American Genetic Association

Journal of Heredity, 2018, 744–756
doi:10.1093/jhered/esy050
Original Article
Advance Access publication September 21, 2018

OXFORD

Original Article

The Influence of Selection on MHC DQA and DQB Haplotypes in the Endemic New Zealand Hector's and Māui Dolphins

Dorothea Heimeier, Alana Alexander, Rebecca M. Hamner, Franz Pichler, and C. Scott Baker

From the School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand (Heimeier, Hamner, Pichler, and Baker); Biodiversity Institute, University of Kansas, 1345 Jayhawk Boulevard, Lawrence, KS 66045 (Alexander); and Marine Mammal Institute and Department of Fisheries and Wildlife, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Drive, Newport, OR 97365 (Hamner and Baker). Dorothea Heimeier is at the Pirbright Institute, Ash Road, Pirbright, Woking GU240NF, UK. Rebecca M. Hamner is now at the Department of Life Sciences, Texas A&M University, Corpus Christi, 6300 Ocean Boulevard, Unit 5802, Corpus Christi, TX 78414. Alana Alexander is now at the Department of Anatomy, University of Otago, PO Box 56, Dunedin 9054, New Zealand.

Address correspondence to Dorothea Heimeier at the address above, or e-mail: dorothea.heimeier@auckland.ac.nz or dorothea.harrison@pirbright.ac.uk.

Received January 31, 2018; First decision May 2, 2018; Accepted September 19, 2018.

Corresponding Editor: William Sherwin

Abstract

Strong balancing selection on the major histocompatibility complex (MHC) can lead to different patterns in gene frequencies and neutral genomic variation within species. We investigated diversity and geographic structure of MHC genes DQA and DQB, as well as their inferred functional haplotypes, from 2 regional populations (East and West Coast) of the endangered Hector's dolphin (*Cephalorhynchus hectori hectori*) and the critically endangered Māui dolphin (*Cephalorhynchus hectori maui*) (West Coast, North Island), and contrasted these results with patterns from neutral microsatellites. The Māui had the lowest number of alleles for DQA (2) and DQB (3), consistent with strong genetic drift acting on this remnant population. However, the 2 retained DQA alleles are among the most divergent combinations of all 4 alleles found across the Hector's metapopulation, potentially reflecting the retention of divergent alleles due to balancing selection. The high frequency of the divergent *DQB*04* allele also gave this population the highest nucleotide diversity for DQB. Strong differentiation was evident for DQA, DQB, and DQA–DQB haplotypes between the regional populations of Hector's dolphins ($F_{ST} > 0.213$) and both subspecies ($F_{ST} > 0.311$). Differentiation was generally greater than observed at neutral microsatellite loci, suggesting the influence of selection between geographically proximate East and West Coast populations. This might be the result of spatial differences in directional selection on those opposite coastlines. In addition, measures of the ratio of nonsynonymous to synonymous substitutions (*dN*/*dS*) were consistent with balancing selection over evolutionary time. Together, these results suggest a complex interplay of balancing selection, directional selection, local fidelity, and genetic drift.

Subject areas: Population structure and phylogeography, Conservation genetics and biodiversity
Keywords: balancing selection, cetacean, MHC diversity, population genetics, small populations

The genes of the major histocompatibility complex (MHC) represent the most variable functional regions in the nuclear genome of vertebrates (Hedrick 1994). The MHC determines an organism's ability to cope with pathogens from the environment and hence may affect the long-term persistence of populations (Hedrick 1998; Paterson et al. 1998). The high level of polymorphism at MHC loci is demonstrated in the number of alleles and sequence variation among them, believed to be shaped and maintained by strong balancing selection as a result of pathogen-mediated selection (Hughes and Nei 1989; Hughes and Yeager 1998; Bernatchez and Landry 2003; Spurgin and Richardson 2010).

In large outbred populations, selection can be the main factor driving and maintaining abundant MHC variation, whereas in small populations, demographic and stochastic effects can have a greater influence on MHC diversity (Bernatchez and Landry 2003). In populations that have gone through a bottleneck, variation at MHC loci can sometimes be maintained (Aguilar et al. 2004, but see Hedrick 2004; van Oosterhout et al. 2006), although selection has to be strong to overcome genetic drift (selection effective population size $N_e s >> 1$, Kimura 1983). In addition, when selection acts on MHC loci prior to a bottleneck event, combined with drift during the bottleneck, the result can be an overall loss of MHC polymorphism that is on average 15% greater than loss of neutral genetic diversity (Robertson 1962; Sutton et al. 2011).

It has been suggested that limited MHC diversity resulting from bottleneck events can lead to isolated populations becoming increasingly vulnerable to infectious diseases (Bowen et al. 2002). This possibility has motivated many studies to investigate MHC functional diversity in small and/or endangered populations (e.g., Gutierez-Espeleta et al. 2000; Hedrick 2003).

There is also a special interest in characterizing MHC diversity for marine mammals, many of which have gone through bottlenecks associated with human activities, such as elephant seals (Weber et al. 2004). Furthermore, in the marine environment, geographic barriers to movements are less obvious than on land, and selective forces such as pathogens are less well described.

The Hector's dolphin (*Cephalorhynchus hectori*), endemic to the coastal waters of New Zealand, is therefore of particular interest with regard to MHC diversity, as it is both marine and characterized by small populations that declined due to anthropogenic activities. It consists of 2 subspecies: the Māui dolphin (*Cephalorhynchus hectori maui*) and the Hector's dolphin (*Cephalorhynchus hectori hectori*), listed by the IUCN as critically endangered and endangered, respectively (Reeves et al. 2013). The Māui dolphin is limited to the northwest coast of the North Island, with a remnant distribution of approximately 140 km (Oremus et al. 2012) (Figure 1), and a total population size estimated in 2011 at just 55 dolphins (95% confidence limits [CL] = 48, 69) over the age of 1 year Hamner, Wade, et al. 2014), with a similar estimate produced in 2016: 63 (95% CL = 57, 75; Baker et al. 2016). Evidence that the population has been limited in size over long timescales includes the presence in all individuals of a single "diagnostic" mitochondrial DNA control region haplotype not found in Hector's dolphins (Pichler and Baker 2000; Hamner et al. 2012, 2017; Hamner, Constantine, et al. 2014). In comparison, the Hector's dolphin was estimated to number 7270 individuals in 2004 (95% CL = 5303, 9966; Slooten et al. 2004), distributed discontinuously around the South Island in at least 3 genetically differentiated regional populations (with occasional interchange between the regions): East Coast, West Coast, and South Coast (Hamner et al. 2012) (Figure 1). The Hector's and Māui dolphin subspecies appear to be genetically isolated, but some Hector's dolphins have recently been documented in the current range of the Māui dolphin, presumably as a result of long-distance dispersal (Hamner et al. 2012; Hamner, Constantine, et al. 2014).

Previously, Heimeier et al. (2009) described MHC sequence diversity and confirmed expression of single-copy DQA and DQB genes encoding for the classical MHC class II DQ molecule from 2 Hector's dolphin individuals. However, the analysis of DQA and DQB diversity is complicated by strong linkage disequilibrium (LD). This nonrandom association of alleles at different genetic loci is a feature of many highly polymorphic MHC loci, including MHC class II DQ. DQA and DQB code for polypeptides that form the functional MHC class II DQ heterodimer. This heterodimer is expressed on antigen-presenting cells as a cell-surface receptor, with the peptide binding cleft for detecting small peptides derived from pathogens formed by both the DQA and DQB peptides (Klein et al. 1990). Given the structural interdependence of the DQA and DQB peptides, it is perhaps not surprising that LD has been reported between the DQA and DQB loci in old world monkeys and humans (Trachtenberg et al. 1995; Gaur et al. 1998). Human DQA and DQB loci are located about 20 kb apart, with an assumed low recombination rate due to their physical proximity (Trowsdale et al. 1991). DQA and DQB are also in close physical proximity in the Yangtzee finless porpoise (Ruan et al. 2016), which shared a common ancestor with the Hector's and Māui dolphins approximately 15 mya (Alexander et al. 2013). This suggests that selection for specific combinations of DQA–DQB haplotypes, due to functional incompatibility of specific DQA and DQB polypeptide combinations (Kwok et al. 1995; Megiorni and Pizzuti 2012) and/or varying levels of protection or susceptibility against different pathogens (Ahmad et al. 2003), could also be occurring in dolphins. Therefore, investigating patterns in DQA–DQB haplotypes as well as considering each locus separately can yield a better understanding of selection and drift in wild cetaceans. In this study, we describe the geographic patterns of genetic diversity for DQA and DQB, and inferred DQA–DQB haplotypes, using samples collected from Hector's and Māui dolphins throughout the subspecies distributions. Our overall objectives were to investigate:

(i) LD between DQA and DQB in the context of a functional MHC class II peptide;
(ii) genetic differentiation of these functional markers between the Māui and Hector's dolphins in comparison with previously described patterns of neutral differentiation (Pichler 2002; Hamner et al. 2012);
(iii) the influence of selection and genetic drift in the Māui dolphin, based on comparison of MHC and microsatellite diversity with the more abundant Hector's dolphin; and
(iv) the influence of balancing selection on an evolutionary timeframe, as inferred by d_N/d_S, including d_N/d_S of specific regions that code for putative peptide-binding sites (PBS).

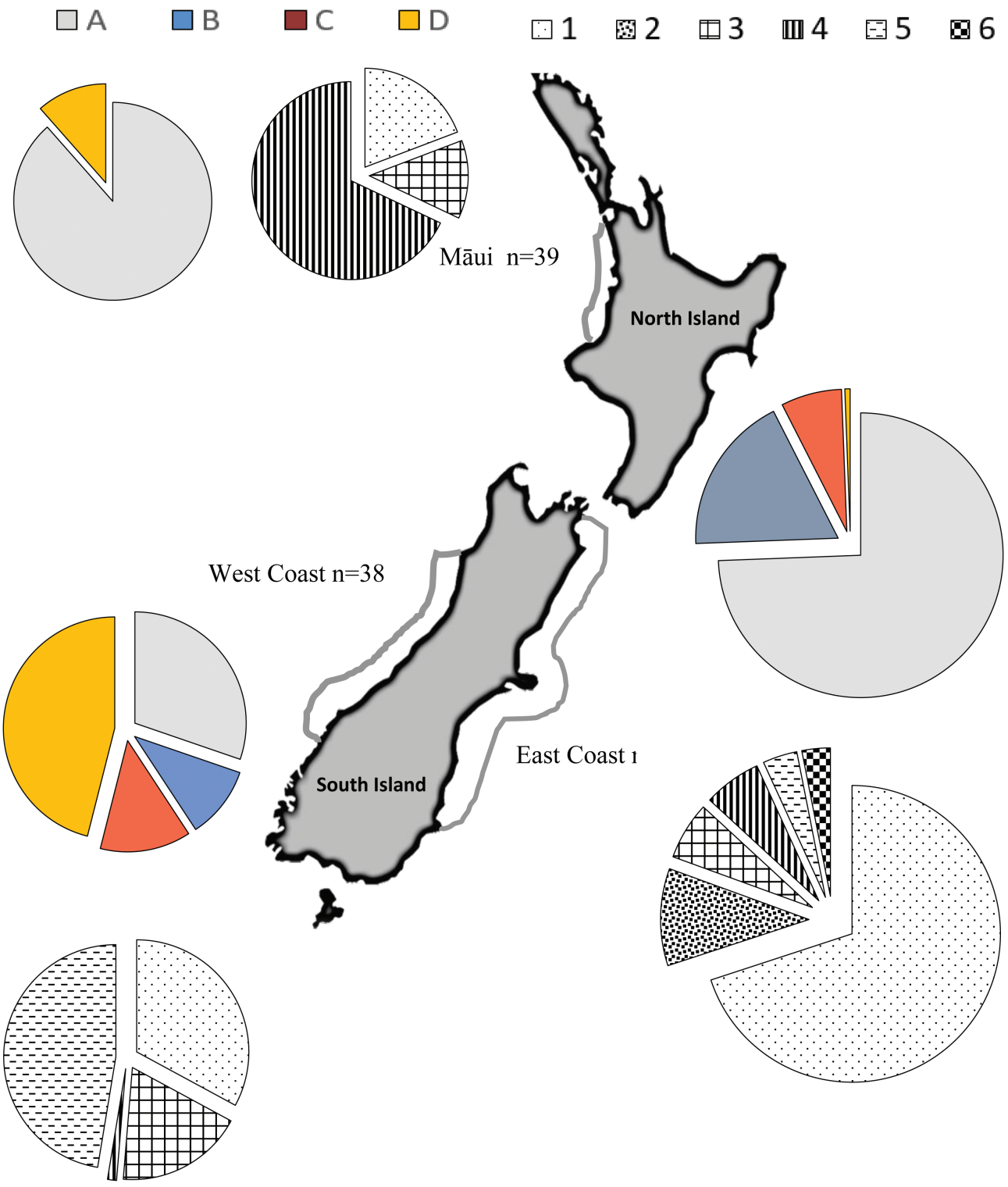

Figure 1. Distribution and frequencies of alleles of DQA and DQB for Māui and Hector's dolphin populations. The coastlines that samples were collected from are indicated by gray lines. Sample size for each population is given by *n*. DQA-DQB haplotypes are shown in Table 1.

Materials and Methods

Sample Collection and DNA Extraction

DNA was available from $n = 118$ Hector's and $n = 39$ Māui dolphin tissue samples held by the University of Auckland Cetacean Tissue Archive. Details of sample collection and DNA extraction are given in previous studies (Pichler et al. 2001; Hamner et al. 2012; Baker et al. 2013). Briefly, skin samples were collected from beachcast dolphins recovered by New Zealand Department of Conservation staff (Hector's $n = 46$, Māui $n = 4$) or from dolphins at sea (Hector's $n = 72$, Māui $n = 35$) using a small, stainless steel, biopsy dart fired from a modified veterinary capture rifle (Krützen et al. 2002). DNA

was extracted following standard phenolchloroform procedures (Sambrook et al. 1989), modified for small sample sizes (Baker et al. 2002). These samples were a subset of a larger data set utilized in previous studies (n = 233; Hamner et al. 2012; Baker et al. 2013) because samples were included in the present study only if they represented unique individuals (based on 12-locus microsatellite genotypes; Hamner et al. 2012; Baker et al. 2013) and were also successfully genotyped at DQA and DQB.

Microsatellite Genotypes

Microsatellite genotypes were available from Hamner et al. (2012) and Baker et al. (2013) for all of the individuals included in the survey of MHC diversity. The microsatellite data set utilized here encompassed 12 microsatellite loci (415/416; DlrFCB6; DlrFCB11; EV14; EV37; EV94; GT23; GT575; KWM9b; KWM12a; MK5; PPHO110). Most individuals were typed for all 12 microsatellite loci, with some missing 1 or 2 microsatellite genotypes.

Polymerase Chain Reaction Amplification of DQA and DQB

The exon 2 region of DQA (amplicon 218 bp) was amplified using the primers DQA1 (5′AGTACACCCAAGAATTTGATGG-3′; Auffray et al. 1987) and DQAex2R (5′-TCCACCATTCTGCCTCTCTT-3′; Heimeier et al. 2009) in a 20 µL volume with 1× Plat*Taq* buffer, 2.5 mM $MgCl_2$, 0.5 U Plat*Taq* (Applied Biosystems), 0.24 µM dNTPs (Amersham Pharmacia Biotech), 0.2 µM each primer, and approximately 10–20 ng of template DNA. The temperature profile was as follows: 94 °C for 3 min, followed by 30 cycles at 94 °C for 20 s, 58 °C for 15 s, 72 °C for 20 s, then followed by a final extension step at 72 °C for 10 min. Initial results for DQA single-stranded conformation polymorphism (SSCP) suggested the presence of null alleles using this primer combination because of significant deviation from Hardy–Weinberg expectations (HWE) for DQA (Heimeier 2009), which was the result of a 2-bp mismatch in the DQA1 primer binding site that did not amplify alleles DQA-C and D (Supplementary Figure S1a). Therefore, direct sequencing was used to confirm SSCP genotypes. Exon 2 of DQA was amplified with DQAex2F (5′-GCTGACCAYGTTGSCTCCTATG-3′; modified from Pfau et al. 1999) and DQA2 (5′-CCAGTGCTCCACCTTGCAGTC-3′; Auffray et al. 1987), which amplified a longer fragment (~900 bp) of exon 2 and exon 3 of DQA, including the intervening intron (Heimeier et al. 2009). The polymerase chain reaction (PCR) reagent conditions for this longer DQA fragment were the same as for the shorter DQA fragment. The temperature profile was as follows: 94 °C for 3 min, followed by 30 cycles at 94 °C for 30 s, 56 °C for 30 s, 72 °C for 60 s, followed by a final extension step at 72 °C for 10 min.

The exon-2 region of DQB (amplicon 172 bp) was amplified using the primers DQB2 (forward primer; 5′-CATGTGCTACTTCACCAACGG-3′) and DQB1 (reverse primer; 5′CTGGTAGTTGTGTCTGCACAC-3′) (Tsuji et al. 1992) in a 20 µL volume with 1× *Taq*Gold buffer, 2.5 mM $MgCl_2$, 0.5 U *Taq*Gold (Applied Biosystems), 0.24 µM dNTPs (Amersham Pharmacia Biotech), 0.2 µM each primer, and approximately 10–20 ng of template DNA. The temperature profile was as follows: 94 °C for 3 min, followed by 5 touchdown cycles at 94 °C for 15 s, annealing temperature (initially 64 °C, decreasing by 1 °C per cycle) for 20 s, 72 °C for 40 s, and then by another 30 cycles with 94 °C for 15 s, 52 °C for 20 s, 72 °C for 40 s, followed by a final extension step at 72 °C for 10 min.

Genotyping of MHC Loci with SSCP

SSCP was used to detect sequence variants from amplified fragments of DQA and DQB exon 2. For each SSCP screening run for each sample, 8–12 µL of PCR product was mixed with an equal volume of denaturing loading buffer, containing 80% formamide, 5% bromophenol blue, 5% xylene cyanol, and 0.1 M EDTA. The mixture was heated at 95 °C for 5 min and snap-cooled on ice, before being loaded onto a 12% nondenaturing polyacrylamide gel (37.5:1; 2.6% crosslinker ratio), and run in 1× Tris-borate-EDTA buffer for 1114 h at 4 °C using a Bio-Rad's vertical gel electrophoresis system (Bio-Rad, Hemel Hempstead, UK). To visualize the bands, we followed the protocol of Bassam et al. (1991) with slight modifications: the gel was fixed in 50% methanol (instead of 10% acetic acid) and stained in a silver solution to visualize allelic conformation. Rather than adding sodium thiosulfate directly to the developing step, the gel was rinsed with distilled water and soaked in 200 mg/L sodium thiosulfate for 1 min, which reduced background staining, followed by an additional rinsing step in distilled water.

At least 2 independent runs of each sample were performed with SSCP. On the first run, genotypes were scored for each sample from the band pattern. For the second run, samples were grouped according to genotype, reamplified and run again on an SSCP gel. Rather than comparing band patterns to reference samples with known genotypes, this procedure made it easier and more efficient to detect occasional genotyping errors (following Xu et al. 2007). Following the second screening run, a selection of samples for each of the unique band patterns observed were excised from the gel, destained, and reamplified by PCR and sequenced in both directions (as described in direct sequencing to confirm SSCP genotypes). We destained the gel by washing it in a 1:1 solution of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate for 5 min. Once the solution was clear, we then washed in 100 mM ammonium bicarbonate for 20 min.

In addition to Sanger sequencing a subset of samples for each unique band pattern, the longer DQA fragment described earlier, was directly sequenced for all samples that had an allelic configuration consistent with the presence of the null DQA allele (homozygous for A or B allele).

Direct Sanger Sequencing to Confirm SSCP Genotypes

Destained SSCP bands and amplicons were, after treatment with ExoSap-IT (Thermo Fisher), sequenced in both directions using the Big Dye™ terminator sequencing kit V3.1 on an ABI 3100 or 3130 automated capillary sequencer (Applied Biosystems). We aligned nucleotide sequences with the software programs Sequencher 4.2 (Genes Codes Corporation) and MacClade v4.0 (Maddison and Maddison 2000). The "call secondary peaks" option of the program Sequencher was set to a cutoff value of 30% to identify potential heterozygotes (Lento et al. 2003), and the quality of the sequence was reviewed using PHRED scores (Ewing et al. 1998). Only nucleotide sequences with an average PHRED score of 20 or greater were considered for analysis.

Allele Nomenclature

We followed the rules of Human Nomenclature for accepting and naming new MHC alleles (Klein et al. 1990; Marsh et al. 2010). In general, an individual allele was considered confirmed if an identical sequence was obtained from at least 2 independent amplifications from the same

individual or from different individuals. New alleles were confirmed to be MHC alleles by a BLAST search against the NCBI nucleotide database. New alleles characterized in this study have been submitted to GenBank under the following accession numbers: MH48167, MH48168, MH48169, MH48170, and MH48171 (previously characterized alleles also found in this study have the following accession numbers: EU024796, EU024797, EU024807, EU024808, and EU024809).

Phasing and Haplotype Inference

The DQA–DQB haplotypes for each individual within each population were reconstructed using Bayesian inference, as implemented in the program Phase 2.1, using the default values for number of iterations, thinning interval, and burn-in iterations (Stephens et al. 2001; Stephens and Scheet 2005). The most probable inferred haplotypes were used as input to test for the presence of LD (i.e., significant association between pairs of DQA and DQB alleles) in MIDAS (Multiallelic Interallelic Disequilibrium Analysis Software; Gaunt et al. 2006). The LD for each possible combination of alleles is presented as both the Pearson correlation coefficient, r^2, and the coefficient of LD, D'. r^2 ranges from 0 for complete linkage equilibrium to nearing 1 for complete LD. Its upper limit is a function of the allele frequencies at the 2 loci and can only reach 1 when they are equal. D' is standardized by the maximum possible value of D given allele frequencies (Lewontin 1964) and can reach the extremes of −1 or 1 when there is complete LD.

LD between MHC and Microsatellite Loci

Overall pairwise LD between all MHC microsatellite loci for each population was calculated in PyPop (Lancaster et al. 2007) using default parameters.

Phylogenetic Trees

Phylogenetic trees were constructed for all DQA and DQB alleles in Mega 6.0 (Tamura et al. 2013) with cow DQA and DQB sequences as outgroups. The Kimura 2-parameter substitution model was selected as the substitution model as this model has been widely used for MHC alleles in the literature. Other substitution models (i.e., Jukes–Cantor model) did result in similar phylogenetic trees (data not shown).

Population Diversity and Differentiation

Major Histocompatibility Complex

We estimated the observed and expected heterozygosities, and nucleotide diversity within each population for each locus, and tested for deviation from HWE using the exact test in Arlequin v3.5 (Excoffier and Lischer 2010). Nucleotide diversity was calculated with the Jukes–Cantor model in Mega 6 (Tamura et al. 2013). We assessed genetic differentiation between pairs of populations by calculating the F_{ST} index and its molecular analogue φ_{ST} (Excoffier et al. 1992) for each MHC locus and their combined haplotypes. G''_{ST} between populations was calculated with GenoDive ver2.0b23 (Meirmans and van Tienderen 2004).

We tested for geographic differences in MHC allelic and nucleotide diversity using a permutation procedure, genetic_diversity_diffs v1.0.1 (Alexander 2016; Alexander et al. 2016). This script randomly resamples individuals over all populations to test whether the observed differences in allelic/nucleotide diversity between specific populations are greater than expected by chance from samples that belong to a single underlying panmictic population. P values for the differences in diversity metrics between *C. hectori* dolphin populations were assessed using 10 000 permutations. Using divergent_allele_test v1.0.0 (Alexander 2015), we then examined how diverse the combination of unique alleles in each population was, by comparing them with combinations of alleles from the total metapopulation (e.g., the divergence of observed unique alleles compared with the divergence of all combinations of $[n_k]$ alleles, where n is total number of alleles observed in the metapopulation of the species and k is number of observed alleles in a given population). This analysis was conducted separately for DQA and DQB.

Microsatellites

We tested for deviations from HWE by estimating observed and expected heterozygosities for all 12 microsatellite loci using Arlequin v3.5. Genetic differentiation between pairs of populations was evaluated by calculating an averaged F_{ST} as well as G''_{ST} index for microsatellite loci that were in HWE and not under selection (see Supplementary Table S1 for included loci). DlrFCB6 locus was omitted by Arlequin for F_{ST} computation because more than 5% of the samples were missing genotypes at this locus.

Testing for Selection by Deviations from Neutrality

For MHC loci, 3 different tests of selection were utilized: 1) deviations from HWE were tested because an excess of heterozygotes could point toward heterozygote advantage at the individual level, 2) the F_{ST} outlier test detects directional or balancing selection in geographic subpopulations on an ecological timescale, and 3) the ratio of nonsynonymous to synonymous substitutions (d_N/d_S) can, besides detecting positive selection, also indicate balancing selection on an evolutionary timescale when d_N substitutions are found in excess.

HWE was calculated from observed and expected heterozygosity frequencies, as described earlier. A significant excess of observed heterozygotes would be observed from the P values.

The outlier test for selection incorporating DQA and DQB alleles, and DQA–DQB haplotypes and microsatellite loci was conducted on a pairwise population basis with the program BayeScan 2.1 (Foll and Gaggiotti 2008; Fischer et al. 2011). BayeScan is based on a logistic regression model that models selection by decomposing F_{ST} coefficients into a population-specific (beta) and a locus-specific (alpha) component. When the alpha component is necessary to explain the observed pattern of diversity, departure from neutrality is assumed. A positive value of alpha suggests diversifying selection, whereas negative values suggest balancing or purifying selection. Posterior probabilities contrasting selection (including the alpha component) versus no selection (alpha = 0) are calculated using a reversible-jump Markov chain Monte Carlo algorithm. BayeScan was run with default parameters. To account for the chance of multiple testing, we set the false discovery rate to 0.05 for our analyses.

The d_N/d_S ratio was computed for putative PBS and non-PBS within exon 2 of DQA and DQB, following published PBS definitions by Paliakasis et al. (1996) for DQA and Seddon and Ellegren (2002) for DQB. We calculated the d_N/d_S ratios using the modified Nei–Gojobori distance (Nei and Gojobori 1986) applying the Jukes–Cantor correction for multiple hits as implemented in MEGA 6.0 (Tamura et al. 2013). The significance of this ratio was computed using the Z-test of selection under the null hypothesis of $d_N = d_S$. Standard error estimates and P values were obtained by a bootstrap procedure of 500 replicates.

Results

Allelic Identity, Divergence, and Population Diversity

A total of 4 DQA and 6 DQB alleles were identified by direct sequencing and SSCP in the sample of 157 *C. hectori* individuals (Table 1; Supplementary Figures S2 and S3).

Table 1. Observed allele frequencies for DQA and DQB and for DQA–DQB haplotypes from each population

DQB / DQA	1	2	3	4	5	6	DQA total ↓ $k = 4$
(a) East Coast ($2n = 160$)							
A	0.68125	0.01250	0	0.03750	0.00625	0.03125	0.76875
B	0.01250	0.09375	0.06250	0	0.00625	0	0.17500
C	0.01250	0	0	0.03750	0.01875	0.00625	0.07500
D	0	0	0	0	0.00625	0	0.00625
DQB total → $k = 6$	0.70625	0.10625	0.06250	0.07500	0.03750	0.03750	
							DQA total ↓ $k = 4$
(b) West Coast ($2n = 76$)							
A	0.21053	0	0.05263	0	0.07895	0	0.34211
B	0.05263	0	0.03947	0	0	0	0.09211
C	0.03947	0	0.03947	0.01316	0.02632	0	0.11842
D	0.03947	0	0.05263	0.01316	0.35526	0	0.46053
DQB total → $k = 4$	0.34211	0	0.18421	0.02632	0.46053	0	
							DQA total ↓ $k = 2$
(c) Māui ($2n = 78$)							
A	0.17949	0	0.01282	0.67949	0	0	0.87179
B	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0
D	0.01282	0	0.11538	0	0	0	0.12821
DQB total → $k = 3$	0.19231	0	0.12821	0.67949	0	0	

$2n$ is total number of alleles within each population; k is number of alleles observed for each population.

Two DQA and 3 DQB alleles were previously found in Hector's dolphins (GenBank accession numbers: EU024796, EU024797, EU024807, EU024808, and EU024809) with the DQA alleles represented by longer previous sequences from mRNA covering exons 1–3 (Heimeier et al. 2009). The remaining alleles identified in this study have been submitted to GenBank with accession numbers: MH48167, MH48168, MH48169, MH48170, and MH48171. Of the 4 DQA and 6 DQB alleles, only *DQB*05* was identical to a previously reported allele from another species, the Pacific white-sided dolphin (*Lagenorhynchus obliquidens*; AB164225.1). All alleles translated to amino acid peptides without stop codons or frameshifts. All resulting inferred protein sequences were unique, except for DQA-C and D, which differed at only a single synonymous site (Supplementary Figure S1a, Supplementary Table S2a). No individual was found to have more than 2 alleles for either DQA or DQB, supporting the finding of previous work that these expressed loci exist as single copies within the *C. hectori* genome (Heimeier et al. 2009).

The East Coast Hector's dolphin was the only population that included all 4 DQA and all 6 DQB alleles (Table 1; Figure 1), and the only population with private alleles for DQB (*Cehe-DQB*02* and **06*). No population had private DQA alleles.

The Māui dolphin population had the fewest alleles for DQA and DQB with just 2 alleles for DQA and 3 alleles for DQB (Table 1). The West Coast Hector's dolphins were represented by all 4 alleles for DQA and 4 (of the total 6) for DQB. Except for a slightly higher expected ($H_{exp} = 0.211$) than observed ($H_{obs} = 0.180$) heterozygosity at DQA for the Māui dolphin ($P = 0.023$) (Supplementary Table S3), no significant differences were detected between observed and expected heterozygosity for either locus in any of the populations.

The mean nucleotide diversity for DQB was about 10 times higher than DQA within each population (Table 2). For DQA, all pairwise geographic comparisons of allelic and nucleotide diversity were significantly different (Table 2), with the West Coast Hector's dolphin showing the highest diversity and Māui dolphin showing the lowest. For DQB, the West Coast Hector's dolphin population had significantly higher nucleotide diversity than the East Coast and the Māui populations; however, the Māui population had the highest observed nucleotide diversity of all populations, significantly higher than the East Coast Hector's dolphin (Table 2).

We attempted to test for a "divergent allele" effect when a population had a lower number of DQA or DQB alleles than the metapopulation; however, the limited number of alleles in the metapopulation did not allow for a statistical test. For example, for DQB of the West Coast population, there are only 15 possible sets of 4 unique alleles that can be taken from the 6 alleles found across the total metapopulation (Table 3). Sixty percent of those possible sets have a higher average pairwise divergence than the observed divergence; 13% an equal and 20% a lower divergence than observed. The Māui dolphin population contained only 2 DQA alleles, but these were among the most divergent combinations over 2 of the 4 DQA alleles found in the metapopulation (Table 3; Supplementary Figure S2a) and also supported by the phylogenetic tree (Supplementary Figure S3a). The set of DQB alleles within the West Coast and Māui populations were not among the most divergent combination of alleles. Instead, the high DQB nucleotide diversity presents in the Māui population is driven by the high frequency of allele *DQB*04* within this population, which showed an average pairwise difference with other DQB alleles of 0.091. In comparison, the next most divergent allele (at high frequency in the West

Table 2. Genetic diversity indices calculated with Arlequin and significance of pairwise comparisons of genetic diversity calculated for DQA and DQB in *Cephalorhynchus hectori* populations (genetic_diversity_diffs v1.0.10)

	DQA			DQB		
	East Coast	West Coast	Māui	East Coast	West Coast	Māui
(a) Expected heterozygosity DQA and DQB						
East Coast	0.412	0.0015	0.0118	0.492	0.0031	0.9947
West Coast		0.677	<0.00001		0.642	0.0086
Māui			0.211			0.491
(b) Nucleotide diversity (%) DQA (228 bp) and DQB (170 bp)						
East Coast	0.798	0.0006	0.0347	2.925	0.1832	0.0020
West Coast		1.431	<0.00001		3.587	0.1098
Māui			0.554			4.561

For (a) expected heterozygosity and (b) nucleotide diversity (%). *P* values are shown above diagonal.

Table 3. Comparison between the observed mean pairwise divergences across the *k* unique alleles within each population, compared with divergence across different sets of *k* unique alleles from alleles available across the metapopulation

Locus (number of alleles in metapopulation)	Population	Number of observed alleles (*k*)	Mean observed pairwise divergence between *k*	Number of possible combinations of *k* alleles drawn from metapopulation alleles	Proportions of allelic combinations where: average pairwise divergence		
					> Obs divergence	== Obs divergence	< Obs divergence
DQA (4)	East Coast	4	0.020	1	NA		
	West Coast	4	0.020	1	NA		
	Māui	2	0.026	6	0	0.6	0.6
DQB (6)	East Coast	6	0.077	1	NA		
	West Coast	4	0.075	15	0.6	0.13	0.27
	Māui	3	0.078	20	0.5	0.05	0.45

Analysis conducted separately for DQA and DQB. NA, not applicable.

Coast population), *DQB*06*, had an average pairwise difference to other DQB alleles of just 0.076 (Supplementary Figure S2b).

Haplotypes and LD

We found evidence of significant LD between DQA and DQB alleles within each population (East Coast Hector's [EC]/Māui $P < 0.001$; West Coast [WC] $P < 0.015$) (Table 4). No microsatellite locus was in LD with either MHC locus in all populations (Supplementary Table S4). In each population, a different haplotype was most frequent, and all 3 of these haplotypes were observed significantly more often than expected based on DQA and DQB allele frequencies (EC: 67% haplotype A1, $D' = 0.89$, $P < 0.001$; WC: 35% haplotype D5, $D' = 0.54$, $P < 0.001$; Māui: 68% haplotype A4, $D' = 1$, $P < 0.001$). This meant that other haplotypes were less frequently observed than expected, such as B1 and A3 in the East Coast Hector's dolphin, A5 and D1 in the West Coast Hector's dolphin, and A3 and D4 in the Māui dolphin. The most frequent haplotype of the East Coast Hector's dolphin, A1 at about 67%, was the only haplotype that was present in all 3 populations and was also found as the second most frequent haplotype in the other 2 populations (WC: 21%; Māui: 18%) (Table 4). Despite having 2 fewer DQB alleles, the West Coast population exhibited similar numbers of DQA–DQB haplotypes as the East Coast population (WC: 12 haplotypes out of a potential for 16 based on allele frequencies, EC: 13 of 24). The Māui dolphin had 5 haplotypes from the 6 possible based on allele frequencies; however, 2 of those were only observed once (Table 4).

Geographic Differentiation and Outlier Test for Selection

DQA and DQB single-locus allele frequencies varied significantly among all populations (F_{ST}: 0.031–0.398, φ_{ST}: 0.06–0.49, $P < 0.05$; Table 5). The differentiation between populations was greater than 0.2 for all comparisons involving MHC loci, except for the East Coast Hector's to Māui dolphin comparison, where differentiation for DQA ($F_{ST} = 0.031$; $\varphi_{ST} = 0.06$) was much lower than that for DQB ($F_{ST} = 0.398$; $\varphi_{ST} = 0.49$; Table 5). Given the significant LD and different prominent haplotypes in populations, we also conducted a test of differentiation for the MHC haplotypes. The differentiation for haplotype frequencies between populations (F_{ST}: 0.213–0.391, $P < 0.001$; Table 5) was comparable to differentiation seen when DQA (except DQA East Coast to Māui) and DQB were analyzed independently (F_{ST}: 0.247–0.398, $P < 0.001$; Table 5). Among DQA, DQB, and DQA–DQB haplotypes, the largest differentiation occurred between East Coast Hector's and Māui dolphins at DQB.

Although microsatellite and MHC loci showed similar levels of F_{ST} differentiation for the West Coast to Māui comparison ($F_{ST} > 0.2$), surprisingly in the comparison of the East Coast to the Māui dolphin, DQA showed a lower level of differentiation than the other markers/haplotypes analyzed in this study (Table 5). In contrast, the microsatellite differentiation between the East Coast and West Coast Hector's populations was much weaker than that observed for most MHC comparisons (e.g., DQA–DQB haplotype $F_{ST} = 0.213$; microsatellite $F_{ST} = 0.027$; Table 5). This estimate did not include microsatellite loci out of HWE (Supplementary Table S6). The G''_{ST} index accounting for differing levels of heterozygosity showed similar levels of differentiation and showed the same pattern as F_{ST} (Supplementary Table S5). All pairwise estimates of population differentiation based on microsatellites were similar to values calculated from a larger data set and 13 microsatellite loci (Supplementary Table S7).

The outlier analyses provided evidence of balancing selection for the DQA–DQB haplotypes between East Coast and West Coast

Table 4. Inferred (through the program phase) observed "obs" and expected numbers "exp" (HWE assuming no LD) of DQA–DQB haplotypes for each population

DQA–DQB	East Coast ($2n$ = 160)				West Coast ($2n$ = 76)				Māui ($2n$ = 78)			
	Overall D' = 0.84 (P < 0.0001)				Overall D' = 0.51 (P < 0.015)				Overall D' = 1 (P < 0.001)			
	obs	exp	D'	r^2	obs	exp	D'	r^2	obs	exp	D'	r^2
A 1	107	83	0.89 (<0.001)	0.63	16	8	0.54 (<0.001)	0.26	14	13	0.46 (0.897)	0.01
A 2	2	13	−0.81 (<0.001)	0.22								
A 3	0	7	−1 (<0.001)	0.19	4	4	−0.16 (0.15)	0	1	9	−1 (<0.001)	0.89
A 4	4	7	−0.20 (0.429)	0.01	0	0			53	47	1 (<0.001)	0.28
A 5	1	4	−0.73 (<0.001)	0.06	6	11	−0.46 (0.003)	0.08				
A 6	5	4	0.73 (0.182)	0.01								
B 1	3	20	−0.88 (<0.001)	0.40	4	3	0.34 (0.276)	0.03				
B 2	15	3	0.84 (<0.001)	0.38								
B 3	10	2	1 (<0.001)	0.30	3	1	0.24 (0.015)	0.03				
B 4	0	2	−1 (0.277)	0.01	0	0						
B 5	1	1	−0.08 (0.925)	0	0	4	−1 (0.005)	0.11				
B 6	0	1	−1 (0.285)	0.01								
C 1	2	8	−0.8 (<0.001)	0.11	3	3	−0.19 (0.834)	0				
C 2	0	1	−1 (0.236)	0.01								
C 3	0	1	−1 (0.375)	0	3	2	0.1 (0.311)	0.01				
C 4	6	1	0.55 (<0.001)	0.28	1	0	1 (0.010)	0.09				
C 5	3	0	0.43 (<0.001)	0.10	2	5	−0.49 (0.238)	0.03				
C 6	0	0										
D 1	0	1	−1 (0.125)	0.01	3	12	−0.74 (<0.001)	0.23	1	2	−0.46 (0.897)	0.01
D 2	0	0										
D 3	0	0			4	6	−0.38 (0.390)	0.03	9	1	1 (<0.001)	0.89
D 4	0	0			1	0	1 (0.352)	0.02	0	6	−1 (<0.001)	0.28
D 5	1	0	1 (<0.001)	0.16	27	17	0.55 (<0.001)	0.29				
D 6	0	0										

P values for D' are given in parentheses.

Table 5. Pairwise F_{ST} values for DQA, DQB, DQA–DQB haplotype, and microsatellites between populations of Hector's and Māui dolphins (φ_{ST} also provided for DQA and DQB)

Pop1 (n)–Pop2 (n)	F_{ST} DQA	φ_{ST} DQA	F_{ST} DQB	φ_{ST} DQB	F_{ST} DQA–DQB haplotype	Mean F_{ST} microsatellite
East Coast (70)–West Coast (39)	0.294 (<0.001)	0.38 (<0.001)	0.247 (<0.001)	0.24 (<0.001)	0.213 (<0.001)	0.027 (<0.001)
East Coast (70)–Māui (38)	0.031 (<0.027)	0.06 (<0.001)	0.398 (<0.001)	0.49 (<0.001)	0.391 (<0.001)	0.188 (<0.001)
West Coast (39)–Māui (38)	0.377 (<0.001)	0.40 (<0.001)	0.374 (<0.001)	0.46 (<0.001)	0.311 (<0.001)	0.227 (<0.001)

P values are given in parentheses.

populations, between East Coast and Māui dolphins, as well as between West Coast and Māui dolphins (Figure 2; Supplementary Figure S4).

The significant negative alpha value indicated balancing selection for the DQA–DQB haplotype for one or even both of the populations. Neither DQA nor DQB showed significant alpha values when investigated as separate loci.

Influence of Balancing Selection

The putative PBS for both the DQA and DQB loci showed a d_N/d_S ratio significantly greater than 1 in each of the 3 populations (Table 6). At non-PBS, most d_N/d_S ratios were not significantly greater than 1, with the exception of DQB for the West Coast Hector's dolphin population.

Discussion

We have shown strong geographic structure at the MHC (DQA and DQB) between the Māui and Hector's subspecies, as well as 2 regional populations of Hector's dolphin that inhabit opposite coasts of the South Island of New Zealand. We have also shown strong LD between individual alleles of DQA and DQB loci and the prevalence of differing haplotypes between populations, despite sharing the same MHC alleles, and limited differentiation at neutral markers between some pairs of populations. To the best of our knowledge, this is the first characterization of LD between functional loci of a cetacean. The analysis of functional haplotypes indicated balancing selection within the East Coast population, suggestive of a recent selective event that potentially increased the frequency of the prevalent A1 haplotype in this population. These patterns were not

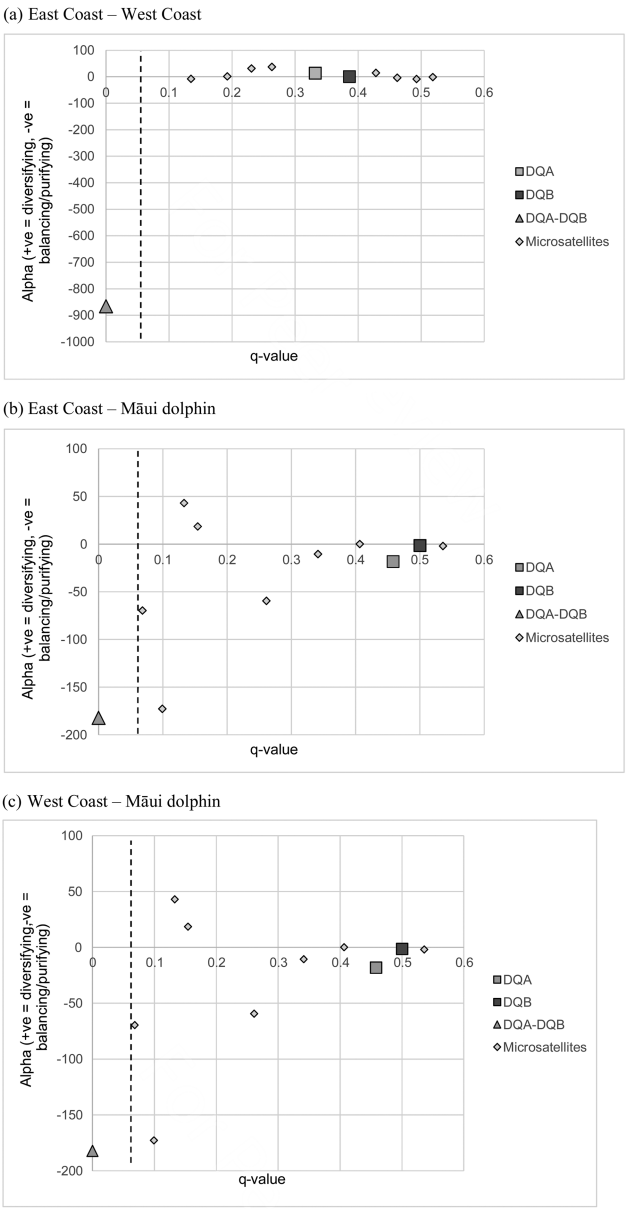

Figure 2. F_{ST} outlier analysis of *Cephalorhynchus hectori* dolphin populations East Coast versus West Coast (a), East Coast versus Māui dolphin (b), and West Coast versus Māui dolphin (c) excluding microsatellite loci out of HWE DlrFCB11, EV14, EV37, and EV94. The alpha value has been plotted against the *q*-value, which is the false discovery rate (FDR). FDR = 0.05 shown by the dashed line.

detected from MHC loci analyzed separately, demonstrating we can achieve a much deeper understanding of selection processes in relatively small populations when we consider functional haplotypes.

Selection and Genetic Drift in the Māui Dolphin

The Māui dolphin population numbers less than 100 individuals (Hamner, Wade, et al. 2014) and has probably undergone a population decline since the 1960s (Slooten et al. 2004). This subspecies is also reproductively isolated from the South Island Hector's dolphin populations (Hamner, Constantine, et al. 2014). On this basis, we would expect to see low overall levels of MHC diversity in the Māui dolphin population, consistent with the low genetic diversity observed at neutral markers such as microsatellites (H_{obs} = 0.427), and the mtDNA control region (only a single fixed haplotype; Baker et al. 2002; Hamner et al. 2012; Hamner, Constantine, et al. 2014; Hamner, Wade, et al. 2014). In this study, we did indeed find that the Māui dolphin population had fewer DQA and DQB alleles, and DQA–DQB haplotypes, in comparison with the Hector's dolphin populations. Although none of the DQA alleles retained in the Māui population were especially divergent in comparison to other alleles in the metapopulation, the *combination* of the 2 alleles were among the most divergent possible from the total alleles found in the metapopulation. Surprisingly, the Māui dolphin showed higher DQB nucleotide diversity than either of the Hector's dolphin populations (Māui π = 4.6%; East Coast π = 2.9%; and West Coast π = 3.6%), even though it retained only 3 of the 6 alleles found in the total metapopulation. The nucleotide diversity of the Māui dolphin DQB was also high compared with other cetacean species with larger population sizes such as the beluga (Murray et al. 1995) and finless porpoise (Hayashi et al. 2006). The loss of alleles but the preservation of nucleotide diversity could be explained by a "divergent allele advantage," where combinations of divergent alleles are selected for resistance to a wider range of potential pathogens (Potts and Wakeland 1990; Jeffrey and Bangham 2000). Highly divergent alleles have been found to be retained at the related class II DRB gene in other mammal populations affected by severe bottlenecks, such as red wolves (*Canis rufus*) (Hedrick et al. 2002), Arabian oryx (*Oryx leucoryx*) (Hedrick et al. 2000), and European bison (*Bison bonasus*) (Radwan et al. 2007). We found strong evidence for balancing selection from the outlier test, on an evolutionary timescale, and the retention of divergent alleles in Māui provides some additional evidence for relatively recent balancing selection. Furthermore, haplotypes A4 and D3 were observed significantly more than expected. Taken together, these findings could suggest the role of some form of balancing selection, in maintaining specific DQA and DQB alleles, as well as specific haplotypes. This can be observed despite the pressures of severe genetic drift in the critically endangered Māui dolphin, similar to the process of maintaining MHC variation in other severely bottlenecked mammalian species.

Balancing Selection on the South Island's East and West Coast

Differentiation of MHC loci between the Māui and Hector's dolphin subspecies was expected based on their reproductive isolation (also indicated by differentiation at neutral microsatellites and a fixed difference in the mtDNA control region). However, the strong differentiation between the East and West Coast Hector's dolphin populations was unexpected under the assumption of balancing selection in response to presumably similar oceanographic/environmental conditions for the 2 populations (Muirhead 2001), as well as gene flow between these regions (Hamner et al. 2012). Although we do not discount the influence of drift in these geographically, and to at least some degree reproductively, isolated populations, the outlier test suggests that balancing selection is contributing to the pattern of differentiation. Balancing selection can maintain genetic polymorphism by mainly 3 not mutually exclusive mechanisms, which are heterozygote advantage, frequency-dependent selection, and selection that varies in time and space. Spatially variable selection at MHC loci has also been shown in populations of various fish species, for example, Atlantic salmon (Landry and Bernatchez 2001), sockeye salmon at small geographic scales (Miller et al. 2001), and Californian steelhead (Aguilar and Garza 2006); in birds, for example, the great snipe (Ekblom et al.

Table 6. Proportion of nucleotide differences (Nei–Gojobori correction: nucl diff), proportion of amino acid differences (Aa diff), and estimated nonsynonymous (d_N) and synonymous (dS) substitutions per site for antigen PBS and non-PBS for the Hector's East Coast, West Coast, and Māui dolphin populations

		DQA						DQB				
		nucl diff	Aa diff	d_N	d_S	d_N/d_S	*P* value	nucl dif	Aa dif	d_N	d_S	d_N/d_S	*P* value
East Coast	PBS	0.025 (0.012)	0.077 (0.036)	0.034 (0.015)	0	∞	0.025	0.088 (0.025)	0.233 (0.076)	0.124 (0.041)	0.001 (0.001)	124	<0.001
	Non-PBS	0.008 (0.003)	0.009 (0.007)	0.004 (0.003)	0	∞	ns	0.013 (0.005)	0.028 (0.012)	0.015 (0.007)	0.008 (0.009)	1.88	ns
West Coast	PBS	0.035 (0.017)	0.110 (0.052)	0.049 (0.022)	0	∞	0.014	0.117 (0.045)	0.331 (0.128)	0.162 (0.041)	0.001 (0.001)	162	0.001
	Non-PBS	0.014 (0.005)	0.016 (0.011)	0.008 (0.006)	0.010 (0.009)	0.8	ns	0.014 (0.007)	0.042 (0.021)	0.020 (0.009)	0.001 (0.001)	20	0.041
Māui	PBS	0.012 (0.007)	0.038 (0.022)	0.017 (0.009)	0.000 (0.000)	∞	0.044	0.134 (0.050)	0.323 (0.119)	0.185 (0.082)	0.015 (0.015)	12.3	0.001
	Non-PBS	0.006 (0.002)	0.007 (0.005)	0.003 (0.003)	0.004 (0.004)	0.75	ns	0.023 (0.009)	0.048 (0.021)	0.028 (0.013)	0.012 (0.013)	2.33	ns

Standard error is given in parentheses. ns, not significant.

2007); and also in mammals, for example, black rat populations on Madagascar that differ in their resistance to plague (Tollenaere et al. 2012). This type of selection can shape MHC allele frequencies during contemporary timeframes, as has been shown for water voles (Bryja et al. 2007). Further evidence for differing selection regimes between the East Coast and West Coast population comes from the detection of balancing selection for the functional DQA–DQB haplotype based on the outlier test, which potentially led to the high frequency of the A1 haplotype in the East Coast population. Ninety-one percent of individuals have at least one copy of the A1 haplotype in the East Coast, whereas in the West Coast population, only 63% of individuals have at least one copy of the most common West Coast haplotype D5. This finding might indicate that the A1 haplotype has been important for survival in the East Coast population in recent decades and could be an indication for frequency-dependent selection, where a rare allele has a selective advantage and increases in frequency over time until it reaches an intermediate frequency at which it will be maintained. The initial stages between frequency-dependent and directional selection are difficult to distinguish. In the case of directional selection, the increase in allele frequency would continue until the allele becomes fixed. In addition, balancing selection is supported by the d_N/d_S test, which can detect long-term balancing selection that has to be consistent over many generations to leave a significant signal. In contrast, the outlier test assesses selection over shorter timeframes based on allele frequency changes from generation to generation. Therefore, short bursts of directional selection of a specific allele or haplotype are not in disagreement with balancing selection over longer evolutionary timescales.

The underlying selective forces influencing MHC loci in the Hector's and Māui dolphin are unclear. Hector's and Māui dolphins are known to be susceptible to infectious diseases, such as toxoplasmosis and infection with *Brucella* sp. (Roe et al. 2013, 2017), but the distribution of these pathogens or the geographic pattern of mortality has not been studied. A study by van Wormer et al. (2016) showed that increased marine mammal *Toxoplasma gondii* infection aligned closely with coastal regions with highest levels of oocyst runoff, linked to higher levels of coastal development and larger domestic cat populations. At present, no studies exist to quantify overall differences of coastal pathogen diversity (or describe the general occurrence of pathogens for that matter) on the 2 opposite coastlines of the South Island of New Zealand. It is known, however, that the coasts differ in human population densities (higher on the East Coast), intensity of agricultural activities (higher on the East Coast), and freshwater runoff (higher on the West Coast; de Sherbinin et al. 2007; http://www.fao.org/countryprofiles/index/en/?iso3=NZL).

All of these factors could influence the diversity of pathogen communities along the coastlines (Yan et al. 2016), creating contrasting pathogen distributions and/or abundance levels between the East and West Coasts. However, more research is needed to investigate pathogen distributions and the potential for pathogen-mediated fluctuating selection in cetacean populations.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

This work was supported by the New Zealand Department of Conservation (DoC), Auckland Conservancy, the New Zealand

Worldwide Fund for Nature (WWF), and a grant to C.S.B. from the Marsden Fund of the New Zealand Royal Society.

Acknowledgments

This work was made possible by access to the University of Auckland Cetacean Tissue Archive, maintained and curated by C.S.B. and R. Constantine. Biopsy samples were collected under permit to C.S.B. from the New Zealand Department of Conservation (DoC) and animal ethics protocols AEC/022002/R9 and AEC/02/2005/R334 from the University of Auckland. The collection of beachcast samples was conducted by staff from the Department of Conservation. We thank K. Russell and M. Krützen for assistance with collection of biopsy samples and P. Brown and the Auckland Conservancy for support of sampling of Māui dolphins. Special thanks to R. Constantine for supporting resources essential for writing this manuscript.

Data Availability

MHC alleles are available on GenBank. Individual microsatellite, DQA, DQB, and DQA–DQB phased haplotype data are available at Dryad doi:10.5061/dryad.42nk23k

References

Aguilar A, Garza JC. 2006. A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). *Mol Ecol*. 15:923–937.

Aguilar, A, Roemer, G, Debenham, S, Binns, M, Garcelon, D, Wayne RK. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc Natl Acad Sci USA*. 101:3490–3494.

Ahmad T, Neville M, Marshall SE, Armuzzi A, Mulcahy-Hawes K, Crawshaw J, Sato H, Ling KL, Barnardo M, Goldthorpe S, *et al.* 2003. Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. *Hum Mol Genet*. 12:647–656.

Alexander AM. 2015. *Divergent_allele_test v1.0.0*. Available from: https://github.com/laninsky/divergent_allele_test

Alexander AM. 2016. *Genetic_diversity_diffs v1.0.0*. Available from https://github.com/laninsky/genetic_diversity_diffs

Alexander A, Steel D, Hoekzema K, Mesnick SL, Engelhaupt D, Kerr I, Payne R, Baker CS. 2016. What influences the worldwide genetic structure of sperm whales (*Physeter macrocephalus*)? *Mol Ecol*. 25:2754–2772.

Alexander A, Steel D, Slikas B, Hoekzema K, Carraher C, Parks M, Cronn R, Baker CS. 2013. Low diversity in the mitogenome of sperm whales revealed by next-generation sequencing. *Genome Biol Evol*. 5: 113–129.

Auffray A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK. 1987. Structure and expression of HLA-DQ alpha and -DX alpha genes: interallelic splicing of the HLA-DQ gene and functional splicing of the HLA-DQ alpha gene using a retroviral vector. *Immunogenetics*. 26:63–73.

Baker CS, Hamner RM, Cooke J, Heimeier D, Vant M, Steel D, Constantine R. 2013. Low abundance and probable decline of the critically endangered Maui's dolphin estimated by genotype capture-recapture. *Anim Conserv*. 16:224–233.

Baker A, Smith A, Pichler F. 2002. Geographical variation in Hector's dolphin: recognition of new subspecies of *Cephalorhynchus hectori*. *J R Soc N Z*. 32:713–727.

Baker CS, Steel D, Hamner RM, Hickman G, Boren L, Arlidge W, Constantine R. 2016. *Estimating the abundance and effective population size of Maui dolphins using microsatellite genotypes in 2015–16, with retrospective matching to 2001–16*. Auckland (New Zealand): Department of Conservation.

Bassam BJ, Caetano-Anollés G, Gresshoff PM. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal Biochem*. 196:80–83.

Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol*. 16:363–377.

Bowen L, Aldridge BM, Gulland F, Woo J, Van Bonn W, DeLong R, Stott JL, Johnson ML. 2002. Molecular characterization of expressed DQA and DQB genes in the California sea lion (*Zalophus californianus*). *Immunogenetics*. 54:332–347.

Bryja J, Charbonnel N, Berthier K, Galan M, Cosson JF. 2007. Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles. *Mol Ecol*. 16:5084–5097.

de Sherbinin A, Carr D, Cassels S, Jiang L. 2007. Population and environment. *Annu Rev Environ Resour*. 32:345–373.

Excoffier L, Lischer HEL. 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.

Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, Grahn M, Kålås JA, Höglund J. 2007. Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Mol Ecol*. 16:1439–1451.

Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*. 8:175–185.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 131:479–491.

Fischer MC, Foll M, Excoffier L, Heckel G. 2011. Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). *Mol Ecol*. 20:1450–1462.

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.

Gaunt TR, Rodriguez S, Zapata C, Day IN. 2006. MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. *BMC Bioinf*. 7:227.

Gaur LK, Pandarpurkar M, Anderson J. 1998. DQA-DQB linkage in Old World monkeys. *Tissue Antigens*. 51:367–373.

Gutierez-Espeleta GA, Kalinowski ST, Boyce WM, Hedrick PW. 2000. Genetic variation and population structure in desert bighorn sheep: implications for conservation. *Conserv Genetics*. 1:3–15.

Hamner RM, Constantine R, Mattlin R, Waples R, Baker CS. 2017. Genotype based estimates of local abundance and effective population size for Hector's dolphins. *Biol Conserv*. 211:150–160.

Hamner RM, Constantine R, Oremus M, Stanley M, Brown P, Baker CS. 2014. Long-range movement by Hector's dolphins provides potential genetic enhancement for critically endangered Maui's dolphin. *Mar Mamm Sci*. 30:139–153.

Hamner RM, Pichler FB, Heimeier D, Constantine R, Baker CS. 2012. Genetic differentiation and limited gene flow among fragmented populations of New Zealand endemic Hector's and Maui's dolphins. *Conserv Genetics*. 13:987–1002.

Hamner RM, Wade P, Oremus M, Stanley M, Brown P, Constantine R, Baker CS. 2014. Critically low abundance and limits to human-related mortality for the Maui's dolphin. *Endang Species Res*. 26:87–92.

Hayashi K, Yoshida H, Nishida S, Goto M, Pastene LA, Kanda N, Baba Y, Koike H. 2006. Genetic variation of the MHC DQB locus in the finless porpoise (*Neophocaena phocaenoides*). *Zool Sci*. 23:147–153.

Hedrick PW. 1994. Evolutionary genetics of the major histocompatibility complex. *Am Nat*. 143:945–964.

Hedrick PW. 1998. Balancing selection and MHC. *Genetica*. 104:207–214.

Hedrick PW. 2003. The major histocompatibility complex (MHC) in declining populations: an example of adaptive variation. In: Holt WV, Pickard AR, Rodger JC, Wildt DE, editors. *Reproduction science and integrated conservation*. Cambridge (UK): Cambridge University Press. p. 97–113.

Hedrick PW. 2004. Evolutionary genomics: foxy MHC selection story. *Heredity (Edinb)*. 93:237–238.

Hedrick PW, Lee RN, Garrigan D. 2002. Major histocompatibility complex variation in red wolves: evidence for common ancestry with coyotes and balancing selection. *Mol Ecol*. 11:1905–1913.

Hedrick PW, Lee RN, Parker KM. 2000. Major histocompatibility complex (MHC) variation in the endangered Mexican wolf and related canids. *Heredity (Edinb)*. 85:617–624.

Heimeier D. 2009. *Comparative diversity at the major histocompatibility complex in two dolphin species*. *Society*. Auckland (New Zealand): University of Auckland.

Heimeier D, Baker CS, Russell K, Duignan PJ, Hutt A, Stone GS. 2009. Confirmed expression of MHC class I and class II genes in the New Zealand endemic Hector's dolphin (*Cephalorhynchus hectori*). *Mar Mamm Sci*. 25:68–90.

Hughes AL, Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA*. 86:958–962.

Hughes AL, Yeager M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. *Annu Rev Genet*. 32:415–435.

Jeffrey KJ, Bangham RM. 2000. Do infectious diseases drive MHC diversity? *Microbes Infect*. 2:1335–1341.

Kimura M. 1983. *The neutral theory of molecular evolution*. Cambridge (UK): Cambridge University Press.

Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK, Watkins DI. 1990. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics*. 31:217–219.

Krützen M, Barré LM, Möller LM, Heithaus MR, Simms C, Sherwin WB. 2002. A biopsy system for small cetaceans: darting success and wound healing in *Tursiops* spp. *Mar Mamm Sci*. 18:863–878.

Kwok WW, Nepom GT, Raymond FC. 1995. HLA-DQ polymorphisms are highly selective for peptide binding interactions. *J. Immunol*. 155:2468–2476.

Lancaster AK, Single RM, Solberg OD, Nelson MP, Thomson G. 2007. PyPop update—a software pipeline for large-scale multilocus population genomics. *Tissue Antigens*. 69(Suppl 1):192–197.

Landry C, Bernatchez L. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Mol Ecol*. 10:2525–2539.

Lento GM, Baker CS, David V, Yuhki N, Gales NJ, O'Brien SJ. 2003. Automated single-strand conformation polymorphism reveals low diversity of a major histocompatibility complex class II gene in the threatened New Zealand sea lion. *Mol Ecol Notes*. 3:346–349.

Lewontin RC. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics*. 49:49–67.

Maddison DR, Maddison WP. 2000. *MacClade 4: analysis of phylogeny and character evolution*. Sunderland (MA): Sinauer Associates.

Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Fernández-Viña M, Geraghty DE, Holdsworth R, Hurley CK. 2010. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*. 75:291–455.

Megiorni F, Pizzuti A. 2012. HLA-DQA1 and HLA-DQB1 in celiac disease predisposition: practical implications of the HLA molecular typing. *J Biomed Sci*. 19:88.

Meirmans PG, van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes*. 4:792–794.

Miller KM, Kaukinen KH, Beacham TD, Withler RE. 2001. Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. *Genetica*. 111:237–257.

Muirhead CA. 2001. Consequences of population structure on genes under balancing selection. *Evolution*. 55:1532–1541.

Murray BW, Malik S, White BN. 1995. Sequence variation at the major histocompatibility complex locus DQ beta in beluga whales (*Delphinapterus leucas*). *Mol Biol Evol*. 12:582–593.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 3:418–426.

Oremus M, Hamner R, Stanley M, Brown P, Baker CS, Constantine R. 2012. Distribution, group characteristics and movements of the critically endangered Maui's dolphin *Cephalorhynchus hectori maui*. *Endang Species Res*. 19:1–10.

Paliakasis K, Routsias J, Petratos K, Ouzounis C, Kokkinidis M, Papadopoulos GK. 1996. Novel structural features of the human histocompatibility molecules HLA-DQ as revealed by modeling based on the published structure of the related molecule HLA-DR. *J Struct Biol*. 117:145–163.

Paterson S, Wilson K, Pemberton JM. 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries*). *Proc Natl Acad Sci USA*. 95:3714–3719.

Pfau RS, Van Den Bussche RA, McBee K, Lochmiller RL. 1999. Allelic diversity at the Mhc-DQA locus in cotton rats (*Sigmodon hispidus*) and a comparison of DQA sequences within the family Muridae (Mammalia: Rodentia). *Immunogenetics*. 49:886–893.

Pichler FB. 2002. *Genetic assessment of population boundaries and gene exchange in Hector's dolphin*. DOC Science Internal Series 44. Wellington (New Zealand): Department of Conservation.

Pichler FB, Baker CS. 2000. Loss of genetic diversity in the endemic Hector's dolphin due to fisheries-related mortality. *Proc Biol Sci*. 267:97–102.

Pichler FB, Robineau D, Goodall RN, Meÿer MA, Olivarría C, Baker CS. 2001. Origin and radiation of Southern Hemisphere coastal dolphins (genus *Cephalorhynchus*). *Mol Ecol*. 10:2215–2223.

Potts WK, Wakeland EK. 1990. The maintenance of MHC polymorphism. *Immunol Today*. 11:1989–1990.

Radwan J, Kawałko A, Wójcik JM, Babik W. 2007. MHC-DRB3 variation in a free-living population of the European bison, *Bison bonasus*. *Mol Ecol*. 16:531–540.

Reeves RR, Dawson SM, Jefferson TA, Karczmarski L, Laidre K, O'Corry-Crowe G, Rojas-Bracho L, Secchi ER, Slooten E, Smith BD, Wang JY, Zhou K. 2013. *Cephalorhynchus hectori. The IUCN Red List of Threatened Species*. 2013: e.T4162A44199757. doi:10.2305/IUCN.UK.2013-1.RLTS.T4162A44199757.en.

Robertson A. 1962. Selection for heterozygotes in small populations. *Genetics*. 47:1291–1300.

Roe WD, Howe L, Baker EJ, Burrows L, Hunter SA. 2013. An atypical genotype of *Toxoplasma gondii* as a cause of mortality in Hector's dolphins (*Cephalorhynchus hectori*). *Vet Parasitol*. 192:67–74.

Roe WD, Michael S, Fyfe J, Burrows E, Hunter SA, Howe L. 2017. First report of systemic toxoplasmosis in a New Zealand sea lion (*Phocarctos hookeri*). *N Z Vet J*. 65:46–50.

Ruan R, Ruan J, Wan XL, Zheng Y, Chen MM, Zheng JS, Wang D. 2016. Organization and characteristics of the major histocompatibility complex class II region in the Yangtze finless porpoise (*Neophocaena asiaeorientalis asiaeorientalis*). *Sci Rep*. 6:22471.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Seddon JM, Ellegren H. 2002. MHC class II genes in European wolves: a comparison with dogs. *Immunogenetics*. 54:490–500.

Slooten E, Dawson SM, Rayment WJ. 2004. Aerial surveys for coastal dolphins: abundance of Hector's dolphins off the South Island West Coast, New Zealand. *Mar Mamm Sci*. 20:477–490.

Spurgin LG, Richardson DS. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc Biol Sci*. 277:979–988.

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet*. 76:449–462.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*. 68:978–989.

Sutton JT, Nakagawa S, Robertson BC, Jamieson IG. 2011. Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Mol Ecol*. 20:4408–4420.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis, version 6.0. *Mol Biol Evol*. 30:2725–2729.

Tollenaere C, Susi H, Nokso-Koivisto J, Koskinen P, Tack A, Auvinen P, Paulin L, Frilander MJ, Lehtonen R, Laine AL. 2012. SNP design from 454

sequencing of *Podosphaera plantaginis* transcriptome reveals a genetically diverse pathogen metapopulation with high levels of mixed-genotype infection. *PLoS One*. 7:e52492.

Trachtenberg EA, Erlich HA, Rickards O, DeStefano GF, Klitz W. 1995. HLA class II linkage disequilibrium and haplotype evolution in the Cayapa Indians of Ecuador. *Am J Hum Genet*. 57:415–424.

Trowsdale J, Ragoussis J, Campbell RD. 1991. Map of the human MHC. *Immunol Today*. 12:443–446.

Tsuji K, Aizawa M, Sasazuki T. 1992. *HLA 1991: Proceedings of the 11th International Histocompatibility Workshop and Conference*. Oxford: Oxford University Press.

van Oosterhout C, Joyce DA, Cummings SM, Blais J, Barson NJ, Ramnarine IW, Mohammed RS, Persad N, Cable J. 2006. Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). *Evolution*. 60:2562–2574.

van Wormer E, Carpenter TE, Singh P, Shapiro K, Wallender WW, Conrad PA, Largier JL, Maneta MP, Mazet JAK. 2016. Coastal development and precipitation drive pathogen flow from land to sea: evidence from a *Toxoplasma gondii* and felid host system. *Sci Rep*. 6:1–9.

Weber DS, Stewart BS, Schienman J, Lehman N. 2004. Major histocompatibility complex variation at three class II loci in the northern elephant seal. *Mol Ecol*. 13:711–718.

Xu S, Sun P, Zhou K, Yang G. 2007. Sequence variability at three MHC loci of finless porpoises (*Neophocaena phocaenoides*). *Immunogenetics*. 59:581–592.

Yan C, Liang LJ, Zheng KY, Zhu XQ. 2016. Impact of environmental factors on the emergence, transmission and distribution of *Toxoplasma gondii*. *Parasit Vectors*. 9:137.