	Coastal	F	0.183333	-80.5	25.1	1998	BIOPSY-SWFSC-SPAM
34	12039	Satt11	Coastal	Coastal	F	0.183333	-80.5	25.1	1998	BIOPSY-SWFSC-SPAM
35	12040	Satt12	Coastal	Coastal	M	0.183333	-80.5	25.1	1998	BIOPSY-SWFSC-SPAM
36	14346	Satt16	Pantropical-ETP	Offshore-Western+Southern	U	3.333333	-87.2	27.71	1979	FISHERY-ETP TUNA
37	14352	Satt17	Pantropical-ETP	Offshore-Western+Southern	M	-7.716666	-86.066666	24.07	1980	FISHERY-ETP TUNA
38	16028	Satt18	Coastal	Coastal	M	13.05	-89.75	28.83	1999	BIOPSY-SWFSC-STAR99
39	16030	Satt19	Coastal	Coastal	M	13.05	-89.75	28.83	1999	BIOPSY-SWFSC-STAR99
40	16127	Satt14	Coastal	Coastal	M	7.2	-78.483333	26.87	1999	BIOPSY-SWFSC-STAR99
41	16128	Satt7	Coastal	Coastal	M	7.2	-78.483333	26.87	1999	BIOPSY-SWFSC-STAR99
42	16251	Satt20	Pantropical-ETP	Offshore-Northern	F	13.63333	-106.233333	28.41	1999	BIOPSY-SWFSC-STAR99
43	16255	Satt21	Pantropical-ETP	Offshore-Northern	M	14.48333	-105.683333	28.69	1999	BIOPSY-SWFSC-STAR99
44	16258	Satt22	Pantropical-ETP	Offshore-Northern	M	14.6	-105.633333	28.72	1999	BIOPSY-SWFSC-STAR99
45	17428	Satt23	Pantropical-ETP	Offshore-Western+Southern	F	10.23333	-125.05	28.5	1985	FISHERY-ETP TUNA
46	17429	Satt24	Pantropical-ETP	Offshore-Western+Southern	F	8.916666	-129.966666	28.19	1985	FISHERY-ETP TUNA
47	17431	Satt25	Pantropical-ETP	Offshore-Western+Southern	M	10.35	-129.583333	28.43	1985	FISHERY-ETP TUNA
48	18138	Satt26	Pantropical-ETP	Offshore-Northern	U	13.01666	-101.283333	28.56	2000	BIOPSY-SWFSC-STAR00
49	18139	Satt26	Pantropical-ETP	Offshore-Northern	U	13.01666	-101.283333	28.56	2000	BIOPSY-SWFSC-STAR00
50	18198	Satt26	Pantropical-ETP	Offshore-Northern	U	11.61666	-100.05	27.44	2000	BIOPSY-SWFSC-STAR00
51	18267	Satt8	Coastal	Coastal	M	7.8	-82.4	27.9	2000	BIOPSY-SWFSC-STAR00

Table A1-**S2.** Sample information and mitochondrial genome haplotype assignment for *S. longirostris* (cont).

#	LabID	Hap	Subspecies	Stock Strata	SEX	Latitude	Longitude	SST	YR	SOURCE
52	18268	Satt8	Coastal	Coastal	M	7.766666	-82.233333	27.89	2000	BIOPSY-SWFSC-STAR00
53	18353	Satt27	Pantropical-ETP	Offshore-Northern	U	8.733333	-100.166666	27.23	2000	BIOPSY-SWFSC-STAR00
54	18354	Satt28	Pantropical-ETP	Offshore-Northern	U	8.733333	-100.166666	27.23	2000	BIOPSY-SWFSC-STAR00
55	18357	Satt29	Pantropical-ETP	Offshore-Northern	U	8.666666	-100.4	27.23	2000	BIOPSY-SWFSC-STAR00
56	18397	Satt3	Coastal	Coastal	M	14.43333	-92.4	28.02	2000	BIOPSY-SWFSC-STAR00
57	24047	Satt37	Pantropical-ETP	NA	M	5.866666	-90.066666	29.29	2001	FISH-IDCPA NECROPSY
58	24049	Satt38	Pantropical-ETP	Offshore-Northern	M	9.066666	-108.35	27.75	2000	FISH-IDCPA NECROPSY
59	24052	Satt39	Pantropical-ETP	Offshore-Western+Southern	F	8.15	-128.6	27.6	2001	FISH-IDCPA NECROPSY
60	24848	Satt40	Pantropical-ETP	Offshore-Northern	M	9.633333	-102.116666	28.65	2001	FISH-IDCPA NECROPSY
61	24849	Satt7	Pantropical-ETP	Offshore-Western+Southern	F	10.366666	-127.666666	27.86	2001	FISH-IDCPA NECROPSY
62	33842	Satt42	Pantropical-HI	NA	U	21.41666	-158.35	25.7	2003	BIOPSY-HI-BAIRD
63	38131	Satt43	Pantropical-ETP	Offshore-Northern	F	13.38333	-101.433333	29.4	2003	BIOPSY-SWFSC-STAR03
64	38133	Satt44	Pantropical-ETP	Offshore-Northern	F	12.91666	-101.3	29.24	2003	BIOPSY-SWFSC-STAR03
65	38135	Satt45	Pantropical-ETP	Offshore-Northern	M	12.91666	-101.3	29.24	2003	BIOPSY-SWFSC-STAR03
66	38286	Satt7	Pantropical-ETP	Offshore-Northern	M	14.58333	-106.183333	28.87	2003	BIOPSY-SWFSC-STAR03
67	38287	Satt46	Pantropical-ETP	Offshore-Northern	M	14.58333	-106.183333	28.87	2003	BIOPSY-SWFSC-STAR03
68	38288	Satt47	Pantropical-ETP	Offshore-Northern	F	14.58333	-106.183333	28.87	2003	BIOPSY-SWFSC-STAR03
69	55197	Satt48	Pantropical-HI	NA	F	19.28333	-156	24.33	2006	BIOPSY-CASCADIA-BAIRD
70	67068	Satt50	Pantropical-ETP	Offshore-Northern	F	13.95	-106.383333	28.48	2006	BIOPSY-SWFSC-STAR06
71	67119	Satt51	Pantropical-HI	NA	M	19.48333	-156.15	26.53	2006	BIOPSY-SWFSC-STAR06
72	72686	Satt52	Pantropical-ETP	Offshore-Northern	F	13.75	-106.1	28.93	2007	BIOPSY-SWFSC-STARLITE07
73	72687	Satt52	Pantropical-ETP	Offshore-Northern	F	13.75	-106.1	28.93	2007	BIOPSY-SWFSC-STARLITE07
74	72695	Satt53	Pantropical-ETP	Offshore-Northern	U	13.916666	-106.05	28.09	2007	BIOPSY-SWFSC-STARLITE07
75	73893	Satt42	Pantropical-HI	NA	M	19.4	-156.033333	24.56	2008	BIOPSY-CASCADIA-BAIRD
76	75657	Satt42	Pantropical-HI	NA	M	19.4	-156.033333	26.73	2008	BIOPSY-CASCADIA-BAIRD

> SRY AB275398.1 Slon gene for testis determining factor, complete cds
GAAGTGAAGAAACAAGGATCATGATTAGCGTGAATTCAGTATCCTGTGTGTTCTAATGGCTAACAGGTTCA
GGTTGTCTTTAAGCTTTGTGTTTAAGATAATACTCACTTACATTCATAACGATAAAAAATTCATTAGTATCT
CTGCTGCACCTTCATCCTTTGCATTAAGGGGATAATAAAATAGCTTTATAATGACAAAGTTTGTATTTTAAA
GCAGGTGTTTTAGCACCTTCAGCGGTTTTGATTAGATATAAACAAAGCAGAAAGCAGAGCGTTACAATCGT
ATTAAATAGTTTCCCTTGAGAATAGGTAGGTTGGCTGGGCGGGCTTTGGCTGAGGGCCGGGGGCGGTATTG
AGGGGAGGTGTTGGGGGCGGAGAGGTACGTATTTCACTGCAAAAGCTGTAGTCAGTCAGTGCCTAGTAGA
AATAACCCTTGAATAGCAAGATAATTTTCGTAATACTTGCACCTTCGAATTTTGATTCCTCACCCTTTTTAA
ACGGTGCAATCATACGCTTGTGCCATGTTCAGAACTGTGAACGGTGAGGATTACAGCCCAGCGGTACAGC
AACGAAATATCCTCGACTTTGGGAAAGCACATTCCCTACTGTGGACGGACAACGGTAGCGCGAATGATCG
GTGTGAAACCGGAGGAAATTGTAGAGAGAGCGGCCAGGACCGCGTCAAGCGACCCATGAACGCTTTCATT
GTGTGGTCTCGTGATCAAAGGCGAAAGGTGGCTCTAGAGAATCCCCAAATGCAAAACTCAGAGATCAGTA
AGCGCCTGGGATACGACTGGAAAATGCTTACAGAAGCCGAAAAGCAGCCATTCTTCGAGGAGGCACAGA
GACTACGAGCCATGCACCGAGACAAATACCCGGGCTATAAATATCGACCTCGTCGGAAGGCCAAAGAGGC
CACAGAAATCGCTTCCCGCAGACTCTTCAGTACTGTGCAGCCGAATGCACATAGAGGAGACGTTGTACCC
CTTCCCATCCAAGGACGGTTGCGCCAAGGCCACACGTTCACGAATGGAAAGCCGGTTAAGCCACTCACAG
CCCACGAACATAAGCAGCTCACTCCTGCCACAGGAGCATCGCAGCAGCTGGACAAGCCTGCGCCACAATA
GGGTAACACAGGCTACGCAGACCTGCGGACGTCCCCTTTTACTATAGTTCAGAGCCCGAACTTTCTCACGT
TTATTTTCGATGCTGGTTTCCTTACTCTCGCTATCAAAGGCCCTATTCATCTCAATTTTACTATTATTTCACC
TGTGACTTAATTTCAAAATAAGTCACATAAACATGTTTAACATATAAAGAATTAGGACCTTACAATATGAC
TGAACCTTTGTTTATGACTGCTTGAAAGAATACTAACTATTAAGAAAGTATCTTAACAGCACTGAAACTGC
TTGAATTCAAAGGCCATCTGTTTTTCCTATCGGTGGAAATATTTTCATATCTAATTTTAGTTGTTCCCGAGA
TTGGCCATTAGATAGGTGTTCGTATTTGATAGTAGTCCGGTAATCGTCAGCATAGACAATAGAATTCTCAC
TCTTTATTTTAAATACTGTAACTCAAAACTATAGTGCTTCGAGAAACACTCACGGATTTATGGTATAGGGA
AAAAATGTCGCACTTCGATGGAAATCTTCCTAACTCCTAAACCACTCCCTAATTAGTGAATACATATGTAC
ATTCTATCTTCTTTGACAGTTTATGTTGGGTCCTATTCCCGTTGTGGAAAAACCAGTTTTGCTGCCATTTGC
AACTCAATTCCTTCAGTCCGCACCAACTCTGCACGCTTCGACCTCTATTTTGAACCAATTATAACGTCTATC
GGGTAAAAAGAAATCTTTAATTAGCTCCTATGGGACTCACGGTTTCGCCTCTTTTCCGAAGTTTTTCTTTTA
ACAATGGCAA

>ASIP_CapratoTtruExon1genescaffold:turTru1:GeneScaffold_536:33601:34831:1
GAGGGCCAAACCACAGAGGGCCTCAAGGACCAAAGTAACGGGGAGGAGGGTGTGCTTTAAGCCCGAGTG
TGATGCCATCATAGGTCTTCTGAGAGGTCCCTTTTGTGCACTGGAGGTCGGACTGTAAGGGGGCAAAAGCC
GCAGACAGGGGCCTAGCCCAGGGTAGCAGCGGTAGTGAAGCAGTGGGTAGACTCCAGATCTATTGTGAAG
GCAGAACTGACAGGGATTGTGGCTACACTGGAAGGGGAAGACAACAGAGACGGTAGTTAAAAACGGTGG
AGCTGAGTGGGATGCCCTAGGGAGAGAAGAGGCCAGGCCGGAGCCAGTTAAGTTTTGATGAAGTGGGGC
ATTGTTGGGGGAGCAGGCGGGGACCCGCTGGGATGCCCTGTGGGCATCAGTCGAGCGGGCGGTGGGCAGA
AGGTGGGGATGTCTGGTCGGGGGCCTCCATAGTCCAAGGAGTCTCCAGGCAGAGGTGAGGACGAGTGGGG
CGGACGTGGACTGCTCGGGCAGACCCGGCGTTTCCCCGCAGAAAAAGGCTTCGATGAAGAAGGTGGCACA
GCCCCGGCCCCCGCCGCCCGGCCCCTGCGTGGCCAGCCGCGACAGCTGCAAGCCGCCGGCGCCCGCCTGC
TGCGACCCGTGCGCCTTCTGCCAGTGCCGCTTCTTCCGCAGCGCCTGTTCCTGCCGCGTGCTCAACCCCACC
TGCTGAGCGCGCCTCTGGGTGGTGGGCGGGGTGGGGCGGGGCTTCCACGGGCGGGGCTCCCCGGACGGGC
CCTTCTGGGGCGGGCAATCTCTAGTAGGTGTGACTACCTAAAACAGGGAGTGGGCGTAGCTACTGACGTT
GGGCGGAGCTTTTAGGAGGCGGGGCTTCAAGGAGACTTGGCTTGGGCTGGGCTAAAATCCAAATATATAT
AGGCTATTGAAAGGTGTGTGGCTGTTTCTTTAAAGAACCCGAAAGATCCTTTTCCCTTCGCCGAGTCCCCTT
CACCGCAGAAGTTCCGCGGGCTGCGCACTTGCTCTCTCCTGGTCCTGGGGAAACGTGGTTACCTGGCCCTG
AGGCTGCCCTCCTGGTAAGCCCATCTCTGGGTGTGGCCGCCAGAGACAATGCTGCCCCCTTCCTTTAAGCC
TGGGAACACCCGAAGCTGCCTACAAAACCAGGGAGGCAGAAAGGAACCATTAAGGAACTACCTTTCAGT
GTTTCACCCTTTCCCCGTTCTAAGAATCTCCCCTCAGC

Figure A1-S1. List of additional nuclear gene sequences used in capture array design.

>TYRP1 genescaffold:turTru1:GeneScaffold_710:164613:165948:1
GGGGACTTTAAATCCTGGATATTAACTTGCTAGAAATCTGCCCCCCATTTCTTCCTTCAGCTCCATGGAT
AAATGTAATTAACGACAGTGACCCGATAATGAGTGCTTTTGATAAATGACATCCAGGATAAGTATTTTTT
TATATAAATAGAAATTTAAAATGAAGAAAATTAAATGACGAGGACAGGGGATTTAATTATGTAGGTATT
ATACATGTCAAATTTTTAAAGGTAAGTTACTATTTTGATTAGTTTTTGAAATGTCATTTTACAGATATGA
AAAGCATTTTGAAGGTGGGTTGGGAAGGGTCTCATGCCATTTATAACAGTACAAAACAACTTGCATAAT
TTCATCTTGCTTTTTCCTTTTCAGCTTGATTTTTATCTATGTGCCTCGGGCTCTACACAAAGGGCTGAAAA
CCAGAGCCTGTACTGCATTCTTGCTTCAAGCAGAAATGAAAGCTCCCAAACTCCTCTCTCTGGGCTATAC
CTTCTTGGCCCTGCTTTTTTTCCAGCAGGCCTGGGCTCAATTCCCAAGACAGTGTGTCACCATCGAGGCT
TTGAGAAATGGCGTGTGTTGCCCAGACCTGTTCCCACTGTCTGGGCCTGGGACTGACCGCTGTGGTTCCT
CATCGGGAAGGGGCAGGTGTGAAGTGGTGACAGCTGACTCCCGACCCCACGGCCCCCAGTACCCACAT
GATGGCAGAGATGACCGAGAGGGCTGGCCCACGCGCTTCTTCAACAGGACATGCCACTGCATTGGCAAT
TTCTCAGGATACAACTGTGGAACCTGCCGTCCTGGATGGAGAGGAGCTGCCTGTGACCAGAGGGTTCTC
ATAGGTAAGTGGGGTGTGAATGGATGCACAGTTCTGCGTGGAACTCAACACACTTACAGAATCCTGAAC
CAGTTGAACTTTCAGAACTCCTAGAAATCATAAAGCTCAAGACTTACTTTTCACAGCTGAGGAAACTGA
GGCTTAGAGAGGCTGAGAAATGTATCTGATTTAAGGAGTTGAAGGTGAACCTCAAGTCTTCTAATAACA
TTTCCTTTCAAAAGATTTGTTGAGCAACTCCTTTATACTAGGAACTGTGCTAGGTGTTTATGACTGGACC
TTTCAGTAACTTTGCAAGATAGTTATGTTAATTGTGAGGTTCTGAAAGGTCAAGTCCATGTAGCTGGTAT
GTGGTAAACCTGGAATTCACACTGAGATGTGTCCAGCTCTAATCGATGAGCTTCTTCCACTACCCTTGCT
GACCACAAGGATGGTTCCTGCTCTGATGATCAAGTTCAGCCATTTTACAAGATGAGCCTGGCCTGCAGC
CTAAATGACAATAAA

>TYRP1 Exon 4 genescaffold:turTru1:GeneScaffold_710:169647:171223:1
ACAAAACAGAAACAAAATAAAAGGCACTGCTTCACATAATATATCTCAAGGCTGAGTCTGGCGCCAGA
CCAGGAGTTAGTCATGCTTGAACTAAAGATAAGAAAATGGAAGGAAGTGGAAATGTCAAAGTGTAGAA
GTGATTATTGATAATTTTTAGTAAAACCAACAGAAATGTAACTACTTGCCTAAAAAAGAAAAGGGAGAC
ACCTCCCCAAAGCATTTGCAGAAACTTCTGTATAACAAGTTGAGTTTCATTCCCTCTAGTGCCCAAGTGA
CTTCTTGTAGTAAATTTGAGGCTTTGTTAAATAAGGTATTTAAGTCCTCTAGCCCTTCAGACACCACTGA
CACACTAACTGGTTCCTAAGTGGCTGAAGAGAACTAATAGAAACAGAATGCCAGGGAGTAAACCAAGC
AAAGGGAGAATTTTAAACAAAAGCAGAGAATGCTAACAGAGTTTCTGTGATCTAGGAAATGTTGCAGG
ATACTTCTTTCTCCCTTCCTTACTGGAATTTTGCCACTGGGAAGAACACCTGTGACATTTGCACCGATGA
CTTGATGGGATCAAGAAGCAACTTTGATTCCACTCTTATAAGCCCGAACTCCGTCTTTTCTCAATGGCGA
GTGGTCTGCGAATCCTTGGAAGATTATGATACCCTGGGAACCCTTTGTAACAGTAAGTTCCAAATGACA
ACTAGGATTCAGAATTGCCTGTTATGTAAAGTGATTAAATGTGCTGCCTGAAGGCTCTTCATTCTACTAA
GGACTTCAGACTAAAATCCACGTTTATTATGGAGAGTTGATGTACAAATATTCACTAAGTACCTAGGAC
AGTCAAGGCAATCTGGGGTGCTCCAGGGGAAAAAGAAAGTGCCTATACATTTTTCCATTGCTTTCCTTG
AACAACTTGGTCAGTGCATTGCTCACTTTCTTATTTTGGAAATGTCTATTTTGTATTAATTAATCCTCTTA
GTAAGCTCAGCACTGCTTAGTAGTCATGTGTCTTGTGTTGGAATTTCACAGAAAATGTTTCCCTAGGAAA
TGTGAAATATGCAGAAAGAGTTGGAAGTGCCCTGGGAGTAAAAAATACATATTATAATAATAATTACTA
TTATTATTGTATAATGATTATTACTATTATTTTATGTTATTATAATATATTATTTTTACTATTATGTTAATC
TGGTCAGAAGGTTAAAAAACCGAAAGGTGACATGGAGGATGCCTTTAATAAAAGAAAGATACTGTACA
GAAGATGTCTGCCAACTAGCTCTTTCCAGCTCTCCCAAAGTGAATACAAAAAGAGAATGATTAACCAGT
TTTCCTAGCATCCTTTCTCCAGTGGTCTGAATGCCATCTCTGCCACACGTCACGTTTCTGTCCTATGTGGA
TCAGCTTCTGTGAAGCTTTCTGTTCTCTTCCTTAGATCCTATTTTTTACCGTCTCTCTTGCTGCTAACAGAA
GGTCTTAATTATCAGAGTCTTTTTATCCCATCTTATTTTTCTTATTCAGGCATGTCTGGCTATTTTGAACTT
TTGTTCTTCCATATAAAAACTAGAATTATTGGAATTTATTTTTTA

Figure A1-S1. List of additional nuclear gene sequences used in capture array design (cont.).

>MC1R genescaffold:turTru1:GeneScaffold_3491:64839:66374:1
CGGTCTCGTAGGACCGAAGGTAGCAAGAGGGCTAGAAATGTCCGGATCTGGGCGACCGCGCCTCCAGG
GAAGAGGGGAGGTGGACTGAGTGACCAGGACAGAGACACTGAGCCCGCAAGGCTGGCAACTGGGAGC
TTGGGGTCCTGGCTGGGAAGTGACGACTGGTGGAGAGGCCCGGGGCCGTGCCTGGGCCGACATTTGTCC
AGCCCCGGAGGACGGAGTGTGAGGGGCCTCCCGGTGGAGCCATGAGCTGGGCGGGACCCTGAGAGCAA
GGACCCCTTCCTGCTCCCTGGCAGGACGATGCCTGTGCTCGGCCCCCAGAGGCAGCTGCTGAGTCCCCT
CAACTCCACGTCCCCAGCCACCCCCCGCCTCGGGCTGGCCACCAACCAGACGGGGCCCCAGTGCCTGGA
GGTGTCCGTTCCCGAAGGGCTGTTCCTCAGCTTGGGGCTGGTGAGTCTCGTGGAGAACATGCTGGTGGT
GGCCGCCATCATGAAGAACCGCAACCTGCACTCGCCCATGTACTACTTCATCTGCTGCCTGGCCGCGTCC
GACCTGCTTGTGAGCATCAGCAACGTGCTGGAGACGGCCGTCATGCTGCTGCTGGAGGCCGGCGCCCTG
GCCCCTCGGGCCGCCGTGGTGCAGCAGCTGGACGACGTCATTGATGTGCTCATCTGCAGCTCCATGGTG
TCCAGCCTCTGCTTCCTGGGCGCCATCGCCGTGGACCGCTACATCTCCATCTTCTACGCCCTGCGGTACC
ACAGCCTCGTGACATTGCCCAGGGCGTGGCGGGCCATCGCGGCCATCTGGGTGGCCAGCATCCTTGCCA
GCACCCTCTTCATCGCCTACTACAACCACGCGGCCGTCCTGCTCTGTCTCGTGAGCTTCTTCGTAGCCAT
GCTGGTGCTCATGGCGGTGCTGTACGTCCACATGCTGGCCCGGGCGTGCCAGCATGCCCGGGGCATTGC
CCGGCTCCACAAGAGGCAGCGCTCCACCCAGCAGGGCTTCGGCCTCAAGGGCGCAGCCACCCTCACCAT
CCTGCTGGGCACCTTCTTCCTCTGCTGGGGCCCCTTCTTCCTCCACCTCTTGCTCATCGTCCTCTGCCCTC
AACACCCCACCTGTGTCTGCATCTTCAAGAACTTCAACCTCTTCCTCACCCTCATCATCTGCAACGCCAT
TGTGGACCCTCTCATCTACGCCTTCCGCAGCCAGGAGCTCCGGAAGACGCTCCAAGAGGTGCTGCAGGG
CTCCTGGTGAGCTGCTGGGCAGAGGGAGGAGGTGCTGTCGTGTGGCCCAAGCCCTGGGTGGCCGGGGC
GGTCCCTTGGCGGAGAGGACGGGCTAGGCCATCGCTGAAGGTGTGGCTGCACAGACCTTCTGGGGCCCG
GGAGGGGAATGGCCCGAAGCTCCAGGAGGCGCTGTGGGGGATGGGTGAGGCTGGGAGACTGCGGGGC
ACGGCCACCCCGTCCACTGGGATGTGCAGGGTGGTGGCAAAAGCTGTCGGCAGCGCTGCTCCAAGGACT
TTGTGACCAGCAAGGAGGA

>Kit Ligand genescaffold:turTru1:GeneScaffold_136:166877:168036:-1

AAAAAAAAAAAACACGCACGAGAATTCATCTATGAAGAAGCCTTGAAATCTATCAAGGTCTTTAGGAA
AATAATCTATTATTAATCCACTGATTCATTTTCCAATTCACCAAAAGACTGGAAGTTCCTTGAAGACTAG
GACTCTCTCCCACAATAATTTTTTTGTAGTTTCCTCTGGGCCTAAAGTTCTTTTAAATTCCATTAGTATCT
GAACCTTGTCAAGGACTCTGTCTTGGCTCTTCTTGTCTTAATTTTACTTCCATGGAAGAGCAAAATCTTTA
CCATCAACAGTTTCCTCTGATCTTTATGCATTTGTAAATTTTGAAACAGATGTTGGGTATGACTTAAACA
AAACATAAGAACTAACCTGTTCTGATATAAGGAAATCCTTTGTCCGGGAGTCCATTACTGAAACATTGA
TGGCTTTGAAGGGTAATCCACAGGTAATCCACCTGTACTTAAAAAATAGGTGCTTTATCTTCTTTTCTTTC
AATTCTTACAGCCTAGTCATTGTTGGATAAGCGAGATGGTGGAACAATTGTCAGTCAGCTTGACTGATCT
TCTGGACAAGTTTTCAAATATTTCTGAAGGCTTGAGTAATTACTCTATCATAGACAAACTTGTGAAAATA
GTCGATGACCTTGTGGAGTGCATGGAAGAACACTCATTTGAGGTAACTTGGTATTCATTGGGATTGCTTT
TTATTACTCTTCTCTAAAACCTGTGCTGATTGGTCATGGCTGGGATGGGGGCATATTTGTGAGATTTCTA
ATTTATAAATTTTTAATTAAAAAATTTTTAAGATGAGCACATCTCTTAGAGTTAACATCTCTTTTTATGTC
AATCGTATCAGTACAAATAGAGTGCTAATCGTATCAGTACAAATAGTGTGCTAATGGCATTGATAAGAA
GTAGAGGTCACAAAAAGTAGGGTATATTAGACGAGGTATCATGAAGGAAATAACATACGCATTGGGGC
ATGAAAGATGAGTTAGGTGGTCAGTAAATGGCAGTGTCTGGAAGAGCATTCCAAATACCAGAGAATGT
CCGTTTACTAATCATAGAGTAGTACATAGTAAGTGCTACATAGGTGTTAACTCTTATTATTACCTGCTTTT
ATTTATAATCATGTATTTATGAATTCCTAAAATAGACCTTTGA

Figure A1-S1. List of additional nuclear gene sequences used in capture array design (cont.).

>beta-defensin 103 (CBD103)scaffold:turTru1:scaffold_83571:21060:22135:1
CTTTAATACCAAGAGGAAAAATTGACAGTCGAGAACAGAAGCACAGACTTCCCTTTTATTAGCTTAGTG
GTTGGTGGTAGAGTCTTTTGGCCTCTAGAGAGTCCCGGGGGACCACCGTGGCCCCGGCCTCCCAGGTGT
CAGTTCACGAGGGGGGCTGCCTGTGGTCATGCCACTTTCTCAGCACTCCGGCGGTTTATGACTGAGCTGA
ACCCATGGACGCGTTGTTAGTTGCTTAGTCCCACCTTCTACTTTACTCCTGATGGCATAGAAAAAGGAAA
AAGAAATAAAAGAAAGGCCACAGATCAGTTACGTGAGGAATTCCAAAAGCCTAGGTCTTGCACTTCCTA
AAACCTTCCCGTGCTTTCCGTTTCTTCATTGCAGGCAGTGGAACAATCGTAAATATATTACAAAAGTATT
ATTGCAGAGTGAGAGGCGGCCGGTGTGCTCTGTTTGGCTGTCTTCCGAGGGAGATACCCTTAGGCCTTTG
TTCCCACAGTGGTCGAAAATGCTGCCGAAGGGAGAGATGAAAAATCCAGAAGCGCGATGAGAACGTTG
CCAAGTGTGAAAATGCTCCTGCGGGGTCTATAGAAGCCAAATCAAATTAAAGATTTTTCACAAAAAAAT
TAGAAGGTGGATTATTTTTTTTTCTTTTAATTCTGGTTGTTGTGCTAGTTGTTACCAAACAGGCAACTTAG
GAAGTGCTGGTGACTCACCTTTAATTCAAATAAAGTATTATCTCTAGAGCAAGAGGATCTGAACTGATG
TTCCCAGGGTTATCAGACTTTATTCCTCTTTTGCAAATCTTGCACCCCAGTGACACAGGACTCCTCGTTCT
TCCTGAAATATTTTCTTCACTTACTTGTCACCCATTTTGAGTTTCCAGTTCATCTATAAGCCATAGCCTCT
CCCCATCAAATCTATGTCTCATTTTTTGATGATGGCCATTCTGACTGGTGTGAGATGATATCTCATTGTGC
TTTTGATTTGCATTTCTCTAATGATTAGTGATGTTGAGCATCCTTTCATGTGCTTGTTGGCAATCTGTATA
TCTTCTCTGGAGAAATGTCTATTTAGGT

>Pax7 scaffold:turTru1:scaffold_114323:104712:105911:-1
TCTTCCCGGACCGGGGCACGAACCCGTGTTCCGTGCATCGGCAGGCGGACTCTCAACCACTGCGCCACC
AGGGATACCCCCTTGCCTTCTTCTTAAAGTCAGGTTTATTGAGGTGCAGTGAATATTCAGTAAAATTCTC
CCCCTTTTGGAGAATATAGTTCTAGGAATTTTGACAAGTGTGTTTCCTCATGAAACCATCCCCTGAGTGC
CATGTCATGGACATCCTTCCGTGTCAGTAGAAATAACCTAAATGCCCAACTACAGGGGTTATTTTTTTAA
AGGAATGGCTGGGCCATAAGGTATTTAACCAATCCCCTAAGGTTGGGCAGAGAGGTTGCTTCCCAGCCT
TCACGCTGATAAATAGTGCCGGGGGGGGAACATCCTTCCTGTGCTCAGGAGGGATCTGAGGCTCAGTGCC
GTGGTGGCCTGCAGGTATGCACAAAGCACCCGTCCATCTCACTGGGCCCCCACCCTCTTCCTCCCTCTCA
CCTGGCTGTAGGCAACCGTCTGGATGAGGGCTCGGACGTGGAGTCAGAACCCGACCTCCCTCTGAAACG
TAAGCAGCGCCGCAGTCGGACCACGTTCACAGCCGAGCAGCTGGAGGAGCTGGAGAAGGCCTTCGAGA
GGACTCACTACCCGGACATCTACACCCGGGAGGAGCTGGCGCAGAGGACCAAGCTGACGGAGGCACGC
GTGCAGGTGAGGGGGCGCCTGGGGCTTTGTGCTGTGGGCATCATCCGGAGACCAGGATTGGGAAGGTG
GGAGTGGCCAGGGTCAGAGGGTTTAGTGGGCACTCCCCAGGGTGGGGATGGAGGTTCTGCAGCAAGTG
ATCCAGGAACCGTGTAGGCAAGTATTGCCCTGGCACTTGGTACAGAGTAAATGGGCAGCGAATGTTGAC
TGATTAACTAATCTCTAGGCAGTCCCAGGGAAGCCAGGAGCGGCAGTGAGGACTTGGATCGTGGGGGG
GCGGGGAAGGGGGGAAGGAGGGATGGAGGAGAGACAGAAAGGGAGAGGAGAAAAGGGTGAGGGAGG
GAGGAAGGAAGGACATTGGTTCAGCCACAAGAGAGGGAGCTCGGGGAGTAGAAAAGTCTCCAAATCTC
TGAGGAGGGGAGGAAGTAATCAGGACACACATTCCAGAAGAGACGTAGCTTCCTAAGATGGGTAATGA
GTCTCCTGTTATGAGAAGCATGCAAGCCGACTG

Figure A1-S1. List of additional nuclear gene sequences used in capture array design (cont.).

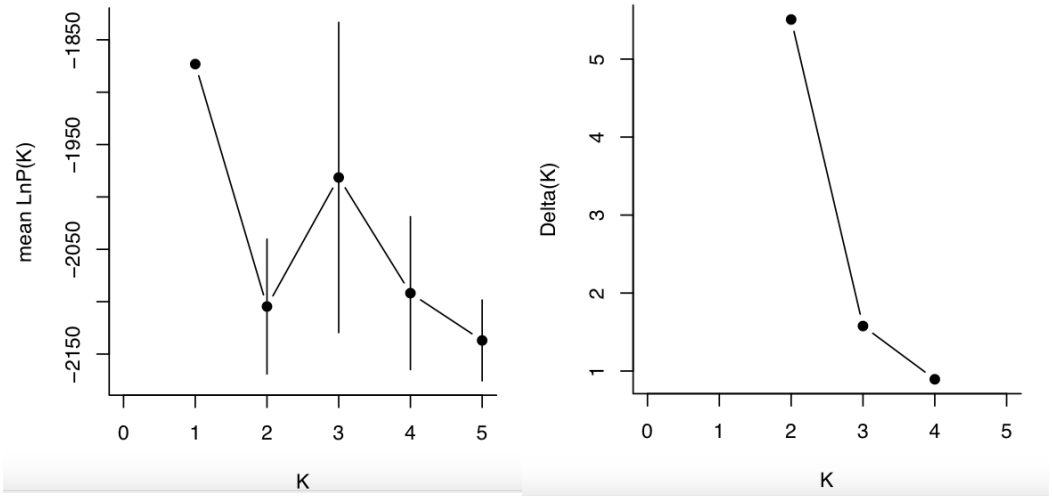
>OCA2 genescaffold:turTru1:GeneScaffold_613:187212:188330:-1

CAGGCATGGTGTGAGACACTGGGTGGCGGTGCTTTATGCTCTAGCCTCATCCTATTCAGAGGGGTCTGGT
GGCAAAGATGCATTCATCTGTGCCAGGGCCTTCCCTGCACAGGGTCCTGCACCCTGAGTGCTTCCTCCAG
ACCTGTGGGGGCGAGGCAGGAGTATATTTTCACATGGATGGTTCAGATGGGTTTTATGTTAATTTTCAGA
GAAATCTAATGCTGAGAGTCAAATTTATCCAGTAGCTCTGTAGCACTGTTGTCGTTTTTTACCCCTCAAC
ATTGCCCTTCTCACAGGTGTGTGAGTGTGCGTGCACATGTGCACAAGCACATGCACACACATGCCTCCCT
TGTACAAGCAAATGCTTTAAAAAAAATTAAGTCCCACATGGGCAATTCACAATGTGAATAGTGGGACTG
TTTTCATTTTCCTCCATTCGTGACAGATTGTGTGAGGTGCTCAATCTCGATCCAAGACAAGTCCTGATTG
CAGAAGTGATCTTCACAAACATTGGAGGAGCTGCCACTGCCATTGGGGACCCACCAAATGTCATTATTG
TTTCCAACGAGGAGTTGAGGAAAATGGTATGTATCAGTATAACAGGGTTGCTTCCAATAAACATAAGTG
TGCAGCTTACCTGCACACTTAATAAAATTTTACCCTAAATGGATGTCACTTCTTCATATGAAGTGCTTTT
GGTCATCTTGAGAGTGGGAAGTTAGGTGACTGATCCGCTAATTTAGAGTTAAGTGGTCTGAATTGATTTC
ATTCTTTTTTTTAAACATCTTTATTGGAGTATAATTGCATTACAATGGTGTGTTAGTTTCTGCTGTATAAC
AAAGTGAATCAGCTATACATACACATATATCCCCATATCTCCTCCCTTTTGCGTCTCCCTCCCACCCCTCT
AGATGGACACAAAGCACTGAGCTGATCTCCCTGTGCTATGCAGCTGCTTCCCACTAGCTATCTATTTTAA
ATTTGGTAGTTTATATATGTCAATGCCACTCCCTCACTTTGTCCCAGCTTATCCTTCCACATCCCCGTGTC
CTCAAGTCCATTCTCTACATCTGCATCTTTATTCCTGTCCTTTCCCAGGTTCTTCAGAACCTGGTTTTT

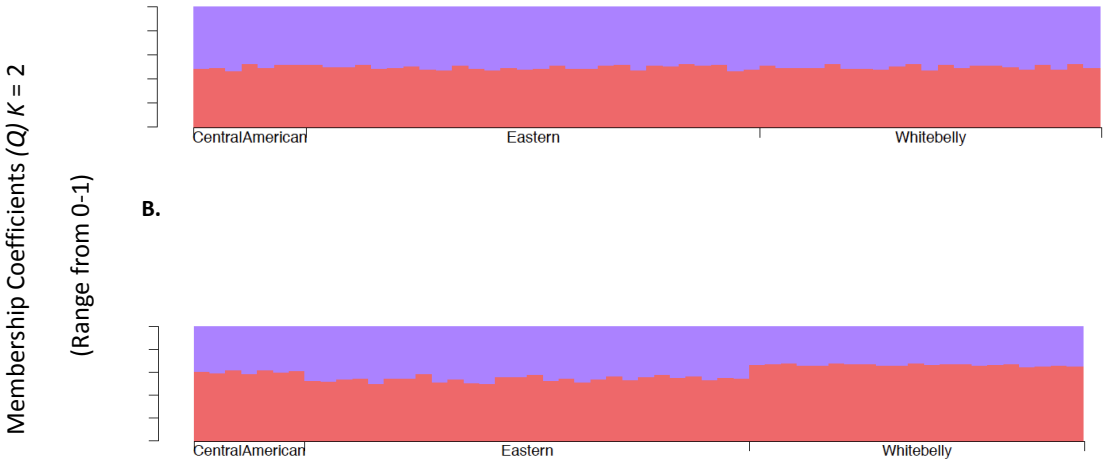
>BCDO2 - genescaffold:turTru1:GeneScaffold_3620:729737:730941:1

ATATGGACGAAAATGATTATCCAGTAAGGGTAGTCATGGGTAGTCATCTTACTTGCTTACTTTAGCATAG
TTCTGATAAATTTGAAAAAATTTTAAATAAATAGAAAATGTAATATGTATAAGATATGAAGATAACTAT
CTCTAAAAGTTTTGGAATCCATTTTAATTTGAGAAATGTCAAGCATTCCATATGATGATATCTTGGCTAT
AGGTTAAAGACAGAATTATCTTAGGCATCACCAGCTTAGGGTTTTTCTTTAGCCCTCTGATGTGTGAGCT
TTTTGAGAAGGGGTAAGGGTCAAGAGAATGTAAGTTGCTCTTTCATTGTTGTACTCCAGGATATACTCAA
AAGCTCTCAACCCTTAATTTCAAGTTCTTTGACTGAGCAGGTTCTGCTGGGAAGCCATCCCATTTCCTTA
TTTCTGTATTTGAAATCTCTCATTTTGACCCCATGTTTATCTTCCTAATATTTATTTTTCTATTTTTGCTCCT
GTTTAAAGGTACAATCACTGGTTTGATGGAATGGCTTTACTTCACCAGTTCAAGGTGGAGAAGGGCACA
GTGACATACAGGAGCAAGTTTCTACAGAGTGATACATATAAGGCCAACAGTGTTCATGATCGAATTGTG
ATCTCAGAATTTGGCACATTGGCTCTCCCCGATCCATGCAAGAATGTTTTTGAACGTTTCCTGTCAAAAT
TTGAACCGCCTGGTAAGCAGCCTTTAATATGCTCAGAATTTGCAGGGGATAGTGATGATTCTTTAAATTA
TTATGATATATGTCAAGCTATATAAATATGAATATTGAAATATATATTTTTTTGATATGAGAGGATGTTT
AAACTGCTTTCTCCCTTGAGAAAGCTGTGGTGTTTAAATAACTCTGGGGTTTGATTTCAGCAATTACTGA
CAACACCAATGTCAACTATGTGCAGTACAAGGGTGATTACTACATTAGTACGGAAACCAACTTCATGAA
CAAAGTGGACATTGAAACCCTGGAAAAAACAGAAAAGGTAAAATATAGGTTAAAAACGAATATAATAA
ACCCTAAAATTTTTCTCCATTCCAAATACTTTGTAGAGTAGCTAATATCTGAAGAAAAAATTGGATCCTT
AAACAAAACCAAAGTTACCCACTGATTCATATCATTGTGTACCCCCTGTCCTGGTTAGAAGGGCCTGTG
GGGTTGGGGTTTTCCTTAA

Figure A1-S1. List of additional nuclear gene sequences used in capture array design (cont.).



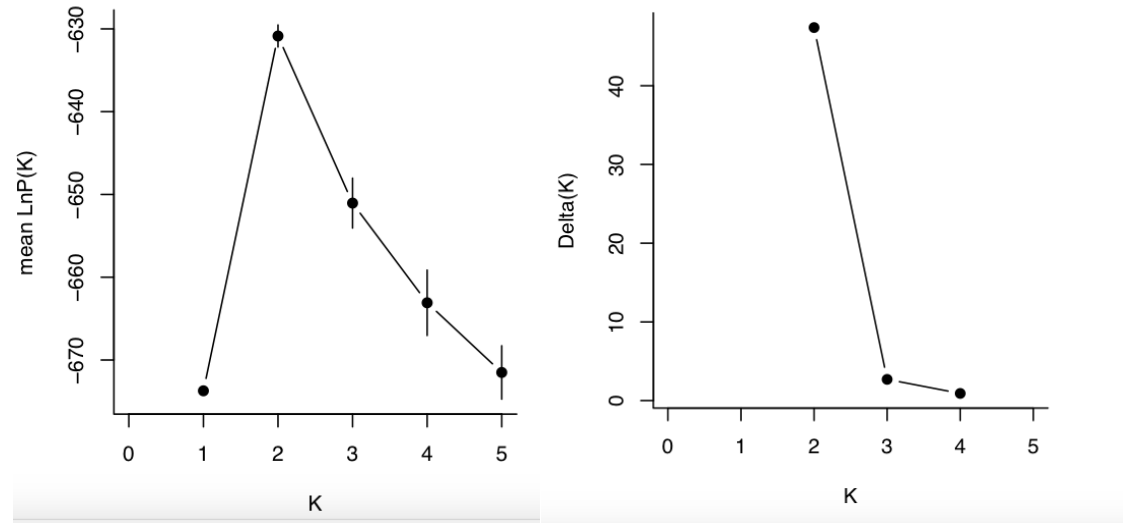
A.



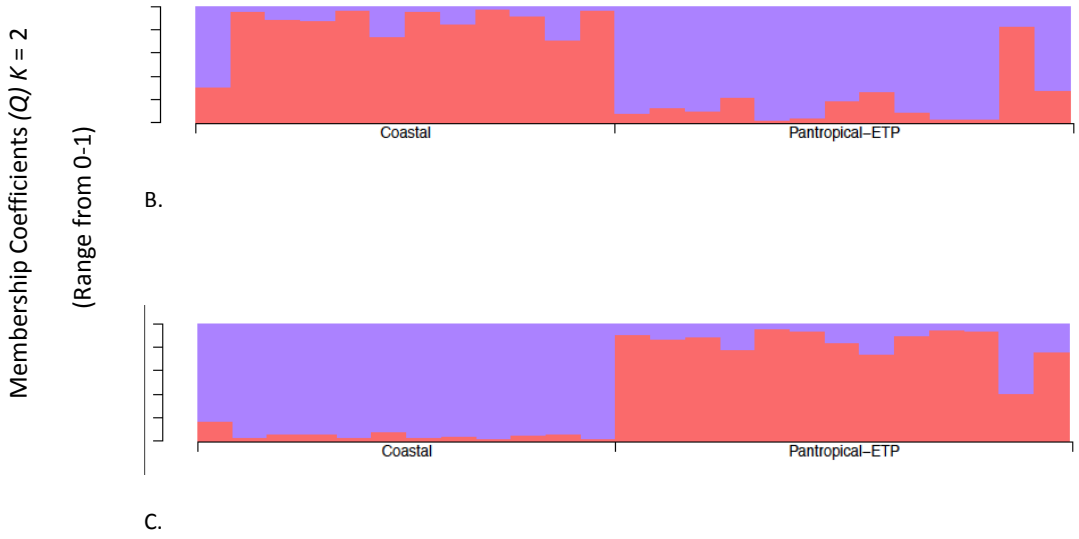
B.

C.

Figure A1-S2. A. The mean log likelihood of the spinner dolphin SNP data (LnP(K)) and the second-order rate of change in LnP(K) (DeltaK), with respect to K. Estimates calculated in STRUCTURE for K = 2 through 5. B) Estimated membership coefficients (*Q*) plot for K = 2 without a location prior. C. Estimated membership coefficients (*Q*) plot for K = 2 with location prior.



A.



B.

C.

Figure A1-S3. A. The mean log likelihood of the spotted dolphin SNP data (LnP(K)) and the second-order rate of change in LnP(K) (DeltaK), with respect to K. Estimates calculated in STRUCTURE for K = 2 through 5. **B.** Estimated membership coefficients (*Q*) plot for K = 2 with no location prior. **C.** Estimated membership coefficients (*Q*) plot for K = 2 with location prior.

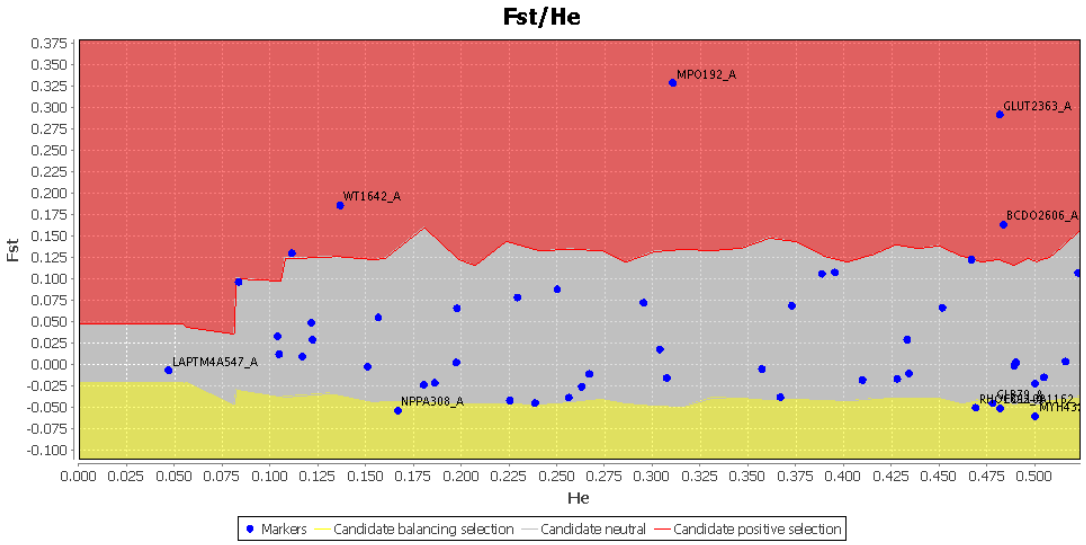


Figure A1-S4. F_{ST} outlier test for selection in ETP spinner dolphins using the *Lositan* workbench based on 75M simulations. F_{ST} for all 51 loci are plotted against the heterozygosity. Four loci were estimated to be under positive selection (BCDO, GLUT, WT, MPO), shown in red section; five were shown to be under balancing, shown in yellow. BCDO is a gene with known function in coloration. Candidate neutral loci are shown in the grey area.

Table A1-S3. Test for selection on 51 nuclear SNPs in spinner dolphins (*S. longirostris*).

Locus Name	Het	Fst	P(Simul Fst<sample	Result	Origin
Actin274	0.111111	0.130435	0.973628	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
ADH2758	0.409641	-0.017609	0.307102	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
AMBP1143	0.121296	0.049073	0.805247	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
ASIP326	0.372727	0.069385	0.870835	Neutral	Coloration (reviewed in Hubbard et al. 2010; Rieder et al. 2001;)
BCDO2606	0.483333	0.163708	0.989492	Positive	Coloration (Eriksson et al. 2008)
BGN165	0.083333	0.096774	0.966524	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
BTN554	0.25	0.088337	0.91075	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
C51138	0.433048	0.029608	0.701873	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
CGA166	0.490029	0.003092	0.593723	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
CHRNA1305	0.116667	0.009667	0.589659	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
CHY146	0	-100	0.5	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
CK222	0.434028	-0.009689	0.399385	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
CKMM167	0.52246	0.107512	0.910514	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
COL10A1162	0.481703	-0.050251	0.000151	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
CYO19346	0.229167	0.078788	0.88794	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
DRD2279	0.307217	-0.015057	0.344578	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
FES561	0.295113	0.072542	0.886081	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
Fibrinogen24	0.121981	0.029326	0.714508	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
G6PD298	0.489035	-0.000534	0.56403	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
GLB79	0.477778	-0.044578	0.007902	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
GLUT2363	0.481481	0.291855	0.999784	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
GNAS614	0.180093	-0.023055	0.1964	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
HGBA905	0.156381	0.055353	0.804158	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
HOXc8692	0.504561	-0.01426	0.273789	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
IFNG885	0.150727	-0.002025	0.374916	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
Lactalbumin5	0.266782	-0.010418	0.405378	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
LAPTM4A54	0.046667	-0.006195	-100	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
LF582	0.388406	0.106418	0.949297	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
MATR3691	0.515873	0.004103	0.478604	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
MC1R446	0.197181	0.002978	0.439682	Neutral	Coloration (Rieder et al. 2001; Hoekstra et al. 2006; reviewed in Hubbard et al. 2010)
MPO192	0.310417	0.328859	0.999276	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
MYH4328	0.5	-0.059797	0	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
MYL4984	0.427824	-0.016146	0.322161	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
NPPA308	0.166667	-0.053169	0.000244	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
OCA2407	0.5	-0.021485	0.34478	Neutral	Coloration (reviewed in Hubbard et al. 2010)
Pax7238	0.303624	0.018339	0.634601	Neutral	Coloration (reviewed in Hubbard et al. 2010)
PGK1511	0.357087	-0.004643	0.455093	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
PIM171	0.262787	-0.025274	0.208521	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
PKM223	0.225238	-0.041276	0.051437	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
PLP509	0.197531	0.066216	0.788129	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
RDS611	0.395202	0.107948	0.951117	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
RHO1692	0.468842	-0.049717	0.011647	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
RYR2465	0.451471	0.066757	0.872432	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
SPTBN1465	0.185897	-0.020844	0.219142	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
TCRA1174	0.103704	0.033752	0.653326	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
TCRB178	0.256039	-0.037932	0.072035	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
TOP1893	0.366667	-0.037397	0.051564	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
TPI1236	0.466667	0.12289	0.966634	Neutral	Coloration (reviewed in Hubbard et al. 2010; Rieder et al. 2001)
TYRP11007	0.238462	-0.044331	0.036015	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.
UBE1553	0.104497	0.012658	0.542809	Neutral	Y chromosome – T. truncatus (Aitken et al. 2004; Lyons et al. 1997)
WT1642	0.136364	0.186046	0.995041	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al.

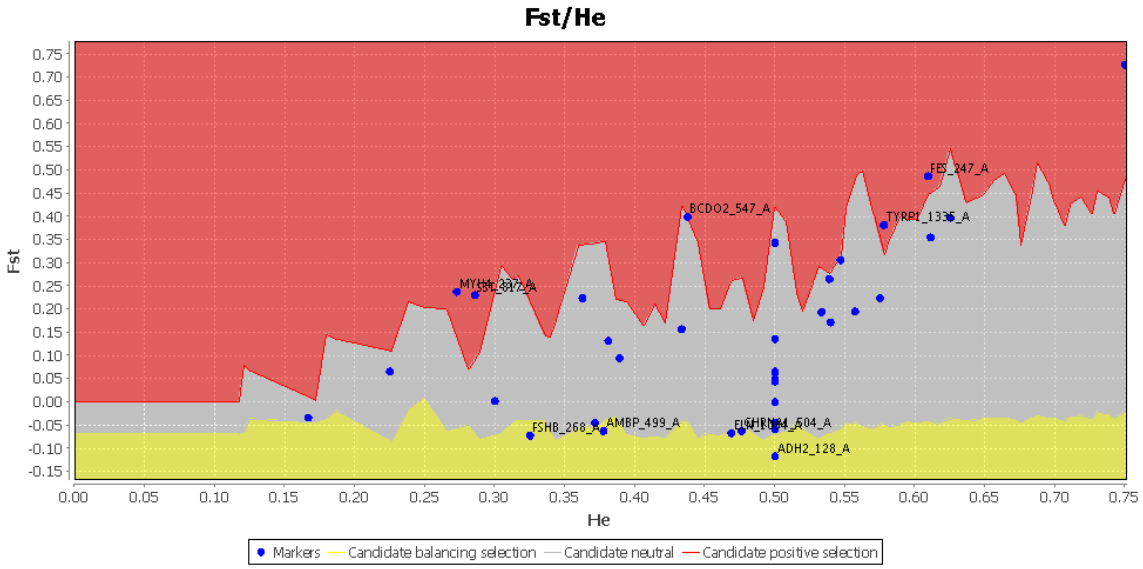


Figure A1-S5. F_{ST} outlier test for selection in ETP spotted dolphins using the *Lositan* workbench based on 75M simulations. F_{ST} for all 36 loci are plotted against the heterozygosity. Six loci were estimated to be under positive selection (BCDO, FES, HOXc8, MYH4, SST, TYRP1), shown in red section; five were shown to be under balancing (ADH2, AMBP, CHRNA, ELN, FSHB), shown in yellow. BCDO and TYRP1 are genes for coloration. Candidate neutral loci are shown in the grey area.

Table A1-S4. Results of tests for selection on 36 nuclear SNPs in our spotted dolphin (*S. attenuata*) dataset.

Locus	Het	Fst	P(SimFst<sample Fst)	Simulation Result	Origin
ADH2128	0.5	-0.116883	0	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
AMBP499	0.377551	-0.06237	0	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
ASIP598	0.575	0.224983	0.84332	Neutral	Coloration (reviewed in Hubbard et al. 2010; Rieder et al. 2001)
BCDO2547	0.4375	0.4	0.991676	Positive	Coloration (Eriksson et al. 2008)
BTN434	0.611111	0.355372	0.915438	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
CAT352	0.166667	-0.033333	0.231214	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
CHRNA1504	0.47619	-0.062937	0	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
CHY146	0.538889	0.265807	0.959939	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
CKMM403	0.5	0	0.450598	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
DRD2200	0.388889	0.0953	0.706798	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
EDN1354	0.546875	0.306667	0.884835	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
ELN1004	0.46875	-0.066667	0	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
FES247	0.609375	0.487179	0.977508	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
FSHB268	0.325	-0.072296	0	Balancing	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
GBA134	0.557143	0.195924	0.761232	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
GNAS199	0.539683	0.172212	0.834206	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
HGBA688	0.5	0.044444	0.546065	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
HOXc8395	0.75	0.727273	0.99607	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
INT548	0.5	0.064593	0.5802	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
LAPTM4A106	0.533333	0.194014	0.872747	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
LF466	0.5	-0.058333	0.056894	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
MPO1006	0.5	0.343434	0.880361	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
MYH4237	0.272727	0.238095	1	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
MYL4784	0.3	0.002392	0.320866	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
ODC128	0.5	-0.043478	0.246566	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
Pax7439	0.5	0.065934	0.582349	Neutral	Coloration (reviewed in Hubbard et al. 2010)
PIM172	0.625	0.398039	0.931291	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
PIT1301	0.433333	0.157343	0.962393	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
RDS637	0.380952	0.132212	0.760149	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
SPTBN1386	0.225	0.066277	0.464187	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
SST817	0.285714	0.230769	1	Positive	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
TCRA1293	0.371429	-0.045157	0.235925	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
TCRB583	0.3625	0.22444	0.679178	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
TPI1778	0.5	0.136264	0.679374	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)
TYRP11335	0.577778	0.381758	1	Positive	Coloration (reviewed in Hubbard et al. 2010; Rieder et al. 2001)
WT1529	0.5	0.048848	0.553843	Neutral	Tursiops truncatus Conserved Anchor Tag orthologs (Aitken et al. 2004; Lyons et al. 1997)

Table A1-S5. Pairwise F_{ST} estimates for coloration genes found to be subject to positive selection.

Species	Locus	strata.1	strata.2	n.1	n.2	*Fst*	Fst.p.val
S. attenuata	BCDO2547	Coastal	Pantropical-ETP	7	8	0.3891	0.0047
	TYRP11335	Coastal	Pantropical-ETP	9	10	0.3704	0.0059
S. longirostris	BCDO2602	Central American	Eastern	5	21	0.3298	0.0259
	BCDO2602	Central American	Whitebelly	5	10	0.1522	0.1699
	BCDO2602	Eastern	Whitebelly	21	10	-0.0231	0.5353

APPENDIX 2.

SUPPLEMENTARY MATERIAL

CHAPTER 3

Table A2-S1. Alpha-scores simulated using datasets for testing hypotheses of structure in ETP spinner (*k* = 4) and spotted dolphins (*k* = 3). An a-score close to one is a sign that the DAPC solution is both strongly discriminating and stable, while low values (toward 0 or lower) indicate either weak discrimination or instability of the results. "Method for # of PCs" indicates the method used to choose the number of PCs for the DAPC analysis ("N/3" is the highest recommended and is equal to the sample size divided by 3; "Simulated" is the number of PC that yielded the highest mean a-score based on 10 simulations of the dataset.

Species	Pop1	Pop2	Pop3	Pop4	mean	#PCs	#DAs	Method for # PCs
S.longirostris	**whitebelly**	**eastern**	**Central American**	**Tres Marias**				
	0.38000000	0.07777778	0.16666667	0.34166667	0.2583333	24	3	*N*/3
	0.7133333	0.0250000	0.1444444	0.7916667	0.4186111	5	3	Simulated
S.attenuata	**northeastern**	**western-southern**	**coastal**					
	0.2333333	0.2312500	0.2111111	*NA*	0.2252315	19	3	*N*/3
	0.76666667	0.56250000	-0.05555556	*NA*	0.424537	1	1	Simulated

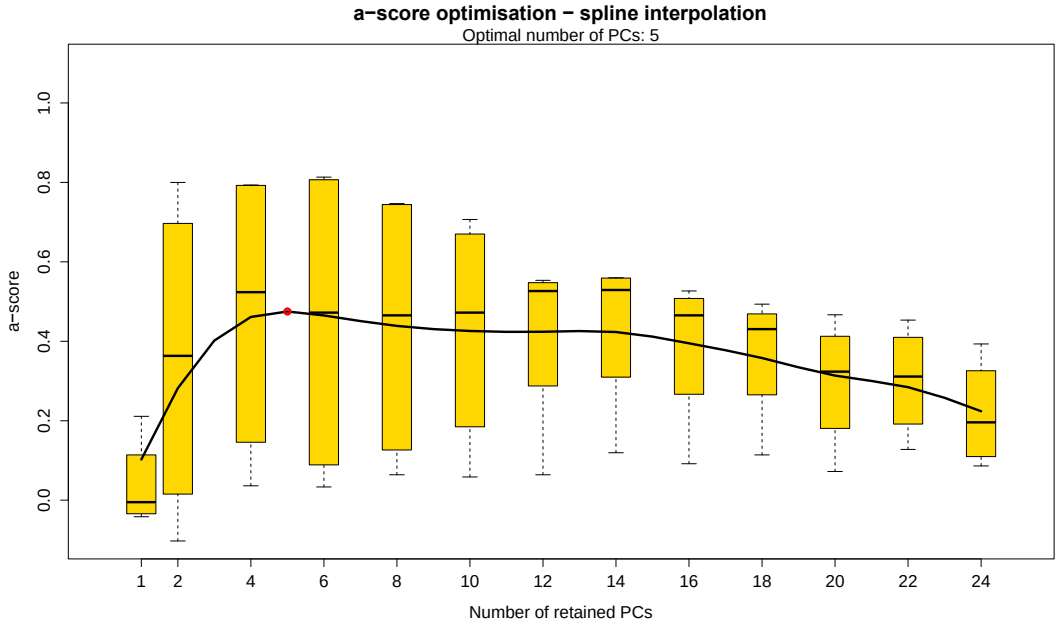


Figure A2-S1. Alpha-score optimization for the spinner dolphin dataset – spline interpolation for PCs 1 through 24. Box and whiskers show the overall mean and variance of individual population a-score for each of the PCs representing the ETP spinner dolphin dataset based on 10 simulations. Five PCs had the highest estimated mean a-score for the dataset and was chosen as the "optimal" is subsequent analyses of DAPC. Twenty-four PCs was also chosen as a "highest recommended" number of PCs (the number of PCs should be <N/3) because it showed only marginally lower mean a-scores and less variance.

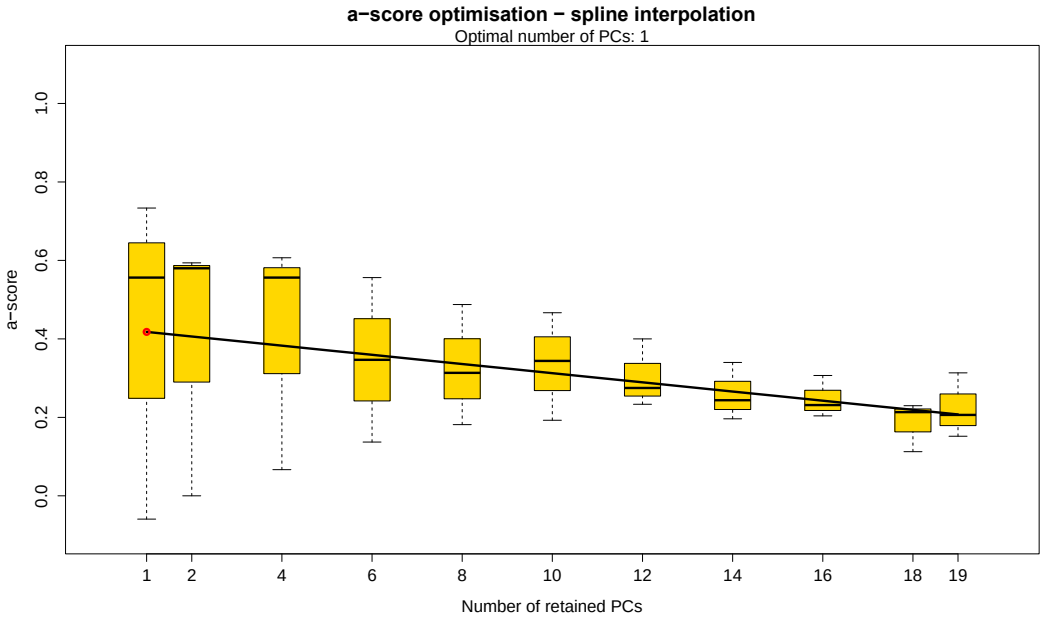


Figure A2-S2. a-score optimization – spline interpolation for PCs 1 through 19 for ETP spotted dolphin data. Box and whiskers show the overall mean and diversity of individual population a-score for each of the PCs representing the ETP spotted dolphin dataset based on 10 simulations. One PC had the highest estimated mean a-score for the dataset and was chosen as the "optimal" is subsequent analyses of DAPC. Seventeen PCs was also chosen as a "highest recommended" number of PCs (the number of PCs should be <N/3) because it showed only marginally lower mean a-scores and less variance.

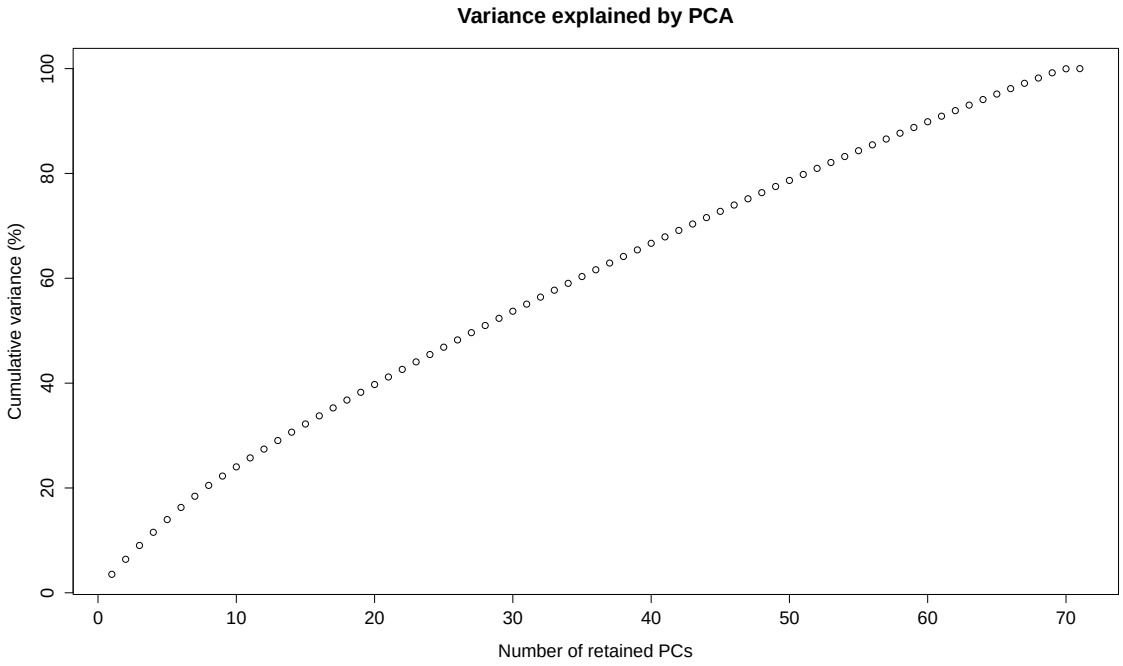


Figure A2-S3. Cumulative genomic variation explained by PC 1 through 70 in the ETP spinner dolphin dataset (maximum *k* set to 20).

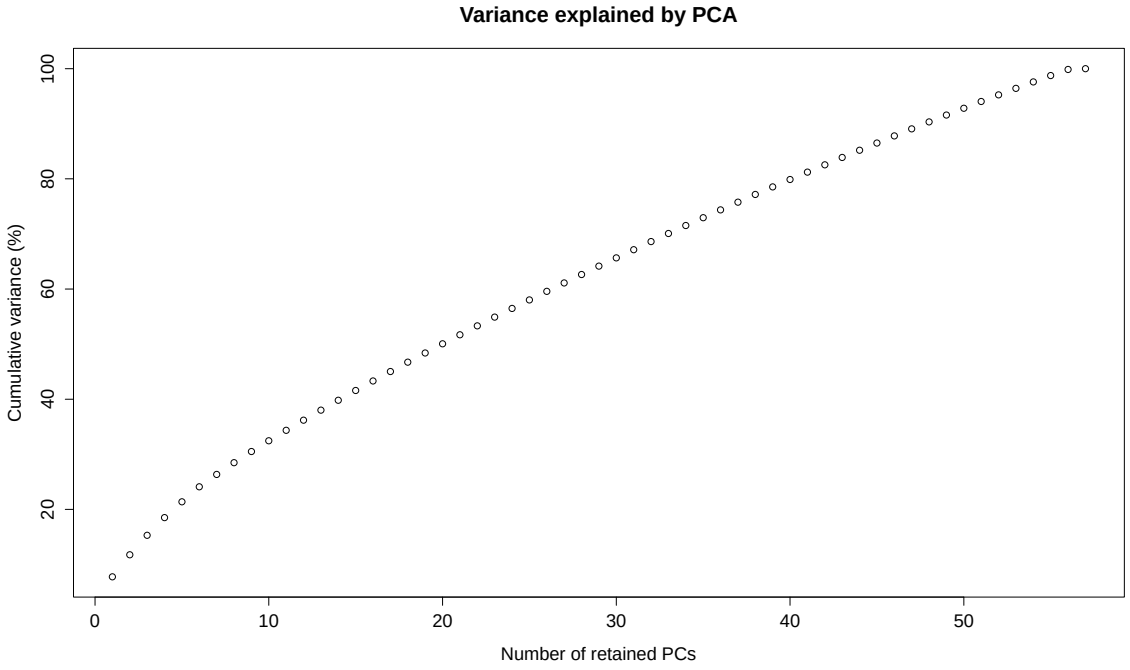


Figure A2-S4 Cumulative genomic variation explained by PC 1 through 60 in the ETP spotted dolphin dataset (max *k* set to 20).

APPENDIX 3.

SUPPLEMENTARY MATERIAL

CHAPTER 4

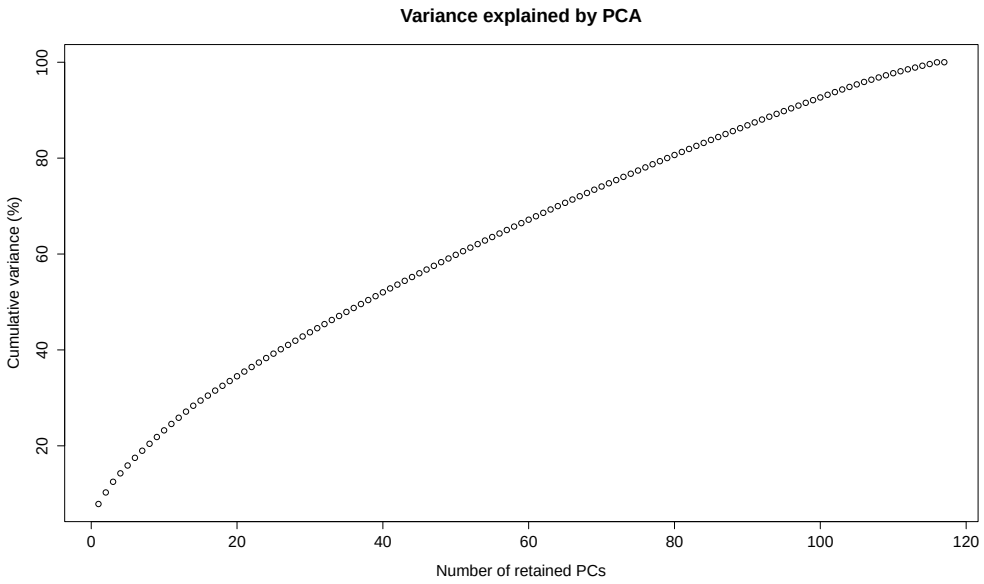


Figure A3-S1. Cumulative genomic variation explained by principle components 1 through 120 in the global spinner dolphin dataset.

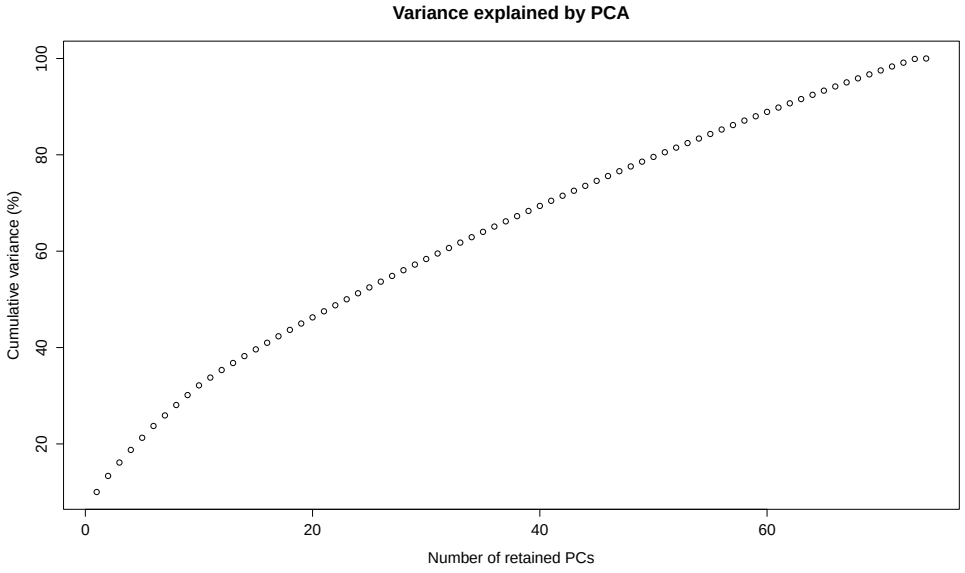


Figure A3-S2. Cumulative genomic variation explained by principle components 1 through 70 in the spotted dolphin dataset.

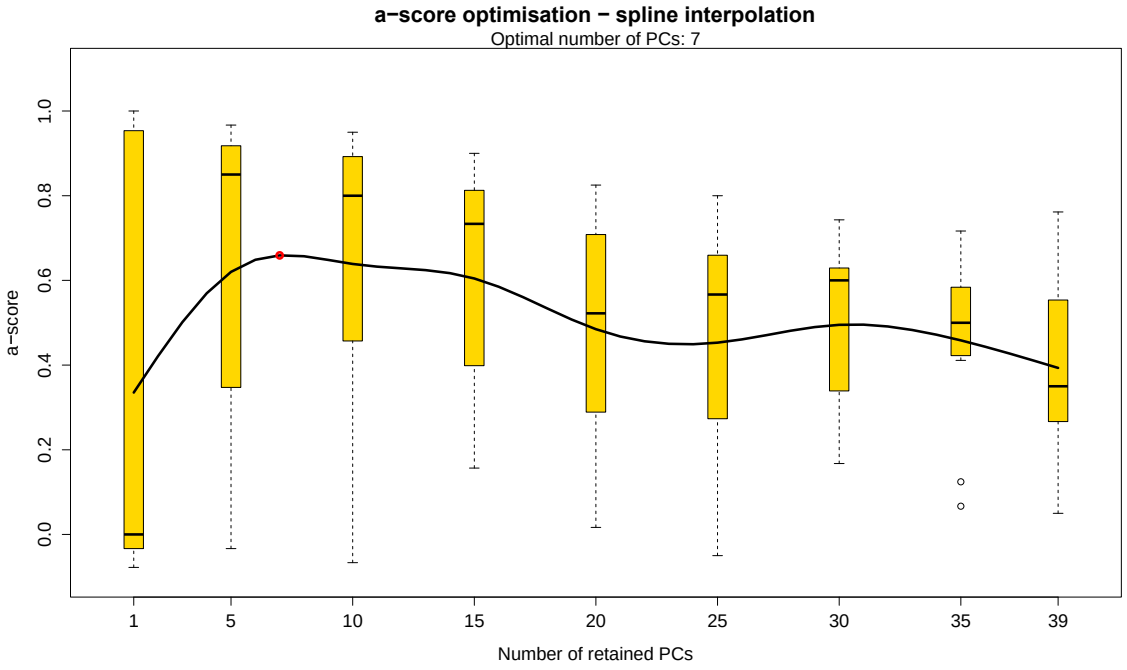


Figure A3-S3. Alpha (or a-score) optimization for the spinner dolphin dataset – spline interpolation for PCs 1 through 39. Box and whiskers show the overall mean and variance of individual population a-score for each of the PCs representing the global spinner dolphin dataset based on 10 simulations. Seven PCs had the highest estimated mean a-score for the dataset and was chosen as the "optimal" is subsequent analyses of DAPC. Thirty-nine PCs was also chosen as a "highest recommended" number of PCs (the number of PCs should be <N/3).

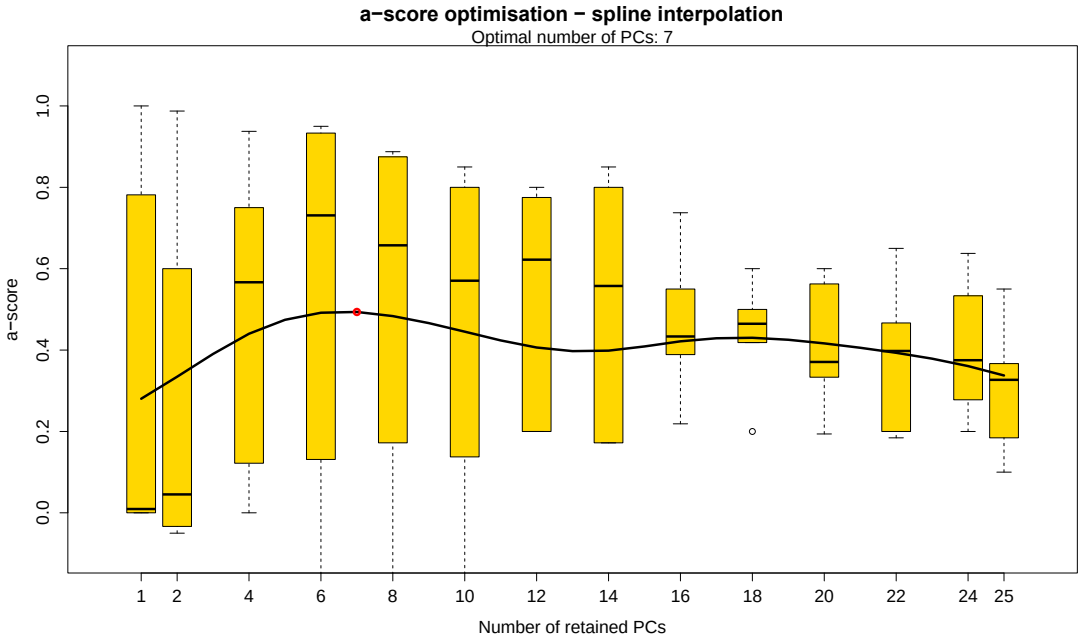


Figure A3-S4. Alpha (or a-score) optimization – spline interpolation for PCs 1 through 25 for spotted dolphin data. Box and whiskers show the overall mean and diversity of individual population a-score for each of the PCs representing the ETP spotted dolphin dataset based on 10 simulations. Seven PCs had the highest estimated mean a-score for the dataset and was chosen as the "optimal" is subsequent analyses of DAPC. Twenty-five PCs was also chosen as a "highest recommended" number of PCs (the number of PCs should be <N/3).

Table A3-S1. List of spinner dolphin samples subset from the total dataset and used for global phylogeographic analyses in *SNAPP*.

LabID	subspecies	Slon_GBL_22	Latitude	Longitude	SEX
367	roseiventris	DWF_AUS	-12.633333	126.4	M
372	roseiventris	DWF_AUS	-12.6	127.3	M
392	longirostris	PHIL	8.85	123.116666	F
400	longirostris	PHIL	8.85	123.116666	U
462	longirostris	ATL	36.383333	-75.816666	M
2138	longirostris	WB	8.6	-138.216666	M
2757	longirostris	ATL	28.666666	-87.766666	M
4095	longirostris	WB	8.6	-138.216666	M
7185	longirostris	HI	20	-155.833333	M
7202	longirostris	HI	19.666666	-156.083333	M
9847	longirostris	MALDIVES	4.766666	73.5	F
9854	longirostris	MALDIVES	6.933333	73.233333	M
11664	centroamericana	CA	15.866666	-94.833333	F
24928	orientalis	EAST	10.2	-98.533333	F
38017	orientalis	TM	23.633333	-107.383333	F
38018	orientalis	TM	22.25	-106.4	F
38050	orientalis	EAST	21.45	-106.066666	M
38191	centroamericana	CA	12.583333	-88.533333	M
47281	longirostris	ZAN	-4.866666	39.65	U
47283	longirostris	ZAN	-4.866666	39.65	U
79908	roesiventris	DWF_INDO	2.266666	118.3	U
79923	roseiventris	DWF_INDO	1.3	118.75	U

Table A3-S2. List of spotted dolphin (*Stenella attenuata*) samples subset from the total dataset and used for global phylogeographic analyses in SNAPP.

LabID	subspecies	Satt_GBL_reduced	Latitude	Longitude	SEX
130	attenuata	ETP.Off	14	-96.083333	U
2081	attenuata	ETP.Off	13.283333	-128.766666	F
2092	attenuata	ETP.Off	12.8	-131.183333	F
9860	attenuata	MAL	7.183333	73	U
11371	graffmani	Coastal	22.666666	-106.5	M
11381	graffmani	Coastal	16.45	-99.3	M
11393	graffmani	Coastal	14.833333	-93.183333	M
11921	graffmani	Coastal	11.466666	-86.566666	F
11931	graffmani	Coastal	9.433333	-85.466666	M
11950	graffmani	Coastal	7.65	-82.333333	M
12039	graffmani	Coastal	0.183333	-80.5	F
24047	attenuata	ETP.Off	5.866666	-90.066666	M
24937	attenuata	ETP.Off	10.233333	-98.566666	F
30485	attenuata	HI	21.7	-157.616666	F
55197	attenuata	HI	19.283333	-156	F
56750	attenuata	HI	NA	NA	M
75657	attenuata	HI	19.4	-156.033333	M
79906	attenuata	INDO	2.3	118.466666	U
79912	attenuata	INDO	2.183333	118.533333	U
79915	attenuata	INDO	2.366666	118.433333	U
108200	attenuata	Guam	13.7	144.816666	U
108201	attenuata	Guam	13.7	144.816666	U
116842	attenuata	NMI	14.126452	145.059145	U
116844	attenuata	NMI	14.127214	145.076016	U
116849	attenuata	NMI	14.067994	145.209951	U

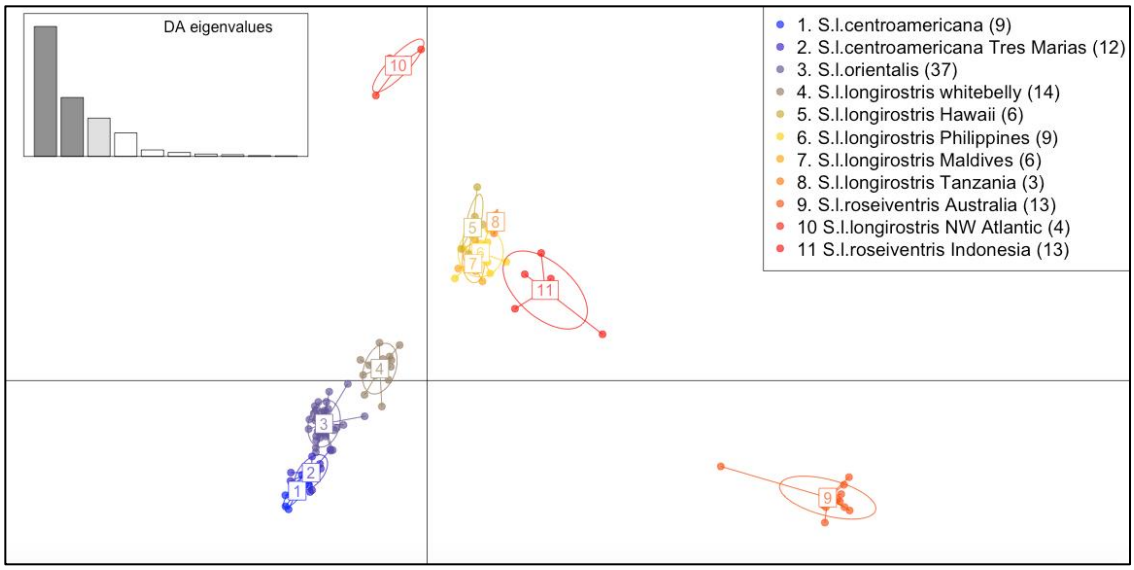


Figure A3-S5. Genomic variation across individuals and populations of spinner dolphins using 39 principle components and 3 discriminant analyses: Scatter plot of individuals based on the first two eigenvalues of the Discriminant Analysis of Principle Components (DAPC). Inset shows the amount of variation represented by the DA eigenvalues. Ellipses represent 67% of the variation for each population. CA is Central American spinner, TM is Tres Marias, WB is whitebelly spinner, ATL is Atlantic, IND is Indian Ocean (Zanzibar), PHIL are samples from the Philippines, DWF_INDO are samples of the dwarf subspecies from Indonesia, DWF_AUS are samples of the dwarf subspecies from Australia, HI is Hawaii.

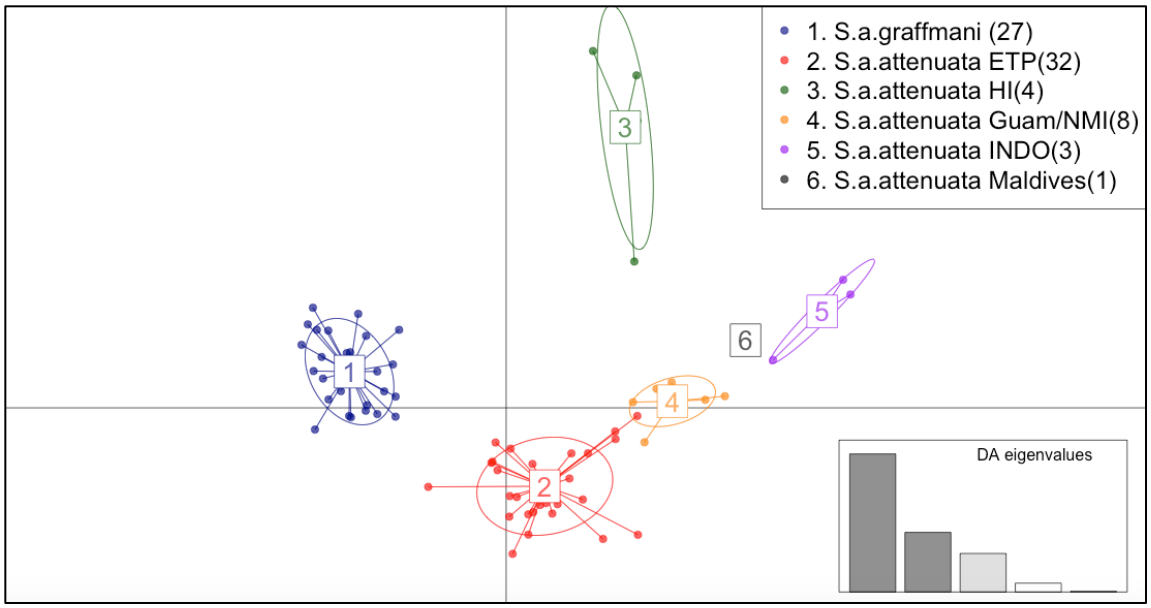
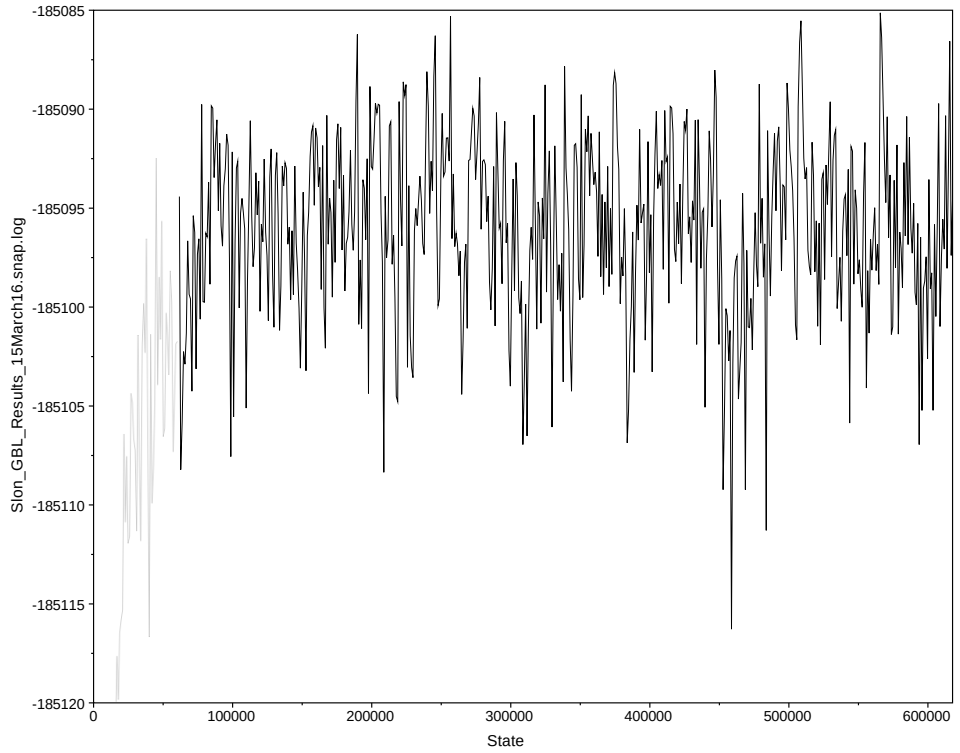
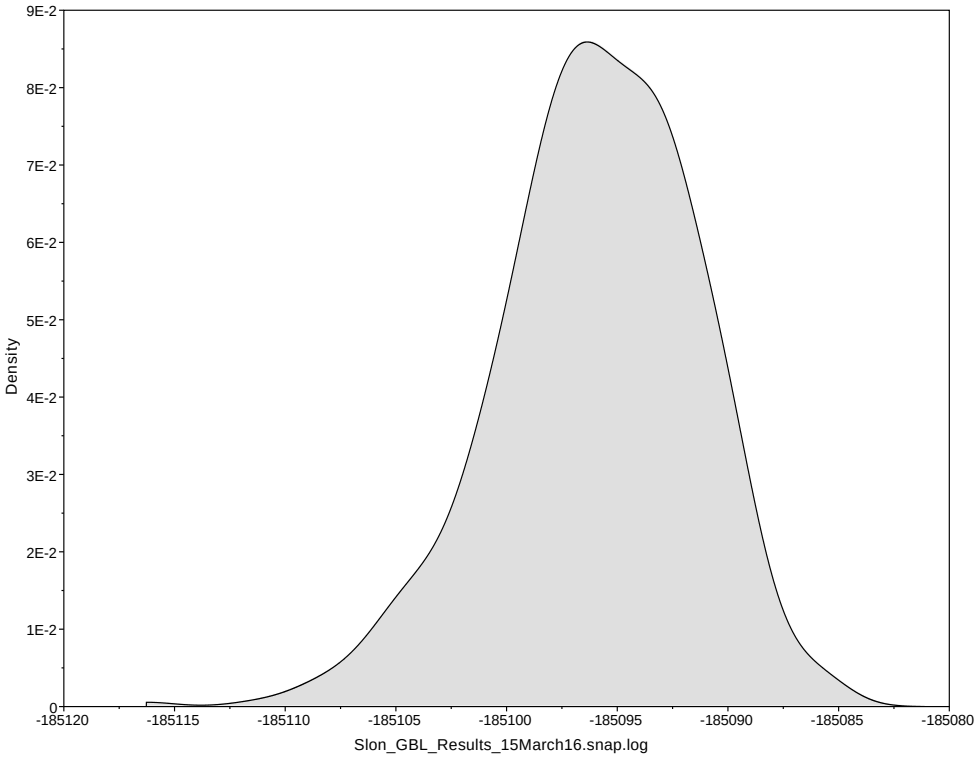


Figure A3-S6. Genomic variation across individuals and populations of spotted dolphins using 25 principle components and three discriminant analyses: Scatter plot of individuals based on the first two eigenvalues of the Discriminant Analysis of Principle Components (DAPC). Ellipses represent 67% of the variation for each population. Inset shows the amount of variation represented by the DA eigenvalues. 1 is Coastal ETP, 2 is Offshore ETP, 3 is Hawaii, 4 is Guam/NMI, 5 is Indonesia, 6 is Maldives. Maldives is represented by only one sample.

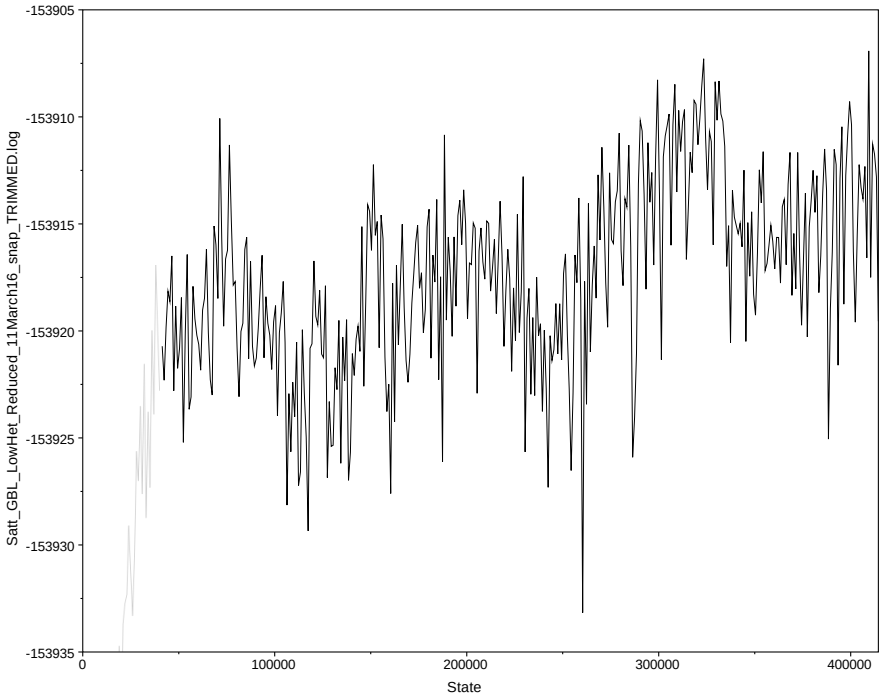


a)

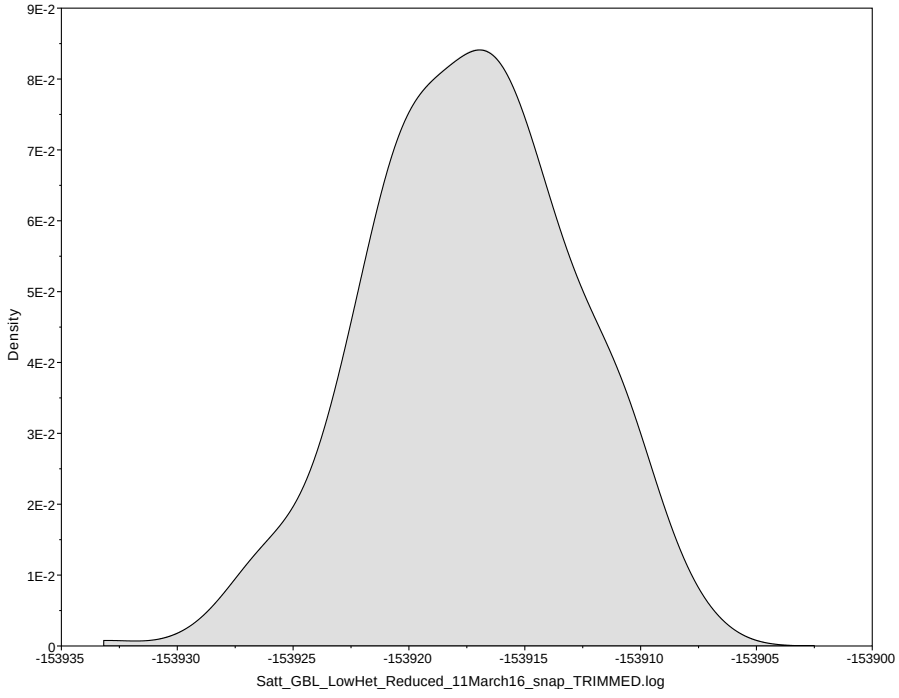


b)

Figure A3-S7. Trace (a) and density distribution (b) for spinner dolphin Bayesian phylogeographic analysis. MCMC run for 617,000 chains.



a)



b)

Figure A3-S8. Trace (a) and density distribution (b) for spotted dolphin Bayesian phylogeographic analysis. MCMC run for 414,000 chains.