Phylogenomics and species delimitation of mobulid rays reveals cryptic diversity and a new species of manta ray

Jane Hosegood1,2,11 *, Emily Humble2,3, Rob Ogden3,4, Mark de Bruyn1,5, Si Creer1, Guy Stevens2, Mohammed Abudaya6, Kim Bassos-Hull7, Ramon Bonfil8, Daniel Fernando2,9,10, Andrew D. Foote1, Helen Hipperson11, Rima W. Jabado12, Jennifer Kaden13, Muhammad Moazzam14, Lauren Peel2,15,16,17, Stephen Pollett2, Alessandro Ponzo18, Marloes Poortvliet19, Jehad Salah20, Helen Senn13, Joshua Stewart2,21, Sabine Wintner22,23 and Gary Carvalho1

1 Molecular Ecology and Fisheries Genetics Laboratory, Bangor University, Bangor, LL57 2UW, UK
2 The Manta Trust, Catemwood House, Norwood Lane, Dorset, DT2 0NT, UK
3 Royal (Dick) School of Veterinary Studies and the Roslin Institute, University of Edinburgh, Easter Bush Campus, EH25 9RG, UK
4 TRACE Wildlife Forensics Network, Edinburgh, EH12 6LE, UK
5 The University of Sydney, School of Life and Environmental Sciences, Sydney 2006, NSW, Australia
6 National Research Center, Gaza City-Palestine
7 The Center for Shark Research, Mote Marine Laboratory, 1600 Ken Thompson Parkway, Sarasota, FL 34236, USA
8 Océanos Vivientes A. C. Cerrada Monserrat 9, Col. La Candelaria, CDMX 04380 Mexico.
9 Department of Biology and Environmental Science, Linnaeus University, SE 39182 Kalmar, Sweden
10 Blue Resources Trust, Colombo, Sri Lanka
11 NERC Biomolecular Analysis Facility, Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK
12 Gulf Elasmo Project, P.O. Box 29588, Dubai, UAE
13 RZSS WildGenes Lab, Royal Zoological Society of Scotland, Edinburgh EH12 6TS, UK
14 WWF-Pakistan, 46-K, PECHS, Block 6, Karachi 75400, Pakistan
15 School of Biological Sciences, University of Western Australia, Crawley, WA 6009, Australia
16 The Australian Institute of Marine Science, Crawley, WA 6009, Australia
17 Save Our Seas Foundation – D’Arros Research Centre, CH-1201, Geneva, Switzerland
18 Large Marine Vertebrates Research Institute Philippines, Cagulada compund, Jagna, Bohol, Philippines
19 Tolheksbos 57, 2134 GH Hoofddorp, the Netherlands
20 Ministry of Agriculture Directorate General of Fisheries, Palestine
21 Scripps Institution of Oceanography, La Jolla, CA, United States
22 KwaZulu-Natal Sharks Board, Private Bag 2, Umhlanga Rocks 4320, South Africa
23 School of Life Sciences, University of KwaZulu-Natal, Durban 4000, South Africa

*Corresponding author: bsp420@bangor.ac.uk or jane@mantatrust.org
Abstract

Practical biodiversity conservation relies on delineation of meaningful units, particularly with respect to global conventions and regulatory frameworks. Species delimitation methods have been revolutionised with the advent of next-generation sequencing approaches, allowing diversity within species radiations to be assessed with genome-wide data. Manta and devil rays (Mobula spp.) are threatened globally primarily from targeted and bycatch fishing pressure, resulting in recent protective measures under several global conventions and frameworks. However, a collective lack of representative global samples, ongoing taxonomic ambiguity, and ineffectual traceability measures combine to constrain the development and implementation of a coherent and enforceable conservation strategy for these species. Here we generate genome-wide Single Nucleotide Polymorphism (SNP) data from a globally and taxonomically comprehensive set of mobulid tissue samples, producing the most extensive phylogeny for the Mobulidae to date. We assess patterns of monophyly and combine this with species delimitation based on the multispecies coalescent. We find robust evidence for an undescribed species of manta ray in the Gulf of Mexico, and for the resurrection of a recently synonymised species, Mobula eregodootenkee. Further resolution is achieved at the population level, where geographic population structure is identified in Mobula species. In addition, we estimate the optimal species tree for the group and identify substantial incomplete lineage sorting, where standing variation in extinct ancestral populations is hypothesised to drive taxonomic uncertainty. Our results provide genome-wide data to support a taxonomic review of the Mobulidae, and generate a robust taxonomic framework to support effective management, conservation and law enforcement strategies.

Introduction

The Anthropocene has been characterised by unprecedented human exploitation of natural resources, resulting in global threats to biodiversity and extinction events within many taxa (Dirzo et al., 2014; McGill et al., 2015). Effective measures for the conservation of biodiversity require an understanding and characterisation of diversity within and among species. The field of conservation genetics provides opportunities for quantifying diversity across space and time (Allendorf et al., 2010) and such approaches are increasingly powerful with the growing incorporation of genome-wide data. Species delimitation, the process by which populations of individuals are grouped into reproductively isolated and separately evolving units, is a fundamental application of genomic data to biodiversity conservation.
Accordingly, species delimitation has received increasing attention recently, with numerous methods now available (Carstens et al., 2013; Zhang et al., 2013; Grummer et al., 2014; Leache et al., 2014; Rannala 2015; Yang 2015). Traditional approaches typically relied upon morphological observation, often resulting in artificially broad delineations arising from difficulties in detecting and identifying cryptic species (Frankham et al., 2012). More recently, DNA sequencing has allowed genetic data to be utilised for species delimitation, although early approaches were limited to information from only a few genes or markers. These early approaches therefore left interpretation challenging, particularly in recently diverged groups with substantial incomplete lineage sorting (Maddison 1997; Maddison and Knowles, 2006). Genome-wide multi-locus approaches have increased the resolution of species delimitation studies and have been used to clarify contentious relationships and phylogenies (eg. Leache et al., 2014; Herrera and Shank, 2016), and disclose previously unknown diversity (eg. Pante et al. 2014). Species delimitation remains constrained by the lack of a single universal species concept (De Queiroz, 2007; Frankham et al., 2012). The delineation of monophyletic assemblages underpins the phylogenetic species concept, and the biological species concept where species occur in sympatry (Frankham et al., 2012). This has application in characterisation of both Conservation Units and Evolutionary Significant Units for the purposes of effective conservation (Funk et al., 2012).

Globally, biodiversity conservation is enacted through conventions and regulatory frameworks, including the Convention on the International Trade in Endangered Species of Wild Fauna and Flora (CITES), and the Convention on the Conservation of Migratory Species of Wild Animals (CMS). These conventions are implemented through national legislation acting at the species level (Vincent et al., 2014), and effective wildlife protection, management and law enforcement therefore depends on unambiguous species classification. Recent examples of proposed taxonomic revisions having far-reaching consequences for biodiversity conservation include giraffe (Fennessy et al., 2016; Bercovitch et al., 2017; Fennessey et al., 2017) and African elephant (Roca et al., 2001). In these cases, genetic research has led to possible reclassification and consequent changes to the legal status of these threatened megafauna.

In the marine realm, manta and devil rays (Mobula spp.), are circumglobally distributed megafauna of high conservation priority (Lawson et al., 2017) that also carry substantial economic value for tourism (O’Malley et al., 2013). Despite the economic benefits provided through the non-consumptive use of these species (family Mobulidae; collectively, mobulids), this vulnerable group is threatened primarily by intense targeted and bycatch fishing pressure, in part driven by demand for their gill plates, which are utilised in Asian medicines (Couturier et al., 2012; Croll et al., 2016; Lawson et al., 2017; O’Malley et al., 2017). Exploitation of mobulid rays for human consumption is considered unsustainable due to their life history traits; late maturation, low reproductive rates and long generation times, hindering...
their ability to recover from fishing impacts (Dulvy et al., 2014). To alleviate these threats, all species of mobulid ray have recently been listed on the CITES Appendix II to regulate international trade and protection and enforcement. Additionally, several species are regulated under national jurisdictions, with varying levels of protection and enforcement. Unfortunately, a collective lack of representative global samples, ongoing taxonomic ambiguity, and ineffectual traceability measures has constrained the development and implementation of a coherent and enforceable conservation strategy (Stewart, 2018a).

Recently, White et al. (2017) conducted an evaluation of genetic and morphological datasets for 11 previously recognised species of mobulid ray across two genera. Eight species were recognised, and the authors called for the genus Manta (consisting of two species; Manta alfredi and Manta birostris) to be subsumed into Mobula (devil rays); a recommendation that is yet to be reviewed by the International Commission on Zoological Nomenclature (ICZN) at the time of writing. For the purposes of this study, we use the common name ‘manta ray’ to refer to individuals of the species M. alfredi and M. birostris or species identified therein. Although multi-locus genetic datasets were used in the study by White et al. (2017), only a single sample was included per putative species, thereby preventing delineation of monophyletic species groups. Furthermore, the conclusion that M. rochebrunii is a junior synonym of M. hypostoma was based entirely on mitogenome data (White et al., 2017), which is considered unsuitable for species delimitation or phylogenetics when used in isolation (Petit and Excoffier, 2009; Herrera and Shank, 2016). Prior to this study, the most recent major taxonomic change for the Mobulidae came with the resurrection of species status for Manta alfredi, resulting in recognition of two species of manta ray (Marshall et al., 2009). Whilst the validity of this split has been confirmed with genetic data (Kashiwagi et al., 2012), there remains evidence of both historic (Kashiwagi et al., 2012) and modern (Walter et al., 2014) hybridisation between the two species. In addition, a third putative species of manta ray is hypothesised to occur in the Caribbean (Marshall et al., 2009; Hinojosa-Alvarez et al., 2016).

The Mobulidae is a group characterised by recent divergence times, estimated to have diverged from Rhinoptera only 30 million years ago (MYA), and having undergone relatively short bursts of speciation associated with periods of decreased ocean productivity (Poortvliet et al., 2015), of which the most recent known is only 0.5MYA (Kashiwagi et al., 2012). The age of these divergences implies that secondary contact and introgression between separately evolving species is likely to be widespread within the group, further encumbering efforts to define species boundaries.

Such ongoing uncertainties within the Mobulidae demonstrate a requirement for genomic approaches to enable robust species delimitation. Here, we generate double-digest Restriction-site Associated
DNA sequence (ddRAD) data (Peterson et al., 2012) from the largest and most comprehensive geographic sampling of mobulid species (Figure 1), inclusive of taxon replicates within sampling sites to: (1) delimit mobulid species, resulting in identification of cryptic diversity and an undescribed species of manta ray, (2) estimate the optimal species tree for the group, and (3) identify the extent of incomplete lineage sorting.

Figure 1: Map of mobulid sampling locations. Species are denoted by different colour points, scaled for sample size. Total numbers of samples for each species are given in the legend. See Supplementary Table 1 for further details. Note that we use species names that were assigned to samples at the time of collection, some of which are now invalid (White et al., 2017).

Results

Monophyly and clustering

Maximum Likelihood phylogenetic trees based on two genome-wide SNP data matrices (hereafter referred to as datasets p10 and p90, see Supplementary Table 2 for details) of varying size displayed highly congruent patterns (Figure 2 and Supplementary Figure 1). These trees represent the most comprehensive phylogenetic trees in terms of numbers of individuals and geographic coverage for mobulid rays published to date. Putative species fall into reciprocally monophyletic groups with high bootstrap support, and these species groups fall into well supported clades separated by long branch lengths. *Mobula japonica* and *Mobula mobular* form a single monophyletic group with 100% bootstrap support. In contrast, *Mobula kuhlii* and *Mobula eregoodootenkee* were resolved into two distinct
monophyletic groups, each with 100% bootstrap support. Furthermore, two distinct monophyletic groups are reported within *M. kuhlii*; each with 100% bootstrap support (based on dataset p10) corresponding to individuals sampled in the West (South Africa) and East (Sri Lanka eastwards) Indian Ocean. Finally, the manta rays can be resolved into distinct monophyletic groups corresponding to *M. alfredi* and *M. birostris*. Within *M. alfredi*, two well supported groups that correspond to Indian and Pacific Ocean populations are observed, whilst *M. birostris* is split into two groups; an Atlantic and a global group. One individual (sampled in Flower Garden Banks National Marine Sanctuary) was noted to switch between *M. birostris* clades depending on the data matrix used and was placed outside each main group with low bootstrap support (69% for dataset p10).

Principal Components Analyses (PCA) were carried out on each of the clades referred to above using dataset p10 (Figure 3; see Supplementary Table 3 for details of SNPs retained following division of data into clades). For the manta rays (Figure 3A & B), the first principal component (hereafter PC) separates *M. alfredi* from *M. birostris*, whilst the second PC distinguishes between *M. birostris*, and a possible third species of manta ray. The third PC provides clear distinction between *M. alfredi* from the Indian and Pacific Oceans (*F*~ST~ = 0.162). The screeplot shows a steep decline in the amount of variation shown by each axis (Supplementary Figure 2A-B). For *M. mobular* and *M. japanica* (Figure 3C & D), there is no clear separation between the two putative species, although the first PC does provide some evidence to suggest a clustering of individuals into Indo-Pacific and Atlantic (including Mediterranean individuals) groups (*F*~ST~ = 0.061). The screeplot for this clade shows a much shallower decline, and the amount of variation explained by each axis is much lower than for other clades (Supplementary Figure 2C-D). For the *M. thurstoni*, *M. kuhlii* and *M. eregoodootenkee* group (Figure 3E & F), these three species are very clearly differentiated on the first and second PCs, and this variation is reflected in the corresponding screeplot (Supplementary Figure 2E-F). The third PC reflects the geographic separation of *M. kuhlii* referred to above (*F*~ST~ = 0.319). For *M. hypostoma* and *M. munkiana* (Figure 3G & H), only the first PC was found to represent a large portion of the variation in the data (Supplementary Figure 2G-H), which corresponds to the separation of individuals into *M. hypostoma* and *M. munkiana*.

**COI gene phylogeny**

A Maximum Likelihood tree of mobulid species built using COI sequences is presented in Supplementary Figure 3. COI sequencing was unable to resolve the two manta ray species (*M. alfredi* and *M. birostris*), into monophyletic groups, and failed to resolve *M. kuhlii* and *M. eregoodootenkee*. Several species were resolved into reciprocally monophyletic groups with high bootstrap support (*M.
tarapacana, M. mobular, M. hypostoma and M. munkiana), but several multifurcating nodes within the tree indicate poor resolution achieved with this dataset.

Figure 2: Maximum Likelihood Phylogenetic Tree of mobulid species based on dataset p10 (left). Tips represent individuals, with colour indicating species, and shape the geographic origin of the sample as indicated in the legend. Bootstrap values are shown on the branches. Nodes with less than 50% bootstrap support are collapsed. Bayes Factor Delimitation (BFD*) models are also presented (right) where individuals are assigned to species as indicated by coloured bars. Models are ranked in order of performance from left to right. Marginal Likelihood Estimates (MLE) and Bayes Factors (2log_e BF) for runs with a gamma prior on lambda, relative to the null model are shown beneath each model. See Supplementary Table 4 for MLEs and 2log_e BF for runs with alternative prior combinations. Models that included individuals from the sister clade are not shown, as these consistently performed poorly. Note that we use species names that were assigned to samples at the time of collection, some of which are now invalid (White et al., 2017).
Figure 3: Principle Components 1-3 plotted for each mobulid clade. Putative species are represented with colours, and shape represents the geographical location of sampling as indicated in the legend.

A) manta rays, PC1 and 2, B) manta rays, PC1 and 3, C) *M. mobular* and *M. japanica*, PC1 and 2, D) *M. mobular* and *M. japanica*, PC1 and 3, E) *M. thurstoni*, *M. kuhlii* and *M. eregoodootenkee*, PC1 and 2, F) *M. thurstoni*, *M. kuhlii* and *M. eregoodootenkee*, PC1 and 3, G) *M. hypostoma* and *M. munkiana*, PC1 and 2, H) *M. hypostoma* and *M. munkiana*, PC1 and 3. Note that we use species names that were assigned to samples at the time of collection, some of which are now invalid (White *et al.*, 2017).
Species Delimitation

Species models (see Supplementary Table 4 for details) were tested following the Bayes Factor Delimitation with genomic data (BFD*) method of Leache et al. (2014), and Bayes Factors calculated relative to a null model of mobulid species as defined by White et al. (2017). Marginal Likelihood estimates did not differ considerably between chains with different priors on lambda (Supplementary Table 4). For the manta rays (Figure 2), we find decisive support for models that recognise the Gulf of Mexico and global M. birostris clades referred to above as two separate species (2log BF = -775.82), and that recognise geographically separated populations of M. alfredi as separate species (2log BF = -1063.58).

The M. mobular and M. japanica clade was best described by models that were more similar in their performance (Figure 2). The null model performed poorly in comparison to three models that split individuals based on geographical information (indeed, prior to White et al. (2017); M. mobular was considered to be restricted to the Mediterranean Sea, whilst M. japanica was considered circumglobal). The model that split individuals into these two previously recognised species performed best (2log BF = -119.58 relative to null model) but was only marginally better than a model that split individuals into Atlantic (including the Mediterranean) and Indo-Pacific groups (2log BF = -119.34 relative to null model).

Decisive support was found for the M. thurstoni, M. kuhlii and M. eregoodootenkee clade (Figure 2), in models that resurrect M. eregoodootenkee as a valid species, and that further split M. kuhlii based on geographical information (2log BF = -1007.04 and -1263.8 respectively).

Finally, within the M. hypostoma and M. munkiana clade, we find decisive support for the null model, that recognises M. hypostoma and M. munkiana as distinct species (Figure 2).

In all clades, models assessing support for interaction from higher up the tree, as well as models testing random assignment of individuals to species, perform comparatively poorly (Supplementary Table 4).

Relationships among the Mobulidae

Maximum Likelihood trees using the two species level data matrices containing varying amounts of missing data were highly congruent (Figure 4 and Supplementary Figure 4). Both data matrices support the findings of White et al. (2017); that manta rays are nested within Mobula, and sister to M. mobular (≥95% bootstrap support) and hereafter all species of manta ray are referred to with genus name Mobula. In addition, these trees strongly suggest that the undescribed third species of manta ray is
most closely related to *M. birostris* (100% bootstrap support). Finally, *M. tarapacana* is tentatively placed on the first lineage to diverge from the remaining Mobulidae (84% bootstrap support with dataset p10).

![Maximum Likelihood tree of mobulid species based on dataset p10. Bootstrap values are shown on the branches. Illustrations © Marc Dando. The drawing of *Mobula* sp. 1 is based on images of dozens of individuals off the Yucatan Peninsula, Gulf of Mexico.]

With respect to Bayesian tree estimation under the multispecies coalescent, the consensus tree topology and estimates of theta were relatively consistent across independent runs that included different individuals from each species (Supplementary Table 5). This suggests that there was no major effect of subsampling on topology of the species trees inferred with SNAPP. In trees inferred with SNAPP, *M. tarapacana* was consistently placed within a clade separate to the ingroup of *M. hypostoma* and *M. munkiana* (highest posterior density (HPD) = 1.0). Other nodes within the tree were generally poorly supported. This topological uncertainty is apparent when visualised as a cludogram of gene trees sampled from the posterior distribution (Figure 5 and Supplementary Figures 5-7). The number of alternative topologies inferred per subsampling and within the 95% HPD ranged from 9-25 (Supplementary Table 6). In all inferred topologies within the 95% HPD, the topology within the clades
separated by long branches, previously discussed, remains the same, and the main difference was the placement of *M. mobular* relative to the other clades.

Figure 5: SNP phylogeny of 30 individuals assigned to ten species based on dataset p90, individual subsample 1 (see Supplementary Table 5 for details). Tree cloud produced using DENSITREE of sampled trees (representing samples taken every 1000 MCMC steps from 5,000,000 iterations) from SNAPP analysis to visualise the range of alternative topologies.

TreeMix inferred an admixture graph with the same topology as that inferred with RAxML (see Supplementary Figure 8). This model was found to explain 99.86% of the variance in the data, indicating that species placement is unaffected by admixture, where species may be more closely related than the tree suggests, or where species may be forced closer together due to unmodeled migration (Pickrell and Pritchard, 2012). Furthermore, three-population tests were all positive
We therefore found no evidence of introgression between clades containing *M. alfredi*, *M. mobular* and *M. thurstoni*.

### Discussion

Our analyses of a globally and taxonomically comprehensive set of mobulid tissue samples produced the most extensive phylogeny for the Mobulidae to date. Genome-wide SNP data provided a high degree of resolution compared to analysis of a single gene. Combined with results from analyses based on the multispecies coalescent, our findings provide robust support for several changes to be made to mobulid taxonomy, including the recognition of a new species of manta ray, and have implications for management, conservation and law enforcement.

It is important to recognise speciation as a continuous process, where lineage splitting does not necessarily correspond to speciation events. When this is explicitly modelled, the multispecies coalescent has been shown to overestimate species numbers, recovering all structure both at the level of the species and the population (Sukumaran and Knowles, 2017). In contrast to previous studies of mobulid taxonomy, the global nature of our dataset allows for this conflict to be resolved, where in many cases, individuals from pairs of putative species are sampled within sites, thereby allowing this distinction to be made.

We find strong evidence supporting the existence of a third, undescribed species of manta ray in the Gulf of Mexico (hereafter referred to as ‘*Mobula* sp. 1’). Samples were collected at two sites within the Gulf of Mexico; offshore of the Yucatan Peninsula and Flower Garden Banks National Marine Sanctuary, and were provisionally identified as *M. birostris*. When these Gulf of Mexico samples were analysed alongside *M. birostris* samples collected elsewhere (Sri Lanka, Philippines and Mexico Pacific), individuals were found to fall within two distinct groups; one containing only individuals from the Gulf of Mexico sites, and the other containing additional individuals from the same Gulf of Mexico sites as well as *M. birostris* individuals sampled elsewhere. In addition, we find decisive support for two models which recognise these groups as distinct species through Bayes Factor Delimitation (BFD*; Figure 2). Given that samples from both groups were collected within Gulf of Mexico sites, *M. birostris* can be considered to occur in sympathy with *Mobula* sp. 1, constituting separately evolving lineages (De Queiroz, 2007). Monophyly of groups supports these as separate species under the phylogenetic species concept (Frankham *et al.*, 2012). Furthermore, sympathy of populations suggests reproductive isolation driven either by a factor other than geographical separation, or historical separation followed by modern secondary contact (as hypothesised by Hinojosa-Alvarez *et al.* (2016)), and these species
are therefore further supported under the Biological Species concept (Frankham et al., 2012). In addition, we report on a single individual which could be considered as genetically intermediate between the two groups (Figures 2 and 3), indicating that hybridisation may occur between the two species, as between M. alfredi and M. birostris (Walter et al., 2014).

Novel mtDNA haplotypes have previously been reported from manta rays off the Yucatan Peninsula, and a speciation event hypothesised (Hinojosa-Alvarez et al., 2016), in addition to previous morphological observations (Marshall et al., 2009). Our study is the first analyses of genome-wide data to suggest that there are two species of manta ray present in the Gulf of Mexico; a finding that is consistent with previous studies (Hinojosa-Alvarez et al., 2016; Stewart et al., 2018b). Monophylly of groups indicate that some M. birostris individuals using sites in the Gulf of Mexico are more closely related to M. birostris in Sri Lanka and the Philippines than to individuals of Mobula sp. 1 using those same Gulf of Mexico sites. It is likely that these species occur in a state of mosaic sympatry, as with M. alfredi and M. birostris elsewhere (Kashiwagi et al., 2011). For effective conservation it will be necessary to formally describe this new species and determine the extent of its range.

A recent taxonomic review concluded that M. eregoodootenkee is a junior synonym of M. kuhlii based on mitogenome and nuclear data for a single sample per putative species (White et al., 2017). In direct contrast, our phylogenetic analysis of genome-wide SNPs which included multiple individuals per species from multiple geographic locations, placed individuals of M. kuhlii and M. eregoodootenkee into discrete monophyletic clades with very high bootstrap support (Figure 2). This pattern was also mirrored in the results of our Principal Components Analysis (Figure 3). In addition, BFD* models that recognised M. eregoodootenkee as a distinct species from M. kuhlii are consistently favoured over the null model (Figure 2). Given that both species groups included samples that were collected within the same ~120km stretch of South African coastline, the divergence reported here between M. kuhlii and M. eregoodootenkee cannot be attributed to geographic population structure (Sukumaran and Knowles, 2017). There is evidence to suggest that periods of speciation within the Mobulidae correspond to episodes of global warming and associated changes in upwelling intensity and productivity, and it is hypothesized that this led to fragmentation and subsequent divergence with respect to feeding strategies (Poortvliet et al., 2015). Differences in morphology between M. kuhlii and M. eregoodootenkee (Notarbartolo Di Sciara, 1987; Notarbartolo di Sciara et al., 2017), and particularly the suggestion of differences in the length of the cephalic fins and gill plate morphology (Paig-Tran et al., 2013), that relate directly to the filter feeding strategy of mobulid rays, may lend support to this hypothesis. Notwithstanding, the present study provides the best available evidence regarding the species status of this group, and as such we resurrect Mobula eregoodootenkee as a distinct species.
In agreement with the conclusion of White et al. (2017), we find no evidence to support *M. japonica* as a distinct species to *M. mobular*. Individuals provisionally identified as *M. mobular* as it was formerly recognised (with a distribution that was restricted to the Mediterranean Sea), do not form a reciprocally monophyletic group to the exclusion of individuals belonging to *M. japonica* (a species previously considered to be circumglobally distributed with the exception of the Mediterranean Sea), and instead these individuals form a single clade, with high bootstrap support (Figure 2). Clustering analyses indicate a degree of population structure, with some modest differentiation between Indo-Pacific and Atlantic (including Mediterranean) groups ($F_{ST} = 0.06$). Results from BFD* are far less conclusive than those for other clades (Figure 2), and support for split models being driven by geographic segregation of populations cannot be ruled out (Sukumaran and Knowles, 2017). We therefore uphold *M. mobular* as a single species, with *M. japonica* considered a junior synonym of the same.

With respect to species delimitation of the final clade examined, we find strong evidence to support *M. hypostoma* and *M. munkiana* as distinct species (Figures 2 and 3). Whilst these species are geographically segregated in the Atlantic and Eastern Pacific Oceans respectively, the divergence is of a similar magnitude to that of other species groups within the Mobulidae (Figures 2 and 3, Supplementary Figure 2) and morphological differences between the two species are considered sufficient to recognise two species (Notarbartolo Di Sciara 1987; Stevens et al. 2018). As such we find no evidence to support any modification to the taxonomy of this clade.

Previous studies found morphological differences sufficient to consider *M. rochebrunei* (a pygmy devil ray species described off the coast of West Africa) a distinct species (Cadenat, 1960); summarised in Notarbartolo Di Sciara (1987). In this study, we were unable to generate molecular data representing *M. rochebrunei* (now considered to be a junior synonym of *M. hypostoma* (White et al., 2017)). However, the revision published by White et al. (2017) is based on low mitochondrial sequence divergence between single representative samples of the two putative species, and is consistent with sequence divergence estimates for other mobulid groups where further study has resolved separate species status: *M. alfredi* and *M. birostris* (Marshall et al., 2009; Kashiwagi et al., 2012; this study), and *M. kuhlii* and *M. eregoodootenkee* (this study). Therefore, given the high vulnerability to extinction which exists for any mobulid species with a restricted range in this region (Atta-Mills et al., 2004; Doumbouya 2009) efforts to resolve this taxonomic uncertainly should be given a high priority (see Stewart 2018a).

Through phylogenetic and clustering analyses, we identify substantial geographically-mediated population structure within *M. kuhlii* and *M. alfredi*. In both cases, individuals fall into monophyletic
groups corresponding to the East and West Indian Ocean \( (F_{ST} = 0.32) \), and Indian and Pacific Oceans \( (F_{ST} = 0.16) \), respectively, with high bootstrap support. This pattern is consistent in our clustering analysis, and BFD* supports models that recognise these populations as distinct species. Indeed, there are anecdotal suggestions of morphological differences occurring in *M. kuhlii* across the Indian Ocean (Stevens et al., 2018). However, given that we cannot rule out a geographic driver of these patterns, *M. kuhlii* and *M. alfredi* must currently be maintained as singular species. Further study is required to investigate this pattern, and to assess the population genetic structure of both species to support effective management.

The inference of relationships within the Mobulidae provided largely congruent results across Maximum Likelihood and Bayesian analyses, with an exception of the placement of *M. tarapacana*. Our ML analysis placed *M. tarapacana* on the oldest mobulid lineage, as result consistent with similar ML analysis based on nuclear data (White et al., 2017). Yet our Bayesian analyses consistently placed *M. tarapacana* as sister species to *M. hypostoma* and *M. munkiana*. Analyses employing mitochondrial data support *M. tarapacana* as a sister species to the manta rays and *M. mobular* (Poortvliet et al., 2015; White et al., 2017), an observation that we were unable to reproduce with our data. Discordant trees in phylogenomic studies may be attributed to a small number of genes or loci, either driven by positive selection resulting in convergent evolution, or by evolutionary processes such as incomplete lineage sorting or hybridisation (Shen et al., 2017). Coalescent-based approaches, such as the independent analysis of unlinked SNPs completed here, account for each gene trees history, and are therefore less likely to be influenced by single genes (Shen et al., 2017), lending support to the hypothesis that *M. tarapacana* is sister to *M. hypostoma* and *M. munkiana*.

Application of a multispecies coalescent-based approach to our data allowed visualisation of the uncertainty in species tree topology and incomplete lineage sorting. Our Maximum Likelihood phylogenetic analysis indicates that the previously recognised genus *Manta* is nested within *Mobula*, and provides further justification for the associated change in nomenclature implemented by White et al. (2017). However, concatenated approaches can be prone to converge to an incorrect phylogeny (Kubatko and Degnan, 2007), whilst ignoring heterozygous sites can effect estimates of divergence times (Lischer et al., 2014). Whilst our Bayesian multispecies coalescent analyses do not specifically refute the observation that *Manta* is nested within *Mobula*, we find substantial uncertainty in the placement of *M. mobular*. Trees within the 95% HPD that place *M. mobular* with the manta rays are present in approximately equal proportions to trees placing the species with the remaining devil rays (Supplementary Table 6), thereby producing trees where the two formerly recognised genera are reciprocally monophyletic. In groups that have undergone a rapid speciation process and had large ancestral effective population size, the effects of incomplete lineage sorting on species tree estimation
are particularly prominent (Flouri et al., 2018). The Mobulidae are known to have undergone recent rapid bursts of speciation (Poortvliet et al., 2015), and our estimates of theta (mutation-scaled effective population size), were larger on the deeper branches of the tree reflecting the large effective population size of the extinct shared ancestral species of the contentious extant taxa (Supplementary Figure 9). Thus, standing variation in ancestral populations of mobulid rays is likely to drive taxonomic uncertainty with respect to the validity of Manta as a genus. Since there is no evidence of admixture driving these patterns (Supplementary Table 7), this uncertainty can be attributed to incomplete lineage sorting. Given that recently separated populations or species will pass through stages of polyphyly and paraphyly before becoming reciprocally monophyletic in the absence of additional introgression (Avise 1990; Patton and Smith, 1994), it is reasonable to hypothesise we are observing this process here. Based on current information however, we support *Mobula alfredi* and *Mobula birostris* as being taxonomically valid (White et al., 2017).

Our proposed changes to the taxonomy of the mobulid rays will have profound implications for practical conservation of the Mobulidae on an international scale, as conventions designed to regulate and effect conservation measures rely on systematic review at the species level (Shafer et al., 2015). Furthermore, many of these administrations rely on experts to evaluate the literature and assess priorities for species conservation, for example, under the IUCN’s Red List framework. Of particular importance from this study is the distinction of *M. eregoodootenkee* from *M. kuhlii*, given that they share a similar geographic range across a region with intensive fishing pressures (Notarbartolo di Sciara et al., 2017). Although each species is still treated as a single stock across the Indo-Pacific due to limited data available on their population structure, inference from related species suggest that their low reproductive output likely results in population numbers that will not withstand heavy fishing pressure (Dulvy et al., 2014; Croll et al., 2016). As such, their conservation status would be considered quite critical, requiring very specific management measures. In contrast, species such as *M. mobular* will now likely face lower conservation concerns given that *M. japonica* is a junior synonym. However as with other mobulid species, further investigations into population structure are warranted in order to conduct clear stock assessments for fisheries management.

Similarly, for conservation conventions such as CITES and CMS, and fisheries management bodies, management plans are drafted and approved at a species level and can severely impact anthropogenic pressures on a species. It is therefore imperative that decisions on species status are based upon the best available evidence.
Conclusions

This study represents the most comprehensive phylogenomic study in terms of numbers of individuals and geographic coverage for mobulid rays published to date and makes use of genome-wide SNP data to evaluate the taxonomy of the group and relationships between species. We present genome-wide evidence to support ten species within the Mobulidae: *Mobula alfredi*, *Mobula birostris*, *Mobula mobular*, *Mobula thurstoni*, *Mobula kuhlii*, *Mobula eregoodootenkee*, *Mobula hypostoma*, *Mobula munkiana*, *Mobula tarapacana* and a currently undescribed species of manta ray (*Mobula* sp. 1) in the Gulf of Mexico. In addition, we advocate the recognition of *Mobula rochebrunei* for conservation purposes until more data is available. We emphatically urge policy-makers, particularly the large conventions (such as the CITES and CMS) and the relevant specialist group within the IUCN to evaluate these as separate units in their assessments and when implementing conservation policy.

Future work in this area will necessarily involve formal description of the third species of manta ray (*Mobula* sp. 1), shown here to be present in the Gulf of Mexico. In addition, population level studies on individual species will allow more informed management by delineating conservation units. In the case of the Mobulidae, a group known to be vulnerable to overexploitation, assessment of stock structure within fisheries will allow for effective management.

This significant increase in the resolution of species diversity within the global evolutionary radiation of the Mobulidae was achieved through an international collaboration of researchers, contributing to a global collection of representative samples, combining multiple genome-wide markers with a combinatorial approach to data analysis. As such, the study provides a framework for molecular genetic species delimitation which is relevant to other wide-ranging taxa of conservation concern and highlights the potential for applied research to support conservation, management and law enforcement.

Materials and Methods

Sample collection, DNA extraction and Sanger sequencing

Tissue samples were collected representing all described species of mobulid ray, including the recently invalid species’ *Mobula japonica*, *Mobula eregoodootenkee* and *Mobula rochebrunei*, currently considered to be junior synonyms of *Mobula mobular*, *Mobula kuhlii* and *Mobula hypostoma* respectively (White et al., 2017), and an outgroup species, *Rhinoptera bonasus*. Where possible, samples were collected from a broad geographical range, and with multiple samples per site. Samples
were identified to species level based on morphological characters described in Stevens et al. (2018).

Samples included in the analyses described below (those yielding high quality DNA), totalling 20 countries and 31 sites, are shown in Figure 1, and details given in Supplementary Table 1. We use the original species names that were assigned to samples at the time of collection, some of which are now considered invalid following White et al. (2017).

Genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer’s instructions and eluted in nuclease-free water. DNA yield was measured using a Qubit 3.0 Broad Range Assay, and quality assessed on a 1% agarose gel stained with SafeView. The single sample of *Mobula rochebrunei*, from the Musee de la Mer, Goree, Senegal, had been stored in formalin, yielded no detectable DNA, and was therefore not sequenced.

To investigate the utility of traditional markers for mobulid species delimitation, PCR amplification of an approximately 650bp portion of the COI gene was carried out using universal Fish primers (Ward *et al.*, 2005) or, where these primers failed to amplify, as was the case for *M. munkiana* and *M. hypostoma* samples, primers MunkF1 (GGGATAGTGGGTACTGGCCT) and MunkR1 (AGGCGACTACGTGGGAGATT) were designed in-house using Primer-BLAST (Ye *et al.*, 2012). PCR was carried out in 15µl reactions, consisting of: 5.6µl nuclease-free water, 7.5µl of ReddyMix PCR Master Mix (ThermoFisher), 0.45µl of each primer, and 1µl DNA. PCR cycling conditions consisted of: 95°C for 2 min, followed by 35 cycles of 94°C for 30s, 54°C for 30s and 72°C for 1 min, with a final extension of 72°C for 10 mins. Sanger sequencing was carried out by Macrogen Europe, and raw sequences edited using the software Chromas Lite, yielding 110 high quality sequences (see Supplementary Table 1). Data was imported into MEGA7 (Kumar *et al.*, 2016), aligned using ClustalW, and the alignment checked for stop codons. The HKY+G model was identified as most suitable for this dataset using the Find Best Model option in MEGA7, and a Maximum Likelihood tree built with 1000 bootstrap replicates.

ddRAD library preparation and sequencing
ddRAD libraries were prepared in-house using a modified version of the protocol published by Peterson *et al.* (2012), and fully described in Palaiokostas *et al.* (2015). For each sample, 21ng of genomic DNA was digested with the restriction enzymes *SbfI* and *SphI* (NEB). Unique P1 and P2 barcode combinations were ligated to the resulting fragments for individual identification before samples were pooled. DNA fragments between 400 and 700bp were size-selected using gel
electrophoresis and PCR amplified. Individual sample replicates within and among libraries were included to assess error rates following the method described by Mastretta-Yanes et al. 2015.

A pilot ddRAD library was sequenced on the Illumina MiSeq at the Institute of Aquaculture, University of Stirling. Subsequent ddRAD libraries were sequenced by Edinburgh Genomics, University of Edinburgh on Illumina HiSeq High Output v4, with the 2 x 125PE read module.

### Data quality control and filtering

Data quality was assessed with FastQC software (Andrews 2010) with particular interest in the per base sequence quality module for SNP calling and the overrepresented sequences module to check for adapter contamination. Stacks (version 1.46; (Catchen et al., 2011)) was used for demultiplexing, quality filtering and assembling raw read data. Data were demultiplexed using the process_radtags.pl module and due to an indication of adapter contamination, adapter sequences were filtered out at this stage, with two mismatches allowed in the adapter sequence. In addition, the score limit was raised to 20 (99% probability) within the process_radtags sliding window to remove low quality sequence reads. Reads with an uncalled base were also discarded at this stage.

To minimise the level of linkage in our SNP data, only forward reads were included in the next stages of analysis. To remove any short fragments that were not successfully filtered out at the size-selection stage of the wet-lab protocol, a custom bash script was used to remove any sequence reads that contained a cut site for the SphI enzyme. This amounted to 8.5% of reads across samples.

In order to assemble loci and call SNPs, the denovomap.pl program was executed in Stacks (Catchen et al., 2011). The three main parameters for assembly were set as those that generated the largest number of new polymorphic loci shared across 80% of individuals, following the method for parameter testing described by Paris et al. (2017). Four identical reads were required to build a stack (-m), stacks that differed by up to four nucleotides were merged into putative loci (-M) and putative loci across individuals that differed by up to five nucleotides were written to the catalog (-n). This resulted in an average coverage of 105x across loci and samples. Allele and SNP error rates, as defined by Mastretta-Yanes et al. (2015), were below 6% and 2.5% respectively.

To generate a SNP matrix at the individual level, the populations.pl program in Stacks (Catchen et al., 2011) was used to output a VCF file containing all discovered SNPs across every polymorphic locus that was shared across more than a specified minimum number of individuals (10 or 90). This generated two matrices of varying size and with varying levels of missing data (see Supplementary Table 2). In order to remove possible paralogous loci from these matrices, VCFtools (Danecek et al.,
2011) was used to generate information on the average coverage at each locus across individuals. Those loci that were sequenced at more than double the standard deviation of coverage were assumed likely to be paralogous loci and were excluded. In addition, loci that were sequenced at less than one-third the standard deviation of coverage were excluded to mitigate for the effects of allele dropout (Arnold et al., 2013; Gautier et al., 2013). Moreover, loci were assessed for excess heterozygosity due to mapping artefacts, where those loci that were identified as having a high probability of heterozygote excess in one or more species were excluded from the entire dataset. Finally, loci were assessed for excess heterozygosity due to mapping artefacts, where those loci that were identified as having a high probability of heterozygote excess in one or more species were excluded from the entire dataset. To output final quality controlled SNP matrices for downstream analysis, the remaining loci and SNPs were written to a whitelist, and passed back to the populations.pl program in Stacks (Catchen et al., 2011). The –write_random_snp option was enabled at this stage to output a single random SNP per locus, thereby minimising the risk of genetic linkage, since this is a fundamental assumption of some of our downstream analyses. This resulted in two final matrices, p10 and p90, with 7926 and 1762 SNPs and 47.1% and 14% missing data respectively (summarised in Supplementary Table 2).

At the species level, these same whitelists were passed to populations.pl along with a population map assigning individuals to species based on the best-supported species model. The resultant matrices (summarised in Supplementary Table 2) were used for the species level analyses described below.

Reduced numbers of SNPs reported are due to a population (or species in this case) having incompatible loci – those with more than two alleles – which becomes possible when grouping individuals together.

**Assessment of monophyly and clustering**

To infer relationships among mobulid individuals, Maximum Likelihood (ML) phylogenetic analysis was carried out on concatenated ddRAD loci using RAxML version 8.2.11 (Stamatakis 2014). Analyses were run for both datasets since missing data is known to influence aspects of phylogenetic inference such as branch length (Leaché et al., 2015). The GTRGAMMA model of rate heterogeneity was implemented following assessment of best fit models in jModelTest (Darriba et al., 2015). Support for clades was assessed with 1000 bootstrap replicates and *Rhinoptera bonasus* was used as the outgroup to root the tree.

Once clades had been delimited with RAxML, the data were split into four groups, corresponding to four highly supported clades that were separated by long branch lengths. These four groups
correspond to the manta rays (*M. alfredi* and *M. birostris*), *M. mobular* (including specimens identified as *M. japanica* prior to the taxonomic revision published by (White et al., 2017)), *M. thurstoni* and *M. kuhlii* (including specimens identified as *M. eregoodootenkee* prior to the taxonomic revision published by (White et al., 2017)) and *M. hypostoma* and *M. munkiana*. See Supplementary Table 3 for details of numbers of SNPs sampled within each clade.

To assess how individuals cluster together, Principal Components Analysis (PCA) was performed on dataset p10 using the Adegenet package in R (Jombart 2008). After assessment of up to ten axes, three axes were retained in all cases. The populations.pl program in Stacks (Catchen et al., 2011) was used to calculate pairwise *F_{ST}* values among inferred clusters.

### Bayes Factor Delimitation of species

Species delimitation was carried out using the Bayes Factor Delimitation method with genomic data (BFD*) (Leache et al., 2014), which allows for direct comparison of Marginal Likelihood Estimates (MLE) for alternative species delimitation models under the multispecies coalescent. This analysis was carried out using the modified version of SNAPP (Bryant et al., 2012), implemented as a plug-in to BEAST (version 2.4.8; (Bouckaert et al., 2014)). Path sampling was carried out with 10 steps, (1,000,000 MCMC iterations, 20% burnin), implementing the log-likelihood correction available in the program (Leache et al., 2014). Since marginal likelihood estimates are affected by improper prior distributions, a gamma distribution was implemented on the lambda (tree height) parameter. To ensure that the ranking order of models was not affected by the priors, a second round was carried out retaining the default 1/X distribution on lambda, implementing upper and lower bounds of 10,000 and 0.00001 respectively, so that the prior becomes proper. Bayes Factors (2\(\log_e BF\)) were calculated from the MLE from each model for comparison (Kass and Raftery, 1995; Leache et al., 2014), using the formula:

\[
2\log_e BF = 2*(\text{MLE}_{null} - \text{MLE}_{test})
\]

Where positive 2\(\log_e BF\) values indicate support for the null model, whilst negative BF values favour the tested model. 2\(\log_e BF\) values < 10 are considered decisive support (Leache et al., 2014).

Due to the high computational requirements of running SNAPP, this analysis was carried out on the smaller dataset, p90, and the data was split up into clade specific datasets, as described above. For each clade however, four random individuals from the sister clade were included, to assess support for interaction from higher up the tree. See Supplementary Table 4 for details of numbers of SNPs sampled within each clade.
Alternative species delimitation models for each clade were informed both by the literature and by our own phylogenetic and clustering analyses (see Supplementary Table 4 for details). In addition, a model that randomly assigns individuals to two or three species was included for each clade, to assess relative support for other models. In all clades, the null model was considered as those species defined by White et al. (2017), and all Bayes Factors were calculated relative to this null model.

Species tree inference

To estimate relationships among the Mobulidae, phylogenetic analyses of individuals belonging to each of the best supported species was carried out using both Maximum Likelihood and Bayesian methods. Maximum Likelihood phylogenetic analysis was carried out on concatenated ddRAD loci for both species-level datasets, as described above for the individual-level datasets.

To test the tree topology and evaluate uncertainty, for example, due to incomplete lineage sorting, species tree inference was also carried out in SNAPP (Bryant et al., 2012), which allows each SNP to have its own history under the multispecies coalescent whilst bypassing the need to sample each individual gene tree. Due to the computational constraints associated with running SNAPP on a dataset as large as ours, dataset p90 was used, and three individuals per species were randomly selected following (Foote and Morin, 2016), whilst maximising geographical coverage within species. This process was repeated a further three times, randomly sampling individuals with replacement, resulting in four subsampled alignments (individual-specific details of each subsample, as well as details of numbers of SNPs retained with each subsample are provided in Supplementary Table 5). These four independent runs were carried out with an MCMC chain of 5,000,000 iterations, sampling every 1000 and retaining default priors on lambda and theta. Similar runs with different prior combinations produced similar results. Convergence to stationary distributions were assessed by visual inspection after 20% burnin in TRACER (Rambaut et al., 2018). The distribution of trees was visualised after 20% burnin in DensiTree (version 2.2.6; Bouckaert 2010)). The maximum clade credibility tree was drawn using TreeAnnotator (version 2.4.7; Bouckaert et al., 2014)).

Multi species coalescent approaches, such as SNAPP used in this study, assume that any discordance of topologies among loci results from incomplete lineage sorting and do not consider introgression as a source of discordance. Therefore, to investigate the extent to which the variation in these data is best explained by a single bifurcating tree, TreeMix (Pickrell and Pritchard, 2012) was used to evaluate whether there is evidence for significant introgression events within the Mobulidae. TreeMix involves building a maximum likelihood tree of user defined groups and calculating how much of the variance
in the data this fixed tree model accounts for. TreeMix was run on dataset p10. Given patterns observed using SNAPP with respect to uncertainty in the placement of *M. mobular*, the three-population test (Reich *et al.*, 2009) was additionally used to test for ‘treeness’ between clades. Similar to TreeMix, the three-population test estimates the covariance of allele frequencies between populations, but is a simple and less parameterised model than TreeMix, and thus can be a more powerful tool for identifying introgression. In addition to *M. mobular*, *M. alfredi* and *M. thurstoni* were randomly chosen from their respective clades for this test.

Acknowledgements

We are very grateful to the Save Our Seas Foundation (SOSF) and The People’s Trust for Endangered Species (PTES) for providing generous support for this work. JH is supported by a NERC CASE studentship through the ENVISION DTP (CASE partner - Royal Zoological Society of Scotland) and has received additional grants from the Fisheries Society of the British Isles and the Genetics Society. Data analysis was supported by the UK Natural Environment Research Council (NERC) Biomolecular Analysis Facility at the University of Sheffield.

The authors are very grateful to the following people and organisations for their help and support sourcing and collecting tissue samples; Julia L.Y. Spaet, Alec Moore, Rachel Brittain, Akazul, Grace Phillips, Jon Slayer, Framadou Doumbouya, West Africa Musee de la mer a Dakar, Dan Bowling, Heather Pacey, the Barefoot Collection, Dr. Bernard Seret, Prof. Dr. D. A. Croll, Kelly Newton, Mr. Hamid Badar Osmany (Marine Fisheries Department), Silvia Hinojosa, Planeta Oceano, all LAMAVE staff and volunteers and field team Captains Dean Dougherty, Captain Peter Hull, Captain Greg Byrd, Krystan Wilkinson and Breanna DeGroot. We would also like to thank all the staff at Atlantis-The Palm Dubai for giving us access to specimens brought in by fishermen and for their valuable help with data collection and dissections.

Blue Resources Trust (BRT) would like to thank the Department of Wildlife Conservation and the Department of Fisheries and Aquatic Resources for all their support provided to the field work carried out in Sri Lanka. BRT also acknowledges the generous support provided by the Save Our Seas Foundation (SOSF) and the Marine Conservation and Action Fund (MCAF) that enabled field work in Sri Lanka.

We thank Disney Conservation Fund, Save Our Seas Foundation and Mote Scientific Foundation for supporting sample collection in Florida. Special thanks also to the Local Government Unit of Jagna, the Philippines Bureau of Fisheries and Aquatic Resources Region 7. The SOSF D’Arros Research Centre is
a main affiliate of the Seychelles Manta Ray Project, which is funded by the SOSF. Sample collection in the Seychelles was approved by, and conducted with the knowledge of, the Ministry of Environment, Energy, and Climate Change, Seychelles.

The National Commission for Fisheries and Aquaculture of Mexico (CONAPESCA) allowed RB the collection of samples in Mexico through research permit PPF/DGOPA-091/15; the National Commission for Natural Protected Areas (CONANP) of Mexico and authorities of the Biosphere Reserve of Whale Sharks, kindly gave permission for work in the reserve. The Save Our Seas Foundation and the Marine Conservation Action Fund provided funding for research in Mexico. The Perfect World Foundation generously funded RB for the replacement of a drone used to locate manta rays. The Mexican CITES authority, Secretary of Environment and Natural Resources (SEMARNAT) provided CITES export permit for tissue samples through permit MX 80544.

We also thank Dr John Taggart for his support with the ddRAD library preparation protocol, and for his help sequencing a pilot ddRAD library. Gustavo Colucci assisted with DNA extractions and COI amplifications. In addition, we thank Marc Dando for kindly agreeing for us to reproduce his mobulid illustrations.

AF was funded by the Welsh Government and Higher Education Funding Council for Wales through the Sêr Cymru National Research Network for Low Carbon, Energy and Environment, and from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 663830.

Author Contributions

JH, EH, GC, MdB, RO, SC and GS designed and conceived of the study and secured funding for consumables relating to laboratory work. EH, GS, DF, AP, MA, JS, SP, SW, RJ, MP, MM, KBH, RB, JS and LP were responsible for sourcing and collecting samples. JH, HS and JK carried out laboratory work. JH, EH, GC, MdB, RO, SC, HH, AF and HS contributed to analysis of genome-wide SNP data. Figures were designed by EH and JH and produced by EH. All authors contributed to writing and editing the manuscript.

References

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics* 11:697–709.
A critical analysis of genealogical concordance in species delimitation using life history theory.

Bennett, A., Mccauley, Croll DA, Dewar H, and Genotyping Loci De Cadenat J. 2014. Directly from biallelic genetic markers: Bypassing gene tree 

Bryant, R, Bouckaert, R, Heled, J, Kühnert, D, Vaughan, T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology, 10:1–6.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, Roychoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution 29:1917–1932.

Cadenat J. 1960. The Mobulidae of the West Coast of Africa. Notes d’Ichtyologie ouest-africaine 22A:1053-1084.

Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. Molecular Ecology 22:4369–4383.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH, De Koning DJ. 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3: Genes, genomes, genetics 1:171-182.

Couturier LE, Marshall AD, Jaine FR, Kashiwagi T, Pierce JP, Townsend AK, Weeks SJ, Bennett MB, Richardson AJ. 2012. Biology, ecology and conservation of the Mobulidae. Journal of Fish Biology, 80:1075–1119.

Croll DA, Dewar H, Dulvy NK, Fernandez D, Francis MP, Galván-Magaña F, Hall M, Heinrichs S, Marshall A, Mccauley D, et al. 2016. Vulnerabilities and fisheries impacts: the uncertain future of manta and 

devil rays. Aquatic Conservation: Marine and Freshwater Ecosystems, 26:562–575.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Darriba D, Taboada GL, Doallo R, Posada D. 2015. jModelTest 2: more models , new heuristics and high-performance computing. Nature Methods 9:6–9.

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B. 2014. Defaunation in the Anthropocene. Science 345:401–406.

Doumbouya F. 2009. Rapport sur l’actualisation des etudes sure les raies mantas en Guinee. Centre National des Sciences Halieutiques de Boussoura. Ministère de la Pêche et de l’Aquaculture. Republique de Guinee. Halieutiques Boussoura. Ministère la Pêche l’Aquaculture. Repub. Guinee

Dulvy NK, Pardo SA, Simpfendorfer CA, Carlson JK. 2014. Diagnosing the dangerous demography of 

manta rays using life history theory. PeerJ 2:e400.

Fennessy J, Bidon T, Reuss F, Kumar V, Elkan P, Nilsson MA, Vamberger M, Fritz U, Janke A. 2016.
Multi-locus Analyses Reveal Four Giraffe Species Instead of One. *Current Biology* 26:2543–2549.

Fennessy J, Winter S, Reuss F, Kumar V, Nilsson MA, Vambeperger M, Fritz U, Janke A. 2017. Response to "How many species of giraffe are there?". *Current Biology* 27:R137–R138.

Flouri T, Jiao X, Rannala B, Yang Z. 2018. Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. *Molecular Biology and Evolution* 35:2585-2593.

Foote AD, Morin PA. 2016. Genome-wide SNP data suggest complex ancestry of sympatric North Pacific killer whale ecotypes. *Heredity* 117:316–325.

Frankham R, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Mendelson JR, Porton U, Ralls K, Ryder OA. 2012. Implications of different species concepts for conserving biodiversity. *Biological Conservation* 15325–31.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. 2012. Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution* 27:489–496.

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

Grummer JA, Bryson RW, Reeder TW. 2014. Species delimitation using bayes factors: Simulations and application to the sceloporus scalaris species group (Squamata: Phrynosomatidae). *Systematic Biology* 63:119–133.

Herrera S, Shank TM. 2016. RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution* 100:70–79.

Hinojosa-Alvarez S, Walter RP, Diaz-Jaimes P, Galván-Magaña F, Paig-Tran EM. 2016. A potential third Manta Ray species near the Yucatán Peninsula? Evidence for a recently diverged and novel genetic Manta group from the Gulf of Mexico. *PeerJ* 4:e2586.

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

Kashiwagi T, Marshall AD, Bennett MB, Ovenden JR. 2012. The genetic signature of recent speciation in manta rays (Manta alfredi and M. birostris). *Molecular Phylogenetics and Evolution* 64:212–218.

Kashiwagi T, Marshall AD, Bennett MB, Ovenden JR. 2011. Habitat segregation and mosaic sympathy of the two species of manta ray in the Indian and Pacific Oceans: Manta alfredi and M. birostris. *Marine Biodiversity Records* 4.

Kass RE, Raftery AE. 1995. Bayes factors. *Journal of the american statistical association* 90:773-795.

Kubatko LS, Degnan JH. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* 56:17-24.

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

Lawson JM, Fordham S, O’Malley MP, Davidson LNK, Walls RHL, Heupel MR, Stevens G, Fernando D, Budziak A, Simpfendorfer CA, et al. 2017. Sympathy for the devil: a conservation strategy for devil and manta rays. *PeerJ* 5:e3027.

Leaché AD, Banbury BL, Felsenstein J, De Oca ANM, Stamatakis A. 2015. Short tree, long tree, right tree, wrong tree: New acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology*
64:1032–1047.

Leache AD, Fujita MK, Minin VN, Bouckaert RR. 2014. Species delimitation using genome-wide SNP data. Systematic Biology 63:534–542.

Lischer HEL, Excoffier L, Heckel G. 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: Implications for the evolutionary history of Microtus voles. Molecular Biology and Evolution 31:817–831.

Maddison W, Knowles L. 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55:21-30.

Maddison WP. 1997. Gene trees in species trees. Systematic Biology 46:523-526.

Marshall AD, Compagno LJV, Bennett MB. 2009. Redescription of the genus Manta with resurrection of Manta alfredi (Krefft, 1868) (Chondrichthyes; Myliobatoidei; Mobulidae). Zootaxa 2301:1–28.

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. Molecular Ecology Resources 15:28–41.

McGill BJ, Dornelas M, Gotelli NJ, Magurran AE. 2015. Fifteen forms of biodiversity trend in the anthropocene. Trends in Ecology and Evolution 30:104–113.

Notarbartolo di Sciara G, Fernando D, Adnet S, Cappetta H, Jabado RW. 2017. Devil rays (Chondrichthyes: Mobula) of the Arabian Seas, with a redescription of Mobula kuhlii (Valenciennes in Müller and Henle, 1841). Aquatic Conservation: Marine and Freshwater Ecosystems 27:197-218.

Notarbartolo Di Sciara G. 1987. A revisionary study of the genus Mobula Rafinesque, 1810 (Chondrichthyes: Mobulidae) with the description of a new species. Zoological Journal of the Linnean Society 91:1–91.

O’Malley MP, Townsend KA, Hilton P, Heinrichs S, Stewart JD. 2017. Characterization of the trade in manta and devil ray gill plates in China and South-east Asia through trader surveys. Aquatic Conservation: Marine and Freshwater Ecosystems 27:394-413.

O’Malley MP, Lee-Brooks K, Medd HB. 2013. The Global Economic Impact of Manta Ray Watching Tourism. PLoS ONE 8:e65051.

Paig-Tran EWM, Kleinteich T, Summers AP. 2013. The filter pads and filtration mechanisms of the devil rays: Variation at macro and microscopic scales. Journal of Morphology 274:1026–1043.

Palaiokostas C, Bekaert M, Khan MGQ, Taggart JB, Gharbi K, McAndrew BJ, Penman DJ. 2015. A novel sex-determining QTL in Nile tilapia (Oreochromis niloticus). BMC Genomics 16:1–10.

Pante E, Abdelkrim J, Viricel A, Gey D, France SC, Boisselier MC, Samadi S. 2014. Use of RAD sequencing for delimiting species. Heredity. 114:450–459.

Paris JR, Stevens JR, Catchen JM. 2017. Lost in parameter space: A road map for stacks. Methods in Ecology and Evolution 8:1360-1373.

Patton JL, Smith MF. 1994. Paraphyly, polyphyly, and the nature of species boundaries in pocket gophers (genus Thomomys). Systematic Biology 43:11-26.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135.

Petit RJ, Excoffier L. 2009. Gene flow and species delimitation. Trends in Ecology and Evolution 24:386–
De Queiroz K. 2007. Species concepts and species delimitation. Systematic Biology 56:879–886.

Rannala B. 2015. The art and science of species delimitation. Current Zoology 61:846–853.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population history. Nature 461:489.

Roca AL, Georgiadis N, Pecon-Slattery J, O’Brien SJ. 2001 Genetic evidence for two species of elephant in Africa. Science 293:1473–1477.

Shafer ABA, Wolf JBW, Alves PC, Bergström L, Bruford MW, Brännström I, Colling G, Dalén L, De Meester L, Ekblom R, et al. 2015. Genomics and the challenging translation into conservation practice. Trends in Ecology and Evolution 30:78–87.

Shen XX, Hittinger CT, Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nature Ecology and Evolution 1:1–10.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Stevens G, Fernando D, Dando M, Notarbartolo Di Scia GN. 2018. Guide to the Manta and Devil Rays of the World. Princeton University Press.

Stewart JD, Jaime FRA, Armstrong AJ, Armstrong AO, Bennett MB, Burgess KB, Couturier LIE, Croll DA, Cronin MR, Deakos MH, et al. 2018a. Research Priorities to Support Effective Manta and Devil Ray Conservation. Front Mar Sci 5:314.

Stewart JD, Nuttall M, Hickerson EL, Johnston MA. 2018b. Important juvenile manta ray habitat at Flower Garden Banks National Marine Sanctuary in the northwestern Gulf of Mexico. Marine Biology 165:111.

Sukumaran J, Knowles LL. 2017. Multispecies coalescent delimits structure, not species. Proceedings of the National Academy of Sciences 114:1607–1612.

Vincent ACJ, Sadovy de Mitcheson YJ, Fowler SL, Lieberman S. 2014. The role of CITES in the conservation of marine fishes subject to international trade. Fish and Fisheries 15:563–592.

Walter RP, Kessel ST, Alhasan N, Fisk AT, Heath DD, Chekchak T, Klaus R, Younis M, Hill G, Jones B, et al. 2014. First record of living Manta alfredi × Manta birostris hybrid. Marine Biodiversity 44:1–2.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia’s fish species. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 360:1847–1857.

White WT, Corrigan S, Yang L, Henderson AC, Bazinet AL, Swofford DL, Naylor GJP. 2017. Phylogeny of the manta and devilrays (Chondrichthyes: Mobulidae), with an updated taxonomic arrangement for the family. Zoological Journal of the Linnean Society, 182:50-75.
Yang Z. 2015. The BPP program for species tree estimation and species delimitation. *Current Zoology* 61:854–865.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics* 13:134.

Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29:2869–2876.