Historical introgression drives pervasive mitochondrial admixture between two species of pelagic sharks

Shannon Corrigan^{1*}, Pierpaolo Maisano Delser^{2,3,4}, Corey Eddy⁵, Clinton Duffy⁶, Lei Yang¹, Chenhong Li^{1,7}, Adam L. Bazinet^{8,9}, Stefano Mona^{2,3} and Gavin J. P. Naylor^{1,10}

¹ *Department of Biology, College of Charleston, Charleston SC, USA.*

² *Institut de Systématique, Évolution, Biodiversité, ISYEB - UMR 7205 - CNRS, MNHN, UPMC, EPHE, Ecole Pratique des Hautes Etudes, Paris, France.*

³ *EPHE, PSL Research University, Paris, France.*

⁴ *Present address: Department of Zoology, University of Cambridge, Cambridge, UK*

⁵ *Department of Biology, University of Massachusetts Dartmouth, Dartmouth MA, USA.*

⁶ *Marine Ecosystems, Science and Policy, Department of Conservation, Auckland, New Zealand.*

⁷ *Present address: Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Shanghai Ocean University, Ministry of Education, Shanghai, China.*

⁸ *Laboratory of Molecular Evolution, Center for Bioinformatics and Computational Biology, University of Maryland, College Park MD, USA.*

⁹ *Present address: National Biodefense Analysis and Countermeasures Center, Fort Detrick, MD, USA.*

¹⁰ *College of Graduate Studies, Medical University of South Carolina, Charleston SC, USA.*

*Corresponding author: Shannon Corrigan email: corriganl@cofc.edu

ABSTRACT

We use a genomic sampling of both nuclear and mitochondrial DNA markers to examine a pattern of genetic admixture between *Carcharhinus galapagensis* (Galapagos sharks) and *Carcharhinus obscurus*, (dusky sharks), two well-known and closely related sharks that have been recognized as valid species for more than 100 years. We describe widespread mitochondrial-nuclear discordance in which these species are readily distinguishable based on 2152 nuclear single nucleotide polymorphisms from 910 independent autosomal regions, but show pervasive mitochondrial admixture. The species are superficially morphologically cryptic as adults but show marked differences in internal anatomy, as well as niche separation. There was no indication of ongoing hybridization between the species. We conclude that the observed mitochondrial-nuclear discordance is likely due to historical mitochondrial introgression following a range expansion.

KEYWORDS

Mitonuclear discordance; hybridization; dusky shark; Galapagos shark; population genomics; gene capture

1. INTRODUCTION

Carcharhinus galapagensis (Snodgrass and Heller, 1905) and *C. obscurus* (Lesueur 1818) are closely related shark species that have been recognized as distinct for more than 100 years. While the two species look superficially similar as adults, they have different numbers of pre-caudal vertebrae (Garrick, 1982), as well as marked and consistent

41 differences in onset of maturity (Natanson et al., 1995; Simpfendorfer et al., 2002;
42 Wetherbee et al., 1996). They also occupy different niches: *Carcharhinus galapagensis*
43 prefers shelf and reef habitats, most often occurring at offshore islands and seamounts,
44 while *C. obscurus* are primarily found along continental margins (Ebert et al., 2013).
45 Notwithstanding these differences, preliminary molecular analysis (see Naylor et al.,
46 2012b) based on the mitochondrial DNA NADH2 marker (which is widely used to
47 identify elasmobranch species) showed that *C. galapagensis* individuals sampled near
48 Hawaii were admixed with *C. obscurus* individuals that were sampled from throughout
49 the Pacific Ocean. While these species are known to co-occur at Norfolk Island,
50 sympatric occurrences are generally uncommon and *C. obscurus* has not been recorded
51 off Hawaii. This precludes taxonomic misidentification as a potential cause for the
52 observed mitochondrial admixture and suggests that either the NADH2 marker is
53 failing to resolve these species because there is mitochondrial introgression between
54 them (or perhaps because of incomplete lineage sorting or some form of selection at the
55 mitochondrial level), or that the two forms are not actually valid species, but are instead
56 ecologically driven morphological variants of the same species.

57 Accurately inferring taxonomic boundaries is fundamental to our understanding of
58 biodiversity and also has implications for the conservation management of exploited
59 species such as *C. galapagensis* and *C. obscurus*, which are targeted in artisanal,
60 recreational and commercial fisheries. Molecular data can be powerful for
61 discriminating species, but it is well-recognized that inferring relationships based on a

small number of molecular markers is problematic because of the potential for discordance between gene trees and species trees (Maddison, 1997). In this study we explored the pattern of genetic admixture between *C. galapagensis* and *C. obscurus* using a genomic-scale sampling of both nuclear and mitochondrial DNA markers, sampled from individuals throughout the respective ranges of both species, including their type localities and one location where they occur in sympatry. We used these data to explore the possibilities that these species hybridize or are conspecific.

2. MATERIALS AND METHODS

2.1 Sample collection and sequencing

Muscle samples were collected from 128 individuals nominally identified as *C. galapagensis* (N = 53) or *C. obscurus* (N = 75) based on their external morphology and sampling location (offshore islands/seamounts versus coastal, Fig. 1). Samples were collected from across the distribution range of both species, including sites in the Atlantic Ocean basin and the Indo-Pacific (Fig. 1, Table S1). The type locality for each species was sampled (Galapagos Islands for *C. galapagensis*, east coast U.S.A. for *C. obscurus*), as well as Norfolk Island where these species are known to occur in sympatry (Table S1). Tissue was stored in 95% ethanol prior to DNA extraction using the E.Z.N.A.[®] Tissue DNA Kit (Omega Bio-Tek, Inc., Georgia USA). Universal primers were used to amplify the mitochondrial NADH2 fragment by Polymerase Chain Reaction (PCR) for all samples prior to purification and Sanger sequencing, following Naylor et al. (2012a).

A subset of samples (N = 11 *C. galapagensis*, N = 15 *C. obscurus*) that were representative of the overall sampling effort (including individuals sampled in sympatry at Norfolk Island and the individual sampled off Senegal which possessed an unusual mitochondrial haplotype, see below) were subjected to targeted DNA hybridization capture for the purpose of collecting nuclear sequence data (Table S1). Following Li et al. (2013), Illumina sequencing libraries (500 bp) were prepared and amplified by PCR prior to two rounds of target gene capture (1077 independent autosomal regions). Enriched libraries were pooled and sequenced (paired-end 250 bp reads) on an Illumina MiSeq bench top sequencer (Illumina, Inc, San Diego, CA).

2.2 Data matrix assembly

NADH2 sequences were aligned using Geneious® Pro v.6.1.7 (Biomatters Ltd, Auckland, New Zealand), yielding a nucleotide alignment that was 1044 bp in length.

De novo reference sequence assembly, read mapping, variant calling, and filtering for the nuclear data is detailed in Maisano Delser et al. (2016). Adapters and low-quality reads were removed using cutadapt and FastQC, available in the wrapper script Trim Galore! v0.3.1 (Krueger, 2012). Trimmed sequence reads from several individuals were first used to build a haploid reference sequence for the 1077 target exons and their associated introns. Contigs were assembled de novo using ABySS v1.3.6 (Birol et al., 2009). On-target material was identified by comparing the assembled contigs to a set of single-copy core-orthologs that were found to be conserved across model vertebrates using HaMStR v13.2.3 (Ebersberger et al., 2009). For each exon and intron, the longest

contig across all sequenced individuals was retained, which resulted in an 836,916 bp reference sequence including exons plus 5' and 3' introns (see White et al., accepted for details). Trimmed reads from all individuals were then mapped to the reference sequence using the Burrows-Wheeler Aligner (BWA, Li and Durbin, 2009). Duplicate read marking was performed with Picard v.1.129 (<http://broadinstitute.github.io/picard/>), local realignment with the Genome Analysis ToolKit (GATK) v.3.3.0 (McKenna et al., 2010) and variant calling with SAMtools v.1.1 (Li, 2011). Raw variants were filtered for strand bias, minimum 6× depth of coverage, triallelic sites, sites with heterozygous calls in more than 80% of samples and missing data (Table S2) using custom Perl (v.5.18.2) and R (v.3.0.2) scripts (Table S3). This resulted in data for 2152 high-quality nuclear SNPs from 910 independent loci (Table S2).

2.3 Data analysis

Network v.5 (Fluxus Technology Ltd.) was used to generate a median-joining network of NADH2 haplotypes, with MP post-processing (substitution weights = 10 and epsilon = 0). In addition, maximum likelihood analysis (GTR+G) was conducted using RAxML v.7.3.0 (Stamatakis et al., 2005) with 1000 rapid bootstrap replicates (Stamatakis et al., 2008). *Carcharhinus longimanus* and *C. perezii* were used as outgroups.

Hierarchical Analysis of Molecular Variance (AMOVA; Excoffier et al., 1992; Weir and Cockerham, 1984; Weir, 1996) was used to quantify interspecific divergence relative to divergence among ocean basins, for both the mitochondrial NADH2 and the nuclear

SNP datasets. Four populations were defined: *Carcharhinus galapagensis* from the Atlantic, *C. galapagensis* from Indo-Pacific, *C. obscurus* from the Atlantic and *C. obscurus* from the Indo-Pacific. Two separate AMOVA were conducted in which the four populations were grouped either by species or by ocean basin. Total variance was partitioned into within-populations, among-populations-within-species (or ocean basin)-groups and among-species (or ocean basin)-groups variance components. Principal component analysis (PCA) was performed on the filtered nuclear SNP dataset with the function “prcomp” in the R environment (RCoreTeam, 2014). The number of clusters present in the data was further investigated using model-based clustering in STRUCTURE v.2.3.4 (Pritchard et al., 2000) with 100,000 iterations (30,000 burn-in). The number of assumed populations, K , varied from two to four with 10 independent runs for each. Replicate runs were merged and visualized using CLUMPAK (Kopelman et al., 2015).

3. RESULTS

Forty-one unique mitochondrial NADH2 haplotypes that are defined by 39 polymorphic sites were detected across the 128 sampled individuals (53 *C. galapagensis*, 14 haplotypes, 9 polymorphic sites; 75 *C. obscurus*, 29 haplotypes, 29 polymorphic sites). Haplotypic diversity was 0.848 ± 0.030 for *C. galapagensis* and 0.904 ± 0.024 for *C. obscurus* (0.893 ± 0.020 combined). Nucleotide diversity was 0.004 ± 0.002 for *C. galapagensis* and 0.005 ± 0.003 for *C. obscurus* (0.005 ± 0.003 combined). *Carcharhinus galapagensis* and *C. obscurus* were not distinguished by hierarchical

146 AMOVA of the mitochondrial data ($F_{CT} = -0.14$; -13.77 % of total variation), while there
147 was substantial divergence between ocean basins ($F_{CT} = 0.23$; 23.25 % of total variation;
148 Table S4).

149 Haplotypes were linked together by 14 or fewer mutations. The network consisted of
150 two clusters that were separated by six mutations (Fig. 2a). These clusters largely
151 corresponded to haplotypes sampled within the Atlantic Ocean versus the Indo-Pacific
152 (Fig. 2a), except for two individuals nominally identified as *C. obscurus* that were
153 sampled from Atlantic waters off Senegal and France that clustered in the otherwise
154 Indo-Pacific clade (Figs. 2a, S1). Haplotypes were generally closely related within the
155 Atlantic cluster, separated by three or fewer mutations. The Atlantic cluster was
156 dominated by one abundant haplotype that was sampled at all Atlantic locations, except
157 for off of Senegal and France (Fig. S1). The remaining haplotypes in the cluster were
158 singletons, or were associated with a small number of individuals from distinct locations
159 (Fig. S1). The Indo-Pacific cluster was more diverse (Fig. 2a), but also contained a
160 predominant haplotype that was sampled from most Indo-Pacific locations (Fig. S1).
161 Four other haplotypes were encountered across multiple locations. The remaining
162 haplotypes were singletons, or were unique to a single location (Fig. S1). Haplotypes
163 within the Indo-Pacific clade were separated by eight or fewer mutations (Fig. 2a).
164 The two most abundant haplotypes in the network were shared by *C. galapagensis* and
165 *C. obscurus*. No other haplotypes were shared among species (Fig. 2a); however, the
166 Indo-Pacific and Atlantic clades consisted of individuals from both species. While

haplotypes did not cluster according to species, ocean basins were readily distinguishable (Fig. 2a). The two clades corresponding to ocean basins were also resolved with high bootstrap support in the maximum likelihood analysis, while there was no statistical support for the separation of the two species (Fig. S2).

In contrast to the results based on the mitochondrial data, *C. galapagensis* and *C. obscurus* were readily distinguishable by hierarchical AMOVA of the 2152 nuclear SNP dataset ($F_{CT} = 0.27$; 26.74 % of total variation). A substantial proportion of the nuclear variation was also due to differences between ocean basins within species ($F_{SC} = 0.27$; 19.57 % of total variation; Table S4).

The first two components in the principle component analysis (PCA) explained ~ 50% of the variance and indicated four distinct clusters (Fig. 2b). The first principal component (PC1) separated *C. galapagensis* from *C. obscurus*, while the second principal component (PC2) showed geographical clustering by ocean basin (Fig. 2b). The four clusters persisted when considering the third (PC3) and fourth (PC4) principal components (Fig. S3). Although only accounting for approximately 10% of the total variation, the pattern in PC2 suggests greater similarity between species within ocean basins, than among ocean basins within species (Figs. 2b; S3)

The cluster analysis also supported separation of *C. galapagensis* and *C. obscurus*. The mean estimated log probability of the data was highest for $K = 2$ ($\ln \Pr(X|K) = -20877.90$), which corresponded to each of the species (Figs. 2c; S5). A

small amount of admixture was observed in *C. galapagensis* from the Indo-Pacific and *C. obscurus* from the Atlantic (Figs. 2c; S5).

The *C. galapagensis* and *C. obscurus* individuals that were sampled in sympatry at Norfolk Island, and the individual sampled off Senegal, were all assigned to their nominal species clusters in both the PCA and clustering analysis (Figs. S4, S5).

4. DISCUSSION

Genomic-scale surveys of DNA variation are becoming more common. It is anticipated that genomic-scale sampling should reduce uncertainty arising from coalescent variation among markers (Ballard and Whitlock, 2004; Maddison, 1997). The data deluge that has accompanied the genomics era has also generated a wider appreciation of the potential for conflicting patterns of differentiation between the mitochondrial and nuclear genomes that are not the consequence of incomplete lineage sorting (ILS), but rather caused by factors such as selective differences, demographic asymmetries or hybridization (Hudson and Turelli, 2003). A recent review revealed 126 such cases in animal systems spanning a taxonomic diversity that included invertebrates, amphibians, reptiles, birds, mammals and bony fishes (Toews and Brelsford, 2012).

Here, we describe a case of widespread discordance in signal obtained from the mitochondrial versus nuclear genomes in elasmobranchs. Mitochondrial NADH2 haplotype sharing is ubiquitous between *C. galapagensis* and *C. obscurus* (Figs. 2a, S2).

The two most frequent haplotypes were shared among 27 *C. galapagensis* and 26 *C. obscurus* sampled from across their distribution ranges. In fact, interspecific

208 mitochondrial divergence is lower than the divergence across ocean basins (Fig. 2a,
209 Table S4). One interpretation of this pattern could be that *C. galapagensis* is the oceanic
210 form of *C. obscurus*. However, the results of our analyses of more than 2000 nuclear SNP
211 loci are consistent with the existing taxonomy. *Carcharhinus galapagensis* is readily
212 distinguishable from *C. obscurus* according to the results of the AMOVA (Table S4), PCA
213 and cluster analysis (Figs. 2b, c) of the nuclear data. Intraspecific population structure
214 was also detected by the AMOVA and PCA, showing a clear separation between the
215 Atlantic Ocean and Indo-Pacific samples in both species (Fig. 2b, Table S4). The
216 individuals sampled at Norfolk Island, which represent the only location where the two
217 species are known to occur in sympatry, were all assigned to their nominal species and
218 ocean basin clusters, as was the individual sampled off Senegal which possesses a
219 mitochondrial haplotype that clustered with the Pacific Ocean clade (Figs. S4, S5).
220 There are several possible causes of the admixed mitochondrial signal that we observe,
221 including ILS, and adaptive or biased mitochondrial introgression. In this case, it does
222 not appear that the observed discordance is the consequence of ILS. Mitochondrial
223 lineages are expected to sort more rapidly than nuclear lineages, given the four-fold
224 smaller effective population size of the mitochondrial genome. Contrary to this
225 expectation, nuclear lineages in our dataset appear to be sorted among species while
226 mitochondrial lineages are not. Moreover, ILS typically results in unpredictable
227 geographic patterns of allele sharing (Toews and Brelsford, 2012). Rather, we observe

228 spatially widespread discordance together with a strong geographic pattern of sorting
229 by ocean basin.

230 The mitochondria from one species may be adaptively introgressed into another if
231 selection favors a particular mitochondrial variant (Funk and Omland, 2003). Given that
232 we observe two distinct clusters of mitochondrial variants, it seems unlikely that
233 adaptive introgression would be the sole cause of the mitochondrial admixture, as this
234 would imply selection pressure for alternative mitochondrial variants in each ocean
235 basin.

236 Particularities of the spatial dynamics and demography of species can also result in
237 biased mitochondrial introgression, even without the influence of selection. We prefer a
238 demographic scenario where biased introgression of mitochondrial DNA has occurred
239 in each ocean basin during periods of secondary contact following a range expansion
240 (Currat et al., 2008) to explain the pattern we report here. During a range expansion,
241 alleles from the resident species can become rapidly introgressed into the expanding
242 species during the period of demographic growth that follows colonization. When
243 intraspecific gene flow is ongoing between the newly colonized population and others
244 in the expanding species, the introgression will usually be biased toward markers with
245 lower rates of migration, in this case the mitochondrial genome. These markers are more
246 readily introgressed under such circumstances because they are less likely to be
247 swamped by intraspecific gene flow and therefore more likely to increase in frequency
248 due to genetic drift (Currat et al., 2008; Petit and Excoffier, 2009).

249 Although contemporary *C. galapagensis* and *C. obscurus* exhibit niche separation, we
250 hypothesize that these species came into secondary contact following speciation at some
251 point in the past, allowing their mitochondrial genomes to become biasedly introgressed
252 within each ocean basin. Moreover, male-biased dispersal accentuates biased
253 mitochondrial introgression (Petit and Excoffier, 2009). While male-biased dispersal has
254 not been demonstrated in either species specifically, it is common in sharks (Dudgeon et
255 al., 2012). Further supporting the idea that the admixed mitochondrial pattern is the
256 consequence of historical hybridization events is the fact that only the most frequent
257 NADH2 haplotypes, which are typically inferred to be ancestral, are shared between
258 species. This suggests that hybridization occurred, allowing the mitochondria of
259 *C. galapagensis* and *C. obscurus* to become admixed, after which the species became
260 isolated once more and resumed independent evolutionary trajectories.

261 We did not observe any evidence of ongoing hybridization between species in the
262 nuclear SNP data, though we do not exclude this as a possibility upon obtaining nuclear
263 data from a larger number of samples of both species, particularly from regions where
264 they are sympatric. *Carcharhinus galapagensis* from the Indo-Pacific and *C. obscurus* from
265 the Atlantic, however, did show a small amount of admixture in the cluster analysis
266 (Figs. 2c, S5). Consistent with this, the pattern observed for PC2 in the PCA suggests that
267 for some loci there is less interspecific variation within an ocean basin, than intraspecific
268 variation across ocean basins (Figs. 2b, S3). We used AMOVA to test every locus for a
269 signature of introgression that may be driving these patterns. Briefly, two separate

hierarchical AMOVA were performed, for each locus, grouping the four populations (*C. galapagensis* from the Atlantic, *C. galapagensis* from Indo-Pacific, *C. obscurus* from the Atlantic and *C. obscurus* from the Indo-Pacific) by species or by ocean basin. The basic premise of this test was that introgressed loci should show lower divergence among species groups ($F_{CT\text{species}}$) than among ocean basin groups ($F_{CT\text{oceans}}$; see supplementary material for details). Using this approach, 53 loci were identified as potentially introgressed, falling outside the range of the overall distribution of interspecific and among ocean basin divergences (Supplementary material, Fig. S6). Presumably these loci are driving the subtle signature of admixture that we observed in the PCA and cluster analysis, and possibly represent the nuclear remnants of the same historical introgression event that resulted in the observed widespread mitochondrial admixture. Distinguishing between alternative drivers of introgression is challenging. It is possible that a combination of selection and demography have played a role in generating the patterns that we observe. A comprehensive phylogeographic study of both species that includes extensive phenotypic and environmental data, aimed at elucidating the population and demographic histories of *C. galapagensis* and *C. obscurus*, will be required to distinguish the evolutionary processes that may have been responsible, in addition to the location and direction of historical hybridization events.

ACKNOWLEDGEMENTS

We are grateful to Mark Scott, Tom Trnski, Jimmy Martinez, Heather Marshall, Diego Bernal, Greg Skomal, Severine Hannam, Mauricio Hoyos, Eric Sanders, Colin

291 Simpfendorfer, Ken Graham, Jeremy Cliff, Don De Maria, Roger Bakey, Andy
292 Richardson and Nicola Weber, the Ascension Islands Government Conservation
293 Department, the Natal Sharks Board and the Auckland Museum for providing valuable
294 tissue samples and museum voucher specimens. Kerri A. Matthes-Rosana and Elisabeth
295 Rochel performed data collection for the mitochondrial NADH2 dataset. Robert
296 McGuinn provided GIS expertise and assistance with mapping distribution ranges. This
297 work was supported by a National Science Foundation grant (Jaws and Backbone:
298 Chondrichthyan Phylogeny and a Spine for the Vertebrate Tree of Life; DEB-01132229).
299 GenBank accession numbers for the NADH2 dataset are KX171949 – KX172076. A VCF
300 file of nuclear SNP variants is available via Mendeley Data.

301 **REFERENCES**

302 Ballard, J.W.O., Whitlock, M., 2004. The incomplete natural history of mitochondria. *Mol*
303 *Ecol* 13, 729–744.

304 Birol, I., Jackman, S., Nielsen, C., Qian, J., Varhol, R., Greg Stazyk, G., Morin, R., Zhao,
305 Y., Hirst, M., Schein, J., Horsman, D., Connors, J., Gascoyne, R., Marra, M., Jones, S.,
306 2009. De novo transcriptome assembly with ABySS. . *Bioinformatics* 25, 2872–2877.

307 Currat, M., Ruedi, M., Petit, R.J., Excoffier, L., 2008. The hidden side of invasions:
308 massive introgression by local genes. *Evolution* 62, 1908–1920.

309 Dudgeon, C.L., Blower, D.C., Broderick, D., Giles, J.L., Holmes, B.J., Kashiwagi, T.,
310 Krück, N.C., Morgan, J.A., Tillett, B.J., Ovenden, J.R., 2012. A review of the application

311 of molecular genetics for fisheries management and conservation of sharks and rays. J
312 Fish Biol 80, 1789–1843.

313 Ebersberger, I., Strauss, S., von Haeseler, A., 2009. HaMStR: profile hidden markov
314 model based search for orthologs in ESTs. BMC Evolutionary Biology 9:157, 157.

315 Ebert, D., Fowler, S., Compagno, L., 2013. Sharks of the World. Wild Nature Press,
316 Maurice, Plymouth.

317 Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred
318 from metric distances among DNA haplotypes: Application to human mitochondrial
319 DNA restriction data. Genetics 131, 479–491.

320 Funk, D.J., Omland, K.E., 2003. Species level paraphyly and polyphyly: frequency,
321 causes, and consequences, with insights from animal mitochondrial DNA. Annual
322 Review of Ecology, Evolution and Systematics 34, 397–423.

323 Garrick, J.A.F., 1982. Sharks of the genus *Carcharhinus* NOAA Technical Report NMFS
324 Rockville, p. 194.

325 Hudson, R., Turelli, M., 2003. Stochasticity overrules the “three-times rule”: genetic
326 drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA.
327 Evolution 57, 182–190.

328 Kopelman, N., Mayzel, J., Jakobsson, M., Rosenberg, N., Mayrose, I., 2015. CLUMPAK: a
329 program for identifying clustering modes and packaging population structure
330 inferences across K. Molecular Ecology Resources 5, 1179–1191.

331 Krueger, F., 2012. Trim Galore!

332 Li, C., Hofreiter, M., Straube, N., Corrigan, S., Naylor, G., 2013. Capturing protein-
333 coding genes across highly divergent species. *BioTechniques* 54, 321–326.

334 Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association
335 mapping and population genetical parameter estimation from sequencing data.
336 *Bioinformatics* 27, 2987–2993.

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

339 Maddison, W., 1997. Gene trees in species trees. *Syst Biol* 46, 523–536.

340 Maisano Delser, P., Corrigan, S., Hale, M., Li, C., Veuille, M., Planes, S., Naylor, G.,
341 Mona, S., 2016. Population genomics of *C. melanopterus* using target gene capture data:
342 demographic inferences and conservation perspectives. *Scientific Reports* 6:33753.

343 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A.,
344 Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome
345 Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA
346 sequencing data. *Genome Res* 20, 1297–1303.

347 Natanson, L.J., Casey, J.G., Kohler, N.E., 1995. Age and growth estimates for the dusky
348 shark, *Carcharhinus obscurus*, in the western North Atlantic Ocean. *Fishery Bulletin* 93,
349 116–126.

350 Naylor, G., Caira, J., Jensen, K., Rosana, K., Straube, N., Lakner, C., 2012a. Elasmobranch
351 phylogeny: A mitochondrial estimate based on 595 species. In: Carrier, J., Musick, J.,

352 Heithaus, M. (Eds.), Biology of sharks and their relatives 2nd edition. CRC Press, Boca
353 Raton, FL.

354 Naylor, G., Caira, J., Jensen, K., Rosana, K., White, W., Last, P., 2012b. A DNA sequence-
355 based approach to the identification of shark and ray species and its implications for
356 global elasmobranch diversity and parasitology. Bull Am Mus Nat Hist 367, 1–262.

357 Petit, R., Excoffier, L., 2009. Gene flow and species delimitation. Trends Ecol Evol 24,
358 386–393.

359 Pritchard, J., Stephens, M., Donnelly, P., 2000. Inference of population structure using
360 multilocus genotype data. Genetics 155, 945–959.

361 RCoreTeam, 2014. R: A Language and Environment for Statistical Computing. R
362 Foundation for Statistical Computing.

363 Simpfendorfer, C.A., McAuley, R., Chidlow, J., Unsworth, P., 2002. Validated age and
364 growth of the dusky shark, *Carcharhinus obscurus*, from Western Australian waters. Mar
365 Freshw Res 53, 567–573.

366 Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the
367 RAxML Web servers. Syst Biol 57, 758–771.

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

370 Toews, D., Brelsford, A., 2012. The biogeography of mitochondrial and
371 nucleardiscordance in animals. Mol Ecol 21, 3907–3930.

372 Weir, B., Cockerham, C., 1984. Estimating F-statistics for the analysis of population
373 structure. *Evolution* 38, 1358–1370.

374 Weir, B.S., 1996. *Genetic Data Analysis II: Methods for Discrete Population Genetic*
375 *Data*. Sinauer Associates, Incorporated, Sunderland, MA.

376 Wetherbee, B., Crow, G., Lowe, C., 1996. Biology of the Galapagos shark, *Carcharhinus*
377 *galapagensis*, in Hawaii. *Environ Biol Fishes* 45, 299–310.

378 White, W., Corrigan, S., Yang, L., Henderson, A., Bazinet, A., Naylor, G., accepted.
379 Phylogeny of the manta and devilrays (Chondrichthyes: Mobulidae), with an updated
380 taxonomic arrangement for the family. *Zool J Linn Soc*.

381

FIGURE CAPTIONS

Fig. 1 Distribution ranges of *Carcharhinus galapagensis* (orange) and *Carcharhinus obscurus* (blue). Sampling sites are indicated by black squares (*Carcharhinus galapagensis*) and black circles (*Carcharhinus obscurus*) that give the number of individuals sampled.

Fig. 2 (a) Median joining network of mitochondrial DNA NADH2 haplotypes (1044 bp). Superfluous links have been removed based on parsimony. The area of each haplotype circle is proportional to its frequency. Small solid black circles are intermediate states that were not observed. Numbers indicate mutational steps between haplotypes (one, if not shown). *Carcharhinus galapagensis* is in orange and *Carcharhinus obscurus* in blue. The Indo-Pacific is indicated by solid colors and the Atlantic Ocean by triangles. **(b)**

Principal components analysis based on 2152 nuclear SNPs showing PC1 and PC2.

Carcharhinus galapagensis is in orange and *Carcharhinus obscurus* in blue. The Indo-Pacific is indicated by squares and the Atlantic Ocean by triangles. **(c)** Plot of the estimated

membership coefficients for each individual in each of two clusters (K=2), as inferred using the program STRUCTURE (Pritchard et al. 2000) based on 2152 nuclear SNPs.

Each individual is represented by a vertical column. Individuals are grouped by sampling location and colored by species. *Carcharhinus galapagensis* is in orange and *Carcharhinus obscurus* in blue.

401

402

403

404

Carcharhinus obscurus

Carcharhinus galapagensis

Mitochondrial DNA NADH2 Haplotypes

2152 nuclear SNPs