ORIGINAL ARTICLE

Genetic structure and signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*)

P Momigliano[1,2,3], R Harcourt[1], WD Robbins[4,5], V Jaiteh[6], GN Mahardika[7], A Sembiring[7] and A Stow[1]

With overfishing reducing the abundance of marine predators in multiple marine ecosystems, knowledge of genetic structure and local adaptation may provide valuable information to assist sustainable management. Despite recent technological advances, most studies on sharks have used small sets of neutral markers to describe their genetic structure. We used 5517 nuclear single-nucleotide polymorphisms (SNPs) and a mitochondrial DNA (mtDNA) gene to characterize patterns of genetic structure and detect signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*). Using samples from Australia, Indonesia and oceanic reefs in the Indian Ocean, we established that large oceanic distances represent barriers to gene flow, whereas genetic differentiation on continental shelves follows an isolation by distance model. In Australia and Indonesia differentiation at nuclear SNPs was weak, with coral reefs acting as stepping stones maintaining connectivity across large distances. Differentiation of mtDNA was stronger, and more pronounced in females, suggesting sex-biased dispersal. Four independent tests identified a set of loci putatively under selection, indicating that grey reef sharks in eastern Australia are likely under different selective pressures to those in western Australia and Indonesia. Genetic distances averaged across all loci were uncorrelated with genetic distances calculated from outlier loci, supporting the conclusion that different processes underpin genetic divergence in these two data sets. This pattern of heterogeneous genomic differentiation, suggestive of local adaptation, has implications for the conservation of grey reef sharks; furthermore, it highlights that marine species showing little genetic differentiation at neutral loci may exhibit patterns of cryptic genetic structure driven by local selection.
Heredity (2017) **119**, 142–153; doi:10.1038/hdy.2017.21; published online 19 April 2017

INTRODUCTION

Genetic tools have often been employed to delineate management units and establish conservation priorities in marine organisms; however, the use of genetic data to inform management in the marine environment presents unique and important challenges (Hellberg *et al.*, 2002). Barriers to dispersal are rarely absolute in the sea and marine populations may have very large effective population sizes (N_e), and hence even small migration rates can homogenize genetic variance. Populations, which are demographically independent, can therefore appear panmictic. Genetic differentiation at neutral regions of the genome is often low in marine species, but low genetic drift also means that potentially adaptive mutations are less likely to be stochastically lost (Allendorf *et al.*, 2010; Lamichhaney *et al.*, 2012; Savolainen *et al.*, 2013). Alleles under selection, whose frequencies diverge as a function of the migration rate (m) and the selection coefficient (s) but independently of N_e, can evolve at a much faster rate in marine populations (Allendorf *et al.*, 2010; Lamichhaney *et al.*, 2012; Kelley *et al.*, 2016). Recent work on model organisms such as the three-spined stickleback (*Gasterosteus aculatus*) (DeFaveri *et al.*, 2013; Guo *et al.*, 2015), and commercially fished species such as herring (Lamichhaney *et al.*, 2012) and cod (Bradbury *et al.*, 2013) suggests that genomic divergence of marine populations is often heterogeneous;

differentiation is low across most of the genome with the exception of genomic islands of divergence driven by local selection.

For decades, population genetic studies of non-model organisms were limited to relatively low numbers of neutral markers, yielding low statistical power and difficulty in detecting the low levels of genetic drift expected for many marine populations. The recent advent of reduced representation sequencing techniques, such as RADSeq (restriction site associated DNA sequencing; Baird *et al.*, 2008), ddRADseq (double digest restriction associated DNA sequencing; Peterson *et al.*, 2012) and DArTSeq (Sansaloni *et al.*, 2011), allows thousands of loci to be genotyped *de novo* in non-model organisms. Using large panels of markers greatly increases the statistical power to detect small genetic differences, even with limited numbers of individuals (Willing *et al.*, 2012), and allows the identification of genomic regions of exceptionally high differentiation that may be indicative of local adaptation (Jensen *et al.*, 2016). As a result, marine species previously thought to be panmictic over large geographic areas are starting to reveal complex patterns of cryptic genetic structure (Lamichhaney *et al.*, 2012; Bradbury *et al.*, 2013). These cryptic patterns of differentiation are crucial to the development of effective conservation strategies, as even low levels of neutral genetic structure may reflect demographic independence in marine populations

[1]Department of Biological Sciences, Macquarie University, Sydney, New South Wales, Australia; [2]Sydney Institute of Marine Science, Mosman, New South Wales, Australia; [3]Ecological Genetics Research Unit, Department of Biosciences, University of Helsinki, Helsinki, Finland; [4]College of Marine and Environmental Science, James Cook University, Townsville, Queensland, Australia; [5]Wildlife Marine, Perth, Western Australia, Australia; [6]Centre for Fish and Fisheries Research, Murdoch University, Murdoch, Western Australia, Australia and [7]The Indonesian Biodiversity Research Centre, Udayana University, Denpasar, Bali, Indonesia
Correspondence: Dr P Momigliano, Department of Biosciences, University of Helsinki, Biocenter 3, Viikinkaari 1, FIN-00014 Helsinki, Finland.
E-mail: paolo.momigliano@helsinki.fi

(Ovenden, 2013) and locally adapted populations may warrant higher conservation priorities (Allendorf et al., 2010).

Interestingly, population genetic studies on sharks, arguably some of the ocean's most ecologically important predators, have thus far been almost entirely limited to small sets of neutral genetic markers—for a notable exception, see Portnoy et al. (2015). Like other marine fishes, sharks often have large historical N_e, even though recent declines may be starting to erode genetic diversity in some species (Castro et al., 2007; Portnoy et al., 2009; Nance et al., 2011; Blower et al., 2012). A larger N_e provides a higher chance of mutations arising and, in concert with lower genetic drift, higher levels of standing genetic variation will be maintained, the raw material upon which selection acts (Allendorf et al., 2010; Savolainen et al., 2013). Furthermore, sharks often have wide distribution ranges and as a result may occupy habitats that differ markedly in biotic and abiotic factors. Accordingly, they may be under different selective pressures across their distribution. Knowledge of where localized adaptation has occurred will assist with conserving evolutionary potential, an increasingly important objective for conservation management (Allendorf et al., 2010).

Characterizing genetic structure and connectivity in sharks has revealed complex patterns, likely representing high interindividual variation in dispersal distances that can be further complicated by sex bias (Pardini et al., 2001; Mourier and Planes, 2013). Sex-biased dispersal in sharks could also favor local adaptation. In a recent study on bonnethead sharks (Sphyrna tiburo), Portnoy et al. (2015) found evidence of sex-biased dispersal and local selection and suggested that the dispersing sex can facilitate the movement of potentially adaptive alleles, whereas the philopatric sex could favor local allele sorting. For species with high habitat specificity, genetic connectivity is expected to be limited by expanses of unsuitable habitat—for a perspective on this topic, see Momigliano et al. (2015b).

Within coral reef ecosystems, reef sharks make up the majority of large predator biomass (Sandin et al., 2008; Friedlander et al., 2014; Mourier et al., 2016). The recent declines in reef shark numbers, recorded in multiple ecosystems across the globe (Robbins, 2006; Graham et al., 2010; Ward-Paige et al., 2010), is therefore of concern. Knowledge of biological and environmental factors shaping patterns of genetic structure in coral reef sharks is fundamental to evaluate the risks of anthropogenic change, and for the development of efficient management strategies (Dudgeon et al., 2012). The grey reef shark (Carcharhinus amblyrhynchos) is among the most abundant reef sharks in the Indo-Pacific, contributing up to 50% to the biomass of higher order predators on coral reefs (Friedlander et al., 2014). Grey reef sharks possess features that are likely to yield complex patterns of genetic structure. They have a large geographic distribution spanning most of the tropical Indo-Pacific (Last and Stevens, 2009), suggesting the potential for wide-range dispersal, yet they are strongly associated with coral reef habitats (Espinoza et al., 2014). Grey reef sharks may undertake movements of >100 km crossing deep oceanic waters (Heupel et al., 2010; Momigliano et al., 2015a), but also show reef fidelity for extended periods of time (Espinoza et al., 2015b; Mourier et al., 2016).

Movements of grey reef sharks are influenced by the spatial distribution of coral reefs. They show low levels of reef fidelity in systems where neighboring reefs are close (Heupel et al., 2010), suggesting that coral reefs separated by only a few km may be perceived as continuous habitat (Momigliano et al., 2015b). As the distance between neighboring reefs increases, grey reef sharks show higher residency (Espinoza et al., 2015a), and inhabitants of isolated oceanic reefs rarely venture far (Barnett et al., 2012). Espinoza et al. (2015a) observed that adult males have larger home ranges than females and juveniles, and speculated that male-mediated dispersal may confer an evolutionary advantage by extending genetic and demographic connectivity beyond individual reefs. Nonetheless, the extent to which these movement patterns, observed at reef systems of varying degrees of isolation, reflect patterns of dispersal and gene flow remains largely unknown. The wide distribution of grey reef sharks also means that these animals inhabit coral reef habitats that greatly differ in terms of geomorphology, environmental factors (such as temperature) and biodiversity, suggesting that across their distribution they may experience spatially diversifying selection.

A recent population genetics study carried out in the Australian Great Barrier Reef (GBR), where most reefs are located within a distance of <2 km from their closest neighbor (Almany et al., 2009), revealed no large-scale genetic structure at microsatellite loci across a latitudinal gradient of nearly 1200 km (Momigliano et al., 2015a). The authors found no genetic differentiation between different regions of the GBR using microsatellite and mitochondrial DNA (mtDNA) markers, and no evidence of genotypic spatial autocorrelation at the reef scale, suggesting regular migration between neighboring reefs. However, it remains unclear whether grey reef sharks can maintain genetic connectivity across large oceanic distances. Furthermore, the extent to which patterns of gene flow are reflected by patterns of adaptive variation remains unknown.

In this study we investigate the genetic structure of grey reef sharks at multiple spatial scales across a substantial portion of the species' range using a combination of genome-wide single-nucleotide polymorphisms (SNPs) and mtDNA data. Based on recent findings from telemetry and population genetics studies, we expect gene flow in grey reef sharks to be hindered by large expanses of oceanic waters. In the absence of large oceanic barriers, we predict that grey reef sharks will be able to maintain genetic connectivity via male dispersal through coral reef 'stepping stones'. We also investigate whether the level of fragmentation of coral reef systems is associated with genotypic spatial autocorrelation. Finally, we scan for genomic signatures of local selection to identify locations that might contribute to the species' evolutionary potential.

MATERIALS AND METHODS
Sampling
A total of 180 grey reef shark DNA samples were obtained from 9 locations, spanning nearly 80° of longitude and 20° of latitude in the Indian and Pacific Oceans (Figure 1). Sampling in western Australian sites (Ningaloo Reef and the Rowley Shoals) was conducted in 2013 and 2014 as described by Momigliano et al. (2015a) under a permit from the Western Australia Department of Environment and Conservation (permit number: CE003632). Samples from Scott Reef were donated by the Australian Institute of Marine Science. Samples from the Chagos Archipelago were donated by Stanford University and the Bertarelli Foundation. Samples from Misool were collected in 2012 under an Indonesian research permit (RISTEK, permit 035/SIP/FRP/SM/I/2012). Genetic analyses of samples from Indonesia were undertaken under RISTEK permit 03B/TKPIPA/FRP/SM/III/2014. Individuals from the eastern Australian coast represent a subsample from Momigliano et al. (2015a).

mtDNA sequencing and analyses
We extracted DNA from a total of 180 samples (Figure 1) and amplified 813 bp of the NADH dehydrogenase subunit 4 (ND4) gene as per Momigliano et al. (2015a), using the primer set ND4 (Arèvalo et al., 1994) and HI12293-Leu (Inoue et al., 2003). DNA fragments were sequenced with the forward primer using a commercial service (Macrogen Inc., Seoul, Korea) and aligned by eye in BioEdit v. 7.1 (Hall, 1999). The sequences obtained in this study were analyzed together with the ND4 sequences from the North GBR ($N=32$), South GBR ($N=27$) and the Coral Sea ($N=8$) recently published by Momigliano et al. (2015a). The final analyses included a total of 247 ND4 sequences. Diversity
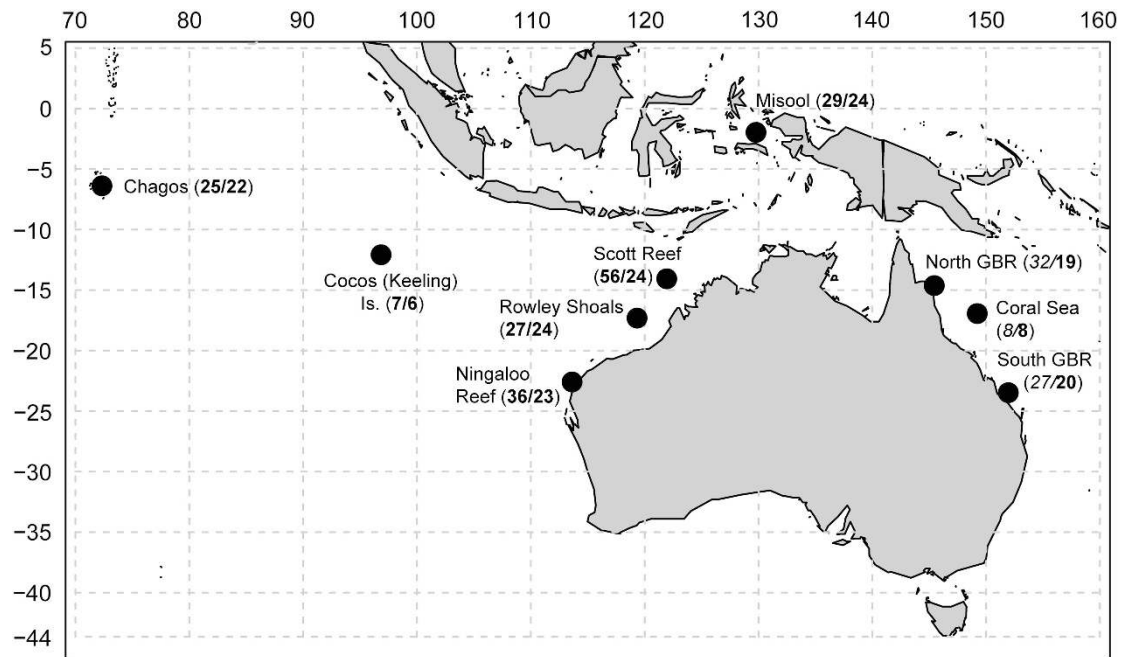
Figure 1 Map showing the sampling locations. In brackets are the numbers of individuals used for the mtDNA analyses and the SNP analyses, respectively. Numbers in italics represent samples from which data were retrieved from Momigliano et al. (2015a). Numbers in bold represent samples for which genetic data were generated in this study.

indices (number of haplotypes N_H and haplotype diversity h) for mtDNA sequences were calculated in the software DnaSP (Librado and Rozas, 2009). Measures of pairwise genetic differentiation (Φ_{ST}, an analog of F_{ST} for DNA sequence data) were estimated in the software Arlequin v. 3.51 (Excoffier and Lischer, 2010). A minimum spanning network of ND4 haplotypes was constructed in the software Popart v. 1.7 (Leigh and Bryant, 2015), using the same parsimony inference method implemented in the software TCS (Clement et al., 2000). Furthermore, we tested for past population expansions using Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) tests in Arlequin v. 3.51.

To test for sex-biased dispersal, we estimated pairwise Φ_{ST} for females and males separately. Although mtDNA is transmitted only maternally, there is a solid rationale for which mtDNA estimates of genetic differentiation in different sexes can be used to test for sex-biased dispersal. Males cannot transmit mtDNA, and hence the haplotype from an immigrant male can only be sampled while it is alive, whereas females have a far greater potential to homogenize mtDNA via migration and successful local reproduction (Lukoschek et al., 2008). As O'Corry-Crowe et al. (1997) and Lukoschek et al. (2008) noted, higher genetic differentiation at mtDNA markers in females is therefore evidence of male-biased dispersal and/or female reproductive philopatry. We did not have sex information for all our samples, and in most locations we did not have enough samples from each sex to accurately estimate Φ_{ST}. We therefore compared three locations for which we had at least 10 males and 10 females: the Rowley Shoals (10 males and 14 females), the North GBR (11 males and 18 females) and the South GBR (10 males and 15 females). Hence, the results regarding sex-biased dispersal should be interpreted with some caution.

SNP discovery and filtering

SNP discovery and genotyping was performed at Diversity Arrays Technology Pty. Ltd (Canberra, Australia), using the standard DArTSeq protocol. DArTSeq genotyping is a SNP genotyping-by-sequencing approach that combines Diversity Arrays (DArT) markers (Jaccoud et al., 2001) and next-generation sequencing on Illumina platforms (Sansaloni et al., 2011) to genotype thousands of SNPs homogenously spaced across the genome. The original DArT method is described in Jaccoud et al. (2001) and its combination with next-generation sequencing for SNPs genotyping is described by Sansaloni et al.

(2011). A detailed description of laboratory protocols and SNP calling and filtering procedures is provided in the Supplementary Materials.

Detection of loci putatively under selection

We utilized outlier tests to identify loci for which genetic differentiation (F_{ST}) is higher than expected by genetic drift alone. As outlier tests only perform well when neutral genetic differentiation is low (Whitlock and Lotterhos, 2015), we only included in these analyses individuals from Australian and Indonesian sampling locations that, as previous analyses revealed, show little genetic differentiation at nuclear loci (see Results). There are various methods that can be used to identify loci under selection, and each has different drawbacks, including limited sensitivities (high rate of type 2 errors) and susceptibility to false discoveries (type 1 errors) (Lotterhos and Whitlock, 2014). Methods that are commonly used include the Bayesian approach implemented in BAYES-CAN (Foll and Gaggiotti, 2008), the hierarchical model, coalescent simulation approach implemented in Arlequin (Excoffier et al., 2009) as well as Bonhomme et al. (2010) extensions of the Lewontin–Krakauer test that accounts for population co-ancestry (FLK). Each of these methods performs well under certain conditions, but both Arlequin and BAYESCAN can identify a large number of false positives when populations are spatially auto-correlated (Lotterhos and Whitlock, 2014). The FLK approach is much less susceptible to false positives under scenarios of evolutionary non-independence among populations (Lotterhos and Whitlock, 2014). A new method named Out-FLANK, which estimates a null distribution of F_{ST} for loci unlikely to be under strong positive selection based on the core distribution of a large set of loci, has also proved to be very robust under a number of demographic history scenarios including isolation by distance (IBD), although it is only suited for the identification of loci under strong spatially diversifying selection (Whitlock and Lotterhos, 2015). In this study we combined all the methods outlined above to identify loci putatively under selection. For the purpose of this study, loci were considered to be outliers if identified jointly by at least three outlier tests as putatively under selection. Furthermore, we tested whether divergence at neutral loci and loci putatively under selection were correlated, and could therefore be the result of the same processes. The outlier tests are described in detail in the Supplementary Materials.

Analysis of genetic diversity and structure

Analyses of neutral genetic variation were performed using a data set of 4798 loci identified by OutFlANK as representing the core distribution of F_{ST} (that is, trimming the upper and lower 5%), and hence unlikely to be affected by strong balancing and diversifying selection (Whitlock and Lotterhos, 2015). All the analyses were also performed on the entire SNP data set (results not reported) that yielded nearly identical results. Analyses of outlier loci were carried out on a set of 8 loci that were identified as putatively under natural selection by at least three outlier tests.

We calculated the following diversity indices across loci for each sampling location using the software GenAlEX v. 6.5 (Peakall and Smouse, 2012): expected heterozygosity (H_E), observed heterozygosity (H_O) and the fixation index $F = 1 - (H_O/H_E)$, along with their standard errors. Pairwise Weir and Cockerham F_{ST} and their 95% confidence intervals were estimated in the R package *diveRsity* (Keenan *et al.*, 2013) using 100 pseudo-replicate data sets created by bootstrapping individuals within each location. Using confidence intervals determined by bootstrapping individuals was deemed a good strategy to determine the significance of F_{ST} estimates given that two of our locations had very low sample sizes (Cocos (Keeling) Islands, $N = 6$, and Coral Sea, $N = 8$).

We investigated genetic structure at neutral loci using the software fastSTRUCTURE (Raj *et al.*, 2014). fastSTRUCTURE implements a fast algorithm for approximate inference of the simple admixture model from STRUCTURE (Pritchard *et al.*, 2000), and given its computational speed it is particularly well suited for large genomic data sets. First, we ran the algorithm with the simple prior, as suggested by the authors of the program, at multiple numbers of K (1–10). We determined the likely range of K using the function *chooseK.py*, which reports the model components that have a cumulative ancestry contribution of at least 99% ($K_øc$), and the value of K for which the log-marginal likelihood lower bound (LLBO) of the data (Raj *et al.*, 2014) is maximized ($K^*ε$). To further investigate K we then ran the fastSTRUCTURE algorithm for K values of 1–6 with both the simple prior and the logistic prior (the latter being more effective in detecting subtle population structure), using fivefold cross-validation. Theoretically, the most likely model is the one that minimizes prediction error. For K values greater than optimal prediction error may still decrease, although it will usually remain within one standard error of estimates of prediction error (Raj *et al.*, 2014). Therefore, the value of K after which estimates of prediction error remain within one standard error is usually considered the most parsimonious model (Raj *et al.*, 2014).

We further investigated patterns of genetic structure at neutral loci by carrying out Discriminant Analysis of Principal Components (DAPC) using sampling location as the grouping factor (Jombart *et al.*, 2010). The analysis was carried out on the whole data set as well as only including the data from Australia and Indonesia. First, we carried out a principal component analysis, and used the principal components (PCs) thus produced as synthetic variables for a discriminant analysis, as outlined in Jombart *et al.* (2010). This first step is necessary as two of the main assumptions of discriminant analysis are that variables are uncorrelated, and that the number of variables is less than the number of observations. Allele frequencies are inevitably correlated, and in large SNPs data sets the number of variables (that is, number of alleles) can be much greater than the number of sampled individuals. Using PCs obtained from the principal component analysis reduces the number of dimensions and creates a set of orthogonal variables that explain exactly the same variation as the original variables. The number of PCs to be retained in the DAPC analyses was determined using cross-validation to avoid overfitting: 80% of the data were used as a training set and we retained the number of PCs for which the obtained mean square error was lowest (number of PCs = 40 for both analyses).

Isolation by distance

We investigated the relationship between genetic distance at both neutral SNP loci (F_{ST}) and mtDNA (Φ_{ST}) and geographic distance using Mantel tests (using 30 000 randomizations) and major axis regression. As genetic differentiation across large oceanic expanses did not show a linear relationship with geographic distance, we only investigated IBD across locations in Australia and Indonesia (that is, we excluded samples from the Cocos (Keeling) Islands and Chagos archipelago). First, geographic distance between locations was estimated as the least-cost path across the sea. We then plotted geographic distance vs genetic distance to ensure the relationship was linear. The significance of the correlation between genetic distance and geographic distance was tested using the Mantel test, thereby taking into account that sampling locations, rather than pairs of sampling locations, are the units of replication. A linear model was then fitted to the data using major axis regression, a more appropriate method than an ordinary linear regression when both variables (genetic and geographic distance) are sampled with error.

Fine-scale spatial genetic structure

Genetic structuring of haplotype and allele frequencies is a function of the migration rate (m) and the effective population (N_e) (Kalinowski, 2002). Therefore, in large populations a small migration rate can effectively slow down or prevent any detectable level of genetic drift. This means that estimates of genetic differentiation based on allele and haplotype frequencies are often a poor proxy of ecological connectivity at fine spatial and temporal scales. In contrast, estimates of genotypic spatial autocorrelation provide a powerful tool to investigate fine-scale spatial structure. Patterns of spatial autocorrelation can appear even in face of migration rates exceeding 10% and are expected to reach quasi-stationarity in as little as 50 generations (Epperson, 2005; Momigliano *et al.*, 2015a).

We tested for genotypic spatial autocorrelation across four distance classes (within reef, within 500 km, within 1000 km and within 1500 km) in two geographic regions where we sampled multiple reefs: eastern Australia and western Australia. The spatial structure of these reef systems is broadly different. Eastern Australian coral reefs are relatively close with an average distance between neighboring reefs of < 2 km (Almany *et al.*, 2009), whereas coral reefs in western Australia are often separated by distances of 100s of km. We estimated pairwise genetic distances between individuals using the codom-genotypic measure of genetic distance implemented in GenAlEX v. 6.5 (Peakall and Smouse, 2012). The spatial autocorrelation coefficient (r) and its confidence intervals were calculated using 999 bootstraps in the software GenAlEX v. 6.5 (Peakall and Smouse, 2012) along with the 95% confidence intervals of the null model of no spatial autocorrelation.

RESULTS

mtDNA analyses

The final alignment included 247 individuals for which an 813 bp fragment of ND4 was obtained with no missing data. There were a total of 26 polymorphic sites within the alignment, of which 19 were parsimony informative. Twenty-six distinct haplotypes were present (Figure 2), and haplotype diversity across all locations was high ($h = 0.78$). Number of haplotypes and h within each sampling location ranged between 2 and 7 and 0.28 and 0.80, respectively (Table 1).
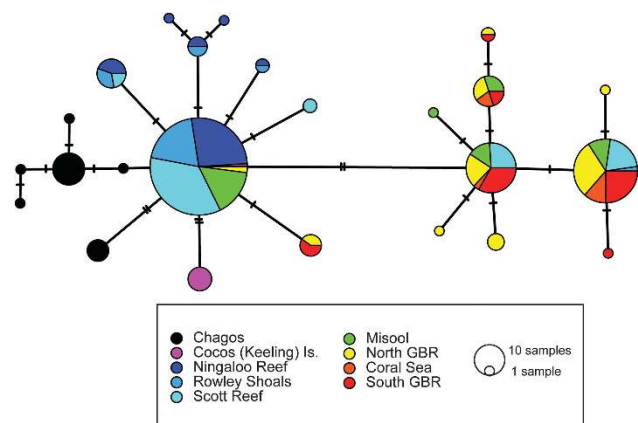

Figure 2 Minimum spanning network obtained from 247 individual partial ND4 sequences comprising 26 distinct haplotypes.

Table 1 Diversity indices for all sampling locations

Location	N (ND4)	N_H	h	N (SNP)	H_O	H_E	F
Chagos	25	7	0.75	22	0.139 ± 0.003	0.133 ± 0.003	− 0.05 ± 0.003
Cocos (Keeling)	7	2	0.28	6	0.312 ± 0.004	0.293 ± 0.003	− 0.14 ± 0.005
Ningaloo Reef	36	6	0.43	23	0.295 ± 0.003	0.300 ± 0.002	− 0.004 ± 0.003
Rowley Shoals	27	5	0.45	24	0.293 ± 0.003	0.298 ± 0.002	− 0.004 ± 0.003
Scott Reef	56	5	0.55	24	0.295 ± 0.002	0.300 ± 0.002	− 0.005 ± 0.003
Misool	29	5	0.66	24	0.301 ± 0.003	0.302 ± 0.002	− 0.017 ± 0.003
North GBR	32	9	0.80	19	0.293 ± 0.003	0.298 ± 0.002	− 0.008 ± 0.003
Coral Sea	8	3	0.61	8	0.294 ± 0.003	0.298 ± 0.003	− 0.053 ± 0.004
South GBR	27	6	0.73	20	0.288 ± 0.003	0.292 ± 0.002	− 0.009 ± 0.003

Abbreviations: F, fixation index $1 - (H_O/H_E)$; h, haplotype diversity; H_E, average unbiased expected heterozygosity across all loci; H_O, average heterozygosity across all loci; mtDNA, mitochondrial DNA; N (ND4), number of samples in the mtDNA data set; N_H, number of haplotypes; N (SNP), number of samples in the neutral SNP data set; SNP, single-nucleotide polymorphism. Diversity indices for the SNP data set are given along with their standard error.

Haplotype diversity was lower in western Australia and the Cocos (Keeling) Islands compared with the rest of the locations. The haplotype network revealed that no haplotypes were shared between the samples from the Chagos Archipelago and any of the other locations. All but one individual from the Cocos (Keeling) Islands shared the same private haplotype (Figure 2). The haplotype network revealed two main groups of haplotypes, one including most haplotypes from western Australia and the other including most haplotypes from eastern Australia (Figure 2). Samples from Indonesia included haplotypes from both main haplotype groups.

Pairwise Φ_{ST} between locations within the same region (eastern and western Australia) were not significant, with the exception of Scott Reef that showed significant pairwise Φ_{ST} with other locations in western Australia (Φ_{ST} 0.105–0.160, Supplementary Table S1 and Figure 2). All pairwise Φ_{ST} between locations in different regions were highly significant and very high (Φ_{ST} range: 0.197–0.83), with the exception of the comparison between Scott Reef and Misool. Strong genetic differentiation was observed not only across large oceanic expanses but also between regions along the continental shelves of Australia and Indonesia. Pairwise Φ_{ST} estimated separately for males and females showed no differentiation between North GBR and South GBR. Pairwise Φ_{ST} estimates between Rowley Shoals and North GBR, and Rowley Shoals and South GBR were significant for both sexes, yet the level of differentiation was higher for females (0.65 and 0.67, respectively) than for males (0.49 and 0.50, respectively).

There was no evidence of recent demographic expansion as Tajima's D and Fu's Fs tests of neutrality were nonsignificant after Bonferroni's corrections (Supplementary Table S2). Rowley Shoals (D − 1.91294, $P = 0.04700$ and $Fs = − 1.71855$, $P = 0.029$) and Ningaloo Reef ($Fs = − 3.19859$, $P = 0.006$) did, however, show some moderate evidence of past population expansion, although not significant after adjusting for multiple comparisons (Supplementary Table S2).

Identification of loci putatively under selection

As expected, the two tests that did not explicitly account for evolutionary nonindependence among sampling locations identified the largest number of outliers: Arlequin identified 25 outliers and BAYESCAN 28 (Figure 3a). Yet, given the evidence of spatial autocorrelation in this species (see Results below), it seems likely that at least some of these loci may represent false positives (Lotterhos and Whitlock, 2014). The FLK approach, which takes into account evolutionary nonindependence, and has been reported as having very low false positive rates under IBD scenarios identified 17 outlier loci, 8 of which were also identified by both BAYESCAN and Arlequin

(Figure 3b). OutFLANK was the most conservative of all approaches, reporting only three outlier loci, all of which, however, were also identified by the other three methods (Figures 3c and d).

Genetic distance at the eight outlier loci identified by at least three outlier tests, and at the three outlier loci identified by all tests, were highest when comparing any pair of locations spanning the Torres Strait (that is, eastern Australia locations vs all other locations; Figures 3e and f, respectively). When all pairwise comparisons were considered there was no significant correlation between genetic distance estimated from neutral loci and from loci under selection (Figures 3e and f), indicating that divergence at putatively neutral and outlier loci is the result of different processes. However, when eastern Australian locations were excluded, genetic differentiation at neutral and outlier loci exhibited a clear, linear and significant correlation (open circles in Figure 3e: $P = 6 \times 10^{-8}$, $R^2 = 0.83$; and Figure 3f: $P = 9 \times 10^{-9}$, $R^2 = 0.87$).

SNP-based analysis of diversity and structure

The final neutral SNP data set included 4798 polymorphic, biallelic SNPs genotypes for 170 individuals. None of the loci deviated from Hardy–Weinberg equilibrium across three or more locations. Observed and expected heterozygosity ranged between 0.139 and 0.294 and between 0.131 and 0.291, respectively (Table 1). Fixation indices revealed no significant heterozygosity deficiency, but a significant heterozygosity excess was evident in the Cocos (Keeling) Islands, and to a lesser extent in the Chagos Archipelago and Coral Sea. However, both Cocos (Keeling) Islands and the Coral Sea had very low sample sizes, and therefore the reported heterozygosity excess could also be due to sampling bias.

Pairwise F_{ST} were low (0.0015–0.0187) and mostly not significant when comparing locations along Australia's and Indonesia's continental shelves (with the exception of some comparisons between eastern and western Australian locations; see Supplementary Table S1 and Figure 4). Between these locations, pairwise F_{ST} were on average two orders of magnitude lower than pairwise Φ_{ST}. Genetic distances across large oceanic expanses however were much higher (0.0688– 0.5148), and F_{ST} between samples from the Chagos Archipelago and all other locations were similar to estimates of Φ_{ST} obtained from the mtDNA data (Figure 4 and Supplementary Table S2).

The difference between Φ_{ST} and F_{ST} was largely dependent upon the scale considered. The largest difference was for pairwise comparisons between locations far apart but located within Australia's and Indonesia's continental shelves, whereas Φ_{ST} and F_{ST} difference were
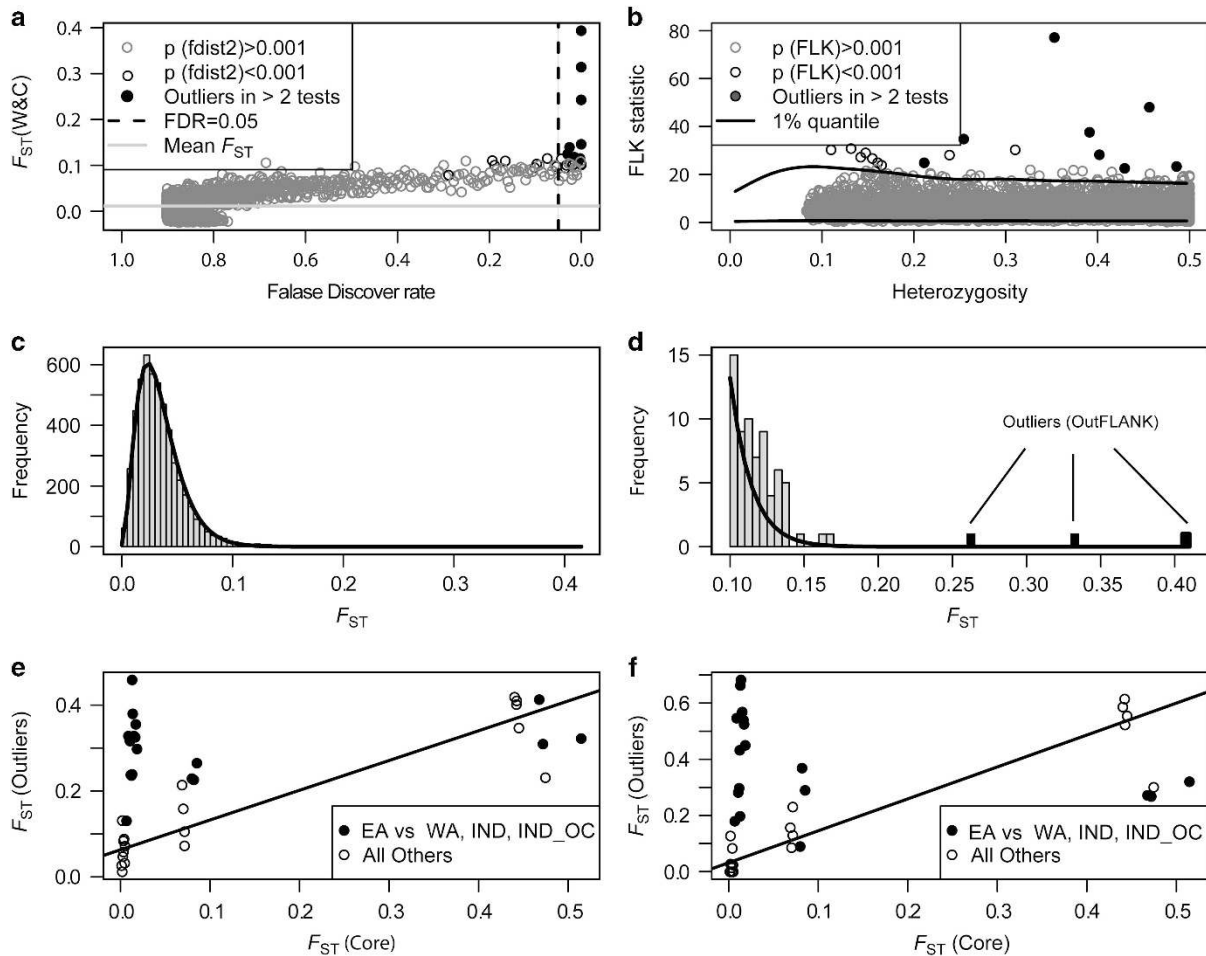
Figure 3 (a) Outlier loci identified as potentially under divergent selection by Arlequin and BAYESCAN. The dashed line represents the false discovery rate of 0.05, black open circles to the right of the dashed line represent the outlier loci that were identified by both methods and black filled circles represent the outliers identified by at least three independent methods. (b) Results from FLK. Black lines define the 99% probability envelope of the neutral distribution of FLK statistics, black open circles show outliers identified by FLK at a critical *P*-value of 0.001 and filled black circles represent the eight outliers that were identified by at least three outlier tests. (c) F_{ST} distribution of neutral loci estimated using the OutFLANK method. (d) Right tail of the neutral F_{ST} distribution, showing the three outliers identified by OutFLANK (that were also identified by all other approaches). (e, f) Relationship between 'neutral' F_{ST} (estimated from the core distribution of F_{ST}) and average F_{ST} of the eight outliers that were identified by at least three outlier tests (e) and the three outliers identified by the OutFLANK approach (f). There was no significant correlation between 'neutral' F_{ST} and F_{ST} estimated from outlier loci when all locations were included (e, f). When all locations from eastern Australia (filled black circles) are excluded from the analyses, there is a positive linear relationship between neutral F_{ST} and F_{ST} estimated from outlier loci ((e): $P = 6 \times 10^{-8}$, $R^2 = 0.83$; (f) $P = 9 \times 10^{-9}$, $R^2 = 0.87$).

within the same order of magnitude when comparing either locations close by or locations separated by large oceanic expanses (Figure 4).

The function chooseK.py reported a $K^*\varepsilon$ of 2 and a $K_\varnothing c$ of 4, suggesting a range of K of 2–4 as useful description of the data. Fivefold cross-validation using both the logistic and simple prior showed the lowest prediction error at $K = 3$, but prediction error fell within 1 standard error of the estimate at $K = 2$ (Supplementary Figure S1a). For the simple prior analysis, LLBO reaches a maximum at $K = 2$ and decreases with increasing model complexity (Supplementary Figure S1b). In the logistic prior analyses, LLBO changes abruptly from $K = 1$ to $K = 2$, and then continues to slightly increase with increasing model complexity.

We report the results from the logistic prior analyses using both $K = 2$ and $K = 3$ (Figures 5a and b, respectively). The $K = 2$ analyses identified two distinct genetic clusters (Figure 5a), representing the Chagos Archipelago and all locations in Australia and Indonesia.

Individuals from the Cocos (Keeling) Islands showed some degree of admixture, sharing ancestry with both genetic clusters. If model complexity was increased ($K = 3$), another genetic cluster representing all locations east of the Torres Strait emerged.

Results from DAPC analyses revealed very similar patterns (Figure 5c). DAPC performed on data from all locations revealed the greatest differentiation in locations separated by large oceanic distances (Figure 5d). The first discriminant function clearly showed greater differentiation between samples from the Chagos Archipelagos and the Cocos (Keeling) Islands and all other locations, whereas locations in different regions of Australia and Indonesia separate along the second discriminant function (Figure 5c). The result of the DAPC performed on samples from Australia and Indonesia further revealed subtler patterns of genetic differentiation (Figure 5d), in particular differentiation between the western Australian locations and Indonesia.
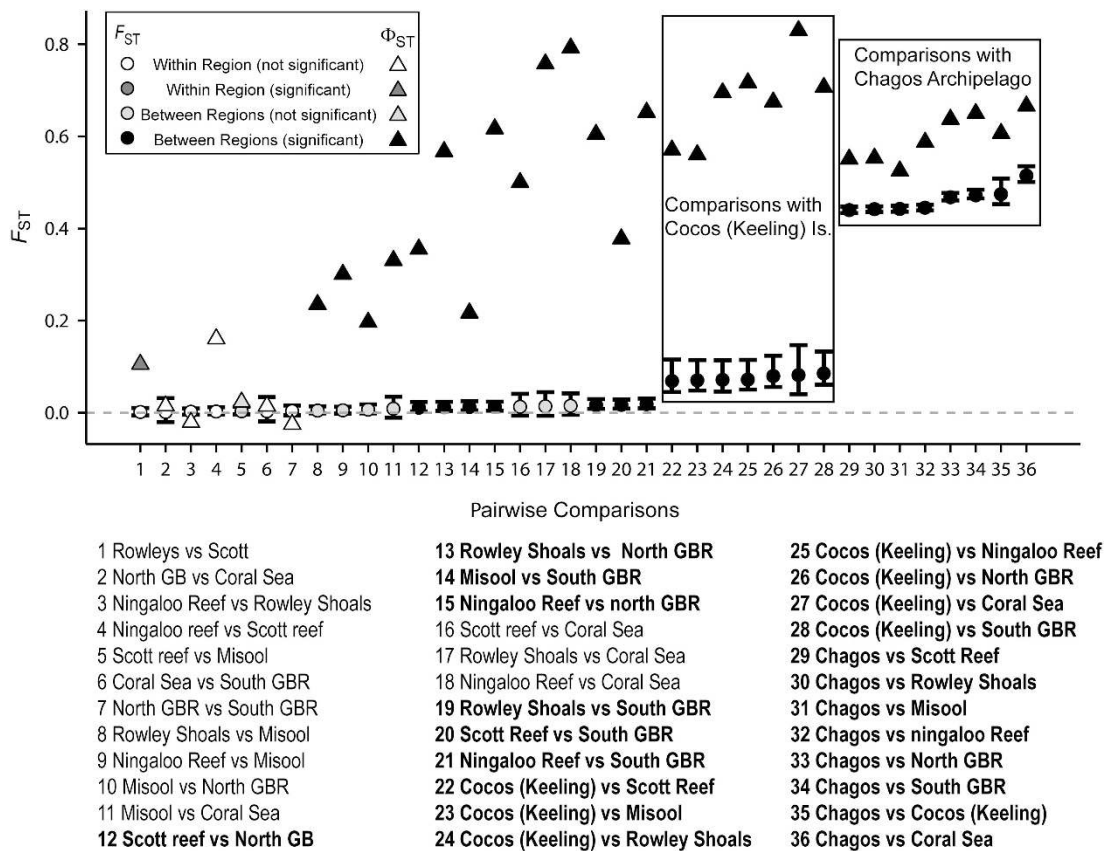
Figure 4 Estimates of pairwise genetic differentiation across all sampled location estimated from SNP data (F_{ST}) and mtDNA data (Φ_{ST}). Comparisons are arranged on the x axis in ascending order of F_{ST} values. The error bars represent 95% confidence intervals estimated by bootstrapping individuals within locations (only for F_{ST}). Symbols are color coded based on whether the comparisons are within or among distinct regions and whether they are or are not statistically significant.

Each sampling location appeared to be most similar to the two geographically closest locations. Furthermore, the densities along the first discriminant function closely matched the longitudinal distribution of the sampling locations (west to east), whereas densities on the second discriminant function were clearly correlated with latitude. If the order of the second discriminant function were reversed, the plot would closely match the geographic distribution of the sampling locations, a pattern that suggests a stepping stone model of dispersal (Jombart et al., 2010).

Isolation by distance
There was a clear, linear and significant relationship between geographic distance and genetic distance as measured by both F_{ST} and Φ_{ST} (Figure 6). Geographic distance explained 92% and 79% of the variation in genetic distance at nuclear loci and mtDNA, respectively. Although genetic differentiation along continental shelves follows an IBD model, this linear relationship breaks down when samples from the Chagos Archipelago and the Cocos (Keeling) Islands are included (data not shown). Both nuclear and mtDNA distances follow an IBD model, but the slope of the fitted linear model is nearly two orders of magnitude higher for mtDNA.

Fine-scale spatial genetic structure
Spatial autocorrelation analyses performed separately for samples collected in eastern Australia and western Australia revealed a sharply contrasting pattern. In eastern Australia there was a trend of

decreasing spatial autocorrelation with increasing geographic distance (Figure 6c), yet this trend did not significantly deviate from the null model, mostly because of the large variance in spatial autocorrelation within reefs. Conversely, there was a clear and significant pattern of spatial autocorrelation in isolated reefs of western Australia, showing that individuals within single reefs are more genetically similar than individuals across any other distance class (Figure 6d).

DISCUSSION
We used thousands of genome-wide SNPs and mtDNA data to investigate both neutral and putatively adaptive genetic variation in a key coral reef predator, the grey reef shark, across a substantial portion of its geographic range. By using both neutral and outlier loci, and sampling at multiple spatial scales across different seascapes, we unveiled complex cryptic patterns of genetic structure. Our findings show that genetic connectivity is maintained along continental shelves by male dispersal, whereas strong genetic structure revealed by mitochondrial markers suggests females are more philopatric. Reef systems separated by large expanses of oceanic water showed high genetic differentiation, and thus appear to receive few transoceanic immigrants. Several SNPs were consistently identified with different analytical approaches as divergent outliers and are therefore suggestive of spatially diversifying selection.

Our results show that large expanses of oceanic water represent strong barriers to both male and female reef shark dispersal. Conversely, along the continental shelves of Australia and Indonesia,
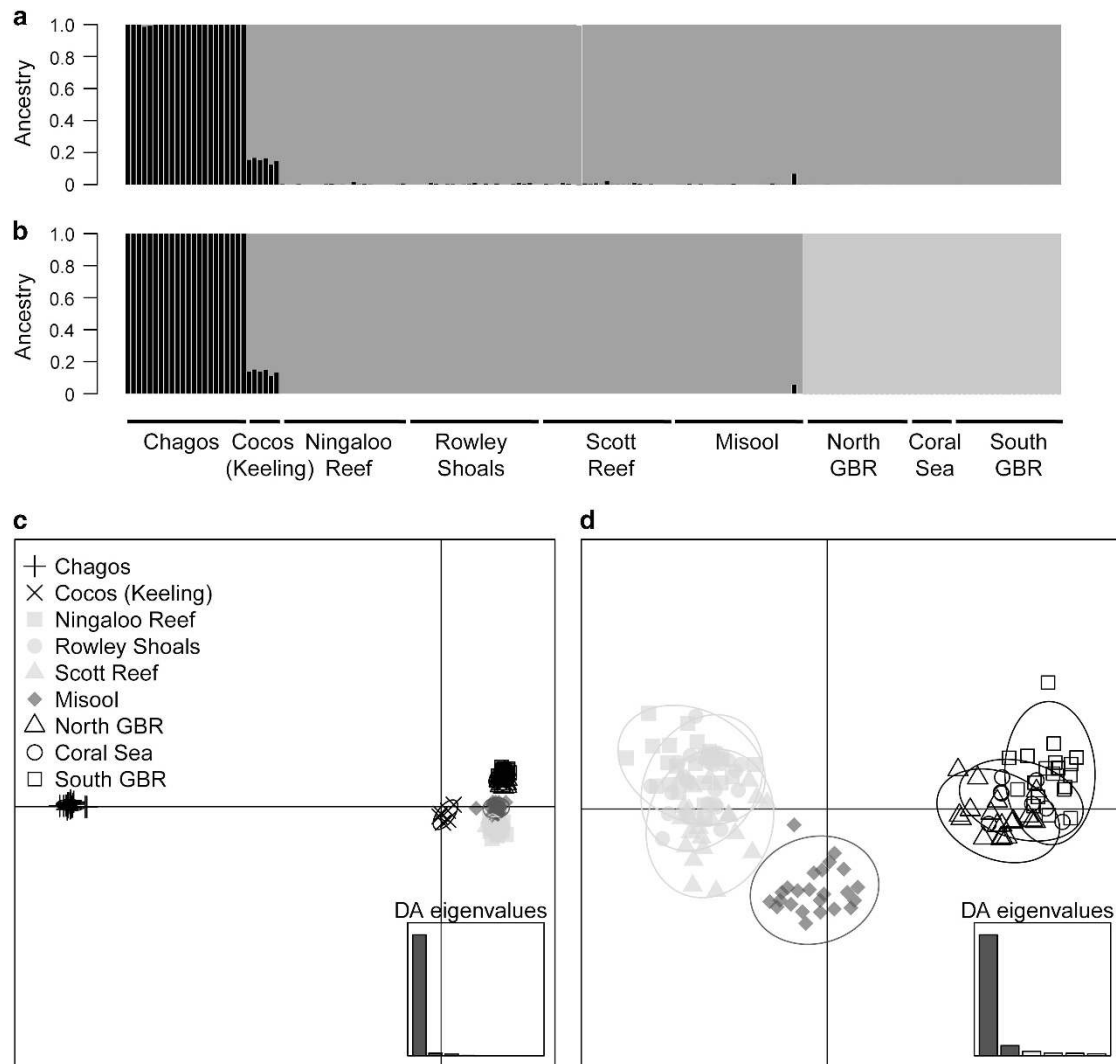
Figure 5 Results from the fastSTRUCTURE and DAPC analyses. (**a**, **b**) Results from fastSTRUCTURE using logistic prior, $K = 2$ and $K = 3$, respectively. (**c**, **d**) Results from DAPC analysis performed on SNP data from all locations (**c**) and only from locations in Australia and Indonesia (**d**). The x and y axes represent the first and second discriminant functions, respectively and ellipses represent 95% confidence inertia ellipses.

genetic differentiation at nuclear SNPs was weak but nevertheless statistically significant. Both the fastSTRUCTURE analyses and the DAPC analyses suggest the presence of continental shelf genetic structure, though it should be noted that the logistic prior in fastSTRUCTURE can tend to overestimate the true value for K, and that increasing K after 2 only resulted in a statistically insignificant decrease in prediction error, suggesting that $K = 3$ may represent an over-parameterization of the model (Raj *et al.*, 2014). Furthermore, the STRUCTURE algorithm assumes individuals are sampled from a set of discrete populations, an assumption that is clearly not met under an IBD scenario. Under these conditions, discrete sampling over a continuous distribution could result in artificial clusters, further complicating the interpretation of the results as evidence of a barrier to gene flow between eastern and western Australia. Grey reef sharks have a continuous distribution along the northern Australian and Indonesian coasts (Last and Stevens, 2009) and there is no contemporary physical barrier to dispersal along the continental shelves of Australia and Indonesia. The Torres Strait was exposed during the last glacial maximum, but it was submerged again 8000 years ago, and given the dispersal potential of grey reef sharks (Heupel *et al.*, 2010;

Momigliano *et al.*, 2015a) and their generation time (Robbins, 2006), it seems likely that grey reef sharks in the region may have reached, or are approaching, drift equilibrium following the Torres Strait's disappearance. Indeed, our data revealed a very clear IBD pattern that explained > 90% of the variation in pairwise F_{ST}, suggesting that geographical distance is the main driver of neutral genetic differentiation along the continental shelves of Australia and Indonesia.

The lack of a clear biogeographic barrier between eastern and western Australia, and between Australia and Indonesia, may seem surprising in light of the fact that a closely related species, the blacktip reef shark (*C. melanopterus*), shows marked genetic structure in the same region (Vignaud *et al.*, 2014). *C. melanopterus*, however, is smaller than *C. amblyrhynchos* and larger sharks can traverse greater distances for comparable energetic expenditure (Parsons, 1990). Furthermore, *C. melanopterus* exhibits strong reproductive philopatry (Mourier and Planes, 2013), a behavior that may weaken genetic connectivity (Dudgeon *et al.*, 2012). The patterns of nuclear genetic structure across Australia and Indonesia for grey reef sharks are more similar to larger species with higher dispersal potential, such as scalloped hammerheads (*Sphyrna lewini*), dusky sharks (*C. obscurus*)
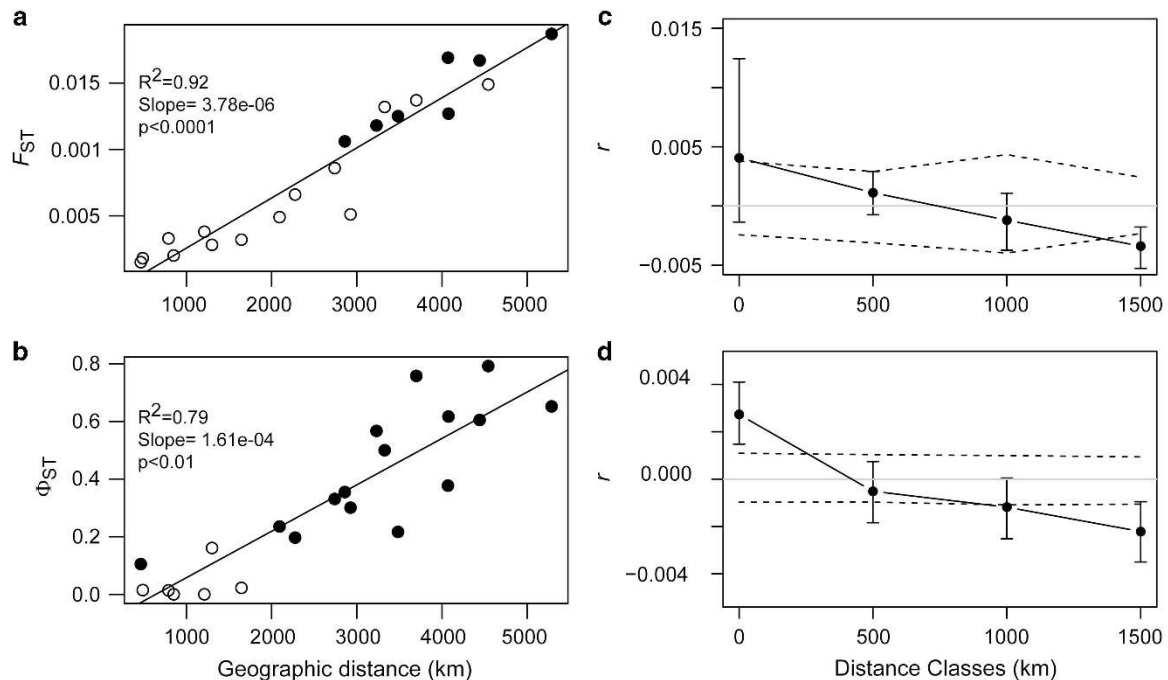
Figure 6 Patterns of isolation by distance (**a**, **b**) using F_{ST} (**a**) estimated from SNP data from the 'core' of the F_{ST} distribution and Φ_{ST} (**b**, mtDNA data). R^2 and slope of the fitted linear model were estimated by major axes regression, whereas the significance of the correlation between genetic and geographic distance (*P*) was estimated via Mantel test using 30 000 randomizations. Open circles represent statistically nonsignificant pairwise distances and filled circles represent significant pairwise comparisons. (**c**, **d**) Estimates of spatial autocorrelation (*r*) across different distance classes (within the same reef and within 500, 1000 and 1500 km, respectively) using samples from eastern Australia (**c**) and western Australia (**d**). Error bars represent 95% confidence intervals estimated from 999 bootstraps, and the dotted lines represent the 95% confidence intervals of the null model of no spatial autocorrelation.

and blue sharks (*Prionace glauca*) (Ovenden et al., 2009). Comparative large-scale patterns have also been reported in the other common reef shark species, the whitetip reef shark *Triaenodon obesus*, although a more passive dispersal mechanism has been suggested for this finding (Whitney et al., 2012). Although genetic differentiation seems to be weak or absent for many shark species across the Torres Strait, many teleosts show striking patterns of genetic differentiation in the same region (Elliott, 1996; van Herwerden et al., 2006; Horne et al., 2011, 2012). Most bony fish, however, have a bipartite life cycle: adults are sedentary and dispersal is achieved by means of pelagic propagules. On both sides of Australia the prevailing currents transport larvae southwards, hindering dispersal across the northern coast. Sharks, on the other end, disperse by active swimming, and hence their dispersal is less affected by prevailing oceanographic features (Momigliano et al., 2015b).

In contrast to the weak genetic differentiation we report for nuclear loci, genetic differentiation at mtDNA markers was strong (Φ_{ST} at mtDNA up to 77 times higher than nuclear F_{ST}), even within locations not separated by large oceanic expanses. Divergence followed the same pattern of IBD revealed by nuclear SNPs, but the slope of the relationship between genetic and geographic distance was nearly two orders of magnitude higher for mtDNA (Figures 6a and b). Despite the strong genetic differentiation, the three most common haplotypes were represented in eastern Australia, western Australia and Indonesia, although at different frequencies. Samples collected from Misool shared both haplotypes common in eastern and western Australia, with intermediate frequencies, as would be expected in the case of isolation by distance in the absence of biogeographic barriers.

The limited genetic structuring at neutral loci is consistent with behavioral data. Grey reef sharks, particularly adult males, have the

potential to undertake movements of >100 km across oceanic waters (Heupel et al., 2010), yet they are found nearly exclusively in coral reef habitats (Espinoza et al., 2014). Therefore, along the continental shelf the most likely mode of dispersal resulting in a pattern of IBD is one where coral reef habitats represent 'stepping stones', allowing for the maintenance of genetic connectivity (at least for the nuclear genome). This interpretation is strengthened by the fact that genetic differentiation across regions separated by large oceanic expanses is high for both the nuclear and mtDNA data sets, suggesting that deep oceanic waters represent a strong barrier to both male and female dispersal.

In contrast to the low genetic differentiation at neutral markers among locations within the Australian and Indonesian continental shelves, genetic differentiation at loci putatively under selection was more than an order of magnitude higher, and showed a clear biogeographic pattern. Divergence at neutral SNPs and SNPs putatively under diversifying selection were not correlated when all pairs of locations were considered. However, when locations from eastern Australia were excluded from analysis, divergence at outlier loci closely mirrored divergence at neutral loci. This pattern suggests that among most locations (that is, pairwise comparisons between all locations with the exception of eastern Australia) divergence at neutral and outlier loci is in both cases likely the product of genetic drift. Divergence at neutral and outlier loci is however not correlated if we include locations from eastern Australia in the analyses, suggesting that here divergence at neutral and outlier loci is the product of fundamentally different processes.

Theoretically, local selection is not the only process that can produce outlier loci. For example, the frequency of mutations arising at the front of a range expansion may change dramatically (Excoffier and Ray, 2008); however, neutrality tests did not show evidence of

recent demographic expansions in grey reef sharks, and hence gene surfing during spatial population expansions is unlikely to be responsible for the observed patterns. Evolutionary nonindependence and unequal levels of genetic differentiation among sampled populations can also result in large numbers of false positives, a problem that has been widely recognized in recent years (Bonhomme et al., 2010; Narum and Hess, 2011; Lotterhos and Whitlock, 2014). The original Lewontin–Krakauer test was particularly sensitive to both these issues.

The analytical approaches employed in this study, however, are exceptionally robust under scenarios of unequal genetic differentiation and autocorrelated spatial structure. The methods employed by both Foll and Gaggiotti (2008) and Excoffier et al. (2009) account for unequal levels of genetic differentiation among populations, yet can still yield a large number of false positives when genetic structure is spatially autocorrelated. However, the method employed by FLK directly accounts for evolutionary branching and heterogeneous drift (Bonhomme et al., 2010), and both FLK (Lotterhos and Whitlock, 2014) and OutFLANK (Whitlock and Lotterhos, 2015) have extremely low false positive rates under most evolutionary scenarios. By applying all of these tests in conjunction, and compiling a consensus set of outlier loci, we minimized the chances of false positives.

The Torres Strait was completely exposed during the last glacial maximum and was not submerged again until 2000 years after the end of the last glaciation, ~ 8000 years ago. Barriers to dispersal are expected to promote local adaptation (Marshall et al., 2010) and it is therefore possible that local selective pressures and historical reproductive isolation have acted in concert to produce the pattern shown here. However, the outliers we identified are not necessarily all under local selection. Although past isolation may have favored local adaptation, it may also have resulted in genetic incompatibilities (that is, ecologically independent, endogenous selection). Bierne et al. (2011) discovered that although the location of such endogenous barriers to gene flow is expected to shift over time, they often become trapped by exogenous barriers because of local selection (the 'coupling hypothesis'). Hence, determining with certainty whether a specific outlier locus is under endogenous or exogenous selection is not possible using outlier tests or genotype-by-environment association tests alone, because both processes can become coupled across the same environmental gradients.

Our finding that genetic differentiation described using mtDNA sequences was substantially larger than for neutral nuclear SNPs, and more pronounced in females than males, especially when comparing locations separated by large distances along the continental shelf (in these comparisons Φ_{ST} was on average nearly 40 times higher than F_{ST}), is indicative of sex-biased dispersal. These results should be interpreted with some caution, as we only had enough samples in three populations to analyze males and females separately. Nevertheless, this pattern of sex-biased dispersal and local adaptation is consistent with the hypothesis advanced by Portnoy et al. (2015) that dispersing males and philopatric females may favor local adaptation by simultaneously facilitating the dispersal and local sorting of adaptive alleles. As male-biased dispersal is a common strategy in both elasmobranchs (Pardini et al., 2001; Daly-Engel et al., 2012) and marine mammals (Escorza-Treviño and Dizon, 2000; Möller and Beheregaray, 2004; Ahonen et al., 2016), future comparative studies testing this hypothesis could shed new light on the evolution of local adaptation in marine predators.

In the absence of a reference genome, it was not possible to identify which regions of the genome are under selection (BLAST searches did not yield any unambiguous match). Furthermore, as the signature of selection follows a clear biogeographic pattern, it may not be feasible with the current sampling design to identify the underlying selective pressure by investigating correlations between SNPs and environmental variables. Nonetheless, our data are consistent with the hypothesis that grey reef sharks in eastern Australia, which have undergone dramatic declines (Robbins et al., 2006), are under different selective pressures than elsewhere throughout the sampled range. Future management strategies for this species in eastern Australia and elsewhere should consider taking into account that grey reef sharks are likely to be locally adapted and that local declining populations are unlikely to be rescued by migrants if they exhibit lower fitness because of phenotype–environment mismatches (Marshall et al. 2010). Furthermore, although there is no evidence to date that recent fishery-induced declines have affected genetic diversity, it is possible that further, more dramatic declines may lead to a stochastic loss of local adaptations and to the erosion of standing genetic variation that may ultimately influence adaptive potential.

We also note the association between the spatial distribution of coral reefs and patterns of fine-scale genotypic spatial structure. In eastern Australia, reefs within the GBR, and to a lesser extent reefs in the Coral Sea, are relatively close. Within the GBR the average between-reef distance is < 2 km, and the oceanic reef in the Coral Sea (the North East Herald Cay) that we sampled is < 200 km offshore from the outer shelf of the GBR and potentially connected through multiple intermediate reefs that may act as stepping stones. In contrast, coral reef habitats in western Australia are far more fragmented, often separated by hundreds of km of unsuitable habitat. Our results showed that in eastern Australia there is no strong evidence of genotypic spatial autocorrelation at the reef scale. However, despite the fact that within each spatial class r was within the expectations of the null model (Figure 6c and d), there was a clear, although very weak, trend of decreasing autocorrelation with increasing distance, compatible with an IBD model of dispersal.

The variance in autocorrelation at the reef scale was very large, suggesting a mix of local individuals and recent migrants within single reefs in eastern Australia. This pattern is consistent with the high potential for short-distance migration in reef sharks, and is also consistent with results obtained by Momigliano et al. (2015a) using microsatellite markers. In western Australia, a clear pattern of genotypic autocorrelation was observed at the reef scale. Genotypic autocorrelation was highest when comparing individuals from the same reef, and dropped abruptly when moving to the next distance class. This pattern is consistent with a higher degree of intergenerational reef fidelity, suggesting that reefs isolated by large distances of unsuitable habitat (even if located within the same continental shelf) may be effectively self-seeding and demographically independent. However, further studies are needed to confirm the effects of habitat on connectivity. Employing a full seascape genomic approach, for example, by developing isolation-by-resistance models (McRae and Beier, 2007) as suggested by Momigliano et al. (2015a, b), will be necessary but is not possible with the current sampling design.

CONCLUSION

We investigated the large-scale and fine-scale genetic structure of grey reef sharks using a combination of genome-wide nuclear SNPs and a mtDNA marker. We revealed that patterns of gene flow are associated with the geographic separation of reef systems and, with the exception of localities separated by oceanic waters, dispersal appears to be male-biased. Large oceanic distances are strong barriers to both male and female dispersal, suggesting that reefs surrounded by oceanic waters are less likely to be replenished by migrant sharks. These results have implications for the effectiveness of spatially discontinuous Marine

Protected Areas in terms of the protection of grey reef sharks, for which replenishment may be dependent upon the level of habitat fragmentation exhibited by different reef systems (Momigliano *et al.*, 2015b). Our finding of signatures of selection, despite generally negligible levels of genetic differentiation at biparentally inherited and neutral markers, reveals the presence of biologically important cryptic genetic structure. These patterns of male-mediated dispersal and signatures of local selection are consistent with the recently advanced hypothesis that sex-biased dispersal may facilitate local adaptation by allowing, simultaneously, dispersal and the localized sorting of adaptive alleles.

DATA ARCHIVING

Mitochondrial DNA sequences were deposited in GenBank (accession numbers: KX712891-KX713065). SNP allele frequency data were deposited in the DRYAD digital repository http://dx.doi.org/10.5061/dryad.mt2m3.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Ahonen H, Lowther AD, Harcourt RG, Goldsworthy SD, Charrier I, Stow AJ (2016). The limits of dispersal: fine scale spatial genetic structure in Australian sea lions. *Front Mar Sci* **3**: 65.

Allendorf FW, Hohenlohe PA, Luikart G (2010). Genomics and the future of conservation genetics. *Nat Rev Genet* **11**: 697–709.

Almany G, Connolly S, Heath D, Hogan J, Jones G, McCook L *et al.* (2009). Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs* **28**: 339–351.

Arèvalo E, Davis SK, Sites JW (1994). Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Syst Biol* **43**: 387–418.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA *et al.* (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3**: e3376.

Barnett A, Abrantes KG, Seymour J, Fitzpatrick R (2012). Residency and spatial use by reef sharks of an isolated seamount and its implications for conservation. *PLoS One* **7**: e36574.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011). The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* **20**: 2044–2072.

Blower DC, Pandolfi JM, Bruce BD, Gomez-Cabrera MdC, Ovenden JR (2012). Population genetics of Australian white sharks reveals fine-scale spatial structure, transoceanic dispersal events and low effective population sizes. *Mar Ecol Prog Ser* **455**: 229–244.

Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J, Blott S *et al.* (2010). Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics* **186**: 241–262.

Bradbury IR, Hubert S, Higgins B, Bowman S, Borza T, Paterson IG *et al.* (2013). Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evol Appl* **6**: 450–461.

Castro A, Stewart B, Wilson S, Hueter R, Meekan M, Motta P *et al.* (2007). Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*). *Mol Ecol* **16**: 5183–5192.

Clement M, Posada D, Crandall KA (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**: 1657–1659.

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ *et al.* (2012). Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). *PLoS One* **7**: e29986.

DeFaveri J, Jonsson PR, Merilä J (2013). Heterogeneous genomic differentiation in marine threespine sticklebacks: adaptation along an environmental gradient. *Evolution* **67**.

Dudgeon CL, Blower DC, Broderick D, Giles JL, Holmes BJ, Kashiwagi T *et al.* (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *J Fish Biol* **80**: 1789–1843.

Elliott NG (1996). Allozyme and mitochondrial DNA analysis of the tropical saddle-tail sea perch, *Lutjanus malabaricus* (Schneider), from Australian waters. *Mar Freshw Res* **47**: 869–876.

Epperson B (2005). Estimating dispersal from short distance spatial autocorrelation. *Heredity* **95**: 7–15.

Escorza-Treviño S, Dizon AE (2000). Phylogeography, intraspecific structure and sex-biased dispersal of Dall's porpoise, *Phocoenoides dalli*, revealed by mitochondrial and microsatellite DNA analyses. *Mol Ecol* **9**: 1049–1060.

Espinoza M, Cappo M, Heupel MR, Tobin AJ, Simpfendorfer CA (2014). Quantifying shark distribution patterns and species-habitat associations: implications of Marine Park Zoning. *PLoS One* **9**: e106885.

Espinoza M, Heupel MR, Tobin A, Simpfendorfer C (2015a). Residency patterns and movements of grey reef sharks (*Carcharhinus amblyrhynchos*) in semi-isolated coral reef habitats. *Mar Biol* **162**: 343–358.

Espinoza M, Lédée EJ, Simpfendorfer CA, Tobin AJ, Heupel MR (2015b). Contrasting movements and connectivity of reef-associated sharks using acoustic telemetry: implications for management. *Ecol Appl* **25**: 2101–2118.

Excoffier L, Hofer T, Foll M (2009). Detecting loci under selection in a hierarchically structured population. *Heredity* **103**: 285–298.

Excoffier L, Lischer HE (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**: 564–567.

Excoffier L, Ray N (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol Evol* **23**: 347–351.

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

Friedlander AM, Caselle JE, Ballesteros E, Brown EK, Turchik A, Sala E (2014). The real bounty: marine biodiversity in the Pitcairn Islands. *PLoS One* **9**: e100142.

Fu Y-X (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.

Graham NA, Spalding MD, Sheppard CR (2010). Reef shark declines in remote atolls highlight the need for multi-faceted conservation action. *Aquat Conserv* **20**: 543–548.

Guo B, DeFaveri J, Sotelo G, Nair A, Merilä J (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biol* **13**: 19.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.

Hellberg ME, Burton RS, Neigel JE, Palumbi SR (2002). Genetic assessment of connectivity among marine populations. *Bull Mar Sci* **70**: 273–290.

Heupel MR, Simpfendorfer CA, Fitzpatrick R (2010). Large-scale movement and reef fidelity of grey reef sharks. *PLoS One* **5**: e9650.

Horne JB, Momigliano P, Welch DJ, Newman SJ, van Herwerden L (2011). Limited ecological population connectivity suggests low demands on self-recruitment in a tropical inshore marine fish (*Eleutheronema tetradactylum*: Polynemidae). *Mol Ecol* **20**: 2291–2306.

Horne JB, Momigliano P, Welch DJ, Newman SJ, Van Herwerden L (2012). Searching for common threads in threadfins: phylogeography of Australian polynemids in space and time. *Mar Ecol Prog Ser* **449**: 263–276.

Inoue JG, Miya M, Tsukamoto K, Nishida M (2003). Evolution of the deep-sea gulper eel mitochondrial genomes: large-scale gene rearrangements originated within the eels. *Mol Biol Evol* **20**: 1917–1924.

Jaccoud D, Peng K, Feinstein D, Kilian A (2001). Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res* **29**: e25.

Jensen JD, Foll M, Bernatchez L (2016). The past, present and future of genomic scans for selection. *Mol Ecol* **25**: 1–4.

Jombart T, Devillard S, Balloux F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* **11**: 94.

Kalinowski ST (2002). Evolutionary and statistical properties of three genetic distances. *Mol Ecol* **11**: 1263–1273.

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013). diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* **4**: 782–788.

Kelley JL, Brown AP, Therkildsen NO, Foote AD (2016). The life aquatic: advances in marine vertebrate genomics. *Nat Rev Genet* **17**: 523–534.

Lamichhaney S, Barrio AM, Rafati N, Sundström G, Rubin C-J, Gilbert ER *et al.* (2012). Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. *Proc Natl Acad Sci USA* **109**: 19345–19350.

Last PR, Stevens JD (2009). *Sharks and Rays of Australia*, 2nd edn. CSIRO: Collingwood.

Leigh JW, Bryant D (2015). popart: full-feature software for haplotype network construction. *Methods Ecol Evol* **6**: 1110–1116.

Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Lotterhos KE, Whitlock MC (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Mol Ecol* **23**: 2178–2192.

Lukoschek V, Waycott M, Keogh J (2008). Relative information content of polymorphic microsatellites and mitochondrial DNA for inferring dispersal and population genetic structure in the olive sea snake, *Aipysurus laevis*. *Mol Ecol* **17**: 3062–3077.

Marshall DJ, Monro K, Bode M, Keough MJ, Swearer S (2010). Phenotype-environment mismatches reduce connectivity in the sea. *Ecol Lett* **13**: 128–140.

McRae BH, Beier P (2007). Circuit theory predicts gene flow in plant and animal populations. *Proc Natl Acad Sci USA* **104**: 19885–19890.

Möller LM, Beheregaray LB (2004). Genetic evidence for sex-biased dispersal in resident bottlenose dolphins (*Tursiops aduncus*). *Mol Ecol* **13**: 1607–1612.

Momigliano P, Harcourt R, Robbins WD, Stow A (2015a). Connectivity in grey reef sharks (*Carcharhinus amblyrhynchos*) determined using empirical and simulated genetic data. *Sci Rep* **5**: 13229.

Momigliano P, Harcourt R, Stow A (2015b). Conserving coral reef organisms that lack larval dispersal: are networks of Marine Protected Areas good enough? *Front Mar Sci* **2**: 16.

Mourier J, Maynard J, Parravicini V, Ballesta L, Clua E, Domeier ML *et al.* (2016). Extreme inverted trophic pyramid of reef sharks supported by spawning groupers. *Curr Biol* **26**: 2011–2016.

Mourier J, Planes S (2013). Direct genetic evidence for reproductive philopatry and associated fine-scale migrations in female blacktip reef sharks (*Carcharhinus melanopterus*) in French Polynesia. *Mol Ecol* **22**: 201–214.

Nance HA, Klimley P, Galván-Magaña F, Martínez-Ortíz J, Marko PB (2011). Demographic processes underlying subtle patterns of population structure in the scalloped hammerhead shark, *Sphyrna lewini*. *PLoS One* **6**: e21459.

Narum SR, Hess JE (2011). Comparison of F_{ST} outlier tests for SNP loci under selection. *Mol Ecol Resour* **11**: 184–194.

O'Corry-Crowe G, Suydam R, Rosenberg A, Frost K, Dizon A (1997). Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Nearctic revealed by mitochondrial DNA. *Mol Ecol* **6**: 955–970.

Ovenden JR (2013). Crinkles in connectivity: combining genetics and other types of biological data to estimate movement and interbreeding between populations. *Mar Freshw Res* **64**: 201–207.

Ovenden JR, Kashiwagi T, Broderick D, Giles J, Salini J (2009). The extent of population genetic subdivision differs among four co-distributed shark species in the Indo-Australian archipelago. *BMC Evol Biol* **9**: 40.

Pardini AT, Jones CS, Noble LR, Kreiser B, Malcolm H, Bruce BD *et al.* (2001). Sex-biased dispersal of great white sharks. *Nature* **412**: 139–140.

Parsons G (1990). Metabolism and swimming efficiency of the bonnethead shark *Sphyrna tiburo*. *Mar Biol* **104**: 363–367.

Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**: 2537–2539.

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.

Portnoy DS, McDowell JR, McCandless CT, Musick JA, Graves JE (2009). Effective size closely approximates the census size in the heavily exploited western Atlantic population of the sandbar shark, *Carcharhinus plumbeus*. *Conserv Genet* **10**: 1697–1705.

Portnoy DS, Puritz JB, Hollenbeck CM, Gelsleichter J, Chapman D, Gold JR (2015). Selection and sex-biased dispersal in a coastal shark: the influence of philopatry on adaptive variation. *Mol Ecol* **24**: 5877–5885.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Raj A, Stephens M, Pritchard JK (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**: 573–589.

Robbins WD (2006). Abundance, demography and population structure of the grey reef shark (*Carcharhinus amblyrhynchos*) and the white tip reef shark (Triaenodon obesus) (Fam. Charcharhinidae). Doctoral dissertation thesis. James Cook University.

Robbins WD, Hisano M, Connolly SR, Choat JH (2006). Ongoing collapse of coral-reef shark populations. *Curr Biol* **16**: 2314–2319.

Sandin SA, Smith JE, DeMartini EE, Dinsdale EA, Donner SD, Friedlander AM *et al.* (2008). Baselines and degradation of coral reefs in the northern Line Islands. *PLoS One* **3**: e1548.

Sansaloni C, Petroli C, Jaccoud D, Carling J, Detering F, Grattapaglia D *et al.* (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. *BMC Proceedings* **5**: 54.

Savolainen O, Lascoux M, Merilä J (2013). Ecological genomics of local adaptation. *Nat Rev Genet* **14**: 807–820.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

van Herwerden L, Choat JH, Dudgeon CL, Carlos G, Newman SJ, Frisch A *et al.* (2006). Contrasting patterns of genetic structure in two species of the coral trout *Plectropomus* (Serranidae) from east and west Australia: Introgressive hybridisation or ancestral polymorphisms. *Mol Phylogenet Evol* **41**: 420–435.

Vignaud TM, Mourier J, Maynard JA, Leblois R, Spaet JL, Clua E *et al.* (2014). Blacktip reef sharks, *Carcharhinus melanopterus*, have high genetic structure and varying demographic histories in their Indo-Pacific range. *Mol Ecol* **23**: 5193–5207.

Ward-Paige CA, Mora C, Lotze HK, Pattengill-Semmens C, McClenachan L, Arias-Castro E *et al.* (2010). Large-scale absence of sharks on reefs in the Greater-Caribbean: a footprint of human pressures. *PLoS One* **5**: e11968.

Whitlock MC, Lotterhos KE (2015). Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST}*. *Am Nat* **186**: S24–S36.

Whitney NM, Robbins WD, Schultz JK, Bowen BW, Holland KN (2012). Oceanic dispersal in a sedentary reef shark (*Triaenodon obesus*): genetic evidence for extensive connectivity without a pelagic larval stage. *J Biogeogr* **39**: 1144–1156.

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

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)